

Computational Methods for Coupled Problems in Science and Engineering IV

M. Papadrakakis, E. Oñate and B. Schrefler (Eds.)



**Computational Methods for Coupled
Problems in Science and Engineering IV -
“*COUPLED PROBLEMS 2011*”**

Computational Methods for Coupled Problems in Science and Engineering IV

Proceedings of the IV International Conference on Computational Methods for Coupled Problems in Science and Engineering held in Kos, Greece
20-22 June 2011

Edited by:

M. Papadrakakis

National Technical University of Athens, Greece

E. Oñate

Universitat Politècnica de Catalunya, Spain

B. Schrefler

Università degli Studi di Padova, Italy

A publication of:

International Center for Numerical Methods in Engineering (CIMNE)
Barcelona, Spain



Computational Methods for Coupled Problems in Science and Engineering IV

M. Papadrakakis, E. Oñate and B. Schrefler (Eds.)

First edition, June 2011

© The authors

Printed by: Artes Gráficas Torres S.L., Huelva 9, 08940 Cornellà de Llobregat, Spain

Depósito legal: B-42316-2011

ISBN: 978-84-89925-78-6

PREFACE

This volume contains the full length papers accepted for presentation at the IV International Conference on Computational Methods for Coupled Problems in Science and Engineering, COUPLED PROBLEMS 2011, island of Kos, Greece on June 20 – 22, 2011.

COUPLED PROBLEMS 2011 is the fourth International Conference on this subject organized in the framework of Thematic Conferences of the European Community on Computational Methods in Applied Sciences (ECCOMAS) and is a Special Interest Conference of the International Association for Computational Mechanics (IACM).

The objective of COUPLED PROBLEMS 2011 is to become a forum for state of the art presentations and discussions of mathematical models, numerical methods and computational techniques for solving coupled problems of multidisciplinary character in science and engineering. The conference goal is to make a step forward in the formulation and solution of real life problems with a multidisciplinary nature and industrial interest, accounting for all the complex couplings involved in the physical description of the problem.

Previous editions of the conference were held on the Islands of Santorini, Greece (2005), Ibiza, Spain (2007) and Ischia, Italy (2009), with an increasing number of participants in each edition. Coupled Problems 2011 has attracted over 320 participants, coming from all over the world. All together about 300 lectures were presented, including 11 plenary lectures, which reflect the current state of the research and advances in engineering practice in the field. The Proceedings contain the contributions sent directly by the authors and the editors cannot accept responsibility for any inaccuracies and opinions expressed in the text.

The conference was jointly organized by the National Technical University of Athens (Greece), the International Centre for Numerical Methods in Engineering (CIMNE), of the Technical University of Catalonia (UPC) and the University of Padova (Italy). The organizers acknowledge the encouragement and support of ECCOMAS and IACM, under whose auspices this conference is held. Special thanks go to the colleagues who contributed to the organization of the 15 Invited Sessions.

Manolis Papadrakakis

Institute of Structural Analysis &
Antiseismic Research
National Technical University
Athens, Greece

Eugenio Oñate

International Center for Numerical
Methods in Engineering (CIMNE)
Univ. Politècnica de Catalunya
(UPC), Barcelona, Spain

Bernhard A. Schrefler

Dipartimento di Costruzioni
e Trasporti
Università di Padova

ACKNOWLEDGEMENTS

The conference organizers acknowledge the support towards the organization of the COUPLED PROBLEMS 2011 Conference to the following organizations:

- European Community on Computational Methods in Applied Sciences (ECCOMAS)
- International Association for Computational Mechanics (IACM)
- Greek Association for Computational Mechanics (GRACM)
- National Technical University of Athens, Greece
- International Center for Numerical Methods in Engineering (CIMNE), Barcelona, Spain
- Universitat Politècnica de Catalunya, Spain
- Dipartimento di Costruzioni e Trasporti, Università di Padova, Italy

Plenary Speakers and Invited Session Organizers

We would also like to thank the Plenary Speakers and the Invited Session Organizers for their help in the setting up of a high standard Scientific Programme.

Plenary Speakers: Ramon Codina, Leszek Demkowicz, Thomas J. R. Hughes, Sergio R. Idelsohn, George Karniadakis, Lyesse Laloui, Wing Kam Liu, Roger Ohayon, K. C. Park, Tayfun Tezduyar and Wolfgang A. Wall

IS Organizers: Carlos Agelet de Saracibar, Yuri Bazilevs, Gernot Beer, Elias Cueto, Francisco Chinesta, Suvranu De, Eduardo Divo, Carlos Felippa, Massimo Guarnieri, Alain Kassab, Wing Kam Liu, Carmelo Majorana, Bernd Markert, Francesco Marotti de Sciarra, Roger Ohayon, Sebastià Olivella, K. C. Park, Jacques Pèriaux, Jean-Philippe Ponthot, Ernst Rank, Lakshmi Reddi, David Ryckelynck, Valentina Salomoni, Lorenzo Sanavia, Fotis Sotiropoulos, Dörte Carla Sternel, Kenji Takizawa, Tayfun Tezduyar, Zohar Yosibash, Lucy Zhang and Jessica Zhang.

SUMMARY

Preface	5
Acknowledgements	7
Contents	9
Plenary Lectures	23
Invited Sessions	55
Computational Mechanics of Multiphase Porous Materials	55
<i>Invited Session organized by Bernd Markert and Lorenzo Sanavia</i>	
Computational Multiphysics: Methods and Models	100
<i>Invited Session organized by Carlos Felippa, Roger Ohayon and K.C. Park</i>	
Coupled Problems for Biomimetics	147
<i>Invited Session organized by Fotis Sotiropoulos</i>	
Coupled Problems in Geotechnical Engineering	160
<i>Invited Session organized by Sebastia Olivella</i>	
Coupled Thermomechanical Simulation of Material Forming and Impact Problems	242
<i>Invited Session organized by Carlos Agelet de Saracibar and Jean-Philippe Ponthot</i>	
Coupling of Different Numerical Methods	277
<i>Invited Session organized by Gernot Beer</i>	
Fast Single and Multi Discipline Optimization Tools for Aircraft/Aero Engines Design with Uncertainties	313
<i>Invited Session organized by Jacques Pèriaux</i>	
Multiresolution Biomechanics: From Nano to Macro	325
<i>Invited Session organized by Wing-Kam Liu, Yongjie (Jessica) Zhang, Suvranu De and Lucy Zhang</i>	
Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering: Magneto-hydrodynamics and High frequency Coupled Problems	337
<i>Invited Session organized by Massimo Guarnieri and Carmelo Majorana</i>	
Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering: Electro-thermal Modeling	347
<i>Invited Session organized by Massimo Guarnieri and Carmelo Majorana</i>	
Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering: Electrostatic-micro Interactions and Electro-actuators	379
<i>Invited Session organized by Massimo Guarnieri and Carmelo Majorana</i>	
Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering IV: Electrostatic-multiphase Problems	398
<i>Invited Session organized by Massimo Guarnieri and Carmelo Majorana</i>	

Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering: Electro-vibrational and Electro-thermo-mechanical Problems	424
<i>Invited Session organized by Massimo Guarnieri and Carmelo Majorana</i>	
Plasticity and Damage: Experimental and Numerical Simulations	484
<i>Invited Session organized by Valentina Salomoni and Francesco Marotti de Sciarra</i>	
Speeding-up Numerical Simulations: Model Reduction and Other Advanced Discretization Techniques	570
<i>Invited Session organized by Francisco Chinesta, David Ryckelynck and Elias Cueto</i>	
Symposium Celebrating the 70th Birthday of Ahmed Sameh: Fluid-Structure Interactions	590
<i>Invited Session organized by Kenji Takizawa, Yuri Bazilevs and Tayfun Tezduyar</i>	
Contributed Sessions	623
Aeroelasticity	623
Coupled Problems in Geomechanics	686
Coupled Solution Strategies	779
Fluid-Structure Interaction	934
Multiphysics Problems	1022
Multiscale Problems	1062
Numerical Methods for Coupled Problems	1083
Optimum Design and Control of Coupled Problems	1261
Thermo-Hydromechanical Problems	1318
Thermo-Mechanochemical Problems	1400
Authors Index	1416

CONTENTS

Plenary Lectures

Coupled Problems on the Cellular and Sub-Cellular Scale W.A. Wall , C.J. Cyron, T. Klöppel, C. Meier and K. Müller	23
Model Problems in Magneto-hydrodynamics: Individual Numerical Challenges and Coupling Possibilities R. Codina , S. Badia and R. Planas	31
Space-Time FSI Modeling and Dynamical Analysis of Ringsail Parachute Clusters K. Takizawa, T. Spielman and T.E. Tezduyar	43

Invited Sessions

Computational Mechanics of Multiphase Porous Materials

Invited Session organized by Bernd Markert and Lorenzo Sanavia

A Three-phase Model for Damage Induced by ASR in Concrete Structures C. Comi and R. Pignatelli	55
An Efficient Computational Model for CO ₂ Flow in Porous Media M. Talebian, R. Al-Khoury and L.J. Sluys	67
Back Analysis of a Coupled Thermo-Hydro-Mechanical Model Based on Instrumented Constant Volume Column Test T. Schanz, M. Datcheva and L. Nguyen-Tuan	76
Modelling Water Infiltration into Macroporous Hill Slopes using Special Boundary Conditions L. Stadler, C. Adamczak and R. Hinkelmann	90

Computational Multiphysics: Methods and Models

Invited Session organized by Carlos Felippa, Roger Ohayon and K.C. Park

A Partitioned Quasi-Newton Solution Technique for Fluid-structure Interaction Problems using a Coarsened Grid to Accelerate the Convergence of the Coupling Iterations J. Degroote and J. Vierendeels	100
An Iteration-free, Partitioned Method for Solving Coupled Problems J. Rang	112
Finite Element/Boundary Element Coupling for Airbag Deployment T.M. van Opstal and E.H. van Brummelen	124
Immersed Boundary Method: Performance Analysis of Popular Finite Element Spaces D. Boffi, N. Cavallini, F. Gardini and L. Gastaldi	135

Coupled Problems for Biomimetics

Invited Session organized by Fotis Sotiropoulos

A Finite Element Method for Non-linear Hyperelasticity Applied for the Simulation of Octopus Arm Motions
V. Vavourakis, A. Kazakidi, D.P. Tsakiris and J.A. Ekaterinaris 147

Coupled Problems in Geotechnical Engineering

Invited Session organized by Sebastia Olivella

Keynote Lecture: Towards the Generic Conceptual and Numerical Framework for the Simulation of CO₂ Sequestration in Different Types of Georeservoirs

U-J. Görke, J. Taron, A.K. Singh, W. Wang and O. Kolditz 160

A Coupled Model of Mechanical Behaviour and Water Retention for Unsaturated Soils with Double Porosity
A. Koliji, L. Laloui and L. Vulliet 172

Algebraic Multigrid Preconditioning for Mixed Elliptic-Hyperbolic Problems
M. Emans 184

Hydromechanical Analysis in Geotechnical Engineering using the Material Point Method
F. Zabala and E. Alonso 194

Non-isothermal Compositional Gas Flow during Carbon Dioxide Storage and Enhanced Gas Recovery
A.K. Singh, N. Böttcher, W. Wang, U-J. Görke and O. Kolditz 206

Numerical Analysis of CO₂ Injection into Deformable Saline Reservoirs: Benchmarking and Initial Observations
J. Taron, C-H. Park, U-J. Görke, W. Wang and O. Kolditz 218

Numerical Study of the Chemo-poro-mechanical Behaviour of the Cement Sheath during CO₂ Injection
V. Vallin, J.M. Pereira, A. Fabbri, H. Wong and N. Jacquemet 230

Coupled Thermomechanical Simulation of Material Forming and Impact Problems

Invited Session organized by Carlos Agelet de Saracibar and Jean-Philippe Ponthot

Comparison of Data Transfer Methods between Two Different Meshes
P. Bussetta and J-P. Ponthot 242

Thermomechanical Simulation of Roll Forming Process based on the Coupling of Two Independent Solvers
Y. Carretta, R. Boman, A. Stephany, T. Bouache, R. Canivenc, P. Montmitonnet, N. Legrand, M. Laugier and J-P. Ponthot 254

Thixoforging Tools materials: Determination of Appropriate Features and Experimental Evaluation
A. Rassili and J-C. Pierret 265

Coupling of Different Numerical Methods

Invited Session organized by Gernot Beer

Fatigue Analysis of the Structural Components of a Mechanical Press Cutting High-Strength Steels
B-A. Behrens, R. Krimm and C. Wager 277

Finite Volume and Finite Element Schemes for the Euler Equations in Cylindrical and Spherical Coordinates
D. De Santis, G. Geraci and A. Guardone 289

Iterative Coupling of Boundary and Discrete Element Methods using an Overlapping FEM Zone
L. Malinowski, G.F. Karlis, G. Beer and J. Rojek 301

Fast Single and Multi Discipline Optimization Tools for Aircraft/Aero Engines Design with Uncertainties

Invited Session organized by Jacques Pèriaux

A Hardware-Based Multi-Disciplinary Design Optimization Method for Aeronautical Applications
J. Kok, F. Gonzalez, N. Kelson and T. Gurnett313

Multiresolution Biomechanics: From Nano to Macro

Invited Session organized by Wing-Kam Liu, Yongjie (Jessica) Zhang, Suvrana De and Lucy Zhang

Efficient Force Field Calculation in Articulated Multiscale Molecular Simulations
M. Poursina and K.S. Anderson325

Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering: Magneto-hydrodynamics and High frequency Coupled Problems ***Invited Session organized by Massimo Guarnieri and Carmelo Majorana***

Influence of Mechanical Strains on Electromagnetic Signals of a Microstrip Antenna. FEM/BIM model
N. Adnet, F. Pablo, I. Bruant and L. Proslie337

Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering: Electro-thermal Modeling

Invited Session organized by Massimo Guarnieri and Carmelo Majorana

Electromagnetic Plasma Modelling in Circuit Breaker within the Finite Volume Method
L. Rondot, J-P. Gonnet and V. Mazauric347

Investigation of the Purely Hyperbolic Maxwell System for Divergence Cleaning in Discontinuous Galerkin based Particle-In-Cell Methods

A. Stock, J. Neudorfer, R. Schneider, C. Altmann and C-D. Munz355

Numerical Simulation Framework for Weakly Coupled Multiphysical Problems in Electrical Engineering

P. Alotto, M. Jaindl, R. Kutschera, C. Magele and A. Köstinger367

Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering: Electrostatic-micro Interactions and Electro-actuators

Invited Session organized by Massimo Guarnieri and Carmelo Majorana

Numerical Modelling of Coupled Electro-Mechanical Problems for the State Space Controller Design

T. Nestorović and M. Trajkov379

Precise Electromagnetic-Thermoelastic Actuator for Laser Technologies

I. Doležel, V. Kotlan and B. Ulrych388

Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering IV: Electrostatic-multiphase Problems

Invited Session organized by Massimo Guarnieri and Carmelo Majorana

Numerical Analysis of Dielectric Micro-particle Motion in Fluid and Electric Field
M.K. Baek, Y.S. Kim and I.H. Park398

Numerical Simulation of Droplet Shapes in External Electric Field, Gravity and Surface Tension

T.I. Sung, H.S. Choi, Y.S. Kim and I.H. Park404

Numerical Simulations of Tungsten Targets Hit by LHC Proton Beam

M. Scapin, L. Peroni, A. Bertarelli and A. Dallochio413

Numerical Modelling of Coupled Electromagnetic, Hydrodynamic and Thermomechanical Problems in Engineering: Electro-vibrational and Electro-thermo-mechanical Problems

Invited Session organized by Massimo Guarnieri and Carmelo Majorana

A Computational Study of Efficiency of Controllable Semiactive Magnetorheological Dampers Reducing Lateral Vibration of Rotors Working in Chambers Submerged in liquid J. Zapoměl, L. Čermák and P. Ferfecki	424
Design, Simulation and Prototyping of Polymeric Ortho-planar Microvalves E. Bertarelli, E. Bianchi, D. Strohmeier, J.G. Korvink and G. Dubini	436
Electrostatic Diaphragm Micropump Electro-fluid-mechanical Simulation E. Bertarelli, R. Ardito and A. Corigliano	448
Multiphysical Computations of the Electrical Machines using FEM A. Stermecki, O. Bíró, M. Hettegger, H. Lang, G. Ofner, S. Rainer and B. Weilharter	460
The Different Levels of Magneto-mechanical Coupling in Energy Conversion Machines and Devices A. Belahcen, R. Kouhia, K. Fonteyn	472

Plasticity and Damage: Experimental and Numerical Simulations

Invited Session organized by Valentina Salomoni and Francesco Marotti de Sciarra

Keynote Lecture: Coupling of Damage Mechanisms in the Prediction of Creep Failure of a Welded Branched Header Pipe D.R. Hayhurst	484
A Finite Element for Nonlocal Elastic Analyses F. Marotti de Sciarra	496
A Geometric Theory of Plasticity V.P. Panoskaltzis, D. Soldatos and S.P. Triantafyllou	506
Analysis of Rock Massif Based on the Theory of Damage J. Kruis and T. Koudelka	521
Contact–Damage Coupled Modelling of FRP Reinforcements under Variable Loading Times G. Mazzucco, V.A. Salomoni and C.E. Majorana	530
Effect of Aggregates and ITZ on Visco-damaged Response of Concrete at the Meso Scale Level G. Xotta, V.A. Salomoni and C.E. Majorana	538
Non-linear Modelling, Design and Production of Steel Blast-resistant Doors and Windows V.A. Salomoni, G. Mazzucco, G. Xotta, R.L. Fincato, M. Schiavon and C.E. Majorana	548
Subloading Surface Plasticity Model Algorithm for 3D Subsidence Analyses Above Gas Reservoirs V.A. Salomoni and R.L. Fincato	559

Speeding-up Numerical Simulations: Model Reduction and Other Advanced Discretization Techniques

Invited Session organized by Francisco Chinesta, David Ryckelynck and Elias Cueto

Application of the Proper Generalized Decomposition Method to a Viscoelastic Mechanical problem with a Large Number of Internal Variables and a Large Spectrum of Relaxation Times M. Hammoud, M. Beringhier and J-C. Grandidier	570
Thermo-mechanical Modelling of Dynamic Tensile Extrusion Test L. Peroni, M. Scapin and M. Avale	578

Symposium Celebrating the 70th Birthday of Ahmed Sameh: Fluid-Structure Interactions
Invited Session organized by Kenji Takizawa, Yuri Bazilevs and Tayfun Tezduyar

Comparative Patient-Specific FSI Modeling of Cerebral Aneurysms K. Takizawa, T. Brummer, T.E. Tezduyar and P.R. Chen	590
Multi-scale Analysis on Cavitation Damage and its Mitigation for the Spallation Neutron Source K. Okita, K. Ono, S. Takagi and Y. Matsumoto	600
Multiscale Space-Time Computation Techniques K. Takizawa and T.E. Tezduyar	611

Contributed Sessions

Aeroelasticity

A Refined 1D FE Model for Application to Aeroelasticity of Composite Wings A. Varello, M. Petrolo and E. Carrera	623
Experimental Benchmark and Numerical Validation of a Free Heaving Airfoil J.J. Sterenberg, A.H. van Zuijlen and H. Bijl	635
Frequency Parametrization to Numerically Predict Flutter in Turbomachinery M. Philit, L. Blanc, S. Aubert, W. Lolo, P. Ferrand and F. Thouverez	647
Neural Networks as Surrogate Models for Nonlinear, Transient Aerodynamics within an Aeroelastic Coupling-scheme in the Time Domain K. Lindhorst, M.C. Haupt and P. Horst	661
Wind-structure Interaction on Construction Stages for Unbalanced Segmental Bridges A. Hernández and J. G Valdés	677

Coupled Problems in Geomechanics

An Optimized Return Mapping Algorithm for the Barcelona Basic Model M. Pertl, M. Hofmann and G. Hofstetter	686
Dynamic Behaviour of Saturated Poroelastic Layers with Embedded Wall Submitted to Seismic Actions J.C. Grazina and P.L. Pinto	698
Hierarchical Hybrid Simulation of Biofilm Growth Dynamics in 3D Porous Media G. Kapellos, T. Alexiou and S. Pavlou	710
Non-linear Dynamic Soil Response Underneath a Vertical Breakwater Subjected to Impulsive Sea Wave Actions M.M. Stickle, P. de la Fuente and C. Oteo	721
Numerical Modeling of Shallow Landslide Impacts on Flexible Protection Systems and Its Validation with Full-scale Testing A. von Boetticher, A. Volkwein, R. Wüchner, K-U. Bletzinger and C. Wendeler	731
Shear Deformable Beams on Nonlinear Viscoelastic Foundation under Moving Loading E.J. Sapountzakis and A.E. Kampitsis	743
Some Consequences of the Qualitative Analysis of the Point-Symmetric Coupled Consolidation Models E. Imre and S. Fityus	755

Water Retention in Unsaturated Soils Subjected to Wetting and Drying Cycle W. Arairo, F. Prunier and I. Djeran-Maigre	767
--	-----

Coupled Solution Strategies

A Multi-model Incremental Adaptive Strategy to Accelerate Partitioned Fluid-structure Algorithms using Space Mapping T. Scholcz, A.H. van Zuijlen and H. Bijl	779
A New Numerical Scheme for a Linear Fluid-structure Interaction Problem M. González and V. Selgas	792
A Substructuring FE Model for Structural-acoustic Problems with Modal-based Reduction of Poroelastic Interface R. Rumpler, J-F. Deü and P. Göransson	802
Coarse Level Newton-Krylov Acceleration of Sub-iterations in Partitioned Fluid-structure Interaction A.H. van Zuijlen and H. Bijl	814
Comparisons of Coupling Strategies for Massively Parallel Conjugate Heat Transfer with Large EDDI Simulations S. Jauré, F. Duchaine and L.Y. Gicquel	826
Finite Element Methods for Strongly-coupled Systems of Fluid-structure Interaction with Application to Granular Flow in Silos S. Reinstädler, A. Zilian and D. Dinkler	837
Higher Order Time Integration Schemes for Thermal Coupling of Flows and Structures V. Kazemi-Kamyab, A.H. van Zuijlen and H. Bijl	850
Mixing Snapshots and Fast Time Integration of PDEs M-L. Rapún, F. Terragni and J.M. Vega	861
Multi-level and Quasi-Newton Acceleration for Strongly Coupled Partitioned Fluid-structure Interaction J.J. Kreeft, M. Weghs, A.H. van Zuijlen and H. Bijl	873
O-PALM: an Open Source Dynamic Parallel Coupler A. Piacentini, Th. Morel, A. Thévenin and F. Duchaine	885
Partitioned Time Integration Methods for Hardware in the Loop based on Linearly Implicit L-stable Rosenbrock methods O.S. Bursi, C. Jia and Z. Wang	896
Periodic Solutions for Acousto-Elastic Scattering Problems by Controlling the Time-Dependent Equations S. Mönkölä and S. Kähkönen	908
Proper Generalized Decomposition (PGD) to solve Mixed Convection Problem A. Dumon, C. Allery and A. Ammar	919
The Influence of Different Time Integration Schemes in ALE Description Applied to Moving Meshes F. Flitz, D.C. Sternel and M. Schäfer	925

Fluid-Structure Interaction

Computational Fluid-structure Interaction Simulations for Wind Induced Vibrations in Silo Groups J. Hillewaere, J. Degroote, G. Lombaert, J. Vierendeels and G. Degrande	934
FSI Simulations for Explosions Very Near Reinforced Concrete Structures M.A. Price, A. Lee, O. Soto and O.Y. Chong	946
Investigating the Effect of Rotational Degree of Freedom on a Circular Cylinder at Low Reynolds Number in a Cross Flow S.H. Madani, J. Wissink and H. Bahai	954

Modeling of Airblast Propagation through an Enclosed Structure J.A. Sherburn, D.H. Nelson, C.D. Price and T.R. Slawson	962
Multidisciplinary Investigation by Fluid-Structure-Motion Integrated Simulation S. Takahashi and N. Arai	974
Numerical Study of the Fluid-structure Interaction in the Diffuser Passage of a Centrifugal Pump A. Fontanals, A. Guardo, M. Coussirat and E. Egusquiza	986
Onset of Two-Dimensional Turbulence with High Reynolds Numbers in the Navier-Stokes Equations B. Bermudez and A. Nicolas	996
Vortex-induced Vibrations of an Elastically Mounted Cylinder with Low Mass-ratio at $Re=3900$ Y. Jus, E. Longatte, J-C. Chassaing and P. Sagaut	1007
Multiphysics Problems	
Mathematical Formulation of a Coupled Dynamic Model for Environmental Business Action Planning S.A. David, C. de Oliveira and D.D. Quintino	1022
Modernizing Science & Engineering Software Systems L. Favre, L. Martinez and C. Pereira	1033
Multi-Physics Modelling and Simulation of Sand Transfer around Cube K. Matsui, M. Suzuki and M. Yamamoto	1045
Recurrence Plots Analysis of Pressure Fluctuations in Fluidized Beds M.M. Tahmasebpoor, R. Sotudeh-Gharebagh, R. Zarghami and N. Mostoufi	1054
Multiscale Problems	
Fabrication and Multiscale Mechanical Characterization of Ti Alloy/ Al_2O_3 Functionally Graded Materials for Orthopaedic Applications D. Carnelli, E. Bertarelli, D. Gastaldi, T. Villa, F. Casari, A. Molinari and P. Vena	1062
Magnetization Model for Particle-based Simulations of Magnetorheological Fluids H. Lager, J. Peguiron, C. Bierwisch and M. Moseler	1071
Numerical Methods for Coupled Problems	
A Three-phase Finite Element Model of Water-infiltrated Porous Materials Subjected to Freezing M. Zhou and G. Meschke	1083
Adaptive Calibration of Nonlocal Coupled Damage-Plasticity Model for Aluminium Alloy AA6082 T0 J.P. Belnoue and A.M. Korsunsky	1095
Advanced Stochastic FEM-Based Artificial Neural Network for Crack Damage Detection C. Sbarufatti, A. Manes and M. Giglio	1107
Computational Study of the Interaction between a Newtonian Fluid and a Cellular Biological Medium in a Straight Vessel T.S. Alexiou, G. Kapellos and S. Pavlou	1120
Crack Propagation in Civil Engineering Bridge Cables: Coupled Phenomena of Fatigue and Corrosion V. Périer, L. Dieng and L. Gaillet	1128
Development and Implementation of an Eulerian Approach for Efficient Simulation of Frictional Heating in Sliding Contacts N. Strömberg	1136
Dynamically Coupled Models of the Sliding and Spinning Friction based on Padé Expansions A. Kireenkov	1147

Method for Estimating Parameters of Coupled Problem of Interaction of Gas Flows Loaded by Solid Particles with Solids A.V. Nenarokomov, O.M. Alifanov, E.A. Artiukhine, I.V. Repin and D.M. Titov	1155
Modelling of Simultaneous Induction Hardening in Monolithic Formulation P. Karban, I. Doležel and B. Ulrych	1167
Numerical Simulation of Aileron Buzz using an Adaptive-grid Compressible Flow Solver for Dynamic Meshes G. Forestieri, A. Guardone, D. Isola, F. Marulli and G. Quaranta	1176
Numerical Simulations of Particles in a Shear Flow N. Verdon, P. Laure, A. Lefebvre-Lepot and L. Lobry	1188
Reduced-order Modeling of Parametrized Finite Element Solutions by the POD-ISAT Technique. Application to Aircraft Air Control Systems D. Bui, M. Hamdaoui and F. De Vuyst	1200
Simulation of Residual Stresses in an Induction Hardened Roll C. Groth and L. Hellenthal	1212
Solution of Dendritic Growth in a Binary Alloy by a Novel Point Automata Method A.Z. Lorbicka and B. Šarler	1220
Solution of Integral Equation in Scattering Analysis of Bodies of Revolution by MoM with First Type Elliptic Integrals C. Vidal and U. Resende	1232
Some Aspects of Dynamic Computational Modelling of Direct Current Plasma Arc Phenomena Q.G. Reynolds and B.D. Reddy	1239
The Influence of Stem Surface in Micromobility and Cement Bone Stresses A. Ramos, C. Relvas, A. Completo and J. A Simões	1251
Optimum Design and Control of Coupled Problems	
Coefficients Identification for Ship Manoeuvring Simulation Model based on Optimization Techniques K.T. Tran, A. Ouahsine, H. Naceur, F. Hissel and A. Pourplanche	1261
Evolutionary, Iterative Aerodynamical Global Optimization, with Weak Aerodynamics/Structure Interactions A. Nastase	1273
Model Order Reduction of Systems for Active Vibration and Noise Control M. Kurch, H. Atzrodt, V. Carli, O. Heuss and J. Mohring	1282
Sensor Network Optimization for Damage Detection on Aluminium Stiffened Helicopter Panels C. Sbarufatti, A. Manes and M. Giglio	1294
Solving the Dilemma of Contradictory Goals O. Frommann	1306
Thermo-Hydrromechanical Problems	
A Finite Element Modeling of Thermo-hydro-mechanical Behaviour and Numerical Simulations of Progressive Spalling Front M.T Phan, F. Meftah, S. Rigobert, P. Autuori, C. Lenglet and S. Dal Pont	1318
Aero-thermo-mechanical Coupling for Flame-wall Interaction B. Baqué, M.P. Errera, A. Roos, F. Feyel, E. Laroche and D. Donjat	1326
Application of a Viscoplastic-Damage Model for the Failure Prediction of Regeneratively Cooled Nozzle Structures V. Tini, I.N. Vladimirov and S. Reese	1332

Coupled Analysis of Transport Processes and Mechanical Behaviour of Concrete at High Temperatures F. Cramer, L. Ostermann, U. Kowalsky and D. Dinkler	1342
Coupled Simulation of Process-integrated Powder Coating by Radial Axial Rolling of Rings R. Kebriaei, J. Frischkorn and S. Reese	1350
Evaluation and Improvement of the THM Modelling Capabilities for Rock Salt Repositories A. Pudewills	1362
Mass, Heat and Momentum Transfer in Natural Draft Wet Cooling Tower with Flue Gas Discharge A.F. Klimanek and R.A. Bialecki	1371
Process Modelling of Linear Friction Welding (LFW) between AA2124/SiC _p Composite and Unreinforced Alloy X. Song, N. Baimpas, S. Harding and A.M. Korsunsky	1379
Three Domain Thermal and Mechanical Fluid-Structure Interaction Analysis Applied to Cooled Rocket Thrust Chambers D. Kowollik, M.C. Haupt and P. Horst	1388
Thermo-Mechanochemical Problems	
Coupled Reactive Transport Modeling - The Program Transport V. Žabka and J. Šembera	1400
Thermoelastic Stress Analysis for a Tube under General Mechanochemical Corrosion Conditions Y. Pronina	1408
Authors Index	1416

PLENARY LECTURES

COUPLED PROBLEMS ON THE CELLULAR AND SUB-CELLULAR SCALE

WOLFGANG A. WALL, CHRISTIAN J. CYRON, THOMAS KLÖPPEL,
CHRISTOPH MEIER, KEI MÜLLER

Institute for Computational Mechanics, Technische Universität München
Boltzmannstr. 15, 85748 Garching b. München, Germany
e-mail: wall@lrm.mw.tum.de, www.lrm.mw.tum.de

Key words: biomechanics, cell mechanics, red blood cell, polymer network, Brownian dynamics, finite elements

Abstract. Coupled problems arising in biomechanics on the cellular and sub-cellular scale are discussed based on two model systems: the human red blood cell (RBC) and biopolymeric networks. We present a novel finite-element based model on cellular level for RBC and demonstrate its potential to realistically simulate various experimental results with one set of parameters. In addition to that we present a finite-element approach to the computer aided analysis of biopolymer networks on sub-cellular level.

1 INTRODUCTION

Relevant advances and insights in biomechanics are more and more obtained on smaller scales. In this contribution we will present some of our recent work on such scales including a variety of coupled problems. For this purpose, we will concentrate on two selected subjects - one on the cellular and one on the sub-cellular scale.

Human red blood cells (RBC), that can be sketched as a liquid capsule enclosed by a biological membrane, often serve as model system for living cells. The great interest in RBC behaviour also comes from the fact that they govern the rheology of blood and the avoidance of RBC damage can govern the design of medical devices. A comprehensive numerical model of the RBC and its interaction with the surrounding plasma first necessitates a realistic model of the cell membrane. The two-layered continuum based model we have presented in [1] accounts for the characteristics of the lipid-bilayer and the cytoskeleton separately. Second, the membrane has to be coupled to the interior and to the surrounding fluid. Here, we propose a dual mortar based monolithic fluid-structure interaction (FSI) approach [2] to deal with the resulting balloon-type problem. This approach can handle nonconforming FSI interfaces that arise from different solution requirements of the physical fields and can easily be extended to allow for the simulation of rotating

structures. We show that our RBC model seems to be the first model that is able to accurately match different experimental results with one set of realistic parameters.

As the mechanics of RBC on the cellular scale, the mechanics of biopolymer networks is a model problem on the sub-cellular scale: the viscoelastic response of living cells is well-known to be largely determined by the so-called structural polymorphism [3] of the cytoskeleton, which is thus of crucial importance for cell motility, cell division and mechano-transduction in cells and as a consequence also for the development of new techniques in tissue engineering and cancer therapy. The cytoskeleton itself is a biopolymer network in the interior of cells. A major difficulty in the simulation of biopolymer networks on sub-cellular scale is the great impact the laws of statistical mechanics have on this length scale. As pointed out in [4, 5, 6], this impact has to be accounted for by a space-time white noise term if the single biopolymers are modelled as rod-like Cosserat continua. In this part we give a brief introduction into a novel finite element approach to the simulation of biopolymer networks (including the essential coupling with the stochastic field) and its application to biologically relevant problems such as structural polymorphism of such networks

2 THE HUMAN RED BLOOD CELL

2.1 The red cell membrane

The erythrocyte membrane can be considered to mainly consist of a phospholipid bilayer and a cytoskeleton network. The behavior of the RBC membrane is greatly affected by its two-layered construction and the different material characteristics found in each layer. Thus, the proposed approach models the RBC membrane by two distinct layers, the modelling of which has been discussed in great detail in [1] and is only briefly addressed here.

On the one hand the outer lipid bilayer is very thin and contributes little to the membrane in-plane shear elasticity, but on the other hand it provides the significant bending elasticity of the membrane. Moreover, it shows a fluid-like characteristics namely viscosity, incompressibility and large resistance against area dilatation. The basic idea to include this difficult characteristic property of the lipid bilayer in the present membrane model, is to use an incompressible solid-shell element, an algebraic global surface area constraint and an anisotropic viscoelastic constitutive law. The latter adds virtual rods initially aligned in the thickness direction of the incompressible solid shell element. Since a bending deformation of an element in this layer will stretch these virtual rods, their stiffness provides the necessary bending elasticity to the membrane.

The cytoskeleton of the RBC is a spectrin network tethered to intramembrane proteins in the lipid bilayer [7]. Spectrin is a highly extensible elastic molecule. Thus the cytoskeleton is the source of shear elasticity, its network topology implying a hardening effect for large deformations. Numerical, this is accounted for by a compressible, third-order material.

Modelling two distinct membrane layers accounts for the required properties of healthy RBC as stated above. Furthermore, will allow us to study certain diseases and to investigate certain damaging effects independently: defects of the lipid bilayer, defects of the cytoskeleton and damaged connection between the two membrane layers.

2.2 Coupling between membrane, surrounding fluid and cytoplasm

Considering the cell's mechanical response to dynamic deformation requires a reliable fluid-structure interaction (FSI) framework that allows for the interaction between membrane and cytoplasm as well as for the interaction between membrane and outer blood plasma. The balloon-like problem emerging from the membrane and the enclosed incompressible cytoplasm necessitates the application of monolithic FSI coupling schemes. Different resolution requirements demands independent mesh generation for blood plasma and RBC membrane, resulting in a nonconforming FSI interface.

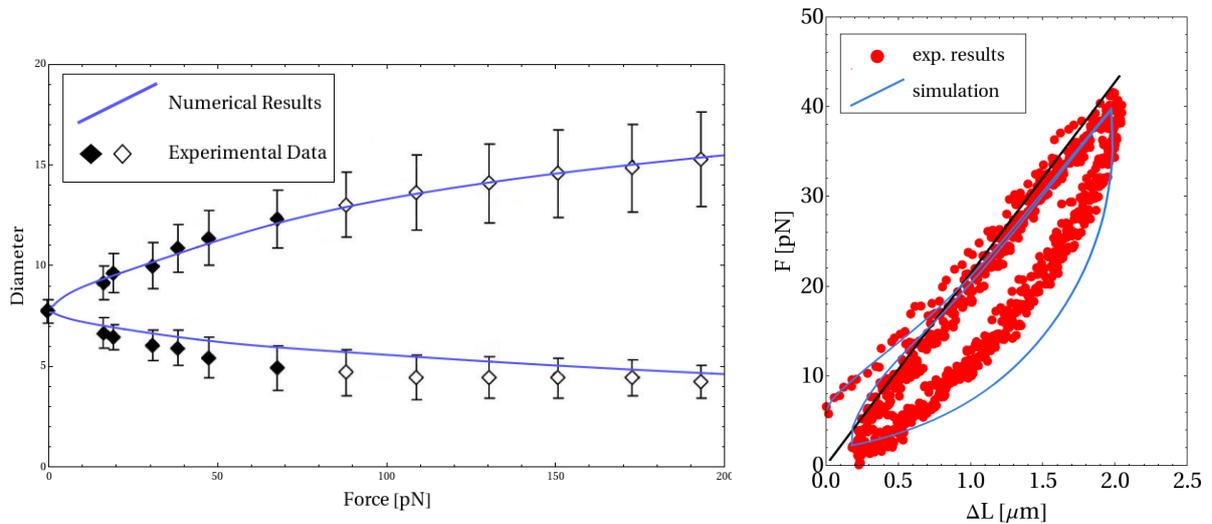
To meet these requirements, we developed a novel dual mortar based monolithic FSI scheme [2]. The additional Lagrange multiplier degrees of freedom are eliminated from the global linear monolithic system by condensation, which requires the inversion of one of the coupling matrices. Due to the dual mortar approach this inversion is trivial because of the diagonal shape of the considered matrix. The final system matrices have the same block structure as the matrices obtained in the conforming case. Therefore, the iterative solution methods recently proposed in [8] for conforming FSI interfaces can be used without any conceptual changes.

Since the coupling matrices do not change during the simulation, the computational cost associated with their evaluation is negligible. Compared to the monolithic FSI approach used in [8], the elimination of the Lagrange multipliers requires additional explicit matrix-matrix products on some sub-blocks of the global monolithic system of equations. In [2] we show that the increase in computation time caused by these operations is very low.

If subjected to shear flow the cell undergoes large rotations, which is also to be dealt with by the FSI framework. Owing to its generality and flexibility the proposed approach can be easily extended to allow for the simulation of rotating structures, by decoupling the fluid grid displacements and the fluid velocities at the interface. Using a projection algorithm the fluid grid interface does not follow the rotation of the structure, but accounts for its translation and deformation.

2.3 Numerical results

The model of the human red blood cell is validated based on results from laser trap experiments, which allow to investigate the deformation of a single erythrocyte under a well-controlled loading condition. In these experiments two rigid silica beads are attached at the diametrically opposite ends of the cell. With the help of a laser, a force is applied to the beads and the cell can be stretched.



(a) Displacement over force diagram. Top curve for axial, bottom curve for transverse diameter. Compared with experimental data from [9] (b) Hysteresis curve at $5 \mu\text{m/s}$ trap speed in comparison with results from [11]

Figure 1: Comparison between numerical and experimental results for laser trap experiments.

To investigate the elastic behavior, which is governed by the cytoskeleton for large deformations, we focus on static laser trap experiments presented in [9]. This work has been the experimental basis for several numerical studies. In [1] we show that the model predict the outcome of a static laser trap experiment very well. An excellent agreement of the force over displacement diagrams is shown in Fig. 1(a). The numerical results are obtained with a total initial shear modulus μ_M of $4.8 \text{ pN}/\mu\text{m}$, which is in the range stated in literature. Other computational studies in literature propose a higher shear modulus. In [1] it has been discussed that the difference can be explained by the fact, that all available models in literature either neglect the area conservation of the lipid bilayer or the particular loading conditions in optical tweezers. These shortcomings in the models have to be compensated for by higher values of the shear modulus.

Besides static experiments, laser traps also allow for a dynamic deformation of the cell. In [11] a triangle wave-like displacement of one bead was realised for different velocities, while the second bead was fixed. Force over displacement curves showed a hysteresis behavior and the energy dissipation can be related to the strain rate in the cell. Employing values for fluid and membrane viscosities taken from literature, numerical results show good agreement to the experimental data, as can be seen in Fig. 1(b).

3 BIOPOLYMER NETWORKS

3.1 Importance of biopolymer networks on sub-cellular scale

As the mechanics of RBC on the cellular scale, the mechanics of biopolymer networks is a model problem on the sub-cellular scale: shape and mechanical properties of living cells are largely determined by the structure of their cytoskeleton. The cytoskeleton is a biopolymer network mainly consisting of actin filaments, microtubules and intermediate filaments. In response to different incentives it can adjust its structure in order to promote cell motility or cell division thereby enabling directed growth or repair of tissue in living organisms. The precise mechanisms of this dynamic structural polymorphism of the cytoskeleton have been unknown to a large extent so far. For their experimental investigation usually simplified in-vitro model systems are studied such as filament networks consisting only of one type of biopolymer - typically actin - in aqueous solution. Filament length in such in-vitro systems typically ranges between $1\mu m$ and $20\mu m$ and the filaments either form an entangled polymer solution or are connected by so-called crosslinker molecules, which establish transient elastic links between different filaments. Despite the simplicity of these in-vitro systems as compared to the cytoskeleton in living cells, their experimental or theoretical investigation is yet a challenging problem owing to their small length scale and still tremendous complexity. Thus, computer simulations have become increasingly popular in this field.

3.2 Finite element simulation of biopolymer networks

The mechanics of single biopolymers such as actin can be modeled on a micrometer scale in many cases by the laws of continuum mechanics. The laws of quantum mechanics do not matter on this length scale and even molecular mechanics can be neglected in many cases. Thus, single biopolymers can be modeled as rod-like Cosserat continua and represented in computer simulations by nonlinear three dimensional finite beam elements [6]. Due to the extreme slenderness ratios of typical biopolymers in cytoskeletal networks (e.g. ~ 1000 for a $1\mu m$ long actin filament) and the slow motion of the fluid they are embedded into, the dynamics of the solvent surrounding the biopolymers has usually not to be simulated explicitly. Rather the effect of the solvent to the biopolymers can usually be accounted for by distributed line loads on the rod-like Cosserat continua representing the biopolymers in our model. Unlike it would be the case in simulations of macroscopic systems, these line loads are not only viscous damping terms owing to the relative motion between biopolymer and fluid, but rather they have to include also a so-called Brownian sheet, i.e. a stochastic excitation with zero correlation time and zero correlation length in space [4, 5, 6]. This stochastic excitation term is a direct consequence of the laws of statistical mechanics, which play an important role for biopolymers on the micrometer scale. These laws imply the so-called fluctuation-dissipation theorem, which says that in a system each degree of freedom subject to a damping force or moment at the same time experiences a stochastic excitation proportional to the square root of its damping

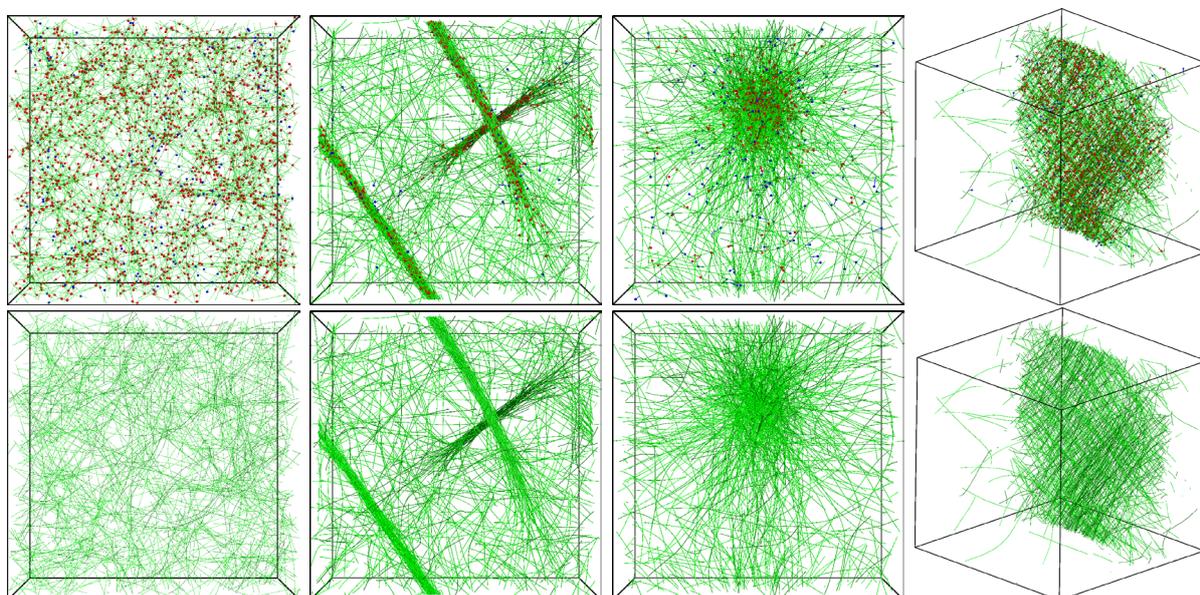


Figure 2: Isotropic-homogeneous, bundle, cluster and layer (from left to right) network phase reproduced in a finite element simulation, illustrated with (top) and without (bottom) crosslinker molecules

coefficient and the system temperature. In principle this stochastic excitation is present in mechanical systems on any scale, i.e. also for example in the wall of a building or the crankshaft of a motor. Nevertheless, in the analysis of such macroscopic systems it can usually be neglected as it scales at constant temperature with the square root of the damping coefficient. The damping coefficient itself, however, scales usually with the surface area of a body in case external friction and with its stiffness or mass in case of internal friction. Thus the impact of the stochastic excitation decreases with increasing characteristic length of the system so that for typical problems in structural engineering it can usually be neglected. However, on the micrometer scale stochastic excitation plays a crucial role.

3.3 Numerical results

In experiments so far mainly four thermodynamical phases have been observed in crosslinked actin networks [13, 14]: isotropic-homogeneous phases, bundle phases, cluster phases, and layer phases. So far only little is known about the exact conditions under which these different phases arise. Using finite element simulations with nonlinear finite beam elements and coupling the stochastic field in the fluid surrounding the biopolymers to the dynamics of the biopolymers by means of distributed stochastic line loads as described in section 3.2, all the four phases could be reproduced in finite element simulations as shown in Fig. 3.3. In addition to that, the simulations revealed that orientation constraints of crosslinker molecules and the crosslinker density determine which phases arise.

4 CONCLUSIONS

Both on cellular and sub-cellular level finite element methods allow for the computer aided analysis of both in-vitro and in-vivo systems. Not only are they capable of capturing the mainly deterministic dynamics predominant on cellular-level, but also the stochastic Brownian dynamics of macromolecules on sub-cellular level. Therefore, finite element simulations are expected to be a highly valuable tool for the future scientific and technological progress in biomechanics, bioengineering and biomedical engineering.

REFERENCES

- [1] Klöppel, T. and Wall, W.A. A novel two-layer, coupled finite element approach for modelling the nonlinear elastic and viscoelastic behavior of human erythrocytes *Biomech. Model. Mechanobiol.* (2010) DOI:10.1007/s10237-010-0246-2.
- [2] Klöppel, T., Popp, A., Küttler, U. and Wall, W.A. Fluid-structure interaction for non-conforming interfaces based on a dual mortar formulation. *Comput. Methods Appl. Mech. Engrg.* (2010) submitted.
- [3] Lieleg, O., Schmoller, K.M., Cyron, C.J., Luan, Y., Wall, W.A. and Bausch, A.R. Structural polymorphism in heterogeneous cytoskeletal networks, *Soft Matter* (2009) **5**(9): 1796-1803.
- [4] Cyron, C.J. and Wall, W.A. Finite-element approach to Brownian dynamics of polymers. *Phys. Rev. E* (2009) **80**(6): 066704
- [5] Cyron, C.J. and Wall, W.A. Consistent finite-element approach to Brownian polymer dynamics with anisotropic friction. *Phys. Rev. E* **82**(6): 066705
- [6] Cyron, C.J. and Wall, W.A. Numerical method for the simulation of the Brownian dynamics of rod-like microstructures with three dimensional nonlinear beam elements. *Int. J. Numer. Meth. Eng.* (2010) submitted
- [7] Heinrich, V., Ritchie, K., Mohandas, N. and Evans, E.A. Elastic Thickness Compressibility of the Red Cell Membrane *Biophys. J.* (2001) **81**:1452-1463
- [8] Gee, M.W., Küttler, U. and Wall, W.A. Truly monolithic algebraic multigrid for fluid-structure interaction *Int. J. Numer. Meth. Engng.* (2011) **85**(8):987–1016
- [9] Mills, J.P., Qie, L., Dao, M., Lim, C.T. and Suresh, S. Nonlinear elastic and viscoelastic deformation of the human red blood cell with optical tweezers. *Mech. Chem. Biosyst.* (2004) **1**:169-180
- [10] Henon, S., Lenormand, G., Richert, A. and Gallet, F. A New determination of the shear modulus of the human erythrocyte membrane using optical tweezers. *Biophys. J.* (1999) **115**:515-519

- [11] Yoon, Y.-Z., Kotar, J., Yoon, G. and Cicuta, P. The nonlinear mechanical response of the red blood cell. *Phys. Biol.* (2008) **5**:036007
- [12] Svitkina, T.M., Verkhovsky, A.B., McQuade, K.M. and Borisy, G.G. Analysis of the Actin-Myosin II System in Fish Epidermal Keratocytes: Mechanism of Cell Body Translocation. *The Journal of Cell Biology* (1997) **139**:397-415
- [13] Wong, G.C.L., Lin, A., Tang, J.X., Li, Y., Janmey, P.A. and Safinya, C.R. Lamellar Phase of Stacked Two-Dimensional Rafts of Actin Filaments. *Physical Review Letters* (2003) **91**:018103–
- [14] Lieleg, O., Claessens, M.M.A.E. and Bausch, A.R. Structure and dynamics of cross-linked actin networks. *Soft Matter* (2010) **6**:218–225
- [15] Cyron, C.J. and Wall, W.A. Dynamic simulations of biopolymer networks with finite elements. To be submitted

MODEL PROBLEMS IN MAGNETO-HYDRODYNAMICS: INDIVIDUAL NUMERICAL CHALLENGES AND COUPLING POSSIBILITIES

RAMON CODINA*, SANTIAGO BADIA† AND RAMON PLANAS†

*Universitat Politècnica de Catalunya
Jordi Girona 1-3, Edifici C1, 08034 Barcelona, Spain
e-mail: ramon.codina@upc.edu

†Centre Internacional de Mètodes Numèrics a l'Enginyeria (CIMNE)
Parc Mediterrani de la Tecnologia, UPC
Esteve Terrades 5, 08860 Castelldefels, Spain
e-mail: sbadia@cimne.upc.edu, rplanas@cimne.upc.edu

Key words: MHD, compatibility conditions, stabilized finite element methods, Stokes' problem, Maxwell's problem, Darcy's problem, iterative schemes

Abstract. In this work we discuss two model problems appearing in magneto-hydrodynamics (MHD), namely, the so called full MHD problem and the inductionless MHD problem. The first involves as unknowns the fluid velocity and pressure, the magnetic (induction) field and a pseudo-pressure introduced to impose the zero-divergence restriction of this last unknown. The building blocks of this model are the Stokes problem for the velocity and the pressure and the Maxwell problem for the magnetic field and pseudo-pressure. We discuss the numerical challenges of the approximation of these two model problems having in mind the need to couple them in the full problem, where additional coupling terms appear. The second model we consider is the inductionless MHD approximation. Instead of the magnetic induction and pseudo-pressure, the magnetic unknowns are now the current density and the electric potential. The building blocks are the Stokes problem for the fluid and the Darcy problem (in primal form) for the current density and electric potential. We discuss also the numerical challenges involved in the approximation of this last problem, particularly considering that it has to be coupled to the former. Once the building blocks have been analysed independently, the possibilities of dealing with the fully coupled problems are discussed. Iterative schemes that can be shown to be stable are presented in the stationary case, showing that a segregated solution for the flow and the magnetic problem is not possible. Most of the results presented are elaborated independently in previous works. Our objective in this paper is to present the different problems with a unified perspective.

1 INTRODUCTION

The objective of this work is to discuss some aspects related to the finite element approximation of two model problems in MHD, namely, the so called full MHD approximation and the inductionless model. The two main issues to be addressed are the compatibility conditions required for the unknowns and the iterative schemes that may be used (at least for the stationary problem), as well as the links between both aspects.

Let us present the two models to be discussed. The general MHD approximation can be stated as follows. Given a domain $\Omega \subset \mathbb{R}^d$ and a time interval $(0, T)$, find a velocity $\mathbf{u} : \Omega \times (0, T) \rightarrow \mathbb{R}^d$, a pressure $p : \Omega \times (0, T) \rightarrow \mathbb{R}$, a magnetic (induction) field $\mathbf{B} : \Omega \times (0, T) \rightarrow \mathbb{R}^d$ and magnetic pseudo-pressure $r : \Omega \times (0, T) \rightarrow \mathbb{R}$ as the solution of the problem:

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p - \frac{1}{\mu_m \rho} (\nabla \times \mathbf{B}) \times \mathbf{B} = \mathbf{f}_f, \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (2)$$

$$\partial_t \mathbf{B} + \frac{1}{\mu_m \sigma} \nabla \times (\nabla \times \mathbf{B}) - \nabla \times (\mathbf{u} \times \mathbf{B}) + \nabla r = \mathbf{f}_m, \quad (3)$$

$$\nabla \cdot \mathbf{B} = 0. \quad (4)$$

In these equations, ρ is the fluid density, ν is the kinematic viscosity, \mathbf{f}_f and \mathbf{f}_m are the body forces for the momentum and magnetic field equations (zero in the applications), μ_m is the the magnetic permeability and σ the conductivity. Appropriate initial and boundary conditions need to be appended. Note that the pseudo-pressure r has been added ($r = 0$ is the exact solution).

The second model to be considered is the inductionless MHD approximation. Now \mathbf{B} is assumed to be given, causing an unknown current density $\mathbf{j} : \Omega \times (0, T) \rightarrow \mathbb{R}^d$ and an unknown electric potential $\phi : \Omega \times (0, T) \rightarrow \mathbb{R}$. The equations to be solved are:

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p - \frac{1}{\rho} (\mathbf{j} \times \mathbf{B}) = \mathbf{f}_f, \quad (5)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (6)$$

$$\mathbf{j} + \sigma \nabla \phi - \sigma (\mathbf{u} \times \mathbf{B}) = \mathbf{0}, \quad (7)$$

$$\nabla \cdot \mathbf{j} = 0. \quad (8)$$

As before, appropriate initial and boundary conditions need to be appended.

Apart from the complex physics represented by equations (1)-(4) and (5)-(8), its finite element approximation has several difficulties. The purpose of this paper is to touch two of them, as mentioned above. We will present here a summary of previous works, presented in a unified manner and showing their computational implications. In particular, basic compatibility conditions and the use of stabilized formulations as a means to avoid them are analyzed in [3, 1, 2], the finite element approximation of the general MHD problem

in [4] and of the inductionless model in [5]. The reader is referred to these articles for additional references, where appropriate credit to previous developments is indicated. Due to space restrictions, no further references will be included.

2 BUILDING BLOCKS

Let us first consider problem (1)-(4). Deleting nonlinearities, coupling terms and time derivatives, we are left with a Stokes problem for \mathbf{u} and p and a Maxwell problem for \mathbf{B} and r . Therefore, the inf-sup conditions for both problems are *necessary* conditions to be met when the Galerkin finite element approximation of the problem is undertaken. It is easily shown that these conditions are also *sufficient*.

Let us turn our attention to (5)-(8). Deleting again nonlinearities, coupling terms and time derivatives, the problems we now find are a Stokes problem for \mathbf{u} and p and a Darcy problem for \mathbf{j} and ϕ . The inf-sup conditions associated to both problems are required if the problem is approximated using the standard Galerkin method.

From these observations it is clear that the building blocks of a finite element approximation of the general MHD model and the inductionless approximation are the Stokes, the Maxwell and the Darcy problems. These are the problems whose approximation is discussed in this section. First of all, let us write them in a unified format. They consist in finding $\mathbf{u} : \Omega \rightarrow \mathbb{R}^d$ and $p : \Omega \rightarrow \mathbb{R}$ such that

Stokes:

$$\begin{aligned} -\nu \Delta \mathbf{u} + \nabla p &= \mathbf{f}, \\ \nabla \cdot \mathbf{u} &= 0. \end{aligned}$$

Maxwell:

$$\begin{aligned} \lambda \nabla \times \nabla \times \mathbf{u} + \nabla p &= \mathbf{f}, \\ \nabla \cdot \mathbf{u} &= 0. \end{aligned}$$

Darcy:

$$\begin{aligned} \sigma \mathbf{u} + \nabla p &= \mathbf{f}, \\ \nabla \cdot \mathbf{u} &= 0. \end{aligned}$$

In these equations, ν , λ and σ are appropriate physical parameters.

To write down the variational formulation of these problems, let $V_X \times Q_X$ be the functional spaces where the pair $[\mathbf{u}, p]$ is sought, and let

$$a_X(\mathbf{u}, \mathbf{v}) = \begin{cases} a_S(\mathbf{u}, \mathbf{v}) := \nu(\nabla \mathbf{u}, \nabla \mathbf{v}) & \text{for the Stokes problem} \\ a_M(\mathbf{u}, \mathbf{v}) := \lambda(\nabla \times \mathbf{u}, \nabla \times \mathbf{v}) & \text{for the Maxwell problem} \\ a_D(\mathbf{u}, \mathbf{v}) := \sigma(\mathbf{u}, \mathbf{v}) & \text{for the Darcy problem} \end{cases}$$

Spaces V_X and Q_X are determined by requiring that

	V_X	Q_X	$\langle \nabla p, \mathbf{v} \rangle$
Stokes	$V_S = H_0^1(\Omega)^d$	$Q_S = L_0^2(\Omega)$	$-(p, \nabla \cdot \mathbf{v})$
Maxwell	$V_{M1} = H_0(\mathbf{curl}; \Omega)$	$Q_{M1} = H_0^1(\Omega)$	$+(\nabla p, \mathbf{v})$
	$V_{M2} = V_{M1} \cap H(\text{div}; \Omega)$	$Q_{M2} = L_0^2(\Omega)$	$-(p, \nabla \cdot \mathbf{v})$
Darcy	$V_{D1} = L^2(\Omega)^d$	$Q_{D1} = H^1(\Omega)$	$+(\nabla p, \mathbf{v})$
	$V_{D2} = H_0(\text{div}; \Omega)$	$Q_{D2} = L_0^2(\Omega)$	$-(p, \nabla \cdot \mathbf{v})$

Table 1: Functional setting

	$\ [\mathbf{v}, q]\ _X$
Stokes	$\nu^{\frac{1}{2}} \ \nabla \mathbf{v}\ + \nu^{-\frac{1}{2}} \ q\ $
Maxwell 1	$\lambda^{\frac{1}{2}} \ \nabla \times \mathbf{v}\ + \frac{\lambda^{\frac{1}{2}}}{L_0} \ \mathbf{v}\ + \lambda^{-\frac{1}{2}} \ q\ + \lambda^{-\frac{1}{2}} L_0 \ \nabla q\ $
Maxwell 2	$\lambda^{\frac{1}{2}} \ \nabla \times \mathbf{v}\ + \lambda^{\frac{1}{2}} \ \nabla \cdot \mathbf{v}\ + \frac{\lambda^{\frac{1}{2}}}{L_0} \ \mathbf{v}\ + \lambda^{-\frac{1}{2}} \ q\ $
Darcy, primal	$\sigma^{\frac{1}{2}} \ \mathbf{v}\ + \sigma^{-\frac{1}{2}} \ \nabla q\ $
Darcy, dual	$\sigma^{\frac{1}{2}} \ \mathbf{v}\ + \sigma^{\frac{1}{2}} L_0 \ \nabla \cdot \mathbf{v}\ + \frac{\sigma^{-\frac{1}{2}}}{L_0} \ q\ $

Table 2: Working norms

- $a_X(\mathbf{u}, \mathbf{v})$ is continuous.
- The term $\langle \nabla p, \mathbf{v} \rangle$, obtained by testing ∇p by \mathbf{v} , is well defined under the minimum regularity conditions.

In case of the Stokes problem, the first condition implies that \mathbf{u}, \mathbf{v} need to be in $H^1(\Omega)^d$, and thus the minimum regularity for p corresponds to take $\langle \nabla p, \mathbf{v} \rangle = -(p, \nabla \cdot \mathbf{v})$, which requires $p \in L^2(\Omega)$. However, for both the Maxwell and the Darcy problem we may choose either $\langle \nabla p, \mathbf{v} \rangle = -(p, \nabla \cdot \mathbf{v})$, $p \in L^2(\Omega)$, $\nabla \cdot \mathbf{v} \in L^2(\Omega)$, or $\langle \nabla p, \mathbf{v} \rangle = (\nabla p, \mathbf{v})$, $\nabla p \in L^2(\Omega)^d$, $\mathbf{v} \in L^2(\Omega)^d$. The possibilities for the functional setting of the different problems are summarized in Table 1. The choice of the functional setting has important practical consequences, both physical and numerical.

The norms in the product space $V_X \times Q_X$ depending on the choice of the functional setting are indicated in Table 2. All have been written to ensure a correct scaling. Note that the two possibilities for the Maxwell problem have been simply indicated as Maxwell 1 and 2, whereas for the Darcy problem they correspond to the well known primal and dual formulations [3].

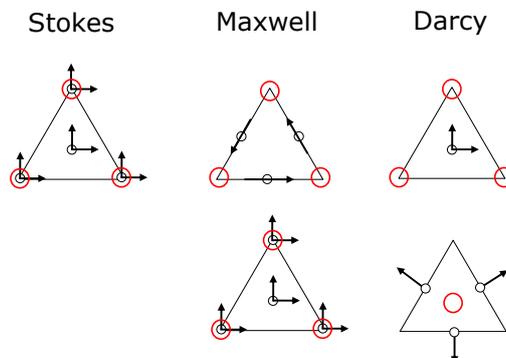


Figure 1: inf-sup stable elements

3 GALERKIN VS STABILIZED FINITE ELEMENT APPROXIMATIONS

When the Galerkin finite element approximation of the problem is attempted, the finite element spaces $V_h \subset V_X$ and $Q_h \subset Q_X$ need to satisfy the discrete counterpart of the inf-sup conditions that hold at the continuous level. These inf-sup conditions are different for the three problems considered, leading to different requirements for V_h and Q_h . The simplest inf-sup stable elements in 2D are schematically shown in Fig. 1, where nodes to interpolate the velocity components are printed in black and pressure nodes in red. For the Maxwell and the Darcy problems, the first row corresponds to the choice $p \in H^1(\Omega)$ and the second to $p \in L^2(\Omega)$.

There are several inconveniences in the use of the interpolations of Fig. 1. For example, if we consider a combined problem of the form

$$\begin{aligned} -\nu \Delta \mathbf{u} + \lambda \nabla \times \nabla \times \mathbf{u} + \sigma \mathbf{u} + \nabla p &= \mathbf{f}, \\ \nabla \cdot \mathbf{u} &= 0, \end{aligned}$$

it is clear that spaces satisfying the inf-sup condition for the Stokes problem need to be used if $\nu > 0$. However, we might be interested in letting $\nu \rightarrow 0$, or $\lambda \rightarrow 0$ or $\sigma \rightarrow 0$. From the numerical point of view, oscillations will show up if the correct interpolation is not used in each case.

Another inconvenience is faced in the case of a coupled problem, such as the MHD models discussed earlier. From the implementation point of view, it is much simpler to have all the unknowns at the same nodes of the finite element mesh.

As an alternative to the use of inf-sup stable elements, our approach is the use of stabilized formulations, in which *any conforming \mathbf{u} - p interpolation is allowed*. No stability problems will be found in the limits $\nu \rightarrow 0$, $\sigma \rightarrow 0$, $\lambda \rightarrow 0$ and, if Lagrangian interpolations

are used, coupling of different problems will be easy. For example, in contrast to the different interpolations shown in Fig. 1, it will be possible to use the simplest continuous P_1 interpolation for \mathbf{u} and p .

The key ingredients to design the stabilization methods presented in the following are

- A two scale decomposition of \mathbf{u} and p , within the variational multiscale framework (VMS).
- A proper scaling of the problem, which requires the introduction of a length scale.
- A closed form expression for the subscales based on an approximate Fourier analysis of the problem.

These ingredients will not be elaborated here. The methods proposed will be stated without (heuristic) derivation. For simplicity, we will take $\mathbf{f} \in L^2(\Omega)^d$.

4 STOKES' PROBLEM

This is the problem for which stabilized finite element methods are best known. Let us start writing the variational form of the problem, which is: find $[\mathbf{u}, p] \in H_0^1(\Omega)^d \times L_0^2(\Omega)$ such that

$$B_S([\mathbf{u}, p], [\mathbf{v}, q]) = L_S([\mathbf{v}, q]) = (\mathbf{f}, \mathbf{v}),$$

for all $[\mathbf{v}, q]$, where

$$B_S([\mathbf{u}, p], [\mathbf{v}, q]) = \nu(\nabla \mathbf{u}, \nabla \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) + (q, \nabla \cdot \mathbf{u}),$$

The stabilized finite element approximation we propose is: find $[\mathbf{u}_h, p_h] \in V_h \times Q_h$ such that

$$B_{S,h}([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) = L_{S,h}([\mathbf{v}_h, q_h]) \quad \forall [\mathbf{v}_h, q_h] \in V_h \times Q_h,$$

where $B_{S,h}$ and $L_{S,h}$ depend on the particular stabilized formulation. In particular, for the so called Algebraic Subgrid Scale (ASGS) method $B_{S,h}$ and $L_{S,h}$ are given by:

$$\begin{aligned} B_{S,h}([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) &= B_S([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) + \sum_K \tau_p \langle \nabla \cdot \mathbf{u}_h, \nabla \cdot \mathbf{v}_h \rangle_K \\ &\quad + \sum_K \tau_u \langle -\nu \Delta \mathbf{u}_h + \nabla p_h, \nu \Delta \mathbf{v}_h + \nabla q_h \rangle_K, \\ L_{S,h}([\mathbf{v}_h, q_h]) &= (\mathbf{f}, \mathbf{v}_h) + \sum_K \tau_u \langle \mathbf{f}, \nu \Delta \mathbf{v}_h + \nabla q_h \rangle_K, \end{aligned}$$

where τ_p and τ_u are the stabilization parameters, that we compute as

$$\tau_p = c_1 \nu, \quad \tau_u = (c_1 \nu)^{-1} h^2,$$

with c_1 an algorithmic constant and h the element size of the mesh, which we consider quasi-uniform for simplicity.

For the Orthogonal Subgrid Scales (OSS) method $B_{S,h}$ and $L_{S,h}$ are given by:

$$\begin{aligned} B_{S,h}([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) &= B_S([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) \\ &+ \sum_K \tau_p \langle P_h^\perp(\nabla \cdot \mathbf{u}_h), P_h^\perp(\nabla \cdot \mathbf{v}_h) \rangle_K \\ &+ \sum_K \tau_u \langle P_h^\perp(-\nu \Delta \mathbf{u}_h + \nabla p_h), P_h^\perp(\nu \Delta \mathbf{v}_h + \nabla q_h) \rangle_K, \\ L_{S,h}([\mathbf{v}_h, q_h]) &= (\mathbf{f}, \mathbf{v}_h), \end{aligned}$$

where P_h^\perp is the projection orthogonal to the finite element space and τ_p and τ_u are the same as for the ASGS method.

The numerical analysis of both the ASGS and the OSS methods shows that they have the same stability and convergence properties. Let us define the mesh dependent norm:

$$\|[\mathbf{v}_h, q_h]\|_{S,h}^2 = \nu \|\nabla \mathbf{v}_h\|^2 + \frac{1}{\nu} \|q_h\|^2 + \frac{h^2}{\nu} \sum_K \|\nabla q_h\|_K^2.$$

We also define the error function

$$E_S^2(h) = \nu \varepsilon_1^2(\mathbf{u}) + \frac{1}{\nu} \varepsilon_0^2(p),$$

where $\varepsilon_i(\cdot)$ denotes the interpolation error in the $H^i(\Omega)$ -seminorm. We then have:

Theorem 4.1 (Stability) *Suppose that the constant c_1 is large enough. Then, there exists a constant $C > 0$ such that*

$$\inf_{[\mathbf{u}_h, p_h]} \sup_{[\mathbf{v}_h, q_h]} \frac{B_{S,h}([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h])}{\|[\mathbf{u}_h, p_h]\|_{S,h} \|[\mathbf{v}_h, q_h]\|_{S,h}} \geq C > 0.$$

Theorem 4.2 (Convergence) *Let $[\mathbf{u}, p]$ be the solution of the continuous problem and $[\mathbf{u}_h, p_h]$ the solution of the discrete one. Suppose that c_1 is large enough. Then*

$$\|[\mathbf{u} - \mathbf{u}_h, p - p_h]\|_{S,h} \lesssim E_S(h).$$

5 MAXWELL'S PROBLEM

The variational formulation of Maxwell's problem can be written as: find $[\mathbf{u}, p] \in V \times Q$ such that

$$B_M([\mathbf{u}, p], [\mathbf{v}, q]) = L_M([\mathbf{v}, q]) = (\mathbf{f}, \mathbf{v}), \quad (9)$$

for all $[\mathbf{v}, q]$, where

$$B_M([\mathbf{u}, p], [\mathbf{v}, q]) = \lambda(\nabla \times \mathbf{u}, \nabla \times \mathbf{v}) + \langle \nabla p, \mathbf{v} \rangle - \langle \nabla q, \mathbf{u} \rangle,$$

and $\langle \nabla q, \mathbf{v} \rangle$ has two possible expressions according to the functional setting chosen:

$$\text{Formulation M1 : } \langle \nabla q, \mathbf{v} \rangle = (\nabla q, \mathbf{v}), \quad V = H_0(\mathbf{curl}, \Omega), \quad Q = H_0^1(\Omega).$$

$$\text{Formulation M2 : } \langle \nabla q, \mathbf{v} \rangle = -(q, \nabla \cdot \mathbf{v}), \quad V = H_0(\mathbf{curl}, \Omega) \cap H(\text{div}, \Omega), \quad Q = L^2(\Omega).$$

We will refer to M1 as the *curl formulation* and to M2 as the *curl-div formulation*. The main theoretical interest of Maxwell's problem is that there are solutions that can be found approximating M1 but not from the approximation of M2. These are the so called *singular solutions*. This fact, known as *the corner paradox*, follows from the following results:

Lemma 5.1 *If Ω is not convex, $H^1(\Omega)^d$ is a closed proper subspace of $H(\mathbf{curl}, \Omega) \cap H(\text{div}, \Omega)$ (all with tangential boundary conditions).*

Corollary 5.1 *If Ω is not convex, \mathbf{u} is the solution of (9) and \mathbf{u}_h its finite element approximation, then*

$$\lim_{h \rightarrow 0} \|\mathbf{u} - \mathbf{u}_h\|_{H(\mathbf{curl}, \Omega) \cap H(\text{div}, \Omega)} \neq 0,$$

in general.

One could argue whether this Lemma implies an approximability problem when using standard Lagrangian finite elements. The reason is that it can be shown that

$$\text{If } \mathbf{u}_h \in H^1(\Omega)^d \Rightarrow \|\nabla \mathbf{u}_h\| \lesssim \|\nabla \times \mathbf{u}_h\| + \|\nabla \cdot \mathbf{u}_h\|.$$

Thus, if \mathbf{u}_h is C^0 , its H^1 norm will be bounded, and thus \mathbf{u}_h will converge to a function in $H^1(\Omega)^d$, whereas the exact solution may belong to $H(\mathbf{curl}, \Omega) \cap H(\text{div}, \Omega)$.

The problem however cannot be attributed to C^0 interpolations, but to the curl-div formulation. It is not true that C^0 spaces cannot approximate the solution to (9). This only happens if, for some reason, $\nabla \cdot \mathbf{u}_h$ happens to be uniformly bounded. Thus, we have proposed a stabilized finite element method using C^0 spaces but able to reproduce the curl formulation [1]. This reads as follows: find $[\mathbf{u}_h, p_h] \in V_h \times Q_h$ such that

$$B_{M,h}([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) = L_{M,h}([\mathbf{v}_h, q_h]) \quad \forall [\mathbf{v}_h, q_h] \in V_h \times Q_h,$$

where

$$\begin{aligned} B_{M,h}([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) &= B_M([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) \\ &+ \sum_K \tau_p \langle \tilde{P}(\nabla \cdot \mathbf{u}_h), \tilde{P}(\nabla \cdot \mathbf{v}_h) \rangle_K \\ &+ \sum_K \tau_u \langle \tilde{P}(\nabla p_h), \tilde{P}(\nabla q_h) \rangle_K, \\ L_{M,h}([\mathbf{v}_h, q_h]) &= (\mathbf{f}, \mathbf{v}_h) + \sum_K \tau_u \langle \tilde{P}(\mathbf{f}), \tilde{P}(\nabla q_h) \rangle_K, \end{aligned}$$

and where

$$\tilde{P} = \begin{cases} I & \text{for the ASGS method} \\ P_h^\perp & \text{for the OSS method} \end{cases}$$

The stabilization parameters are given by:

$$\tau_p = c_2 \lambda \frac{h^2}{\ell^2}, \quad \tau_u = \frac{\ell^2}{\lambda},$$

where

$$\ell = \begin{cases} L_0 \text{ (characteristic length)} & \text{for the curl formulation (M1)} \\ h & \text{for the curl-div formulation (M2)} \end{cases}$$

It can be shown that *it is possible to switch from the functional setting M1 to M2 just by a proper scaling of the stabilization parameters.*

The numerical analysis we have performed shows that the formulations proposed are stable and optimally convergent in the norm

$$\|[\mathbf{v}_h, q_h]\|_{M,h} = \lambda^{\frac{1}{2}} \|\nabla \times \mathbf{v}_h\| + \lambda^{\frac{1}{2}} \frac{h}{\ell} \|\nabla \cdot \mathbf{v}_h\| + \frac{\ell}{\lambda^{\frac{1}{2}}} \|\nabla q_h\|.$$

Note that

- $\ell = L_0$ yields the discrete $H(\mathbf{curl}, \Omega) \times H^1(\Omega)$ norm.
- $\ell = h$ yields the discrete $H(\mathbf{curl}, \Omega) \cap H(\text{div}, \Omega) \times L^2(\Omega)$ norm.

Moreover, when $\ell = L_0$ if the continuous solution is singular ($\mathbf{u} \in H^r(\Omega)^d, r < 1$), $\mathbf{u}_h \rightarrow \mathbf{u}$ in H^r .

6 DARCY'S PROBLEM

As for the Maxwell problem, there are two possible functional settings, now called primal and dual. The variational formulation of the primal problem reads: find $\mathbf{u} \in L^2(\Omega)^d$ and $p \in H^1(\Omega)$ such that

$$\begin{aligned} \sigma(\mathbf{u}, \mathbf{v}) + (\nabla p, \mathbf{v}) &= (\mathbf{f}, \mathbf{v}), & \forall \mathbf{v} \in L^2(\Omega)^d, \\ -(\nabla q, \mathbf{u}) &= (g, q), & \forall q \in H^1(\Omega), \end{aligned}$$

whereas for the dual problem the variational formulation is: find $\mathbf{u} \in H_0(\text{div}; \Omega)$ and $p \in L_0^2(\Omega)$ such that

$$\begin{aligned} \sigma(\mathbf{u}, \mathbf{v}) - (p, \nabla \cdot \mathbf{v}) &= (\mathbf{f}, \mathbf{v}), & \forall \mathbf{v} \in H_0(\text{div}; \Omega), \\ (q, \nabla \cdot \mathbf{u}) &= (g, q), & \forall q \in L_0^2(\Omega). \end{aligned}$$

A forcing term g has been included in the continuity equation for generality. Let B_D and L_D the appropriate forms of the problem, written as find $[\mathbf{u}, p] \in V_D \times Q_D$ such that

$$B_D([\mathbf{u}, p], [\mathbf{v}, q]) = L_D([\mathbf{v}, q]) \quad \forall [\mathbf{v}, q] \in V_D \times Q_D.$$

The stabilized approximations that we propose read as follows: find $[\mathbf{u}_h, p_h] \in V_{D,h} \times Q_{D,h}$ such that

$$B_{D,h}([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) = L_{D,h}([\mathbf{v}_h, q_h]) \quad \forall [\mathbf{v}_h, q_h] \in V_{D,h} \times Q_{D,h}$$

where

$$\begin{aligned} B_{D,h}([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) &= B_D([\mathbf{u}_h, p_h], [\mathbf{v}_h, q_h]) \\ &+ \sum_K \tau_p \langle \tilde{P}(\nabla \cdot \mathbf{u}_h), \tilde{P}(\nabla \cdot \mathbf{v}_h) \rangle_K \\ &+ \sum_K \tau_u \langle \tilde{P}(\sigma \mathbf{u}_h + \nabla p_h), \tilde{P}(-\sigma \mathbf{v}_h + \nabla q_h) \rangle_K \\ L_{D,h}([\mathbf{v}_h, q_h]) &= (\mathbf{f}, \mathbf{v}_h) + (g, q_h) + \sum_K \tau_p \langle \tilde{P}(g), \tilde{P}(\nabla \cdot \mathbf{v}_h) \rangle_K \\ &+ \sum_K \tau_u \langle \tilde{P}(\mathbf{f}), \tilde{P}(-\sigma \mathbf{v}_h + \nabla q_h) \rangle_K, \end{aligned}$$

and where \tilde{P} is defined as for Maxwell's problem. The stabilization parameters are computed as

$$\tau_p = c_3 \sigma \ell^2, \quad \tau_u = (c_3 \sigma \ell^2)^{-1} h^2$$

where

$$\ell = \begin{cases} h & \text{for the primal formulation (D1).} \\ L_0 \text{ (characteristic length)} & \text{for the dual formulation (D2).} \end{cases}$$

We will be able to switch from the functional setting D1 to D2 just by a proper scaling of the stabilization parameters.

The numerical analysis indicates that the formulations proposed are stable and optimally convergent in the norm

$$\|[\mathbf{v}_h, q_h]\|_{D,h}^2 = \sigma \|\mathbf{v}_h\|^2 + \sigma \ell^2 \|\nabla \cdot \mathbf{v}_h\|^2 + \frac{1}{\sigma L_0^2} \|q_h\|^2 + \frac{h^2}{\sigma \ell^2} \sum_K \|\nabla q_h\|_K^2.$$

Note that

- $\ell = L_0$ yields the discrete $H(\text{div}, \Omega) \times L^2(\Omega)$ norm (dual problem).

Method	Primal mixed $k = 0, l = 1$	Stabilized $k = l = 1$	Dual mixed $k = 1, l = 0$
$\ \mathbf{e}_u\ $	h	h^2	h
$\ e_p\ $	h^2	h^2	h
$\ \nabla \cdot \mathbf{e}_u\ $	1	h	h
$\ \nabla e_p\ $	h	h	1

Table 3: Convergence order using the lowest order interpolations

- $\ell = h$ yields the discrete $L^2(\Omega)^d \times H^1(\Omega)$ norm (primal problem).
- $\ell = (hL_0)^{1/2}$ yields a norm that happens to be optimal for equal order interpolations.

Just as an example, in Table 3 we have indicated the convergence orders that can be found using the Galerkin method and inf-sup stable elements for the primal formulation (left column), for the dual formulation (right column) and using the stabilized method we propose (central column). In this table, k and l refer to the interpolation order of \mathbf{u} and p , respectively. The gain in accuracy using the stabilized formulation we propose is clear.

7 COUPLING AND CONCLUSIONS

Let us consider now the linearization of the stationary version of problems (1)-(4) and (5)-(8), restricting our attention to fixed-point-type schemes. Starting with the former, it can be shown that the only stable scheme is [4]:

$$\mathbf{a} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p - \frac{1}{\mu_m \rho} (\nabla \times \mathbf{B}) \times \mathbf{b} = \mathbf{f}_f, \quad (10)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (11)$$

$$\frac{1}{\mu_m \sigma} \nabla \times (\nabla \times \mathbf{B}) - \nabla \times (\mathbf{u} \times \mathbf{b}) + \nabla r = \mathbf{f}_m, \quad (12)$$

$$\nabla \cdot \mathbf{B} = 0. \quad (13)$$

where \mathbf{a} is the velocity and \mathbf{b} the magnetic field of the previous iteration. It is observed that all the variables *need* to be computed in a coupled way.

Let us move our attention to problem (5)-(8). Now it can be shown that the only stable scheme is [5]:

$$\mathbf{a} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p - \frac{1}{\rho} (\mathbf{j} \times \mathbf{B}) = \mathbf{f}_f, \quad (14)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (15)$$

$$\mathbf{j} + \sigma \nabla \phi - \sigma (\mathbf{u} \times \mathbf{B}) = 0, \quad (16)$$

$$\nabla \cdot \mathbf{j} = 0. \quad (17)$$

Remember that now \mathbf{B} is given. Again, all the variables have to be computed in a coupled way.

Let us conclude with some remarks related to the finite element approximation of problems (10)-(13) and (14)-(17). The first point to remark is that the stabilized finite element approximation that we propose allows the use of arbitrary, and in particular equal, interpolation for all the unknowns. This implies an important ease of implementation, since only one data structure (nodal connections, derivatives of shape functions) is required for all the variables. For example, when computing the element matrices arising from (10), contributions multiplying the arrays of \mathbf{u} , p and \mathbf{B} can be computed within the same loops, over the elements, over the nodes and over the integration points.

An aspect that we have not explored here is that the extension of the stabilized formulation to (10)-(13) and (14)-(17) allows to deal with combined problems and all the range of physical parameters. In particular, instabilities due to small viscosity (or large magnetic permeability or conductivity) are avoided.

A very important aspect of our formulation is the adaptation to the appropriate functional setting by a proper design of the scaling in the stabilization parameters. This is not a mathematical divertimento, but has important consequences. For example, for the magnetic field we may want or need to capture singular solutions. For problem (14)-(17) we may choose a better approximation for ϕ or for \mathbf{j} . Nevertheless, we have also shown that there is a version of the stabilization parameters that yields better accuracy that both the primal and the dual formulation of the Darcy problem of which ϕ and \mathbf{j} are solution.

REFERENCES

- [1] S. Badia and R. Codina. A nodal-based finite element approximation of the Maxwell problem suitable for singular solutions. *Submitted*.
- [2] S. Badia and R. Codina. Stokes, Maxwell and Darcy: a single finite element approximation for three model problems. *Submitted*.
- [3] S. Badia and R. Codina. Unified stabilized finite element formulations for the Stokes and the Darcy problems. *SIAM Journal on Numerical Analysis*, 47:1971–2000, 2009.
- [4] R. Codina and N. Hernández. Approximation of the thermally coupled MHD problem using a stabilized finite element method. *Journal of Computational Physics*, 230:1281–1303, 2011.
- [5] R. Planas, S. Badia, and R. Codina. Approximation of the inductionless MHD problem using a stabilized finite element method. *Journal of Computational Physics*, 230:2977–2996, 2011.

SPACE–TIME FSI MODELING AND DYNAMICAL ANALYSIS OF RINGSAIL PARACHUTE CLUSTERS

KENJI TAKIZAWA*, TIMOTHY SPIELMAN† AND TAYFUN E. TEZDUYAR†

*Department of Modern Mechanical Engineering and
Waseda Institute for Advanced Study, Waseda University
1-6-1 Nishi-waseda, Shinjuku-ku, Tokyo 169-8050, JAPAN

†Mechanical Engineering, Rice University – MS 321
6100 Main Street, Houston, TX 77005, USA

Key words: Fluid–structure interaction, Parachute clusters, Ringsail parachute, Space–time technique, Geometric porosity, Contact

Abstract. Computer modeling of ringsail parachute clusters involves fluid–structure interaction (FSI) between the parachute canopy and the air, geometric complexities created by the construction of the parachute from “rings” and “sails” with hundreds of gaps and slits, and the contact between the parachutes. The Team for Advanced Flow Simulation and Modeling (T★AFSM) has successfully addressed the computational challenges related to the FSI and geometric complexities, and recently started addressing the challenges related to the contact between the parachutes of a cluster. This is being accomplished with the Stabilized Space–Time FSI technique, which was developed and improved over the years by the T★AFSM and serves as the core numerical technology, and the special techniques developed by the T★AFSM to deal with the geometric complexities and the contact between parachutes. We present the results obtained with the FSI computation of parachute clusters and the related dynamical analysis.

1 INTRODUCTION

Computer modeling of large ringsail parachutes by the Team for Advanced Flow Simulation and Modeling (T★AFSM) was first reported in [1, 2]. The two major computational challenges successfully addressed were the fluid–structure interaction (FSI) between the parachute canopy and the air and the geometric complexities created by the construction of the “ringsail” parachute from “rings” and “sails” with hundreds of gaps and slits. These large parachutes are typically used in clusters of two or three parachutes. The contact between the parachutes creates another major challenge, which the T★AFSM started addressing very recently (see [3, 4]).

The core technology used in the T★AFSM parachute FSI computations (see, for example, [5, 1, 2, 6, 7, 3, 4]) is the stabilized space–time FSI (SSTFSI) technique, which was introduced in [5]. The SSTFSI technique is based on the new-generation Deforming-Spatial-Domain/Stabilized Space–Time (DSD/SST) formulations, which were also introduced in [5], increasing the scope and performance of the DSD/SST formulations developed earlier [8, 9, 10, 11] for computation of flows with moving boundaries and interfaces, including FSI. A number of special FSI techniques were introduced in [5, 1, 12, 6, 7, 3] in conjunction with the SSTFSI technique and the DSD/SST formulation. Many of these special techniques are in the category of interface projection techniques, such as the FSI Geometric Smoothing Technique (FSI-GST) [5], Separated Stress Projection (SSP) [1], Homogenized Modeling of Geometric Porosity (HMGP) [1], adaptive HMGP [6], “symmetric FSI” technique [6], accounting for fluid forces acting on structural components (such as parachute suspension lines) that are not expected to influence the flow [6], a new version of the HMGP that is called “HMGP-FG” [7], and other interface projection techniques [12]. These special FSI techniques address many of the challenges involved computer modeling of ringsail parachutes.

In FSI modeling of parachute clusters, the contact between the parachutes is one of the major challenges. In a contact algorithm we need for the parachute FSI computations, the objective is to prevent the structural surfaces from coming closer than a predetermined minimum distance we would like to maintain to protect the quality of the fluid mechanics mesh between the structural surfaces. The Surface-Edge-Node Contact Tracking (SENCT) technique was introduced in [5] for this purpose. It had two versions: SENCT-Force (SENCT-F) and SENCT-Displacement (SENCT-D). In the SENCT-F technique, which is the relevant version here, the contacted node is subjected to penalty forces that are inversely proportional to the projection distances to the contacting surfaces, edges and nodes. For FSI problems with incompatible fluid and structure meshes at the interface, it was proposed in Remark 1 of [1] to formulate the contact model based on the fluid mechanics mesh at the interface. This version of the SENCT was denoted with the option key “-M1”. The contact algorithm used in the parachute cluster computations reported in [3] has some features in common with the SENCT-F technique but is more robust. Also, compared to the SENCT-F technique, the forces are applied in a conservative fashion. For that reason, we call the new technique “SENCT-FC”, where the letter “C” stands for “conservative”. The new technique was used with option M1 in [3]. The SENCT-FC technique was described in detail in [4] and was used with option M1 also in the cluster computations reported in that article. This article is a short version of [4]. We present the computational results together with the related dynamical analysis.

2 CLUSTER COMPUTATIONS

The objective in these two-parachute cluster computations is to determine how the parameters representing the payload models and starting-conditions affect long-term cluster dynamics. The parachute clusters reported in this paper are used with a payload that is

about 19,000 lbs. Each parachute has 80 gores and a nominal diameter of about 120 ft. It has 4 rings and 9 sails. More information on the parachutes can be found in [2, 6, 7]. The parameters selected for testing are the payload-model configurations and initial coning angles (θ_{INIT}) and parachute diameters (D_{INIT}). We also investigate two scenarios to approximate the conditions immediately after parachute disreefing. This is explained in more detail in a later paragraph. In all cases, θ_{INIT} is the same for both parachutes.

The first set of computations investigates the effect of the payload model. In drop tests, the parachutes are connected to a rectangular pallet that is weighted to represent the mass and inertial properties of a proposed crew capsule. The preliminary parachute cluster computations reported in [3] modeled the payload as a point mass located at the confluence of the risers. We will refer to this as the payload at the confluence (PAC) configuration. Two new computational payload models were created to see how they would influence parachute behavior. The payload lower than the confluence (PLC) configuration adds another cable element below the confluence and models the payload as a point mass at the location of the pallet center of gravity. The payload as a truss element (PTE) configuration further enhances the model by distributing the payload mass at 9 different points to match the mass, center of gravity, and six components of the inertia tensor of the pallet. This is accomplished by adding 5 cable elements and 26 truss elements below the confluence. All of the payload comparison computations use $\theta_{\text{INIT}} = 35^\circ$.

The second set of computations investigates the effect of θ_{INIT} . Three values of θ_{INIT} are tested: 15° , 25° , and 35° . It should be noted that 35° is greater than the θ values seen in drop tests. The average θ during normal descent is around 15° , and the maximum θ does not usually exceed 25° . We use $\theta_{\text{INIT}} = 35^\circ$ only to cause a large perturbation in order to analyze the dynamic response of the parachute cluster. All of the θ_{INIT} comparison computations use the PTE configuration.

The parachute described in this paper uses a reefing technique to permit incremental opening of the canopy. The parachute skirt is initially constricted by reefing lines and the reefing lines are cut at predetermined time intervals to allow the canopy to “disreef” to larger diameters. In the third set of computations, we compute two scenarios to analyze how conditions immediately after disreefing could have an effect on long-term dynamics. In the first scenario, which we call “simulated disreef”, $\theta_{\text{INIT}} = 10^\circ$, and for both parachutes $D_{\text{INIT}} = 70$ ft. These values represent the approximate θ during final disreefing and the average minimum D during nominal descent. The second scenario represents an “asynchronous disreef” by using, with $\theta_{\text{INIT}} = 35^\circ$, for one parachute $D_{\text{INIT}} = 70$ ft, and for the other $D_{\text{INIT}} = 90$ ft. These values represent the average minimum and maximum parachute diameters during nominal descent, respectively. Both scenarios use the PTE configuration.

2.1 Starting conditions

We first build a starting condition for a single parachute. We begin with a parachute shape obtained with the symmetric FSI computation reported in [6]. We do another

symmetric FSI computation with a horizontal inflow velocity of $24.0 \sin(\theta_{\text{INIT}})$ ft/s. This results in an angle of attack of θ_{INIT} , and we compute for three breathing cycles. We use the parachute shape and position corresponding to the time when the parachute skirt diameter is at its average value and assemble the cluster structural mechanics mesh with the parachutes at θ_{INIT} . After that we generate a fluid mechanics mesh. With the cluster mesh, holding the parachute shapes and positions fixed, we first do a fluid mechanics computation. The inflow velocity is 31.0 ft/s. Next, we do a fluid mechanics computation with a prescribed, time-dependent shape for both parachutes. The time-dependent shape comes from the single-parachute symmetric FSI computation carried out earlier at an angle of attack of θ_{INIT} . We use the solution from the fluid mechanics computation with prescribed parachute motion as the starting condition for the FSI computation.

2.2 Computational conditions

Figure 1 shows, for a single parachute, the canopy structure mesh and the fluid mechanics interface mesh. The fluid mesh is cylindrical with a diameter of 1,740 ft and a

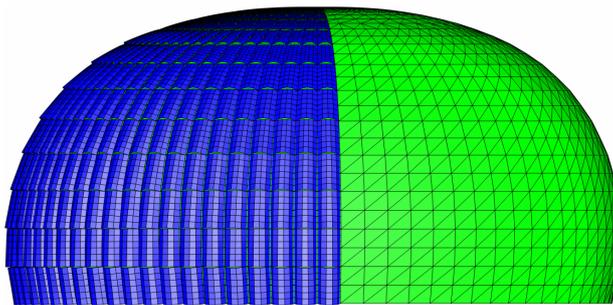


Figure 1: Canopy structure mesh (left) and fluid mechanics interface mesh (right) for a single parachute. The structure has 30,722 nodes, 26,000 four-node quadrilateral membrane elements, and 12,521 two-node cable elements. There are 29,200 nodes on the canopy. The fluid mechanics interface mesh has 2,140 nodes and 4,180 three-node triangular elements.

height of 1,566 ft. It has four-node tetrahedral elements. The fluid interface mesh has three-node triangular elements. The number of nodes and elements are given in Table 1.

All computations are carried out using air properties at standard sea-level conditions. The geometry and material properties are the same as they are described in Section 3.2 of [7]. In addition to moving our reference frame vertically with a reference descent speed, as originally proposed in [6], we move the mesh horizontally and vertically, with the average displacement rate for the structure. The horizontal motion of the mesh becomes particularly helpful when the parachute glides significantly. With a mesh that moves horizontally, we use the velocity form of the free-stream conditions also at the lateral boundaries. All computations are carried out in a parallel computing environment using PC clusters. The meshes are partitioned to enhance the parallel efficiency of the computations. Mesh partitioning is based on the METIS [13] algorithm. In solving

Structure		<i>nn</i>	61,443
	Membrane	<i>ne</i>	52,000
	Cable	<i>ne</i>	25,042
	Payload	<i>ne</i>	1
	Interface	<i>nn</i>	58,400
<i>ne</i>		52,000	
Fluid	Interface	<i>nn</i>	4,280
		<i>ne</i>	8,360
	Volume (15°, 80/80 ft)	<i>nn</i>	197,288
		<i>ne</i>	1,210,349
	Volume (25°, 80/80 ft)	<i>nn</i>	280,601
		<i>ne</i>	1,739,739
	Volume (35°, 80/80 ft)	<i>nn</i>	289,679
		<i>ne</i>	1,797,003
	Volume (10°, 70/70 ft)	<i>nn</i>	352,861
		<i>ne</i>	2,199,472
	Volume (35°, 70/90 ft)	<i>nn</i>	289,221
		<i>ne</i>	1,795,542

Table 1: Number of nodes and elements for the two-parachute clusters before any payload modifications. Here *nn* and *ne* are number of nodes and elements, respectively. The fluid mechanics volume mesh is tabulated for different combinations of θ_{INIT} and D_{INIT} values. The PLC configuration has 1 more structure node and 1 more cable element. The PTE configuration has 10 more structure nodes, 5 more cable elements, 26 more truss elements, and 8 more payload elements.

the linear equation systems involved at every nonlinear iteration, the GMRES search technique [14] is used with a diagonal preconditioner.

The stand-alone fluid mechanics computations are done in two parts. The first part uses the semi-discrete formulation given in [11]. We compute 1,000 time steps with a time-step size of 0.232 s and 7 nonlinear iterations per time step. The number of GMRES iterations per nonlinear iteration is 90. The second part uses the DSD/SST-TIP1 technique (see Remark 5 in [5]), with the SUPG test function option WTSA (see Remark 2 in [5]). The stabilization parameters used are those given in [5] by Eqs. (9)–(12), (14)–(15) and (17), with the τ_{SUGN_2} term dropped from Eq. (14). The porosity model is HMGP-FG. We compute 600 time steps with a time-step size of 0.0232 s, 6 nonlinear iterations per time step, and 90 GMRES iterations per nonlinear iteration. For the fluid mechanics computations with prescribed, time-dependent shapes, we again use the DSD/SST-TIP1 technique, with the same SUPG test function option and stabilization parameters as those described above. We compute roughly 300 time steps with a time-step size of 0.0232 s, 6 nonlinear iterations per time step, and 90 GMRES iterations per nonlinear iteration.

For FSI computations, we use the SSTFSI-TIP1 technique (see Remarks 5 and 10 in [5]), again with the same SUPG test function option and stabilization parameters as those described above. The fully-discretized, coupled fluid and structural mechanics and mesh-moving equations are solved with the quasi-direct coupling technique (see Section 5.2 in [5]). The time-step size is 0.0232 s, and the number of nonlinear iterations per time step

is 6. The porosity model is HMGP-FG. We use SSP. We use selective scaling (see [5]), with the scale for the structure part set to 100. The SENC-T-FC contact algorithm is used with $\epsilon_A^S = \epsilon_A^C = 1.45$ ft (see [4] for the terminology), which is approximately equal to the radial distance between the valley nodes and the outermost part of the sails at the parachute skirt. The number of GMRES iterations per nonlinear iteration is for most of the time steps 140 for the fluid+structure block, and 30 for the mesh-moving block. When the parachutes are close to each other, the number of GMRES iterations per nonlinear iteration for the fluid+structure block is increased as needed to control the residuals, especially those corresponding to the structural mechanics part. The maximum number of GMRES iterations used per nonlinear iteration for the fluid+structure block is 1,400.

We compute each parachute cluster for a total of about 75 s, and we remesh as needed to preserve mesh quality. The frequency of remeshing varies for each computation and usually depends on how often the parachutes collide, how much the cluster rotates about the vertical axis, and how much each parachute rotates about its own axis. Depending on the computation, remeshing is needed every 170 to 370 time steps.

2.3 Results

The critical measure of performance for the parachute system described in this paper is the descent speed of the payload. The maximum payload descent speed ultimately determines if the system meets mission requirements. Another common measure of performance is the drag coefficient, which is calculated as $C_D = W/(\frac{1}{2}\rho U^2 S_o)$. Here W is the payload weight, ρ is the density of the air, U is the payload descent speed, and S_o is the nominal area of the parachute. Figures 2–5 show the computational results for the parachute cluster computations.

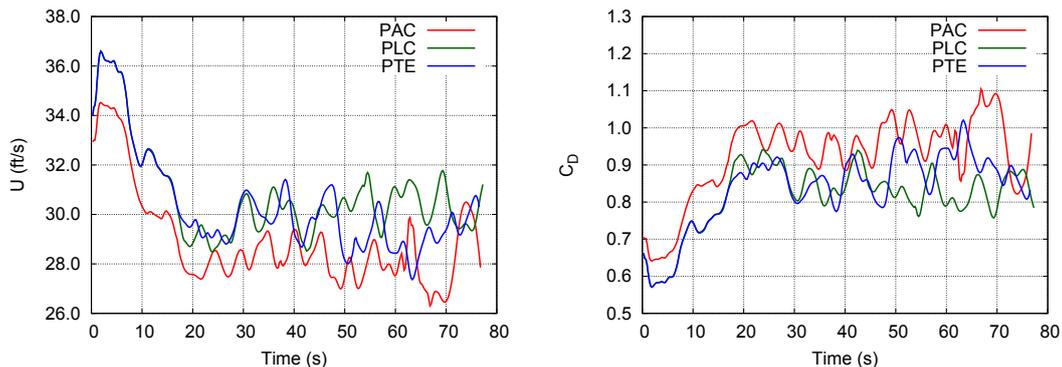


Figure 2: Cluster computations for different payload models and $\theta_{\text{INIT}} = 35^\circ$.

The geometry of parachute clusters usually forces individual parachutes to fly at angles of attack that are higher than the angle of attack at which they would fly as single parachutes. If the forced angle of attack in the cluster is not a stable one for the parachutes,

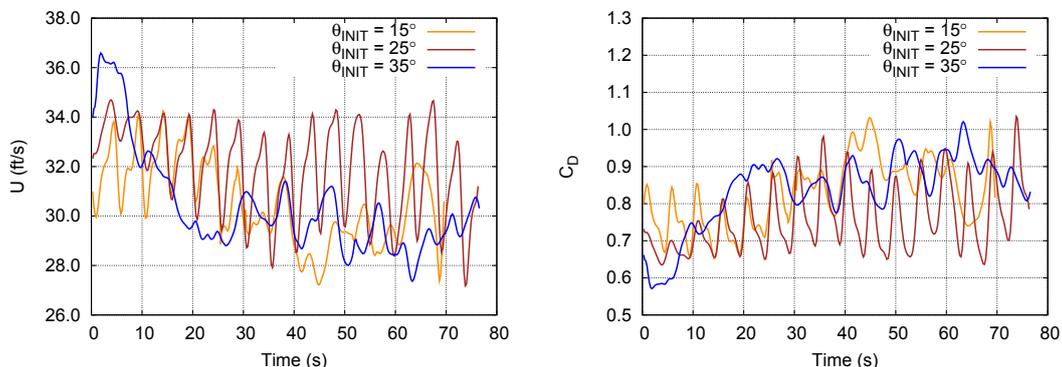


Figure 3: Cluster computations for PTE and different θ_{INIT} values.

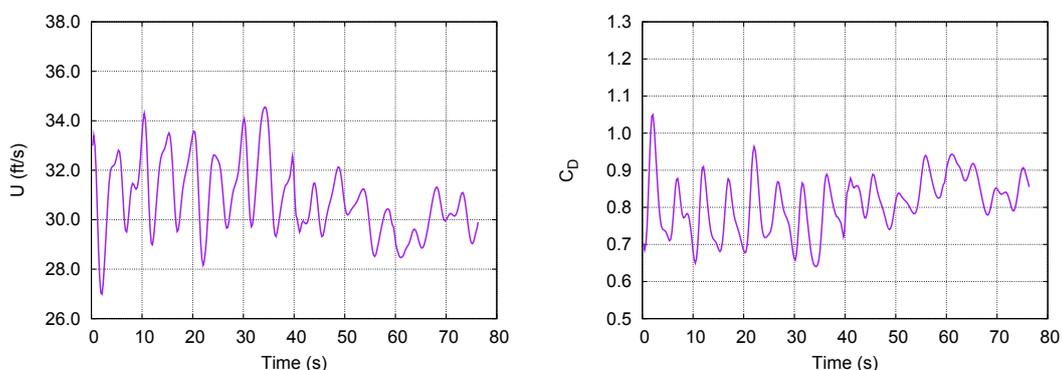


Figure 4: Cluster computations for simulated disreef.

they tend to collide with each other as they attempt to reach an angle of attack that is stable. Figures 6–7 show the contact between two parachutes from the asynchronous-disreef computation. Parachute clusters often experience reductions in drag due to this mutual interference between parachutes. The oscillatory motion of parachutes in the cluster and the frequency of collisions between parachutes can be used to characterize cluster stability. Figures 8–11 show the vent-separation distance (“ L_{VS} ”) for all cluster computations. The horizontal black line on each plot shows the approximate vent-separation distance when the parachutes are in contact. Tables 2–4 summarize the payload descent speeds and drag coefficients for all of the cluster computations.

One of the goals of this computational analysis is to assist parachute design engineers in determining which factors contribute to the payload descent speed oscillations seen in drop tests. For example, collisions between parachutes are usually associated with increased payload descent speed, but this is not always true. Previous analyses have also noted some correlation between parachute coning angles and payload descent speed. However, the correlation between these parameters is not strong enough to conclude that coning angle is the only, or even the most important, factor. The payload descent speed

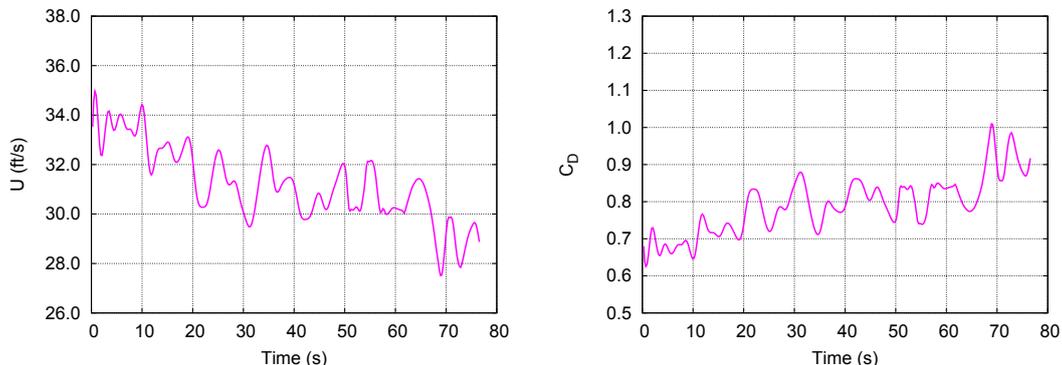


Figure 5: Cluster computations for asynchronous disreef.



Figure 6: Parachutes at $t = 52.20$ s, $t = 53.36$ s and $t = 54.52$ s during the asynchronous-disreef computation modeling the contact between parachutes.

is composed of several overlapping frequencies caused by various parachute dynamics. The overlap makes it difficult to determine which individual parachute behaviors and parachute cluster behaviors are contributing to changes in payload descent speed. In order to address this complex problem, we have developed a technique to decompose the payload descent speed into components. This technique is described in Section 4 in [4].

3 CONCLUDING REMARKS

We have presented our FSI computations of large ringsail parachute clusters. The core technology used is the SSTFSI technique, supplemented with a number of special FSI techniques, many of which were developed to address the computational challenges involved in computer modeling of ringsail parachutes. One of the special techniques, which was fully developed recently, addresses the major computational challenges created



Figure 7: Parachutes at $t = 55.68$ s, $t = 56.84$ s and $t = 58.00$ s during the asynchronous-disreef computation modeling the contact between parachutes.

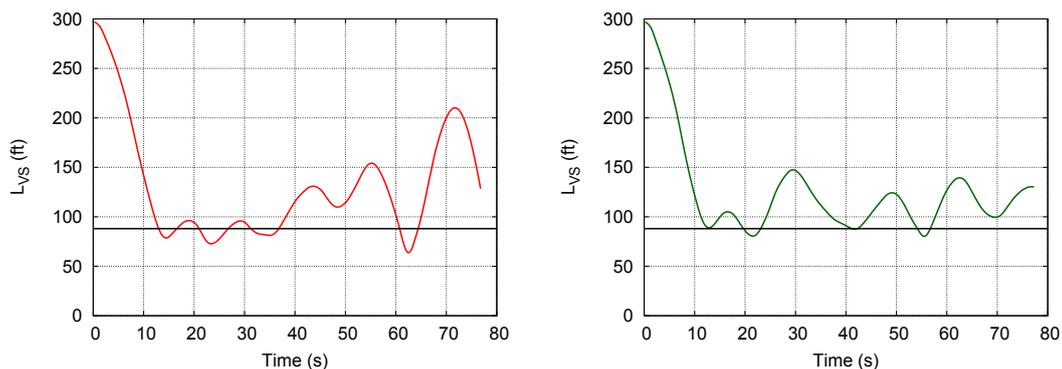


Figure 8: Vent-separation distance. Left: PAC and $\theta_{\text{INIT}} = 35^\circ$, Right: PLC and $\theta_{\text{INIT}} = 35^\circ$.

by the contact between the parachutes of a cluster. We have also presented a dynamical analysis of the computed results. Overall, we demonstrated that the core and special FSI techniques developed by the T★AFSM give us a powerful tool for computer modeling of ringsail parachute clusters.

ACKNOWLEDGMENT

This work was supported by NASA Grant NNX09AM89G, and also in part by the Rice Computational Research Cluster funded by NSF Grant CNS-0821727.

REFERENCES

- [1] T.E. Tezduyar, S. Sathe, J. Pausewang, M. Schwaab, J. Christopher, and J. Crabtree, “Interface projection techniques for fluid–structure interaction modeling with

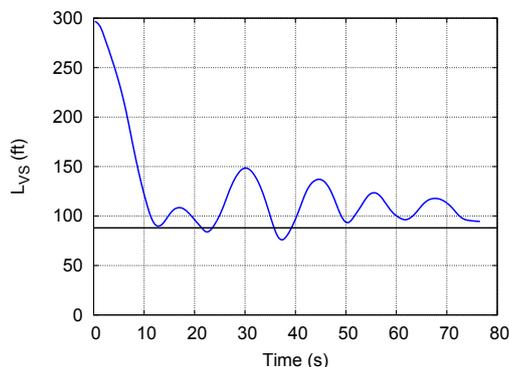


Figure 9: Vent-separation distance. PTE and $\theta_{\text{INIT}} = 35^\circ$.

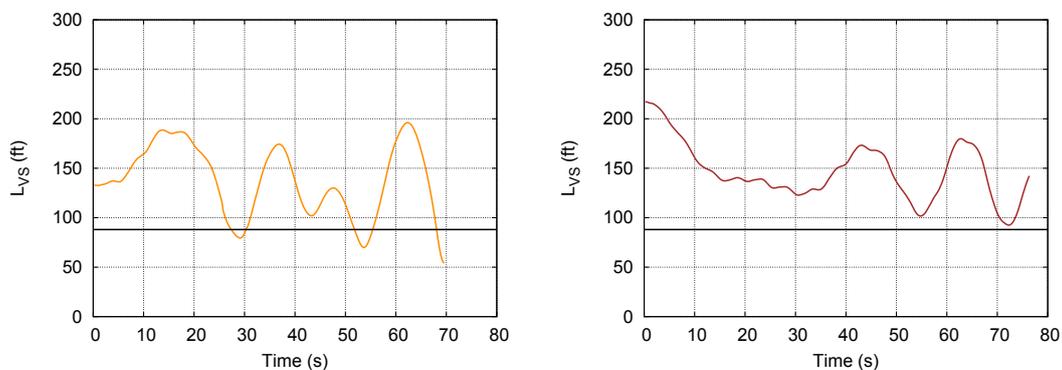


Figure 10: Vent-separation distance. Left: PTE and $\theta_{\text{INIT}} = 15^\circ$, Right: PTE and $\theta_{\text{INIT}} = 25^\circ$.

moving-mesh methods”, *Computational Mechanics*, **43** (2008) 39–49.

- [2] T.E. Tezduyar, S. Sathe, J. Pausewang, M. Schwaab, J. Christopher, and J. Crabtree, “Fluid–structure interaction modeling of ringsail parachutes”, *Computational Mechanics*, **43** (2008) 133–142.
- [3] K. Takizawa, S. Wright, C. Moorman, and T.E. Tezduyar, “Fluid–structure interaction modeling of parachute clusters”, *International Journal for Numerical Methods in Fluids*, **65** (2011) 286–307.

Payload Model	U (ft/s)	C_D
PAC	28.1	0.97
PLC	30.1	0.85
PTE	29.5	0.88

Table 2: Average U and C_D for different payload models with $\theta_{\text{INIT}} = 35^\circ$. Statistical analysis begins 20 s after the start of the computation.

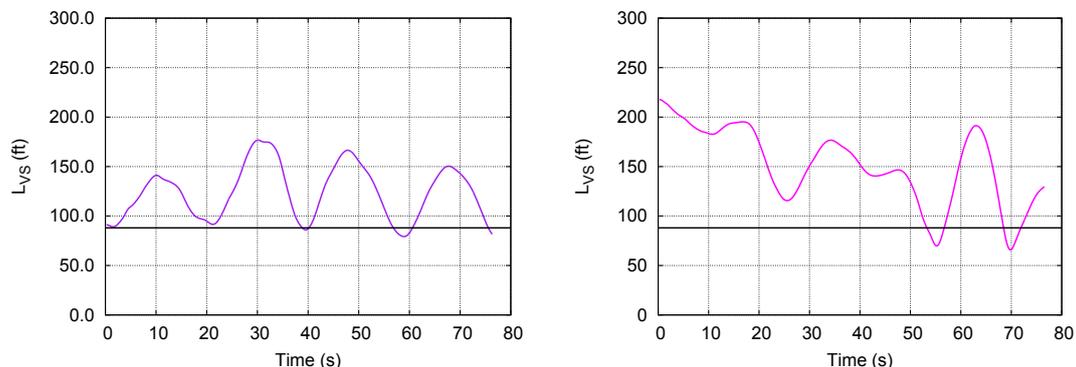


Figure 11: Vent-separation distance. Left: Simulated-disreef, Right: Asynchronous-disreef.

θ_{INIT}	U (ft/s)	C_D
15°	29.9	0.86
25°	31.4	0.78
35°	29.5	0.88

Table 3: Average U and C_D for PTE and different values of θ_{INIT} . Statistical analysis begins 20 s after the start of the computation.

- [4] K. Takizawa, T. Spielman, and T.E. Tezduyar, “Space–time FSI modeling and dynamical analysis of spacecraft parachutes and parachute clusters”, *Computational Mechanics*, to appear, 2011.
- [5] T.E. Tezduyar and S. Sathe, “Modeling of fluid–structure interactions with the space–time finite elements: Solution techniques”, *International Journal for Numerical Methods in Fluids*, **54** (2007) 855–900.
- [6] T.E. Tezduyar, K. Takizawa, C. Moorman, S. Wright, and J. Christopher, “Space–time finite element computation of complex fluid–structure interactions”, *International Journal for Numerical Methods in Fluids*, **64** (2010) 1201–1218.
- [7] K. Takizawa, C. Moorman, S. Wright, T. Spielman, and T.E. Tezduyar, “Fluid–structure interaction modeling and performance analysis of the Orion spacecraft parachutes”, *International Journal for Numerical Methods in Fluids*, **65** (2011) 271–285.

	U (ft/s)	C_D
Simulated Disreef	30.6	0.82
Asynchronous Disreef	30.8	0.81

Table 4: Average U and C_D for the disreef cases. Statistical analysis begins 5 s after the start of the computation for the simulated-disreef case, and 20 s after the start of the computation for the asynchronous-disreef case.

- [8] T.E. Tezduyar, “Stabilized finite element formulations for incompressible flow computations”, *Advances in Applied Mechanics*, **28** (1992) 1–44.
- [9] T.E. Tezduyar, M. Behr, and J. Liou, “A new strategy for finite element computations involving moving boundaries and interfaces – the deforming-spatial-domain/space–time procedure: I. The concept and the preliminary numerical tests”, *Computer Methods in Applied Mechanics and Engineering*, **94** (1992) 339–351.
- [10] T.E. Tezduyar, M. Behr, S. Mittal, and J. Liou, “A new strategy for finite element computations involving moving boundaries and interfaces – the deforming-spatial-domain/space–time procedure: II. Computation of free-surface flows, two-liquid flows, and flows with drifting cylinders”, *Computer Methods in Applied Mechanics and Engineering*, **94** (1992) 353–371.
- [11] T.E. Tezduyar, “Computation of moving boundaries and interfaces and stabilization parameters”, *International Journal for Numerical Methods in Fluids*, **43** (2003) 555–575.
- [12] K. Takizawa, C. Moorman, S. Wright, J. Christopher, and T.E. Tezduyar, “Wall shear stress calculations in space–time finite element computation of arterial fluid–structure interactions”, *Computational Mechanics*, **46** (2010) 31–41.
- [13] G. Karypis and V. Kumar, “A fast and high quality multilevel scheme for partitioning irregular graphs”, *SIAM Journal of Scientific Computing*, **20** (1998) 359–392.
- [14] Y. Saad and M. Schultz, “GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems”, *SIAM Journal of Scientific and Statistical Computing*, **7** (1986) 856–869.

A THREE-PHASE MODEL FOR DAMAGE INDUCED BY ASR IN CONCRETE STRUCTURES

CLAUDIA COMI, ROSSELLA PIGNATELLI

Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano
claudia.comi@polimi.it, pignatelli@stru.polimi.it

Key words: Damage, ASR, Humidity, Concrete

Abstract. In the present work we propose a three-phase chemo-elastic damage model for the description of the mechanical degradation of concrete induced by the Alkali Silica Reaction (ASR). The model accounts for ASR kinetics, depending on temperature and humidity conditions, and the swelling pressure exerted by the ASR reaction products on the skeleton. The model is validated on the basis of experimental tests on concrete reactive beams performed at varying ambient humidity conditions.

1 INTRODUCTION

The alkali-silica reaction occurring in concrete composed by certain type of aggregates and cement paste forms a hydrous alkali-calcium-silica gel. In the presence of water this gel swells, causing expansion and micro-cracking. Water in concrete is always present in little quantities as the rest of the process of hydration, but environmental conditions can change the water content.

The structural consequences of this phenomenon can be very serious, especially in massive concrete structure such as dams or bridges built some decades ago. For this reason in recent years many experimental campaigns has been performed to assess the influence of environmental factors, such as temperature and humidity conditions, on the ASR development [1, 2] and several mathematical models have been proposed to simulate the mechanical effects of ASR. In early proposed models [3, 4] the gel expansion was taken into account as an incompatible strain. Subsequently, more refined models have been proposed (see e.g. [5, 6]), with focus on the kinetics of the reaction. In [6] the kinetics law, based on Larive's proposal, also includes, in a heuristic way, the effect of the stress state. Other models address the mechanical modeling of ASR by considering a two-phase material with the concrete skeleton and the expansive gel acting in parallel [7, 8, 9, 10]. Very recently a model was developed taking into account concrete creep, the stress induced by the formation of ASR gel and the mechanical damage [11]. More in general multi-phase models have been developed to model other chemo-physical phenomena in concrete, like e.g. calcium leaching [12].

In the present work the phenomenological bi-phase isotropic damage model proposed in [13] is extended in order to catch the structural effects induced by the alkali-silica reaction (ASR) in concrete structures when ambient humidity conditions change and moisture gradients occur. The quantity of water present in the structure is taken into account through the degree of saturation of concrete and, in the framework of Biot's theory of multi-phase

porous materials, concrete is here conceived as a three-phase material constituted by the superposition of three homogenized phases: the concrete skeleton, the gel produced by the chemical reaction and the water. The assumption of three different phases allows to take into account also the effects of shrinkage at difference from the two-phase model previously proposed by the Authors [14].

The ASR kinetic is described by using an intrinsic time (as proposed in [1]), expressed as a function of the latency time and the characteristic time depending on both temperature and degree of saturation. When external humidity boundary conditions are not homogeneous, the moisture gradients are taken into account through a diffusion analysis of moisture in its liquid form. The mechanical part of the model is based on the damage theory in order to assess the decrease of stiffness of the concrete due to cracking caused by ASR.

The proposed model has been validated through the comparison with experimental results in [15] for free expansion tests on unconstrained cylindrical specimens and plain concrete beams.

2 CHEMO-DAMAGE MODEL

2.1 Field equations

The present model for concrete affected by ASR is formulated within the theory of multi-phase porous materials [16]. At the meso-scale three different phases are considered: the concrete skeleton, (*s*) including cement paste and aggregates, the gel (*g*) and the water (*w*) present in the connected porosity.

In the proposed model gel and water are considered as distinct phases, occupying two distinct and not interacting porosities. This hypothesis is in accordance to the theoretical models based on surface-chemistry principles, that attribute the expansion of the product gel in the presence of water to an electrical double-layer repulsion (see e.g. [17]), similarly to what happens for clays (see e.g. [18]).

Let us consider a representative volume element RVE of total volume V and denote by V_s, V_g, V_w the volume occupied by solid, gel and water phases, and by $V_v = V_{vg} + V_{vw}$ the void volume, formally divided into the volume of voids V_{vg} and V_{vw} , which can be filled by the gel and by the water respectively. The total porosity $\varphi = V_v/V$ is thus divided into two non interacting porosities $\varphi_g = V_{vg}/V$ and $\varphi_w = V_{vw}/V$, and the degrees of saturation for the gel and the water are defined as $S_g = V_g/V_{vg}$ and $S_w = V_w/V_{vw}$. With the above definitions, the volumetric fractions of gel and water are expressed as

$$\zeta_g = V_g/V = \varphi_g S_g, \quad \zeta_w = V_w/V = \varphi_w S_w \quad (1)$$

Petrographic investigations show that ASR gel forms between grains and cement past consuming the silica particles necessary for the reaction, so we assume that the gel porosity evolves with the gel formation and be always filled by the gel, so that $S_g=1$. The very low permeability of concrete with respect to gel than to water justifies the assumption that the forming gel does not fill the connected porosity of concrete, initially available for water.

Assuming small strains and quasi-static conditions the compatibility, equilibrium and mass balance equations for the multi-phase solid read

$$\boldsymbol{\varepsilon} = \frac{1}{2} (\text{grad} \mathbf{u}_s + \text{grad}^T \mathbf{u}_s) \quad (2)$$

$$\text{div} \boldsymbol{\sigma} + \rho \mathbf{b} = 0 \quad (3)$$

$$\text{div} (m_\beta \dot{\mathbf{w}}_\beta) + \frac{\partial m_\beta}{\partial t} = \pm \gamma_{\rightarrow\beta} \quad \beta = s, g, w \quad (4)$$

where: $\boldsymbol{\varepsilon}$ is the total strain in the skeleton, \mathbf{u}_s is the skeleton displacement, $\boldsymbol{\sigma}$ is the Cauchy stress in the combined solid and fluid mix, $\rho = (1-\phi)\rho_s + \sum \phi_\alpha \rho_\alpha S_\alpha$ is the density of the assembly, ρ_β is the mass density of phase β , $\rho \mathbf{b}$ is the body force of the solid and fluid mix, m_β is the mass variation of phase β , $\mathbf{w}_\alpha = \phi_\alpha S_\alpha (\mathbf{u}_\alpha - \mathbf{u}_s)$, $\alpha = g, w$, is pseudo-displacement of the fluid phase α relative to the skeleton, $\mathbf{w}_s = \mathbf{0}$ and $\gamma_{\rightarrow\beta}$ is the rate of mass formation $+\gamma_{\rightarrow\beta}$ or consumption $-\gamma_{\rightarrow\beta}$ for the phase β which coincides with the rate of mass exchanged among phase β and other two phases.

The mass variation m_α of the fluid phase α can be expressed as the product of the mass density ρ_α times the volumetric fractions of gel and water ζ_α ($\alpha = g, w$) which can be interpreted as the variation of fluid phase content with respect to initial conditions, i.e. the volume change of fluid phase per unit total volume:

$$m_\alpha = \rho_\alpha \zeta_\alpha = \rho_\alpha \phi_\alpha S_\alpha, \quad \alpha = g, w \quad (5)$$

In order to adapt the general equations of mixture theory to our problem, some simplifying assumptions can be introduced at the meso-scale. First of all the rate of solid mass dissolved by water and intervening in the ASR reaction can be neglected in the solid mass balance equation, while its contribution is relevant for the evolution of gel mass. Moreover the rate of mass of water absorbed by the gel is negligible with respect to the concentration of water in the porosity and the relative velocity of gel with respect to solid skeleton $\dot{\mathbf{w}}_g$ is negligible with respect to the relative velocity of water $\dot{\mathbf{w}}_w$. Finally, the rate of mass of water transformed into vapor can be neglected because of the small porosity of concrete. With these approximations, using eq.(5), the conservation laws (4) for the gel the liquid and water can be written as

$$\frac{\partial (\rho_g \phi_g S_g)}{\partial t} = \gamma_{s \rightarrow g} \quad (6)$$

$$\text{div} (\rho_w \phi_w S_w \dot{\mathbf{w}}_w) + \frac{\partial (\rho_w \phi_w S_w)}{\partial t} = 0 \quad (7)$$

Proper transport law should be specified for the fluid phases. As already remarked the low permeability of concrete with respect to gel allows to neglect the transport of gel. For the water phase, we adopt the simplified equation of moisture transport proposed in [19], valid for slightly porous material and obtained by the combination of the Darcy's law for fluid flow in porous media with conservation law (7). The Darcy's law reads:

$$\varphi_w S_w \dot{\mathbf{w}}_w = -\frac{K}{\eta_w} k_{rw}(S_w) \text{grad}(p_w) \quad (8)$$

where K , η_w , k_{rw} and p_w denote respectively intrinsic permeability of concrete (independent from the saturating fluid), dynamic viscosity, relative permeability and pressure of water. Substitution of (8) into the conservation law (7) leads to the following non-linear transport law for moisture in its liquid form

$$\varphi_w \frac{\partial S_w}{\partial t} + \text{div}(D_w(S_w) \text{grad} S_w) = 0 \quad (9)$$

where $D_w(S_w)$ is the permeability of concrete, dependent on the degree of saturation. By taking into account the approximations introduced in the conservation laws (6) and (7), the expression introduced in [21] for the relative permeability $k_{rw}(S_w)$ and the procedure proposed in [20] the following expression for the permeability of concrete can be obtained:

$$D_w(S_w) = m_1 m_2 \left(1 - \frac{1}{m_2}\right) \frac{K}{\eta_w} \frac{\left[1 - (1 - S_w^{1/m_3})^{m_3}\right]^2}{S_w^{1/2+m_2} (S_w^{-m_2} - 1)^{1/m_2}} \quad (10)$$

where m_1 , m_2 and m_3 are assumed as in [20].

2.2 Constitutive equations

The state equations relating the static variables (total Cauchy stress $\boldsymbol{\sigma}$, chemical potential of each fluid phase μ_α and entropy \mathcal{S}) to the conjugate kinematic variables (strain of the skeleton $\boldsymbol{\varepsilon}$, variation of fluid phase contents ζ_α and temperature variation $\theta = T - T_0$, being T_0 the local reference temperature) are derived from the free energy potential Ψ . In order to model the concrete skeleton degradation induced by the ASR, we also introduce in the free energy an internal damage variable D . This damage variable affects the solid skeleton elastic behavior and depends on two scalar variables, D_t and D_c , which describe damage mechanisms under prevailing tension and compression conditions: $D = 1 - (1 - D_t)(1 - D_c)$.

The following explicit expression for the free energy is proposed:

$$\begin{aligned} \Psi &= \Psi(\boldsymbol{\varepsilon}, \theta, \zeta_w, \zeta_g, D) \\ &= \frac{1}{2}(1-D) \left[2G\boldsymbol{\varepsilon} : \mathbf{e} + K\text{tr}^2\boldsymbol{\varepsilon} + M_g b_g^2 \left(\text{tr}\boldsymbol{\varepsilon} - \frac{\zeta_g}{b_g} \right)^2 + M_w b_w^2 \left(\text{tr}\boldsymbol{\varepsilon} - \frac{\zeta_w}{b_w} \right)^2 \right] + \\ &+ (1-D) \left[\frac{1}{2} \left(-\frac{C}{T_0} + M_g \alpha_g^2 + M_w \alpha_w^2 \right) \theta^2 - K\text{tr}\boldsymbol{\varepsilon} \alpha \theta \right] + \\ &+ (1-D) \left[-M_g b_g \left(\text{tr}\boldsymbol{\varepsilon} - \frac{\zeta_g}{b_g} \right) \alpha_g \theta - M_w b_w \left(\text{tr}\boldsymbol{\varepsilon} - \frac{\zeta_w}{b_w} \right) \alpha_w \theta \right] + \psi_g \rho_g \zeta_g + \psi_w \rho_w \zeta_w \end{aligned} \quad (11)$$

In the above equation \mathbf{e} is the deviatoric strain tensor, G and K are respectively the shear and bulk moduli of the homogenized concrete skeleton, M_g and M_w are the Biot moduli referred to gel and water, b_g and b_w are the Biot coefficients, α , α_g , α_w are respectively the volumetric coefficients of thermal expansion for concrete skeleton, gel and water. The terms $\psi_\alpha \rho_\alpha \zeta_\alpha$ account for free energy supply associated with the mass variations $\rho_\alpha \zeta_\alpha$.

The state equations are obtained by partial derivation and read:

$$\boldsymbol{\sigma} = \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}} = (1-D) \left\{ 2G\mathbf{e} + \left[K\text{tr}\boldsymbol{\varepsilon} - K\alpha\theta + M_g b_g (b_g \text{tr}\boldsymbol{\varepsilon} - \zeta_g - \alpha_g \theta) + M_w b_w (b_w \text{tr}\boldsymbol{\varepsilon} - \zeta_w - \alpha_w \theta) \right] \mathbf{1} \right\} \quad (12)$$

$$\rho_g \mu_g = \frac{\partial \Psi}{\partial \zeta_g} = -(1-D) M_g (b_g \text{tr}\boldsymbol{\varepsilon} - \zeta_g - \alpha_g \theta) + \rho_g \psi_g \quad (13)$$

$$\rho_w \mu_w = \frac{\partial \Psi}{\partial \zeta_w} = -(1-D) M_w (b_w \text{tr}\boldsymbol{\varepsilon} - \zeta_w - \alpha_w \theta) + \rho_w \psi_w \quad (14)$$

$$S = -\frac{\partial \Psi}{\partial \theta}; Y = -\frac{\partial \Psi}{\partial D} \quad (15)$$

The chemical potential of the fluid phases can be expressed in terms of the pressure p_g and p_w that gel and water exert on the solid skeleton and of the specific free energies of gel and water ψ_g and ψ_w as

$$\mu_\alpha = p_\alpha / \rho_\alpha + \psi_\alpha \quad (16)$$

Substituting (16) into equations (13) and (14) one obtains the following forms for the gel and water pressure

$$p_g = -(1-D) M_g (b_g \text{tr}\boldsymbol{\varepsilon} - \zeta_g - \alpha_g \theta) \quad (17)$$

$$p_w = -(1-D) M_w (b_w \text{tr}\boldsymbol{\varepsilon} - \zeta_w - \alpha_w \theta) \quad (18)$$

Consequently the macroscopic stress, equation (12), can be written as

$$\boldsymbol{\sigma} = \frac{\partial \Psi}{\partial \boldsymbol{\varepsilon}} = (1-D) \left[2G\mathbf{e} + K(\text{tr}\boldsymbol{\varepsilon} - \alpha\theta) \mathbf{1} \right] - b_g p_g \mathbf{1} - b_w p_w \mathbf{1} \quad (19)$$

The constitutive model should be completed by evolution equations for the variation of fluid volume contents ζ_α and damage D . For the evolution of the water content we assume the simple form

$$\dot{\zeta}_w = A \varphi_w \dot{S}_w \quad (20)$$

where the constant A is calibrated on the basis of experimental data [20] as explained in Section 3. The evolution of the gel content is assumed to be proportional to the rate of the reaction extent ξ

$$\dot{\xi}_g = c \dot{\xi} \quad (21)$$

The constant c is proportional to the free asymptotic volumetric expansion due to the ASR in the fully saturated case ($t \rightarrow \infty, S_w = 1$) $\mathcal{E}_{ASR, S_w=1}^\infty$

$$c = \frac{K_u}{M_g b_g} \mathcal{E}_{ASR, S_w=1}^\infty \quad (22)$$

with $K_u = K + M_g b_g^2 + M_w b_w^2$ denoting the undrained bulk modulus.

Considering a first order reaction kinetics [9], the following form for the rate of reaction is proposed

$$\dot{\xi} = \frac{\langle f(S_w) - \xi \rangle^+}{\tilde{t}(\xi, S_w, T)} \quad (23)$$

where \tilde{t} is the intrinsic time of the reaction which depends on the local temperature and degree of saturation histories $T(t)$ and $S_w(t)$, and on the reaction extent $\xi(t)$. As suggested by experiments [22] and assumed by various authors, we consider the simultaneity of gel formation and swelling. The intrinsic time is expressed in terms of latency time τ_{lat} and characteristic time τ_{ch} registered for the swelling of reactive specimens. In this work both the dependence on temperature T and degree of saturation S_w is introduced

$$\frac{1}{\tilde{t}} = \frac{\xi + \exp[-\tau_{lat}/\tau_{ch}]}{\tau_{ch}(1 + \exp[-\tau_{lat}/\tau_{ch}])} \quad (24)$$

with

$$\tau_i(T, S_w) = \left\{ \tau_i(\bar{T}, 1) + \frac{\tau_i(\bar{T}, 0) - \tau_i(\bar{T}, 1)}{1 + c_{1i} \exp\left[-\frac{c_{2i}(1 - 2S_w)}{S_w(1 - S_w)}\right]} \right\} \exp\left[U_i \left(\frac{1}{T} - \frac{1}{\bar{T}} \right) \right], \quad i = ch, lat \quad (25)$$

The expressions (25) here proposed combine the Arrhenius law, describing the dependence on the temperature, with a dependence on the moisture, based on experimental results at the reference temperature $\bar{T} = 38^\circ\text{C}$ (see e.g. [9]). In the equation (25) U_{lat} and U_{ch} are the activation energies [13], while the parameters c_{1i} , c_{2i} , $\tau_i(\bar{T}, 0)$ and $\tau_i(\bar{T}, 1)$ are calibrated with experimental data [14]. The function $f(S_w)$ takes into account the influence of water content on the final extent of reaction and assumes the following form

$$f(S_w) = [1 + b_1 \exp(-b_2)] / [1 + b_1 \exp(-b_2 S_w)] \quad (26)$$

where parameters b_1 and b_2 are calibrated with experimental data and reported in Section 3.

The evolution of the damage variables D_t and D_c is governed by loading-unloading conditions defined in terms of the macroscopic stress σ , the gel pressure p_g and the water pressure p_w through the ‘inelastic effective stress’ σ'' defined as

$$\sigma'' \equiv \sigma + \beta_g p_g \mathbf{1} + \beta_w p_w \mathbf{1}$$

The non-dimensional coefficients $\beta_g \leq b_g$ and $\beta_w \leq b_w$ govern the damage level achievable in a concrete specimen under free expansion due ASR and the pressure exerted on skeleton by gel and water.

The loading-unloading conditions read

$$f_t \leq 0 \quad \dot{D}_t \geq 0 \quad f_t \dot{D}_t = 0; \quad f_c \leq 0 \quad \dot{D}_c \geq 0 \quad f_c \dot{D}_c = 0$$

where f_t and f_c are the damage activation function in tension and compression, defined as

$$f_t(\sigma'', D_t) = \frac{1}{2} \mathbf{s} : \mathbf{s} - a_{t0} (\text{tr} \sigma'')^2 + a_{t1} \text{tr} \sigma'' h_t - a_{t2} h_t^2$$

$$f_c(\sigma'', D_c) = \frac{1}{2} \mathbf{s} : \mathbf{s} - a_{c0} (\text{tr} \sigma'')^2 + a_{c1} \text{tr} \sigma'' h_c - a_{c2} h_c^2$$

where \mathbf{s} is the deviatoric stress, h_t and h_c are the hardening-softening functions and a_{i0} , a_{i1} , a_{i2} ($i=t,c$) are non-negative material parameters governing the shape and the size of the elastic domain (see [23] for details).

3 MODEL CALIBRATION

The proposed model requires the identification of four distinct sets of material parameters, namely: (i) elastic parameters for concrete skeleton, gel and water (K , G , M_g , b_g , M_w , b_w) (ii) parameters governing the damage response (a_{i0} , a_{i1} , a_{i2} , $i=t,c$), (iii) parameters defining the expansion due to ASR (U_i , $\tau_i(\bar{T}, 1)$, $\tau_i(\bar{T}, 0)$, c_{i1} , c_{i2} , with $i=lat, ch$, b_1 , b_2 , β_g , β_w) and (iv) parameters governing the hygroscopic behavior of concrete (m_1 , m_2 , m_3 and A). The procedure for the choice of these parameters is briefly illustrated in the following. Table 1 collects the values identified for the concrete used in the experimental test performed by [15], simulated in Section 4.

Table 1: Parameters of the model

a_{t0}	0.3	-	a_{t1}	5.01	MPa	a_{t2}	21.33	MPa ²	K	1e-21	m ²
a_{c0}	0.0025	-	a_{c1}	3.8	MPa	a_{c2}	349.29	MPa ²	η_1	0.001	Kg/ms
$\tau_{lat}(S_w=1)$	80	days	$\tau_{ch}(S_w=1)$	70	days	b_1	130000	-	m_1	37.55	MPa
$\tau_{lat}(S_w=0)$	200	days	$\tau_{ch}(S_w=0)$	150	days	b_2	17	-	m_2	2.17	-
φ_w	0.16	-	φ_g	0.003	-	$\epsilon_{ASR, Sw=1}$	0.0035	-	m_3	0.46	-
b_w	0.41	-	b_g	0.0089	-	β_w	0.025	-	A	0.025	-
M_w	11923.85	MPa	M_g	612496.35	MPa	β_g	0.0008	-			
c_{1lat}	0.0145	-	c_{2lat}	3.43	-	c_{1ch}	0.049	-	c_{2ch}	2.2984	-

(i) The elastic parameters for concrete can be computed from the experimental values of Young’s modulus and Poisson ratio; Biot’s parameters M_w and b_w can be obtained from the

initial porosity of concrete through the homogenized relations proposed in [25]; M_g and b_g can be estimated from the experimental data on the volume fraction occupied by the gel (data in [11] are used).

(ii) The second set of parameters can be identified on the basis of standard uniaxial and biaxial mechanical tests on concrete. Feasible intervals for the model parameters, together with their units and experimental data required for their identification can be found in [23].

(iii) The activation energies for ASR kinetics in (25) have been estimated by Larive on the basis of experiments on reactive concrete specimens, at different temperatures. The parameters c_{1i} , c_{2i} , $\tau_i(\bar{T}, 0)$ and $\tau_i(\bar{T}, 1)$ in (25) and b_1 and b_2 in (26) have been calibrated by using isothermal free expansion tests on reactive concrete cylindrical specimens at different moisture conditions as shown in [14]. Figure 1 shows the model variation of characteristic time, latency time and asymptotic axial ASR expansion with degree of saturation and experimental data. The values of coefficients β_g and β_w can be obtained by calibration on the basis of experimental data or by a proper micromechanical model (e.g. [24] for a bi-phase model).

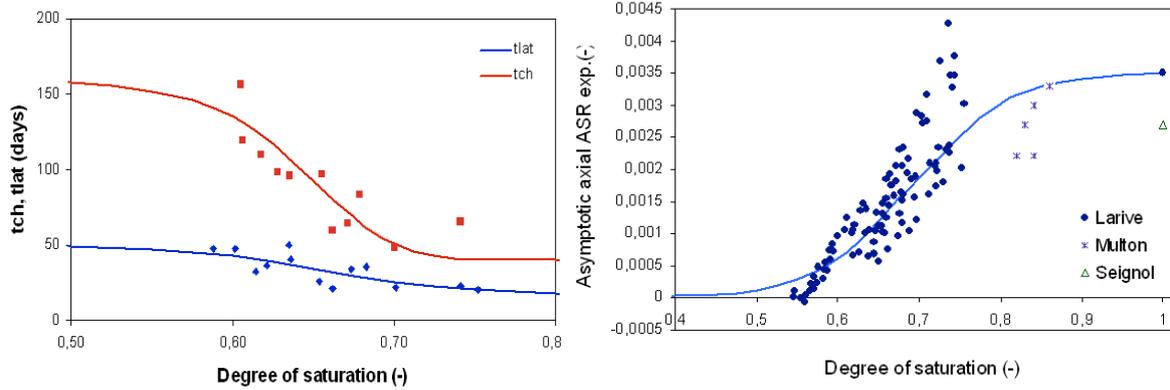


Figure 1: (a) Variation of characteristic time and latency time and (b) asymptotic axial ASR expansion with degree of saturation S_w : experimental points from [1, 2] and proposed model

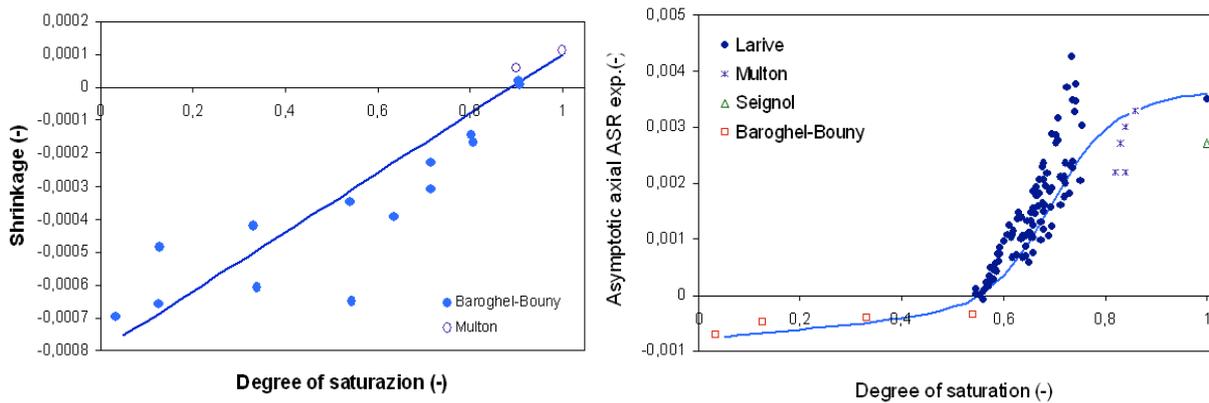


Figure 2: (a) Drying shrinkage of hardened concrete: experimental points from [2, 20] and proposed model (b) Total axial deformation of cylindrical specimens of reactive concrete with degree of saturation: experimental points from [1, 2] and proposed model

(iv) For the calibration the data reported in [20] for ordinary and high performance hardened concrete can be used since experimental studies [2, 15] show the same drying behavior for reactive and non-reactive concrete. The constant A governing the evolution of the water content in (20) is calibrated on the basis of experimental drying and adsorption tests reported in Figure 2a. Parameters m_1 , m_2 and m_3 , governing the dependence of concrete permeability on the degree of saturation in (10), has been taken from [20] for an ordinary concrete.

Figure 2b shows the total axial deformation for a reactive-concrete obtained with the model together with the experimental data used for calibration.

The model has been validated by simulating the experimental tests reported in [2] for a reactive cylindrical specimen kept in water for two years and then exposed to drying air at 30% RH, as shown in Figure 3.

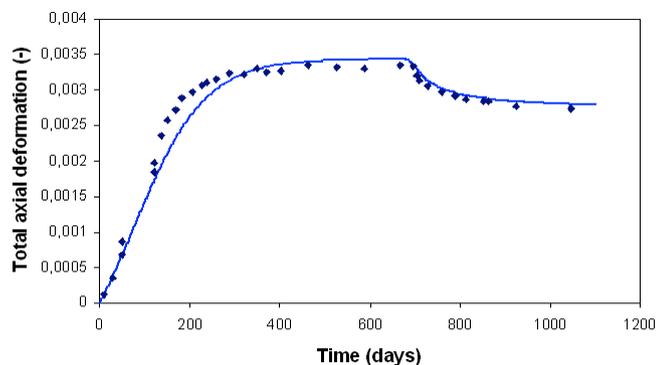


Figure 3: Free expansion tests at $T=38^{\circ}\text{C}$ for a reactive specimen kept in water for two years and then exposed to drying air at 30% RH: experimental points from [15] and model prediction

4 REACTIVE CONCRETE BEAMS SUBMITTED TO MOISTURE GRADIENTS

The proposed model has been implemented in a finite element code and has been employed to simulate the experimental test reported in [15] performed on reactive and non-reactive plain concrete beams subjected to controlled imposed histories of humidity and temperature.

To compute the degradation effects of ASR on concrete structures when temperature and moisture gradients vary in time, the mechanical analysis is preceded by two uncoupled diffusion analyses: the heat diffusion analysis, governed by Fourier law, and the liquid moisture diffusion analysis, governed by equation (9). A 2D (plane strain) finite element discretization of the beam section has been used. The experimental conditions adopted in [15] are reproduced in the analysis by proper initial and boundary conditions: after curing under aluminium sealing, the lower face of the beam was immersed in water, while the upper face was in contact with air at 30% RH for 14 months and then submitted to permanent water supply for 9 months. Figure 4 shows the computed patterns of degree of saturation and reaction extent after 14 and 23 months respectively. The comparison between experimental results and numerical analysis is shown in Figure 5 in terms of strain evolution at different depths of the beam. At 0.08 m from the top the model predicts shrinkage due to the decrease of external humidity conditions. With the model parameters calibrated in Section 3, on the

basis of the experimental results in [1, 2], no ASR expansion occurs at that level of humidity. On the contrary the experimental data show a limited expansion in the upper part of the beam which seems to indicate that ASR is not completely stopped despite the low humidity conditions.

Figure 6 shows the corresponding damage evolution at four depths of the reactive beam. In the upper part (0.08m depth) a limited damage develops during the first months due to drying, while damage induced by ASR develops after the water supply. The damage evolution is faster in the lower part of the beam, where the degree of saturation is high and ASR evolves rapidly.

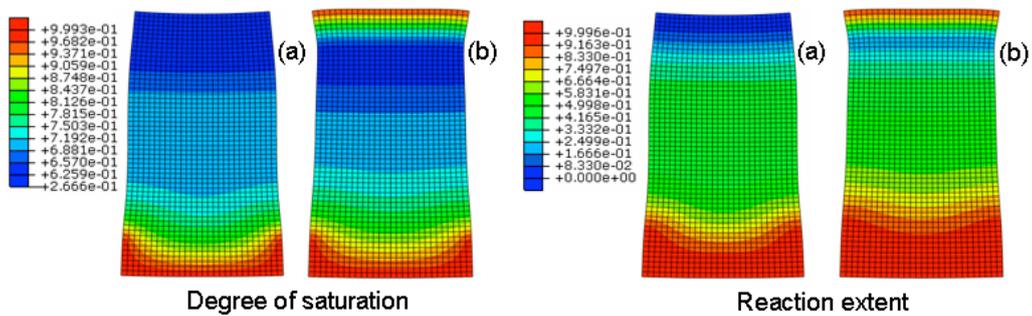


Figure 4: Patterns of degree of saturation and reaction after (a) 14 and (b) 23 months

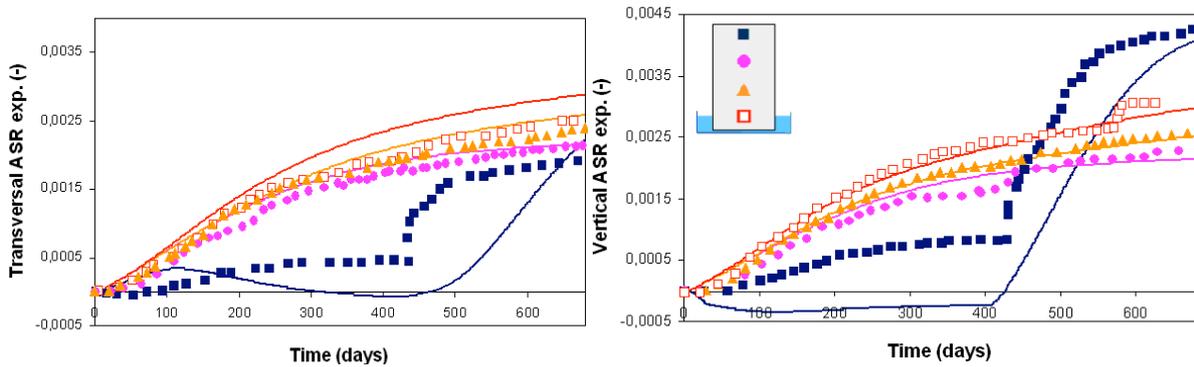


Figure 5: Vertical and transverse strains of reactive beams measured at four depths (0,08, 0,17, 0,27, 0,37 m from the upper face): experimental points from [15] and model prediction

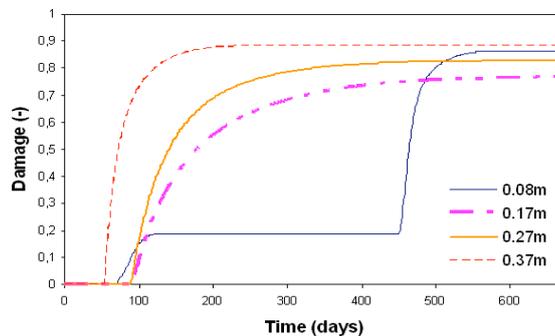


Figure 6: Damage of reactive beams measured at four depths (0,08, 0,17, 0,27, 0,37 m): model prediction

Figure 7 shows vertical and transversal deformations for a non-reactive beam submitted to the same moisture gradient history. A good qualitative agreement is observed. The quantitative discrepancy between experimental data and model results could be caused by a disagreement between the moisture pattern obtained with the theoretical expression in (6) and the real adsorption-desorption behavior of the beam. This aspect needs a closed examination.

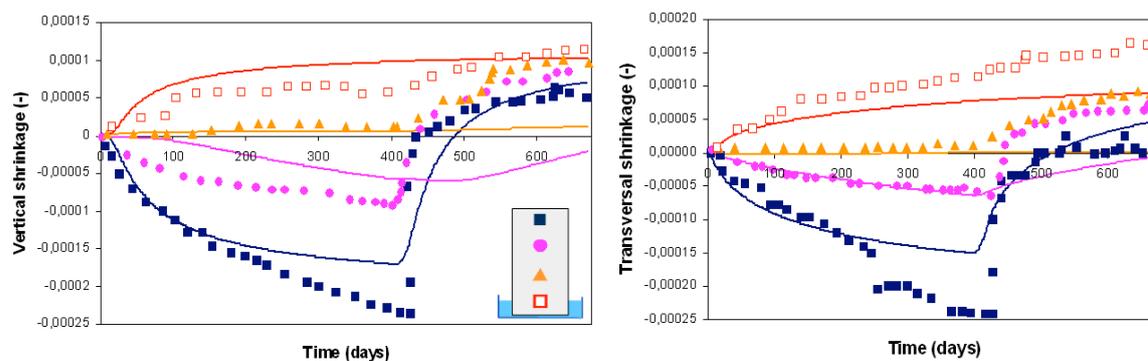


Figure 7: Vertical and transverse strains of non-reactive beams measured at four depths (0,08, 0,17, 0,27, 0,37 m from the upper face): experimental points from [15] and model prediction

5 CONCLUSIONS

In this study a three-phase chemo-elastic damage model for the description of the mechanical degradation of concrete induced by the ASR has been presented. The model takes into account the influence of temperature and humidity histories on the kinetic of the reaction and on the final expansion of the gel produced by the ASR. The consequent mechanical degradation of concrete is described by introducing an isotropic damage variable. The application of the model to beams subjected to moisture gradients shows the capability of the isotropic model to predict the structural degradation due to ASR.

Although the attention has been focused on the rigorous thermodynamic formulation, the model is simple enough to be used in structural analysis.

With respect to the two-phase model previously proposed by the Authors [14], the present model permits to take into account shrinkage effects in a straightforward way.

REFERENCES

- [1] Larive, C. *Apports combinés de l'expérimentation et de la modélisation la compréhension de l'alcali-réaction et de ses effets mécaniques*. Ph.D. Thesis, LCPC, (1998).
- [2] Multon, S. and Toutlemonde, F. Effect of applied stresses on alkali-silica reaction-induced expansions. *Cement and Concrete Research* (2006) **36**:912-920.
- [3] Charlwood, R.G. A review of alkali aggregate in hydro-electric plants and dams. *Hydropower Dams* (1994) 5:31-62.
- [4] Thompson, G.A., Charlwood, R.G., Steele, R.R., Curtis, D. Mactaquac generating station intake and spillway remedial measures. *Proceedings of the 18th International Congress on Large Dams, Durban, South Africa* (1994), 1(Q-68, R. 24): 347-368.
- [5] Léger, P., Côte, P., and Tinawi, R. Finite Element Analysis of concrete Swelling due to Alkali-Aggregate Reactions in Dams. *Computers & Structures* (1996), 60(4):601-611.

- [6] Saouma, V. and Perotti, L. Constitutive Model for Alkali_aggregate Reactions. *ACI Material Journal* (2006) **103**:194-202.
- [7] Farage, M.C.R., Alves, J.L.D. and Fairbairn, E.M.R. Macroscopic modelling of concrete subjected to alkali-aggregate reaction. *Cement and Concrete Research* (2004) **34**:495-505.
- [8] Fairbairn, E., Ribeiro, F., Lopes, L., Toledo-Filho and R., Silvano, M. Modelling the structural behaviour of a dam affected by alkali-silica reaction. *Communications in Numerical Methods in Engineering* (2006) **22**:1-12.
- [9] Ulm, F.J., Coussy, O., Kefei, L. and Larive, C. Thermo-chemo-mechanics of asr expansion in concrete structures. *ASCE Journal of Engineering Mechanics* (2000) **126**(3):233-242.
- [10] Ulm, F.J., Peterson and M., Lemarchand, E.. Is ASR-expansion caused by chemo-poroplastic dilatation? *Journal of Concrete Science in Engineering* (2002) **4**:47-59.
- [11] Grimal E., Sellier, A. Le Pape, Y. and Bourdarot E. Creep, shrinkage, and anisotropic damage in alkali-aggregate reaction swelling mechanism – Part I: a constitutive model. *ACI Material Journal* (2008) **105**(3):227-235.
- [12] Gawin, D., Pesavento, F. and Schrefler B. A. Modeling of cementitious materials exposed to isothermal calcium leaching, considering process kinetics and advective water flow. Part 1: Theoretical model. *international Journal of Solids and Structures* (2008) **45**:6221-6240
- [13] Comi, C., Fedele, R. and Perego, U. A chemo-thermo-damage model for the analysis of concrete dams affected by alkali-silica reaction. *Mechanics of Materials* (2009) **41**:210-230.
- [14] Comi, C. and Pignatelli, R. On damage modeling of concrete affected by alkali-silica reaction in the presence of humidity gradient. In: *XVIII GIMC Conference. Siracusa (Italy)*, (2010), p. 1-4.
- [15] Multon, S. and Toutlemonde, F. Effect of moisture conditions and transfer on alkali silica reaction damaged structures. *Cement and Concrete Research* (2010) **40**:924-934.
- [16] Coussy, O. *Poromechanics*. John Wiley and Sons, New York (2004).
- [17] Prezzi, M., Monteiro, P. J. M. and Sposito, G. The alkali-silica reaction, part I: use of the double-layer theory to explain the behavior of reaction-product gels. *ACI Material Journals* (1997) **94**:10-17.
- [18] Coussy, O., Dangla, P., Dormieux, L. and Lemarchand, E. A two-scale modelling of a swelling clay. *Journal de Physique IV*(1999) **9**(9):21-31.
- [19] Mainguy, M., Coussy, O. and Baroghel-Bouny, V. Role of air pressure in drying of weakly permeable materials. *Journal of engineering mechanics* (2001) **127**(6):582-592.
- [20] Baroghel-Bouny, V., Mainguy, M., Lassabatere, T. and Coussy, O. Characterization and identification of equilibrium and transfer moisture properties for ordinary and high-performance cementitious materials. *Cement and Concrete Research* (1999) **29**:1225-1238.
- [21] Van Genuchten, M. Th. A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils. *Soil Science Society of America Journal* (1980) **44**:892-898.
- [22] Ben Haha, M., Gallucci, E., Guidoum A. and Scrivener K. L. Relation of expansion due to alkali silica reaction to the degree of reaction measured by SEM image analysis. *Cement and Concrete Research* (2007) **37**:1206-1214.
- [23] Comi, C. and Perego, U. Fracture energy based bi-dissipative damage model for concrete. *International Journal of Solids and Structures* (2001) **38**:6427-6454.
- [24] Dormieux, L., Kondo, D. and Ulm, F.J. A micromechanical analysis of damage propagation in fluid-saturated cracked media. *Comptes Rendus Mecanique* (2006) **334**:440-446.
- [25] Heukamp, F.H., Ulm, F.J. and Germaine, J.T. Mechanical properties of calcium-leached cement pastes triaxial stress states and the influence of pore pressures. *Cement and Concrete Research* (2001) **31**:767-774.

AN EFFICIENT COMPUTATIONAL MODEL FOR CO₂ FLOW IN POROUS MEDIA

M. TALEBIAN, R. AL-KHOURY, L. J. SLUYS

Faculty of Civil Engineering and Geosciences, Delft University of Technology
Stevinweg 1, 2628CN Delft, the Netherlands
e-mail: m.taleblian@tudelft.nl

Key words: X-FEM, Level set, Fluid flow, Porous media, Petrov-Galerkin

Abstract. This paper presents an efficient computational model for the simulation of fully coupled two-phase flow in a deformed partially saturated formation. Focus is placed on modeling the flow of a CO₂ plume in a porous medium. The numerical procedure is based on coupling between the Level-set method (LS) and the extended finite element method (XFEM). The level set is employed to define the location of the front between injected CO₂ and existing formation water. A streamline upwind Petrov-Galerkin method is utilized to stabilize the possible occurrence of spurious oscillations in the advection of the CO₂ front. The XFEM is employed to model the gradient in the degree of saturation at the front. This is done by decomposing the saturation field into a continuous part and a discontinuous part. The latter is enhanced by the use of a local enrichment function which is calculated on the basis of the Level-set function. Numerical implementation of the method is discussed and numerical examples are given.

1 INTRODUCTION

Carbon dioxide sequestration is considered to be a promising method for reducing the emission of CO₂ into the atmosphere. This process can be achieved by capturing CO₂ from sources, e.g. flue gases in power plants, and then injecting it into a variety of underground formation, including active or depleted oil and gas reservoirs, subsurface aquifers and coal beds.

The injection of CO₂ into underground formation involves coupled multi-physical processes including fluid flow in porous media and hydro-mechanical processes. The simulation of CO₂ flow and its propagation in porous media is essential for understanding these processes, designing field parameters and monitoring the process.

To attain a high level of accuracy and a stable result for the convection-dominated flow process especially in a large domain, conventional simulators require a large number of finite elements and significant CPU time.

In this work the level set method and the partition of unity method are used to model CO₂ injection in a saline aquifer. The level set method is used to capture the interface between CO₂ and water and the evolution of CO₂ plumes during injection. A streamline upwind Petrov-Galerkin method is utilized to stabilize the possible occurrence of spurious oscillations in the advection of the CO₂ front. The partition of unity finite element method is employed to model the sharp front within the element. This is done by decomposing the saturation field into a continuous part and a discontinuous part. The latter is enhanced by the use of a local enrichment function which is calculated on the basis of the level set function.

The Level set method was first introduced by Osher and Sethian [8] for capturing moving fronts. This method makes use of a distance function, referred to as the level-set function, which labels every point with a sign and a value. The sign indicates the fluid domain and the value represents the shortest distance to the interface. As such this function is equal to zero at the interface and non-zero elsewhere. The interface is advected with a local flow velocity, using any of the known advection equations.

However, when using the standard Galerkin finite element method to solve the LS equation, the advection term creates a skew matrix which is the source of non-physical oscillations. One of the most popular solutions of this problem is known as the Streamline Upwind Petrov-Galerkin method (SUPG). It was proposed by Hughes and Brooks in [3]. In SUPG method the shape function is perturbed in the direction of the flow. There is a vast amount of literature regarding the design and amount of this perturbation. The formulation presented in [11] is used in this study.

In XFEM the standard finite element space is enriched with special functions to capture the discontinuity or high gradient in the problem. The method was first utilized for the simulations of cracks in structures without the need of re-meshing (see for example Belytschko and Black [1] and Moes et al [6]). Sukumar et al. [9] were the first to combine the XFEM with a level set method to make the enrichment function related to interfaces.

In this paper we present some preliminary results of the proposed model.

2 GOVERNING EQUATIONS

The physical model which is used in this approach is the same as that presented by Lewis and Schrefler [5] for deforming multiphase flow in porous media. The macroscopic field variables are displacement, wetting fluid pressure and non-wetting saturation. The governing equations of the coupled multiphase flow are briefly presented hereafter.

2.1 Equilibrium equations

For a three-phase medium, i.e. solid, water and CO2 under static loading, the linear momentum equation in terms of total stress tensor, $\boldsymbol{\sigma}$ can be presented as:

$$\nabla \cdot \boldsymbol{\sigma} + \rho \mathbf{g} = 0 \quad (1)$$

where \mathbf{g} is the gravity force vector per unit mass and ρ is the density given by:

$$\rho = (1-n)\rho_s + nS_w\rho_w + nS_g\rho_g \quad (2)$$

in which ρ_s is the solid density, ρ_w is the water density, ρ_g is the CO2 density, S_w is water saturation, S_g is CO2 saturation and n is the porosity.

The effective stress principle used is given by:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}'' - \mathbf{m}^T \alpha (S_w P_w + S_g P_g) \quad (3)$$

in which, $\boldsymbol{\sigma}''$ is a general form of the effective stresses, P_w and P_g are water and gas pressure, α is the Biot constant and $\mathbf{m}^T = [1, 1, 1, 0, 0, 0]^T$.

2.2 Continuity Equations

The continuity equation for the water phase and the gas phase can be respectively expressed as:

$$\left(\frac{\alpha-n}{K_s} S_w^2 + \frac{nS_w}{K_w} \right) \frac{\partial P_w}{\partial t} + \frac{\alpha-n}{K_s} S_w S_g \frac{\partial P_g}{\partial t} + \alpha S_w \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} \quad (4)$$

$$+ \left(\frac{\alpha-n}{K_s} P_w S_w - \frac{\alpha-n}{K_s} P_g S_w + n \right) \frac{\partial S_w}{\partial t} + \nabla T \left[\frac{\mathbf{k}k_{rw}}{\mu_w} (-\nabla P_w + \rho_w \mathbf{g}) \right] = 0$$

$$\frac{\alpha-n}{K_s} S_w S_g \frac{\partial P_w}{\partial t} + \frac{\alpha-n}{K_s} S_g^2 \frac{\partial P_g}{\partial t} - \left(\frac{\alpha-n}{K_s} S_g (P_g - P_w) + n \right) \frac{\partial S_w}{\partial t} \quad (5)$$

$$+ \alpha S_g \mathbf{m}^T \mathbf{L} \frac{\partial \mathbf{u}}{\partial t} + \frac{nS_g}{\rho_g} \frac{\partial}{\partial t} \left(\frac{P_g M_g}{\theta R} \right) + \nabla T \left[\frac{\mathbf{k}k_{rg}}{\mu_g} (-\nabla P_g + \rho_g \mathbf{g}) \right] = 0$$

in which K_s and K_w are bulk modulus of solid and water, respectively, \mathbf{k} is the absolute permeability of porous media, k_{rw} and k_{rg} are water and gas relative permeability, \mathbf{L} is the differential operator, \mathbf{u} is displacement, μ_g is gas viscosity, M_g is the molecular weight of gas, θ is absolute temperature and R is universal gas constant.

2.3 Constitutive Relations

The systems of equations are nonlinear because the relative permeabilities and the degree of saturation S_w are a function of capillary pressure which is defined as: $P_c = P_g - P_w$ and can be determined experimentally or analytically.

Substituting the relation $\frac{\partial S_w}{\partial t} = \frac{\partial S_w}{\partial P_c} \left(\frac{\partial p_g}{\partial t} - \frac{\partial p_w}{\partial t} \right)$, equation (4) and (5) are converted to a

Pressure-Saturation formulation. The constraint $S_w + S_g = 1$ is considered. There are a variety of relationships that could be used for the relative permeability curves and capillary pressure calculation. In this study the Van Genuchten equation [10] and the Brook and Corey's relationship [2] are used.

3 LEVEL SET METHOD

The level set function is defined as a signed distance function Φ and given by:

$$\Phi(x) = \begin{cases} \min_{x_i \in \Gamma} \|x - x_i\|, & x \in \Omega_1 \\ -\min_{x_i \in \Gamma} \|x - x_i\|, & x \in \Omega_2 \end{cases} \quad \Phi(x) \in \mathbb{R}, x \in \Omega \quad (6)$$

In each time step, the level-set values are advected with the velocity field of the fluid, as

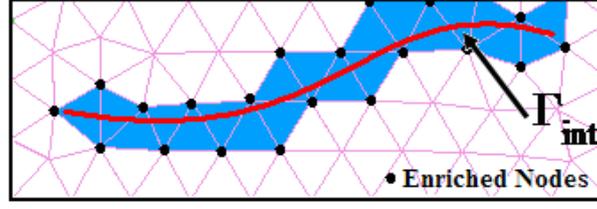


Figure 1: A typical finite element mesh with interface, showing enriched nodes and a front

$$\frac{\partial \Phi}{\partial t} + \mathbf{u} \cdot \nabla \Phi = 0 \quad (7)$$

An important issue in the level-set method is the reinitialization of the level-set function in each time step. This is necessary because otherwise the distance property of the level-set function is no longer maintained after the transport. The reinitialization procedure used in [7] is used in this study.

The SUPG (Streamline Upwind Petrov-Galerkin) method is used in this study to stabilize the level set equation. The numerical scheme is stabilized by adding a perturbation to the weighting functions. This perturbation is proportional to the gradient of the standard interpolation functions as:

$$\tilde{N} = N + \tau \mathbf{u} \cdot \nabla N \quad (8)$$

where N is standard shape function, \mathbf{u} is flow velocity vector, τ is a perturbation parameter and \tilde{N} is the perturbed shape function. The perturbation parameter presented in [11] is used in this study.

4 EXTENDED FINITE ELEMENT

Due to the high gradient in the saturation field at the front between different phases, the XFEM is utilized. The Level set function is utilized for enhancing the element. The enriched approximation is given by:

$$S^h(x, t) = \sum_{I \in N} N_I(x) S_I(t) + \sum_{J \in N^{enriched}} N_J^{enriched}(x, t) a_J \quad (9)$$

where $S(t)$ are the nodal saturations for the standard finite element and $a_j(t)$ are additional nodal parameters at the enriched node J . In constructing the finite element approximation, we distinguish nodes whose element is intersected by the interface Γ_{int} from all others; this set of nodes is indicated by N^{enrich} . Fig. 1 illustrates which nodes are enriched for a typical example. So for the enriched node an extra degree of freedom should be defined. The enriched shape functions can be defined based on a standard shape function and the level set function as:

$$N_J^{enriched}(x, t) = N_j(x) \left(\left| \Phi^h(x, t) \right| - \left| \Phi^h(x_j, t) \right| \right) \quad (10)$$

Figure (2) shows the enrichment functions for a linear two-node element. The enrichment function is calculated by multiplying a standard shape function to a level set function as represented in equation (10).

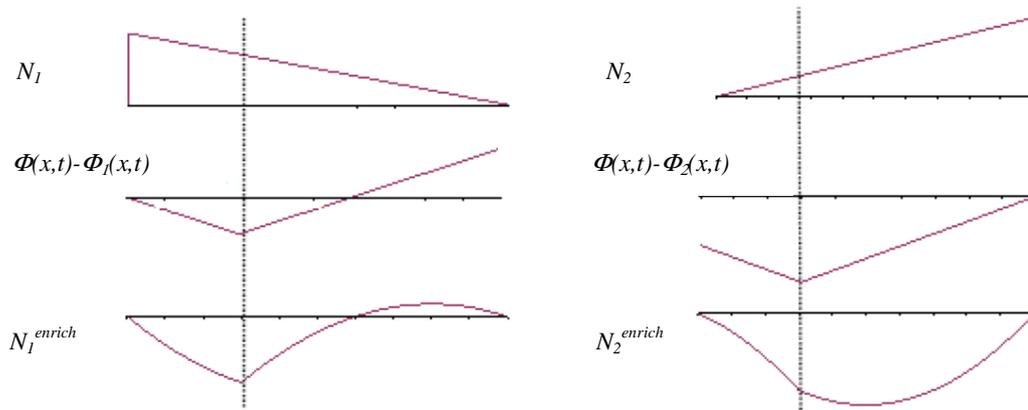


Figure 2: Example of two enriched finite element shape functions for a linear two node element.

5 NUMERICAL EXAMPLE

5.1 Buckley-Leverett

Here, we consider idealized immiscible two phase displacement flow (Buckley-Leverett) in a porous medium. The Buckley-Leverett problem is a simple test problem that has an analytical solution. This problem describes two-phase flow of two immiscible and incompressible fluids in a porous medium, where capillary pressure, gravitational force and deformation are ignored.

The geometry, initial and boundary conditions and model parameter are given in Fig. 3. The CO₂ displaces a wetting phase from left to right. Fig. 4 shows the motion of the saturation front in time. In Fig. 5, the numerical solutions obtained by the standard Galerkin method with 125 elements, by adding artificial diffusion and the PUM method using a coarse mesh with 25 elements are compared with the analytical solution of the Buckley-Leverett problem.

Obviously, the standard Galerkin method exhibits spurious oscillations due to convection. The Petrov Galerkin is highly dissipative, and produces unrealistically smeared sharp fronts. The proposed model produces a sharper front and a smoother profile.

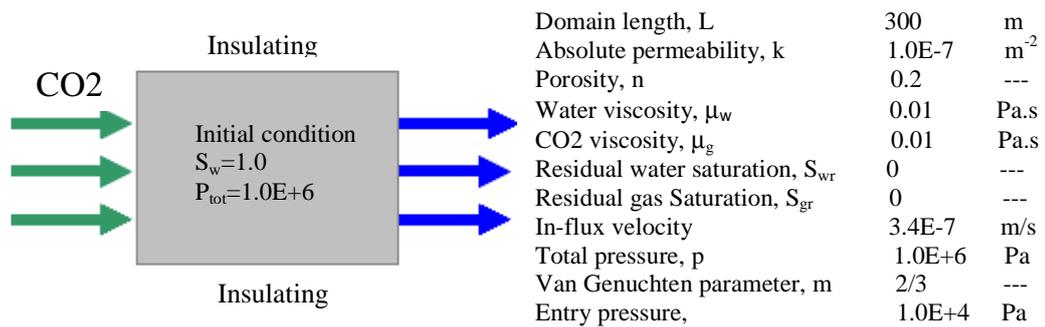


Figure 3. The 2D geometry, initial and boundary conditions used for solving the two-phase flow formulation.

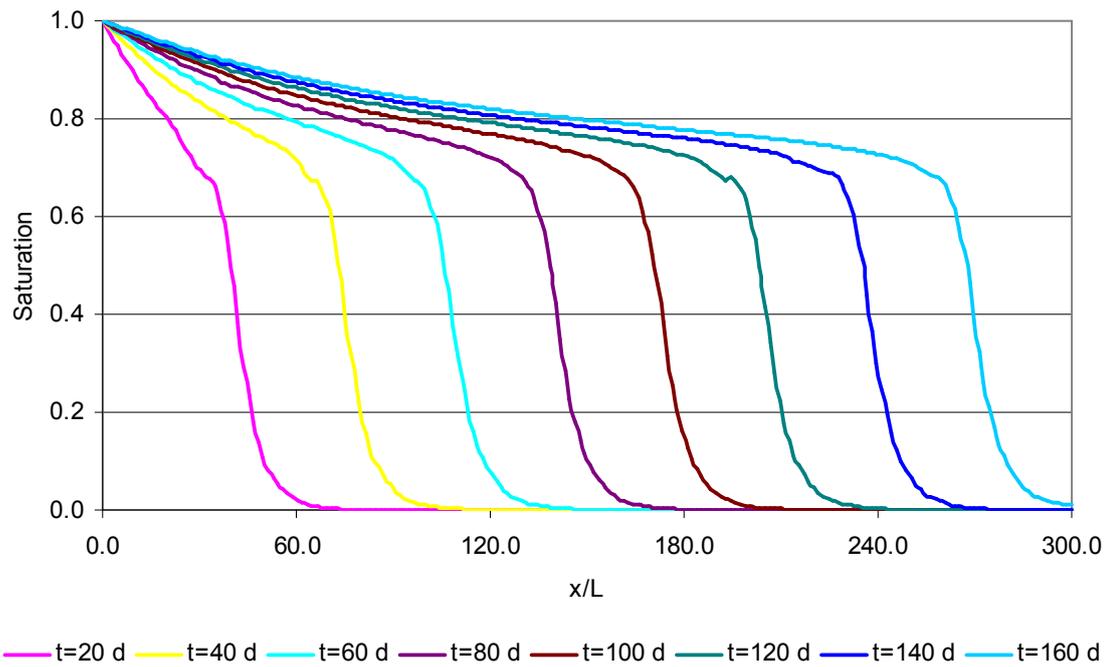


Figure 4. Saturation profiles for different time using the new approach

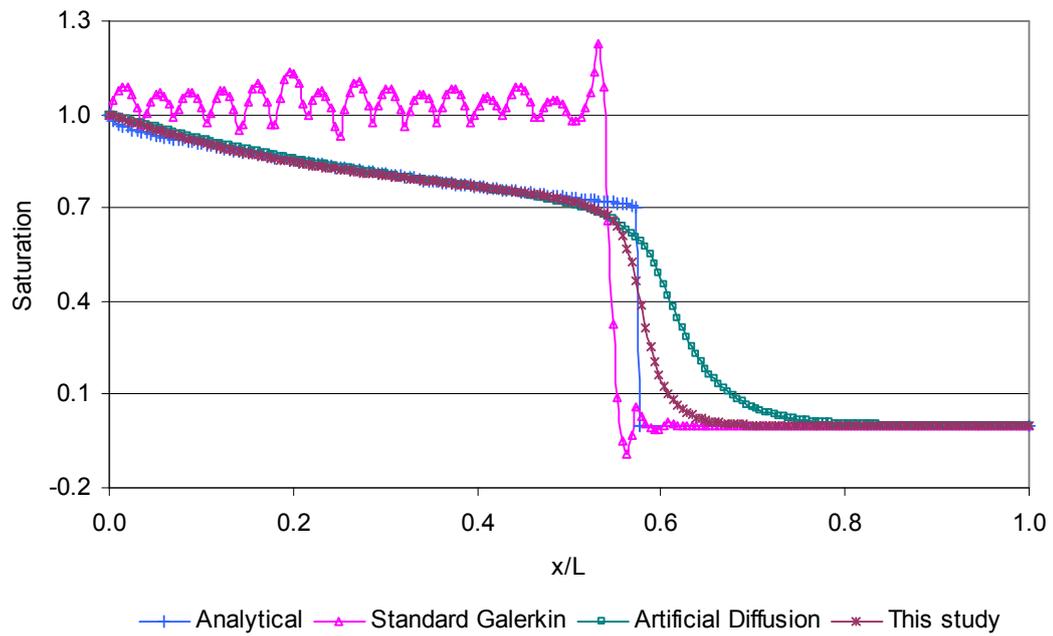


Figure 5. Comparison of saturation profile at t=100 day for the different methods.

5.2 McWhorter Benchmark Problems

In this test, the capillary pressure is taken into consideration. The deformation of porous media is ignored. The initial saturation of CO₂ is assumed to be 0.1 in the medium. The domain length is assumed to be 2.6 meter. Here, flow is governed by a capillary force when CO₂ saturation at the left end of the medium is kept to one and the CO₂ injection pressure is assumed to be 2.0×10^5 Pa.

Figure (6) shows the motion of the saturation front. In Figure 7, numerical results obtained from standard Galerkin method using coarse ($\Delta x=0.1$) and fine meshes ($\Delta x=0.02$) are compared with those obtained from this study and from the semi-analytical solution of the McWhorter problem. The saturation profile at time $t = 4000$ s for the three methods and the analytical solution is plotted in this figure.

In this case, unlike the Buckley-Leverett problem, which is a convection dominant problem, the capillary effect inserts a diffusive effect in the solution; therefore the front between the two phases is no longer sharp. Consequently the difference between the standard procedure and the extended procedure is less pronounced for this case.

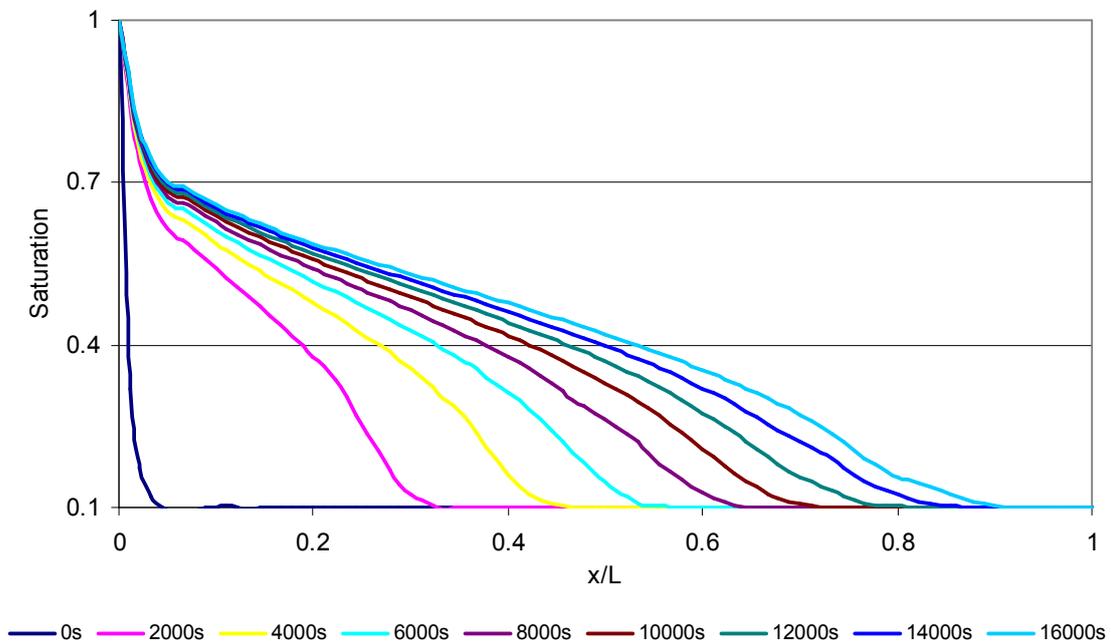


Figure 6. Saturation profiles for different times.

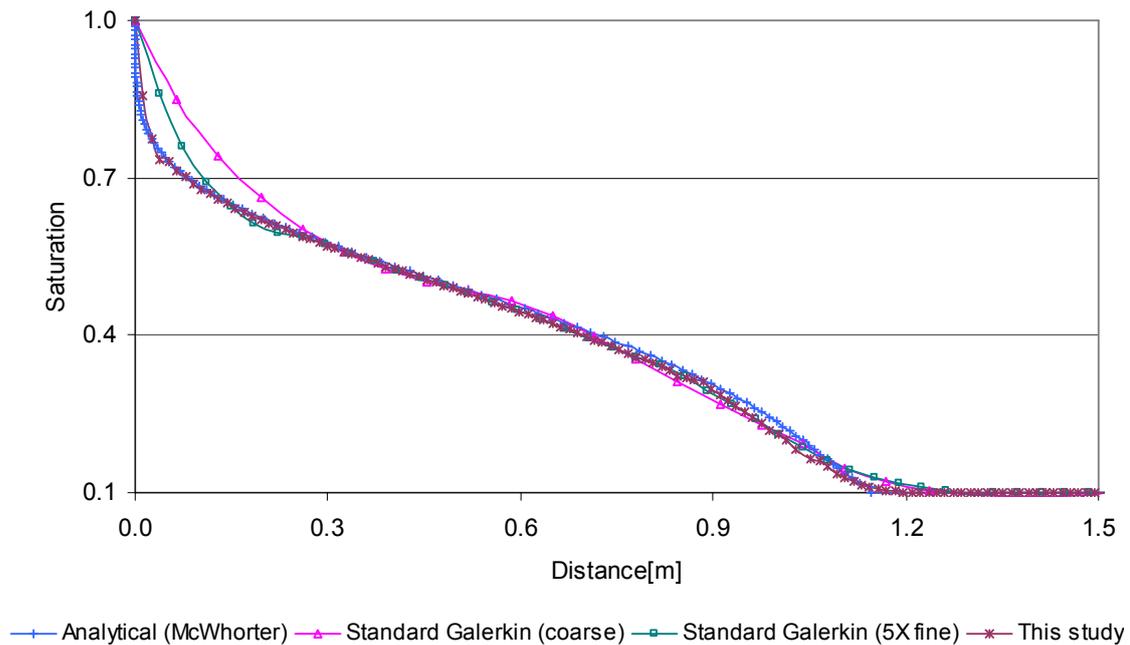


Figure 7. Comparison of saturation profile at $t=4000s$ for the different methods.

6 CONCLUSIONS

A computational model based on the level set method and the extended finite element method has been developed and utilized for modeling two-phase fluid flow in unsaturated porous media. The level set method is used to track the interface between the injected gas and the resident fluid. A streamline upwind Petrov-Galerkin procedure is used to stabilize the level set equation. The high gradient in the saturation profile is modeled by decomposing it into continuous and discontinuous parts based on the partition of unity method. This makes the model capable of capturing the discontinuity or high gradient in the saturation and stabilizes the result using relatively coarse meshes.

REFERENCES

- [1] T. Belytschko and T. Black (1999), "Elastic Crack Growth in Finite Elements With Minimal Remeshing," *International Journal for Numerical Methods in Engineering*, Vol. 45, Number 5, pp. 601-620
- [2] Brooks, R. J., and A. T. Corey, Hydraulic properties of porous media, *Hydrol. Pap. 3*, Colo. State Univ., Fort Collins, 1964.
- [3] A.N. Brooks and T.J.R. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comput. Methods Appl. Mech. Engrg.* 32 (1982), pp. 199-259.
- [4] J. Chessa, T. Belytschko: An extended finite element method for two-phase fluids, *ASME J. Appl. Mech.*, 70, 10-17, 2003.
- [5] Lewis, R.W. and Schrefler, B.A., 2000: "The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media". 2nd edition, John Wiley & Sons.
- [6] N. Moës, J. Dolbow and T. Belytschko (1999), "A Finite Element Method for Crack Growth Without

- Remeshing," *International Journal for Numerical Methods in Engineering*, Vol. 46, Number 1, pp. 131-150
- [7] E. Olsson, G. Kreiss, S.Zahedi, A conservative level set method for two phase flow II *Journal of Computational Physics*, Volume 225, Issue 1, 1 July 2007, Pages 785-807,
- [8] S. Osher, J.A. Sethian, Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulation, *J. Computat. Phys.* 79 (1988) 12–49.
- [9] N. Sukumar, D. L. Chopp, N. Moës, "Modeling Holes and Inclusions by Level Sets in the Extended Finite–Element Method," *Comput. Meth. Appl. Mech. Eng.*, 2001, Vol. 190, Number 46–47, pp. 6183–6200
- [10] Van Genuchten, M.Th. A Closed-Form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils, *Soil Sci. Soc.* , Vol. 44, pp. 892 - 898, 1980.
- [11] Volker John, Ellen Schmeyer Finite element methods for time-dependent convection diffusion- reaction equations with small diffusion, *Comput. Meth. Appl. Mech. Engrg.*, 198, 475 - 494, 2008.

BACK ANALYSIS OF A COUPLED THERMO-HYDRO-MECHANICAL MODEL BASED ON INSTRUMENTED CONSTANT VOLUME COLUMN TEST

TOM SCHANZ*, MARIA DATCHEVA† AND LONG NGUYEN-TUAN*

*Chair for Foundation Engineering, Soil and Rock Mechanics
Ruhr-Universität-Bochum, Germany
Universitätsstr. 150, IA 4, 44780 Bochum
e-mail: [tom.schanz; long.nguyentuan]@rub.de, web page: <http://www.gbf.rub.de/>

†Institute of Mechanics - Bulgarian Academy of Sciences
Acad. G. Bontchev, block 4, 1113 Sofia Bulgaria
e-mail: datcheva@imbm.bas.bg - web page: <http://www.imbm.bas.bg/>

Key words: Coupled Thermo-Hydro-Mechanical (THM), sand-bentonite mixture, sensitivity analysis, optimization, back analysis.

Abstract. This study contributes to identification of the constitutive model parameters for coupled THM models for unsaturated sand-bentonite mixtures via back analysis approach. The approach strategy consists of: definition of the forward model, sensitivity analysis, selection of optimization algorithm, selection of a set of parameters to be optimized, setup of the parameter's constraints, and assessing the reliability and accuracy of the identified model and material parameters. For this analysis the iterative direct approach based on numerical solution of the direct problem and minimization of an objective function has been selected. It is given an example of application of the selected inverse analysis procedure to identification of parameters involved in a modified Barcelona Basic Model taking into account of variation of temperature.

1 INTRODUCTION

Current solution for the radioactive waste disposal is to place the canisters containing the waste in a tunnel system located deep in the host rock. The canisters are surrounded by expansive clay that composes the buffer. The behavior of the buffer needs to be well understood in order to guarantee the safety and the efficiency of the radioactive waste repository. The clay in the buffer, initially unsaturated, is subjected to high temperature emitted by the radioactive waste and to hydraulic gradients induced by water permeating from the host rock. As a consequence swelling and shrinking phenomena take place with the variation of water content and temperature.

During the past decade a number of numerical simulations were carried out in order to assess the physical processes and predict the behavior of the buffer soil in a real environment. Because the coupled THM phenomena are complicated the constitutive models were gradually gaining complexity. This material model complexity often invokes the need of large sets of model parameters that are not simple for determining experimentally. Several researchers contributed with their experimental studies for deriving unsaturated soil models parameters. Particularly for the Barcelona Basic Model (BBM), contributions are made by e.g. Lloret et al [8] for FEBEX bentonite, Geiser et al [5] for sandy silt, Agus [1] for sand-bentonite mixtures. However due to device and sensor restrictions it may not be always possible directly to measure and provide sufficient and reliable laboratory test data for determining the material model parameters, especially, the parameters with the effect of temperature. The available experimental data may request back analysis procedure for identification of model parameters by minimizing the error function between measurement and e.g. numerical simulation results. For instance, Schanz et al (2008) [11] determined couple hydro-mechanical parameters for the modified BBM model [4] from measurement in swelling pressure cell. In the present paper, the back analysis procedure is introduced to identify the coupled THM model parameters for sand-bentonite mixture based on constant volume column test data. The approach strategy consists of: definition of the forward model, sensitivity analysis, selection of optimization algorithm, selection of a set of parameters to be optimized, setup of the parameter's constraints, and assessing the reliability and accuracy of the identified model parameters.

2 CONSTITUTIVE EQUATIONS OF THE COUPLED THM MODEL

Following the two stress variable concept in unsaturated soil mechanics, the elastic part of the strain increment is taken to be a sum of the increments of suction induced $\boldsymbol{\varepsilon}^{s-e}$, net stress induced $\boldsymbol{\varepsilon}^{\sigma-e}$ and the strain increment due to temperature change $d\boldsymbol{\varepsilon}^{T-e}$. The final relation for the elastic strain increment reads:

$$d\boldsymbol{\varepsilon}^e = d\boldsymbol{\varepsilon}^{\sigma-e} + d\boldsymbol{\varepsilon}^{s-e} + d\boldsymbol{\varepsilon}^{T-e} \quad (1)$$

The nonlinear elastic law for the volumetric strain induced by the net stress is expressed in Eq. 2.

$$d\varepsilon_v^{\sigma-e} = \frac{\kappa_i(s)}{1+e} \frac{dp'}{p'} \quad \text{and} \quad p' = p - \max(p_g, p_l) \quad (2)$$

$$\kappa_i(s) = \begin{cases} \kappa_{io} (1 + \alpha_i s) & \text{if } 1 + \alpha_i s \geq 0.001 \\ 0.001 \kappa_{io} & \text{if } 1 + \alpha_i s < 0.001 \end{cases} \quad (3)$$

where p is mean total stress, p' is the mean net stress in unsaturated state or effective stress in saturated state, p_g and p_l are gas pressure and liquid pressure, e is the void ratio,

κ_{io} and α_i are model parameters. For deviatoric elastic strains, a constant Poisson's ratio is used.

Suction and temperature induce only volumetric strains with constitutive equations given as following:

$$d\varepsilon_v^{s-e} = \frac{\kappa_s(p', s)}{1+e} \frac{ds}{s+p_{at}} \quad ; \quad d\varepsilon_v^{T-e} = \alpha_o dT \quad (4)$$

with

$$\kappa_s(p', s) = \kappa_{so} \kappa_{sp} \exp(\alpha_{ss} s) \quad (5)$$

and

$$\kappa_{sp} = \begin{cases} 1 + \alpha_{sp} \ln\left(\frac{10^{-20}}{p_{ref}}\right) & \text{if } p' \leq 10^{-20} \\ 0 & \text{if } p' \geq p_{ref} \exp\left(\frac{-1}{\alpha_{sp}}\right) \\ 1 + \alpha_{sp} \ln\left(\frac{p'}{p_{ref}}\right) & \text{elsewhere} \end{cases} \quad (6)$$

The parameters involved are: α_o for the elastic thermal strain; κ_{so} is the elastic stiffness parameter in changing of suction at zero net stress; p_{at} is the atmospheric pressure; α_{ss} and α_{sp} are model parameters. The elastic modulus κ_i and κ_s may be considered not dependent on temperature in case of moderate temperature gradients.

The yield surface in BBM model is given in the deviatoric plane $q-p$ via the following equation:

$$F = q^2 - M^2 (p' + p_s) (p_o - p') = 0 \quad (7)$$

where $q = \sqrt{\frac{3}{2} \boldsymbol{\sigma}^D : \boldsymbol{\sigma}^D}$, with deviatoric stress defined as $\boldsymbol{\sigma}^D = \boldsymbol{\sigma}' - \frac{1}{3} \boldsymbol{\sigma}' : \mathbf{I}$. The preconsolidation pressure p_o depends on suction and according to Alonso (1990) [2] it is defined as:

$$p_o = p^c \left(\frac{p_o^*}{p^c} \right)^{\frac{\lambda(0) - \kappa_{io}}{\lambda(s) - \kappa_{io}}} \quad (8)$$

where p^c is a reference pressure, p_o^* is the preconsolidation pressure for a saturated state, $\lambda(0)$ is a plastic stiffness parameters for changes in effective stress at saturated state. The stiffness parameter for changes in the mean net stress at given suction is defined by:

$$\lambda(s) = \lambda(0) [(1-r) \exp(-\beta s) + r] \quad (9)$$

where r and β are model parameters.

The tensile strength p_s , follows a linear relationship with suction and is a function of temperature:

$$p_s = p_{s0} + k s \exp(-\rho \Delta T) \quad \text{and} \quad \Delta T = T - T_{ref} \quad (10)$$

where k is parameter that takes into account the increase of tensile strength due to suction, p_{s0} is tensile strength in saturated state, ρ is a parameter that takes into account the decrease of the tensile strength due to temperature increase, T_{ref} is a reference temperature.

The isotropic hardening depends on the plastic volumetric strain according to:

$$dp_o^* = \frac{1 + e}{\lambda(0) - \kappa_{io}} p_o^* d\varepsilon_v^p \quad (11)$$

For hydraulic process, advective flow of the water phase is described by the generalized Darcy's law:

$$\mathbf{q}_l = -\frac{\mathbf{k}k_{rl}}{\mu_l} (\nabla p_l - \rho_l \mathbf{g}) \quad (12)$$

where μ_l is the dynamic viscosity of the pore liquid, \mathbf{g} is the gravity acceleration, ρ_l is the liquid density. The tensor of intrinsic permeability \mathbf{k} , is supposed to depend on porosity according to the Kozeny's model:

$$\mathbf{k} = \mathbf{k}_o \frac{\phi^3}{(1 - \phi)^2} \frac{(1 - \phi_o)^2}{\phi_o^3} \quad (13)$$

where ϕ is the porosity, ϕ_o is a reference porosity, \mathbf{k}_o is the intrinsic permeability for matrix with porosity ϕ_o . The relative permeability k_{rl} , is derived from Mualem-van Genuchten closed form model, [7]:

$$k_{rl} = \sqrt{S_e} \left(1 - (1 - S_e^{1/\lambda})^\lambda \right)^2 \quad (14)$$

where λ is a shape parameter for retention curve and S_e is defined as:

$$S_e = \frac{S_l - S_{rl}}{S_{ls} - S_{rl}} = \left(1 + \left(\frac{p_g - p_l}{P_0} \right)^{\frac{1}{1-\lambda}} \right)^{-\lambda} \quad (15)$$

where S_l , S_{ls} and S_{rl} are the current, the maximum and the residual liquid degree of saturation, P_0 is a model parameter.

Fick's law is adopted to define the diffusive flux of water vapour \mathbf{i}^v :

$$\mathbf{i}^v = -(\phi \rho_v S_l D_m \mathbf{I}) \nabla \omega^v \quad (16)$$

where ρ_v is the vapour density, ω^v is the mass fraction of the vapour, \mathbf{I} is the identity matrix and D_m is the diffusion coefficient of vapour in m^2/s is defined by:

$$D_m = \tau D \frac{(273.15 + T)^n}{P_g} \quad (17)$$

where τ is the tortuosity, D is the molecular diffusion coefficient at temperature 273.15K and $P_g = 101\text{kPa}$, and n is a coefficient.

Fourier's law is adopted for heat conduction flux, \mathbf{i}_c , of heat:

$$\mathbf{i}_c = -\lambda_T \nabla T \quad \text{where} \quad \lambda_T = \lambda_{sat}^{S_i} \lambda_{dry}^{(1-S_i)} \quad (18)$$

where λ_T is the soil thermal conductivity, λ_{sat} and λ_{dry} are soil thermal conductivity at the saturated and dry state, respectively.

In summary, there are total 26 parameters to describe the behaviour of coupled THM model.

- Parameters involved in modelling net stress driven processes ($d\sigma \neq 0$):
 $\mathcal{M} = \{\kappa_{io}, \alpha_i, p_{ref}, \lambda(0), r, \beta, k, p_{s0}, p^c, M, \alpha, e_o, p_o^*\}$
- Parameters involved in modelling suction driven processes ($ds \neq 0$):
 $\mathcal{H} = \{P_0, \lambda, \phi_0, \kappa_o, \kappa_{s0}, \alpha_{ss}, \alpha_{sp}\}$
- Parameters involved in modelling temperature driven processes ($dT \neq 0$):
 $\mathcal{T} = \{\tau, D, \lambda_{sat}, \lambda_{dry}, \alpha_0, \rho\}$

The total parameters are summarised in vector: $\mathbf{x} = \{\mathcal{H}, \mathcal{T}, \mathcal{M}\}$

3 IDENTIFICATION OF CONSTITUTIVE PARAMETERS FOR COUPLED THM MODEL VIA BACK ANALYSIS

The back analysis approach strategy consists in the following steps: definition of the forward model, parameter sensitivity analysis; selection of a set of parameters to be optimized; selection of optimization algorithm; setup of the parameter constraints, and assessing the reliability and accuracy of the identified model parameters.

3.1 Sensitivity analysis

The influence of model parameters on the model response is determined based the following fundamental equations.

1- *Determination of scaled sensitivity (SS)*: The *SS* analysis indicates the amount of information provided by the *i-th* type of observation for the estimation of *j-th* parameter. We define *SS* for each particular observation *k* of the *i-th* type of observation:

$$SS_{i,j}^k = \frac{x_j}{y_i^k} \frac{\partial y_i^k}{\partial x_j} \quad (19)$$

Thus we use k to denote different observations done during the experiment at different time, e.g. t_k , $k = 1 \dots N$

2- *Determination of composite over time scaled sensitivity (CSS)*: CSS is used to measure the i -th type of observation sensitivity to a given parameter x_j over the whole time of measurement:

$$CSS_{i,j} = \sqrt{\frac{1}{N} \sum_{k=1}^N (SS_{i,j}^k)^2} \quad (20)$$

3- *Determination of sensitivity factor $\gamma_{i,j}$ for each of parameters*: The sensitivity factor is used to normalise $CSS_{i,j}$ in the range from zero to one.

$$\gamma_{i,j} = \frac{CSS_{i,j}}{\max_j \{CSS_{i,j}\}} \quad (21)$$

In order to understand the response of the model in different time intervals, the sensitivity analysis later done for the THM model is performed for three different time intervals, namely at the end of the experiment ($T100$), at 50% of the total time of the experiment ($T50$), and at the first time step of calculation process, ($T0$). In our particular case the vector of model response is $\mathbf{y} = \{S_l(t), T(t), \sigma_{yy}(t)\}$.

3.2 Model-parameter optimization via direct inverse approach

The direct inverse approach, which consists of an automated iterative procedure correcting the trial values of the unknown parameters by minimizing an error function, is applied here for the back analysis of the instrumented constant volume column test. The optimization algorithm uses the simplex Nelder-Mead optimization method [10]. The objective function employed here is the absolute mean error F_{AM} (Eq. 22). The solution of the optimization problem is considered against one of following criteria: $F_{AM} \leq \epsilon$, the maximum number of iteration, or Eq. 23. The optimization routine is executed by means of the VARO²PT [13] optimization tool.

$$F_{AM}(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n |y_i^{meas} - y_i(\mathbf{x})| \quad (22)$$

$$\Delta F_{AM} = |F_{AM} - F_{AM}^{prev}| \leq \Delta_\epsilon \quad (23)$$

where \mathbf{y}^{meas} and \mathbf{y} are the vectors of the measurement and numerical (model) observations, F_{AM}^{prev} is the value of the objective function from the previous step, ϵ and Δ_ϵ are critical values to stop the optimization iterations.

3.3 Assessment of the quality of the optimal set of model parameters

The values of the sensitivity factors $\gamma_{i,j}$ are used to assess the reliability of optimal set of parameters. The mean value ν_j gives the overall model sensitivity to j -th parameter:

$$\nu_j = \frac{1}{n} \sum_i^n \gamma_{i,j} \quad (24)$$

3.4 Assessment of the goodness of the fit

We may use several strategies to back calculate the model parameters and depending on the strategy we may obtain different optimal parameter sets. In order to compare different solutions and assess the goodness of the fit we calculate the following statistical measures: the mean error, the standard deviation and skewness, defined as:

$$\mu = \frac{1}{n} \sum_{i=1}^n \varepsilon_i \quad ; \quad \sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (\varepsilon_i - \mu)^2} \quad ; \quad \gamma_1 = \frac{\mu_3}{\sigma^3} \quad (25)$$

where n is a number of measured samples and ε_i is the error between i -th measurement and simulation values, μ_3 is the third moment about the mean.

4 DEFINITION OF THE FORWARD PROBLEM

A series of hydration test and heating test were performed in the newly developed THM apparatus, [9]. The experimental data obtained is used for back analysis and the boundary and initial conditions of the tests are used in building the model for the forward calculation.

4.1 Hydration test

During the hydration test we measured the water absorbed by sand-bentonite mixture (SBM) and the vertical stress. The sample was hydrated from the top and development of swelling pressure with time was measured at top and bottom ends of the sample. The evolution of water front along the vertical axis of specimen is recorded by three Time Domain Reflectometry (TDR) sensors.

For numerical simulation of hydration test, the numerical model is built in the X-Y plane, Fig. 1a. At the top and the bottom of the model liquid flux boundary condition is applied. The distances from points 1, 2, 3 to the top of the model are 50 mm, 150 mm, 250 mm, respectively (see Fig. 1). Points 1, 2 and 3 correspond to the position of the measurement devices. Point 4 is located at the bottom of the sample where the load cell measurements are recorded.

4.2 Heating test

The heating test series was carried out to investigate the behavior of SBM sample by heating. The temperature boundary conditions is a prescribed temperature at the bottom

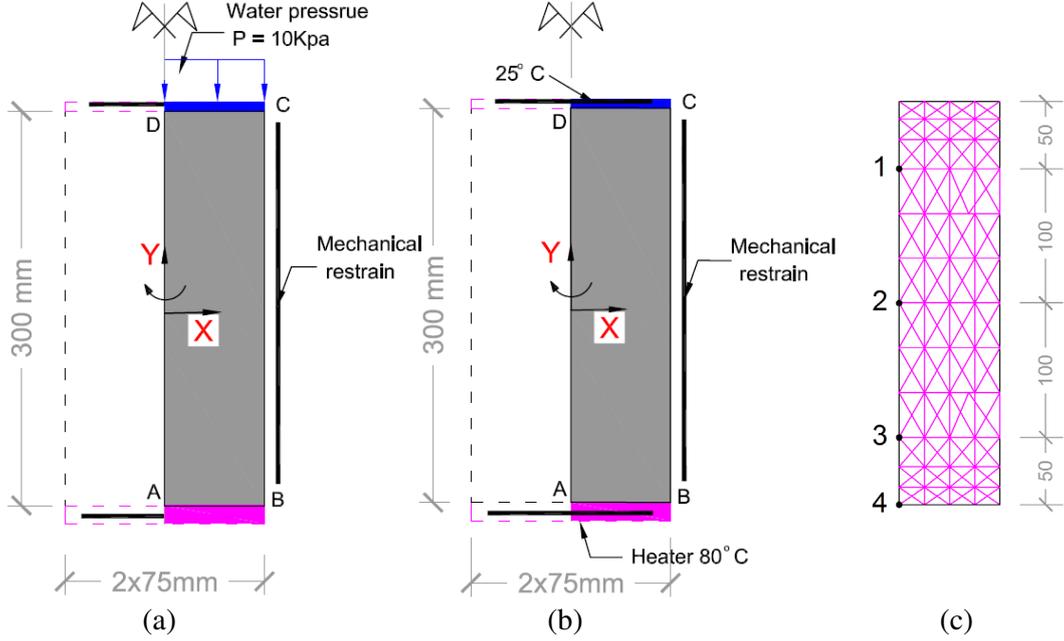


Figure 1: Numerical model: (a)– Hydration test model; (b)– Heating test model; (c)– FE-discretization with the observation points

(80 °C) and at the bottom (25 °C) and zero heat flux at the lateral boundaries (see Fig.1b). The change of humidity was measured by the RH sensors, temperature sensors are also installed at the same place where RH sensors are located. Water content within the specimen is measured by TDR sensor inserted in the soil specimen and thermocouple sensors are placed nearby to the TDRs to measure the current local temperature. No water is supplied. The independent variables for the model are displacement vector \mathbf{u} , liquid pressure p_l and temperature T .

5 RESULTS AND DISCUSSION

5.1 Results of the sensitivity analysis

For the hydration test model, the material response is characterised in the terms of degree of saturation (S_l) and vertical stress (σ_{yy}). The vector of model parameters is now $\mathbf{x}^H = \{\mathcal{H}, \mathcal{M}\}$ and the vector of model response is $\mathbf{y} = \{S_l(t), \sigma_{yy}(t)\}$. The results of the sensitivity analysis are presented in Fig. 2. No data is presented for the parameters that have very low influence to the model responses.

For the heating test model, the vector of model parameters is $\mathbf{x}^T = \{\mathcal{H}, \mathcal{T}, \mathcal{M}\}$ and the vector of the model response is $\mathbf{y} = \{S_l(t), \sigma_{yy}(t), T(t)\}$. The results of sensitivity analysis are presented in Fig. 3. The analysis of temperature response indicates that λ_{sat} is the parameter influencing the most the heat conduction process. The analysis of

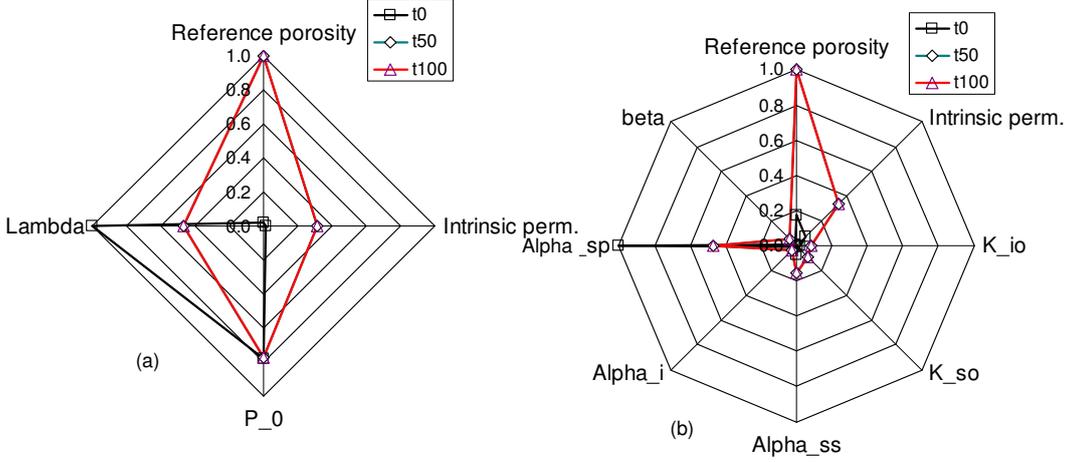


Figure 2: Hydration test: (a)– γ_j of degree of saturation. (b)– γ_j of vertical stress

the vertical stress indicates that the parameter α_{sp} is significantly influencing the model response at the beginning of the test and its influence is reduced with time. Beside that, the parameter k_o is significantly influencing the vertical stress evolution.

5.2 Results of the optimization

For the initial forward calculation we used parameter set with parameter values found in literature. The mechanical parameters (\mathcal{M}) are taken from [1]. The values of thermal conductivity λ_{sat} and λ_{dry} are used as for the FEBEX bentonite reported in [12]. The value for the molecular diffusion coefficient of vapour in the air is taken from [6]. The parameters for the retention curve (Eq. 15) are obtained via regression analysis from test on SBM [3].

Figure 4 presents the comparison between the calculated and measured S_l in the hydration test before and after optimization based on the data from only this test. Fig. 5(a) presents the evolution of the temperature and S_l obtained using the initial model parameter set. Fig. 5(b) depicts the model response after the parameter set optimization using only data from this test.

There are two types of tests to calibrate the coupled THM properties of SBM. When consider the boundary conditions during our tests it can be concluded that there are four possible strategies for identification of the model parameters: (1) to identify the \mathcal{H} and \mathcal{M} parameters using solely the hydration test back analysis; (2) to identify \mathcal{H} , \mathcal{M} and \mathcal{T} parameters via back analysis base of solely the heating test; (3) to identify \mathcal{T} model parameters based on heating test when using \mathcal{H} and \mathcal{M} calibrated via (1); and (4) to identify \mathcal{M} model parameters in hydration test based on \mathcal{H} obtained by inverse modelling of the heating test. The results of the application of these four strategies are given in the Table 3.

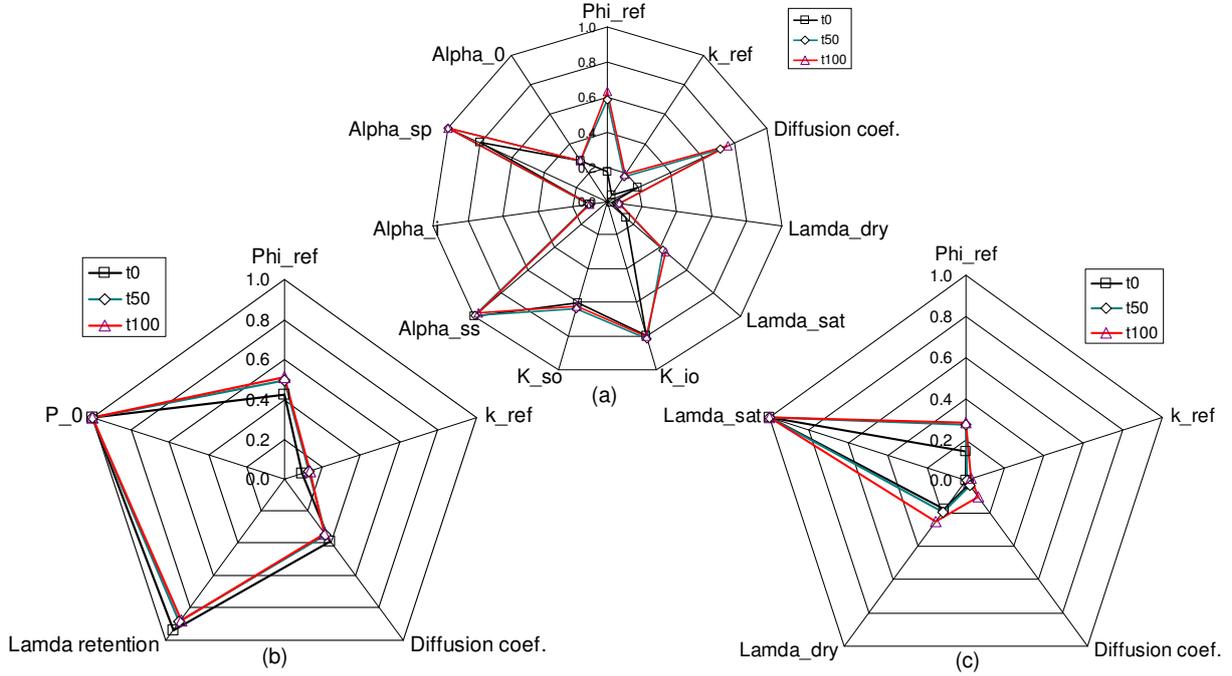


Figure 3: Heating test: (a)– γ_j of vertical stress (σ_{yy}). (b)– γ_j of degree of saturation. (c)– γ_j of temperature

5.3 Assessment of the quality of the optimized parameters

Table 1 presents the assessment of the quality of the identified parameters. The results show that P_0 in hydraulic equation and α_{sp} have a strong influence on the model response. Therefore these two parameters are calibrated most reliably.

5.4 Assessment of the goodness of the fit

In order to assess the goodness of the fit in each result after optimization, residual analysis method is adopted. The mean value, standard deviation and skewness are computed to assess the normality of the residuals.

Table 2 presents the result of the residual analysis. The results indicated that the optimization of hydration test obtained the best fit between measurement and numerical simulation.

6 CONCLUSION

The paper presents the strategy to identify the parameters for coupled THM model of sand bentonite mixture (SBM). Direct back analysis to two types of experiments is applied to identify the model parameters to which the model response is the most sensitive. Further, the quality of the obtained optimal set of parameters is assessed. One may expect

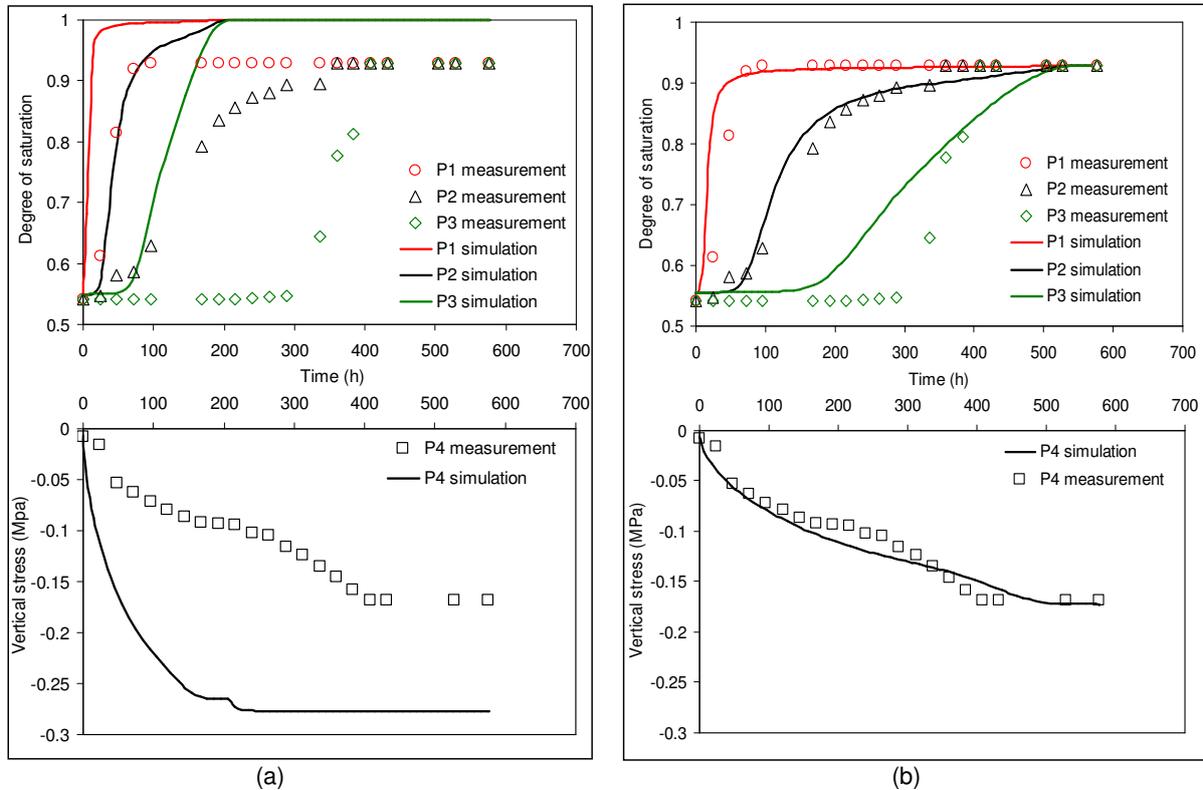


Figure 4: Hydration test–simulation vs. measurement: (a) Before optimization, (b) After optimization

that the best model calibration can be obtained combining the hydration and the heating test data, however the result of out investigation show that the independent back analysis of the heating test provides the best and most reliable set of model parameters.

Acknowledgements

Maria Datcheva acknowledges the financial support of the Bulgarian National Science Fund for a sabbatical leave to Ruhr-Universität Bochum under grant DSAB 02/6.

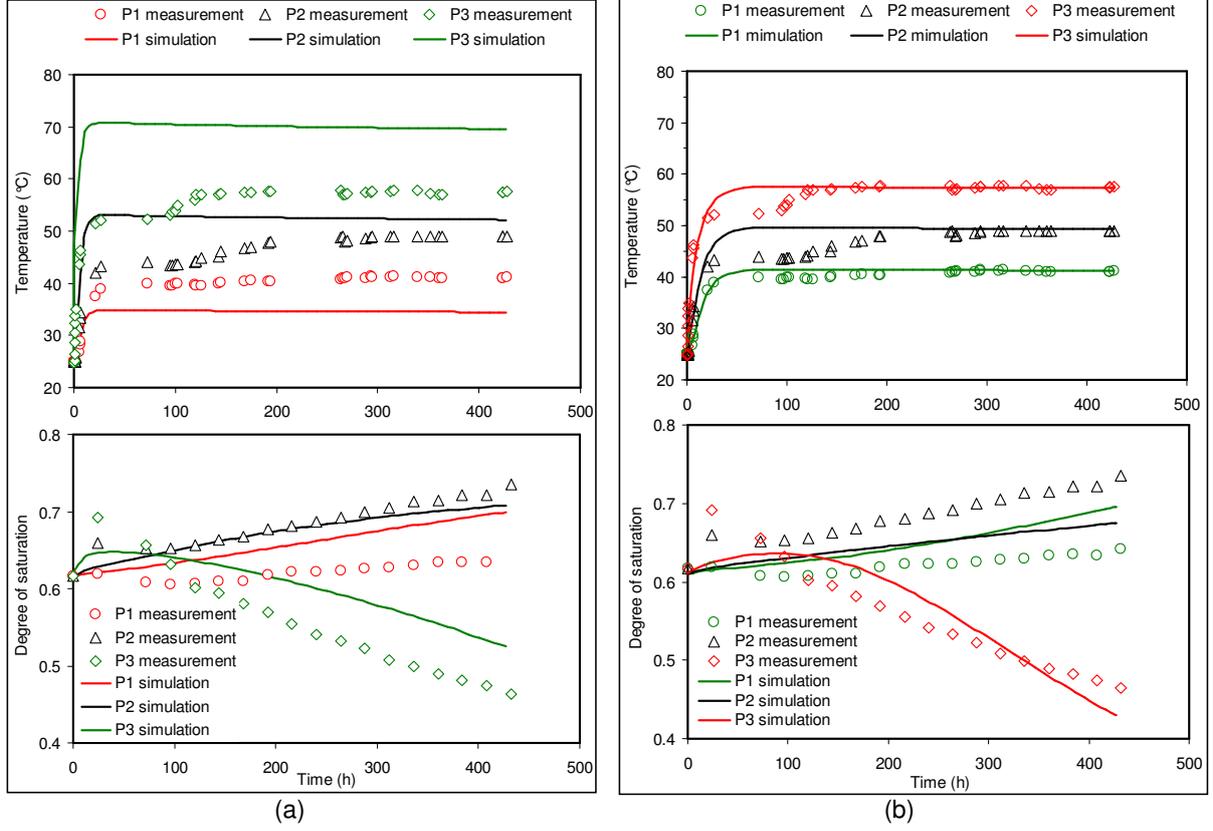


Figure 5: Heating test-simulation vs. measurement: (a) Before optimization, (b) After optimization

Table 1: The quality of the optimized parameters

$\{y_i\}$	$\{x_j\}$										
	k_{io}	k_{so}	α_{ss}	α_i	α_{sp}	P_0	λ	k_o	D	λ_{dry}	λ_{sat}
THM:											
S_l	0.000	0.000	0.000	0.000	0.000	1.000	0.895	0.118	0.355	0.000	0.000
T	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.014	0.059	0.214	1.000
σ_{yy}	0.810	0.623	1.000	0.103	0.911	0.000	0.000	0.116	0.478	0.046	0.291
HM:											
S_l	0.000	0.000	0.000	0.000	0.000	1.000	0.828	0.271	0.000	0.000	0.000
σ_{yy}	0.100	0.099	0.182	0.035	1.000	0.000	0.000	0.392	0.000	0.000	0.000
ν	0.182	0.144	0.236	0.028	0.382	0.400	0.345	0.182	0.178	0.052	0.258

Table 2: Residual analysis for the back analyzes

	Hydration	Heating	Heating →Hydrat.	Hydrat. →Heating
Mean value	-0.0035	0.0127	0.0216	0.00785
Standard deviation	0.0704	0.0400	0.0837	0.04140
Skewness	0.0498	0.0793	0.5108	-0.13469

Table 3: Summary of the parameters before and after back analysis

Par.	Unit	Initial	Constrain Min.	Constrain Max.	Hydrat.	Heating	Heating => Hydration	Hydrat. => Heating
<i>TEP Elastic Parameters</i>								
k_{io}	-	0.0029	0.0019	0.0039	0.0033	0.0029	0.0037	0.0033
k_{so}	-	0.1426	0.1	0.2	0.1468	0.1426	0.181	0.1468
a_{ss}	-	-0.1128	-0.09	-0.18	-0.103	-0.1128	-0.141	-0.103
a_i	-	-0.006	-0.003	-0.009	-0.0063	-0.006	-0.0069	-0.0063
a_{sp}	-	-0.3	-0.15	-0.50	-0.333	-0.3	-0.327	-0.333
<i>Hydraulic and thermal parameters</i>								
P_0	MPa	15	7.00	25.00	16.35	15	19.73	16.35
λ	-	0.53	0.40	0.80	0.564	0.543	0.543	0.564
k_o	(m^2)	2.07E-19	5.0E-22	1.00E-18	5.62E-20	3.59E-21	3.59E-21	5.62E-20
D	(*)	5.90E-06	1.0E-06	1.00E-05	5.90E-06	6.36E-06	6.36E-06	6.10E-06
τ	-	0.8	0.70	1.10	0.8	0.79	0.79	0.83
n	-	2.3	1.50	3.00	2.3	2.5	2.5	2.48
λ_{sat}	-	1.507	1.20	1.80	1.507	1.749	1.749	1.560
λ_{dry}	-	1.00	0.70	1.20	1.00	1.18	1.18	1.00

(*) : $m^2 s^{-1} K^{-n} Pa$

References

- [1] Agus, S. S. *An experimental study on hydro-mechanical characteristics of compacted bentonite-sand mixtures*. PhD thesis Bauhaus-University Weimar, (2005).
- [2] Alonso, E., Gens, A., and Josa, A. A constitutive model for partially saturated soils. *Geotechnique* (1990) **40(3)**: 405–430.
- [3] Arifin, Y. F. *Thermo-Hydro-Mechanical Behavior of Compacted Bentonite-Sand Mixtures: An Experimental Study*. PhD thesis Bauhaus-University Weimar, (2008).
- [4] DIT-UPC. *CODE_BRIGBT user's guide*. Universitat Politècnica de Catalunya, Barcelona, Spain 2009.
- [5] Geiser, F., Laloui, L., and Vulliet, L. Constitutive modelling of unsaturated sandy silt. *Computer Methods and Advances in Geomechanics, Rotterdam* (1997) 899–907.
- [6] Gens, A., Garcia-Molina, A. J., Olivella, S., Alonso, E. E., and Huertas, F. Analysis of a full scale in situ test simulating repository conditions. *Int. J. Numer. Anal. Meth. Geomech.* (1998) **22**: 515–548.
- [7] Genuchten, M. T. V. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* (1980) **44**: 892 – 898.
- [8] Lloret, A., Villar, M., Sanchez, M., Gens, A., Pintado, X., and Alonso, E. E. Mechanical behaviour of heavily compacted bentonite under high suction changes. *Geotechnique* (2003) **53(1)**: 27 – 40.
- [9] Manju, M., Schanz, T., and Tripathy, S. A column device to study THM behaviour of expansive soils. In *The 12th International Conference of International Association for Computer Methods and Advances in Geomechanics (IACMAG)* (October, 2008, Goa, India).
- [10] Nelder, J. A., and Mead, R. A simplex method for function minimisation. *The Computer Journal* (1965) **7**: 308 – 313.
- [11] Schanz, T., Zimmerer, M., and Datcheva, M. Identification of coupled hydro-mechanical parameters with application to engineered barrier systems. In *Unsaturated Soils: Advances in Geo-Engineering - Toll et al. (eds) 1st European Conference on Unsaturated Soils, Durham, UK* (2008) pp. 797–803.
- [12] Sánchez, M., Gens, A., and Olivella, S. Thermo-hydro-mechanical modelling of low permeability media using a double-porosity formulation. *Mecanica Computacional* (2004) **XXIII (7)**: 733 – 754.
- [13] Zimmerer, M. M. *VARO²PT. User Manual*. (2010).

MODELLING WATER INFILTRATION INTO MACROPOROUS HILL SLOPES USING SPECIAL BOUNDARY CONDITIONS

L. STADLER^{*}, C. ADAMCZAK AND R. HINKELMANN

Department of Civil Engineering, Chair of Water Resources Management and Modeling of Hydrosystems, TU Berlin, sec. TIB1-B14, Gustav-Meyer-Allee 25, 13355 Berlin, Germany
e-mail: leopold.stadler@wahyd.tu-berlin.de, www.wahyd.tu-berlin.de

Key words: Dual-Permeability Model, System Depended Boundary Conditions, Macropores, Two-Phase Flow.

Abstract. The formulation of suitable boundary conditions is a very crucial task when modeling water infiltration into macroporous hill slopes. The processes of water infiltration and exfiltration vary in space and time and depend on the flow on the surface as well as in the subsurface. In this contribution we have purposed special system process dependent boundary conditions can be formulated for a two-phase dual-permeability model to simulate infiltration and exfiltration processes. The presented formulation analyses the saturation conditions of the dual-permeability model (e.g. saturation) at the boundary nodes and adopts the boundary conditions depending on the processes at the soil surface such as rainfall intensity. Using a simplified macroporous hill slope and a heavy rainfall event we demonstrate the functionality of our formulation.

1 INTRODUCTION

Fast water infiltration into hill slopes during rainfall is an important issue since it can reduce slope stability and act as trigger for landslides. Modelling the water infiltration is an important key to understand the processes that can lead to a slope failure. The soil of natural slopes may be highly strongly heterogeneous and contains often macropores which strongly affect water flow¹. For this reason, dual-permeability models are frequently used to simulate the coupled flow processes in such macroporous soils, where the soil is separated into two coupled overlaying domains, a matrix domain containing the small matrix pores and a macropore domain containing larger pores (e.g. earthworm channels, fissures and fractures). Separate balance equations are defined for each domain and mass transfer functions are introduced to describe the fluid exchange depending, for example on the pressure differences between macropore and matrix and the resistance along its interface. Dual-permeability models that simulate the water infiltration into the unsaturated zone are typically based on the Richards equation for both domains as found in Gerke and van Genuchten². A more general dual-permeability model can be obtained when applying the two-phase flow equations instead of the Richards equation. This is necessary when the mobility of the soil air must be taken into account and the air pressure deviates from atmospheric pressure. Typical examples for that are strongly heterogeneous and layered soils where water is ponding and soil air escape is limited³. As such a case is investigated here, we decided to use the two-phase (water/air) dual-

permeability model developed by Stadler et al.⁴ of the multi-scale multi-physics toolbox DuMux⁵ for our work.

A special difficulty for modelling the water infiltration processes with a dual permeability model is the definition of reasonable boundary conditions. Most of the rainwater will usually infiltrate via the matrix into the soil until the infiltration capacity of the matrix is exceeded. When surface runoff occurs, water will also infiltrate directly into surface connected macropores. Consequently, the formulation of boundary conditions must be flexible since those for matrix and macropores are coupled and depend on the pressure and saturation conditions which vary in space and time. Such special boundary conditions are also called system-dependent boundary conditions⁶. In this paper we discuss all the different cases which can occur during infiltration and exfiltration and we describe the way they are implemented as boundary conditions in an external module of our two-phase dual-permeability model within DuMu^x.

2 DUAL-PERMEABILITY MODEL

Our dual-permeability model is based on the separation of the soil pores into matrix pores and macropores. Mass balance equations combined with the extended Darcy's law are first defined for both pore systems (domains) separately. They are then linked by a mass transfer equation to describe the fluid exchange between matrix and macropore domain. A detailed review of models and concepts for dual-permeability models can be found in Šimůnek et al.⁷.

2.1 Model equations

The balance equations for a two-phase flow dual-permeability system for the wetting phase w and the non-wetting phase n can be written for the matrix domain m and the macropore domain f as:

$$\begin{aligned}
 \phi^m \frac{\partial(S_w^m \rho_w^m)}{\partial t} + \nabla \cdot (\rho_w^m \vec{v}_w^m) - q_w^m &= \Gamma_w \\
 \phi^m \frac{\partial(S_n^m \rho_n^m)}{\partial t} + \nabla \cdot (\rho_n^m \vec{v}_n^m) - q_n^m &= \Gamma_n \\
 \phi^f \frac{\partial(S_w^f \rho_w^f)}{\partial t} + \nabla \cdot (\rho_w^f \vec{v}_w^f) - q_w^f &= -\Gamma_w \\
 \phi^f \frac{\partial(S_n^f \rho_n^f)}{\partial t} + \nabla \cdot (\rho_n^f \vec{v}_n^f) - q_n^f &= -\Gamma_n.
 \end{aligned} \tag{1}$$

Where ϕ^i [$L^3 L^{-3}$] is the porosity of a domain i (matrix/macropore), S_α^i [$L^3 L^{-3}$] the fluid saturation of a phase α , ρ_α^i [$kg m^{-3}$] the density, \vec{v}_α^i [$m s^{-1}$] the vector of the Darcy velocity, q_α^i [$kg s^{-1}$] is a source/sink term and Γ_α [$kg s^{-1}$] a mass transfer term that describes the exchange between matrix and macropore domain. In a two-phase water/air system the water will be the wetting phase and the air the non wetting phase. The pressure difference between both phases in each domain (i) is equivalent to the capillary pressure p_c^i [Pa] and can be

described as function of the effective saturation S_e^i .

$$p_c^i(S_e^i) = p_n^i - p_w^i \quad (2)$$

$$S_e^i = (S_w^i - S_{wr}^i)/(1 - S_{wr}^i - S_{nr}^i). \quad (3)$$

$S_{\alpha r}^i$ is the residual saturation of a fluid. In the presented work we use the formulation after van Genuchten⁸ to compute the capillary pressure relationship:

$$p_c^i = \frac{1}{\alpha} [(S_e^i)^{-1/m} - 1]^{1/n}. \quad (4)$$

The van Genuchten parameters α, m, n are parameters that describe the shape of the relationship. The parameters depend on the soil properties and are different for each domain of the dual-permeability model. The fluids in each domain fill the full pore space of the domain so that the sum of both saturations in each domain is equal to one:

$$S_w^i + S_n^i = 1. \quad (5)$$

The Darcy velocity of a phase in a domain can be computed with the extended Darcy law:

$$\vec{v}_\alpha^i = -\underline{\underline{K}} \frac{k_{r\alpha}^i}{\mu_\alpha^i} (\text{grad } p_\alpha^i - \rho_\alpha^i \mathbf{g}), \quad (6)$$

where $\underline{\underline{K}}$ [m^2] is the intrinsic permeability tensor, μ_α^i [$\text{kg m}^{-1} \text{s}^{-1}$] the dynamic viscosity and k_α^i [-] is the relative permeability which can be calculated with the van Genuchten relationship in combination with the model of Mualem⁹:

$$k_{rw}^i = \sqrt{S_e^i} \left[1 - (1 - (S_e^i)^{1/m})^m \right]^2 \quad (7)$$

$$k_{rn}^i = (1 - S_e^i)^{1/3} \left[1 - (S_e^i)^{1/m} \right]^{2m}. \quad (8)$$

Where m, n are again the van Genuchten parameters which depend on the considered domain. The mass transfer between both domains is approximated by a first-order transfer equation¹⁰ that depends on the pressure differences between both domains

$$\Gamma_\alpha = s \rho_\alpha \lambda_\alpha \beta (p_\alpha^m - p_\alpha^f), \quad (9)$$

s [m^{-1}] is a scaling factor between the regarded soil volume [m^3] and the macropore surface. β [m] is a surface resistance parameter and λ_α [$\text{kg m}^{-1} \text{s}^{-2}$] the mobility (relative permeability over dynamic viscosity).

2.2 Numerical model

The four balance equations (eq. 1) of the two-phase dual-permeability model together with supplementary and further conditions (eqs. 2-9) yield to a strongly coupled system of four

non-linear partial differential equations with mixed parabolic / hyperbolic character. DuMu^x applies a local and global mass conservative box method (mixture of Finite-Element and Finite-Volume Method) for the spatial discretization of the dual permeability model. The time discretization is done with a full implicit Euler scheme¹¹. Further, the Newton-Raphson Method is used for the linearization of the system¹².

We selected the pressures of the non-wetting phase and the saturation of the wetting phase as primary variables. The switch of the boundary conditions presented in the following is determined by an analysis of the values on the actual time level while the primary variables are computed on the new time level.

3 DEFINITION OF SYSTEM DEPENDENT BOUNDARY CONDITIONS

Natural systems like hill slopes are characterized by a strong temporal and spatial variation of subsurface (e.g. saturated/unsaturated) and surface-water flow conditions (e.g. overland flow/dry conditions). It is urgently required to simulate subsurface flow in a natural slope with varying boundary conditions since soil and surface are representing a coupled system. The most common cases where boundary conditions must be adopted are water infiltration during rainfall and water exfiltration during saturated conditions (see Fig. 1).

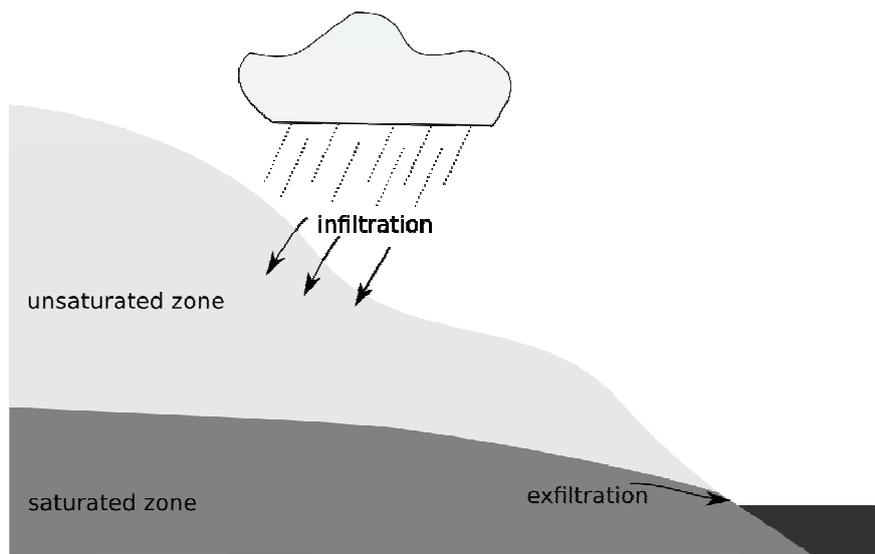


Figure 1: Water infiltration and exfiltration.

Surface runoff can occur in steep hill slopes and when the permeability and / or rainfall intensity are high. However, surface runoff is not taken into account here. The impact of this simplification is reduced when the hill slope gets flatter and when the permeability and rainfall intensity are getting smaller. The presented concept can be easily extended and coupled with surface runoff models.

It is possible to define four different inner states of the system (Fig. 2, left) which depend on the soil conditions in the matrix and macropore domain. In combination with the available water for infiltration, this leads to eight possible cases which must be distinguished for the definition of which will be explained in the following. The available water can be water from

a rainfall event, overland flow or ponding surface system-dependent boundary conditions and water.

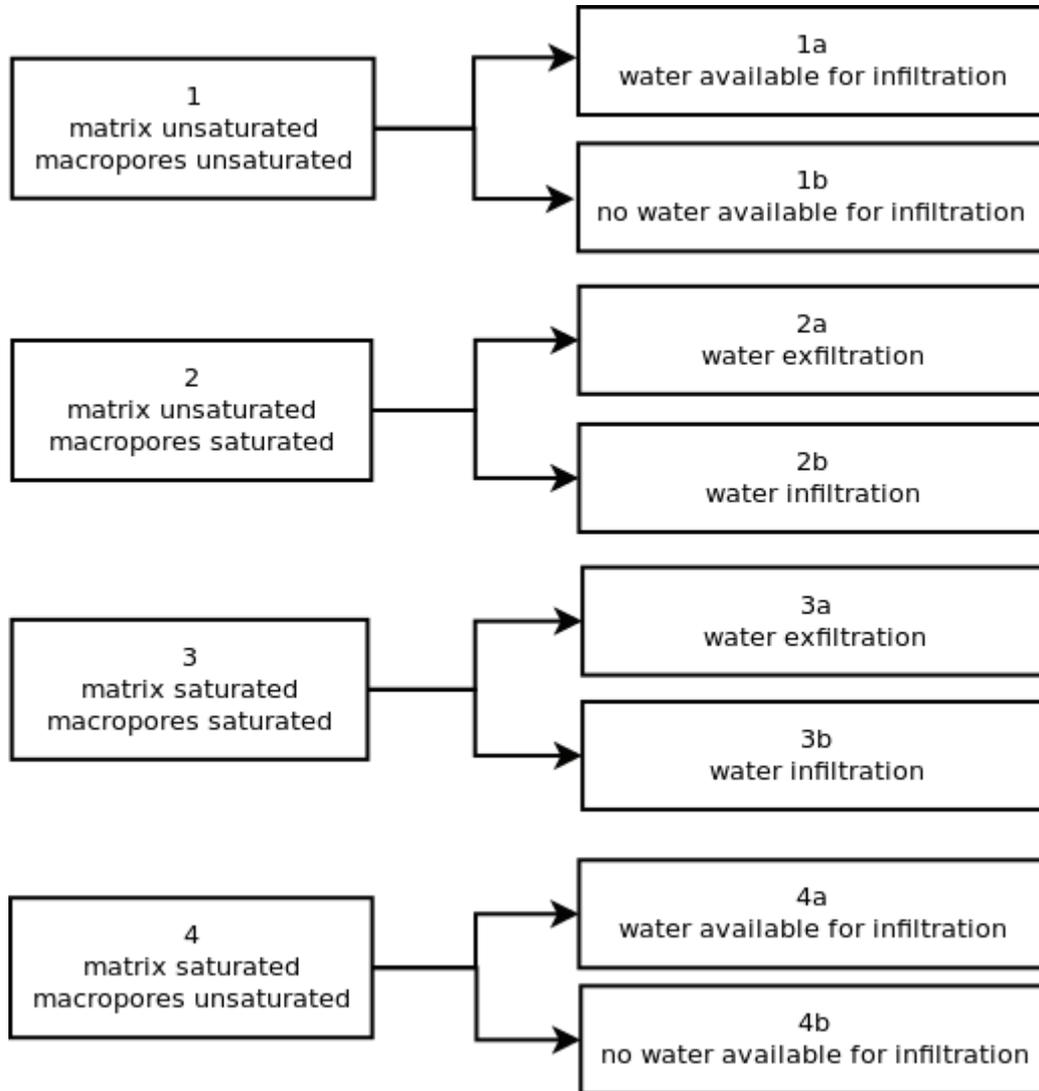


Figure 2: Possible states matrix and macropore domain (left) in combination with possible states at the soil surface (right).

In a first step it is necessary to determine the system state. Therefore, the saturations at boundary nodes are regarded. If a cell is saturated, the mass fluxes (Fig. 3) are additionally computed to analyze whether water is infiltrating or exfiltrating over the boundary surface. The corresponding cell fluxes over the cell surfaces (inner boundaries) are $F_{surface}$ and the fluxes between matrix and macropore domain are mass transfers $F_{transfer}$. If the sum of both fluxes is positive, outflow over the boundary will occur when the element is saturated. A negative sign indicates that water will infiltrate over the surface boundary.

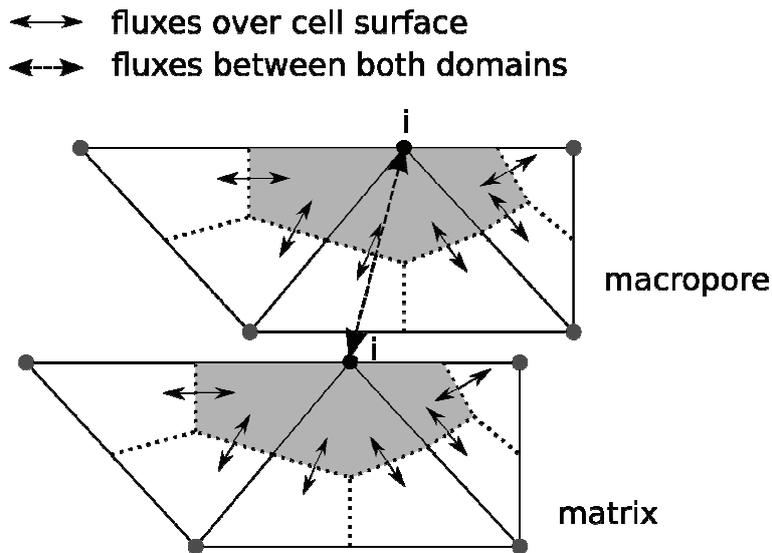


Figure 3: Fluxes at boundary node (i) for matrix and macropore domain and mass transfer fluxes for fluid exchange between both domains.

In the following we discuss the four possible cases and the corresponding sub-cases which can occur. Based on the presented concept Dirichlet and Neumann boundary conditions can be prescribed for both domains. The pressure of the gas phase is chosen as primary variable in both domains and is set to atmospheric pressure as long no water table stands above the soil surface. The saturation of the water phase is variable and Dirichlet or Neumann boundary conditions are set depending on the system state.

3.1 Case 1 – unsaturated matrix and unsaturated macropore domain

No ponding of water occurs as long as both domains are unsaturated. Thus, atmospheric pressure is set as Dirichlet boundary condition for the gas phase of both domains (**1a, 1b**). A Neumann no-flow boundary condition is set for the water phase in both domains (**1b**) if no water is available for infiltration. If water is available, it is checked whether the actual infiltration capacity $q_{w,\max}^m$ of the matrix is exceeded. If yes, the remaining water infiltrates via the macropores ($q_w^m = q_{w,\max}^m$ and $q_w^f = q_w^f - q_{w,\max}^m$). If not, the whole water infiltrates via the matrix ($q_w^m = q_i$ and $q_w^f = 0$). Due to the high macropore conductivity there is usually no limitation for the water infiltration into the macropores until they are saturated.

3.2 Case 2 - unsaturated matrix and saturated macropore domain

Macropore flow will usually only occur when the matrix is saturated. However, in some cases the water can bypass an unsaturated matrix and flow through the macropore domain. This is a special case which may occur at the toe of a slope where water can flow out through saturated macropores. Regardless of whether water infiltrates or exfiltrates (**2a, 2b**), atmospheric pressure is set for the gas phases in both domains (Dirichlet boundary condition).

The definition of boundary conditions is very complex for case 2 since a high non-equilibrium between matrix and macropore domain exists. For single domain concepts water

will only exfiltrate when the soil is fully saturated. This can be also prescribed for dual-permeability models when a Neumann no-flow boundary condition for both domains is set. However, the water can bypass an unsaturated matrix and escape to the surface. Thus, a Neumann boundary condition is set for the water phase of the matrix and a Dirichlet boundary condition is set for the macropore domain (**2a**). The water can also infiltrate from the surface into saturated macropores during infiltration (**2b**). If more water than available infiltrates via the macropores the Dirichlet boundary condition for the water phase is switched to a Neumann boundary condition.

3.3 Case 3 – saturated matrix and saturated macropore domain

For this case it is necessary to check the mass fluxes at the boundary node to control whether water is infiltrating or exfiltrating. As mentioned above, surface runoff can be neglected for macroporous slopes and the pressure in both domains is set as Dirichlet boundary condition, assuming atmospheric pressure during ex-filtration (**3a**). The influence of the pressure increase due the water level will be negligible for small water depths. However, when water stands above the surface (e.g. river), the pressure must be adopted. For infiltration (**3b**) it is checked whether if the infiltration capacity exceeds the available water and in case a switch to a Neumann boundary condition for the water phases in one or both domains (depending on the infiltration rates) is carried out.

3.4 Case 4 – saturated matrix and unsaturated macropores

The last case generally occurs during infiltration (**4a**) if the infiltration capacity $q_{w,max}^m$ of the matrix is lower than the available water for infiltration. Then the rest of the available water will infiltrate via the macropores ($q_i - F_{w,transfer} - F_{w,matrix} = q_w^f$). This case is implemented by a Neumann boundary condition for the water phase of the macropore domain and Dirichlet boundary conditions for the water phase in the matrix domain (fully saturated). The pressure of the gas phase is set again to atmospheric pressure. If no water is available for infiltration, outflow may occur via the matrix pores (**4b**). However, the water will directly infiltrate into the macropores. A Neumann no-flow boundary condition is set for the water phase in the matrix domain. This leads to an increasing saturation in the macropore domain and avoids water exfiltration until the macropores are saturated.

4 EXAMPLE

Common examples where system-dependent boundary conditions can be demonstrated are small slopes where the water infiltration leads to an increasing groundwater table during infiltration. The model domain (Figure 4) for our example is a simplified macroporous hill slope similar as shown in Figure 1. The soil parameters for the study are given in Table 1. A rainfall event with a intensity of 40mm/h and a duration of two hours is investigated to study water infiltration. The groundwater table at the initial state (Figure 3a) is influenced by the water body at the right side where a water level (e.g. lake) is imposed using Dirichlet boundary conditions for the matrix and macropore domain. The level of the water body is assumed to be constant during the whole simulation time. The initial saturation in the matrix and macropore domain is very low (~ 0.3). The rainfall intensity will exceed the infiltration

capacity of the matrix so that macropore flow will occur directly. Soil parameters and initial conditions are chosen to test most of the possible cases during this extreme situation.

Starting with a low saturated zone (Figure 3a), the water table increases mainly in the right part of the hill slope during water infiltration, the nodes on the right boundary above the water table get saturated and switch to Dirichlet conditions (Figure 3b, right). Most of the water infiltrates via the macropores and bypasses the matrix because of the low permeability of the matrix domain. After the rainfall event has finished, the saturation of the boundary nodes at the right side is reduced and they switch back to Neumann no-flow boundary conditions (Figure 3c, right).

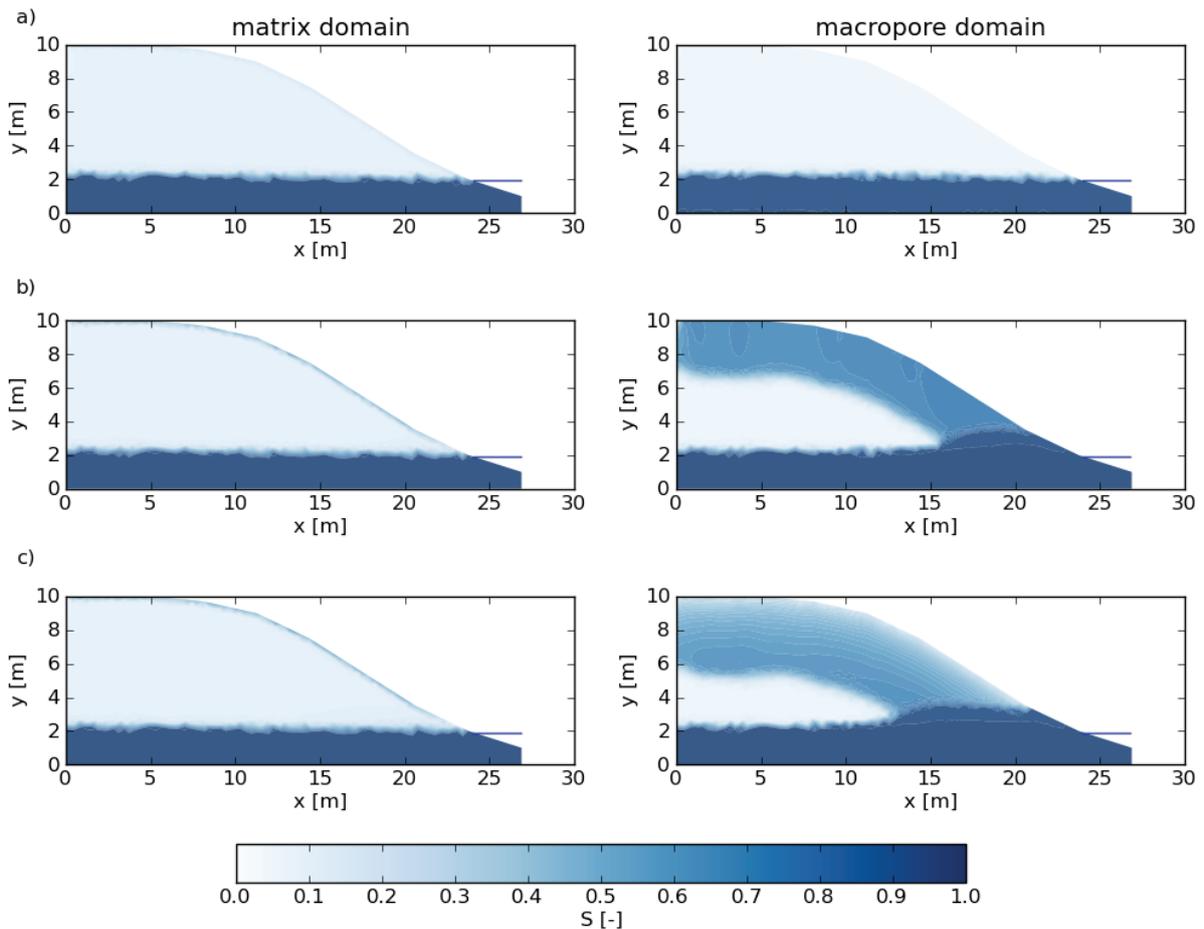


Figure 4: Water saturation of the matrix (left) and macropore domain (right) for various times (a = 0h, b = 2h and c = 2.5h) for an infiltration rainfall rate of 40mm/h and a period of two hours.

Table 1: Soil parameters

	matrix domain	macropore domain
S_{wr} [-]	0.14	0.05
S_{nr} [-]	0.05	0.05
K [m^2]	1.0E-13	1.0E-11
van Genuchten n [-]	2.2	4.2
van Genuchten α [Pa^{-1}]	5.0E-4	8.0E-3
Φ [-]	0.4	0.08
β [m]	2.0E-13	2.0E-13
s [m^{-1}]	1	1

5 CONCLUSIONS

In this paper we have proposed a concept to simulate simplified interactions between surface and subsurface flow on macroporous hill slopes where the overland flow is not taken into account yet. The two-phase dual-permeability model of the numerical simulator DuMux was extended by special system-dependent boundary conditions to simulate infiltration and exfiltration processes on macroporous hill slopes. An idealized system with a low matrix infiltration capacity was investigated to test the model capabilities during a high rainfall event. Due to the low permeability of the soil matrix the rainwater bypassed the soil matrix and water infiltration and exfiltration mainly occurred via the macropores. In future work the proposed concept will be further extended and coupled with a surface runoff model.

REFERENCES

- [1] Beven, K. and Germann, P. Macropores and water flow in soils, *Water Resour. Res.* (1982) **18**:1311-1325.
- [2] Gerke, H.H. and van Genuchten, M.T. A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media. *Water Resour. Res.* (1993) **29**:305-319.
- [3] Stadler, L., Hinkelmann R. and Zehe E. Two-phase flow simulation of water infiltration into layered natural slopes inducing soil deformation, in *Landslides Processes - From Geomorphologic Mapping to Dynamic Modelling*, edited by J.-P. Malet and Remaitre, (2009) 197-201.
- [4] Stadler, L., Hinkelmann R. and Helmig, R. Development of a Dual-Permeability Model for Two-Phase Flow in Macro-porous Media. *Transport in Porous Media*. In preparation for submission.
- [5] Flemisch, B., Darcis, M., Erbertseder, K., Faigle, B., Lauser, A., Mosthaf, K., Müthing, S., Nuske, P., Tatomir, A., Wolff, M. and Helmig R. DuMux: DUNE for Multi-{Phase, Component, Scale, Physics, ...} Flow and Transport in Porous Media. *Adv. Water. Res* (2010), Submitted.
- [6] Radcliffe, D. and Simunek, J. *Soil physics with hydrus: modeling and applications*. Crc Pr Inc, (2010).

- [7] Šimůnek, J., Jarvis, N.J., van Genuchten M.T. and Gardenas A. Review and comparison of models for describing non-equilibrium and preferential flow and transport in the vadose zone. *Journal of Hydrology* (2003) **272**:14-35.
- [8] Van Genuchten, M.T. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J* (1980) **44**:892-898.
- [9] Mualem Y. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res* (1976) **12**:513-521.
- [10] Gerke, H.H. and van Genuchten, M.T. Evaluation of a 1st-order water transfer term for variably saturated dual-porosity flow models. *Water Resour. Res* (1993):1225-1238.
- [11] Helmig R. *Multiphase flow and Transport Processes in the Subsurface*. Springer, (1997).
- [12] Hinkelmann, R. *Efficient Numerical Methods and Information-Processing Techniques for Modeling Hydro- and Environmental Systems, Lecture Notes in Applied and Computational Mechanics*. Springer, Berlin, Heidelberg, New York, Vol. 21, (2005)

A PARTITIONED QUASI-NEWTON SOLUTION TECHNIQUE FOR FLUID-STRUCTURE INTERACTION PROBLEMS USING A COARSENEDED GRID TO ACCELERATE THE CONVERGENCE OF THE COUPLING ITERATIONS

JORIS DEGROOTE* AND JAN VIERENDEELS*

*Department of Flow, Heat and Combustion Mechanics
Ghent University
Sint-Pietersnieuwstraat 41, B-9000 Ghent, Belgium
e-mail: {Joris.Degroote,Jan.Vierendeels}@UGent.be, www.FSI.UGent.be

Key words: partitioned, fluid-structure interaction, quasi-Newton, least-squares, grid

Abstract. Previous stability analyses on Gauss-Seidel coupling iterations in partitioned fluid-structure interaction simulations have demonstrated that Fourier modes with a low wave-number in the difference between the current and correct interface displacement are unstable. To stabilize these modes, the IQN-ILS technique automatically constructs a least-squares model of the flow solver and structural solver. In this work, the multi-level IQN-ILS technique (ML-IQN-ILS) is presented, which uses a coarsened grid of the fluid and structure subdomains to initialize this least-squares model. As the modes that need to be present in this least-squares model have a low wave-number, they can be resolved on a coarsened grid. Therefore, in each time step, a number of cheap coupling iterations is first performed on the coarsened grid to construct the model, followed by a smaller number of coupling iterations on the fine grid. As the iterations on the coarse grid are fast and fewer iterations are performed on the fine grid, the total duration of the simulation decreases compared to a simulation on the fine grid only.

1 INTRODUCTION

Partitioned fluid-structure interaction (FSI) simulation techniques solve the flow equations and the structural equations separately. In this article, the focus lies on partitioned techniques which couple the flow solver and the structural solver as ‘black boxes’, which means that the discretization and solution techniques of the solvers do not have to be known. Implicit (or strongly coupled) partitioned techniques enforce the equilibrium of the stress and velocity (or displacement) on the fluid-structure interface in each time step. Several strongly coupled partitioned techniques are able to couple ‘black-box’ solvers,

for example Gauss-Seidel iterations with Aitken relaxation [1], the Interface Generalized Minimal Residual method (Interface-GMRES) [2], the Interface Block Quasi-Newton technique with an approximation for the Jacobians from Least-Squares models (IBQN-LS) [3] and the Interface Quasi-Newton technique with an approximation for the Inverse of the Jacobian from a Least-Squares model (IQN-ILS) [4].

Several stability analyses on coupling algorithms have been performed for the incompressible, inviscid flow in a straight, elastic tube [5–8]. Fourier analysis has been performed on the difference between the current and the correct interface displacement during Gauss-Seidel coupling iterations in [7, 8]. From these analyses, two lessons can be learned. While the standard, one-level IQN-ILS technique only takes advantage of the first one, the new multi-level IQN-ILS (ML-IQN-ILS) technique takes advantage of both of them.

The first lesson is that only a fraction of the Fourier modes is unstable during Gauss-Seidel iterations. If a quasi-Newton technique is used, then only a low-rank approximation for the exact Jacobian is required, as long as it represents the behaviour of these unstable and slowly converging modes, which explains the performance of quasi-Newton methods like IQN-ILS. For combinations of Fourier modes that are covered by the least-squares model, IQN-ILS performs Newton iterations; for the other modes, IQN-ILS corresponds to Gauss-Seidel iterations.

The second lesson is that the unstable modes have a low wave number, so their behaviour can be determined on a relatively coarse grid. The new ML-IQN-ILS technique uses more than one grid level, each with a different number of grid points. It first calculates the coupled solution on the coarsest grid level and constructs the low-rank approximation for the inverse of the Jacobian as present in IQN-ILS while doing so. Then coupling iterations are performed on the second grid level, during which the approximation for the inverse of the Jacobian obtained on the coarsest grid level is further improved. This procedure is repeated until the solution on the finest grid level has been found.

The goal of the multi-level IQN-ILS technique is thus to obtain the low-rank approximation for the inverse of the Jacobian required for the convergence of the coupling iterations on the finest grid level at a lower cost, by constructing it partly on coarser grid levels. This new multi-level approach is depicted in Figure 1 for two grid levels. As only data on the fluid-structure interface is exchanged, this partitioned multi-level coupling technique can couple black-box solvers. The name multi-grid is not used because it already refers to a common solution technique [9], which has been used for fluid-structure interaction simulations in for example [10, 11] and which is different from the ML-IQN-ILS technique.

The remainder of this article is organized as follows. Section 2 gives a brief overview of the governing equations, before the detailed explanation of ML-IQN-ILS in Section 3. Numerical results in Section 4 illustrate the performance of ML-IQN-ILS compared to the standard, one-level IQN-ILS. Finally, Section 5 offers the conclusions.

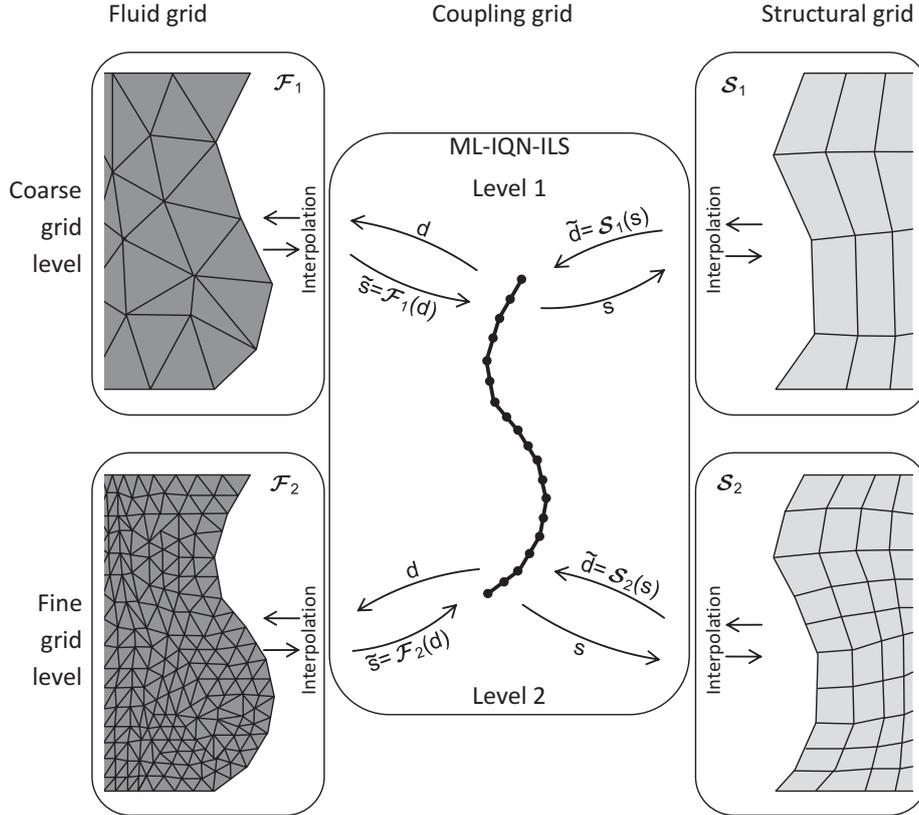


Figure 1: The coarse and fine fluid grid (left) and the coarse and fine structural grid (right), together with the unique coupling grid (centre) in a multi-level simulation with two grid levels. \mathbf{d} represents the displacement of the interface while \mathbf{s} represents the stress distribution on the interface. \mathcal{F} denotes the flow solver and \mathcal{S} the structural solver. The output of a solver is indicated with a tilde as this value is not always directly given as input to the other solver. In the multi-level IQN-ILS algorithm, coupling iterations are first performed on the coarse grid level (level 1) to construct the approximation for the inverse of the Jacobian as present in IQN-ILS at a lower cost. Subsequently, this approximation for the inverse of the Jacobian is used and improved further during the coupling iterations on the fine grid level (level 2), resulting in fewer coupling iterations on the fine grid level. All interface data on the different grid levels are interpolated to and from the unique coupling grid, which determines the dimension of the approximation for the inverse of the Jacobian.

2 GOVERNING EQUATIONS

The fluid and structure subdomains are indicated as Ω_f and Ω_s and their boundaries as Γ_f and Γ_s , respectively. The fluid-structure interface $\Gamma_{fs} = \Gamma_f \cap \Gamma_s$ is the common boundary of these subdomains. This article only considers incompressible fluids as they prove to be most challenging for the partitioned fluid-structure interaction techniques. The unsteady flow of an incompressible fluid is governed by the conservation of mass and the Navier-Stokes equations

$$\nabla \cdot \vec{v} = 0 \quad (1a)$$

$$\rho_f \frac{\partial \vec{v}}{\partial t} + \rho_f \nabla \cdot (\vec{v}\vec{v}) - \nabla \cdot \bar{\sigma}_f = \vec{f}_f. \quad (1b)$$

In these equations, ρ_f is the fluid density, \vec{v} the fluid velocity and t the time. \vec{f}_f represents the body forces per unit of volume on the fluid. The deformation \vec{u} of the structure is determined by the conservation of momentum

$$\rho_s \frac{d^2 \vec{u}}{dt^2} - \nabla \cdot \bar{\sigma}_s = \vec{f}_s \quad (2)$$

with ρ_s the structural density, $\bar{\sigma}_s$ the Cauchy stress tensor and \vec{f}_s the body forces per unit volume on the structure.

The equilibrium conditions on the fluid-structure interface ($\vec{x} \in \Gamma_{fs}$) are

$$\vec{v} = \frac{d\vec{u}}{dt} \quad \text{and} \quad \bar{\sigma}_f \cdot \vec{n}_f = -\bar{\sigma}_s \cdot \vec{n}_s, \quad (3)$$

which stipulate that the velocity and the stress have to be the same on both sides of the interface. The vector \vec{n}_f (\vec{n}_s) is the unit normal that points outwards from the subdomain Ω_f (Ω_s). A Dirichlet-Neumann decomposition of the fluid-structure interaction problem is applied, so the flow equations are solved with a given velocity (or displacement) of the fluid-structure interface and the structural equations are solved with a given stress on the interface.

The flow equations and the structural equations are discretized in space and time with a method of choice. The discrete flow equations are represented by \mathfrak{F} , the discrete structural equations by \mathfrak{S} . The vector \mathbf{v} groups all flow variables (velocity, pressure, etc.) in Ω_f ; the vector \mathbf{u} groups all structural variables (displacement, stress, etc.) in Ω_s . The displacement of the interface Γ_{fs} with respect to the initial geometry is represented by the vector \mathbf{d} and the stress on the interface by the vector \mathbf{s} . In the case of a Dirichlet-Neumann decomposition, the displacement of the interface is considered as a function of the structural degrees of freedom ($\mathbf{d} = \mathbf{d}(\mathbf{u})$) and the stress on the interface as a function of the flow degrees of freedom ($\mathbf{s} = \mathbf{s}(\mathbf{v})$). All variables are at the new time level t^{n+1} ; the dependence of the solution on the variables at t^n, t^{n-1}, \dots is hidden.

The flow solver calculates the flow variables \mathbf{v} that satisfy $\mathfrak{F}(\mathbf{v}, \mathbf{d}(\mathbf{u})) = \mathbf{0}$ for a given interface displacement \mathbf{d} . From the flow field \mathbf{v} , the stress on the interface \mathbf{s} is extracted. Therefore, the flow solver is represented by the function

$$\tilde{\mathbf{s}} = \mathcal{F}(\mathbf{d}). \quad (4)$$

Similarly, the structural solver calculates the structural variables \mathbf{u} that satisfy $\mathfrak{S}(\mathbf{u}, \mathbf{s}(\mathbf{v})) = \mathbf{0}$ for a given stress on the interface \mathbf{s} . The displacement of the interface \mathbf{d} is subsequently extracted from \mathbf{u} , so the structural solver is represented by

$$\tilde{\mathbf{d}} = \mathcal{S}(\mathbf{s}). \quad (5)$$

As the multi-level coupling technique uses several grid levels for the flow equations and the structural equations, data has to be interpolated between different discretizations of the fluid-structure interface. However, even though the discretization of the interface inside the flow solver and the structural solver depends on the grid level, all operations of the coupling algorithm are performed on a unique grid, the so-called ‘coupling grid’ (see Figure 1). In this work, this coupling grid is identical to the interface discretization of the finest fluid grid.

In line with the definition of the flow solver and structural solver as black-box functions, the interpolation on the interface should not require access to the discretization in the solvers. Therefore, interpolation with radial basis functions is applied. A local interpolant is constructed in the neighbourhood of each point on the interface using a basis function introduced by Wendland [12], namely

$$\phi(\|\vec{x}\|/r) = (1 - \|\vec{x}\|/r)_+^4(4\|\vec{x}\|/r + 1), \quad (6)$$

with r the radius and $\|\vec{x}\| = \sqrt{x_1^2 + \dots + x_d^2}$ the Euclidean distance of dimension d . The plus-sign behind the first term denotes that this term is zero if $1 - \|\vec{x}\|/r < 0$ such that ϕ has a compact support.

3 ML-IQN-ILS

In the explanation of this coupling algorithm, a prime denotes the Jacobian matrix of a function and a hat refers to an approximation. The output of a solver is indicated with a tilde as this value is not always directly given as input to the other solver. The grid level is indicated with a subscript i and the coupling iteration within time step $n + 1$ with a superscript k . The superscript $n + 1$ is omitted wherever possible. The standard algorithm with a single grid level is described first, followed by the multi-level algorithm with g grid levels. The first grid level is the coarsest grid level and the g^{th} grid level is the finest one.

The FSI problem reformulated as a set of nonlinear equations in the interface’s displacement

$$\mathcal{R}(\mathbf{d}) = \mathcal{S} \circ \mathcal{F}(\mathbf{d}) - \mathbf{d} = \mathbf{0} \quad (7)$$

can be solved by means of Newton-Raphson iterations

$$\text{solve } \mathcal{R}'^k \Delta \mathbf{d}^k = -\mathbf{r}^k \quad (8a)$$

$$\mathbf{d}^{k+1} = \mathbf{d}^k + \Delta \mathbf{d}^k \quad (8b)$$

with the residual calculated as

$$\mathbf{r}^k = \mathcal{R}(\mathbf{d}^k) = \mathcal{S} \circ \mathcal{F}(\mathbf{d}^k) - \mathbf{d}^k = \tilde{\mathbf{d}}^k - \mathbf{d}^k. \quad (9)$$

\mathcal{R}'^k denotes the Jacobian of \mathcal{R} , evaluated at \mathbf{d}^k . The Newton-Raphson iterations in the time step have converged when $\|\mathbf{r}^k\|_2 \leq \varepsilon$ with ε the convergence tolerance. However, the exact Jacobian of \mathcal{R} is unknown as the Jacobians of \mathcal{F} and \mathcal{S} are unavailable. Moreover, the linear system in Eq. (8a) with as dimension the number of degrees of freedom in the displacement of the fluid-structure interface has to be solved in each Newton-Raphson iteration. If the Jacobian \mathcal{R}' is approximated and quasi-Newton iterations are performed, black-box solvers can be used. However, the linear system in Eq. (8a) still needs to be solved. As will be explained below, it is more advantageous to approximate the *inverse* of the Jacobian by applying the least-squares technique introduced by Vierendeels et al. [3] on a particular set of vectors, which is done by the standard IQN-ILS algorithm. The quasi-Newton iterations with the approximation for the inverse of the Jacobian can be written as

$$\mathbf{d}^{k+1} = \mathbf{d}^k + \widehat{\Delta} \mathbf{d}^k = \mathbf{d}^k + \left(\widehat{\mathcal{R}'^k} \right)^{-1} (-\mathbf{r}^k). \quad (10)$$

It can be seen from Eq. (10) that the approximation for the inverse of the Jacobian does not have to be created explicitly; a procedure to calculate the product of this matrix with the vector $-\mathbf{r}^k$ is sufficient. The vector $-\mathbf{r}^k$ is the difference between the desired residual, i.e. $\mathbf{0}$, and the current residual \mathbf{r}^k and it is further denoted as $\widehat{\Delta} \mathbf{r}^k = \mathbf{0} - \mathbf{r}^k = -\mathbf{r}^k$. The matrix-vector product in Eq. (10) is calculated from information obtained during the previous quasi-Newton iterations. Eq. (9) shows that the flow equations and structural equations are solved in quasi-Newton iteration k , resulting in $\tilde{\mathbf{d}}^k = \mathcal{S} \circ \mathcal{F}(\mathbf{d}^k)$ and the corresponding residual \mathbf{r}^k . So, at the beginning of quasi-Newton iteration $k + 1$, a set of known residual vectors

$$\mathbf{r}^k, \mathbf{r}^{k-1}, \dots, \mathbf{r}^1, \mathbf{r}^0 \quad (11a)$$

and the corresponding set of vectors $\tilde{\mathbf{d}}$

$$\tilde{\mathbf{d}}^k, \tilde{\mathbf{d}}^{k-1}, \dots, \tilde{\mathbf{d}}^1, \tilde{\mathbf{d}}^0 \quad (11b)$$

are available. After each coupling iteration, the difference between the vectors from the current coupling iteration and the vectors from the previous coupling iteration is calculated using

$$\Delta \mathbf{r}^{k-1} = \mathbf{r}^k - \mathbf{r}^{k-1} \quad \text{and} \quad \Delta \tilde{\mathbf{d}}^{k-1} = \tilde{\mathbf{d}}^k - \tilde{\mathbf{d}}^{k-1}. \quad (12)$$

This yields a set of differences $\Delta \mathbf{r}^j$ and the corresponding set of differences $\Delta \tilde{\mathbf{d}}^j$ which both grow in each coupling iteration ($j = 0, \dots, k-1$). These vectors are stored as the columns of the matrices \mathbf{V}^k and \mathbf{W}^k . The number of columns in \mathbf{V}^k and \mathbf{W}^k is indicated with v which is not always equal to k as will be explained further and which is generally much smaller than the number of rows u . Nevertheless, in simulations with a low number of degrees of freedom on the interface, it is possible that the number of columns has to be limited to u by discarding the rightmost columns.

The vector $\widehat{\Delta \mathbf{r}}^k = \mathbf{0} - \mathbf{r}^k$ is approximated as a linear combination of the known $\Delta \mathbf{r}^j$

$$\widehat{\Delta \mathbf{r}}^k \approx \mathbf{V}^k \mathbf{c}^k \quad (13)$$

with $\mathbf{c}^k \in \mathbb{R}^{v \times 1}$ the coefficients of the decomposition. Because $v \leq u$, Eq. (13) is an overdetermined set of equations for the elements of \mathbf{c}^k and hence the least-squares solution to this linear system is calculated. For that reason, the so-called economy-size QR-decomposition of \mathbf{V}^k is calculated using Householder transformations

$$\mathbf{V}^k = \mathbf{Q}^k \mathbf{R}^k, \quad (14)$$

with $\mathbf{Q}^k \in \mathbb{R}^{u \times v}$ an orthogonal matrix and $\mathbf{R}^k \in \mathbb{R}^{v \times v}$ an upper triangular matrix. The coefficient vector \mathbf{c}^k is then determined by solving the triangular system

$$\mathbf{R}^k \mathbf{c}^k = \mathbf{Q}^{kT} \widehat{\Delta \mathbf{r}}^k \quad (15)$$

using back substitution. If a $\Delta \mathbf{r}^i$ vector is (almost) a linear combination of other $\Delta \mathbf{r}^j$ vectors, one of the diagonal elements of \mathbf{R}^k will (almost) be zero. Therefore, the equation corresponding to that row of \mathbf{R}^k cannot be solved during the back substitution. If a small diagonal element is detected, the corresponding columns in \mathbf{V}^k and \mathbf{W}^k are removed. Subsequently, the QR-decomposition (Eq. (14)) is performed again. This procedure is repeated until none of the diagonal elements is too small. The tolerance ε_s for the detection of small diagonal elements depends on how accurately the flow equations and structural equations are solved. An appropriate value for ε_s can be determined by analyzing the change of the vector $\tilde{\mathbf{d}}$ due to a small perturbation of the vector \mathbf{d} . If the perturbation is too small, the resulting change will be numerical noise. The value of ε_s should be chosen so that the change of $\tilde{\mathbf{d}}$ has a physical meaning if the perturbation of \mathbf{d} has an L^2 -norm larger than ε_s . If the solution of the flow equations and the structural equations is calculated with more significant digits, for example by using stricter convergence criteria inside the solvers, then a smaller value of ε_s can be used.

The $\widehat{\Delta \tilde{\mathbf{d}}}^k$ that corresponds to $\widehat{\Delta \mathbf{r}}^k$ is subsequently calculated as a linear combination of the previous $\Delta \tilde{\mathbf{d}}^j$, analogous to Eq. (13), giving

$$\widehat{\Delta \tilde{\mathbf{d}}}^k = \mathbf{W}^k \mathbf{c}^k. \quad (16)$$

From Eq. (9), it follows that $\Delta \mathbf{r} = \Delta \tilde{\mathbf{d}} - \Delta \mathbf{d}$ and substitution of Eq. (16) results in

$$\widehat{\Delta \mathbf{d}}^k = \mathbf{W}^k \mathbf{c}^k - \widehat{\Delta \mathbf{r}}^k. \quad (17)$$

Because the coefficients \mathbf{c}^k are a function of $\widehat{\Delta \mathbf{r}}^k$, Eq. (17) shows how $\widehat{\Delta \mathbf{d}}^k$ can be approximated for a given $\widehat{\Delta \mathbf{r}}^k$. Hence, Eq. (17) can be seen as a procedure to calculate the product of the approximation for the inverse of the Jacobian and a vector $\widehat{\Delta \mathbf{r}}^k = -\mathbf{r}^k$

$$\widehat{\Delta \mathbf{d}}^k = \left(\widehat{\mathcal{R}}'^k \right)^{-1} \widehat{\Delta \mathbf{r}}^k = \mathbf{W}^k \mathbf{c}^k + \mathbf{r}^k. \quad (18)$$

Algorithm 1 shows the Multi-Level IQN-ILS (ML-IQN-ILS) algorithm in detail. Lines 8 to 18 are the standard IQN-ILS algorithm as described above. Around the standard algorithm, an additional loop over the grid levels is added (line 5). First, the coupled solution is calculated on the coarsest grid level. Then, starting from that solution, coupling iterations on the following, finer grid level are performed. These steps are subsequently repeated for all grid levels until the solution on the finest grid has been found. The variable ℓ ensures that at least one coupling iteration is performed on each grid level.

The displacement and the residual are not changed when the grid level i changes, as both are defined on the coupling grid. As explained above, the coupling algorithm itself works with a unique coupling grid, which determines the dimension of the approximation for the inverse of the Jacobian. The different grid levels that are used for the multi-level technique are only present inside the flow solver and structural solver. The solvers have to interpolate the data from the boundary of their grid to the coupling grid of the coupling code. In this way, the acceleration of the coupling iterations and the interpolation of the data on the fluid-structure interface are completely separated, which facilitates the implementation.

Because the coupling algorithm operates on the coupling grid, the difference between \mathbf{r} and $\tilde{\mathbf{d}}$ in consecutive coupling iterations is always interpolated to a fixed number of grid points, regardless of the current grid level. As a result, the modes that have been generated on a coarse grid level can be used to accelerate the coupling iterations on the finer grid levels. The same least-squares model is used for all grid levels so the number of columns in the matrices \mathbf{V}^k and \mathbf{W}^k increases on each grid level. Because the matrices \mathbf{V}^k and \mathbf{W}^k have to contain at least one column to perform a quasi-Newton step, a Gauss-Seidel step using relaxation with factor ω (line 9) is performed in each time step, but only on the coarsest grid level.

The numerical experiments in Section 4 indicate that vectors $\Delta \mathbf{r}^j$ and $\Delta \tilde{\mathbf{d}}^j$ from a coarse grid level can accelerate the coupling iterations on a fine grid level. However, it should be noted that the difference between \mathbf{r} and $\tilde{\mathbf{d}}$ in the last coupling iteration on a certain grid level i and the first coupling iteration on the following grid level $i + 1$,

$$\Delta \mathbf{r}^{j-1} = \mathcal{R}_{i+1}(\mathbf{d}^j) - \mathcal{R}_i(\mathbf{d}^{j-1}) \quad (19a)$$

$$\Delta \tilde{\mathbf{d}}^{j-1} = \mathcal{S}_{i+1} \circ \mathcal{F}_{i+1}(\mathbf{d}^j) - \mathcal{S}_i \circ \mathcal{F}_i(\mathbf{d}^{j-1}), \quad (19b)$$

Algorithm 1 The multi-level IQN-ILS (ML-IQN-ILS) algorithm.

```

1:  $k = \ell = 0$ 
2:  $\mathbf{d}^k = \frac{5}{2}\mathbf{d}^n - 2\mathbf{d}^{n-1} + \frac{1}{2}\mathbf{d}^{n-2}$ 
3:  $\tilde{\mathbf{d}}^k = \mathcal{S}_1 \circ \mathcal{F}_1(\mathbf{d}^k)$ 
4:  $\mathbf{r}^k = \tilde{\mathbf{d}}^k - \mathbf{d}^k$ 
5: for  $i = 1$  to  $g$  do
6:   while  $\|\mathbf{r}^k\|_2 > \varepsilon_i$  or  $\ell = 0$  do
7:      $\ell = 1$ 
8:     if  $k = 0$  then
9:        $\mathbf{d}^{k+1} = \mathbf{d}^k + \omega\mathbf{r}^k$ 
10:    else
11:      construct  $\mathbf{V}^k$  and  $\mathbf{W}^k$ 
12:      calculate QR-decomposition  $\mathbf{V}^k = \mathbf{Q}^k\mathbf{R}^k$ 
13:      solve  $\mathbf{R}^k\mathbf{c}^k = -\mathbf{Q}^{k\top}\mathbf{r}^k$ 
14:       $\mathbf{d}^{k+1} = \mathbf{d}^k + \mathbf{W}^k\mathbf{c}^k + \mathbf{r}^k$ 
15:    end if
16:     $\tilde{\mathbf{d}}^{k+1} = \mathcal{S}_i \circ \mathcal{F}_i(\mathbf{d}^{k+1})$ 
17:     $\mathbf{r}^{k+1} = \tilde{\mathbf{d}}^{k+1} - \mathbf{d}^{k+1}$ 
18:     $k = k + 1$ 
19:  end while
20:   $\ell = 0$ 
21: end for
22: for  $i = 1$  to  $g - 1$  do
23:   synchronize  $\mathcal{F}_i$  and  $\mathcal{S}_i$  with  $\mathcal{F}_g$  and  $\mathcal{S}_g$ 
24: end for

```

should not be added to \mathbf{V}^k and \mathbf{W}^k . Otherwise, the approximation for the inverse of \mathcal{R}' would not only relate a change of the residual to a change of the interface's displacement, but would also represent the additional features that become visible due to a change of the grid level. If these differences are added to \mathbf{V}^k and \mathbf{W}^k nonetheless, the convergence of the coupling iterations on grid level $i + 1$ is hampered in the numerical experiments. When the differences in Eqs. (19) are not used, the number of columns in \mathbf{V}^k and \mathbf{W}^k at the end of the time step is less than or equal to the number of coupling iterations minus the number of grid levels.

Lines 22 to 24 show that synchronization is necessary at the end of the time step. Once the solution has been found on the finest grid level, all degrees of freedom on the coarser grid levels have to be corrected. A possible approach to the synchronization is to interpolate the data in the entire fluid and solid domain from the finest grid level to all other grid levels. If no such mechanism is available because the solvers are black boxes, the interface displacement and stress calculated during the coupling iterations on the finest grid level can be applied to the interface of the coarser grid levels and the flow equations

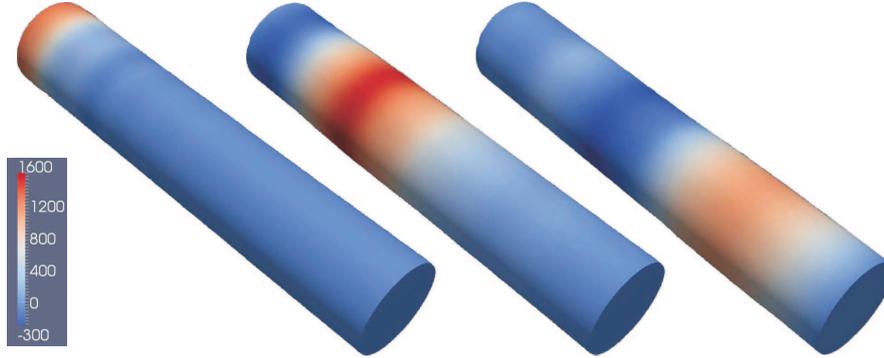


Figure 2: The pressure contours (in Pa) on the fluid-structure interface of the finest grid level for the propagation of a pressure wave in a 3D tube after 10^{-3} s (left), 5×10^{-3} s (centre) and 9×10^{-3} s (right).

and structural equations can be solved once more on all but the finest grid level, with this displacement and stress as boundary condition.

4 NUMERICAL RESULTS

As an example, the propagation of a pressure wave in a straight flexible tube is simulated [13]. This tube with radius 0.005 m and length 0.05 m is a simplified model for a large artery. The finite volume flow solver uses linear interpolation for the pressure and first-order upwind discretization for the momentum. It solves the Navier-Stokes equations in arbitrary Lagrangian-Eulerian (ALE) formulation with PISO pressure-velocity coupling and the first-order backward Euler time integration scheme. The grid of the fluid domain is adapted to the displacement of the fluid-structure interface by replacing the cell edges with springs. The finite element structural solver uses implicit Hilber-Hughes-Taylor time integration of shell elements with 8 nodes and takes into account the geometric nonlinearities due to the large deformation of the structure.

The tube's wall is a linear elastic material with density 1200 kg/m^3 , Young's modulus $3 \times 10^5 \text{ N/m}^2$, Poisson's ratio 0.3 and thickness 0.001 m. The structure is clamped in all directions at the inlet and outlet. The fluid is incompressible and has a density of 1000 kg/m^3 and a viscosity of $0.003 \text{ Pa}\cdot\text{s}$. Both the fluid and the structure are initially at rest. During the first 3×10^{-3} s, an overpressure of 1333.2 N/m^2 is applied at the inlet. The wave propagates through the tube during 10^{-2} s, simulated with time steps of 10^{-4} s. The pressure contours on the fluid-structure interface as shown in Figure 2 correspond well with those in [13].

For this simulation, two grid levels are used and the convergence tolerance is $\varepsilon_i = 10^{-3} \|\mathbf{r}^0\|_2$ for both grid levels. The coarse grid level contains 34944+1824 degrees of freedom for the flow and the structure, respectively. For the fine grid level, each direction is refined with a factor 4, giving 2247168+28032 degrees of freedom. IQN-ILS required

on average 13.2 coupling iterations per time step on the fine grid, whereas ML-IQN-ILS performed on average 12.1 coupling iterations on the coarse grid and 7.0 on the fine grid. The number of coupling iterations on the fine grid is thus reduced by approximately 50 % in the simulation with two grid levels, compared to a simulation with a fine grid only. As the cost of the coupling iterations on the coarse grid level is relatively small, the duration of the simulation also decreases by approximately 50 %.

5 CONCLUSIONS

A new multi-level coupling technique for partitioned simulation of fluid-structure interaction has been presented. This technique is based on the fundamental insight from stability analyses on Gauss-Seidel coupling iterations that in the difference between the current and the correct interface displacement, the Fourier modes with a low wave number are most unstable. ML-IQN-ILS first calculates the coupled solution on the coarsest grid level and subsequently uses that solution as the starting point for the coupling iterations on the following, finer grid level. Moreover, the approximation for the inverse of the Jacobian constructed on the coarser grid levels accelerates the convergence of the coupling iterations on the finer grid levels. The numerical results show that this multi-level algorithm can reduce the duration of a partitioned fluid-structure interaction simulation.

ACKNOWLEDGMENT

J. Degroote gratefully acknowledges a post-doctoral fellowship of the Research Foundation - Flanders (FWO).

REFERENCES

- [1] D. Mok, W. Wall, E. Ramm, Accelerated iterative substructuring schemes for instationary fluid-structure interaction, in: K.-J. Bathe (Ed.), *Computational Fluid and Solid Mechanics*, Elsevier, 1325–1328, 2001.
- [2] C. Michler, E. van Brummelen, R. de Borst, An interface Newton-Krylov solver for fluid-structure interaction, *International Journal for Numerical Methods in Fluids* 47 (10-11) (2005) 1189–1195.
- [3] J. Vierendeels, L. Lanoye, J. Degroote, P. Verdonck, Implicit coupling of partitioned fluid-structure interaction problems with reduced order models, *Computers & Structures* 85 (11–14) (2007) 970–976.
- [4] J. Degroote, K.-J. Bathe, J. Vierendeels, Performance of a new partitioned procedure versus a monolithic procedure in fluid-structure interaction, *Computers & Structures* 87 (11–12) (2009) 793–801.
- [5] P. Causin, J.-F. Gerbeau, F. Nobile, Added-mass effect in the design of partitioned

- algorithms for fluid-structure problems, *Computer Methods in Applied Mechanics and Engineering* 194 (42–44) (2005) 4506–4527.
- [6] S. Badia, F. Nobile, C. Vergara, Fluid-structure partitioned procedures based on Robin transmission conditions, *Journal of Computational Physics* 227 (14) (2008) 7027–7051.
- [7] J. Degroote, P. Bruggeman, R. Haelterman, J. Vierendeels, Stability of a coupling technique for partitioned solvers in FSI applications, *Computers & Structures* 86 (23–24) (2008) 2224–2234.
- [8] J. Degroote, S. Annerel, J. Vierendeels, Stability analysis of Gauss-Seidel iterations in a partitioned simulation of fluid-structure interaction, *Computers & Structures* 88 (5–6) (2010) 263–271.
- [9] A. Brandt, Multilevel adaptive solutions to boundary-value problems, *Mathematics of Computation* 31 (138) (1977) 333–390.
- [10] E. van Brummelen, K. van der Zee, R. de Borst, Space/time multigrid for a fluid-structure-interaction problem, *Applied Numerical Mathematics* 58 (12) (2008) 1951–1971.
- [11] A. van Zuijlen, S. Bosscher, H. Bijl, Two level algorithms for partitioned fluid-structure interaction computations, *Computer Methods in Applied Mechanics and Engineering* 196 (8) (2007) 1458–1470.
- [12] H. Wendland, Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree, *Advances in Computational Mathematics* 4 (1) (1995) 389–396.
- [13] M. Fernandez, M. Moubachir, A Newton method using exact Jacobians for solving fluid-structure coupling, *Computers & Structures* 83 (2–3) (2005) 127–142.

AN ITERATION-FREE, PARTITIONED METHOD FOR SOLVING COUPLED PROBLEMS

JOACHIM RANG*

*Institute of Scientific Computing
TU Braunschweig
Hans-Sommer Str. 65, 38106 Braunschweig, Germany
e-mail: j.rang@tu-bs, web page: <http://www.wire.tu-bs.de>

Key words: Coupled problems, partitioned methods, linear-implicit time integration, Rosenbrock-Wanner methods

Abstract. Coupled problems consist of two or more problems which in most cases describe different physical phenomena. An example of such a problem is the interaction of fluid and structure. Usually, the most accurate way to solve coupled problems is the monolithic approach. But often, due to different reasons, a partitioned method is used, where the subproblems are solved with different software packages and there may be different discretisation methods. One reason for partitioning a coupled problem is that existing codes and the best discretisation schemes can be used. In this note we introduce an iteration-free, partitioned method which is based on a linear-implicit time integration method.

1 INTRODUCTION

Coupled problems appear in different research areas. One common example is the interaction of structure and fluid [DR08], e.g. the numerical simulation of offshore wind turbines, see [MM04], or of biomechanical processes. Coupled problems consist of two or more different physical problems which are in general space and time dependent. The discretisation in space leads to a high dimensional system of ordinary differential equations (ODEs). The computation of the numerical solution needs the simultaneous solution of the strong coupled equations of each problem. But often for each subproblem different discretisation schemes are used. In the case of fluid-structure interaction the fluid is discretised with Finite Volumes and the structure with Finite Elements. For building a monolithic solver [RB00], it is often difficult to find a free available software system which processes different discretisation methods for different problem classes.

This is one reason to use a modular approach and partitioned methods [RB00, FP80, MW01, PFL95, MS02, MNS06], i.e. the subproblems are solved by different codes which

communicate with each other. The communication between the solvers can be realised with the help of the Component Template Library (CTL), i.e. the solvers are transformed into software components and are controlled from outside with a central unit. In [RSM09] the CTL is used to solve FSI problems.

For the time discretisation of parabolic differential equations, the heat equation or the incompressible Navier–Stokes equations, often implicit methods are used [GS00, JR10] to obtain a stable numerical solution. To use an implicit time integration method for a coupled problem leads to difficulties in solving the final non-linear system since each solver processes only a part of the system. Therefore iterative methods as the staggered scheme, the Block-Gauß-Seidel- or the Block–Newton scheme are used to solve this non-linear system.

In this paper we introduce two classes of time stepping schemes. First the diagonally implicit Runge–Kutta methods (DIRK–methods). We formulate the Block-Gauß-Seidel- and the Block–Newton scheme. The second class are the linear implicit Runge–Kutta methods, the so-called Rosenbrock–Wanner methods. This class of methods needs only the solution of a linear system. It is possible to formulate the Block–Gauß–Seidel method in such a way that we get an iteration-free partitioned method, and the Block–Newton method reduces to a Block–Gauß method, i.e. only one iteration step is needed.

The paper is structured as follows: First we give a short introduction into the time discretisation schemes. Then the Block–Gauß–Seidel and the Block–Newton methods for both discretisation schemes are formulated. In chapter 4 we present a numerical result.

2 TIME DISCRETISATION

In this note we are considering strongly coupled problems of ODEs which are given by

$$M_1 \dot{\mathbf{u}} = \mathbf{f}(t, \mathbf{u}, \mathbf{v}), \quad \mathbf{u}(0) = \mathbf{u}_0 \tag{1}$$

$$M_2 \dot{\mathbf{v}} = \mathbf{g}(t, \mathbf{u}, \mathbf{v}), \quad \mathbf{v}(0) = \mathbf{v}_0, \tag{2}$$

where the matrices $M_1 \in \mathbb{R}^{n_1, n_1}$ and $M_2 \in \mathbb{R}^{n_2, n_2}$ are regular. Problems which can be formulated in the form (1)–(2) arise in the simulation of mechanical problems, in the case of semi-discretised Dirichlet-Neumann-problems and in the simulation of FSI problems (see [MS03]).

In practical applications the systems (1) and (2) are stiff, i. e. explicit time discretisations schemes need arbitrarily small time steps to compute a stable numerical solution. Therefore we consider linear-implicit and diagonally implicit Runge–Kutta methods which have no step length restriction to produce a stable numerical solution.

2.1 Diagonally implicit Runge–Kutta methods

Application to ODEs. First we consider an implicate ODE of the form

$$M \dot{\mathbf{u}} = \mathbf{F}(t, \mathbf{u}), \quad \mathbf{u}(0) = \mathbf{u}_0, \tag{3}$$

where M is a regular matrix. A Runge–Kutta method (RK–method) with s internal stages, [HW96, SW92], is a one–step–method for solving (3) of the form

$$M\mathbf{k}_i = \mathbf{F}(t_m + c_i\tau_m, \mathbf{U}_i), \quad \mathbf{U}_i = \mathbf{u}_m + \tau_m \sum_{j=1}^s a_{ij}\mathbf{k}_j, \quad i = 1, \dots, s, \quad (4)$$

$$\mathbf{u}_{m+1} = \mathbf{u}_m + \tau_m \sum_{i=1}^s b_i\mathbf{k}_i. \quad (5)$$

The coefficients of an RK–method are usually represented with the help of a Butcher–table,

$$\begin{array}{c|ccc} c_1 & a_{11} & \dots & a_{1s} \\ c_2 & a_{21} & \dots & a_{2s} \\ \vdots & \vdots & & \vdots \\ c_s & a_{s1} & \dots & a_{ss} \\ \hline & b_1 & \dots & b_s \end{array} = \frac{\mathbf{c} \mid A}{\mid \mathbf{b}^\top}.$$

The vector \mathbf{c} includes the grid points of the time discretisation and \mathbf{b} is a vector with weights. The coefficients a_{ij} , b_i and c_i should be chosen in such a way that some order conditions are satisfied to obtain a sufficient consistency order.

In this paper, the coefficients of the RK–method (4)–(5) satisfy $a_{ij} = 0$ for $i < j$, $i, j \in \{1, \dots, s\}$ and $a_{ii} \neq 0$ for $i \in \{2, \dots, s\}$. RK–methods satisfying these conditions are called diagonal–implicit RK–methods (DIRK–methods). These methods are discussed in several papers and books, e.g. in [SW92, HW96]. Applications to fluid problems can be found in [JGR06, JR10] and to structural problems in [HH10].

Application to strongly coupled problems. Next we apply the RK–method (4)–(5) on our strongly coupled problem (1)–(2). Then the method reads as

$$M_1\mathbf{k}_i = \mathbf{f}(t_m + c_i\tau_m, \mathbf{U}_i, \mathbf{V}_i), \quad \mathbf{U}_i = \mathbf{u}_m + \tau_m \sum_{j=1}^s a_{ij}\mathbf{k}_j, \quad i = 1, \dots, s, \quad (6)$$

$$M_2\mathbf{l}_i = \mathbf{g}(t_m + c_i\tau_m, \mathbf{U}_i, \mathbf{V}_i), \quad \mathbf{V}_i = \mathbf{v}_m + \tau_m \sum_{j=1}^s a_{ij}\mathbf{l}_j, \quad i = 1, \dots, s, \quad (7)$$

$$\mathbf{u}_{m+1} = \mathbf{u}_m + \tau_m \sum_{i=1}^s b_i\mathbf{k}_i, \quad \mathbf{v}_{m+1} = \mathbf{v}_m + \tau_m \sum_{i=1}^s b_i\mathbf{l}_i. \quad (8)$$

In each timestep s non-linear systems have to be solved. One possibility for the solution of these systems is the simplified Newton method [SW92, HW96] which reads for the equation $\mathbf{F}(\mathbf{x}) = 0$ as follows

$$\mathbf{x}^{(\nu+1)} = \mathbf{x}^{(\nu)} - (\partial_{\mathbf{x}}\mathbf{F}(\mathbf{x}^{(0)}))^{-1}\mathbf{F}(\mathbf{x}^{(\nu)}). \quad (9)$$

Since inverting a matrix is very expensive, we multiply by $\partial_{\mathbf{x}}\mathbf{F}(\mathbf{x}^{(0)})$ and get the linear system

$$(\partial_{\mathbf{x}}\mathbf{F}(\mathbf{x}^{(0)}))(\mathbf{x}^{(\nu+1)} - \mathbf{x}^{(\nu)}) = -\mathbf{F}(\mathbf{x}^{(\nu)}).$$

In the case $\gamma := a_{ii}$ for $i = 1, \dots, s$, the system-matrix on the left-hand side in (6)–(7) does not change during a timestep. Therefore we can make one LU-decomposition in each time-step and then solve all non-linear systems by forward and backward substitutions which reduce the cost for the linear algebra. In our case the linear systems reads as

$$\begin{pmatrix} M_1 - \tau a_{ii} \partial_{\mathbf{u}} \mathbf{f}_m & -\tau a_{ii} \partial_{\mathbf{v}} \mathbf{f}_m \\ -\tau a_{ii} \partial_{\mathbf{u}} \mathbf{g}_m & M_2 - \tau a_{ii} \partial_{\mathbf{v}} \mathbf{g}_m \end{pmatrix} \left[\begin{pmatrix} \mathbf{k}_i^{(\nu+1)} \\ \mathbf{l}_i^{(\nu+1)} \end{pmatrix} - \begin{pmatrix} \mathbf{k}_i^{(\nu)} \\ \mathbf{l}_i^{(\nu)} \end{pmatrix} \right] \\ = \begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} \begin{pmatrix} \mathbf{k}_i^{(\nu)} \\ \mathbf{l}_i^{(\nu)} \end{pmatrix} - \begin{pmatrix} \mathbf{f} \left(t_m + c_i \tau_m, \mathbf{U}_i^{(\nu)}, \mathbf{V}_i^{(\nu)} \right) \\ \mathbf{g} \left(t_m + c_i \tau_m, \mathbf{U}_i^{(\nu)}, \mathbf{V}_i^{(\nu)} \right) \end{pmatrix}, \quad (10)$$

where $\nu > 0$, $\partial_{\mathbf{u}} \mathbf{f}_m := \partial_{\mathbf{u}} \mathbf{f}(t_m, \mathbf{u}, \mathbf{v})$, and

$$\mathbf{U}_i^{(\nu)} := \mathbf{u}_m + \tau_m \sum_{j=1}^{i-1} a_{ij} \mathbf{k}_j + \tau_m a_{ii} \mathbf{k}_i^{(\nu)}, \quad i = 1, \dots, s$$

Adaptive time step control. RK-methods have the advantage that they allow an easy implementation of an adaptive time steplength control. Consider a RK-method of order $p \geq 2$. An adaptive time step control employs a second RK-method which has the coefficients a_{ij} , \hat{b}_i and c_i , $i, j = 1, \dots, s$, and order $p - 1$. The solution of the second method at t_{m+1} is given by

$$\hat{\mathbf{u}}_{m+1} = \mathbf{u}_m + \sum_{i=1}^s \hat{b}_i \mathbf{k}_i.$$

Now, the next time step τ_{m+1} is proposed to be

$$\tau_{m+1} = \rho \frac{\tau_m^2}{\tau_{m-1}} \left(\frac{TOL \cdot r_m}{r_{m+1}^2} \right)^{1/p}, \quad (11)$$

where $\rho \in (0, 1]$ is a safety factor, $TOL > 0$ is a given tolerance and

$$r_{m+1} := \|\mathbf{u}_{m+1} - \hat{\mathbf{u}}_{m+1}\|. \quad (12)$$

This step size selection rule is called PI-controller [GLS88], and details on the numerical error and the implementation of the automatic steplength control can be found in [HW96, Lan01].

2.2 Rosenbrock–Wanner methods

Application to ODEs. As in the case of DIRK schemes we start our considerations with an implicit ODE of the form (3). A Rosenbrock–Wanner–method (ROW method) with s internal stages is given by

$$M\mathbf{k}_i = \mathbf{F}\left(t_m + \alpha_i\tau_m, \tilde{\mathbf{U}}_i\right) + \tau_m J \sum_{j=1}^i \gamma_{ij}\mathbf{k}_j + \tau_m \gamma_i \dot{\mathbf{F}}(t_m, \mathbf{u}_m), \quad (13)$$

$$\begin{aligned} \tilde{\mathbf{U}}_i &= \mathbf{u}_m + \tau_m \sum_{j=1}^{i-1} a_{ij}\mathbf{k}_j, \quad i = 1, \dots, s, \\ \mathbf{u}_{m+1} &= \mathbf{u}_m + \tau_m \sum_{i=1}^s b_i\mathbf{k}_i, \end{aligned} \quad (14)$$

where $J := \partial_{\mathbf{u}}\mathbf{F}(t_m, \mathbf{u}_m)$, α_{ij} , γ_{ij} , b_i are the parameters of the method,

$$\alpha_i := \sum_{j=1}^{i-1} \alpha_{ij}, \quad \gamma_i := \sum_{j=1}^{i-1} \gamma_{ij}, \quad \gamma := \gamma_{ii} > 0, \quad i = 1, \dots, s.$$

If the parameters α_{ij} , γ_{ij} , and b_i are chosen appropriately, a sufficient consistency order can be obtained. Additional consistency conditions arise if J is only an approximation to $\partial_{\mathbf{u}}\mathbf{F}(t_m, \mathbf{u}_m)$, or if J is an arbitrary matrix. This class of methods are called W–methods, [SW92]. If a ROW method is applied to semidiscretized partial differential equation, further order conditions should be satisfied to avoid order reduction, see [LO95]. The same stability concepts apply for ROW methods as for DIRK–methods.

The ROW method (13)–(14) requires the successive solution of s linear systems of equations with the same matrix $M - \gamma\tau_m J$. Note, J depends only on \mathbf{u}_m . The right hand side of the i –th linear system of equations depends on the solutions of the first to the $(i - 1)$ –st system. Thus, a main difference of ROW methods to DIRK methods is that it is not necessary to solve a nonlinear system of equations in each discrete time but a fixed number of linear systems of equations, i.e. there appears no iteration loop for solving the nonlinear system and so the method can be interpreted as iteration-free.

Again, as in the last section about diagonal–implicit RK–methods, an automatic step length control can be implemented with the help of an embedded method. Common Rosenbrock methods as ROS3P, ROS3Pw have an embedded method.

Application to strongly coupled systems. An ROW method applied on the strongly

coupled system (1)–(2) reads as

$$\begin{aligned}
 & \begin{pmatrix} M_1 - \tau\gamma\partial_{\mathbf{u}}\mathbf{f} & -\tau\gamma\partial_{\mathbf{v}}\mathbf{f} \\ -\tau\gamma\partial_{\mathbf{u}}\mathbf{g} & M_2 - \tau\gamma\partial_{\mathbf{v}}\mathbf{g} \end{pmatrix} \begin{pmatrix} \mathbf{U}_i \\ \mathbf{V}_i \end{pmatrix} \\
 &= \begin{pmatrix} \mathbf{f} \left(t_m + \alpha_i\tau_m, \hat{\mathbf{U}}_i, \hat{\mathbf{V}}_i \right) \\ \mathbf{g} \left(t_m + \alpha_i\tau_m, \hat{\mathbf{U}}_i, \hat{\mathbf{V}}_i \right) \end{pmatrix} + \tau_m \sum_{j=1}^{i-1} \gamma_{ij} \begin{pmatrix} \partial_{\mathbf{u}}\mathbf{f} & \partial_{\mathbf{v}}\mathbf{f} \\ \partial_{\mathbf{u}}\mathbf{g} & \partial_{\mathbf{v}}\mathbf{g} \end{pmatrix} \begin{pmatrix} \mathbf{U}_i \\ \mathbf{V}_i \end{pmatrix} \\
 &+ \tau_m \gamma_i \begin{pmatrix} \dot{\mathbf{f}} \left(t_m + \alpha_i\tau_m, \hat{\mathbf{U}}_i, \hat{\mathbf{V}}_i \right) \\ \dot{\mathbf{g}} \left(t_m + \alpha_i\tau_m, \hat{\mathbf{U}}_i, \hat{\mathbf{V}}_i \right) \end{pmatrix}, \tag{15}
 \end{aligned}$$

$$\hat{\mathbf{U}}_i = \mathbf{u}_m + \tau_m \sum_{j=1}^{i-1} a_{ij} \mathbf{U}_j, \quad \hat{\mathbf{V}}_i = \mathbf{v}_m + \tau_m \sum_{j=1}^{i-1} a_{ij} \mathbf{V}_j, \quad i = 1, \dots, s, \tag{16}$$

$$\mathbf{u}_{m+1} = \mathbf{u}_m + \sum_{i=1}^s b_i \mathbf{U}_i, \quad \mathbf{v}_{m+1} = \mathbf{v}_m + \sum_{i=1}^s b_i \mathbf{V}_i. \tag{17}$$

3 THE PARTITIONED APPROACH

Next we consider the partitioned approach, i.e. the ODEs (1)–(2) are solved with different codes and methods. Considering the DIRK– and ROW–methods the systems (10) and (15)–(16) are solved on different computers. Therefore we start our considerations with the strongly coupled, non-linear system

$$0 = \mathbf{f}(\mathbf{x}, \mathbf{y}), \quad \mathbf{x} \in \mathbb{R}^m \tag{18}$$

$$0 = \mathbf{g}(\mathbf{x}, \mathbf{y}), \quad \mathbf{y} \in \mathbb{R}^n, \tag{19}$$

where $\mathbf{f} : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $\mathbf{g} : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ are sufficiently smooth functions. We discuss two partitioned approaches to solve the problem (18)–(19). First the Block–Gauß–Seidel method and second the Block–Newton method.

3.1 The Block–Gauß–Seidel method

The idea of the Block–Gauß–Seidel method is the following. First we solve equation (18) with fixed \mathbf{y} w.r.t. \mathbf{x} . Then we insert this solution \mathbf{x} into equation (19) and solve it w.r.t. \mathbf{y} . This procedure is repeated until convergence and a graphical illustration of the method can be found in Figure 1. For the convergence of the Block–Gauß–Seidel method we refer to [Axe96].

Application to DIRK–methods In this section we apply the Block–Gauß–Seidel method on our coupled system (6)–(7). As explained before we use a simplified Newton iteration to solve the two non-linear systems. Our partitioned method reads then as follows:

1. Set $\nu := 0$, $\mathbf{k}_i^{(\nu)} := 0$ and $\mathbf{l}_i^{(\nu)} := 0$.

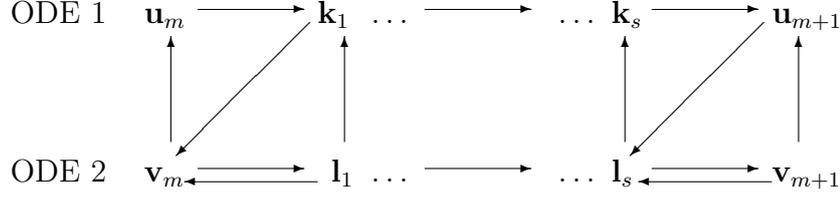


Figure 1: Block-Gauß-Seidel method

2. Compute

$$\mathbf{V}_i^{(\nu)} := \mathbf{v}_m + \tau_m \sum_{j=1}^{i-1} a_{ij} \mathbf{l}_j + \tau_m a_{ii} \mathbf{l}_i^{(\nu)}$$

and communicate it to the first solver.

3. Compute $\mathbf{k}_i^{(\nu+1)}$ by solving (6), set

$$\mathbf{U}_i^{(\nu+1)} := \mathbf{v}_m + \tau_m \sum_{j=1}^{i-1} a_{ij} \mathbf{k}_j + \tau_m a_{ii} \mathbf{k}_i^{(\nu+1)},$$

and communicate $\mathbf{U}_i^{(\nu+1)}$ to the second solver.

4. Compute $\mathbf{l}_i^{(\nu+1)}$ by solving (7).

5. Set $\nu := \nu + 1$.

6. If the values $\mathbf{k}_i^{(\nu+1)}$ and $\mathbf{l}_i^{(\nu+1)}$ are not sufficiently accurate then go to Step 2

Application to ROW-methods In this section we apply the Block-Gauß-Seidel method on our coupled system (15). In this case to a simplified, iteration free Block-Gauß method since the diagonal blocks of the Jacobian, i.e. J_{12} and J_{21} , are set two zero. To get a better approximation we can manipulate the time derivative of the right-hand side, but as we will see later in the section on the numerical example our coupled Rosenbrock methods have order reduction.

Our partitioned method then reads as follows. Compute \mathbf{U}_i by solving

$$(M_1 - \gamma \tau_m \partial_{\mathbf{u}} \mathbf{f}) \mathbf{U}_i = \mathbf{f} \left(t_m + \alpha_i \tau_m, \hat{\mathbf{U}}_i, \hat{\mathbf{V}}_i \right) + \tau_m \sum_{j=1}^{i-1} \gamma_{ij} \partial_{\mathbf{u}} \mathbf{f} \mathbf{U}_j + \gamma_i \tau_m \left(\dot{\mathbf{f}}(t_m, \mathbf{u}_m) + \partial_{\mathbf{v}} \mathbf{f} \mathbf{g}_m \right)$$

and

$$(M_2 - \gamma \tau_m \partial_{\mathbf{v}} \mathbf{g}) \mathbf{V}_i = \mathbf{g} \left(t_m + \alpha_i \tau_m, \hat{\mathbf{U}}_i, \hat{\mathbf{V}}_i \right) + \tau_m \sum_{j=1}^{i-1} \gamma_{ij} \partial_{\mathbf{v}} \mathbf{g} \mathbf{V}_j + \gamma_i \tau_m \left(\dot{\mathbf{g}}(t_m, \mathbf{u}_m) + \partial_{\mathbf{u}} \mathbf{g} \mathbf{f}_m \right)$$

with $\mathbf{f}_m := \mathbf{f}(t_m, \mathbf{u}_m, \mathbf{v}_m)$, $\mathbf{g}_m := \mathbf{g}(t_m, \mathbf{u}_m, \mathbf{v}_m)$, $\hat{\mathbf{U}}_i$ and $\hat{\mathbf{V}}_i$, $i = 1, \dots, s$, given by (16). The terms $\partial_{\mathbf{v}} \mathbf{f} \mathbf{g}_m$ and $\partial_{\mathbf{u}} \mathbf{g} \mathbf{f}_m$ result from the time derivative of \mathbf{f} and \mathbf{g} .

3.2 The Block–Newton method

First we apply Newton’s method on our nonlinear system (18) and (19). We obtain the linear system

$$\begin{pmatrix} \mathbf{f}_x & \mathbf{f}_y \\ \mathbf{g}_x & \mathbf{g}_y \end{pmatrix} \begin{pmatrix} \Delta \mathbf{x}^{(k+1)} \\ \Delta \mathbf{y}^{(k+1)} \end{pmatrix} = - \begin{pmatrix} \mathbf{f}(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}) \\ \mathbf{g}(\mathbf{x}^{(k)}, \mathbf{y}^{(k)}) \end{pmatrix}, \quad (20)$$

where \mathbf{f}_x, \dots are the Jacobians and $\Delta \mathbf{x}^{(k+1)} := \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$. In the next step we apply one Gauß-step on the system (20), i.e. we resolve the first equation of (20) w.r.t. $\Delta \mathbf{x}^{(k+1)}$, i.e.

$$\Delta \mathbf{x}^{(k+1)} = -\mathbf{f}_x^{-1}(\mathbf{f}_y \Delta \mathbf{y}^{(k+1)} + \mathbf{f})$$

and insert this result into the second equation, i.e.

$$-\mathbf{g}_x \mathbf{f}_x^{-1}(\mathbf{f}_y \Delta \mathbf{y}^{(k+1)} + \mathbf{f}) + \mathbf{g}_y \Delta \mathbf{y}^{(k+1)} = -\mathbf{g}$$

or

$$(\mathbf{g}_y - \mathbf{g}_x \mathbf{f}_x^{-1} \mathbf{f}_y) \Delta \mathbf{y}^{(k+1)} = \mathbf{g}_x \mathbf{f}_x^{-1} \mathbf{f} - \mathbf{g}.$$

For abbreviation we set

$$S := \mathbf{g}_y - \mathbf{g}_x \underbrace{\mathbf{f}_x^{-1} \mathbf{f}_y}_{=: C}, \quad \mathbf{p} := \mathbf{f}_x^{-1} \mathbf{f}.$$

The matrix S is often called Schur complement. Now our Block Elimination Algorithm has the following form

1. Compute \mathbf{p} , i.e. solve $\mathbf{f}_x \mathbf{p} = \mathbf{f}$ for \mathbf{p} .
2. Compute $C = \mathbf{f}_x^{-1} \mathbf{f}_y$, i.e. solve the matrix equation $\mathbf{f}_x C = \mathbf{f}_y$ for C .
3. Compute the Schur complement $S = \mathbf{g}_y - \mathbf{g}_x C$.
4. Compute the modified right-hand side $\mathbf{g}_x \mathbf{p} - \mathbf{g} =: \tilde{\mathbf{g}}$.
5. Solve $S \Delta \mathbf{y} = \tilde{\mathbf{g}}$ for $\Delta \mathbf{y}$.
6. Compute $\Delta \mathbf{x} = -(\mathbf{p} + C \Delta \mathbf{y})$.

Application to DIRK–methods In this paragraph we apply our Block–Newton method on the non-linear equation (6)–(7). Therefore, as in the previous section, we first apply a Newton method on (6)–(7) leading to the linear system (10) which is then solved by a Block–Gauß algorithm. Our Block–Newton method reads then as follows

1. Compute \mathbf{p} , i.e. solve $(M_1 - \tau a_{ii} \partial_{\mathbf{u}} \mathbf{f}_m) \mathbf{p} = \mathbf{f} \left(t_m + c_i \tau_m, \mathbf{U}_i^{(\nu)}, \mathbf{V}_i^{(\nu)} \right) - M_1 \mathbf{k}_i^{(\nu)}$.

2. Compute $C = -\tau a_{ii} (M_1 - \tau a_{ii} \partial_{\mathbf{u}} \mathbf{f}_m)^{-1} \partial_{\mathbf{v}} \mathbf{f}_m$, i.e. solve the matrix equation

$$(M_1 - \tau a_{ii} \partial_{\mathbf{u}} \mathbf{f}_m) C = -\tau a_{ii} \mathbf{f}_m \partial_{\mathbf{v}} \mathbf{f}_m$$

for C .

3. Compute the Schur complement $S = (M_2 - \tau a_{ii} \partial_{\mathbf{v}} \mathbf{g}_m) + \tau a_{ii} \partial_{\mathbf{u}} \mathbf{g}_m C$.

4. Compute the modified right-hand side $\tau a_{ii} \partial_{\mathbf{u}} \mathbf{g}_m \mathbf{p} - M_2 \mathbf{l}_i^{(\nu)} + \mathbf{g} =: \tilde{\mathbf{g}}$.

5. Solve $S \Delta \mathbf{y} = \tilde{\mathbf{g}}$ for $\Delta \mathbf{y}$.

6. Compute $\Delta \mathbf{x} = \mathbf{p} - C \Delta \mathbf{y}$.

Application to ROW-methods In the case of ROW-methods the Block-Newton method simplifies to a Block-Gauß method applied on the linear system (15). The method reads as

1. Compute \mathbf{p} , i.e. solve $(M_1 - \tau a_{ii} \partial_{\mathbf{u}} \mathbf{f}_m) \mathbf{p} = \mathbf{f} \left(t_m + c_i \tau_m, \mathbf{U}_i^{(\nu)}, \mathbf{V}_i^{(\nu)} \right)$.

2. Compute $C = \tau a_{ii} (M_1 - \tau a_{ii} \partial_{\mathbf{u}} \mathbf{f}_m)^{-1} \partial_{\mathbf{u}}$, i.e. solve the matrix equation

$$(M_1 - \tau a_{ii} \partial_{\mathbf{u}} \mathbf{f}_m) C = \tau a_{ii} \mathbf{f}_m^{-1} \partial_{\mathbf{u}}$$

for C .

3. Compute the Schur complement $S = (M_2 - \tau a_{ii} \partial_{\mathbf{v}} \mathbf{g}_m) - \tau a_{ii} \partial_{\mathbf{u}} \mathbf{g}_m C$.

4. Compute the modified right-hand side $\tau a_{ii} \partial_{\mathbf{u}} \mathbf{g}_m \mathbf{p} - \mathbf{g} =: \tilde{\mathbf{g}}$.

5. Solve $S \Delta \mathbf{y} = \tilde{\mathbf{g}}$ for $\Delta \mathbf{y}$.

6. Compute $\Delta \mathbf{x} = -(\mathbf{p} + C \Delta \mathbf{y})$.

4 NUMERICAL EXAMPLES

As an example we consider a simple predator-prey model which is given by

$$\begin{aligned} \dot{u} &= 10u(1 - v), & u(0) &= 3 \\ \dot{v} &= v(u - 1), & v(0) &= 1. \end{aligned}$$

For determining the numerical solution we apply the trapezoidal rule, DIRK3 [Ran07], ROS3P [RA05], and ROS3Pw [RA05] with equidistant time steps $\tau = 1/(10 \cdot 2^N)$, $N = 0, 1, \dots, 5$.

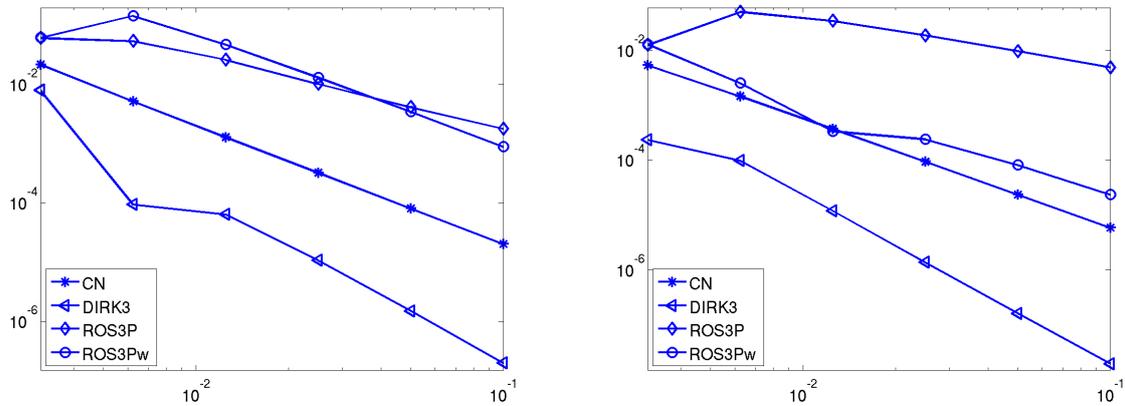


Figure 2: Block–Gauß–Seidel method: τ versus error

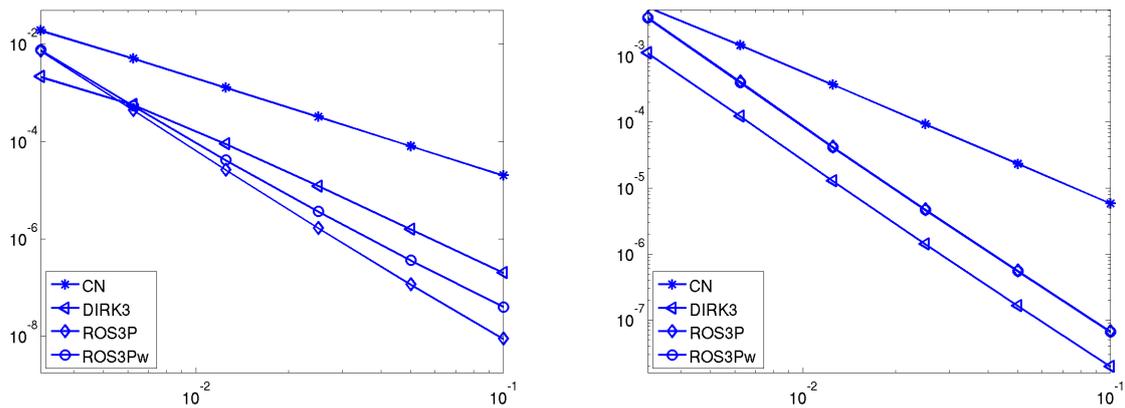


Figure 3: Block–Newton method: τ versus error

In Figure 2 we present the results which are obtained with the Block–Gauß–Seidel method. It can be observed that the ROW–methods have order reduction since the Jacobian is not evaluated exactly. The DIRK–methods give better results since these methods need not the evaluation of a Jacobian.

Different results appear if the Block–Newton method is used to solve the final system (see Figure 3). In this case all methods reach the desired order and the ROW–methods give much better results as then with the Block–Gauß–Seidel method.

5 SUMMARY AND OUTLOOK

In this note we have introduced the Block–Gauß–Seidel and the Block–Newton method for DIRK– and ROW–methods. In the case of ROW–methods we get an iteration-free, partitioned method.

In the next steps we first have to consider more complicated ODEs, e.g. the coupling of parabolic differential equations. Moreover the coupling of differential algebraic equations should be analysed since it is well-known that the Block–Gauß–Seidel method may fail for this class of problems.

REFERENCES

- [Axe96] Owe Axelsson. *Iterative solution methods*. Cambridge Univ. Press., Cambridge, 1996.
- [DR08] D. Dinkler and J. Rang, editors. *Wechselwirkung von Struktur und Fluid - Abschlussbericht eines Graduiertenkollegs*. Wolfram Schmidt Buchbinderei & Druckerei, Braunschweig, 2008.
- [FP80] C.A. Felippa and K.C. Park. Staggered transient analysis procedures for coupled mechanical systems: Formulation. *Comput. Methods Appl. Mech. Eng.*, 24:61–111, 1980.
- [GLS88] K. Gustafsson, M. Lundh, and G. Söderlind. A PI stepsize control for the numerical solution of ordinary differential equations. *BIT*, 28(2):270–287, 1988.
- [GS00] P.M. Gresho and R.L. Sani. *Incompressible Flow and the Finite Element Method*. Wiley, Chichester, 2000.
- [HH10] S. Hartmann and A.-W. Hamkar. Rosenbrock-type methods applied to finite element computations within finite strain viscoelasticity. *Computer Methods in Applied Mechanics and Engineering*, 199(23-24):1455–1470, 2010.
- [HW96] E. Hairer and G. Wanner. *Solving ordinary differential equations. II: Stiff and differential-algebraic problems.*, volume 14 of *Springer Series in Computational Mathematics*. Springer, Berlin, 1996.
- [JGR06] V. John, Matthias G., and J. Rang. A comparison of time-discretization/linearization approaches for the incompressible Navier–Stokes equations. *Comput. Methods Appl. Mech. Eng.*, 195:5995–6010, 2006.
- [JR10] V. John and J. Rang. Adaptive time step control for the incompressible Navier–Stokes equations. *Comput. Methods Appl. Mech. Eng.*, 199:514–524, 2010.
- [Lan01] J. Lang. *Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems*, volume 16 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2001.
- [LO95] C. Lubich and A. Ostermann. Linearly implicit time discretization of non-linear parabolic equations. *IMA J. Numer. Anal.*, 15(4):555–583, 1995.

- [MM04] M. Meyer and H. G. Matthies. State-space representation of instationary two-dimensional airfoil aerodynamics. *Journal of Wind Engineering and Industrial Aerodynamics*, 92(3-4):263–274, 2004.
- [MNS06] Hermann G. Matthies, Rainer Niekamp, and Jan Steindorf. Algorithms for strong coupling procedures. *Comput. Methods Appl. Mech. Eng.*, 195(17-18):2028–2049, 2006.
- [MS02] H. G. Matthies and J. Steindorf. Partitioned but strongly coupled iteration schemes for nonlinear fluid-structure interaction. *Comput. Struct.*, 80:1991–1999, 2002.
- [MS03] H. G. Matthies and J. Steindorf. Partitioned strong coupling algorithms for fluid-structure interaction. *Comput. Struct.*, 81:805–812, 2003.
- [MW01] D. P. Mok and W. A. Wall. Partitioned analysis schemes for the transient interaction of incompressible flows and nonlinear flexible structures. In K. Schweizerhof W. A. Wall, K.-U. Bletzinger, editor, *Trends in Computational Structural Mechanics*, pages 689–698. CIMNE, Barcelona, 2001.
- [PFL95] S. Piperno, C. Farhat, and B. Larrouturou. Partitioned procedures for the transient solution of coupled aeroelastic problems. I: Model problem, theory and two-dimensional application. *Comput. Methods Appl. Mech. Eng.*, 124(1-2):79–112, 1995.
- [RA05] J. Rang and L. Angermann. New Rosenbrock methods for partial differential algebraic equations of index 1. *BIT*, 45(4):761–787, 2005.
- [Ran07] Joachim Rang. Design of DIRK schemes for solving the Navier-Stokes-equations. Informatik-Bericht 2007-02, TU Braunschweig, Braunschweig, 2007.
- [RB00] S. Rugonyi and K.-J. Bathe. On the analysis of fully-coupled fluid flows with structural interactions — a coupling and condensation procedure. *Int. J. of Comp. Civil and Struct. Eng.*, 1:29–41, 2000.
- [RSM09] Joachim Rang, Johannes Schön, and Hermann G. Matthies. Solving FSI problems with high resolution and using a component framework in parallel. In Stefan Hartmann, Andreas Meister, Michael Schäfer, and Stefan Turek, editors, *International Workshop on Fluid-Structure Interaction: Theory, Numerics and Applications*, Kassel, 2009. Kassel Univ. Press.
- [SW92] K Strehmel and R. Weiner. *Linear-implizite Runge-Kutta-Methoden und ihre Anwendung*, volume 127 of *Teubner-Texte zur Mathematik*. Teubner, Stuttgart, 1992.

FINITE ELEMENT/BOUNDARY ELEMENT COUPLING FOR AIRBAG DEPLOYMENT

T.M. van Opstal^{†,*} and E.H. van Brummelen[†]

[†]Multi-Scale Engineering Fluid Dynamics (MEFD),
Eindhoven, University of Technology,
P.O. Box 513, 5600MB, Netherlands.

*email: t.m.v.opstal@tue.nl, web page: <http://w3.tue.nl>

Key words: Fluid–structure interaction, boundary element method, strong coupling, dynamic

Abstract. Fluid–structure interaction of inflatables comprizes a family of applications, of which airbags are one. Challenges in this domain are complex geometries requiring relatively high resolution; large displacements possibly entailing severe mesh distortion; and strong coupling. A promising approach seems to be coupling a classical finite element formulation for the airbag fabric to a boundary element formulation for the enclosed fluid. Together with an appropriate time-integration scheme this method answers aforementioned challenges, as demonstrated by numerical simulation.

1 Introduction

In a small percentage of airbag deployments, out-of-position impact occurs, usually resulting in severe injuries. To understand and improve the inflation process, a precise understanding of the airbag dynamics is required. This can be provided by accurate numerical simulations. These simulations are a complicated endeavor however, mainly on account of the large displacements and length-scale disparities inherently involved. On the one hand, a realistic stowed airbag constitutes a labyrinth of intricate folds. On the other, the final configuration is a relatively simple bulb. To date, the complex behavior on the small scales has been overly simplified (e.g. [9]) rendering the results inappropriate for the analysis of out-of-position situations.

The approach proposed here is to decompose the fluid domain according to the above-mentioned length-scales. The flow inside the geometrically complex folded region is described by a simple, linear potential flow model. This enables analysis using the boundary element method (BEM), which offers significant advantages over domain-discretization approaches. Most prominently: the anticipated large displacements do not entail mesh

skewing problems; the solution is calculated exclusively at the coupling interface; and the structure mesh can be inherited by the fluid.

To assess the aforementioned approach, we consider a fluid-structure-interaction problem consisting of a potential-flow model coupled to a string model described by a nonlinear wave equation. The latter is discretized with the finite element method. The system requires the imposition of volume conservation on the structure subproblem to account for the incompressibility of the fluid, and setting of an arbitrary additive constant on the fluid solution, both by means of a Lagrange multiplier. Also, due to the typically low mass of the membrane and the incompressibility of the fluid, the added-mass effect [2, 3, 5] requires the use of an implicit time-integration scheme.

The remainder of this paper is organized as follows. In sec. 2, the mathematical problem is introduced. From the governing equations suitable variational forms are derived. Some properties pertaining to the numerical solution of the system are elucidated. In sec. 3, the discretization in both space and time of the variational forms is treated. Ensuing, sec. 4 presents numerical verification and demonstration of the proposed method. This leads to the conclusions, which are drawn in sec. 5.

2 Problem statement

The structure is assumed to behave according to a large-displacement string model. The fluid is assumed inviscid and irrotational. In this section variational forms for the structural and fluidic subsystems are given, followed by their coupling and a discussion on the aggregate system.

Membrane The airbag, henceforth designated the membrane, is modeled by a curve in the plane parametrized by $\mathbf{x} : (0, L) \times (0, T) \longrightarrow \mathbb{R}^2$. This membrane is defined by the time dependent set $\Gamma_w := \mathbf{x}((0, L), t)$. Together with the inflator opening Γ_i , it encloses the fluid domain Ω , which is consequently also time dependent. The boundary $\Gamma := \overline{\Gamma_w} \cup \Gamma_i$ is assumed to be $C^{1,1}$ -continuous almost everywhere. Time- and space derivatives are denoted $\partial_t(\cdot)$ and $\partial_s(\cdot)$, respectively. The motion of the membrane is governed by the momentum-balance equation [1, 13]:

$$\varrho_0 \partial_t^2 \mathbf{x} - \partial_s (E[1 - J^{-1}] \partial_s \mathbf{x}) - Jp\mathbf{n} = 0, \quad (1)$$

supplemented with initial conditions $\mathbf{x}(\cdot, 0) = \mathbf{x}_0(s)$, $\mathbf{x}'(\cdot, 0) = \mathbf{x}_1(s)$; and boundary conditions $\mathbf{x}(0, \cdot) = \mathbf{x}_0(0)$, $\mathbf{x}(L, \cdot) = \mathbf{x}_0(L)$. The second term in (1) is derived from $\Psi := -E(J - 1)^2/2$, the strain energy of a *linearly elastic* material. Furthermore, $J := \|\partial_s \mathbf{x}\|$ denotes the determinant of the Jacobian of \mathbf{x} . The normal \mathbf{n} can also be expressed as $\mathcal{R} \partial_s \mathbf{x} J^{-1}$ with \mathcal{R} the $\pi/2$ negative rotation. The parameters ϱ_0 and E represent the structural density and Young's modulus respectively. The variational form becomes *find* $\mathbf{x} - \mathbf{x}_0 \in H_0^1(0, L; \mathbb{R}^2)$ *s.t.* $\forall \mathbf{w} \in H_0^1(0, L; \mathbb{R}^2)$:

$$(\mathbf{w}, \varrho_0 \partial_t^2 \mathbf{x})_{L^2} + (\partial_s \mathbf{w}, E[1 - J^{-1}] \partial_s \mathbf{x})_{L^2} - (\mathbf{w}, p \mathcal{R} \partial_s \mathbf{x})_{L^2} = 0, \quad (2)$$

with $(\cdot, \cdot)_{L^2}$ the $L^2(0, L; \mathbb{R}^2)$ inner product.

Fluid The Laplace problem with pure Neumann boundaries governs the fluid response, described by ϕ ,

$$\begin{cases} \Delta\phi = 0 & \text{in } \Omega, \\ \partial_{\mathbf{n}}\phi = g & \text{at } \Gamma, \end{cases} \quad (3)$$

with g the normal velocity at the boundary. These equations can be cast into a boundary integral formulation, as derived in e.g. [8, 12]:

$$c(\mathbf{x})\phi(\mathbf{x}) + \oint^* H(\mathbf{x}; \mathbf{y})\phi(\mathbf{y})d\Gamma_{\mathbf{y}} = \oint^* G(\mathbf{x}; \mathbf{y})g(\mathbf{y})d\Gamma_{\mathbf{y}}, \quad (4)$$

where $\mathbf{x} \in \Gamma$ and

$$G(\mathbf{x}; \mathbf{y}) = -(2\pi)^{-1} \log r, \quad H(\mathbf{x}; \mathbf{y}) = \partial_{\mathbf{n}_{\mathbf{y}}}G(\mathbf{x}; \mathbf{y}), \quad (5)$$

are the single- and double-layer potentials respectively, with $r := \|\mathbf{x} - \mathbf{y}\|$ the distance. The coefficient $c(\mathbf{x}) = \alpha/2\pi$ with α the local angle of Γ . Recalling our regularity assumptions, $c = 1/2$ a.e. The integrals are singular and, accordingly, the asterisk indicates that these should be interpreted in the *Cauchy principal value* sense. Employing the Galerkin technique, the variational problem corresponding to (4) is: *find* $\phi \in H^{1/2}(\Gamma)$ *s.t.* $\forall w \in H^{1/2}(\Gamma)$:

$$(w, \phi/2)_X + (w, H * \phi)_X = (w, G * g)_X. \quad (6)$$

The appropriate inner product would be that of $X = H^{1/2}(\Gamma)$ (see [10]), however, the choice $X = L^2(\Gamma)$ is more convenient and is taken as a first attempt.

Coupling The membrane and fluid are coupled kinematically by imposing

$$g = \partial_t \mathbf{x} \cdot \mathbf{n} \text{ at } \Gamma_w \quad (7)$$

on the fluid and dynamically through the pressure acting on the membrane, by Bernoulli's law,

$$p = \|\nabla\phi\|^2/2. \quad (8)$$

Closure To ensure uniqueness of the coupled problem, two issues are addressed in this paragraph. The first pertains to the nullspace of the Laplace-Neumann equation, the second to a compatibility condition on the boundary data.

It is well-known that (3) is ill-posed, a property that carries over to the boundary integral equation (BIE) (4), see [7]. It follows by substitution that an arbitrary constant can be added to the solution. Thus, the *nullspace* is nontrivial. Physically, the equation is derived to find a velocity field, which is the gradient of ϕ . This constant is therefore

an artifact and it is dealt with by imposing the auxiliary condition $\oint \phi d\Gamma = 0$ through a Lagrange multiplier approach. Accordingly, we add to (6): *find* $\lambda \in \mathbb{R}$ *s.t.* $\forall \mu \in \mathbb{R}$:

$$\lambda \oint_{\Gamma} w d\Gamma + \mu \oint_{\Gamma} \phi d\Gamma = 0. \quad (9)$$

This freedom in the nullspace comes at the cost of freedom in the choice of valid boundary data, i.e., g has to satisfy a *compatibility condition*. Indeed, by Gauss' theorem, it is apparent that the boundary integral of g should vanish. The boundary data g depends on the structural displacement, however, via the kinematic condition (7), and the structural solution generally does not comply with the compatibility condition. To avoid incompatibility of the structural displacement, a constraint is imposed on the (instantaneous) volume \mathcal{V} of Ω , viz.,

$$C(\mathbf{x}; t) := \mathcal{V}(\mathbf{x}(t)) - \mathcal{V}(\mathbf{x}(0)) - \int_0^t \int_{\Gamma_i} v_n d\Gamma d\tau = 0. \quad (10)$$

The constraint (10) is again imposed weakly through a Lagrange multiplier approach, by adding to (2): *find* $\lambda \in \mathbb{R}$, *s.t.* $\forall \mu \in \mathbb{R}$:

$$\lambda C(\mathbf{x}; t) + \mu C'(\mathbf{x}; t)(\mathbf{w}) = 0. \quad (11)$$

We remark that (11) is derived from the constrained minimization problem corresponding to (2), see [11].

3 Discretization

Spatial semi-discretization The parametric space $(0, L)$ is partitioned into N_e segments κ^e with maximum length h . This generates the computational mesh $\mathcal{T}_h := \{\kappa^e\}_{e < N_e}$.¹ A $(p + 1)$ -dimensional polynomial basis $\{N_i(\xi)\}_{i \leq p}$ is defined on the reference segment $\hat{\kappa} := (-1, 1)$ and mapped by M^e to each κ^e , such that a solution $\mathbf{u} : (0, L) \rightarrow \mathbb{R}^d$ can be approximated by

$$\mathbf{u}(s) \approx \sum_{e,i} N_i \circ [M^e]^{-1} \hat{\mathbf{u}}_i^e,$$

where $\hat{\mathbf{u}}_i^e \in \mathbb{R}^d$. Following the isoparametric concept, a similar relation holds for the geometry, \mathbf{x} , of the fluid subproblem. We employ a piecewise linear basis $\{N_i(\xi)\} := \{(1 - \xi)/2, (1 + \xi)/2\}$. This way, the discrete approximation of Γ satisfies the regularity assumption.

The fluid approximation space is easily derived from that of the structure by reducing the dimension d and appending a discretization of the inflow boundary Γ_i . Where element

¹Indices i, j and e are taken to run in \mathbb{N} from 0.

integrals are singular, they are evaluated analytically. On the polygonal boundary with linear N_i we find, if $s, s_{\mathbf{y}} \in \kappa^e$:

$$\begin{aligned} (N_i \circ [M^e]^{-1}(s), H(s; s_{\mathbf{y}}) * N_j \circ [M^e]^{-1}(s_{\mathbf{y}}))_{L^2} &= 0, \\ (N_i \circ [M^e]^{-1}(s), G(s; s_{\mathbf{y}}) * N_j \circ [M^e]^{-1}(s_{\mathbf{y}}))_{L^2} &= (6 + (-1)^{i+j} - 4 \log 2J) \frac{\mathbf{n}_{\mathbf{y}} J^2}{8\pi}. \end{aligned}$$

Time As elucidated in [5, 3], strong coupling involving an incompressible fluid imposes strong stability requirements on the time discretization. Usually, this entails implicit schemes, such as the 1st-order backward Euler scheme proposed here.² The solution vector \mathbf{u} of the 2nd-order structural semi-discretization is advanced with time-step τ to time-level $n + 1$ by

$$\mathbf{M}(\mathbf{u}^{n+1} - 2\mathbf{u}^n + \mathbf{u}^{n-1})\tau^{-2} + \mathbf{K}(\mathbf{u}^{n+1}) = \mathbf{f}^{n+1}, \quad (12)$$

with \mathbf{M} , \mathbf{K} and \mathbf{f} the mass, stiffness and forcing derived from (11). A Newton procedure is employed to solve this nonlinear system. Our selection of the Euler scheme is motivated by its simplicity and good numerical damping.

4 Results

Two notes are made on the solution of the discrete system presented in sec. 3. Firstly, to solve the nonlinear weak form of the structure, the Newton procedure is employed. This requires second derivatives of (11) with respect to \mathbf{x} . The only non-linear term in (2) pertains to the stiffness, and its linearization yields

$$\begin{aligned} \partial_{\mathbf{x}} [(\partial_s \mathbf{w}, J^{-1} \partial_s \mathbf{x})_{L^2}] (\cdot) &= \left(\partial_s \mathbf{w}, \frac{\|\partial_s \mathbf{x}\| \partial_s(\cdot) - \partial_s \mathbf{x} \frac{1}{2} \|\partial_s \mathbf{x}\|^{-1} 2(\partial_s \mathbf{x} \cdot \partial_s(\cdot))}{\|\partial_s \mathbf{x}\|^2} \right)_{L^2} \\ &= (\partial_s \mathbf{w}, [J^{-1} \text{Id} - J^{-3} \partial_s \mathbf{x} \otimes \partial_s \mathbf{x}] \partial_s(\cdot))_{L^2} \end{aligned}$$

Boundedness of J^{-1} is guaranteed if $\mathbf{x}(0, \cdot) \neq \mathbf{x}(L, \cdot)$. Note that the tangent stiffness $\partial_{\mathbf{x}}(\partial_s \mathbf{w}, E[1 - J^{-1}] \partial_s \mathbf{x})_{L^2}$ is possibly singular, posing problems for the steady case, such as the hoop stress test presented in this section. In this case the tangent above is approximated by *perfect elasticity* [13], i.e., by letting $J \rightarrow \infty$.

Secondly, the volume of the airbag, that is required to impose the compatibility condition 10, can be computed by the determinant rule, according to:

$$\mathcal{V}(\mathbf{x}) = \sum_{\mathcal{T}_h} |\hat{\mathbf{x}}^e|, \quad (13)$$

where the correct numbering should be noted, i.e., $\forall e < N_e : M^e(-1) < M^e(1)$.

The proposed discretization has been verified by various numerical experiments:

²However, recently a 1st-order explicit scheme with defect correction was proposed in [4].

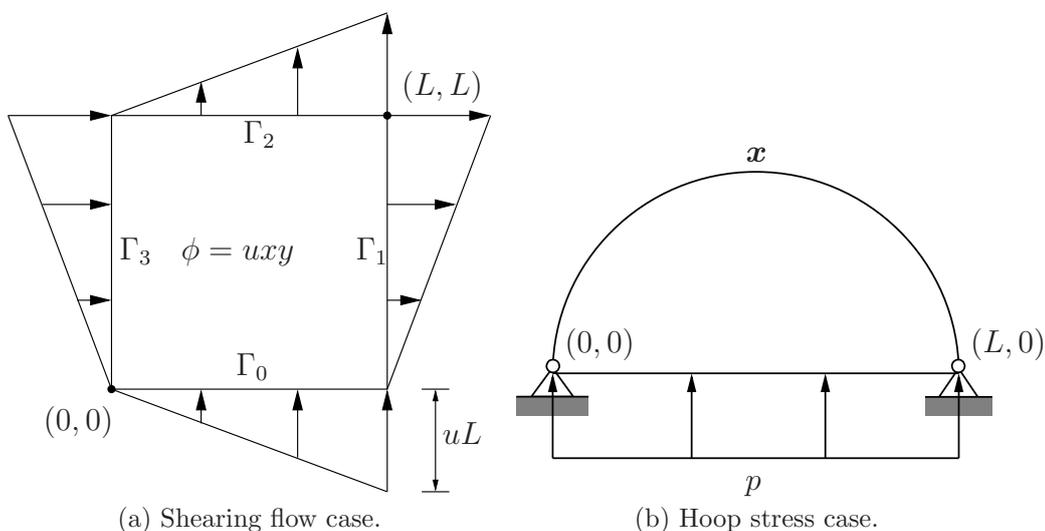


Figure 1: Schematics of the test cases.

- i) the fluid discretization, by a known solution representing *shearing* motion;
- ii) the structural semi-discretization, by finding the *hoop stress*;
- iii) the dynamic fluid–structure–interaction formulation, by the inflation of a *pancake*-shaped domain.

Shearing flow To verify the fluid discretization, the solution $\phi = uxy$ is assumed on a unit square. The resulting Neumann conditions suggest the evolution of the boundary from a diamond-shape to a square, see fig. 1a.

The potential and pressure solutions are given in fig. 2. The solution ϕ is very accurate for both quadrature orders. However, in pressure, the tangential derivative clearly reduces the regularity of the solution at the corners of the domain. From the convergence plots in figure 3, we observe the theoretical quadratic convergence [6] in ϕ and the residual, provided quadrature is high enough. If numerical quadrature is used on an element close to the singularity, the quadrature should be sufficiently high. As N_e increases, the neighboring element moves toward this singularity and these errors start to deteriorate the solution, as can be seen from the low-order quadrature curve for p in fig. 3b.

If the inflow is locally perturbed in an incompatible fashion, it can be seen that pressure changes remain local, except increased oscillations at the corners. Thus the pressure distribution does not cause the membrane movement to accommodate the extra inflow, so volume conservation will have to be enforced explicitly, as implied in sec. 2.

Hoop stress To verify the structure discretization, a constant pressure is applied such that a taut string is elongated to a semi-circle. The strain then equals $\pi - 1$, a straight-

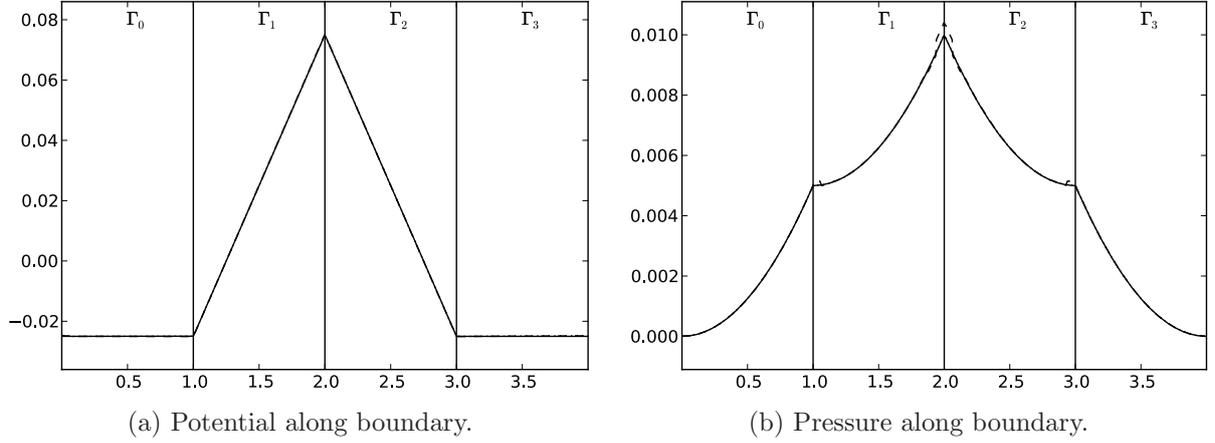


Figure 2: Shearing flow, $N_e = 64$. Plotted are the analytic solution (—), 4th-order (---) and 32nd-order (···) Gaussian quadrature approximations. Errors of these solutions are given along (— · —) in figure 3b.

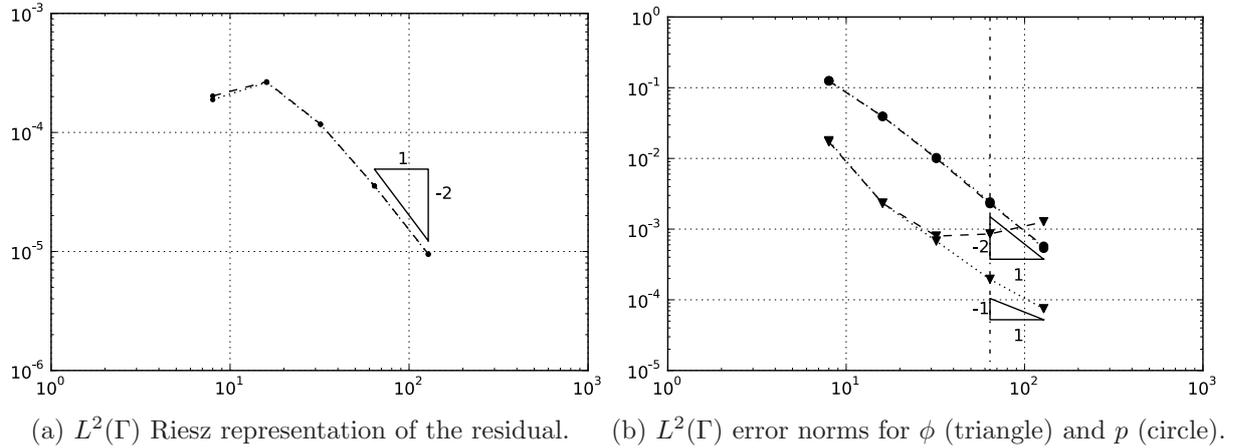
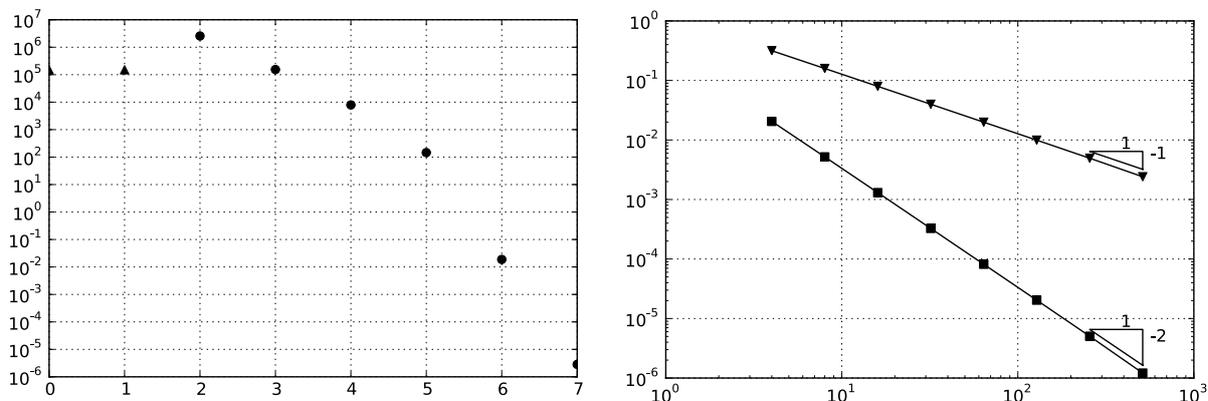


Figure 3: Convergence with N_e (along horizontal axes) for shearing flow w.r.t. the analytical solution of the 4th-order (---) and 32nd-order (···) Gaussian quadrature approximations.



(a) Subiteration residual decay for $N_e = 512$, approx-
imate tangent (triangle), exact tangent (circle). (b) Spatial convergence (N_e varies along the horizontal axis) w.r.t. the analytical solution, $H^1(0, L; \mathbb{R}^2)$ norm (triangle), $L^2(0, L; \mathbb{R}^2)$ norm (square).

Figure 4: Subiteration- and convergence behavior for the hoop stress case.

forward calculation conveys the required internal pressure $p = (\pi - 2)E/L$, cf. fig. 1b.

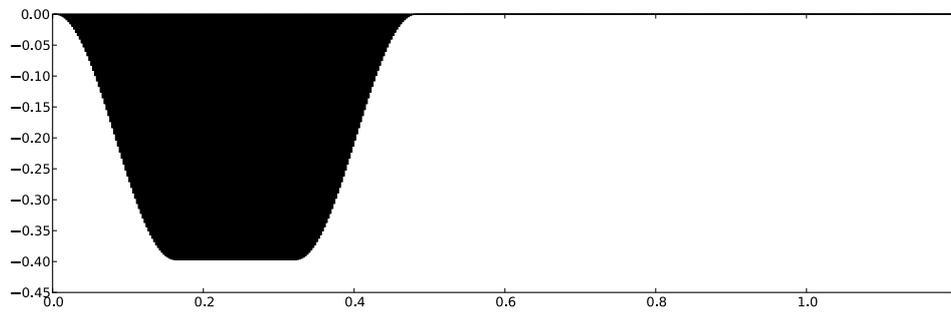
We restrict ourselves to the steady, nonlinear problem. In the Newton method, the residual (plotted in fig. 4a) should decay quadratically where the exact tangent is applied, as it is seen to do. Figure 4b shows that the formulation provides linear and quadratic convergence in the H^1 and L^2 norms, respectively. These convergence rates are optimal.

Pancake Finally, we verify the coupled problem qualitatively, by considering a pancake-shaped unstressed initial configuration at rest. An inflow is then applied, following $\partial_t \mathcal{V} = \overline{\partial_t \mathcal{V}} S(s) T(t)$, with

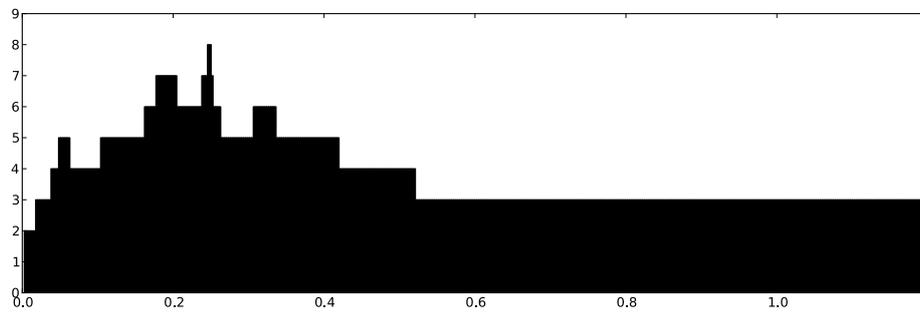
$$S(s) = 4s(L_i - s)/L_i^2$$

$$T(t) = \begin{cases} (1 - \cos(\pi t/T_1))/2, & 0 < t \leq T_1, \\ 1, & T_1 < t \leq T_2, \\ (1 + \cos(\pi(t - T_2)/T_1))/2, & T_2 < t \leq T_1 + T_2, \\ 0, & T_1 + T_2 < t < T. \end{cases}$$

In these relations we have the mean inflation flux $\overline{\partial_t \mathcal{V}} = (\mathcal{V}(0) - \mathcal{V}(T_1 + T_2))/T_2$, inflater opening L_i and two time instances $0 < T_1 \leq T_2$. Note that the mean flux has a negative sign as it is directed into the enclosure. The inflow has a quadratic profile and its magnitude is gradually in- and decreased in time, see fig. 5, which also shows the subiteration history. Snapshots at different time-levels are given in fig. 6. Near the inflater, the membrane is initially concave which causes it to compress under the action of pressure. It is observed that the subsequent wrinkling does not impede the numerical process.



(a) Influx per time-interval.



(b) Subiteration count.

Figure 5: Evolution of total inflow (top) and subiteration count (bottom) in time for the pancake test case.

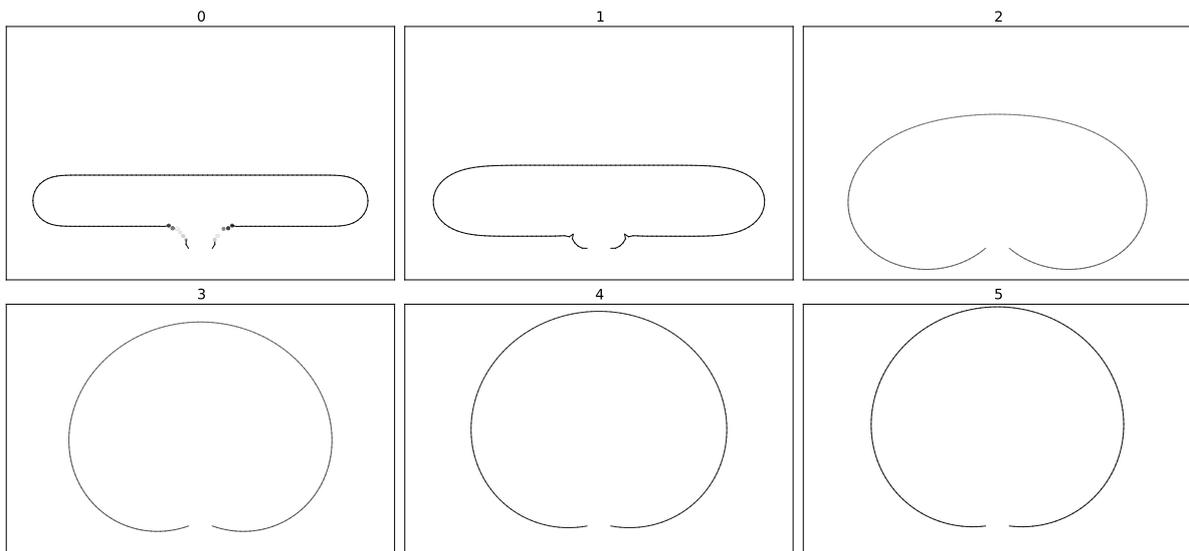


Figure 6: Snapshots of the response in the pancake test case.

5 Conclusions

We have presented an investigation of a computational approach for the simulation of a nonlinear membrane in interaction with an inviscid, incompressible and irrotational fluid, based on the boundary element method. We showed that the compatibility condition on the boundary velocity originating from the incompressibility of the fluid, must be imposed explicitly on the structural motion as an auxiliary constraint. Numerical experiments conveyed optimal convergence of the finite-element approximations of the fluid and structure subsystems, and of the coupled fluid-structure-interaction problem. Our results indicate that BEM provides a very effective approach to large-displacement fluid-structure interactions on geometrically complex domains, because it avoids the mesh degradation inherent to standard volumetric discretization methods. Moreover, it significantly reduces the computational complexity, because the fluid solution is only computed on the boundary.

REFERENCES

- [1] Stuart S. Antman. The equations for large vibrations of strings. *The American Mathematical Monthly*, 87(5):359–370, 1980.
- [2] E. H. Brummelen, van. Partitioned iterative solution methods for fluid–structure interaction. *International Journal for Numerical Methods in Fluids*, 65(13):3–27, 2002.
- [3] E. H. Brummelen, van. Added mass effects of compressible and incompressible flows in fluid-structure interaction. *Journal of Applied Mechanics*, 76(2), 2009.
- [4] Erik Burman and Miguel A. Fernández. Stabilization of explicit coupling in fluid-structure interaction involving fluid incompressibility. *Computer Methods in Applied Mechanics and Engineering*, 198(5-8):766–784, 2009.
- [5] P. Causin, J. F. Gerbeau, and F. Nobile. Added-mass effect in the design of partitioned algorithms for fluid-structure problems. *Computer Methods in Applied Mechanics and Engineering*, 194(42-44):4506–4527, 2005.
- [6] W. Dijkstra, G. Kakuba, and R.M.M. Mattheij. Condition numbers and local errors in the boundary element method. Technical report, Eindhoven, University of Technology, P.O. Box 513, 5600MB, Eindhoven, Netherlands, 2008.
- [7] W. Dijkstra and R.M.M. Mattheij. A relation between the logarithmic capacity and the condition number of the bem-matrices. *Communications in Numerical Methods in Engineering*, 23:665–680, 2007.
- [8] S. Liapis. An adaptive boundary element method for the solution of potential flow problems. *Engineering Analysis with Boundary Elements*, 18:29–37, 1996.

- [9] P.-O. Marklund and L. Nilsson. Simulation of airbag inflation processes using a coupled fluid structure approach. *Computational Mechanics*, 29(4-5):289–297, 2002.
- [10] W. McLean. *Strongly elliptic systems and boundary integral equations*. Cambridge University Press, 2000.
- [11] Brian Moran Ted Belytschko, Wing Kam Liu. *Nonlinear Finite Elements for Continua and Structures*. Wiley, 2000.
- [12] L.C. Wrobel. *The Boundary Element Method*, volume 1. Applications in Thermo-Fluids and Acoustics. John Wiley & Sons, 2002.
- [13] D. Yong. Strings, chains and ropes. *SIAM*, 48(4):771–781, 2006.

IMMERSED BOUNDARY METHOD: PERFORMANCE ANALYSIS OF POPULAR FINITE ELEMENT SPACES

DANIELE BOFFI*, NICOLA CAVALLINI*, FRANCESCA GARDINI*
AND LUCIA GASTALDI†

*Dipartimento di Matematica “F. Casorati”
Università di Pavia
Via Ferrata 1, 27100 Pavia, Italy
{daniele.boffi,nicola.cavallini,francesca.gardini}@unipv.it

†Dipartimento di Matematica
Università di Brescia,
via Valotti, 9 - 25133, Brescia, Italy,
lucia.gastaldi@ing.unibs.it

Key words: Finite Elements, Immersed Boundary Method, Fluid-Structure Interactions, Mass conservation.

Abstract. The aim of this paper is to understand the performances of different finite elements in the space discretization of the Finite Element Immersed Boundary Method. In this exploration we will analyze two popular solution spaces: *Hood-Taylor* and *Bercovier-Pironneau* (P1-iso-P2). Immersed boundary solution is characterized by pressure discontinuities at fluid structure interface. Due to such a discontinuity a natural enrichment choice is to add piecewise constant functions to the pressure space. Results show that $P_1 + P_0$ pressure spaces are a significant cure for the well known “boundary leakage” affecting IBM. Convergence analysis is performed, showing how the discontinuity in the pressure is affecting the convergence rate for our finite element approximation.

1 INTRODUCTION

Several applications involve the dynamical interaction of solids and fluids. Compatibility and dynamical conditions are set to quantify the mechanical phenomenology involving the solid and fluid phases. Solid equations are naturally written in a Lagrangian framework, fluid equations are written in an Eulerian framework. The duality between Lagrangian and Eulerian coordinates is the first issue regarding fluid-structure interactions.

There are different strategies for the coupling of Lagrangian and Eulerian descriptions. Among those we recall the Arbitrary Lagrangian Eulerian approach [13, 14, 12, 15] and

the fictitious domain method [8, Ch. 8].

The subject of this paper is the immersed boundary method (IBM). The IBM was first introduced by Peskin in the 70's in order to simulate the heart blood flow during a cardiac beat. In the case of immersed boundary method the structure is fully immersed in the fluid. We refer to [16] for a review of the method. The main idea is to consider the structure as a part of the fluid. At the fluid and solid overlap, the resulting stress tensor and density consist of a combination of the fluid and solid ones. In this manner fluid and solid can be treated in their natural formulation, Eulerian and Lagrangian respectively. In its original finite difference formulation the IBM takes into account fluid-solid interface conditions by means of a Dirac delta function. The finite difference approximation of Dirac delta functions is characterized by an intrinsically diffusive behavior. A great effort is produced to stabilize this effect and to minimize the well know "boundary leakage" phenomena [10, 11, 9].

Recently, a finite element approximation of the IBM has been introduced [3, 6, 5, 7, 2]. The finite element discretization of IBM treats variationally the interface between the fluid and the solid, resulting in a natural embedding of the solid equations into the fluid ones. The variational treatment of the discontinuity results in a sharper separation of fluid and solid, see [3, 4].

In this paper we explore the performances related to different popular solution spaces for the Navier-Stokes equations solution. We start analyzing the *Bercovier–Pironneau* (P1-iso-P2) and the *Hood–Taylor* finite elements. Then we take advantage of the variational treatment of the fluid solid interface, and we add piecewise constant functions to the pressure space. In this way we give a variational interpretation of the discontinuity between fluid and solid. Adding a P_0 to the pressure solution space results in a five times reduced "boundary leakage". Moreover the oscillations due to the approximation of discontinuous solution with continuous pressure functions are reduced. On the other hand we will show that the solution spaces for the velocities are characterized by non intuitive behavior. The *Hood–Taylor* element is more accurate than the *Bercovier–Pironneau*, but the resulting diffusivity is higher.

This paper is organized as follows. In Section 2 we recall the formulation of our problem, in Section 3 we recall the discrete formulation of our problem, Section 4 is dedicated to the numerical experiments. Careful attention will be paid to the results of choices in Section 3.1. Finally in Section 5 we draw our conclusions.

2 PROBLEM FORMULATION

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a region containing a fluid and an immersed solid material. In particular, we focus on viscous incompressible fluids and incompressible viscoelastic structures. The natural framework for the fluid dynamics is Eulerian, the solid mechanics is usually described using Lagrangian coordinates. The key feature of the IB is assuming that the stress tensor at the overlapping between fluid and solid equals the sum of the fluid and structure stress tensors. In this way fluid and structure can be properly treated

in their natural framework.

At time t the solid body is located in $\mathcal{B}_t \subset \Omega$, which is the image of a reference domain $\mathcal{B} \subset \mathbb{R}^m$, $m = d, d - 1$, through a mapping \mathbf{X} . We set up our Lagrangian framework denoting \mathbf{s} the variable in the reference domain \mathcal{B} and \mathbf{x} as a material point in the current domain \mathcal{B}_t . Moreover, we assume that $\partial\mathcal{B}_t \cap \partial\Omega = \emptyset$. From now on we consider the reference domain coinciding with the initial domain, that is $\mathcal{B} = \mathcal{B}_0$.

The relationship between the two different frameworks is given by the mapping

$$\mathbf{X} : \mathcal{B} \times [0, T] \rightarrow \mathcal{B}_t \quad \text{so that } \mathbf{x} = \mathbf{X}(\mathbf{s}, t) \quad \forall \mathbf{x} \in \mathcal{B}_t. \quad (1)$$

We assume that at any given time the mapping $\mathbf{X}(\mathbf{s}, t)$ is invertible, which implies that the deformation gradient

$$\mathbb{F}_{\alpha i} := \left(\nabla_{\mathbf{s}} \mathbf{X}(\mathbf{s}, t) \right)_{\alpha i} = \mathbf{X}_{\alpha, i}(\mathbf{s}, t) = \frac{\partial \mathbf{X}_{\alpha}(\mathbf{s}, t)}{\partial s_i}, \quad (2)$$

has rank m . We assume that the quantity $|\mathbb{F}|$ defined as

$$|\mathbb{F}| = \begin{cases} \left| \frac{\partial \mathbf{X}}{\partial s} \right| & \text{if } m = 1 \\ \left| \frac{\partial \mathbf{X}}{\partial s_1} \wedge \frac{\partial \mathbf{X}}{\partial s_2} \right| & \text{if } m = 2 \\ \left| \left(\frac{\partial \mathbf{X}}{\partial s_1} \wedge \frac{\partial \mathbf{X}}{\partial s_2} \right) \cdot \frac{\partial \mathbf{X}}{\partial s_3} \right| & \text{if } m = 3, \end{cases} \quad (3)$$

is positive at time $t = 0$, and therefore at any subsequent time.

Velocity and acceleration of a particle at \mathbf{x} at time t are respectively defined as:

$$\mathbf{u}(\mathbf{x}, t) = \frac{\partial \mathbf{X}}{\partial t}(\mathbf{s}, t) \quad \text{for } \mathbf{x} = \mathbf{X}(\mathbf{s}, t), \quad (4)$$

and

$$\dot{\mathbf{u}}(\mathbf{x}, t) = \frac{D\mathbf{u}}{dt}(\mathbf{x}, t) = \frac{\partial^2 \mathbf{X}}{\partial t^2}(\mathbf{s}, t) = \frac{\partial \mathbf{u}}{\partial t}(\mathbf{x}, t) + \mathbf{u}(\mathbf{x}, t) \cdot \nabla \mathbf{u}(\mathbf{x}, t), \quad (5)$$

We assume homogeneous and incompressible fluid and solid, so that the density ρ is piecewise constant:

$$\rho = \begin{cases} \rho_f & \text{in } \Omega \setminus \mathcal{B}_t \\ \rho_s & \text{in } \mathcal{B}_t. \end{cases} \quad (6)$$

Detailed study of stability criteria regarding the ratio ρ_s/ρ_f can be found in [2].

The Cauchy stress tensor $\boldsymbol{\sigma}$, is considered equal to the fluid one, where no solid phase is located. Where the fluid and the solid overlap the fluid and the solid stresses are summed:

$$\boldsymbol{\sigma} = \begin{cases} \boldsymbol{\sigma}_f & \text{in } \Omega \setminus \mathcal{B}_t \\ \boldsymbol{\sigma}_f + \boldsymbol{\sigma}_s & \text{in } \mathcal{B}_t \end{cases} \quad (7)$$

This assumption is satisfied for example by some biological tissues, subject of several fluid structure interaction problems (see, e.g., [18]). The fluid stress tensor is specified for viscous fluids of type:

$$\boldsymbol{\sigma}_f = -p\mathbb{I} + \mu(\nabla \mathbf{u} + (\nabla \mathbf{u})^T) \quad (8)$$

The solid stress tensor $\boldsymbol{\sigma}_s$ is expressed in Lagrangian variables by means of the first Piola-Kirchhoff stress tensor \mathbb{P} :

$$\mathbb{P}(\mathbf{s}, t) = |\mathbb{F}(\mathbf{s}, t)| \boldsymbol{\sigma}_s(\mathbf{X}(\mathbf{s}, t), t) \mathbb{F}^{-T}(\mathbf{s}, t). \quad (9)$$

The previous definitions are combined with the principal of virtual work, the balance of momenta and $|\mathbb{F}| = 1$ to get:

$$\int_{\Omega} (\rho_f \dot{\mathbf{u}} - \nabla \cdot \boldsymbol{\sigma}_f) \cdot \mathbf{v} d\mathbf{x} = -(\rho_s - \rho_f) \int_{\mathcal{B}} \frac{\partial^2 \mathbf{X}}{\partial t^2} \cdot \mathbf{V} ds + \int_{\mathcal{B}} (\nabla_s \cdot \mathbb{P}) \cdot \mathbf{V} ds - \int_{\partial \mathcal{B}} \mathbb{P} \mathbf{N} \cdot \mathbf{V} dA. \quad (10)$$

where $\mathbf{V}(\mathbf{s}, t) = \mathbf{v}(\mathbf{X}(\mathbf{s}, t))$, and \mathbf{N} is the outer normal to the region \mathcal{B} in Lagrangian coordinates. The change of variables between Eulerian and Lagrangian can be avoided using the defining property of the d -dimensional Dirac delta distribution:

$$\mathbf{V}(\mathbf{s}, t) = \mathbf{v}(\mathbf{X}(\mathbf{s}, t)) = \int_{\Omega} \mathbf{v}(\mathbf{x}) \delta(\mathbf{x} - \mathbf{X}(\mathbf{s}, t)) d\mathbf{x} \quad \forall \mathbf{s} \in \mathcal{B}. \quad (11)$$

Taking into account the definition of $\boldsymbol{\sigma}_f$ and noticing that \mathbf{v} is arbitrary we conclude the derivation of our problem: find \mathbf{u} , p and \mathbf{X} which satisfy the Navier–Stokes equations

$$\rho_f \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) - \mu \Delta \mathbf{u} + \nabla p = \mathbf{d} + \mathbf{f} + \mathbf{t} \quad \text{in } \Omega \times]0, T[\quad (12)$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega \times]0, T[\quad (13)$$

with the following source terms

$$\mathbf{d}(\mathbf{x}, t) = -(\rho_s - \rho_f) \int_{\mathcal{B}} \frac{\partial^2 \mathbf{X}}{\partial t^2} \delta(\mathbf{x} - \mathbf{X}(\mathbf{s}, t)) ds \quad \text{in } \Omega \times]0, T[\quad (14)$$

$$\mathbf{f}(\mathbf{x}, t) = \int_{\mathcal{B}} \nabla_s \cdot \mathbb{P} \delta(\mathbf{x} - \mathbf{X}(\mathbf{s}, t)) ds \quad \text{in } \Omega \times]0, T[\quad (15)$$

$$\mathbf{t}(\mathbf{x}, t) = - \int_{\partial \mathcal{B}} \mathbb{P} \mathbf{N} \delta(\mathbf{x} - \mathbf{X}(\mathbf{s}, t)) dA \quad \text{in } \Omega \times]0, T[\quad (16)$$

and the following equation for the immersed boundary with suitable boundary and initial conditions

$$\frac{\partial \mathbf{X}}{\partial t}(\mathbf{s}, t) = \mathbf{u}(\mathbf{X}(\mathbf{s}, t), t) \quad \text{in } \mathcal{B} \times]0, T[\quad (17)$$

$$\mathbf{u}(\mathbf{x}, t) = 0 \quad \text{on } \partial \Omega \times]0, T[\quad (18)$$

$$\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}) \quad \text{in } \Omega \quad (19)$$

$$\mathbf{X}(\mathbf{s}, 0) = \mathbf{X}_0(\mathbf{s}) \quad \text{in } \mathcal{B}. \quad (20)$$

3 FINITE ELEMENT DISCRETIZATION

We now recall how to discretize the problem using finite element in space and a semi-implicit approach in time. We refer to [2] for more details.

Let \mathcal{T}_h be a subdivision of Ω into triangles or rectangles if $d = 2$, tetrahedrons or parallelepipeds if $d = 3$, being K a single element of \mathcal{T}_h . We consider two finite dimensional spaces $V_h \subseteq H_0^1(\Omega)^d$ and $Q_h \subseteq L_0^2(\Omega)$.

Next consider a subdivision \mathcal{S}_h of \mathcal{B} into segments, triangles or tetrahedrons for $m = 1, 2, 3$ respectively. We shall use the following notation: T_k , $k = 1, \dots, M_e$ denotes an element of \mathcal{S}_h , \mathbf{s}_j , $j = 1, \dots, M$ stands for a vertex of \mathcal{S}_h , and \mathcal{E}_h indicates the set of the edges (or faces) e of \mathcal{S}_h . Then S_h is the finite element space of piecewise linear d -vectors defined on \mathcal{B} as follows

$$S_h = \{\mathbf{Y} \in C^0(\mathcal{B}; \Omega) : \mathbf{Y}|_{T_k} \in P_1(T_k)^d, k = 1, \dots, M_e\}, \quad (21)$$

where $P_1(T_k)$ stands for the space of affine polynomials on the element T_k . For an element $\mathbf{Y} \in S_h$ we shall use also the following notation $\mathbf{Y}_j = \mathbf{Y}(\mathbf{s}_j)$ for $j = 1, \dots, M$.

One of the crucial points of our method is the evaluation of the force term. We consider \mathbf{X}_h piecewise linear so that the deformation gradient $\mathbb{F}_h = \nabla_s \mathbf{X}_h$ is piecewise constant. Therefore also $\mathbb{P}_h = \mathbb{P}(\mathbb{F}_h)$ is piecewise constant and we can work as follows for all $\mathbf{v} \in V_h$:

$$\begin{aligned} \langle \mathbf{F}_h(t), \mathbf{v} \rangle &= - \sum_{k=1}^{M_e} \int_{T_k} \mathbb{P}_h|_{T_k} : \nabla_s \mathbf{v}(\mathbf{X}_h) \, ds \\ &= - \sum_{k=1}^{M_e} \int_{\partial T_k} \mathbb{P}_h|_{T_k} \mathbf{N} \cdot \mathbf{v}(\mathbf{X}_h) \, dA \\ &= - \sum_{e \in \mathcal{E}_h} \int_e [[\mathbb{P}_h]] \cdot \mathbf{v}(\mathbf{X}_h) \, dA \end{aligned} \quad (22)$$

where $[[\mathbb{P}_h]]$ is the jump of \mathbb{P}_h across the interelement edge e , defined as:

$$[[\mathbb{P}_h]] = \mathbb{P}_h^+ \mathbf{N}^+ + \mathbb{P}_h^- \mathbf{N}^-, \quad (23)$$

and \mathbf{N}^+ and \mathbf{N}^- are the normals to the interface e pointing outward “+” or inward “−” the element. Moreover we use the convention that, when $e \in \partial \mathcal{B}$, then $[[\mathbb{P}_h]] = \mathbb{P}_h \mathbf{N}$, where \mathbf{N} is the outer normal to \mathcal{B} . We observe that e could be either a face (if $m = 3$) or an edge (if $m = 2$) or a point (if $m = 1$). In the first two cases the integrals are computed with a suitable quadrature formula, while in the latter case there are no integrals in the formula. Moreover, the computation of (22) requires an interpolation procedure in order to evaluate the test function along the structure.

3.1 Stable Finite Elements

In this work we focus on two popular finite elements in two dimensions: *Bercovier–Pironneau* (P1-isoP2) and *Hood–Taylor*. In the IB method the structure imposes a discontinuity in the solution for the stress while the velocity is still continuous. This is the reason why we add piecewise constant functions to the pressure space for both elements.

The numerical experiments will compare *Hood–Taylor* finite element:

$$\begin{aligned} V_h &= \{\mathbf{v} \in H_0^1(\Omega)^2 : \mathbf{v}|_K \in P_2(K)^2 \forall K \in \mathcal{T}_h\}, \\ Q_h &= \{q \in L_0^2(\Omega) : q = q_1, q_1 \in C(\bar{\Omega}), q_1|_K \in P_1(K) \forall K \in \mathcal{T}_h\}, \end{aligned} \quad (24)$$

and its augmented version:

$$\begin{aligned} V_h &= \{\mathbf{v} \in H_0^1(\Omega)^2 : \mathbf{v}|_K \in P_2(K)^2 \forall K \in \mathcal{T}_h\}, \\ Q_h &= \{q \in L_0^2(\Omega) : q = q_1 + q_0, q_1 \in C(\bar{\Omega}), q_1|_K \in P_1(K), q_0|_K \in P_0(K) \forall K \in \mathcal{T}_h\}. \end{aligned} \quad (25)$$

The *Bercovier–Pironneau* will be also tested:

$$\begin{aligned} V_h &= \{\mathbf{v} \in H_0^1(\Omega)^2 : \mathbf{v}|_K \in P_1(K)^2 \forall K \in \mathcal{T}_{h/2}\}, \\ Q_h &= \{q \in L_0^2(\Omega) : q = q_1, q_1 \in C(\bar{\Omega}), q_1|_K \in P_1(K) \forall K \in \mathcal{T}_h\}, \end{aligned} \quad (26)$$

and its augmented version will be considered as well:

$$\begin{aligned} V_h &= \{\mathbf{v} \in H_0^1(\Omega)^2 : \mathbf{v}|_K \in P_1(K)^2 \forall K \in \mathcal{T}_{h/2}\}, \\ Q_h &= \{q \in L_0^2(\Omega) : q = q_1 + q_0, q_1 \in C(\bar{\Omega}), q_1|_K \in P_1(K), q_0 \in P_0(K) \forall K \in \mathcal{T}_h\}, \end{aligned} \quad (27)$$

here $\mathcal{T}_{h/2}$ is the mesh obtained connecting the midpoints of every element of \mathcal{T}_h . In [1] we used the macroelement technique to prove that the inf-sup condition, for the Stokes problem, is satisfied for the augmented finite elements (see also [17, 19]). Introducing the augmented pressure space for the Stokes problem we reduced the difference between analytical and numerical solution. Moreover the *Hood–Taylor* confirmed a more accurate behaviour with respect to *Bercovier–Pironneau*.

In the case of IB we expect two significant improvements in the elements performances:

1. A better mass conservation. Considering a $P_1 + P_0$ test function for the pressure is equivalent to impose the subsequent restriction:

$$\int_K \nabla \cdot \mathbf{u} \, d\mathbf{x} = 0 \quad (28)$$

being K a generic triangle in \mathcal{T}_h .

2. We expect a sharper profile in the pressure solution. The original formulation of *Bercovier–Pironneau* and *Hood–Taylor* approximate the pressure with continuous functions. This causes a Gibbs phenomenon at pressure discontinuity. We expect $P_1 + P_0$ functions to better control this Gibbs phenomenon.

The second order derivative of \mathbf{X}_h with respect to time is discretized with the usual second order finite difference and then approximated using the discrete counterpart of (17). Hence the scheme reads as follows:

Step 1. Compute

$$\langle \mathbf{F}_h^{n+1}, \mathbf{v} \rangle = - \sum_{e \in \mathcal{E}_h} [[\mathbb{P}_h]]^n \cdot \mathbf{v}(\mathbf{X}_h^n(s, t)) \, dA \quad \forall \mathbf{v} \in V_h. \quad (29)$$

Step 2. Solve the Navier-Stokes equations: find $(\mathbf{u}_h^{n+1}, p_h^{n+1}) \in V_h \times Q_h$ such that

$$\begin{aligned} \rho_f \left(\frac{\mathbf{u}_h^{n+1} - \mathbf{u}_h^n}{\Delta t}, \mathbf{v} \right) + b(\mathbf{u}_h^{n+1}, \mathbf{u}_h^{n+1}, \mathbf{v}) + a(\mathbf{u}_h^{n+1}, \mathbf{v}) - (\nabla \cdot \mathbf{v}, p_h^{n+1}) = \\ - (\rho_s - \rho_f) \int_{\mathcal{B}} \frac{\mathbf{u}_h^{n+1}(\mathbf{X}_h^n(s)) - \mathbf{u}_h^n(\mathbf{X}_h^{n-1}(s))}{\Delta t} \cdot \mathbf{v}(\mathbf{X}_h^n(s)) \, ds + \langle \mathbf{F}_h^{n+1}, \mathbf{v} \rangle \quad \forall \mathbf{v} \in V_h \end{aligned} \quad (30)$$

$$(\nabla \cdot \mathbf{u}_h^{n+1}, q) = 0 \quad \forall q \in Q_h.$$

Step 3. Advance the position of the points of the structure:

$$\frac{\mathbf{X}_{hi}^{n+1} - \mathbf{X}_{hi}^n}{\Delta t} = \mathbf{u}_h^{n+1}(\mathbf{X}_{hi}^n) \quad \forall i = 1, \dots, M. \quad (31)$$

4 NUMERICAL EXPERIMENTS

This section is devoted to numerical experiments. We refer to the popular inflated balloon test case. We are considering a two-dimensional domain, the square $[0, 1]^2$, for a Stokes fluid. The immersed boundary is lying along a circle of radius $R = 0.4$. The immersed boundary initial parametric representation is:

$$\mathbf{X}_0(s) = \begin{pmatrix} R \cos(s/R) + 0.5 \\ R \sin(s/R) + 0.5 \end{pmatrix}, \quad s \in [0, 2\pi R].$$

The elastic distributed load is written as follows:

$$\langle \mathbf{F}(t), \mathbf{v} \rangle = -\kappa \int_0^{2\pi R} \frac{\partial \mathbf{X}(s, t)}{\partial s} \frac{\partial \mathbf{v}(\mathbf{X}(s, t))}{\partial s}. \quad (32)$$

where κ represents the structure stiffness.

This parametric representation of the immersed boundary, together with the tension formulation in (32), results in a distributed load per unit length equal to κ/R , radially directed toward the domain center. The resulting Stokes problem is solved using a conjugate gradient method in combination with a Backward Euler time discretization, see [8]. In this simple test case of the inflated balloon we can derive a stationary analytical solution. As soon as we get to a stationary regime the velocity field and pressure map are:

$$\mathbf{u}(\mathbf{x}, t) = \mathbf{0} \quad \forall \mathbf{x} \in \Omega, \quad \forall t \in]0, T[\quad (33)$$

$$p(\mathbf{x}, t) = \begin{cases} \kappa(1/R - \pi R), & |\mathbf{x}| \leq R \\ -\kappa\pi R, & |\mathbf{x}| > R \end{cases} \quad \forall t \in]0, T[. \quad (34)$$

The parameters used for our simulations are: final time ($T = 10^{-1}$), structure density ($\rho_s = 1$), fluid density ($\rho_f = 1$), fluid viscosity ($\mu = 1$), structure stiffness ($\kappa = 1$), fluid mesh amplitude ($h_x = 1/32$), structure mesh amplitude ($h_s = 2\pi R/1024$), time step ($\Delta t = 10^{-4}$). Notice that, for the purpose of this paper we are not interested in the effect of inertia terms, being $\rho_s - \rho_f = 0$. The interested reader can refer to [2] for a detailed analysis of the case $\rho_s \neq \rho_f$.

In Figure 1(a) we compare the performances of the different finite elements in terms of area conservation. We observe that adding piecewise constant functions to the pressure space the area loss is decreased by five times. As predicted in Section 3.1 the imposition of (28) is effective in reducing the “boundary leakage” effect. A deeper understanding of this result is achieved considering Figure 1(b). Here we plot the L^2 -norm of the “real divergence” defined as:

$$\nabla \cdot \mathbf{u}_h = \sum_{j=1}^{df} u_{1j} \frac{\partial \varphi_j}{\partial x} + u_{2j} \frac{\partial \varphi_j}{\partial y}.$$

where φ_j are the basis functions and df is the number of degrees of freedom on $K \in \mathcal{T}_h$ for the *Hood–Taylor* scheme. For the *Bercovier–Pironneau* scheme df are the degrees of freedom of the single triangle in $\mathcal{T}_{h/2}$. The “real divergence” for the *Hood–Taylor* scheme is piecewise linear, while it is piecewise constant for the *Bercovier–Pironneau* one. Plot 1(b) shows that *Hood–Taylor* scheme is affected by slightly greater values of the “real divergence”. This is explaining why the area loss is slightly larger for P_2 finite elements. This result can be at first surprising in the sense that improving the element accuracy provides an overall scheme with slightly more diffusive properties.

Figures 3 and 4 represent the pressure profile at the mid section of the domain for the *Hood–Taylor* and the *Bercovier–Pironneau* schemes, respectively. The numerical and analytical solutions are superimposed. A detailed representation of the Gibbs phenomenon is shown in 3(b), and 4(b). In both cases, the augmented pressure space is effective in reducing the oscillations, resulting in a sharper pressure profile.

Tables 1 and 2 report the convergence analysis for the four finite elements. In the same manner as in [1], augmenting the pressure space is reducing the solution error for both the finite elements. The rate of convergence for *Bercovier–Pironneau* and *Hood–Taylor* elements is comparable, since the discontinuity in the solution provides an upper limit for the achievable accuracy.

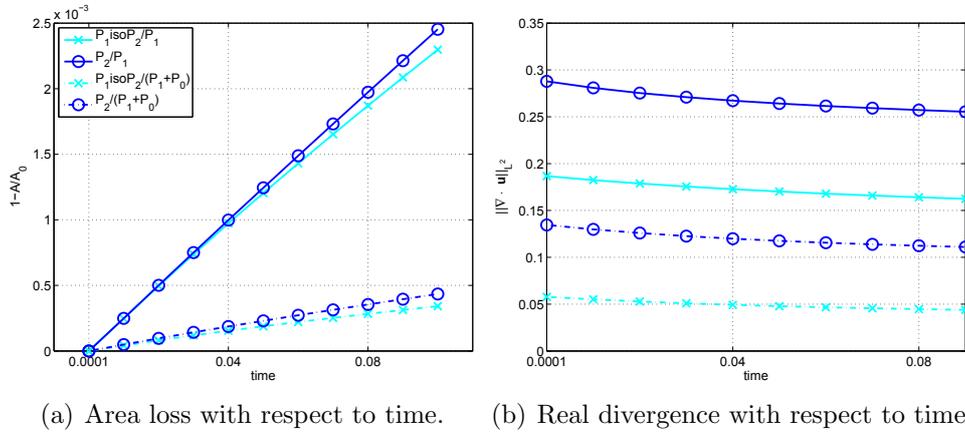


Figure 1: Boundary leakage and real divergence for solution spaces.

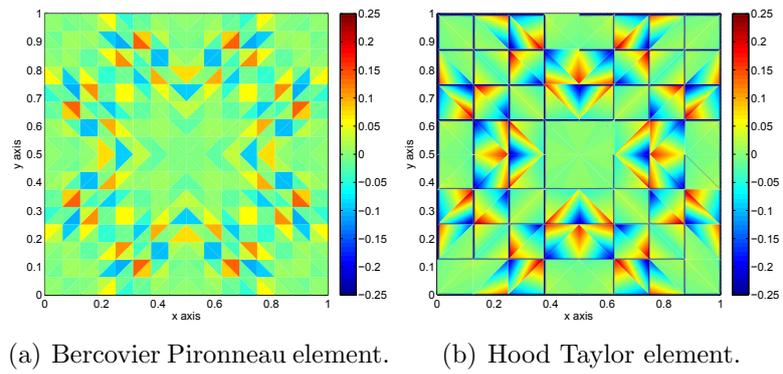


Figure 2: Real divergence for the augmented spaces.

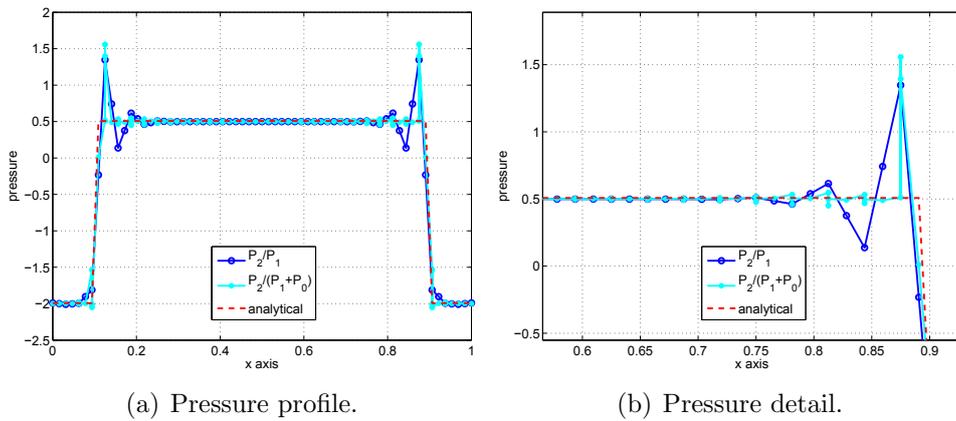
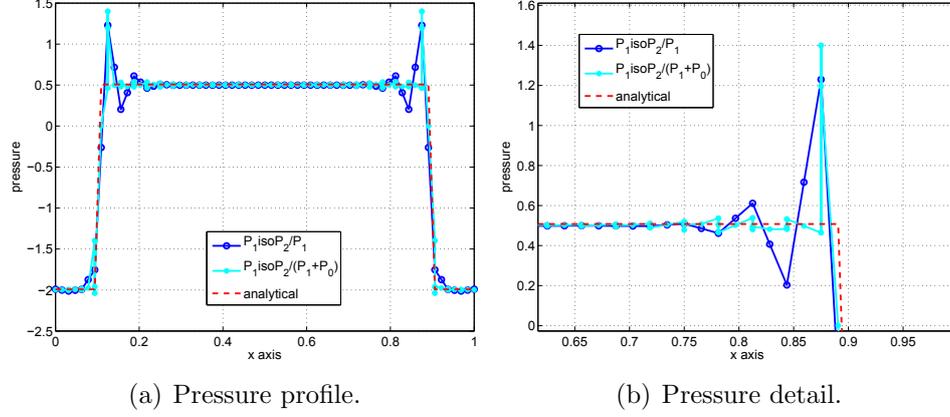


Figure 3: Pressure profile for Hood-Taylor finite element.


Figure 4: Pressure profile for Bercovier-Pironneau finite element.

$P_1isoP_2/P_1 + P_0$					
h_x	DOF	$\ p_h - p\ _{L^2}/\ p\ _{L^2}$	L^2 rate	$\ \mathbf{u}_h - \mathbf{u}\ _{L^2}$	L^2 rate
1/8	289	$2.796936 \cdot 10^{-1}$	-	$1.773827 \cdot 10^{-3}$	-
1/16	1084	$1.896726 \cdot 10^{-1}$	0.5603356	$9.032238 \cdot 10^{-4}$	0.9737103
1/32	4225	$1.349056 \cdot 10^{-1}$	0.4915617	$3.291174 \cdot 10^{-4}$	1.456481
1/64	16641	$1.021229 \cdot 10^{-1}$	0.4016440	$1.191034 \cdot 10^{-4}$	1.466388
P_1isoP_2/P_1					
h_x	DOF	$\ p_h - p\ _{L^2}/\ p\ _{L^2}$	L^2 rate	$\ \mathbf{u}_h - \mathbf{u}\ _{L^2}$	L^2 rate
1/8	289	$4.382848 \cdot 10^{-1}$	-	$9.247036 \cdot 10^{-3}$	-
1/16	1084	$3.055532 \cdot 10^{-1}$	0.5204451	$3.031251 \cdot 10^{-3}$	1.609078
1/32	4225	$2.296133 \cdot 10^{-1}$	0.4122172	$1.045616 \cdot 10^{-3}$	1.535560
1/64	16641	$1.715187 \cdot 10^{-1}$	0.4208407	$3.937636 \cdot 10^{-3}$	1.408951

Table 1: Spatial convergence to the reference stationary solution.

$P_2/P_1 + P_0$					
h_x	DOF	$\ p_h - p\ _{L^2}/\ p\ _{L^2}$	L^2 rate	$\ \mathbf{u}_h - \mathbf{u}\ _{L^2}$	L^2 rate
1/8	289	$2.679253 \cdot 10^{-1}$	-	$1.666593 \cdot 10^{-3}$	-
1/16	1084	$1.873964 \cdot 10^{-1}$	0.5157377	$9.165198 \cdot 10^{-4}$	0.8626639
1/32	4225	$1.323534 \cdot 10^{-1}$	0.5016974	$3.115031 \cdot 10^{-4}$	1.556920
1/64	16641	$9.766083 \cdot 10^{-2}$	0.4385436	$1.083690 \cdot 10^{-4}$	1.523295
P_2/P_1					
h_x	DOF	$\ p_h - p\ _{L^2}/\ p\ _{L^2}$	L^2 rate	$\ \mathbf{u}_h - \mathbf{u}\ _{L^2}$	L^2 rate
1/8	289	$3.015382 \cdot 10^{-1}$	-	$8.617810 \cdot 10^{-3}$	-
1/16	1084	$2.277048 \cdot 10^{-1}$	0.4051762	$3.043257 \cdot 10^{-3}$	1.501705
1/32	4225	$1.590665 \cdot 10^{-1}$	0.5175350	$1.030705 \cdot 10^{-3}$	1.561984
1/64	16641	$1.144642 \cdot 10^{-1}$	0.4747342	$3.776420 \cdot 10^{-4}$	1.448540

Table 2: Spatial convergence to the reference stationary solution.

5 CONCLUSIONS

At the beginning of this paper we recalled the immersed boundary method in its variational formulation as in [7] and [5]. The key idea is to split the Cauchy stress tensor into two components to be treated in Eulerian and in Lagrangian formulation for the fluid and the structure, respectively.

The main result of this paper is the study of the performances of two families of finite elements whose stability was proved in [1]. Being the solution for the pressure discontinuous, we augmented the pressure space with piecewise constant functions. In Section 3.1 we anticipated two improvements in augmenting the pressure space: better mass conservation, and better oscillations control at the interface. Numerical results show the accomplishment of both these goals. On the other hand we experienced a non-intuitive result for the “real divergence”. The *Hood–Taylor* scheme makes the whole IBM algorithm more diffusive than the *Bercovier–Pironneau* one. Considering the “real divergence” it is straightforward to notice that the latter enjoys a more “flexible” shape which is more effective in adjusting the “real divergence” to the structure inclusion (see Figure 2).

Finally, we can conclude that the behavior of both *augmented* elements (*Bercovier–Pironneau* and *Hood–Taylor*) is comparable.

REFERENCES

- [1] D. Boffi, N. Cavallini, F. Gardini, and L. Gastaldi. Local mass conservation of stokes finite elements. *Submitted*.
- [2] D. Boffi, N. Cavallini, and Gastaldi. Finite element approach to immersed boundary method with different fluid and solid densities. *Submitted*.
- [3] D. Boffi and L. Gastaldi. A finite element approach for the immersed boundary method. *Comput. & Structures*, 81(8-11):491–501, 2003. In honour of Klaus-Jürgen Bathe.
- [4] D. Boffi, L. Gastaldi, and L. Heltai. A finite element approach to the immersed boundary method. In Scotland Saxe-Coburg Publications, Stirling, editor, *Progress in Engineering Computational Technology*, B.H.V. Topping and C.A. Mota Soares Eds., pages 271–298, 2004.
- [5] D. Boffi, L. Gastaldi, and L. Heltai. Numerical stability of the finite element immersed boundary method. *Math. Models Methods Appl. Sci.*, 17(10):1479–1505, 2007.
- [6] D. Boffi, L. Gastaldi, and L. Heltai. On the CFL condition for the finite element immersed boundary method. *Comput. & Structures*, 85(11-14):775–783, 2007.
- [7] D. Boffi, L. Gastaldi, L. Heltai, and Charles S. Peskin. On the hyper-elastic formulation of the immersed boundary method. *Comput. Methods Appl. Mech. Engrg.*, 197(25-28):2210–2231, 2008.

- [8] R. Glowinski. Finite element methods for incompressible viscous flow. In *Handbook of numerical analysis, Vol. IX*, Handb. Numer. Anal., IX, pages 3–1176. North-Holland, Amsterdam, 2003.
- [9] B.E. Griffith. On the volume conservation of the immersed boundary method. *Submitted*.
- [10] Boyce E. Griffith, Richard D. Hornung, David M. McQueen, and Charles S. Peskin. An adaptive, formally second order accurate version of the immersed boundary method. *Journal of Computational Physics*, 223(1):10 – 49, 2007.
- [11] Boyce E. Griffith and Charles S. Peskin. On the order of accuracy of the immersed boundary method: Higher order convergence rates for sufficiently smooth problems. *Journal of Computational Physics*, 208(1):75 – 105, 2005.
- [12] G. Guidoboni, R. Glowinski, N. Cavallini, and S. Canic. Stable loosely-coupled-type algorithm for fluid-structure interaction in blood flow. *J. Comput. Phys.*, 228(18):6916–6937, 2009.
- [13] A. Huerta and W. K. Liu. Viscous flow with large free surface motion. *Comput. Methods Appl. Mech. Engrg.*, 69:277–324, 1988.
- [14] T. J. R. Hughes, W. K. Liu, and T. K. Zimmermann. Lagrangian-eulerian finite element formulation for incompressible viscous flows. *Comput. Methods Appl. Mech. Engrg.*, 29:329–349, 1981.
- [15] P. L. Le Tallec and J. Mouro. Fluid structure interaction with large structural displacements. *Comput. Methods Appl. Mech. Engrg.*, 190:3039–3067, 2001.
- [16] C. S. Peskin. The immersed boundary method. In *Acta Numerica 2002*. Cambridge University Press, 2002.
- [17] R. Pierre. Local mass conservation and C^0 -discretizations of the Stokes problem. *Houston J. Math.*, 20(1):115–127, 1994.
- [18] A. Quarteroni, M. Tuveri, and A. Veneziani. Computational vascular fluid dynamics: problems, models and methods. *Comput. Visual Sci.*, 2:163–197, 2000.
- [19] D. M. Tidd, R. W. Thatcher, and A. Kaye. The free surface flow of Newtonian and non-Newtonian fluids trapped by surface tension. *Internat. J. Numer. Methods Fluids*, 8(9):1011–1027, 1988.

A FINITE ELEMENT METHOD FOR NON-LINEAR HYPERELASTICITY APPLIED FOR THE SIMULATION OF OCTOPUS ARM MOTIONS

VASILEIOS VAVOURAKIS^{*,†}, ASIMINA KAZAKIDI[†], DIMITRIOS P.
TSAKIRIS[†] AND JOHN A. EKATERINARIS^{‡,§}

^{*}Department of Civil and Environmental Engineering
University of Cyprus, Nicosia 1678, Cyprus
e-mail: vavourakis.vasileios@ucy.ac.cy

[†]Institute of Computer Science, Foundation for Research and Technology-Hellas
Heraklion, Crete 71110, Greece

[‡]Department of Mechanical and Aerospace Engineering
University of Patras, Rio 26500, Greece

[§]Institute of Applied and Computational Mathematics, Foundation for Research and
Technology-Hellas
Heraklion, Crete 71110, Greece

Key words: FEM; skeletal muscles; muscular hydrostats

Abstract. An implicit non-linear finite element (FE) numerical procedure for the simulation of biological muscular tissues is presented. The method has been developed for studying the motion of muscular hydrostats, such as squid and octopus arms and its general framework is applicable to other muscular tissues. The FE framework considered is suitable for the dynamic numerical simulations of three-dimensional non-linear nearly incompressible hyperelastic materials that undergo large displacements and deformations. Human and animal muscles, consisting of fibers and connective tissues, belong to this class of materials. The stress distribution inside the muscular FE model is considered as the superposition of stresses along the muscular fibers and the connective tissues. The stresses along the fibers are modeled as the sum of active and passive stresses, according to the muscular model of Van Leeuwen and Kier (1997) *Philos. Trans. R. Soc. London*, 352: 551-571. Passive stress distribution is an experimentally-defined function of fibers' deformation; while active stress distribution is the product of an activation level time function, a force-stretch function and a force-stretch ratio function. The mechanical behavior of the surrounding tissues is determined adopting a Mooney-Rivlin constitutive model. The incompressibility criterion is met by enforcing large bulk modulus and by introducing modified deformation measures. Due to the non-linear nature of the problem,

approximate determination of the Jacobian matrix is performed, in order to utilize the full Newton-Raphson iterative procedure within each time-step. In addition, time discretization is performed via the implicit Newmark method. We developed an open-source finite element code that is capable of simulating large deflection maneuvers of muscular hydrostats. The proposed methodology is validated by comparing the numerical results with existing measurements for the squid arm extension. The efficiency and robustness of the proposed numerical method is demonstrated through a series of octopus arm maneuvers, such as extension, compression and bending.

1 INTRODUCTION

Muscular tissues are complex, non-linear, anisotropic, incompressible, viscoelastic materials undergoing large deformations. A wide class of muscular tissues deforms voluntarily and simulation of muscles behavior through the finite element method (FEM) has been the subject of many numerical investigations. Amongst the pioneering works in skeletal muscles' computational mechanics through the FEM, Kojic et al. [1] proposed a FEM numerical algorithm for the determination of muscle response. In this work, Hill's three-element model was used as basis for the mechanical description of fibers and accounted for non-linear force-displacement relation and change of geometrical shape. In the same year, Martins et al. [2] developed a three-dimensional FE methodology for the simulation of skeletal muscles, where the constitutive relation adopted is a generalization of the model proposed by [3], being compatible with the passive and active behavior of skeletal muscles of [4]. Johannsson et al. [5] proposed a mixed total Lagrangian FE formulation for simulating muscular behavior, based on non-linear continuum mechanics, where contractile active and passive properties of skeletal muscles were considered and the stress distribution was assumed to be equal to the superposition of passive and active ones. Yucesoy et al. [6] considered the skeletal muscle in two domains: the intracellular and extracellular matrix domain, represented by two separate meshes linked elastically to account for the trans-sarcolemmal attachments of muscle fibers, which allows force transmission between these domains. Furthermore, Oomens et al. [7] proposed a FE approach where physiological reasoning and "cross-bridge" kinetics via a two-state Huxley model was adopted, in order to examine the mechanical behavior of a tibialis anterior of a rat. In the recent work of Liang et al. [8], the governing equations for a muscle element, were based upon the approach of [9], and incorporated in a commercial general purpose explicit FE program. It was also, Martins et al. [10], in succession to [2], who introduced a multiplicative split of the fiber stretch into contractile and (series) elastic stretches, and they considered the simultaneous presence of the series elastic element, the dependence of the contractile stress on the strain rate and the activation level function. Röhrle and Pullan [11] presented a simulation framework of an anatomically realistic model of the human masseter muscles and associated bones, in order to investigate the dynamics of chewing. In the work of

Tang et al. [12], Hill's muscle theory coupled with fatigue was proposed to describe the mechanical behavior of skeletal muscles, where the force developed by a fatigued muscle was described by a muscle fatigue formula. Stojanovic et al. [13] proposed an extension of Hill's three-component model of [1], in order to take into account for different fiber types. They presented a model consisting of different type sarcomeres coupled in parallel with the connective tissues, where each sarcomere was modeled by one non-linear elastic element connected in series with one non-linear contractile element. Most recently, Tang et al. [14] presented a three-dimensional FEM for skeletal muscles that was developed to simulate their mechanical behavior during lengthening or shortening. The constitutive relation of the muscle was determined by using a strain energy approach and active contraction behavior of the muscle fiber was modeled through the Hill's three-element muscle concept. In addition, Lu et al. [15] developed a visco-hyperelastic model for skeletal muscle, where the constitutive relation was based on the definition of a Helmholtz free-energy function, while their model involves fourteen material parameters.

In the present work, motivated by the OCTOPUS project (FP7-231608) that aims to the development of octopus-like robotic arms, an implicit non-linear finite element numerical procedure has been developed for the accurate simulation of biological muscular tissue dynamic motion. In the next section, the constitutive equations adopted and the FE framework developed in this work are presented in detail. The numerical results obtained with the implementation of these constitutive equations are initially validated, and presentation and discussion of results representative of octopus arm motions concludes the third section. The conclusions of this work are summarized in the last section.

2 FINITE ELEMENT METHODOLOGY

Assume a three-dimensional non-linear, homogeneous elastic continuous medium and denote by ${}_0V$ the volume and by ${}_0S$ the surface in its unloaded state that undergoes large deformations, as shown in Fig. 1. The volume and bounding surface of the body in its current (deformed) state is tV and tS , respectively. The equilibrium equations for a solid, subject to finite deformations, are identical to those for small deformation analysis, i.e.

$$\frac{\partial \sigma_{ij}}{\partial {}^tx_j} = {}^t\rho \ddot{u}_i, \quad (1)$$

where body forces have been neglected and ${}^t\rho$ is the material density. Prescribed displacements are assumed on tS_U and prescribed tractions are assumed on the portion tS_T of the bounding surface to obtain a well-posed boundary value problem.

The stress equilibrium Eq. 1 is replaced by an equivalent principle of virtual work. Application of the variational theorem for finite elasticity in the current configuration yields [16]

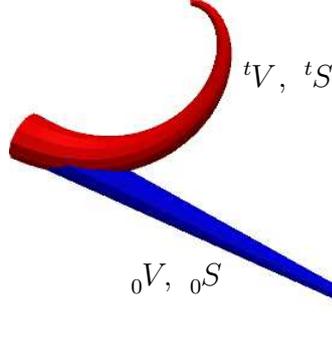


Figure 1: Deformed (red) and undeformed (blue) octopus arm configurations.

$$\delta\Pi = \int_{^tV} \delta v_i \ ^t\rho \ddot{u}_i \ dV + \int_{^tV} \delta L_{ij} \ \sigma_{ij} \ dV - \int_{^tS_T} \delta v_i \ \bar{t}_i \ dS = 0, \quad (2)$$

where \bar{t}_i is the prescribed traction vector on tS_T , δv_i being an admissible velocity variation that satisfies the condition: $\delta v_i = 0$ on tS_U , and $\delta L_{ij} = \partial \delta v_i / \partial ^t x_j$ the virtual velocity gradient matrix. The integrals in the virtual work relation are considered with respect to the current configuration. However, proper transformations can be applied in order to transform the integrals into the reference configuration, since the undeformed elastic body configuration is known. These transformations are the following: ${}_{0}\rho = J \ ^t\rho$, $d \ ^t x_i = F_{ij} d \ _0 x_j$, $d \ ^t S = J_\gamma d \ _0 S$ and $d \ ^t V = J d \ _0 V$, where $F_{ij} = \partial \ ^t x_i / \partial \ _0 x_j$ the deformation gradient tensor (expressed in terms of the initial and current coordinates of a material point ${}_{0}x_i$, ${}^t x_i$, respectively), J the deformation gradient determinant, $B_{ij} = F_{in} F_{jn}$ the left Cauchy-Green deformation tensor, and $J_\gamma = J \sqrt{\hat{n}_i B_{ij}^{-1} \hat{n}_j}$.

In the present work, muscles are assumed as a composite material, comprising of muscle fibers and connective tissues. Therefore, one can safely assume that the stress distribution inside the muscle is the superposition of the stress distribution in the connective tissues and the fibers, respectively, i.e. $\sigma_{ij} = \sigma_{ij}^{(ct)} + \sigma_{ij}^{(f)}$.

Inserting the aforementioned transformations into Eq. 2, the virtual work variational balance equation can be expressed into the reference configuration

$$\delta\Pi = \int_{_0V} \delta v_i \ _0\rho \ddot{u}_i \ J \ dV + \int_{_0V} \delta L_{ij} \ \sigma_{ij} \ J \ dV - \int_{_0S_T} \delta v_i \ \bar{t}_i \ J_\gamma \ dS = 0, \quad (3)$$

where ${}_{0}\rho$ the initial elastic body material density.

Finite element discretization of Eq. 3 using Lagrange polynomial shape functions $\delta v_i = \sum_{\alpha} N^{(\alpha)} \delta v_i^{(\alpha)}$, $\partial \delta v_i / \partial \ ^t x_j = \sum_{\alpha} (\partial N^{(\alpha)} / \partial \ ^t x_j) \delta v_i^{(\alpha)}$ yields:

$$\left\{ \int_{0V} N^{(\alpha)} {}_0\rho\ddot{u}_i J dV + \int_{0V} \frac{\partial N^{(\alpha)}}{\partial t x_j} \sigma_{ij} J dV - \int_{0S_T} N^{(\alpha)} \bar{t}_i J_\gamma dS \right\} \delta v_i^{(\alpha)} = 0 \Leftrightarrow$$

$$\left\{ \int_{0V} N^{(\alpha)} {}_0\rho\ddot{u}_i J dV + \int_{0V} \frac{\partial N^{(\alpha)}}{\partial_0 x_k} F_{kj}^{-1} \sigma_{ij} J dV - \int_{0S_T} N^{(\alpha)} \bar{t}_i J_\gamma dS \right\} \delta v_i^{(\alpha)} = 0. \quad (4)$$

This discrete, non-linear virtual work equation is solved using Newton-Raphson iteration assuming a corrected updated solution $u_i^{(\alpha)} + \Delta u_i^{(\alpha)}$, where $u_i^{(\alpha)}$ is the solution at the end of the previous time increment or the previous iteration solution. After linearization of the integral Eq. 4 and some algebra it can be obtained

$$\left\{ \int_{0V} \left[C_{ijkl} \frac{\partial N^{(\alpha)}}{\partial_0 x_r} F_{rj}^{-1} \frac{\partial N^{(\beta)}}{\partial_0 x_l} F_{lk}^{-1} - \sigma_{ij} \frac{\partial N^{(\alpha)}}{\partial_0 x_m} F_{mk}^{-1} \frac{\partial N^{(\beta)}}{\partial_0 x_n} F_{nj}^{-1} \right] J dV \right\} \Delta u_k^{(\beta)}$$

$$= \int_{0S_T} N^{(\alpha)} \bar{t}_i J_\gamma dS - \int_{0V} \sigma_{ij} \frac{\partial N^{(\alpha)}}{\partial_0 x_p} F_{pj}^{-1} J dV, \quad (5)$$

where C_{ijkl} is the tangent stiffness tensor given by

$$C_{ijkl} = \frac{\partial \sigma_{ij}}{\partial F_{ks}} F_{ls} + \sigma_{ij} \delta_{kl} = C_{ijkl}^{(ct)} + C_{ijkl}^{(f)}, \quad (6)$$

is the sum of the corresponding stiffnesses of the connective tissues and fibers contribution, respectively.

Numerical evaluation of the domain and boundary integrals of Eq. 5 yields a system of linear equations. In the present work, the Newmark implicit method is utilized as time-integration procedure. Given the displacement, velocity and acceleration from the previous time-step, the displacement increment is calculated through Eq. 5 and the updated displacement, velocity and acceleration components are evaluated accordingly [17].

At this point, it is necessary to address the constitutive description of the muscular finite element model. As stated above, the muscle consists of connective tissues with bio-fluids and muscular fibers. Their material description is provided in detail in the following subsections.

2.1 Muscle fibers material description

The fiber is the component that takes active role in the fiber and muscle contraction, resulting into muscle deformation. The fiber comprises of parallel bundles of myofibrils, which in turn are divided longitudinally by the Z-discs into sarcomeres. Sarcomeres are the basic contractile units of a muscle and they are also responsible for the passive properties of the muscle [18].

Assume a direction vector along the fiber, denoted with \hat{n}_i in the undeformed configuration, which is explicitly defined, and denote with $\hat{m}_i = 1/\lambda (F_{ij}\hat{n}_j)$ the direction vector in the deformed (current) configuration, where $\lambda = \sqrt{\hat{n}_i F_{ki} F_{kj} \hat{n}_j}$ the fiber stretch ratio. The nominal fiber strain ε_0^m is defined by the change of length divided by the reference length of the fiber, i.e. $\varepsilon_0^m = \lambda - 1$. Therefore, the corresponding volume preserving fiber strain tensor can be written as $\varepsilon_{ij}^{(f)} = \frac{1}{2} \varepsilon^m (3\hat{m}_i \hat{m}_j + \delta_{ij})$. The corresponding Cauchy stress tensor has the form

$$\sigma_{ij}^{(f)} = \sigma^m \hat{m}_i \hat{m}_j, \quad (7)$$

where the nominal axial stress σ_0^m in the fiber is defined in terms of the Cauchy true fiber stress as: $\sigma^m = \sigma_0^m (\varepsilon_0^m + 1)$, given that the volume of the fiber is preserved.

According to the approach of [19], the fiber nominal axial stress in muscles is defined as the accumulation of passive $\sigma^{(pass)}$ and active axial stress. The latter part is considered as the product of the maximum isometric stress at fiber optimum length $\sigma^{(max)}$, a normalized active state function $f^{(a)}$, a force-length function $f^{(l)}$ and a force-velocity dependent function $f^{(v)}$

$$\sigma_0^m = \sigma^{(pass)} + \sigma^{(max)} f^{(a)} f^{(l)} f^{(v)}. \quad (8)$$

The activation state $f^{(a)}$ describes the activation pattern and is a function of time. The filamentary overlap function $f^{(l)}$ describes the dependence of active stress on the nominal fiber strain ε_0^m , and the force-velocity function $f^{(v)}$ is a rate-dependent function that relates the active muscle stress and the nominal fiber strain rate $\dot{\varepsilon}_0^m$.

Inserting the above stress tensor relations in Eq. 6, one can analytically evaluate the tangent stiffness tensor as follows

$$\begin{aligned} C_{ijkl}^{(f)} = & \left[\lambda \frac{\partial \sigma_0^m}{\partial \varepsilon_0^m} (1 + \varepsilon_0^m) + \lambda \sigma_0^m - 2\sigma_0^m (1 + \varepsilon_0^m) \right] \hat{m}_i \hat{m}_j \hat{m}_k \hat{m}_l \\ & + \sigma_0^m (1 + \varepsilon_0^m) [\delta_{ik} \hat{m}_j \hat{m}_l + \hat{m}_i \delta_{jk} \hat{m}_l + \hat{m}_i \hat{m}_j \delta_{kl}]. \end{aligned} \quad (9)$$

2.2 Connective tissues material description

In the present work the connective tissues are described through a hyperelastic Mooney-Rivlin constitutive relation, where a modified stored energy function \bar{W} is used in order to evaluate stresses, provided by the generalized relation: $\bar{W} = c_1(\bar{I}_1 - 3) + c_2(\bar{I}_2 - 3) + K/2(J - 1)^2$, where c_1, c_2, K are material constants. The modified invariants of the left Cauchy-Green deformation tensor are introduced in \bar{W} : $\bar{I}_1 = I_1 J^{-2/3}$ and $\bar{I}_2 = I_2 J^{-4/3}$ based on the modified deformation gradient tensor $\bar{F}_{ij} = F_{ij} J^{-1/3}$, due to the nearly or fully incompressible behavior of biological tissues the constraint $J \approx 1$ must be satisfied.

The calculation of the Cauchy stress tensor can be obtained by differentiation of the stored energy density function with respect to deformation as

$$\begin{aligned} \sigma_{ij}^{(ct)} = & \frac{2}{J^{5/3}} \left(\frac{\partial \bar{W}}{\partial \bar{I}_1} + \bar{I}_1 \frac{\partial \bar{W}}{\partial \bar{I}_2} \right) B_{ij} - \frac{2}{J^{7/3}} \frac{\partial \bar{W}}{\partial \bar{I}_2} B_{im} B_{mj} \\ & - \frac{2}{3J} \left(\bar{I}_1 \frac{\partial \bar{W}}{\partial \bar{I}_1} + 2\bar{I}_2 \frac{\partial \bar{W}}{\partial \bar{I}_2} \right) \delta_{ij} + \frac{\partial \bar{W}}{\partial J} \delta_{ij}, \end{aligned} \quad (10)$$

with the Cauchy stress tensor consisting of the sum of a purely isochoric contribution and a purely volumetric one. Substituting the modified generalized Mooney-Rivlin constitutive material relation of \bar{W} in Eq. 10, the Cauchy stress tensor for the connective tissues is

$$\sigma_{ij}^{(ct)} = 2(c_1 + c_2 \bar{I}_1) J^{-5/3} B_{ij} - 2c_2 J^{-7/3} B_{im} B_{mj} - [2/3J (c_1 \bar{I}_1 + 2c_2 \bar{I}_2) + K(1 - J)] \delta_{ij}, \quad (11)$$

where for the present analysis the material constant K represents the bulk modulus of elasticity, while c_1 is equal to the shear modulus and $c_2 = 0$.

The corresponding tangent stiffness tensor of the connective tissues can be evaluated analytically through Eq. 6 and is provided below

$$\begin{aligned} C_{ijkl}^{(ct)} = & [4/9J (c_1 \bar{I}_1 + 4c_2 \bar{I}_2) + K(2J - 1)] \delta_{ij} \delta_{kl} + 2c_2 J^{-7/3} (2B_{ij} B_{kl} - B_{il} B_{jk} - B_{ik} B_{jl}) \\ & - 4/3 (c_1 + 2c_2 \bar{I}_1) J^{-5/3} (B_{ij} \delta_{kl} + \delta_{ij} B_{kl}) + 8/3 c_2 J^{-7/3} (B_{ij}^2 \delta_{kl} + \delta_{ij} B_{kl}^2) \\ & + 2(c_1 + c_2 \bar{I}_1) J^{-5/3} (\delta_{ik} B_{jl} + B_{il} \delta_{jk}) - 2c_2 J^{-7/3} (\delta_{ik} B_{jl}^2 + B_{il}^2 \delta_{jk}), \end{aligned} \quad (12)$$

where $B_{ij}^2 = B_{ik} B_{kj}$.

3 NUMERICAL EXAMPLES

In order to validate the finite element methodology, the squid arm extension during the strike to catch prey is simulated. The squid arm is modeled as a simplified cylindrical geometry, consisting of an active stalk and a passive club. The stalk consists of transverse muscles inside the arm and an outer layer consists of longitudinal muscles. Detailed description of the squid musculature is provided by Van Leeuwen and Kier [9], while the material properties of the muscle fibers and surrounding tissues are identical to the ones used by previous FE approaches [5, 8].

Due to symmetry, only one quadrant of the cylindrical arm is modeled. Symmetry boundary conditions are applied on the symmetry planes and the outer surface is considered traction-free. The discretized quadrant consists of 246 four-node hexahedral and 41 six-node triangular base prismatic elements, as seen in Fig. 3(a). The applied activation signal for the current simulation is a step function [9] having a 40 msec activation time for maximum activation level, while the total time duration of the simulation is 100 msec.

In Fig. 3(a) a comparison of the squid arm-length growth in time is presented. The FE numerical results, obtained by the proposed methodology (diamonds), are compared

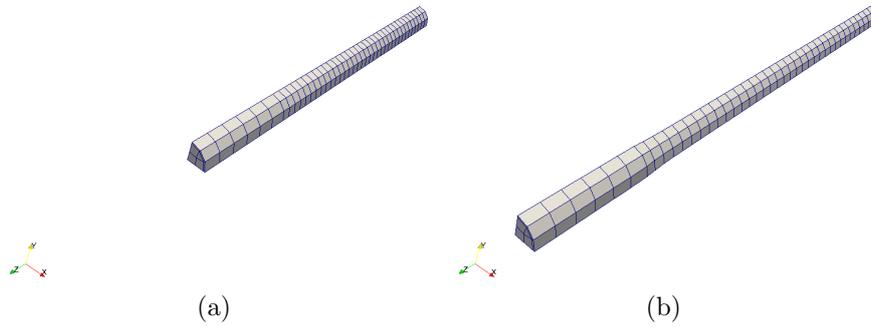


Figure 2: (a) Undeformed and (b) final deformed squid arm.

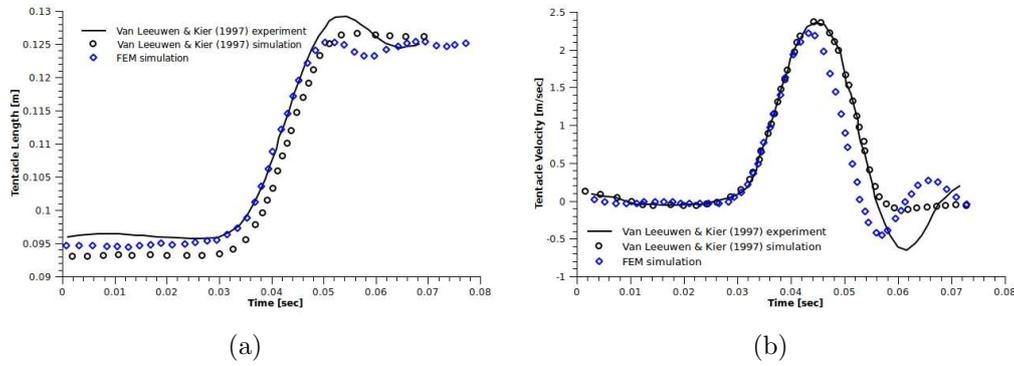


Figure 3: Comparison of experimental and numerical results for the squid arm extension: (a) tentacle length and (b) arm-tip velocity.

with the experimental data (solid line) of the squid arm extension and the corresponding simulations (circles) obtained by Van Leeuwen and Kier [9]. It is observed that the overall agreement, both qualitative and quantitative, of the present FEM numerical results with the experimental measurements is very good. In Fig. 3(b), it can be noticed that a relatively lower arm-tip velocity is predicated through the proposed FE analysis. However, similar observations were made by other investigators [5, 8, 14], who used the same values of the material parameters.

Next, a conical geometry resembling an octopus arm is considered. The arm is 10 cm long, extending along the z axis, and has 1 cm root diameter, as seen in Fig. 4(b). The arrangement of muscles in the octopus muscular hydrostat is very different to that of the squid, and is depicted in Fig. 4(a). The musculature consists mainly of four groups of longitudinal muscles that extend along the arm and the transversal muscles. In addition, oblique muscles are present, which have helically aligned fibers around the arm, and the central axis of the arm is occupied by the axial nerve cord. Due to the lack of experimental data for the octopus muscular hydrostat, the same material parameters utilized for the squid arm simulation are taken for the octopus arm as well.

The finite element mesh of the octopus arm consists of 420 four-node hexahedral and

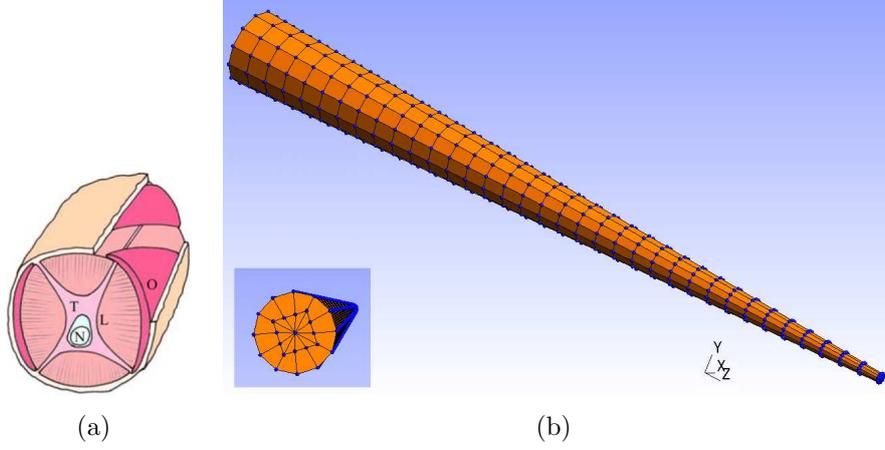


Figure 4: (a) Octopus muscular hydrostat structure: longitudinal muscles (L), transverse muscles (T), oblique muscles (O), axial nerve cord (N), and (b) finite element discretized model of the octopus arm

420 six-node triangular base prismatic elements (see Fig. 4(b)). The core of the truncated conical geometry contains the transverse muscles and the rest of the domain contains the longitudinal muscles. Oblique muscles are neglected in the present analysis because they have minor contribution to bending motion of octopus arms and they are hard to incorporate in robotic arm models. The root of the arm is allowed to move on the x - y plane and is fixed at the origin point $(0,0,0)$, while the rest of the boundary is taken traction-free. The applied activation signal for the current simulation is a step function, as follows

$$f^{(a)} = \begin{cases} 0, & t \leq t_i \\ \left[\frac{1}{2} (1 + \sin(\pi t/t_a - \pi/2)) \right]^{3.5}, & t \leq t_i + t_a \\ 1, & t \leq t_i + t_d \\ 0, & t > t_i + t_d \end{cases}, \quad (13)$$

In Eq. 13 the activation time is set equal to $t_a = 0.5$ sec, while the total simulation duration is one second. Furthermore, t_i and t_d are the initialization and duration time of the activation function.

As seen in the previous example, (Fig. 3(a)) the squid hydrostat can perform an extension maneuver if all transverse muscles are activated simultaneously. In order for the octopus arm to perform a bending or/and reaching move, primarily longitudinal muscles have to be activated. Initially, it is assumed that one group of longitudinal muscles is activated uniformly ($t_i = 0$); then it is assumed that the same muscle is activated non-uniformly ($t_i = \bar{z}/0.2$, $t_i = \bar{z}/0.6$ and $t_i = \bar{z}$), given the normalized axial position \bar{z} of a material point within the muscle. The time duration of the activation level is equal to $t_d = 1$ sec.

In Figs. 5(a) it is shown how the octopus arm deforms when one longitudinal mus-

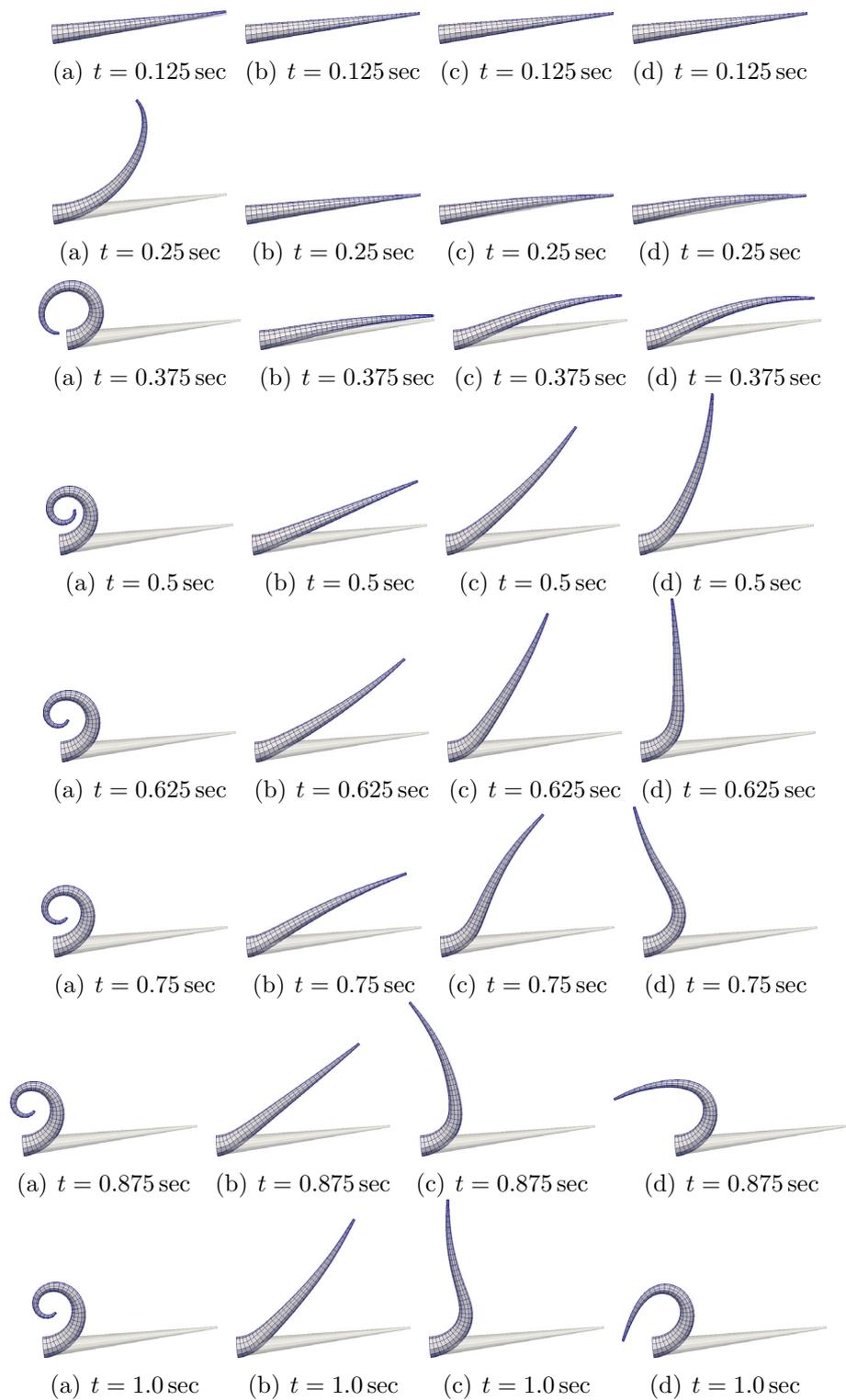


Figure 5: Various snapshots of the octopus arm deformation when uniform and non-uniform activation of one longitudinal muscle occurs.

cle activates uniformly. For this uniform activation case it is evident how the bending maneuver is achieved within 0.4sec approximately. On the other hand, in Figs. 5(b) it is shown how the octopus arm deforms when one longitudinal muscle activates non-uniformly ($t_i = \bar{z}/0.2$) and in Figs. 5(c) the same longitudinal muscle activates non-uniformly ($t_i = \bar{z}/0.6$) faster than in the previous simulation example, and in Figs. 5(d) a longitudinal muscle activates even faster ($t_i = \bar{z}$). Comparing the corresponding snapshots of Fig. 5 one could observe that the bending maneuver is performed more rapidly when uniform activation occurs than in non-uniform activation. Similar arm deformation occurs when two neighboring longitudinal muscles are activated in the same manner (not presented herein), but the final bend is stronger in that case, since two muscles contract instead of one. It is interesting to note here that one can experiment to various activation levels and active muscles, thus, producing various octopus arm deformation regimes. However, it is expected that more impressive simulations can be produced when the oblique muscles are set active that reflect to torsional arm deformation regimes, which is left for future work.

4 CONCLUSIONS

An implicit non-linear FE numerical procedure for the simulation of biological muscular tissues is presented. The proposed methodology is applied to study the behavior of the octopus muscular hydrostat. Within the FEM framework, it is considered a three-dimensional non-linear nearly incompressible hyperelastic fibrous material that undergoes large displacements and deformations. Due to the material and geometrical non-linearity to this problem, the Newton-Raphson technique is utilized and analytical evaluation of the Jacobian matrices is presented. Time discretization of the dynamic problem is done with the aid of the Newmark implicit technique. The proposed FEM is validated by comparing the numerical results with the corresponding experimental ones for the squid arm extension during the strike to catch prey. Furthermore, it has been demonstrated that the FEM can successfully simulate more complex arm motions, such as the octopus arm bending at various activation conditions. However, the presented numerical approach can also be applied to other muscle simulation problems, such as elephant trunks, snakes and skeletal muscles, with various muscle groups and with external loading. Future development to the existing methodology will account for the oblique muscles of the octopus arm, where explicit description of the fibers orientation will be provided, and initial stress distribution will be considered in order for simulations to run when the arm is at deformed initial state.

ACKNOWLEDGEMENTS

The financial support for this work provided by the project OCTOPUS (IP) project with number: FP7-231608 is gratefully acknowledged.

REFERENCES

- [1] Kojic M, Mijailovic S, Zdravkovic N. Modelling of muscle behaviour by the finite element method using Hill's three-element model. *International Journal for Numerical Methods in Engineering* 1998; **43**:941–953.
- [2] Martins JAC, Pires EB, Salvado R, Dinis PD. A numerical model of passive and active behavior of skeletal muscles. *Computer Methods in Applied Mechanics and Engineering* 1998; **151**:419–433.
- [3] Humphrey JD, Yin FCP. On constitutive relations and finite deformations of passive cardiac tissue: I. A pseudostrain-energy function. *Journal of Biomechanical Engineering* 1987; **109**:298–304.
- [4] Zajac FE. Muscle and tendon: Properties, models, scaling and application to biomechanics and motor control. *CRC Critical Reviews in Biomedical Engineering* 1989; **17**:359–411.
- [5] Johansson T, Meier M, Blickhan R. A Finite-Element Model for the Mechanical Analysis of Skeletal Muscles. *Journal of Theoretical Biology* 2000; **206**:131–149.
- [6] Yucesoy CA, Koopman BHFJM, Huijing PA, Grootenboer HJ. Three-dimensional finite element modeling of skeletal muscle using a two-domain approach: linked fiber-matrix mesh model. *Journal of Theoretical Biology* 2002; **35**:1253–1262.
- [7] Oomens CWJ, Maenhout M, Van Oijen CH, Drost MR, Baaijens FP. Finite element modelling of contracting skeletal muscle. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2003; **358**:1453–1460.
- [8] Liang Y, McMeeking RM, Evans AG. A finite element simulation scheme for biological muscular hydrostats. *Journal of Theoretical Biology* 2006; **242**:142–150.
- [9] Van Leeuwen JL, Kier WM. Functional design of tentacles in squid: linking sarcomere ultrastructure to gross morphological dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences* 1997; **352**:551–571.
- [10] Martins JAC, Pato MPM, Pires EB. A finite element model of skeletal muscles. *Virtual and Physical Prototyping* 2006; **1**:159–170.
- [11] Röhrle O, Pullan AJ. Three-dimensional finite element modelling of muscle forces during mastication. *Journal of Biomechanics* 2007; **40**:3363–3372.
- [12] Tang CY, Tsui CP, Stojanovic B, Kojic M. Finite element modelling of skeletal muscles coupled with fatigue. *International Journal of Mechanical Sciences* 2007; **49**:1179–1191.

- [13] Stojanovic B, Kojic M, Rosic M, Tsui CP, Tang CY An extension of Hill's three-component model to include different fiber types in finite element modeling of muscle. *International Journal for Numerical Methods in Engineering* 2007; **71**:801–817.
- [14] Tang CY, Zhang G, Tsui CP. A 3D skeletal muscle model coupled with active contraction of muscle fibres and hyperelastic behaviour. *Journal of Biomechanics* 2009; **42**:865–872.
- [15] Lu YT, Zhu HX, Richmond S, Middleton J. A visco-hyperelastic model for skeletal muscle tissue under high strain rates. *Journal of Biomechanics* 2010; **43**:2629–2632.
- [16] Bowler AF. *Applied Mechanics of Solids*. CRC Press, 2009.
- [17] Bathe K-J. *Finite Element Procedures*. Prentice Hall, 1996.
- [18] Nigg BM, Herzog W. *Biomechanics of the musculo-skeletal system* (2nd edn). Wiley, 1999.
- [19] Van Leeuwen JL. Optimum power output and structural design of sarcomeres. *Journal of Theoretical Biology* 1991; **149**:229–256.

TOWARDS THE GENERIC CONCEPTUAL AND NUMERICAL FRAMEWORK FOR THE SIMULATION OF CO₂ SEQUESTRATION IN DIFFERENT TYPES OF GEORESERVOIRS

Uwe-Jens Görke*, Joshua Taron*, Ashok Kr. Singh*,
Wenqing Wang* and Olaf Kolditz*,[†]

* Department of Environmental Informatics
Helmholtz Centre for Environmental Research - UFZ
Permoserstr. 15, D-04318 Leipzig, Germany
e-mail: uwe-jens.goerke@ufz.de, www.ufz.de

[†] University of Technology Dresden, Germany

Key words: Porous Media Theory, Multiphase Flow, Deformation Processes, Coupled Problems, Finite Element Method, Carbon Capture and Storage

Abstract. In this paper, conceptual and numerical modeling of coupled thermo-hydro-mechanical (THM) processes during CO₂ injection and storage is presented. The commonly used averaging procedure combining the Theory of Mixtures and the Concept of Volume Fractions serves as background for the complex porous media approach presented here. Numerical models are based on a generalized formulation of the individual and overall balance equations for mass and momentum, as well as, in non-isothermal case, the energy balance equation. Within the framework of a standard Galerkin approach, the method of weighted residuals is applied to derive the weak forms of governing equations. After discretizing spatially these weak forms, a system of nonlinear algebraic equations can be obtained. For the required time discretization a generalized first order difference scheme is applied, linearization is performed using Picard or Newton-Raphson methods. The corresponding models are implemented within the scientific open source finite element code OpenGeoSys (OGS) developed by the authors, which is based on object oriented programming concepts. This assists the efficient treatment of different physical processes, whose mathematical models are of similar structure. Thus, the paper is mainly focused on a generic theoretical framework for the coupled processes under consideration. Within this context, CO₂ sequestration in georeservoirs of different type can be simulated (e.g., saline aquifers, (nearly) depleted hydrocarbon reservoirs).

1 INTRODUCTION

Carbon Capture and Storage (CCS) is counted among the most promising transition technologies for the mitigation of anthropogenic greenhouse gas emissions. Recently, CO₂ sequestration as an essential part of CCS procedures has become a subject of worldwide investigation.

Three types of geological formations are particularly considered for the safe storage of CO₂: (nearly) depleted hydrocarbon reservoirs, deep saline aquifers and unminable coal seams. The migration of the carbon dioxide in the subsurface, and its interaction with the formation fluids as well as with the porous reservoir media is characterized by various complex transport, reaction and deformation phenomena. Several trapping mechanisms (e.g., stratigraphic, structural, solubility, mineral effects), which prevent the migration of the buoyant CO₂ back to the surface, are based on these phenomena. Within the context of mechanical loading, the injection of carbon dioxide into the subsurface results in high pressure in the vicinity of the injection well. Due to the injection pressure, the stress distribution in this reservoir region can be changed significantly. To consider high pressure induced medium deformation is very important for the integrity of potential geological storage sites. Fluid and matrix properties of the porous subsurface media are also affected by prevailing thermal conditions. In particular, high temperature differences between injected fluid and reservoir environment have the potential of possible reservoir failure (e.g., thermal matrix damage).

The coupling of multiple physical-chemical processes requires numerical analysis based on physically grounded complex mathematical models. Within this context, the modeling and simulation of the injection and the spreading of carbon dioxide in the underground is essential for the proper understanding of the physical and chemical processes at different length and time scales, to ascertain migration and trapping of CO₂ in the porous formations, and in assessing the capacity as well as the safety of the reservoir. Sophisticated mathematical models, numerical algorithms and software tools have to be developed taking into account all relevant physico-chemical phenomena during migration and storage of CO₂ in the subsurface, such as flow and transport of multiple phases including dissolution and mixing effects, hydrodynamic instabilities (e.g. viscous fingering), rock deformation as well as fracturing, heat transport and phase changes. In most cases it is not necessary to be able to describe all these processes for the whole simulation time period or for the entire model domain. The formulation of specific model concepts has to consider the spatial and temporal scale of the problem and the dominating processes in each case.

While transport and deformation processes in porous media have been studied for several decades considering various applications (e.g. groundwater flow and consolidation problem), the modeling of carbon dioxide migration in geological formations is a fairly new subject of investigations in different research areas (e.g. hydrology, geotechnology, computational mechanics and mathematics). As about ten years ago only a few publications have been dedicated to this topic (cf. [1, 2]), recently an increasing number of

numerical (see e.g., [3–7]), semi-analytical (cf. [8, 9]) and analytical (cf. [10–12]) studies have been published. Within the context of numerical simulations, more complex problems can be treated (e.g. coupling of different processes, consideration of heterogeneities and various geological conditions), although the development of efficient and stable algorithms is challenging. Among hydromechanical site specific simulations of CO₂ injection, hypothetical cases are reported (e.g., [13], saline aquifer) as well as real field studies (e.g., [14], depleted hydrocarbon reservoir). For the state of the art in model and software development for geological CO₂ storage see also [2, 15] and literature cited there. To compare several numerical simulators with respect to their capabilities, efficiency and accuracy code comparison studies have been conducted [16, 17]. Within this context, benchmark problems and real site studies have been defined addressing various aspects of CO₂ storage at different reservoir conditions.

Modeling of CO₂ storage on a reservoir scale for feasibility studies and risk analyses is very demanding with respect to the computational costs due to the complex geometry that needs to be described and due to the diversity of interacting hydraulic, thermal, mechanical, and geochemical processes. According to the evolving processes at various time and length scales, we propose a successive model development with the final goal of a comprehensive coupled simulation of all relevant physico-chemical effects, but starting with the most relevant effects at CO₂ injection near the injection well. Since the model complexity is adapted to these dominating processes, other processes having less contribution to the event under consideration are neglected for computational efficiency (e.g., mineral trapping for short-term processes). Within this context, in this paper, the investigation is restricted to thermo-hydro-mechanical (THM) effects of the CO₂ injection into different types of reservoirs.

In this study, we present the generalized conceptual model, the governing equations and numerical aspects of a finite element approach to analyze THM coupled processes at geological CO₂ storage. To this purpose, we focus on basic ideas and relations, not discussing here specific formulations of governing equations, which are related to typical processes in different storage sites (e.g., multiphase flow in deformable porous media in the case of saline aquifers, non-isothermal compositional gas flow for enhanced gas recovery in gas reservoirs). With the present assumption, the changes in thermal, flow and deformation fields of the model are simulated by using the standard Galerkin finite element method, which is realized in an object oriented scientific tool, OpenGeoSys (an open source finite element code) developed by the authors. Specific model formulations, simulation results and discussion are presented in two complementary conference papers.

In the following, vectors and higher order tensors will be denoted by boldface characters. Their scalar product is characterized by a single dot while double dots indicate the summation product (double inner product) $\mathbf{a} \cdot \cdot \mathbf{b} = a_{ij} b_{ji}$. A superposed dot indicates the material rate of a vector or a tensor. The individual constituents α of a porous material represent the phases of the overall aggregate or components within a phase. Below, superscript $\alpha = s$ marks the solid phase, and $\alpha = \gamma$ denote several pore fluid phases.

2 CONCEPTUAL MODELING

2.1 Preliminary Remarks

Based on Biot's studies on the macroscopic theory of saturated wet soils (cf. [18]), the Theory of Mixtures has been widely accepted to model the complex behavior of porous media (concerning basic assumptions see e.g., [19, 20]). For the realistic modeling of the mutual interactions of the porous media constituents, the Theory of Mixtures has been combined with the Concept of Volume Fractions by, e.g., [21, 22]. Within the context of this enhanced Theory of Mixtures (also known as Theory of Porous Media), all kinematical and physical quantities can be considered at the macroscale as local statistical averages of their values at the underlying microscale. Comprehensive studies about the theoretical foundation and numerical algorithms for the simulation of coupled problems of multiphase continua are given in, e.g., [23–25] and the quotations therein. Recent developments of enhanced numerical approaches for the porous media modeling in geotechnical applications are reported in, e.g., [26–28].

In the following, the subsurface formation designated as carbon dioxide reservoir is considered as a mixture of a solid skeleton and a pore fluid content, which can either be

- a single liquid or gas (single-phase flow in porous media),
- an immiscible fluid mixture of gas and liquids (multiphase flow in porous media) or
- a miscible fluid mixture of different reacting constituents allowing phase transitions due to evaporation, condensation, precipitation (multiphase multicomponent flow in porous media).

Within the framework of the Concept of Volume Fractions, various scalar variables are defined to describe the microstructure of a porous medium in a macroscopic manner neglecting the real topology and distribution of the pores. These variables serve as measures of local fractions of the individual constituents. In this context, the volume fractions n^α represent the ratio of the partial volume of the corresponding constituent of a multiphase body with respect to the overall volume of a representative elementary volume (REV) of the control domain under consideration. One of the most characteristic media properties of a porous material is the porosity n , the local amount of fluid volume fractions.

$$n = \sum_{\gamma} n^{\gamma} = 1 - n^s \quad (1)$$

Since, in general, the overall medium is completely filled with matter, from the definition of the volume fractions follows the saturation condition regarding the overall aggregate.

$$\sum_{\alpha} n^{\alpha} = 1 \quad (2)$$

If multiphase flow occurs, it is more convenient for various applications to use the (partial) fluid saturations S^γ instead of the volume fractions. These local functions are given by

$$S^\gamma = \frac{n^\gamma}{n} \quad \text{with the pore content saturation condition} \quad \sum_{\gamma} S^\gamma = 1 \quad (3)$$

In the following, two different formulations of mass density related to the constituents of a porous medium are introduced. The so-called material (effective, realistic) density $\rho^{\alpha R}$ is defined as the ratio of the mass fraction dm^α of the individual constituent under consideration with respect to its partial volume fraction dv^α .

$$\rho^{\alpha R} = \frac{dm^\alpha}{dv^\alpha} \quad (4)$$

The so-called partial (global, bulk) density is given by the ratio of the mass fraction of the given constituent with respect to the volume fraction of the overall aggregate dv .

$$\rho^\alpha = \frac{dm^\alpha}{dv} \quad \text{with the correlation} \quad \rho^\alpha = n^\alpha \rho^{\alpha R} \quad (5)$$

Obviously, the mass density of the porous medium (homogenized overall aggregate) is defined as the sum of the partial densities of its constituents.

$$\rho = \sum_{\alpha} \rho^\alpha \quad (6)$$

The conceptual idea behind the formulations and relations presented above consists in the assumption that the mass fractions of all constituents of the multiphase medium are simultaneously present and statistically uniformly distributed over the entire control domain. Within this context, the material body under consideration is theoretically substituted by an aggregate completely and continuously filled by superimposed (overlapping) homogenized partial continua with reduced mass densities. Consequently, the motion and physics of the individual constituents as well as the overall aggregate can be specified by well-accepted phenomenological methods of continuum mechanics.

Describing the transport and deformation of the constituents of porous media within the framework of continuum mechanics it is assumed that the geometry of the control domain under consideration is characterized at each time by the solid skeleton, whereas the fluid pore content is able to flow across the boundary of the surface. Thus, it proves to be reasonable not to model the absolute motion state of the pore content, but its motion relative to the motion of the solid phase, considering the porous medium as a local thermodynamic open system with the solid skeleton as volume under observation.

2.2 Governing Equations

The governing field equations for the modeling of transport and deformation processes of non-isothermal multiphase flow in deformable porous media are formulated based on the local individual balance relations of the constituents, particularly mass, momentum and

energy balances. Within this context, the application of the fundamental balance relations on the analysis of multiphase materials is based on Truesdell's *metaphysical principles* [29]. The crucial idea behind these principles is the assumption that the balance relations of the constituents as well as the balance relations of the overall aggregate of a porous medium can be formulated in accordance to the corresponding classical relations of single-phase continuum mechanics. Additionally, to account for the interaction mechanisms between the constituents, so-called production terms are introduced for the individual balance relations of the constituents (cf. [24]). Following, superposition is used to define the balance relations of the overall aggregate based on the individual balance relations of the constituents.

Neglecting mass exchange between the phases (no dissolution and sorption processes), the local mass balance for any individual constituent of the porous medium is given by

$$\frac{d_\alpha \rho^\alpha}{dt} + \rho^\alpha \nabla \cdot \mathbf{v}^\alpha = \frac{\partial \rho^\alpha}{\partial t} + \nabla \cdot (\rho^\alpha \mathbf{v}^\alpha) = 0 \quad (7)$$

with the velocity \mathbf{v}^α of the constituent under consideration, and the usual divergence operator $\nabla \cdot ()$. From the velocity-displacement relation for the solid skeleton follows $\mathbf{v}^s = \dot{\mathbf{u}}^s$ with the solid displacement vector \mathbf{u}^s . The derivative

$$\frac{d_\alpha a}{dt} = \frac{\partial a}{\partial t} + \mathbf{v}^\alpha \cdot \nabla a \quad (8)$$

with the usual gradient operator $\nabla()$ denotes the material time derivative of an arbitrary variable a with respect to the motion of a material point of the constituent. For the further discussion, the relations between the material time derivatives (here, of an arbitrary variable a) with respect to the solid skeleton, and with respect to an individual fluid constituent is of crucial interest in terms of a unified numerical characterization of the different processes.

$$\frac{d_\gamma a}{dt} = \frac{d_s a}{dt} + \mathbf{v}^{\gamma s} \cdot \nabla a \quad (9)$$

Here, $\mathbf{v}^{\gamma s} = \mathbf{v}^\gamma - \dot{\mathbf{u}}^s$ is the so-called seepage velocity describing the fluid motion with respect to the deforming skeleton material.

According to the generalized formulation Eq. (7), considering Eqs. (1) and (5), the local solid phase mass balance is given by

$$\frac{d_s [(1-n)\rho^{sR}]}{dt} + (1-n)\rho^{sR} \nabla \cdot \dot{\mathbf{u}}^s = 0 \quad (10)$$

serving as basis for the calculation of porosity changes. Following the same procedure, additionally considering Eqs. (3) and (9), the mass balance relations for fluid constituents can be defined with respect to the solid phase motion.

$$\frac{d_s (nS^\gamma \rho^{\gamma R})}{dt} + \nabla \cdot (nS^\gamma \rho^{\gamma R} \mathbf{v}^{\gamma s}) + nS^\gamma \rho^{\gamma R} \nabla \cdot \dot{\mathbf{u}}^s = 0 \quad (11)$$

Assuming material incompressibility of the solid phase, i.e. $d_s \rho^{sR}/dt=0$, and applying the solid phase mass balance Eq. (10), Eq. (11) can be represented in a more detailed description.

$$nS^\gamma \frac{d_s \rho^{\gamma R}}{dt} + n\rho^{\gamma R} \frac{d_s S^\gamma}{dt} + \nabla \cdot (\rho^{\gamma R} \mathbf{w}^{\gamma s}) + S^\gamma \rho^{\gamma R} \nabla \cdot \dot{\mathbf{u}}^s = 0 \quad (12)$$

Here $\mathbf{w}^{\gamma s} = nS^\gamma \mathbf{v}^{\gamma s}$ is usually known as filter velocity of the motion of the pore fluid constituent φ^γ . For non-isothermal problems, the saturations S^γ and the mass densities $\rho^{\gamma R}$ represent functions of temperature T and the fluid pressure fractions p^γ

Deformation processes in porous media are described by the momentum balance equation in terms of the total Cauchy stress tensor $\boldsymbol{\sigma}$ referring to the local loading state of the overall aggregate. In geotechnical problems, the internal fluid friction forces can be neglected in comparison to the interaction terms between fluid and skeleton motions. Thus, $\boldsymbol{\sigma}$ is given by the sum of all partial stresses of the constituents. Consequently, the stress tensor is defined according to the well-known effective stress concept [30]:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_E^s - p\mathbf{I} \quad (13)$$

with the solid effective stress tensor $\boldsymbol{\sigma}_E^s$ and the identity tensor \mathbf{I} . Therein, the pore pressure p is given in analogy to Dalton's law.

$$p = \sum_{\gamma} S^\gamma p^\gamma \quad (14)$$

Finally, in non-isothermal case, the specific local overall linear momentum balance equation for fluid saturated porous media is defined as follows:

$$\nabla \cdot \left[\boldsymbol{\sigma}_E^s - \left(\sum_{\gamma} S^\gamma p^\gamma \right) \mathbf{I} - \alpha \Delta T \left(\underset{\mathbf{D}}{\mathbf{D}} \cdot \cdot \mathbf{I} \right) \right] + \rho \mathbf{g} = \mathbf{0} \quad (15)$$

where α is the thermal expansion coefficient, $\underset{\mathbf{D}}{\mathbf{D}}$ is the fourth-order elastic material tensor, and $\rho \mathbf{g}$ is the volume force with the gravity vector \mathbf{g} .

The heat transport equations for individual constituents of a porous medium are based on their local partial energy balance equations.

$$\rho^\alpha C_{p\alpha} \frac{d_\alpha T^\alpha}{dt} - \nabla \cdot (\lambda_\alpha \nabla T^\alpha) + Q_\alpha = 0 \quad (16)$$

Here, $C_{p\alpha}$ denotes the heat capacity and λ_α the thermal conductivity of the given constituent, respectively, T^α is the individual temperature of the constituent, and Q_α characterize source/sink terms. Assuming that the constituents are locally in a thermodynamical equilibrium state, i.e., locally all phases are at the same temperature T at each material point of the homogenized continuum, and using Eq. (9), the set of necessary balance

equations according to the generalized problem under consideration is completed by the heat transport equation for the overall aggregate.

$$(\rho C_p)_{\text{eff}} \frac{d_s T}{dt} + \sum_{\gamma} (C_{p\gamma} n S^{\gamma} \rho^{\gamma R} \mathbf{v}^{\gamma s}) \cdot \nabla T - \nabla \cdot (\lambda_{\text{eff}} \nabla T) + Q_T = 0 \quad (17)$$

Considering the above mentioned Truesdell's metaphysical principles, the effective heat capacity and the thermal conductivity of the overall porous medium are calculated according to the following equations, which allows for the coupling of thermal and hydraulic properties.

$$(\rho C_p)_{\text{eff}} = (1 - n) \rho^{sR} C_{ps} + n \sum_{\gamma} C_{p\gamma} S^{\gamma} \rho^{\gamma R} \quad (18)$$

$$\lambda_{\text{eff}} = \sum_{\alpha} \lambda_{\alpha} \quad (19)$$

Based on specific material functions for the state variables and appropriate formulations of different production terms to characterize the interaction between the constituents of the overall aggregate (e.g., phase transition, mixing), specific formulations for the above mentioned balance equations have to be defined considering the site specific situation.

2.3 Constitutive Relations

If the balance relations characterize fundamental physical and thermodynamical properties of the matter independently of specific material properties, in real applications the response of a physical body on similar interactions with the external environment differs for various materials. Thus, so-called constitutive relations have to be defined to characterize the specific material behavior. In terms of the mathematical modeling of physico-chemical processes this observation is equivalent to the formulation of closed systems of equations, which should consist of balance as well as constitutive relations.

Within the context of the multiphase problem under consideration, constitutive equations are required for selected production terms of the specific balance relations of the individual constituents, as mentioned above, for pore fluid properties like pressure and saturation, for the partial effective stress tensor of the solid skeleton, and for several thermal properties. Among others, specific formulations of Darcy's, Fick's and Fourier's law, equations of state, capillary pressure-saturation functions as well as the relations between relative permeability and saturation, and deformation laws relating the effective stress to the strain and/or strain rate of the solid matrix are substantial constitutive equations required for porous media mechanics.

3 NUMERICAL SCHEME

The numerical treatment of the coupled problem of THM processes in deformable porous media is based on the governing field equations together with discretization meth-

ods in the space and time domains. Some general representations of corresponding numerical approaches to solve the problem under consideration can be found in [25, 31–33].

Usually, the method of weighted residuals is applied to derive the weak formulations of all the governing equations given above. Within the framework of a standard Galerkin procedure, the corresponding local individual or overall balance equations are multiplied by arbitrary test functions, which are defined in appropriate spaces and meet the corresponding homogeneous Dirichlet boundary conditions, and are integrated subsequently over the domain of interest bounded by the solid skeleton.

As usual, in the finite element space the continuous functions of the selected primary variables are interpolated based on their nodal values and appropriately defined shape functions. After discretizing spatially the weak forms of the balance relations a system of nonlinear algebraic equations can be obtained. For the required time discretization we use a generalized first order difference scheme, but in most applications, a fully implicit Euler scheme is used, which is unconditionally stable and can be applied to complex coupled problems. However, numerical diffusion is introduced into the system by time discretization. This problem is reduced by an adaptive time-stepping scheme.

The nonlinear coupled boundary value problem is solved iteratively using the Picard or Newton-Raphson linearization. Within this context, usually all unknowns can be solved at the same time in the context of a so called monolithic scheme. However, solving the whole system of equations monolithically may lead to memory problems when employing a fine mesh due to a large number of freedoms per element node of the coupled problem under consideration. We adopt a mixture of monolithic and staggered schemes to avoid this bottleneck without losing the accuracy of the solutions.

The discussed numerical scheme is realized in an object oriented scientific software tool, OpenGeoSys, developed by the authors, and has been verified by several classic benchmarks. The presented generalized model has been adopted to several site specific applications (e.g., two-phase flow in deformable porous media in the case of saline aquifers, compositional gas flow for enhanced gas recovery in gas reservoirs - for details see [34, 35])

4 CONCLUSIONS

We have presented the conceptual model and the governing equations for the simulation of coupled thermo-hydro-mechanical processes, which are observed for geological storage of carbon dioxide in porous subsurface materials. The presented generalized numerical model is based on balance laws for the overall porous medium as well as for its different fluid and solid phases. Using well-established discretization procedures (in space and time), a coupled multiphysics finite element model is developed from the weak formulations of the governing equations. Monolithic or staggered monolithic/staggered coupling schemes are available to solve the discretized system of equations. The presented generalized approach represents the basis for specific formulations of numerical models for the simulation of different application oriented situations within the context of geological CO₂ storage (e.g., saline aquifers, nearly depleted hydrocarbon reservoirs). For this purpose, the appropriate

choice of relevant primary variables and constitutive relations has a high priority, and is subject of intense scientific studies. This indicates that the analysis of the thermo-hydro-mechanical response is crucial for proper selection of injection sites, injection rates, and total storage capacity of potential CO₂ reservoirs.

5 ACKNOWLEDGEMENTS

This study is funded by the German Federal Ministry of Education and Research (BMBF) as part of joint projects in the framework of the Special Programme GEOTECHNOLOGIEN.

REFERENCES

- [1] Lindeberg, E. Escape of CO₂ from aquifers. *Energy Convers. Mgmt.* (1997) **38**:235–240.
- [2] Pruess, K. and García, J. Multiphase flow dynamics during CO₂ injection into saline aquifers. *Environ. Geol.* (2002) **42**:282–295.
- [3] Basbug, B., Gumrah, F. and Oz, B. Simulating the effects of deep saline aquifer properties for CO₂ sequestration. *J. Can. Petrol. Technol.* (2007) **46**:30–38.
- [4] Ennis-King, J. and Paterson, L. Role of convective mixing in the long-term storage of carbon dioxide in deep saline formations. *SPE J.* (2005) **10**:349–356.
- [5] Ghanbari, S., Al-Zaabi, Y., Pickup, G.E., Mackay, E., Gozalpour, F. and Todd, A.C. Simulation of CO₂ storage in saline aquifers. *Chem. Eng. Res. Des.* (2006) **84**(9):764–775.
- [6] Pruess, K., Xu, T.F., Apps, J. and García, J. Numerical modeling of aquifer disposal of CO₂. *SPE J.* (2003) **8**(1):49–60.
- [7] Vilarrasa, V., Bolster, D., Olivella, S. and Carrera, J. Coupled hydromechanical modeling of CO₂ sequestration in deep saline aquifers. *Int. J. Greenh. Gas Control* (2010) **4**:910-919
- [8] Nordbotten, J.M. and Celia, M.A. Similarity solutions for fluid injection into confined aquifers. *J. Fluid Mech.* (2006) **561**:307–327.
- [9] Saripalli, P. and McGrail, P. Semi-analytical approaches to modeling deep well injection of CO₂ for geological sequestration. *Energy Convers. Mgmt.* (2002) **43**(2):185–198.
- [10] Nordbotten, J.M., Celia, M.A. and Bachu, S. Injection and storage of CO₂ in deep saline aquifers: Analytical solution for CO₂ plume evolution during injection. *Transp. Porous Med.* (2005) **58**:339–360.

- [11] Dentz, M. and Tartakovsky, D.M. Abrupt-interface solution for carbon dioxide injection into porous media. *Transp. Porous Med.* (2008) **51**(7):1–13.
- [12] Riaz, A., Hesse, M., Tchelepi, H.A. and Orr, F.M. Jr. Onset of convection in a gravitationally unstable diffusive boundary layer in porous media. *J. Fluid Mech.* (2006) **548**:87–111.
- [13] Rutqvist, J. and Tsang, C.-F. A study of caprock hydromechanical changes associated with CO₂ injection into a brine formation. *Environ. Geol.* (2002) **42**:296–305.
- [14] Vidal-Gilbert, S., Nauroy, J.-F. and Brosse, E. 3D geomechanical modelling for CO₂ geologic storage in the Dogger carbonates of the Paris Basin. *Int. J. Greenhouse Gas Control* (2009) **3**(3):288–299.
- [15] Metz, B., Davidson, O., de Cominck, H.C., Loos, M. and Meyer L.A. (Eds.) *IPCC special report on carbon dioxide capture and storage*. Prepared by Working Group III of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA (2005).
- [16] Class, H., Ebigbo, A., Helmig, R., Dahle, H.K., Nordbotten, J.M., Celia, M.A., Audigane, P., Darcis, M., Ennis-King, J., Fan, Y., Flemisch, B., Gasda, S.E., Jin, M., Krug, St., Labregere, D., Beni, A.N., Pawar, R.J., Sbai, A., Thomas, S.G., Trenty, L. and Wei, L. A benchmark study on problems related to CO₂ storage in geologic formations: Summary and discussion of the results. *Comput. Geosci.* (2009) **13**(4):409–434.
- [17] Pruess, K., García, J., Kovscek, T., Oldenburg, C., Rutqvist, J., Steefel, C. and Xu, T. Code intercomparison builds confidence in numerical simulation models for geologic disposal of CO₂. *Energy* (2004) **29**:1431–1444.
- [18] Biot, M.A. General theory of three-dimensional consolidation. *J. Appl. Phys.* (1941) **12**:155–164.
- [19] Bowen, R.M. *Theory of mixture*. In: Continuum Physics, Vol. III, Eringen, A.C. (ed.), Academic Press, New York, pp. 1–127 (1976).
- [20] Truesdell, C. and Toupin, R.A. *The classical field theories*. In: Handbuch der Physik, Vol. III/1, Flügge, S. (ed.), Springer, Berlin, pp. 226–902 (1960).
- [21] de Boer, R. and Ehlers, W. On the problem of fluid- and gas-filled elasto-plastic solids. *Int. J. Sol. Struct.* (1986) **22**:1231–1242.
- [22] Prevost, P. Mechanics of continuous porous media. *Int. J. Eng. Sci.* (1980) **18**:787–800.

- [23] de Boer, R. *Theory of porous media: Highlights in historical development and current state*. Springer, Berlin (2000).
- [24] Ehlers, W. and Bluhm, J. *Porous media: Theory, experiments and numerical applications*. Springer, Berlin (2002).
- [25] Lewis, R.W. and Schrefler, B.A. *The finite element method in the static and dynamic deformation and consolidation of porous media*. 2nd Edition. Wiley, New York (1998).
- [26] Kolditz, O. and De Jonge, J. Non-isothermal two-phase flow in low-permeable porous media. *Comput. Mech.* (2004) **33**(5):345–364.
- [27] Sanavia, L., Pesavento, F. and Schrefler, B.A. Finite element analysis of non-isothermal multiphase geomaterials with application to strain localisation simulation. *Comput. Mech.* (2006) **37**:331–348.
- [28] Truty, A. and Zimmermann, T. Stabilized mixed finite element formulations for materially nonlinear partially saturated two-phase media. *Comput. Meth. Appl. Mech. Engrg.* (2006) **195**:1517–1546.
- [29] Truesdell, C. *Thermodynamics of diffusion*. In: Rational Thermodynamics, Truesdell, C. (ed.), 2nd edition, Springer, New York, pp. 219–236 (1984).
- [30] Bishop, A.W. The principle of effective stress. *Teknisk Ukeblad* (1959) **39**:859–863.
- [31] Ehlers, W., Ellsiepen, P. and Ammann, M. Time- and space-adaptive methods applied to localization phenomena in empty and saturated micropolar and standard porous materials. *Int. J. Numer. Methods Eng.* (2001) **52**:503–526.
- [32] Helmig, R. *Multiphase flow and transport processes in the subsurface: A contribution to the modelling of hydrosystems*, Springer, Berlin (1997).
- [33] Korsawe, J., Starke, G., Wang, W. and Kolditz, O. Finite element analysis of poro-elastic consolidation in porous media: Standard and mixed approaches. *Comput. Methods Appl. Mech. Engrg.* (2006) **195**:1096–1115.
- [34] Görke, U.-J., Park, C.-H., Wang, W., Singh, A.K. and Kolditz, O. Numerical Simulation of Multiphase Hydromechanical Processes Induced by CO₂ Injection in Deep Saline Aquifers. *Oil and Gas Science and Technology – Revu. IFP Energies nouvelles* (2011), online first: DOI:10.2516/ogst/2010032.
- [35] Singh, A.K., Böttcher, N., Wang, W., Park, C.-H., Görke, U.-J. and Kolditz, O. *Non-isothermal effects on multi-phase flow in porous medium: CO₂ disposal into a saline aquifer*. In: Proceedings of the International Conference on Greenhouse Gas Control Technologies (GHGT10), Energy Procedia (in print).

A COUPLED MODEL OF MECHANICAL BEHAVIOUR AND WATER RETENTION FOR UNSATURATED SOILS WITH DOUBLE POROSITY

AZAD KOLIJI, LYESSE LALOU AND LAURANT VULLIET

Ecole Polytechnique Fédérale de Lausanne (EPFL), Laboratory for soil mechanics
EPFL-ENAC-LMS, Station 18, CH-1015 Lausanne, Switzerland, <http://lms.epfl.ch>
e-mail: azad.koliji@epfl.ch, lyesse.laloui@epfl.ch, laurent.vulliet@epfl.ch

Key words: Unsaturated soil, constitutive modeling, hydro-mechanical coupling, soil structure; double porosity, plasticity.

Abstract. Many natural soils and engineering geomaterials, such as aggregated soils and compacted clay pallets, exhibit two levels of porosity corresponding to the inter- and intra-aggregate pores within their hierarchical structure. Mechanical behavior of these materials, in particular when unsaturated, is an issue of added complexity which should be described an appropriate constitutive framework. A coupled water retention–mechanical constitutive model for unsaturated soils with double porosity is presented here. Based on the multi-scale experimental results, the model incorporates the inter-particle bonding, fabric and partial saturation effects in a single framework. It is formulated within the framework of hardening elasto-plasticity and is based on the critical state concept. The mechanical model is coupled with the water retention law which itself takes into account the two levels of porosity. The coupling is made through the expression of the effective stress and the evolution of the preconsolidation pressure with suction. On the other hand, the mechanical model at the macro-scale is also coupled with the pore-scale behavior of the materials through an internal variable which accounts for the evolution of the soil structure. The model is used for numerical simulation of the behavior of aggregated and bonded soils. Comparison of numerical simulations and the experimental results show that the model can successfully address the main features in the behavior of unsaturated soils with double porosity.

1 INTRODUCTION

Constitutive modeling of unsaturated soils is a subject of increasing interest owing to its importance in a large number of geotechnical engineering problems. Although the early developments were focusing on a homogeneous continuum concept and a single porosity, further evaluation of structured porous media, such as aggregated soils, fissured clays and fractured rocks, revealed the necessity of considering a structure with inter-particle bonding and at least two distinct values of porosity linked to macro- and micropores. Hence, the concept of double porosity [1,2] usually applies for these materials.

Improvements of constitutive models for natural bonded soils have been proposed by introducing a dependency of the size of the yield limit on the inter-particle bonding [3-9].

These models often ignore the soil fabric effects and use ad hoc relations to describe the soil structure evolution. Moreover, these models are mainly developed for saturated soils and only few works have looked into the coupled effects of soil structure and partial saturation [10, 11].

When unsaturated, the pores are infiltrated by more than one fluid. Hence, beside the improvement of the mechanical constitutive model for capillary effects, an additional constitutive relation is required to describe the liquid retention of the pores. Such a relation becomes of particular importance when the constitutive stress includes a combined function of suction and degree of saturation, as in the models of Bolzon et al. [12] and Wheeler et al. [13]. It is, therefore, important to introduce a coupling between the mechanical and the water retention constitutive relations.

In the present paper, a mechanical constitutive model coupled with the water retention relation is proposed for unsaturated soils with double porosity. It incorporates the inter-particle bonding, fabric and partial saturation effects in a single framework. Also, it allows a two-way coupling between water retention and mechanical behavior. The model development is based on the multi-scale experimental results of unsaturated aggregated soils presented in [14, 15] and that is briefly reviewed in the next section. The two parts of the model corresponding to the mechanical and the water retention parts are then presented in subsequent sections. At the end, the typical response of the model is presented and the numerical simulations are compared the model simulation with the experimental results for unsaturated aggregated and saturated bonded soils.

2 COUPLED EFFECTS OF SUCTION AND SOIL STRUCTRE

The stress-strain behavior of unsaturated aggregated soils at the macro-scale has been studied using suction-controlled oedometer testing method [14]. Results of that study revealed three main class of effects: (i) soil structure effects describing the difference between reconstituted soil (with ‘intrinsic behavior’, after [16]) and aggregated soil at the same suction., (ii) intrinsic suction effects describing reconstituted soil at different suction values, and finally (iii) the coupled effects of soil structure and suction describing unsaturated aggregated soils at different suctions and their comparison to reconstituted soil results.

For the soil structure effects, the experimental results showed an apparent preconsolidation stress in aggregated soils, which depends not only on stress state and stress history, but also on the soil structure. Exhibiting an initial apparent overconsolidated state, aggregated soil appeared to have a normal consolidation line (NCL) located to the right side of that of reconstituted soil at the same suction. It was shown that the two curves tend to converge at higher stresses where the aggregated structure is obliterated (see in Figure 2).

Looking into the intrinsic suction effects, the main feature was found to be the increase of the effective apparent preconsolidation stress increases with suction. The term ‘effective’ will be further discussed in the mechanical constitutive framework in section 3.

Finally, as for the main coupled effect of suction and soil structure, experimental results showed that although the effective (apparent) preconsolidation stress increases with suction in both aggregated and reconstituted soils, the rate of this increase is higher for aggregated soil. In Figure 1, this behavior can be thought of as augmentation of the horizontal separation of the two curves as suction increases.

In addition to the above-mentioned macroscopic results, the pore-scale behavior of unsaturated aggregated soil has been also studied using the combination of mercury intrusion porosimetry, environmental electron scanning microscopy, and neutron computed tomography method [15]. Based on the obtained results, it has been suggested that the structural phenomenon of particle aggregation is the main cause of the macropore formation. As such, the macropores represent the actual state of the soil structure. Any degradation of structures due to mechanical loading or humidity variation results in closure of these extra pores, and brings the state of the soil closer to that of a reconstituted soil. As an important experimental finding, it was evidenced in neutron tomography tests that the change in the volume fraction of macropores is mainly associated with plastic deformations [17]. Those experimental evidences allow quantifying the soil structure in terms of macroporosity as a physical parameter and relate it to the plastic deformations as a state parameter [18,19].

3 MECHANICAL CONSTITUTIVE MODEL

3.1 Modeling framework

The proposed mechanical model, named ACMEG-2S, lies within the framework of hardening elastoplasticity and is based on the critical state concept [20]. It is originally built upon the model of Hujieux [21], and extends the ACMEG model of saturated reconstituted soils [22] to unsaturated structured state. The detailed mathematical formulation of the model can be found in [19]. The model considers the saturated reconstituted state as a reference state to which the suction and structure effects will be added.

The model uses the generalized effective stress defined here as the Bishop's effective stress [23] with the Bishop's coefficient being the degree of saturation [24, 25]:

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} - p_a \mathbf{I} - S_r s \mathbf{I} \quad (1)$$

where $\boldsymbol{\sigma}'$ is the effective stress tensor, $\boldsymbol{\sigma}$ the total Cauchy stress tensor, S_r the degree of saturation, $s = p_a - p_w$ the matric suction with p_a and p_w being, respectively, the air and water pressures, and finally \mathbf{I} is the identity tensor. Adopting the above constitutive stress provides a straightforward transition from saturated to unsaturated state owing to the uniqueness of the yield surface, constitutive stress, and the critical state line at the two states [26].

3.2 Model description

The model uses non-linear elasticity for the revisable part. Based on the experimental results, we assume that the elastic parameters are not affected neither by soil structure nor by partial saturation and are equal to those of the corresponding saturated reconstituted soil.

The plastic response in the model is governed by two plastic mechanisms, isotropic and deviatoric plastic mechanism. Following the concept of multi-mechanism plasticity [27], the total plastic strain increment is induced by the two corresponding dissipative processes. The two-invariant yield functions for the isotropic and the deviatoric mechanisms are:

$$f_{iso} = p' - p'_c r_{iso} = 0 \quad (2)$$

$$f_{dev} = q - Mp' \left(1 - b \ln \frac{p'd}{p'_c} \right) r_{dev} = 0 \quad (3)$$

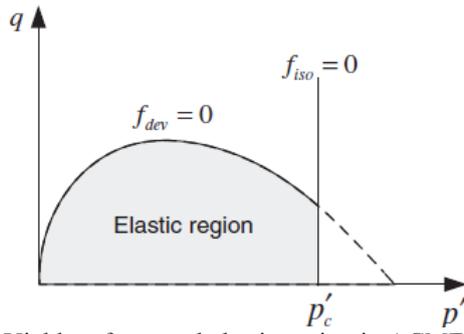


Figure 1 : Yield surfaces and elastic region in ACMEG-2S model

In these equations, p'_c is the apparent effective preconsolidation pressure accounting for the coupled effects of effects of suction and structure. It controls the size of the set union of the elastic regions given by the two yield functions in the stress plane, as depicted in Figure 1. The required material parameters are M , b , d , and those involved r_{iso} and r_{dev} .

Parameter M is the slope of the critical state line in the effective stress plane $q - p'$. This slope is neither affected by partial saturation [26] nor is influenced by the soil structure effects [6, among others]. Parameter b is a material parameter affecting the shape of the deviatoric yield surface. Parameter d , occasionally referred to as the spacing ratio [28], is the ratio of the saturated preconsolidation pressure over the saturated effective critical state pressure in the same yield surface for the soil (d fixed at 2.718 and 2.0 is the original and modified Cam-clay models, respectively). The two variables r_{iso} and r_{dev} represent the degree of mobilization of isotropic and deviatoric plastic mechanisms, respectively, and allow a smooth transition from the elastic to plastic domain without abrupt change in the rate of deformation.

Postulating an identical shape of yield function for unsaturated aggregated and saturated reconstituted soil, parameters b and d are identical to their values in the reference model. Moreover, it is plausible for the mobilization process of plastic mechanisms to be independent of the suction and soil structure and to be governed directly by the provoked plastic strain. This implies no change in r_{iso} and r_{dev} compared with their reference value for reconstituted saturated soils. Accordingly, parameters of the yield limits are directly inherited from the intrinsic values of in the reference model. The model also assumes adopts the plastic flow rule of reference model for reconstituted soil which an associated the isotropic and non-associated for the deviatoric plastic mechanism.

3.3 Apparent preconsolidation pressure

The apparent effective preconsolidation pressure used in the expression of the yield functions (Eqs. 2 & 3) controls the size of the yield limits and it is the main element for taking into account the combined effects of suction and soil structure. The model uses an isotropic hardening rule that allows the change in size but not in the shape of the yield surface and, indeed, describes the evolution of the apparent effective preconsolidation pressure. We introduce the expression of the apparent effective preconsolidation pressure in the form [19]:

$$p'_c = \psi^{st} \psi^s p'_{c0} \quad (4)$$

where p'_{c0} is the reference effective preconsolidation pressure in saturated reconstituted soil, ψ^s is a function of suction accounting for the intrinsic primary suction-induced hardening

effects as in reconstituted soils, and the function ψ^{st} introduces the soil structure effects including the mere soil structure effects and those coupled with suction.

The evolution of the reference effective preconsolidation pressure is governed by the intrinsic strain hardening (or softening) rule of the reference model, which is in the present case a volumetric plastic strain hardening rule of Cam-clay type. For the primary suction effects, the suction-induced hardening relation of reconstituted soil proposed by Nuth and Laloui [29] has been extended for the case of aggregated soils as

$$\psi^s = \begin{cases} 1 & ; 0 < s < s_e^1 \\ 1 + \gamma'_s \text{Log}(s/s_e^1) & ; s_e^1 \leq s < s_{ref} \\ 1 + \gamma_s \text{Log}(s/s_e) & ; s \geq s_{ref} \end{cases} \quad (5)$$

where s_e and s_e^1 are the air entry value suction of the corresponding reconstituted soil and that of the micropores, respectively, γ_s and γ'_s are material parameters which are correlated through $\gamma'_s = \gamma_s \text{Ln}(s_{ref}/s_e^1)/\text{Ln}(s_{ref}/s_e)$, and $s_{ref} > s_e$ is an arbitrary reference suction.

At the next step, the soil structure effects are introduced by recall the definition of the degree of soil structure R , as a physical parameter, which is the ratio of current macrovoid ratio, e^2 , over its initial value, e_i^2 at the intact state (superscript 2 for macrovoids) [18]:

$$R = e^2/e_i^2 \quad (6)$$

The degree of soil structure is, indeed, an internal scaling parameter, which equals unity for an intact aggregated soil in the presence of macropores and zero for a fully destructured soil in their absence. Any structure degradation, irrespective of its cause, might alter the soil structure, the macropores, and consequently the degree of soil structure. This variable is then linked, as a state parameter, to the plastic strain through an equation of the form [18]

$$R = \exp(-\omega \varepsilon^D) \quad (7)$$

where ω is a material parameter controlling the rate of structure degradation with plastic deformation, and ε^D , referred to as destructuring strain, is an invariant of volumetric, deviatoric, or a combination of both plastic strain tensor [17,19].

Using Equation 6 and the compression plane of Figure 5, we can now derive the expression of ψ^{st} . Knowing that the physical definition of R implies $AA''/BB'' \approx R$, one can deduce the following expression for ψ^{st} [18]

$$\psi^{st} = (\psi_i^{st})^R \quad (8)$$

The soil structure function ψ^{st} should also account for the coupled effects of suction with the soil structure. As mentioned previously, the initial apparent preconsolidation pressure increases with suction at a higher rate in aggregated soil rather than in reconstituted soil. This corresponds to an increase in the horizontal distance between the two compression curves of aggregated and reconstituted soils in Figure 2; hence, a higher ψ^{st} as suction increases. Based on the experimental evidences [14], the following reversible function is proposed:

$$\psi^{st} = \psi_{iref}^{st} \left(\frac{s + p_{at}}{s_{ref} + p_{at}} \right)^{n_{st}} ; \psi_{iref}^{st} \neq 1 \quad (9)$$

where p_{at} is the atmospheric pressure, ψ_{iref}^{st} is the initial reference value of the function for structured soil at the reference suction s_{ref} , and the exponent n_{st} is a material parameter.

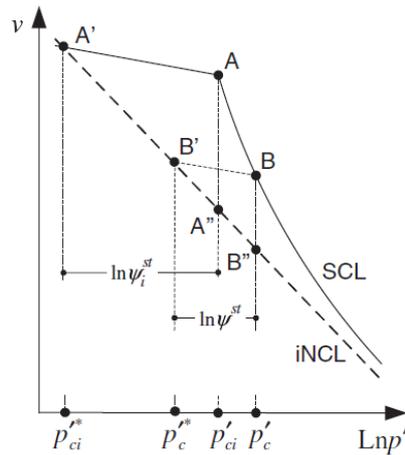


Figure 2 : Normal consolidation curves of aggregated (solid line) and reconstituted (dotted line) soils

The other coupling effect is related to the influence of suction on the rate of structure degradation, i.e. on parameter ω in Eq. 7. This parameter is also allowed to vary with suction to take into account the more brittle yielding of aggregated soils at higher suctions [19].

The coupled effects of suction on the apparent preconsolidation pressure in aggregated (structured) soils are illustrated in Figure 3. The abscissa is the ratio of apparent preconsolidation pressure over the saturated preconsolidation pressure in reconstituted state, p'_c/p_{c0}^* . The increase of apparent preconsolidation pressure due to intrinsic suction effect, $\Delta\psi_1$, given by Eq. 5, is represented by curve *a*. Multiplication of this curve with a reference soil structure function ψ_{ref}^{st} gives the curve *b*, which represents the increase in the apparent preconsolidation pressure due to intrinsic suction, $\Delta\psi_1$, and pure soil structure effects, $\Delta\psi_2$, without considering the suction-hardening of soil structure. Accounting for this latter effect by Eq. 9, the final evolution of apparent preconsolidation pressure with suction in aggregated soils is represented by curve *c*. The gray area between curves *b* and *c* in Figure 6, hence, corresponds to the gain in the apparent effective preconsolidation pressure due to the suction effects on the soil structure, $\Delta\psi_3$.

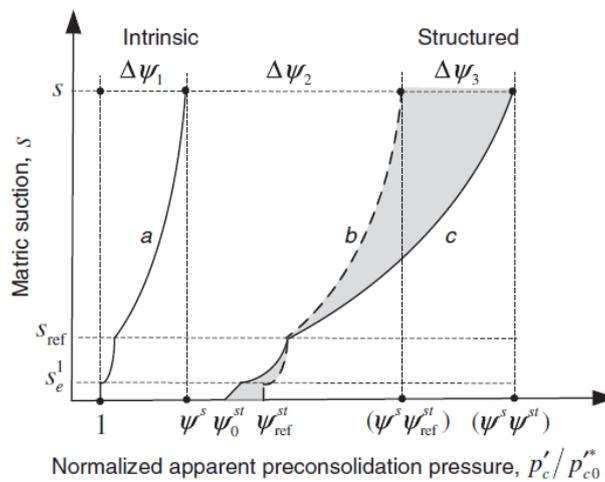


Figure 3 : Combined effects of suction and soil structure on the apparent isotropic preconsolidation pressure

3.4 General model formulation

The vectors of yield functions, plastic potentials, and non-negative plastic multipliers for two mechanisms can be written as

$$(10)$$

where the plastic strain tensor is obtained from the flow rule as

$$d\boldsymbol{\varepsilon}^p = \lambda^p \cdot \mathbf{g}, \quad \mathbf{g} = \frac{\partial \mathbf{G}}{\partial \boldsymbol{\sigma}'} \quad (11)$$

The plastic multipliers are determined using the Prager consistency equation for multi-dissipative plasticity [27] and satisfy the usual Kuhn–Trucker conditions

$$d\mathbf{F} = \mathbf{f} : d\boldsymbol{\sigma}' - (H^* + H^{st}) \cdot \lambda^p = 0, \quad \lambda^p \geq \mathbf{0} \text{ and } \mathbf{F} \leq \mathbf{0} \quad (12)$$

$$\mathbf{f} = \frac{\partial \mathbf{F}}{\partial \boldsymbol{\sigma}'}, \quad \mathbf{H}^* = -\frac{\partial \mathbf{F}}{\partial p_c^*} \cdot \frac{\partial p_c^*}{\partial \lambda^p}, \quad \mathbf{H}^{st} = -\frac{\partial \mathbf{F}}{\partial \psi^{st}} \cdot \frac{\partial \psi^{st}}{\partial \lambda^p} \quad (13)$$

In the above relations, \mathbf{H}^* is the generalized (primary suction effects included) hardening modulus corresponding to the reconstituted model, and \mathbf{H}^{st} is the generalized hardening modulus arising from soil structure effects. For a constant, $\mathbf{H}^{st} < 0$, i.e entries being negative, determines a softening due to structure degradation. However, to determine whether the material is in general hardening, softening, or showing a perfect plastic response, the sign of the total hardening modulus $\mathbf{H} = \mathbf{H}^* + \mathbf{H}^{st}$ should be considered.

Note that in the above form of consistency equation ds does not appear. This means that for any suction increments, all the suction-dependent variables are first updated with the new suction and then derivation is made with suction being held fixed. Hence, as proposed by Borja [30], there will be no return map on the suction during the numerical integrations.

Solving for the plastic multiplier yields

$$\mathbf{D}^{ep} = \mathbf{D}^e - \chi^{-1} \cdot (\mathbf{D}^e : \mathbf{g} \otimes \mathbf{f} : \mathbf{D}^e) \quad , \quad \chi = \mathbf{f} : \mathbf{D}^e : \mathbf{g} - \mathbf{H}^* + \mathbf{H}^{st} \quad (14)$$

where symbol \otimes denotes the dyadic product of two tensors, and \mathbf{D}^{ep} is the general elastoplastic constitutive tensor.

4 WATER RETENTION MODEL

The second part of the model describes the relation between suction and the degree of saturation. A number of well-established phenomenological water retention relations exist for non-deformable homogeneous porous media with single porosity [31, 32]. Such a relation, however, has been rarely investigated for porous media with multi-porosity [33, 34]. Total degree of saturation for double porous media in general is

$$S_r = \sum_{m=1,2} \Psi^m S^{lm} \quad (14)$$

where S^{lm} is the local degree of saturation of the micro ($m=1$) and macropores ($m=2$), and Ψ^m is the volumetric pore fraction of micro/macropores over the entire pores. Assuming two distinct air entry suction values for micro and macropores, the total water retention curve of double porous media exhibits at least two points of inflection as in Figure 4.

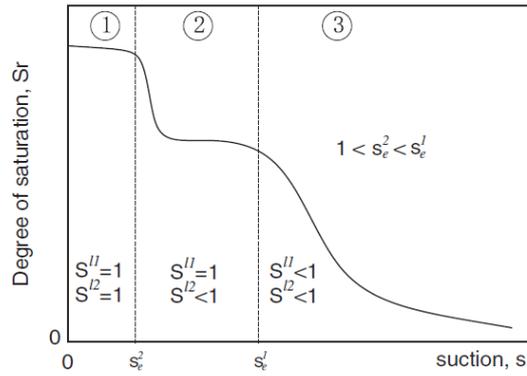


Figure 4 : Schematic water retention curve in soils with double porosity

Three different zones of saturation on this curve are zone 1 where both macro and micropores are saturated, zone 2 where micropores are saturated and macropores are unsaturated, and zone 3 where both macro and micropores are unsaturated.

In reality, however, the air entry value of macropores in aggregated soils is much lower than that of micropores. Therefore, it is plausible to assume that the macropores become drained at the very early stages of desaturation and the total degree of saturation reads

$$S_r \simeq \Psi^1 S^{I1} \quad (15)$$

As for the water retention equation governing S^{I1} , Durner [34] and Carminati et al. [35] considered a van Genuchten-type of equation [31] in non-deformable aggregated soil. A similar equation has been here considered and extended for deformable media by introducing the void ratio:

$$S^{I1} = \begin{cases} 1 & ; s \leq s_e^1 \\ \left[1 + (\alpha_s (e^1)^\beta s)^N\right]^{1/N-1} & ; s > s_e^1 \end{cases} \quad (16)$$

where $\alpha_s [ML^{-1}T^{-2}]$ and $N [-]$ are the van Genuchten's shape parameters, and the exponent β denotes the contribution of void ratio in the variation of degree of saturation. The way the deformation effects are included in this equation is similar to the one proposed by Gallipoli et al. [36], and allows variation of degree of saturation with induced deformation even at a constant value suction as evidenced in the experimental results [13, 14, 36, 37]. Note that in the above form of liquid retention law, the hydraulic hysteresis on drying and wetting paths [38] is not considered.

Using Eq. 6 for the degree of soil structure, we can write $\Psi^1 = 1 - e_i^2 R/e$ yielding the following expression of the total degree of saturation

$$S^{I1} = \begin{cases} 1 & ; s = 0 \\ 1 - e_i^2 R/e & ; s_e^1 > s > 0 \\ (1 - e_i^2 R/e) \cdot \left[1 + (\alpha_s (e - e_i^2 R)^\beta s)^N\right]^{1/N-1} & ; s > s_e^1 \end{cases} \quad (17)$$

The degree of soil structure in the above equation, as a state parameter, is linked to the plastic strains through (21). Accordingly, the above water retention model is coupled with the mechanical model through the degree of soil structure, void ratio and also, the suction and the degree of saturation present in the expression of the effective stress.

5 NUMERICAL SIMULATIONS

The model has been used for numerical simulation of experimental results. Numerical integration of the constitutive equations was made using an existing driver of constitutive equations [39] modified for the equations of ACMEG-2S. Input data includes the material properties, material state and imposed loading. After initialization of parameters and stress states, the model uses a return mapping-type algorithm in which stress and strains of a given step are predicted based on elastic analysis and then corrected using plastic iteration.

The model has been examined for its capability in reproducing the experimental result of suction-controlled oedometer test on unsaturated aggregated Bioley silty clay given in [14]. The results that are presented here (samples USS03) correspond to oedometric compression at the maximum suction of circa 500 kPa. The model parameters are determined based on suction-controlled oedometer tests on corresponding unsaturated reconstituted samples, and some mercury intrusion tests (MIP). Values of Ψ_i^2 have been deduced based on the aggregate porosity 22% obtained from MIP tests of single aggregate [15]. The reference suction is 500 kPa; and the initial soil structure function at this suction $\psi_i^{st} = 175.63$.

This test involves two phases: first is the increase of suction from 50 to 500 kPa, and second is the mechanical loading and unloading where the vertical net stress varies between 15 and 3000 kPa under the constant suction. Figure 7(a) shows the evolution of the degree of soil structure, R , as the effective stress increases. R reduces from 1 at the initial state to 0.26 at the end of the compression, remaining constant during unloading. As observed in Figure 7(b), the model successfully reproduces the effective stress-strain response of the sample.

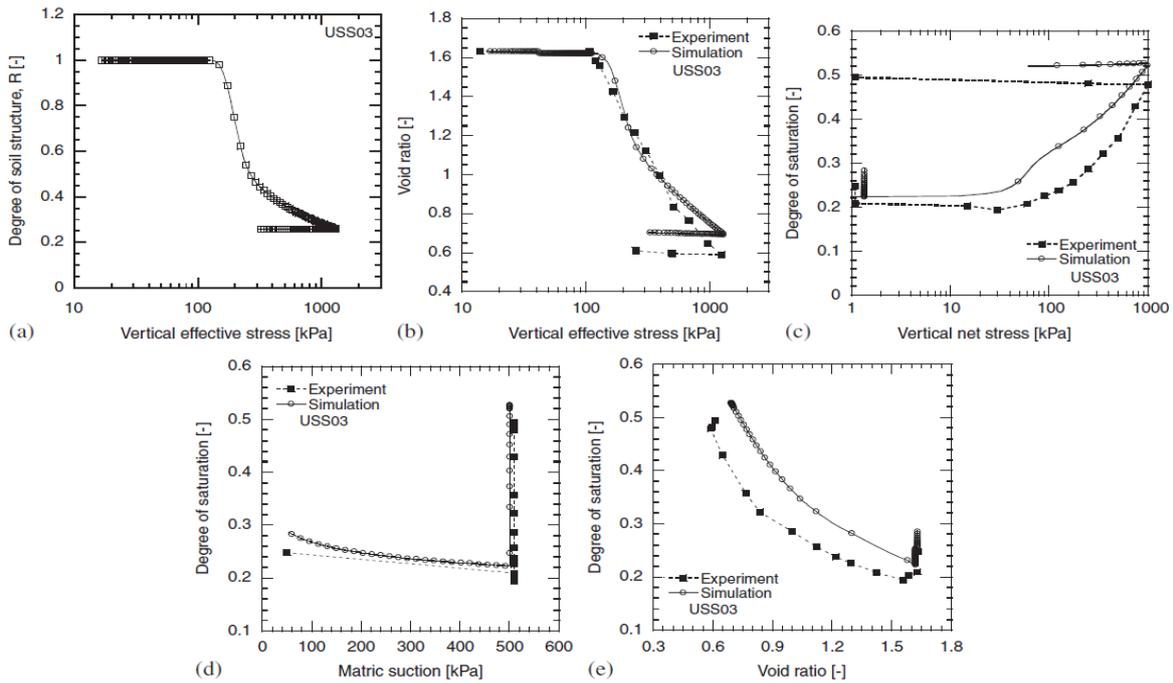


Figure 5 : Model simulation and experimental results for unsaturated oedometer test on aggregated Silty clay

Thanks to the modified water retention relation and its coupling with deformation, the model addresses the main features in the evolution of degree of saturation during the test (Figure 17(c)). For the early steps, despite a relatively low value of initial suction ($s_i=50\text{kPa}$), the model gives a low degree of saturation of about 0.29. This is due to the existence of empty macropores. Similar to the experimental results, simulations also show an insignificant decrease of degree of saturation due to suction increase. During the subsequent mechanical loading at constant suction, however, the model predicts the strong increase of the degree of saturation because of macropore closure. The good correlation between the simulated and experimental data for this phenomenon can be better observed in Figures 7(d) and 7(e) where the evolution of degree of saturation is expressed in terms of suction and void ratio.

6 CONCLUSIONS

A coupled water retention–mechanical constitutive model, named ACMEG-2S, was proposed for unsaturated soils with double porosity. It incorporates the inter-particle bonding, fabric and partial saturation effects in a single framework. The model builds upon multi-scale experimental results and has two interconnected parts: mechanical and water retention parts.

The mechanical part of the model lies within the framework of hardening elastoplasticity and is based on the critical state concept. It uses non-linear elasticity and two isotropic and deviatoric plastic mechanisms. Using the generalized Bishop's effective stress with the Bishop's parameter being the degree of saturation, the model allows a straight forward transition from saturated to unsaturated condition. Partial saturation is considered through the primary effects on soil matrix, and secondary effects coupled with soil structure. A new apparent effective preconsolidation pressure has been introduced which depends not only on stress state and history but also on soil structure and suction. The soil structure is quantified in terms of macropore volume fraction using an experimentally-based parameter, called degree of soil structure, the evolution of which is linked to the plastic strains.

In the water retention part of the model, an improved relation has been proposed accounting for deformation and double porosity effects. The two parts of the model are fully coupled through the expression of the effective stress and the degree of soil structure.

Numerical simulations showed that the model can successfully address the main features including the non-linear stress–stress relationship during the virgin compression and the increase of degree of saturation during compression at constant suction. The proposed model provides an efficient tool for coupled constitutive modeling of unsaturated structured soils behavior in geotechnical problems.

REFERENCES

- [1] Barrenblatt G, Zheltov I, Kochina I. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks. *Journal of Applied Mathematics and Mechanics. Translation of the Soviet Journal: Prik-ladnaja Matematika i Mekhanika (PMM)* 1960; 24(5):852–864.
- [2] Barrenblatt G. On certain boundary value problems for the equation of seepage of liquid in fissured rock. *Journal of Applied Mathematics and Mechanics. Translation of the Soviet Journal: Prik-ladnaja Matematika i Mekhanika (PMM)* 1963; 27:513–518.
- [3] Rouainia M, Wood DM. A kinematic hardening constitutive model for natural clays with

- loss of structure. *Géotechnique* 2000; 50(2):153–164.
- [4] Kavvas M, Amorosi A. A constitutive model for structured soils. *Géotechnique* 2000; 50(3):263–273.
- [5] Baudet B, Stallebrass S. A constitutive model for structured clays. *Géotechnique* 2004; 54(4):269–278.
- [6] Gens A, Nova R. Conceptual bases for a constitutive model for bonded soils and weak rocks. In *Geotechnical Engineering of Hard Soils—Soft Rocks*, Anagnostopoulos et al. (eds). Balkema: Rotterdam, 1993; 485–494.
- [7] Oka F, Leroueil S, Tavenas F. Constitutive model for natural soft clay with strain softening. *Soils and Foundations* 1989; 29(3):54–66.
- [8] Liu MD, Carter JP. Virgin compression of structured soils. *Géotechnique* 1999; 49(1):43–57
- [9] Liu MD, Carter JP. Modelling the destructuring of soils during virgin compression. *Géotechnique* 2000; 50(4):479–483.
- [10] Alonso E, Gens A. On the mechanical behaviour of arid soils. *Conference on Engineering Characteristics of Arid Soils*, London, 1994; 173–205.
- [11] Leroueil S, Barbosa A. Combined effect of fabric, bonding and partial saturation on yielding of soils. *Asian Conference on Unsaturated Soils*, Singapore, 18–19 May 2000; 527–532.
- [12] Bolzon G, Schrefler BA, Zienkiewicz OC. Elastoplastic soil constitutive laws generalized to partially saturated states. *Géotechnique* 1996; 46(2):279–289.
- [13] Wheeler SJ, Sharma RS, Buisson MSR. Coupling of hydraulic hysteresis and stress–strain behaviour in unsaturated soils. *Géotechnique* 2003; 53(1):41–54.
- [14] Koliji A, Laloui L, Vulliet L. Behaviour of unsaturated aggregated soil in oedometric condition. *Soils and Foundations* 2008; 49(3):369–380.
- [15] Koliji A, Vulliet L, Laloui L. Structural characterization of unsaturated aggregated soil. *Canadian Geotechnical Journal* 2010; 47:297–311
- [16] Burland J. On the compressibility and shear strength of natural clays. *Géotechnique* 1990; 40(3):329–378.
- [17] Koliji A, Lehmann P, Vulliet L, Laloui L, Carminati A, Vontobel P, Hassanein R. Assessment of structural evolution of aggregated soil using neutron tomography. *Water Resources Research*, W00C07, vol. 44, 2008; DOI: 10.1029/2007WR006297.
- [18] Koliji A, Vulliet L, Laloui L. New basis for the constitutive modelling of aggregated soils. *Acta Geotechnica* 2008; 3:61–69.
- [19] Koliji A., L. Laloui, L. Vulliet, Constitutive modelling of unsaturated aggregated soils. *Int. J. Numer. Anal. Meth. Geomech.* 2010; 34:1846–1876
- [20] Schofield AN, Wroth C. *Critical State Soil Mechanics*. McGraw-Hill: London, 1968.
- [21] Hujieux J. *Une loi de comportement pour le chargement cyclique des sols*. Génie Parasismique, Paris, Les éditions de l’ENPC, 1985; 287–353.
- [22] François B, Laloui L. ACMEG-TS: a constitutive model for unsaturated soils under non-isothermal conditions *Int. J. Numer. Anal. Meth. Geomech.* 2008; 32:1955–1988.
- [23] Bishop A. The principle of effective stress. *Tecknish Ukeblad* 1959; 106:859–863.
- [24] Nuth M, Laloui L. Effective stress concept in unsaturated soils: clarification and validation of a unified framework. *Int. J. Numer. Anal. Meth. Geomech.* 2008; 32:771–801.

- [25] Laloui L, Nuth M. On the use of the generalised effective stress in the constitutive modelling of unsaturated soils. *Computer and Geotechnics* 2009; 36(1-2):20-23.
- [26] Khalili N, Geiser F, Blight GE. Effective stress in unsaturated soils: review with new evidence. *International Journal of Geomechanics* 2004; 4(2):115-126.
- [27] Rizzi E, Maier G, Willam K. On failure indicators in multi-dissipative materials. *International Journal of Solids and Structures* 1996; 33(20-22):3187-3214.
- [28] Yu HS. Plasticity and Geotechnics. *Advances in Mechanics and Mathematics*. Springer: New York, 2006.
- [29] Nuth M, Laloui L. New insight into the unified hydro-mechanical constitutive modelling of unsaturated soils. In *The 3rd Asian Conference on Unsaturated Soils*, Yin et al. (eds). Science Press: China, 2007; 109-126.
- [30] Borja RI. Cam-Clay plasticity. Part V: a mathematical framework for three-phase deformation and strain localization analyses of partially saturated porous media. *Computer Methods in Applied Mechanics and Engineering* 2004; 193(48-51):5301-5338.
- [31] Van Genuchten MT. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal* 1980; 44: 892-898.
- [32] Brooks R, Corey A. Properties of porous media affecting fluid flow. *Journal of the Irrigation and Drainage Division (ASCE)* 1966; 92:61-68.
- [33] Coppola A. Unimodal and bimodal descriptions of hydraulic properties for aggregated soils. *Soil Science Society of America Journal* 2000; 64(4):1252-1262.
- [34] Durner W. Predicting the unsaturated hydraulic conductivity using multiporosity water retention curves. M.Th. *Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*. University of California: Riverside, 1992; 185-202.
- [35] Carminati A, Kaestner A, Ippisch O, Koliji A, Lehmann P, Hassanein R, Vontobel P, Lehmann E, Laloui L, Vulliet L, Flüher H. Water flow between soil aggregates. *Transport in Porous Media* 2007; 68(2):219-236.
- [36] Gallipoli D, Wheeler SJ, Karstunen M. Modelling the variation of degree of saturation in a deformable unsaturated soil. *Géotechnique* 2003; 53(1):105-112.
- [37] Monroy R. *The influence of load and suction changes on the volumetric behaviour of compacted London Clay*. Ph.D. Thesis, Imperial College, London, 2005.
- [38] Croney D. The movement and distribution of water in soils. *Géotechnique* 1952; 3(1):1-16.
- [39] Modaressi A, Modaressi H, Piccuezzu E, Aubry D. *Driver de la loi de comportement de hujoux*. Ecole Centrale, Paris, 1989.

ALGEBRAIC MULTIGRID PRECONDITIONING FOR MIXED ELLIPTIC-HYPERBOLIC PROBLEMS

Maximilian Emans*

* AVL List GmbH
Hans-List Platz 1, 8020 Graz, Austria
e-mail: maximilian.emans@gmx.at, www.avl.com

Key words: Algebraic Multigrid, Applications, Computing Methods

Abstract. Algebraic multigrid solvers and preconditioners are level of the art solution techniques for many types of linear systems in science and engineering. In this contribution we will compare the computational performance of different algebraic multigrid techniques as preconditioners of Krylov-solvers for coupled systems that reflect the discretisation of problems of mixed elliptic-hyperbolic type. We will report on our experience with different aggregation and cycling strategies as well as on own development and implementation improvements. Our benchmarks are cases of different size from CFD (computational fluid dynamics) applications where the pressure-correction equation is coupled to a transport equation. Very similar systems are those solved in geo-engineering applications, e.g. in oil reservoir simulations. Recently presented k-cycle methods are very efficient and can be readily modified for such linear problems.

1 INTRODUCTION

The efficient solution of linear systems representing the discretisation of mixed elliptic-hyperbolic problems is the kernel of simulations in many areas of science. Some important examples are computational fluid dynamics (CFD), oil reservoir simulations, ground water flow simulations, and semiconductor device simulations. These linear systems reflect the discretisation of coupled differential equations. For subsonic flow problems the pressure usually is ruled by an elliptic differential equation while most of the other variables are essentially determined by equations of hyperbolic type. The common discretisation of problems with both types of equations results in linear systems we refer to as coupled systems.

For large linear problems multigrid techniques are essential parts of effective solvers. In particular for partial differential equations that are discretised on unstructured grids, AMG (algebraic multigrid) methods are attractive since they require for the construction of the grid hierarchy only the information that is stored in the matrix such that the

definition of an interface to the solver part of a program is particularly simple. Since a number of years such solvers are applied with great success to scalar linear problems, i.e. to problems where the components of the unknown vector are associated with a single physical unknown. Sufficiently robust methods are known for a large number of physical problems; appropriate implementations and algorithmic modifications allow to apply algebraic multigrid techniques also on parallel computers with shared or distributed memory without major draw-backs with respect to parallel performance.

For the engineer or scientist who does physical modelling on the level of an own implementation of a simulation program and who wants to solve a particular system, it is a very pleasant fact that a number of very good AMG implementations have been provided to the scientific community through the world wide web. The most prominent examples are hypre, see Falgout and Yang [9], and ML as part of Trilinos, see Gee et al. [10]; rather new is the package AGMG, see Notay [11]. With regard to application to coupled problems and in particular to mixed elliptic-hyperbolic problems, with none of these packages a particularly good performance can be expected. The reason is that for the construction of an appropriate grid the association of the component of the unknown vector to the physical unknowns should be taken into account.

In this contribution we summarise the modifications to existing AMG implementations that are needed for good performance in applications where mixed elliptic-hyperbolic problems are to be solved. We then focus on aggregation AMG and compare the performance of solvers that are obtained by adapting the most important aggregation algorithms to coupled problems.

2 ALGEBRAIC MULTIGRID FOR COUPLED SYSTEMS

We consider a continuous problem with more than one unknown function, i.e. with more than one physical unknown or independent variable. The discretisation leads us to a system of linear equations. Let us denote the linear problem for the moment as

$$A\vec{x} = \vec{b} \tag{1}$$

where $A \in \mathbb{R}^{n \times n}$ is regular and sparse, $\vec{b} \in \mathbb{R}^n$ is some right-hand side vector and $\vec{x} \in \mathbb{R}^n$ the solution where n is the rank of A and thus the number of components of the unknown solution vector \vec{x} . These components are referred to as “nodes” in a large part of the literature on AMG. One node corresponds to one unknown value of one of the function that are approximated by the solution of the linear system. It refers to the vertices of the connectivity graph of the matrix and not to a geometric entity (like a point of the discretisation scheme). Note that for coupled systems this difference is substantial while for scalar systems it is only a question of the word choice.

An AMG algorithm comprises two phases: A setup phase and a solution phase. In the setup phase the operators $A_k \in \mathbb{R}^{n_k \times n_k}$ ($k = 2, \dots, k_{max}$, $A_k = A$) with system size n_k are defined where $n_{k+1} < n_k$ holds for $k = 1, \dots, k_{max} - 1$. Moreover, a smoother S_k as well

as a prolongation operator P_k and a restriction operator R_k have to be determined for each level k . The effects of prolongation and restriction are illustrated in figure 1. It is common practice in algebraic multigrid to choose $R_k = P_k^T$ and to follow the Galerkin approach to generate the coarse-grid hierarchy recursively by

$$A_{k+1} = P_k^T A_k P_k \quad (k = 1, \dots, k_{max} - 1). \quad (2)$$

Various ways have been described in the literature to define P_k ; in fact this is the major difference between the different AMG methods.

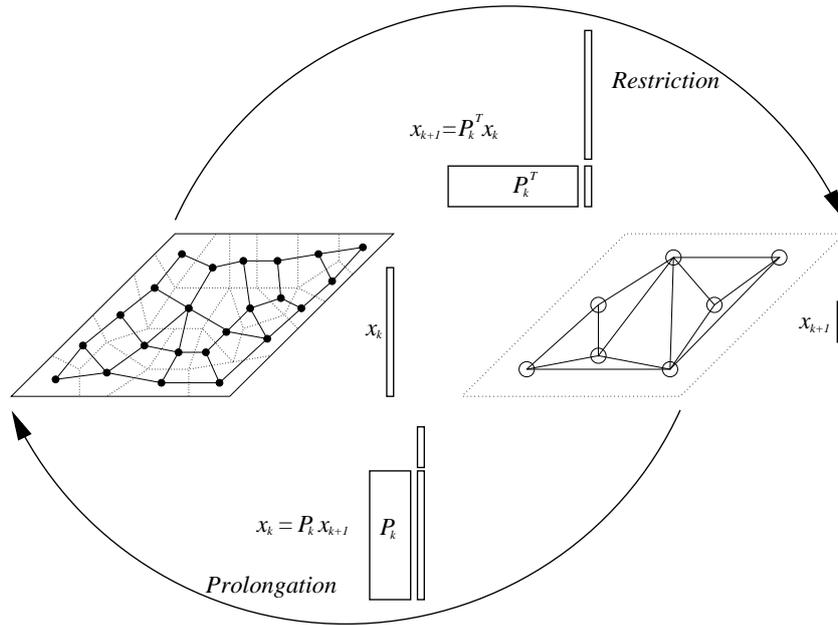


Figure 1: Illustration of prolongation and restriction

The prolongation operator P_l projects a vector of size n_{l+1} which represents a solution on level $l + 1$ to a vector of size n_l representing a solution on level l . Generally, the values on level l are interpolated values where the supporting points are the nodes on level $l + 1$. In a geometric scheme the interpolation relies on the spatial relation of the nodes; in an algebraic scheme, however, the geometric relation of the nodes is not explicitly known and is therefore replaced by the matrix connectivity. Note that in many cases, on the finest level, the geometric relation is reflected by the matrix connectivity. It has to be avoided that the interpolation takes place between two nodes associated with different physical values. For AMG schemes it is therefore not enough to consider only the information contained in the matrix: those connections corresponding to edges in the connectivity graph that link nodes associated with different physical unknowns need to be skipped in the construction of the interpolation scheme. Assuming (for simplicity) an ordering of the



Figure 2: Prolongation operators assuming ordering of nodes by physical unknown (left: scalar systems, right: coupled systems, grey area: first physical unknown out of two)

nodes by physical unknown, this leads to a block structure of the prolongation operators as shown in figure 2.

The blocks in the prolongation operator can now be constructed in exactly the same way as the corresponding operators for scalar systems. The association of the nodes to the physical unknown can either be tracked by a marker field or, as we did in the implementations used for this work, the variables are ordered by physical unknown on each grid; in our opinion the latter way is advantageous since practically no reordering on the coarse-grid is necessary because it is natural to define the coarse-grids for the individual physical unknowns one after the other.

If indeed the coarse-grid of each physical unknown is constructed as if the matrix would describe a scalar problem, then the method is referred to as unknown-based approach, see Clees and Ganzer [3]. The same authors report that the employment of an identical aggregation scheme to all physical unknowns can be beneficial for the convergence. In terms of a geometrical method this approach means that for each physical unknown on any point the same coarse-grid points are involved in the interpolation. This scheme is referred to as point-based approach. Note that this cannot be exploited to simplify the computation of the coarse-grid operators by Galerkin approach (2).

In the solution phase these structures are employed to approach iteratively, starting from an initial guess, the solution. The pseudocode of a standard v-cycle algorithm as one of the easiest cycling strategies is shown in algorithm 1. In practice, the cycling strategy, the smoother and its parameters, and the choice of the prolongation and restriction operator should be adjusted carefully to each other in order to obtain an efficient method.

3 BENCHMARKS

In this section we first describe shortly five algorithms; the selection reflects technically feasible possibilities. For details we will refer the interested reader to the literature. For

Algorithm 1 v-cycle AMG

$$\vec{x}_k^{(3)} = \text{AMG_par}(k, \vec{r}_k, \vec{x}_k^{(0)})$$

Input: level l , right-hand side \vec{r}_k , initial guess $\vec{x}_k^{(0)}$

Output: approximate solution $\vec{x}_k^{(3)}$

-
- 1: **if** $k = k_{max}$ **then**
 - 2: solution of coarse-grid system: $\vec{x}_k^{(3)} = A_k^{-1} \vec{r}_k$
 - 3: **else**
 - 4: pre-smoothing: $\vec{x}_k^{(1)} = S_k(\vec{r}_k, A_k, \vec{x}_k^{(0)})$
 - 5: restriction: $\vec{r}_{k+1} = P_k^T(\vec{r}_k - A_k \vec{x}_k^{(1)})$
 - 6: recursive solution of coarse-grid system: $\vec{x}_{k+1} = \text{AMG_par}(k+1, \vec{r}_{k+1}, \vec{0})$
 - 7: prolongation of coarse-grid solution and update: $\vec{x}_k^{(2)} = \vec{x}_k^{(1)} + P_k \vec{x}_{k+1}$
 - 8: post-smoothing: $\vec{x}_k^{(3)} = S_k(\vec{r}_k, A_k, \vec{x}_k^{(2)})$
 - 9: **end if**
-

the AMG variants we name sources of efficient implementations of the algorithms. Later on we present two benchmarks: the sample systems are two linear systems that reflect the discretisation of a problem of elliptic-hyperbolic type from a CFD application.

3.1 AMG variants

The smoother in all algorithms is a ILU(0) smoother with one sweep before and after the coarse-grid correction. The coarse-grid treatment is done by an agglomeration scheme, i.e. as soon as the grid on one of the processes becomes too small in the course of the coarsening, it is merged to the grid of a neighbour; whenever in the solution phase action on this or a coarser level is taken, this process is idle. The threshold grid size is 200 nodes. An alternative to this procedure is the employment of a sparse direct solver such as MUMPS, see Amestoy et al. [1]. It is easier to implement, but in particular for smaller problems or for runs on distributed systems with slow interconnect the computation might take up to 20 % more time, see Emans [6]. If one starts with one of the mentioned packages, the question which coarse-grid treatment technique to choose is obsolete since these packages cover the parallelisation of the whole algorithm, e.g. AGMG uses MUMPS. The differences between the algorithms are the coarse-grid selection schemes and, related to this, the cycling strategies. This will be outlined in the following.

Smoothed Aggregation AMG This scheme forms aggregates containing typically between 20 and 50 nodes. Formally, a tentative prolongation operator with constant interpolation is constructed. Since it would result in rather poor representation of the fine-grid problem on the coarse-grid, it is smoothed by applying one Jacobi-smoothing step to it. The implemented serial algorithm has been described by Vaněk et al. [15]. In our parallel implementation the aggregation and the smoothing are strictly local processes, i.e. aggregates are not intersected by inter-domain boundaries. More details of the parallel

implementation used here have been discussed in the previous publication [5]. Another similar implementation of this algorithm can be found under the short ML in the solver package Trilinos [10]. A measure for the memory requirement is the operator complexity, i.e. the ratio of the sum of the number of nodes on all grids and the number of nodes on the finest grid. For this method it is typically around 1.5. Here, this algorithm is used as a preconditioner of a GMRES method. The multigrid cycle is a v-cycle with one pre-smoothing and one post-smoothing sweep of the ILU(0) smoother. The algorithm is referred to as **ams1gm**.

Pairwise Aggregation AMG Notay [11] suggested recently an efficient algorithm that produces aggregates of only two nodes. Due to the small aggregate size, a solution on grid $k+1$ is represented with reasonable quality on grid k if constant interpolation is employed. With this interpolation scheme, the prolongation operator P_k is also particularly simple: One row (that corresponds to a node on level k) contains exactly one entry with value one, namely in the column that corresponds to the aggregate on level $k+1$ this node is assigned to; all other elements are zero. The computation of the coarse-grid hierarchy is particularly cheap: Equation (2) reduces essentially to an addition of rows of A_k . Consequently, the computation of the coarse-grid hierarchy in total is inexpensive. It is local to each process since only nodes assigned to the same process are grouped together in an aggregate. However, the number of grids will be rather large since from level l to level $l+1$ the number of nodes is reduced only by a factor of approximately two. It can be seen as a disadvantage that the operator complexity of this method is about 2.5. The cycling strategy is an f-cycle, see Trottenberg et al. [14] for a definition, with one pre-smoothing and one post-smoothing sweep of the ILU(0) smoother. The algorithm is referred to as **amf1gm**.

GCR-accelerated AMG This kind of preconditioner has been suggested by Notay [11]. An implementation (which is quite similar to ours) is available as AGMG from the homepage of the author. The main difference to conventional (fixed cycle) AMG preconditioners is that the coarse-grid system is approximated by one or two iterations (depending on a termination criterion) of an inner GCR-solver [4] (in a modified economic implementation) that is preconditioned by AMG instead of being approximated by the multigrid scheme alone. Since this preconditioning operation is not the same in each iteration, standard Krylov-methods cannot be used as outer iteration. For nonsymmetric problems Notay [11] uses GCR also as outer Krylov-solver. In this algorithm the double-pairwise aggregation suggested by Notay [11] is employed: It is obtained by applying the pairwise aggregation of algorithm **amf1gm** twice. The aggregates usually comprise four nodes, the computation of the coarse-grid system is similarly cheap as that of **amf1gm**. Since only every second grid and its associated operator is stored, the operator complexity is only about 1.6 for this method. We refer to this algorithm as **amk1gm**.

GMRES-accelerated AMG In this algorithm the GCR in `amk1gc` as inner and outer solver is replaced by GMRES. Since the preconditioning operation is not the same in each iteration, GMRES is implemented as FGMRES, see Saad [12]. In our experience this GMRES-based method is significantly more robust than the GCR-based one. A discussion of this issue along with relevant examples can be found in Emans [7]. All other parts of the algorithm, in particular the cycling strategy and the coarsening algorithm, are the same as that of `amk1gc`. A potential disadvantage of this method is that (up to the knowledge of the author) an implementation cannot be directly obtained and consequently some implementation effort will be faced. We refer to this algorithm as **amk1gm**.

ILU-preconditioned BiCGstab The algorithm that serves as a reference is a ILU(0)-preconditioned BiCGstab method. It has been deduced and discussed in detail by Saad [13]. Our implementation requires only to store one additional vector of the size of the unknown vector. We apply two sweeps in each iteration of the Krylov method. Here, we refer to this method as **ilu0bc**.

3.2 Performance

In order to compare the efficiency of the algorithms we have chosen two linear problems from different CFD applications. The SIMPLE algorithm extended by a pressure-enthalpy coupling, see Emans et al.[8], leads to a system of the type

$$\begin{pmatrix} C & S_h \\ S_p & G \end{pmatrix} \begin{pmatrix} \vec{p}' \\ \vec{h}' \end{pmatrix} = \begin{pmatrix} \vec{c} \\ \vec{g} \end{pmatrix} \quad (3)$$

where C is the symmetric positive definite operator of the pressure-correction equation, G the operator of the transport equation for the enthalpy, S_h and S_p are the coupling operators, \vec{c} and \vec{g} are the right-hand sides; the components of the unknown vector are either associated with the pressure correction \vec{p}' or with the enthalpy update \vec{h}' . Since the underlying partial problem of the pressure-correction equation is elliptic and the transport of the enthalpy is essentially a hyperbolic problem, the system is referred to be of mixed elliptic-hyperbolic type. The iterative solution procedure is terminated if the 1-norms of the residuals of both parts of the system have been reduced by a factor of $1.0 \cdot 10^{-6}$.

In problem 012 the flow of cold air (293 K) into a complex engine cylinder geometry with hot walls (650 K) is simulated; the mesh consists of 1.4 mio cells of which approximately 80% are hexagonal while the rest is tetrahedral. In problem 045 a backward-facing step problem, see Armaly et al. [2], is solved where hot gas (1100 K) is entering the domain that is initially filled by cold gas (293 K). The orthogonal mesh of this case has 2.3 mio cells.

For the benchmarks we used up to eight nodes à 2 quad-cores of a Linux cluster; each node of this cluster is equipped with two Intel Xeon CPU X5365 (3.00GHz, main memory 16GB, L2-cache 4MB shared between two cores). The nodes are connected

by an Infiniband interconnect with an effective bandwidth of around 750 Mbit/s and a latency of around 3.3 μ s. The computational part of the program is compiled by the Intel-FORTRAN compiler 10.1, the communication is performed through calls to hp-MPI subroutines (C-binding). The benchmarks were run within the environment of the software AVL FIRE^(R) 2009 with 1, 2, 4, 8, 16, and eventually 32 processes. Computations with 1, 2, and 4 processes were done on a single node, for more processes we used 2, 4, and 8 nodes respectively. This is done to ensure that each process has sole access to one of the L2-caches on the node. It is important to note that the MPI implementation uses the shared memory space of one node for the communication between two processes wherever possible. This means that all intra-node communication (calculations with up to 4 processes) is done without utilisation of the network interconnect.

In figures 3 and 4 the number of iterations, the computing times, and the parallel efficiency $E_p = t_1/(t_p \cdot p)$ (t_p : computing time with p parallel processes) are shown. First it is interesting to see if and at what rate the numerical method converges. Convergence within 250 iterations is observed in all AMG calculations apart from the `ams1gm` runs for problem 012. In this case the algorithm does neither stall nor diverge, but the convergence is so slow that the criterion is not reached within 250 iterations. The diagrams with the iteration count shows further that the AMG methods converge much more rapidly than the ILU(0) method. Apart from `ams1gm` for problem 045, the AMG solvers the parallelisation does not mitigate the convergence behaviour.

The computing times show that the k-cycle methods (`amk1gc` and `amk1gm`) are the fastest methods. They are faster by 50% and by 90% compared to `ilu0bc` for 012 and 045 respectively. The larger advantage in the latter case is due to the problem size. Compared to the pairwise aggregating AMG `amf1gm` the advantage is still considerable. Comparing now `amk1gc` and `amk1gm` the results show that the differences are rather small: for problem 012 `amk1gm` is slightly faster, for problem 045 `amk1gc` is slightly faster. Since both systems here are solved by both algorithms, the main advantage of `amk1gm` stays hidden: It is its larger robustness, see Emans [7]. From the view point of practical usage, `amk1gc` is advantageous since, coming from AGMG, only the modification of the coarse-grid generation for coupled system needs to be implemented.

The parallel efficiency of both algorithms `amk1gc`, `amk1gm`, and `amf1gm` is satisfying and competitive to that of `ilu0bc`. The curves are typical for computations on the used hardware: For up to four processes a decrease is observed. This is due to the fact the the data transfer mechanism on the chip which is a Front-Side Bus (FSB) is the bottleneck; the data transfer between the processes is fast since it relies on the shared memory mechanism. If more processes are involved, inter-node data transfer through the interconnect is required; more important is that the probability of cache misses is reduced since the decomposed problems have become smaller and a larger portion of them fits into the cache. Consequently the curves rise. The rise is more pronounced for smaller problem 012. The observed parallel efficiencies of these coupled solvers correspond well to the parallel efficiencies which can be reached with similar methods for scalar problems.

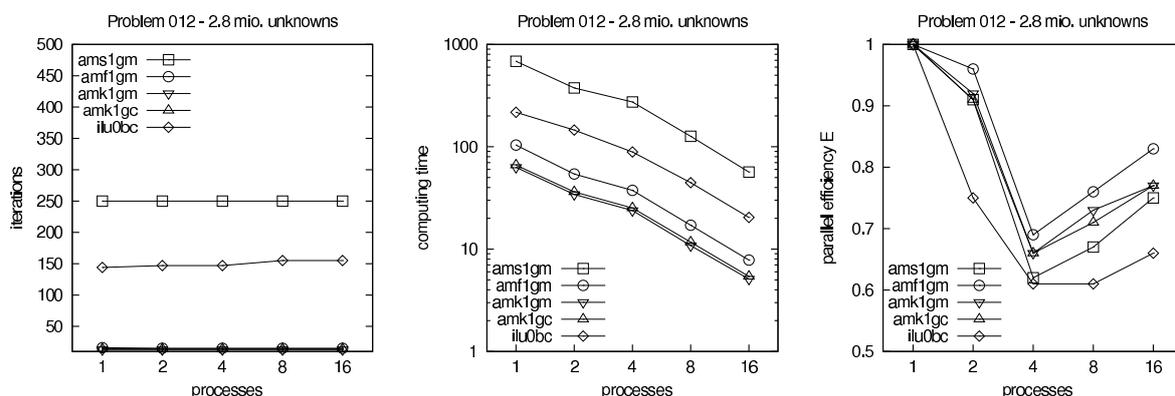


Figure 3: Iteration count (left), computing times (middle), and parallel efficiency (right) for benchmark 012

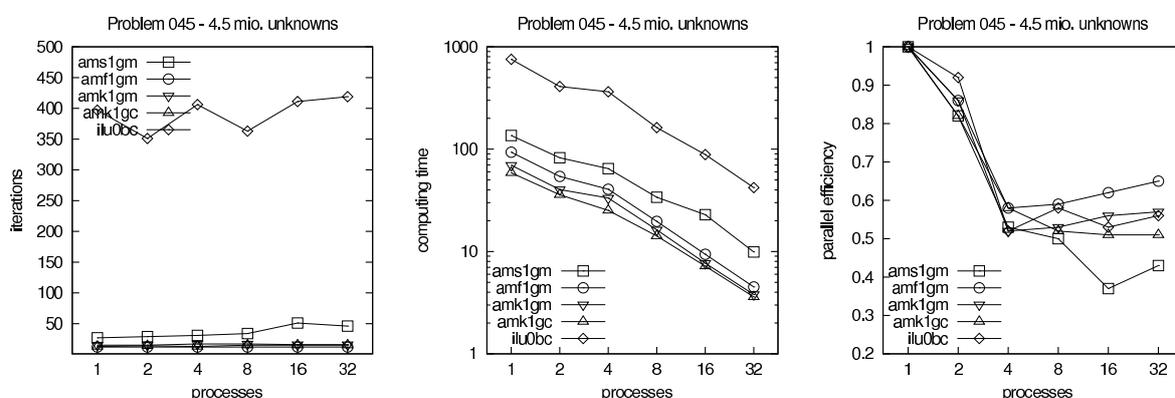


Figure 4: Cumulative iteration count (left), computing times (middle), and parallel efficiency (right) for benchmark 045

4 CONCLUSIONS

The k-cycle methods amk1gc and amk1gm are the fastest AMG solvers for our benchmarks. Practically, amk1gc can be deduced from existing methods which are available to the scientific community where the effort to implement the modifications is small. These relatively new algorithms outperform both, Smoothed Aggregation AMG and ILU(0)-preconditioned BiCGstab.

REFERENCES

- [1] P. R. Amestoy, I. S. Duff, J. Koster, and J.-Y. L'Excellent. A fully asynchronous multifrontal solver using distributed dynamic scheduling. *SIAM Journal on Matrix Analysis and Applications*, 23:15–41, 2001.
- [2] B.F. Armaly, F. Durst, J.C.F. Pereira, and B. Schonung. Experimental and theoretical investigation of backward-facing step flow. *Journal of Fluid Mechanics*, 48:473–

- 496, 1983.
- [3] T. Clees and K. Stüben. Algebraic multigrid for industrial semiconductor device simulation. In E. Bänsch, editor, *Challenges in Scientific Computing – CISC2003*, volume 35 of *Lecture Notes in Computational Science and Engineering*, pages 110–130. Springer-Verlag, 2002.
 - [4] C.S. Eisenstat, H.C. Elman, and M.H. Schultz. Variational iterative methods for nonsymmetric systems of linear equation. *SIAM Journal on Numerical Analysis*, 20:345–357, 1983.
 - [5] M. Emans. Approximate solutions of linear systems in CFD applications. *SIAM Journal on Scientific Computing*, 32:2235–2254, 2010.
 - [6] M. Emans. Coarse-grid treatment in parallel AMG for coupled systems in CFD applications. 2010. submitted.
 - [7] M. Emans. K-cycle AMG for semi-definite and nonsymmetric systems in CFD. 2010. submitted.
 - [8] M. Emans, S. Frolov, B. Lidskii, V. Posvyanskii, Z. Žunič, and B. Basara. Pressure-enthalpy coupling for subsonic flows with density variation. In M. Rahman and C.A. Brebbia, editors, *AFM VIII*, pages 127–136. WIT Press Southampton, 2010.
 - [9] R.D. Falgout and U.M. Yang. hypre: a library of high performance preconditioners. In P.M.A. Sloot, C.J.K. Tan, J.J. Dongarra, and A.G. Hoekstra, editors, *ICCS2002, Part III*, volume 2331 of *Lecture Notes on Computer Science*, pages 632–641. Springer, 2002.
 - [10] M.W. Gee, C.M. Siefert, J.J. Hu, R.S. Tuminaro, and M.G. Sala. *ML 5.0 Smoothed Aggregation User’s Guide*. SAND2006-2009 Unlimited Release, 2006.
 - [11] Y. Notay. An aggregation-based algebraic multigrid method. *Electronic Transactions on Numerical Analysis*, 37:123–146, 2010.
 - [12] Y. Saad. A flexible inner-outer preconditioned GMRES algorithm. *SIAM Journal on Scientific Computing*, 14:461–469, 1993.
 - [13] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, 3 edition, 2003.
 - [14] U. Trottenberg, C. Oosterlee, and Anton Schüller. *Multigrid*. Elsevier Academic Press Amsterdam, 2001.
 - [15] P. Vaněk, J. Mandel, and M. Brezina. Algebraic multigrid by smoothed aggregation for second and fourth order elliptic problems. *Computing* 56, pages 179–196, 1996.

HYDROMECHANICAL ANALYSIS IN GEOTECHNICAL ENGINEERING USING THE MATERIAL POINT METHOD

F. ZABALA* AND E. ALONSO**

* Instituto de Investigaciones Antisísmicas- Universidad Nacional de San Juan
Av. Libertador San Martín Oeste 1290, 5400, San Juan, Argentina
E-mail: fzabala@unsj.edu.ar

** Departamento de Ingeniería del Terreno, Cartográfica y Geofísica
Universidad Politécnica de Cataluña- Edificio D2, Campus Norte UPC
Gran Capitán s/n, 08034 Barcelona, Spain
Email: eduardo.alonso@upc.edu

Key words: Material Point Method, pore pressure, finite strain.

Abstract. The explicit version of the Material Point Method [1] has been extended in order to model coupled hydromechanical saturated problems. MPM discretizes the continuum, which is considered as a saturated soil-fluid mixture, by dividing it into particles or material points. The discrete movement equations are not solved at the material points. Instead a support mesh, built to cover the domain of the problem, is used. In this paper it is assumed that particles carry all the variables needed to represent the state of the continuum including the pore pressure as a variable associated with each particle. The particle pore pressure increment is calculated explicitly using the equation of fluid mass balance, from the particle volumetric deformation and the fluid velocity relative to the soil skeleton, at the particle location.

The shape functions used for the mesh elements are usually the same bi-linear functions of the Finite Element Method and therefore the background mesh elements suffer the same drawbacks. These drawbacks include: volumetric locking for quasi-incompressible materials when four particles per cell are used, which is equivalent to four integration points in the finite element method, pressure instability for quasi-incompressible and low permeability materials and the generation of zero energy modes when one particle per cell is used, which corresponds to reduced integration in the finite element method. The MPM original version has also the disadvantage of generating "noise" in the solution [2] when a particle pass from one cell to another. A simple procedure that can be used to reduce instabilities is to consider constant stress at each cell equal to the stress average of the particles which are in the cell at the instant k . In this case the internal forces are obtained in the same way as in the finite element method when one point of integration is used, using the gradient of the shape functions calculated in the cell center. In this work, to avoid volumetric locking and simultaneously achieve a stable behavior, internal forces and pressure increments at the nodes are calculated using the gradients calculated at the cell center.

The procedure is completely explicit and has proved to be stable for the low permeability values used to model the foundation of Aznalcollar dam. The simulation of Aznalcollar dam progressive failure is presented as an example [3].

1 INTRODUCTION

The method was initially described by Sulsky et al. [1] and by Sulsky and Schreyer [5]. The Material Point Method discretizes the continuum dividing it into particles. A mass is assigned to each particle which remains fixed during all the calculation process, thus assuring mass conservation. Other initial values, such as velocities, strains and stresses, are also assigned to the material points. The discrete movement equations are not solved at the material points. Instead a support mesh, built to cover the domain of the problem, is used. This mesh is composed by elements of the same type as those used in finite element method. For simplicity, it is common to use four nodes bi-linear quadrilateral elements. The boundary conditions are imposed at the mesh nodes and the movement equations are incrementally solved. Then the quantities carried by the material points are updated through the interpolation of the mesh results, using the same shape functions. The information associated with the mesh is not necessary in the following step of the analysis; therefore it can be discarded, taking care of preserving the boundary conditions that may have been established. It is possible to discretize the momentum conservation equation through the application of standard Galerkin weighted residual method ([1], [5]). The derived mass matrix varies in time, and must be calculated for each step of the analysis. To reduce the computational cost of the procedure, a diagonal mass matrix can be used; therefore the movement equations are decoupled, and can be written for a node and for the k time step as:

$$m_i^k \mathbf{a}_i^k = \mathbf{f}_i^{\text{int},k} + \mathbf{f}_i^{\text{ext},k} \quad (1)$$

where $\mathbf{f}_i^{\text{int}}$ and $\mathbf{f}_i^{\text{ext}}$ are the vectors of internal and external forces at the node i , respectively, m_i is the mass of node i and a_i is the node acceleration.

2 MPM COUPLED HYDRO-MECHANICAL ANALYSIS

MPM is very well suited for dealing with finite strains and large displacements that are developed when collapse or near collapse problems are studied. In order to model coupled hydro-mechanical saturated problems the continuum can be considered as a saturated soil-fluid mixture. In this paper it is assumed that particles carry all the variables needed to represent the state of the continuum including the pore pressure as an associated variable.

The formulation of the equations which describe the behavior of a saturated porous medium was first developed by Biot [4] and then extended by Zienkiewicz and others [6], [7]. The governing equations can be simplified if the fluid acceleration relative to the soil skeleton is small as follows.

The balance of momentum is written for the mixture as:

$$\nabla \cdot \boldsymbol{\sigma} + \rho \mathbf{b} = \rho \mathbf{a} \quad (2)$$

$\boldsymbol{\sigma}$ is the total stress tensor, \mathbf{b} are the body forces, \mathbf{a} is the acceleration and ρ is the density of the mixture,

$$\rho = (1-n) \rho_s + n \rho_f \quad (3)$$

n is the porosity, ρ_s is the solid density and ρ_f is the fluid density.

The balance of fluid mass is written,

$$\dot{\epsilon}_v + \nabla \cdot \mathbf{w} + \frac{\dot{p}}{Q} = 0 \quad (4)$$

\mathbf{w} is the Darcy velocity, $\nabla \cdot \mathbf{w}$ the divergence of velocity, $\dot{\epsilon}_v$ is the volumetric strain variation of soil skeleton and \dot{p} the fluid pressure variation. Q is the combined compressibility modulus of the fluid and solid phase.

$$\frac{n}{K_f} + (1-n) \frac{1}{K_s} = 1/Q \quad (5)$$

K_f is the fluid bulk modulus, K_s is the bulk modulus of the grains material.

The Darcy equation is written:

$$\mathbf{w} = -\frac{\mathbf{k}}{\gamma_f} \left[\nabla p - \rho_f \mathbf{b} + \rho_f \mathbf{a} \right] \quad (6)$$

Where \mathbf{k} is the permeability tensor, γ_f the specific weight of fluid and ∇p the pressure gradient.

The relationship between effective stress and total stress is:

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}' - \alpha p \mathbf{m} \quad (7)$$

$\boldsymbol{\sigma}'$ is the effective stress, p the pore pressure (positive for compression), $m_{ij} = \delta_{ij}$,

δ_{ij} : Kronecker Delta, $\alpha = 1 - \frac{K_t}{K_s}$ and K_t is the bulk modulus of soil skeleton.

For soils $K_t \ll K_s$ and $\alpha \cong 1$

These equations together with boundary conditions can be discretized and solved numerically using as unknowns the displacements and the pore pressure.

2.1 MPM discretization of the equations related to the fluid.

The pore pressure increment in a particle is calculated explicitly using the equation of fluid mass balance, from the particle volumetric deformation and the fluid velocity relative to the soil skeleton, at the particle location.

From Equation(4):

$$\Delta p + Q(\dot{\epsilon}_v + \nabla \cdot \mathbf{w}) \Delta t = \Delta p + Q(\Delta \epsilon_v + \nabla \cdot \mathbf{w} \Delta t) = 0 \quad (8)$$

$\Delta\varepsilon_v$ is the volumetric strain increment at a point in the continuum and Δp the pore pressure increment at the end of time interval Δt .

The boundary conditions imposed on the mesh are:

$$\begin{aligned} &\text{in } \Gamma_p : p - \bar{p} = 0 \\ &\bar{p} : \text{Pressure imposed on the boundary } \Gamma_p \end{aligned} \quad (9)$$

$$\begin{aligned} &\text{in } \Gamma_w : k \frac{\partial p}{\partial n} - \bar{w} = w_n - \bar{w} = 0 \\ &\bar{w} : \text{Fluid velocity enforced in } \Gamma_w \\ &w_n : \text{velocity normal to boundary } \Gamma_w \end{aligned} \quad (10)$$

The standard Galerkin method is applied to Equation (8) using weighting functions for the pressures equal to those used for the interpolation of displacements.

$$\int_{\Omega_i} N_i [\Delta p + Q(\Delta\varepsilon_v + \nabla \cdot \mathbf{w}\Delta t)] d\Omega - \int_{\Gamma_w} N_i (w_n - \bar{w})\Delta t d\Gamma = 0 \quad (11)$$

N_i is the shape function for node i.

Integrating by parts,

$$\int_{\Omega_i} N_i \nabla \cdot \mathbf{w}\Delta t d\Omega = \int_{\Gamma_i} N_i \mathbf{w} \cdot \mathbf{n}\Delta t d\Gamma - \int_{\Omega_i} \nabla N_i \cdot \mathbf{w}\Delta t d\Omega \quad (12)$$

results in:

$$\int_{\Omega_i} N_i \Delta p d\Omega + \int_{\Omega_i} N_i Q \Delta\varepsilon_v d\Omega - \int_{\Omega_i} \nabla N_i \cdot \mathbf{w}\Delta t d\Omega + \int_{\Gamma_w} N_i \bar{w}\Delta t d\Gamma = 0 \quad (13)$$

The particle pressure increment is approximated using the pressure increments at the nodes using the same interpolation functions used for displacements:

$$\Delta p_p^k = \sum_{j=1}^{N_n} \Delta p_j^k N_j(\mathbf{x}_p^k) \quad (14)$$

Δp_i^k is the node pressure increment at time k, Δp_p^k the particle pressure increment at time k and \mathbf{x}_p^k is the particle position.

Replacing the integrals in equation 13 by sums of quantities at the material points, the mesh nodes pressure increments at time k can be computed. In these integrals the material points or particles correspond to the integration points in the finite element method. The first of the integrals of equation 13 is equal to:

$$\int_{\Omega_t} N_i \Delta p \, d\Omega = \sum_{p=1}^{N_p} N_i(\mathbf{x}_p^k) \Delta p_p^k V_p = \sum_{p=1}^{N_p} N_i(\mathbf{x}_p^k) \sum_{j=1}^{N_n} \Delta p_j^k N_j(\mathbf{x}_p^k) V_p \quad (15)$$

$$\sum_{j=1}^{N_n} \Delta p_j^k \sum_{p=1}^{N_p} N_i(\mathbf{x}_p^k) N_j(\mathbf{x}_p^k) V_p = \sum_{j=1}^{N_n} \Delta p_j^k V_{ij}^k$$

Where V_p is the particle volume. The element volume matrix V_{ij} can be lumped at the nodes to give a diagonal matrix V_i :

$$V_{ij}^k = \sum_{p=1}^{N_p} N_i(\mathbf{x}_p^k) N_j(\mathbf{x}_p^k) V_p \quad (16)$$

$$V_i^k = \sum_{p=1}^{N_p} N_i(\mathbf{x}_p^k) V_p \quad (17)$$

$$\sum_{p=1}^{N_p} N_i(\mathbf{x}_p^k) \Delta p_p^k V_p = \sum_{j=1}^{N_n} \Delta p_j^k V_{ij} \cong \Delta p_i^k V_i^k \quad (18)$$

Thus the pressure increments at the nodes can be calculated explicitly:

$$\Delta p_i^{k+1} = \frac{1}{V_i} \mathcal{Q} \left[\sum_{p=1}^{N_p} (N_i(\mathbf{x}_p^k) \Delta \varepsilon_{v_p}^k - \nabla N_i(\mathbf{x}_p^k) \cdot \mathbf{w}_p^k \Delta t) V_p + \int_{\Gamma_q} N_i \bar{w} \Delta t \, d\Gamma \right] \quad (19)$$

The Darcy equation is used to calculate the flow velocities at each particle:

$$\mathbf{w}_p^k = -\frac{\mathbf{k}}{\gamma_f} \left[\nabla p(\mathbf{x}_p^k) - \rho_f \mathbf{b} + \rho_f \mathbf{a}_p^k \right] \quad (20)$$

The pressure gradient is evaluated using the field interpolated from the node's pressures at the k instant. This pressure field can be approximated in the same way as the pressure increment:

$$p_p^k = \sum_{j=1}^{N_n} p_j^k N_j(\mathbf{x}_p^k) \quad (21)$$

where p_i^k is the node i pressure for instant k and p_p^k the particle pressure at instant k. Then,

$$\sum_{p=1}^{N_p} N_i(\mathbf{x}_p^k) p_p^k V_p = \sum_{j=1}^{N_n} p_j^k V_{ij} \cong p_i^k V_i \quad (22)$$

$$p_i^k = \frac{1}{V_i} \sum_{p=1}^{N_p} N_i(\mathbf{x}_p^k) p_p^k V_p \quad (23)$$

Thus the pressure gradient can be calculated with the expression:

$$\nabla p(\mathbf{x}_p^k) = \sum_{i=1}^{N_n} p_i^k \nabla N_i(\mathbf{x}_p^k) \quad (24)$$

Finally the following equation is used to update the particle pressures:

$$p_p^{k+1} = p_p^k + \sum_{i=1}^{N_n} \Delta p_i^{k+1} N_i(\mathbf{x}_p^k) \quad (25)$$

The nodal pressures are calculated with equation (23) using the pressures carried by the material points which were obtained in the previous step of the solution. This calculation scheme is similar to the procedure used in MPM to obtain the momentum at the mesh nodes. The pressure gradient is obtained through equation (24) from the mesh pressure field. The fluid flow relative velocity, at the coordinates of particles, is calculated using the Darcy equation(20). The pressure increment at the nodes is calculated using equation(19). The terms in this equation: volumetric strain increment, velocity divergence and external flow are accumulated at the nodes by adding the contribution of the particles. Boundary conditions are imposed on the mesh. The particles pressures are updated through equation (25) for use in the next step. At the end of the time interval the pressures at the nodes are not useful and can be discarded because the particles carry the information on pore pressures.

2.2 Stability of the solution

The shape functions used for the mesh elements are typically the same bi-linear functions used in the finite element method and therefore the background mesh elements suffer the same drawbacks. These drawbacks include: volumetric locking for quasi-incompressible materials when four particles per cell are used, which is equivalent to four integration points in the finite element method; pressure instability for quasi-incompressible and low permeability materials and the generation of zero energy modes when one particle per cell is used, which corresponds to reduced integration in the finite element method.

Mixed displacement-pressure approximations are known to suffer pressure instability if the Babuska-Brezzi condition is not fulfilled. In the framework of the finite element method an element is stable without using special procedures if the number of degrees of freedom used to interpolate the pressure is lower than the number used to interpolate the displacements.

The same shape functions are used in 2.1 to interpolate displacements and pressures in the discretization of the equations for the fluid. Therefore, pressure instability is expected for this formulation.

MPM original version has also the disadvantage of generating "noise" in the solution [2] when a particle pass from one cell to another. This noise is due to the discontinuity of the interpolation function gradient which is involved in the calculation of internal forces [2]. A simple procedure that can be used to reduce this type of instability is to consider a constant stress at each cell equal to the stress average of the particles which are in the cell at the instant k . In this case the internal forces are obtained in the same way as in the finite element method when one point of integration is used, using the gradient of the shape functions calculated in the cell center. When there are more or less particles in a cell than the original amount of

particles per cell, the average density of the cell increases or decreases artificially and hence the internal forces due to the contribution of that cell can be approximated as follows:

$$\mathbf{f}_{i-cell}^{int,k} = - \sum_{p=1}^{N_{p-cell}^k} m_p \mathbf{G}_{ip}^k \mathbf{S}_p^k = - \sum_{p=1}^{N_{p-cell}^k} m_p \mathbf{G}_{ip}^k \frac{\boldsymbol{\sigma}_p^k}{\rho_{cell}^k} = \sum_{p=1}^{N_{p-cell}^k} \mathbf{G}_{ip}^k \boldsymbol{\sigma}_p^k \frac{V_{cell}}{N_{p-cell}^k} = \mathbf{G}_{i0} \sum_{p=1}^{N_{p-cell}^k} \boldsymbol{\sigma}_p^k \frac{V_{cell}}{N_{p-cell}^k} \quad (26)$$

$$\rho_{cell}^k = \frac{N_{p-cell}^k m_p}{V_{cell}}$$

where :

m_p : particle mass

$\mathbf{G}_{ip}^k = \nabla N_i(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_p^k}$: shape function gradient at particle location

\mathbf{G}_{i0} : shape function gradient at cell center

$\mathbf{S}_p^k = \frac{\boldsymbol{\sigma}_p^k}{\rho^k}$: specific stress

ρ_{cell}^k : average cell density at step k

N_{p-cell}^k : number of particles in the cell at step k

V_{cell} : cell volume (constant)

In equation (26) the gradients at particles positions are approximated by the gradient at the center of the element and this is equivalent to the procedure of averaging the stresses of the particles which are in a cell. On the other hand the pressure instability is avoided by imposing a constant pressure increment to the particles in a cell equal to the pressure increment calculated at the element center instead of calculating the pressure increment at positions of the particles (Equation 25).

Figure 1 shows a rigid footing resting on a water-saturated porous material with zero permeability. This problem was solved with 4 particles per cell. The average stress of the particles in the cell and the gradient calculated at the center of the element were used for compute internal forces. Also the particles pressure increment is calculated at the center of the element. A constant vertical velocity was imposed to the particles of the footing. Figure 2.a shows an instant pore pressure distribution for an elastic material and Figure 2.b the same for a Mohr-Coulomb elastic-plastic material with cohesion and friction. The procedure is stable and there are not pressure oscillations. Figure 3 shows contours of equivalent plastic strain for the Mohr-Coulomb material. If the internal forces are calculated using the gradients in the positions of the particles, we obtain the classical pressure instability shown in Figure 4.

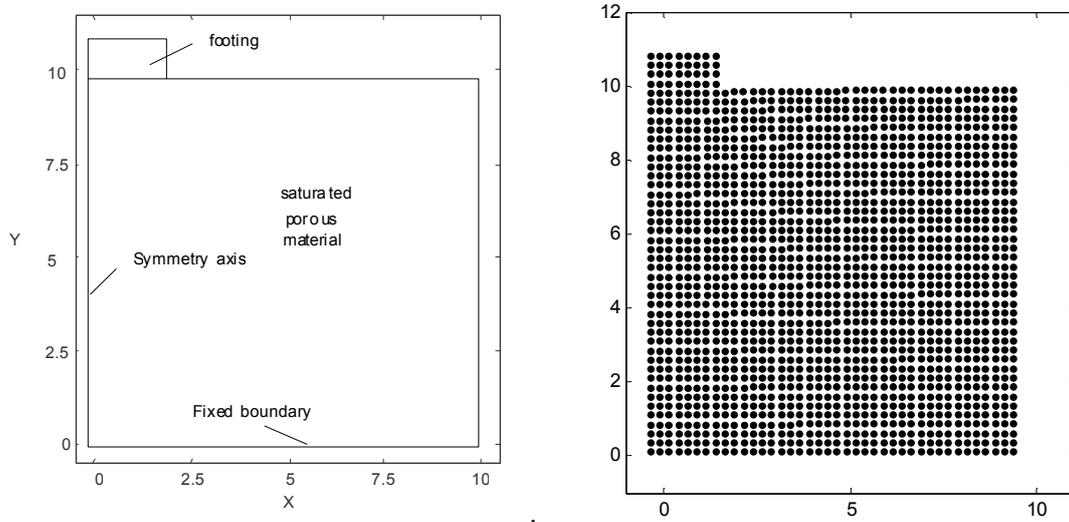


Figure 1. Rigid footing resting on saturated porous material.

a) Geometry b) Particle model.

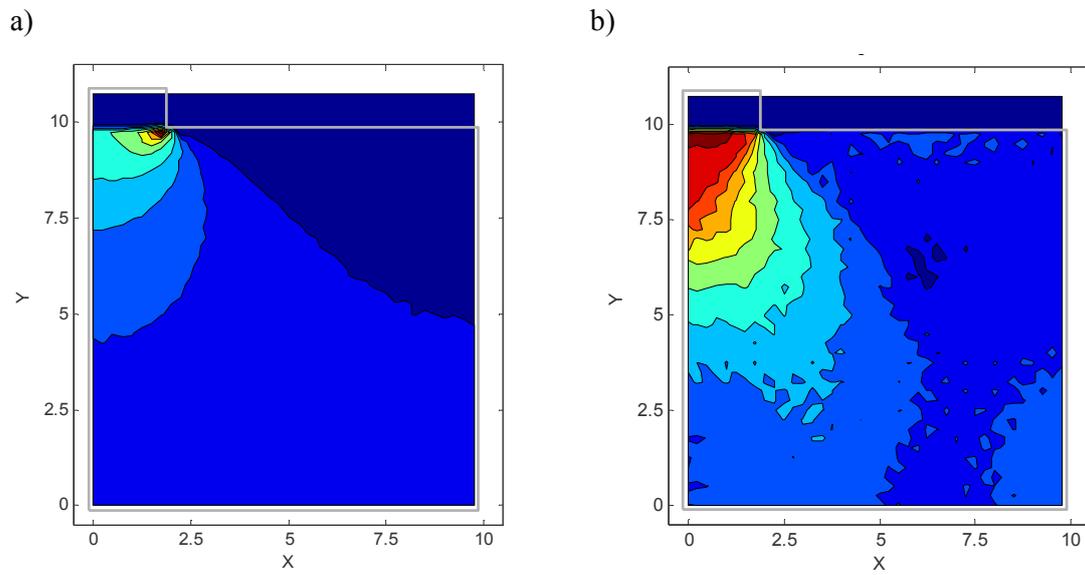


Figure 2. Rigid footing. Pore pressure contours for zero permeability

a) Elastic material b) Mohr Coulomb material.

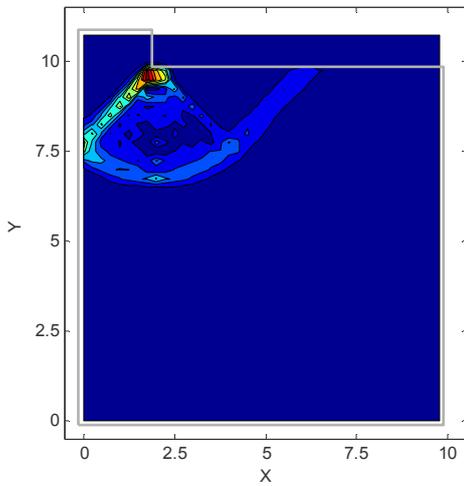


Figure 3. Equivalent plastic shear strain for Mohr Coulomb material

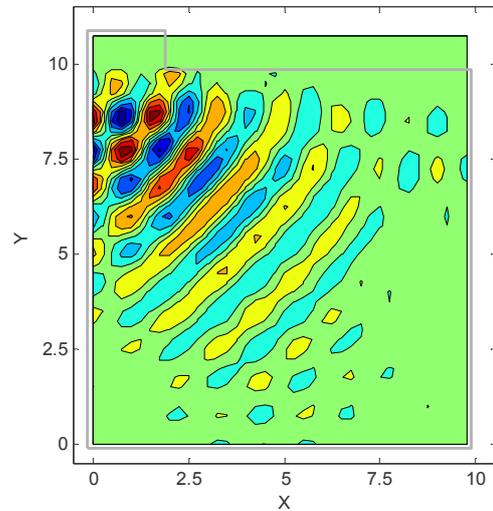


Figure 4. Pressure instability

2.3 Impulsive load applied to a soil column.

This section compares MPM results for a soil column (Figure 5.a) with those obtained by Mira et al. [8] with Q8P4 finite elements (elements with 8 displacement and 4 pressure degrees of freedom) in order to check the stability of the calculation algorithm. Figure 5.b compares results for pressure distribution one second after the application of an impulsive load. Pressure instability occurs if pressure increments at positions of the particles are used for updating the particle pressures.

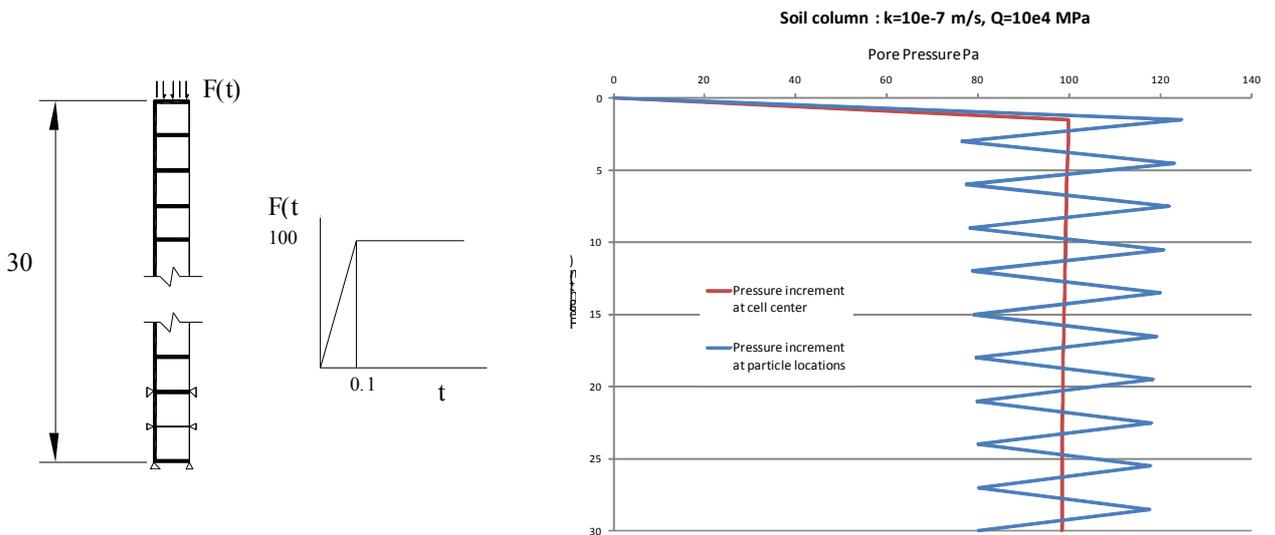


Figure 5. a) Soil column under impulsive load. b) Pore pressure distribution.

3 MODELLING AZNALCOLLAR DAM FAILURE

MPM in its explicit version, including the hydro-mechanical coupling for a water-saturated material, has been used for modeling Aznalcollar dam failure ([9], [10], [3]). It is believed that this dam failed progressively due the fragility of the very low permeability foundation clay, in which high pore pressures persisted during twenty years of construction until failure, and the downstream method of construction which generate high shear stresses under the downstream slope. The dam failed and moved forward as a solid body in the way indicated in Figure 6. The failure was explained by the development of a continuous subhorizontal failure surface located 14 m under the contact between the dam and its foundation at the dam axis.

The construction and filling of the reservoir was simulated in stages (Figure 9). Figure 8 shows the vertical pore pressures distribution for the foundation clay in a zone of the tailings pond, away from the dam, for the 9, 13 and 15 stages of construction. The pressures calculated with MPM are compared with Terzaghi's solution. Also a point is plot, for stage 15, which represents the pressure measured at the approximate depth of the sliding surface in a not failed section located north to the failure zone. A boundary condition of hydrostatic pressure was applied at the bottom of the modeled foundation layer, which is considered fixed during construction. This hypothesis is based on field data from the aquifer that lies under the clay layer. A deep piezometer installed in the sand under the clay measured pore pressure equivalent to the height of the clay layer. The top boundary condition is variable and equal to the height of the tailings, which were always kept under water for each stage. A free water condition was considered in the upper layer of downstream alluvium.

Figure 8 shows contours of equivalent plastic shear strain for different stages of construction. The development of localization bands within the clay foundation and the tailings deposit are shown. The concentration of plastic shearing strains started underneath the downstream slope of the rockfill dam. The shearing band extended first downstream and then in the upstream direction. The shearing band defines a horizontal failure surface that eventually bends upwards below the upstream toe of the dam slope and crosses the tailings' deposit.

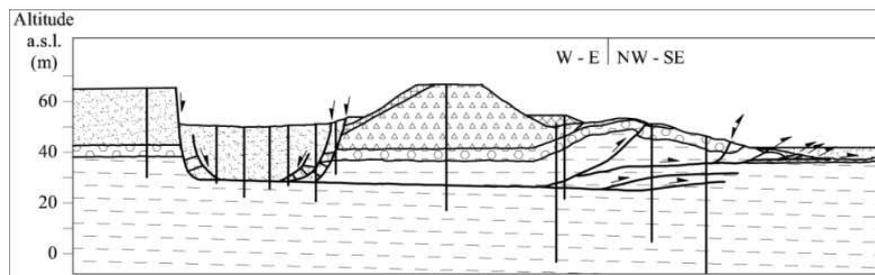


Figure 6. Cross-section of slide. Geometry after the slide, as interpreted from borehole data and surface topography [9]

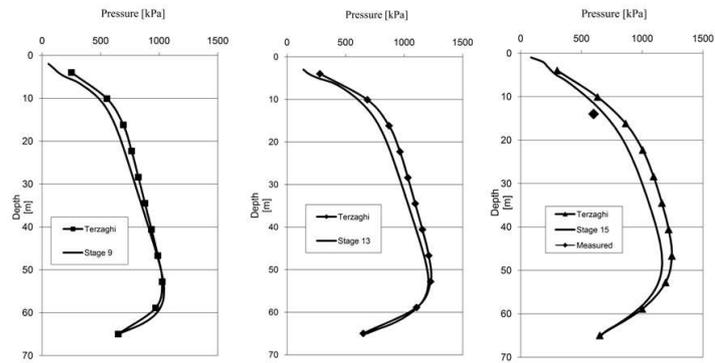


Figure 7. Pore pressure distribution at foundation.

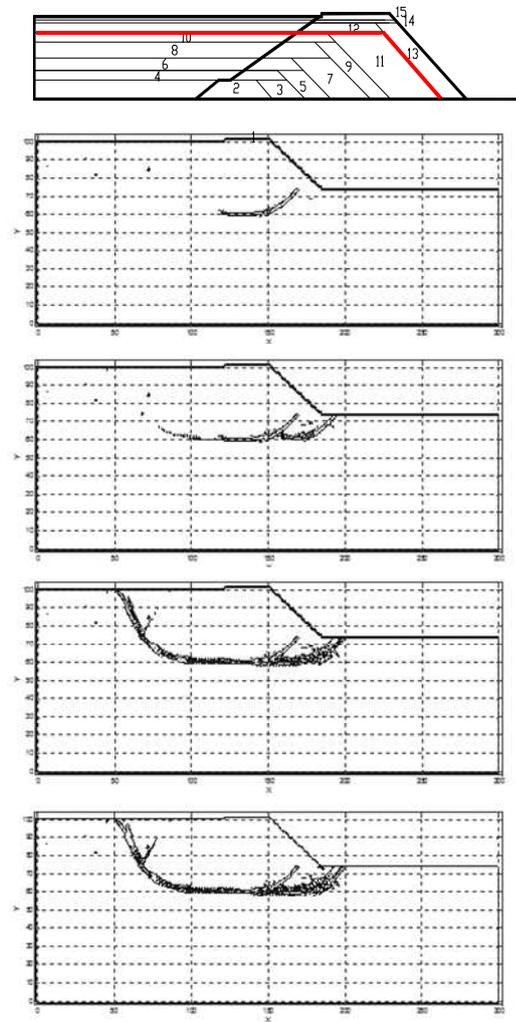


Figure 8. Sequence of contours of equal equivalent plastic strain, 1% and 5%. $c_v = 0.001 \text{ cm}^2/\text{seg}$, $K_0 = 1$

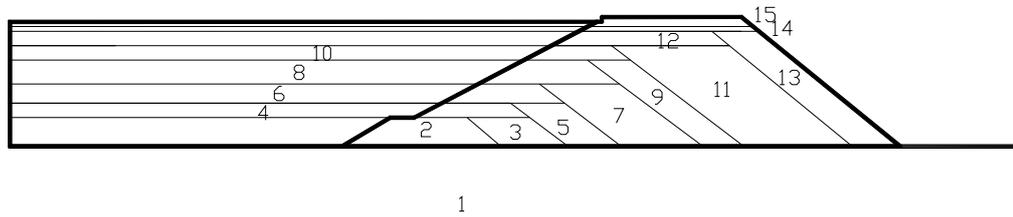


Figure 9. Geometry of the model and construction sequence of the dam

4 CONCLUSIONS

The Material Point Method was extended to model hydro-mechanical geotechnical problems. Pressure instability is circumvented using concepts of the Finite Element Method. A relative large scale problem of dam construction and failure was successfully solved with the procedure shown in this paper.

REFERENCES

- [1] Sulsky, D., Z. Chen, & H. L. Schreyer. A Particle Method for History-Dependent Materials. *Computer Methods in Applied Mechanics and Engineering*. Vol. 118(1), pp 179-96. (1994).
- [2] Bardenhagen, S. G. & E. M. Kober. The Generalized Interpolation Material Point Method. *Computer Modeling in Engineering & Sciences*. Vol.5(6), 477-95.(2004).
- [3] Zabala, F & Alonso, E. E.. Progressive failure simulation of Aznalcollar dam using the Material Point Method. *Géotechnique*, doi: 10.1680/geot. (2011)
- [4] Biot M. A..General theory of three dimensional consolidation. *J. Appl. Phys.* 12, 155-164. (1941)
- [5] Sulsky, D. & H. L. Schreyer. Axisymmetric Form of the Material Point Method with Applications to Upsetting and Taylor Impact Problems. *Computer Methods in Applied Mechanics and Engineering*. 139(1), 409-29. (1996).
- [6] Zienkiewicz O.C., Chang C.T. & Bettess P.. Drained,undrained, consolidating and dynamic behavior assumptions in soils. *Geotechnique* 30, No. 4, 385–395. (1980)
- [7] Zienkiewicz O.C., Shiomi T..Dynamic behavior of saturated porous media: the generalized Biot formulation and its numerical solution, *Int. J. Numer. Anal. Methods Geo-mech.* 8, 71–96. (1984).
- [8] Mira P., Pastor M., Li T., Liu X.. A New Stabilized Enhanced Strain Element with Equal Order of Interpolation for Soil Consolidation Problems. *Computer Methods in Applied Mechanics and Engineering*. 192(37), pp 4257-4277.(2003)
- [9] Alonso, E. E. & Gens, A. (2006a). Aznalcóllar dam failure. Part 1: Field observations and material properties *Geotechnique*, 56, No. 3, 165–183
- [10] Gens A. & Alonso E. E. (2006). Aznalcollar Dam Failure. Part 2: Stability Conditions and Failure Mechanism. *Géotechnique*. 56 No.3, pp 185-201.

NON-ISOTHERMAL COMPOSITIONAL GAS FLOW DURING CARBON DIOXIDE STORAGE AND ENHANCED GAS RECOVERY

A.K. Singh*, N. Böettcher[†], W. Wang *, U.-J. Görke* and O. Kolditz*[†]

*Department of Environmental Informatics
Helmholtz Centre for Environmental Research - UFZ
Permoserstr. 15, D-04318 Leipzig, Germany
e-mail: ashok.singh@ufz.de, www.ufz.de

[†]University of Technology Dresden, Germany

Key words: Non-isothermal compositional gas flow, Enhanced gas recovery, Carbon dioxide sequestration, Joule-Thomson effect, OpenGeoSys project.

Abstract. In this work we present the conceptual modeling and the numerical scheme for carbon dioxide storage into nearly depleted gas reservoirs for enhanced gas recovery reasons. For this we develop non-isothermal compositional gas flow model. We used a combined monolithic / staggered coupling scheme to solve mass balance equation for the gaseous mixture with heat and fractional mass transport equations. Temperature change resulting from fluid expansion and viscous heat dissipation is included in heat transport in addition to advection and conduction. We have used a modified version of the Peng-Robinson equations of state, to determine the density of the real gas mixture along with an empirically extended ideal gas equation. A real behavior of mixture is accounted by using energy and distance parameters.

1 INTRODUCTION

Fossil fuels continue to be the main global energy source. Therefore, squeezing the gas from matured fields will be of positive interest and additionally offering space for injection and storage of CO₂. Sequestration of CO₂ is at present of much importance because CO₂ has been identified to be one of the main greenhouse gases which can potentially lead to global climate change. Nearly depleted natural gas reservoirs are good candidates for storing CO₂ and at the same time recovering the rest gas found in these reservoirs. They offer several advantages for EGR processes namely integrity against gas escape, larger volume capacities, a good reservoir history enhancing its understanding and existing infrastructure. Numerical simulation studies have been performed showing the technical

feasibility of carbon sequestration and enhanced gas recovery (CSEGR) with existing commercial codes (Pusch [1], Oldenburg [2]).

According to Holloway [3] underground storage of industrial quantities of carbon dioxide is technologically feasible. However, combustion of coal produces more carbon dioxide than could be stored in the space from which the coal came from [4, 5]. If one wishes to dispose of carbon dioxide permanently, it has to be confined within depleted hydrocarbon reservoirs or trapped in an aquifer directly analogous to a hydrocarbon reservoir [6]. Hydrocarbon reservoirs have many advantages over aquifers. Firstly, the reservoirs have a seal that has proven capable of retaining liquids or gases for thousands to millions of years. Secondly, some of the equipment installed on the surface or underground for oil or gas recovery may be re-used for carbon dioxide disposal. The underground reservoir needs a sealed trap to prevent the carbon dioxide from reaching the earth's surface due to percolation upwards through inhabitant ground water. This is because supercritical or liquid carbon dioxide is lighter than water and tends to force its way up through the water-saturated rocks.

Starting from appropriate local formulations expressing the balance laws of continuum mechanics, we develop the governing equations for non-isothermal miscible displacement of compressible mixture of nonpolar gases through a non-deformable porous medium. The compressibility of the gaseous mixture is again defined according to the Peng and Robinson [7] equation of state. The mass balance equations are then written to describe the macroscopic behavior of the gaseous mixture with heat transport. The energy conservation equation is solved to account for heat loss due to gas expansion and viscous heat dissipation along with advection and conduction. Macroscopic fractional mass transport equations are then obtained for each component of the mixture. Equation of state for density and viscosity is introduced to obtain the general field variables. We select gaseous mixture pressure p , temperature T and mass-fraction of each component x_k as primary field variables.

A numerical module has been developed for compositional gas flow coupled with heat and fractional mass transport. The finite element module is embedded in the object-oriented framework of the scientific open source code OpenGeoSys ([8, 9]). Therefore, allows an easy extension of the numerical model for multi-dimensional problems. A combined, monolithic and staggered coupling scheme is used (i.e., monolithic for the pressure and temperature fields and staggered for the component mass-fraction) with variable time stepping.

2 GOVERNING EQUATIONS

The differential form of the mass balance equation for the k^{th} component of the gaseous mixture is

$$\frac{\partial (n\rho_k)}{\partial t} + \nabla \cdot [\mathbf{J}_{\text{FA}_k} + \mathbf{J}_{\text{FD}_k}] = Q_{\rho_k} \quad (1)$$

Here, advective mass-flux (mass rate per unit flow area) of the k^{th} component is $\mathbf{J}_{\text{FA}_k} = \rho_k n \mathbf{v}$. In presence of concentration gradients, diffusive mass-flux (mass rate per unit flow area) of the k^{th} component w.r.t. \mathbf{v} (velocity of center of the gravity) is $\mathbf{J}_{\text{FD}_k} = x_k \rho n (\mathbf{v}_k - \mathbf{v})$ where \mathbf{v}_k is the average velocity of the k^{th} component. It is clear from this relationship that sum of diffusive flux over all components (total diffusive mass-flux) is zero. Density ρ , is defined as mass of gas per unit volume which varies with pressure and temperature. Q_{ρ_k} is an external source/sink term due to the k^{th} component.

The linear momentum balance equation for the mixture can be expressed in the form of the extended Darcy's law. This is a reasonable approximation of slow laminar flow arising in reservoir engineering problems. According to Darcy's law, the advective part of the total mass-flux due to all components is $\mathbf{J}_{\text{FA}} = \rho n \mathbf{v}$ and can be written as

$$\rho n \mathbf{v} = -\rho \frac{\mathbf{k}}{\mu} (\nabla p - \rho \mathbf{g}); \text{ where } \mathbf{v} = \sum_k x_k \mathbf{v}_k. \quad (2)$$

Where \mathbf{k} is intrinsic permeability, μ is mixture dynamic viscosity and \mathbf{g} is gravity constant. Summarizing the mass balance Eq. (1) corresponding to each component results in the mass balance equation for the gaseous mixture,

$$\frac{\partial (n\rho)}{\partial t} + \nabla \cdot (\rho n \mathbf{v}) = Q_\rho; \text{ where } \rho = \sum_k \rho_k. \quad (3)$$

Here, ρ is calculated by the extended ideal gas equation which is given by

$$\rho = \frac{pM}{z_{\text{sc}}RT} \quad (4)$$

where M is the molecular weight of the gaseous mixture, R is the universal gas constant; the mixture super compressibility factor z_{sc} is a non-dimensional constant which varies with pressure, temperature and composition.

Following the assumption of local thermal equilibrium, the energy balance equation of porous media can be obtained as

$$\begin{aligned} (\rho c_p)_{\text{eff}} \frac{\partial T}{\partial t} + c_p \rho n \mathbf{v} \cdot \nabla T - \nabla \cdot [\kappa_{\text{eff}} \nabla T] = \\ \underbrace{n \beta_T T \frac{\partial p}{\partial t} + \beta_T T n \mathbf{v} \cdot \nabla p}_{JTC} - \underbrace{n \mathbf{v} \cdot \nabla p}_{VHD} + Q_T \end{aligned} \quad (5)$$

The 1st term of the right hand side is related to Jule-Thomson Cooling (JTC) and the 2nd term is related to Viscous Heat Dissipation (VHD). where κ_{eff} , is the effective thermal conductivity tensor of the porous media, with coordinates defined as $(\kappa_{\text{eff}})_{ij} = (1-n)\kappa_{ij}^s + n\kappa_{ij}$. κ is the thermal conductivity of the gaseous mixture. $(\rho c_p)_{\text{eff}}$ is the effective heat

capacity of the porous medium defined by $(\rho c_p)_{\text{eff}} = (1-n)\rho^s c_p^s + n\rho c_p$. Here, specific heat capacity and thermal conductivity of the gaseous mixture are approximated by averaging over its components.

Analogous to the derivation of the mass balance equation, the transport equation results from the balancing of each component's mass. The characteristic quantity for mass transport is $n\rho x_k$ which is conserved by definition. With macroscopic dispersion, we obtain the mass transport equation for each component in the porous medium,

$$\frac{\partial [n\rho x_k]}{\partial t} + \nabla \cdot \mathbf{J}_{\text{MA}_k} + \nabla \cdot \mathbf{J}_{\text{MD}_k} = Q_{x_k}. \quad (6)$$

where x_k represents the mass-fraction of the k^{th} component of the gaseous mixture. Q_{x_k} is the mass source/ sink term. Advective and diffusive mass-fluxes of the k^{th} components are $\mathbf{J}_{\text{MA}_k} = -(x_k \rho \mathbf{k} / \mu)(\nabla p - \rho \mathbf{g})$ and $\mathbf{J}_{\text{MD}_k} = -n\tau \mathbf{D}_k \nabla (\rho x_k)$, respectively.

Using the componential mass balance Eq. (1) with mass transport Eq. (6), the divergent form of the fractional-mass transport equation is given by

$$n\rho \frac{\partial x_k}{\partial t} + \rho n \mathbf{v} \cdot \nabla x_k + x_k Q_{\rho_k} - \nabla \cdot [n\rho \mathbb{D} \nabla x_k] = Q_{x_k}. \quad (7)$$

3 CONSTITUTIVE RELATIONS

Density and viscosity of the carbon dioxide change abruptly as it passes through critical pressure near critical temperature. Here, we describe the procedure which we have used to calculate the mixture density and viscosity for the necessary range of pressure and temperature.

In reservoir engineering, the empirically extended ideal gas law is well known. Oldenburg et al.[2] have used this in their numerical model for CSEGR. Here, the degree of non-ideality shown by a gaseous mixture can be expressed by the super compressibility factor z_{sc} . For the gaseous mixture z_{sc} is obtained by solving constitutive equations, which are, e.g., derived by inserting $z_{\text{sc}} = pV/(N_n RT)$ in the modified Peng-Robinson equation of state (PR-EOS). For the gaseous mixture, PR-EOS needs critical constants and the acentric factor to get the attraction $a(T)$ and repulsion b parameters values. The critical pressure and temperature of the mixture are approximated as

$$p^{\text{cr}} = M \sum_k x_k M_k^{-1} p_k^{\text{cr}}; \quad T^{\text{cr}} = M \sum_k x_k M_k^{-1} T_k^{\text{cr}} \quad (8)$$

As we have values of $a(T)$ and b , we can solve the following cubic equation analytically.

$$z_{\text{sc}}^3 + z_{\text{sc}}^2 (B - 1) + z_{\text{sc}} (A - 3B^2 - 2B) + (B^3 - B^2 - AB) = 0 \quad (9)$$

where, $A = p a(T)/(RT)^2$ and $B = p b/(RT)$. The parameters $a(T)$ and b are calculated from empirical relations

$$a(T) = 0.4572 \frac{R^2 T^{\text{cr}2}}{p^{\text{cr}}} \left[1 + a_0 \left(1 - \sqrt{\frac{T}{T^{\text{cr}}}} \right) \right]^2; \quad b = 0.07779 \frac{RT^{\text{cr}}}{p^{\text{cr}}} \quad (10)$$

where, $a_0 = 0.3746 + 1.5422\omega - 0.2699\omega^2$; ω is the acentric factor of the gaseous mixture. $a(T)$ accounts for attractive forces between the molecules of the mixture and b takes into account the reduction in free volume of the particular molecule.

The above cubic equation can have three positive roots, with the largest positive root corresponding to the gaseous mixture. To guarantee the single phase i.e., gas phase we choose the reservoir temperature 393.15 K, which is greater than the critical temperature of any component.

The dynamic viscosity of the gaseous mixture is calculated from a temperature and density dependent empirically correlated function. This predicts the viscosity for non-polar gases with an error of $\leq 2\%$.

$$\mu(\rho, T) = d_0 \left(\frac{1}{G_2} + A_6 Y \right) + f_0 A_7 Y^2 G_2 \exp \left(A_8 + \frac{A_9 \epsilon}{kT} + \frac{A_{10} \epsilon^2}{k^2 T^2} \right) \quad (11)$$

where

$$G_2 = \frac{\frac{A_1}{Y}(1 - e^{-A_4 Y}) + A_2 G_1 e^{A_5 Y} + A_3 G_1}{A_1 A_4 + A_2 + A_3}; \quad G_1 = \frac{1 - 0.5Y}{(1 - Y)^3} \quad Y = 1.0 \times 10^{-3} \frac{\rho V^{\text{cr}}}{M};$$

$$d_0 = 4.0785 \times 10^{-7} \frac{(MT)^{\frac{1}{2}}}{V^{\text{cr} \frac{2}{3}} \Omega} (1 - 0.2756\omega); \quad f_0 = 3.6344 \times 10^{-7} \frac{(MT)^{\frac{1}{2}}}{V^{\text{cr} \frac{2}{3}}}$$

where k is the Boltzmann constant and the reduced collision integral Ω depends on the intermolecular potential. Empirical equation for Ω and constants $A_0 - A_{10}$ are given in [10]. The acentric factor and the molecular weight of the mixture are calculated by

$$\omega = \frac{1}{\sigma^3} \sum_i \sum_j x_i x_j (\omega_i \omega_j)^{\frac{1}{2}} (\sigma_i \sigma_j)^{\frac{3}{2}} \quad (12)$$

and

$$M^{\frac{1}{2}} = \frac{1}{\epsilon \sigma^2} \sum_i \sum_j x_i x_j \sigma_i \sigma_j (\epsilon_i \epsilon_j)^{\frac{1}{2}} \left(\frac{2M_i M_j}{M_i + M_j} \right)^{\frac{1}{2}}. \quad (13)$$

Here, i and j stand for the i^{th} and j^{th} components.

The parameters of the intermolecular potential for the gaseous mixture, σ and ϵ , are approximated by averaging over its components.

$$\sigma^3 = \sum_i \sum_j x_i x_j (\sigma_i \sigma_j)^{\frac{3}{2}}; \quad \epsilon = \frac{1}{\sigma^3} \sum_i \sum_j x_i x_j (\sigma_i \sigma_j)^{\frac{3}{2}} (\epsilon_i \epsilon_j)^{\frac{1}{2}} \quad (14)$$

σ and ϵ are the parameters of the intermolecular potential describing the interaction between molecules of components of the mixture and are expressed, respectively, in Angstrom units and degrees Kelvin. An approximation of these two constants for a respective component is provided by $\epsilon_k = 0.77kT_k^{\text{cr}}$ and $\sigma_k = 0.0833 \left(\frac{V_k^{\text{cr}}}{M} \right)^{\frac{1}{3}}$.

To obtain the mixture viscosity at any temperature and any pressure, we first calculate the corresponding density of the mixture by using Eq. (4). This requires the values of M, σ, ϵ , and ω , which can be obtained from Eqs. (12)-(14) by using the critical constants and acentric factor of the components. Finally the mixture viscosity is calculated using Eq. (11), where the mixture critical volume V^{cr} is estimated from the relation $V^{\text{cr}} = M \left(\frac{\sigma}{0.0833} \right)^3$. Material parameters used in this simulation are given in Table 1.

Table 1: Material parameters

Symbol	Meaning	Value	Unit
h_p	Layer thickness	4, 6, 19	m
l_p	Reservoir column length	1000	m
z_{sc}	Super compressibility factor	Max root of Eq. (9)	—
μ	Dynamic viscosity	Eq. (11)	Pas
ρ	Density	Eq. (4), 2460, 2700, 2940	kg m ⁻³
c_p	Heat capacity	1839.3, 960, 920, 880	Jkg ⁻¹ K ⁻¹
κ	Thermal conductivity	0.0466, 3.0, 2.5, 2.2	Wm ⁻¹ K ⁻¹
M	Mixture molecular weight	Eq. (13)	kg kmol ⁻¹
\mathbf{k}	Intrinsic permeability	$1.0 \times 10^{-17, -18, -20}$	m ²
n	Porosity	0.085, 0.014, 0.01	—
τ	Tortuosity	$n^{\frac{1}{3}}$	—
D_0	Self diffusion coefficient	0.965, 2.0, 1.78×10^{-5}	m s ⁻²
α_L, α_T	Dispersivity	0.1, 0.01	m
ϵ	Energy parameter	$0.77kT^{\text{cr}}$	m ² kgs ⁻²
σ	Distance parameter	$0.0833 \left(\frac{V^{\text{cr}}}{M} \right)^{\frac{1}{3}}$	Å
ω	Acentric factor	0.228, 0.0114, 0.04	—

4 NUMERICAL SCHEME

The numerical treatment of the coupled problem of gas flow in undeformable porous media is based on the governing field equations together with discretization methods in the space and time domains. The method of weighted residuals is applied to derive the weak formulations of all the governing equations given. Within the framework of a standard Galerkin procedure, a spatially discretized formulation of the weak forms is defined in the finite element space.

For the required time discretization of the global system of governing equations we use the generalized first order difference scheme. The nonlinear coupled boundary value problem is solved iteratively using the Picard linearization within the context of the finite element method [9]. We adopt a staggered scheme without losing the accuracy of the solutions. Within this context, at first the gas flow problem is solved, then heat and

mass transport problem are solved separately considering the results from the gas flow. This iteration cycle is repeated until the converged solution of all primary variables is achieved. For the sake of simplicity, at the current stage, gas flow to transport coupling is conducted in one way only, i.e. pore pressure changes influence transport, without the inverse feedback.

5 BENCHMARK DEFINITION

In the present study of non-isothermal compressible flow, we consider three main components (carbon dioxide, nitrogen and methane), which appear in the natural gas mixture. The accurate calculation of density of nonpolar light hydrocarbon mixtures is done by using a modified version of the Peng-Robinson equations of state. The pressure provided by carbon dioxide during expansion upon natural gas, enhances gas recovery. Concentration gradients in the gas phase lead to mixing by molecular diffusion-dispersion. The extent to which carbon dioxide could mix with natural gas is a primary concern for technical and economic feasibility of carbon sequestration with enhanced gas recovery. Pressure gradients arise due to mixing and introduce minor advective flow at rates depending on the permeability and porosity of the porous reservoir medium.

5.1 System geometry and conditions

- Geometry: We used middle and bottom layers as reservoir and top one for seal. Injection and production well are 500 m apart. 2-dimensional depleted gas reservoir model is taken in $r - z$ plane which is 19 m thick in z -direction and 1000 m long in r -direction (see Fig. 1).
- IC: We assume that caprock-reservoir system is filled with 75% methane and 25% nitrogen at pressure of 4.0×10^6 Pa and temperature 393.15 K.
- BC: Through injection point I, we perform injection of CO_2 gas with pressure 6.0×10^6 Pa and temperature 353.15 K. At production running well at pressure 3.0×10^6 Pa, we are producing natural gas which is a mixture of 75% methane and 25% nitrogen at temperature 393.15 K.

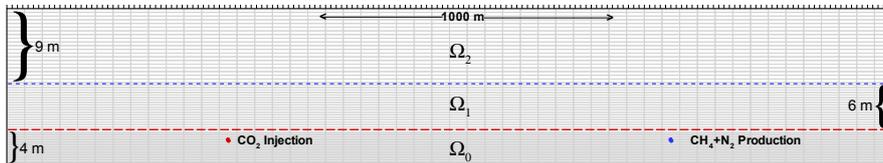


Figure 1: Model setup for CSEGR and grid arrangement

5.2 Results

In Figs. 2a, b we have shown the time evolution of carbon dioxide mass fraction and temperature at certain observation points. The observation points are O_1 and O_2 , which are at the perimeter of circle around the injection point I. Larger value of x_{CO_2} and temperature at observation point below to injection point than above is shown by Figs. 2a, b, and is due to advection enhanced by gravity force. At beginning of the simulation rate of gas expansion is high due to relatively low formation pressure. Hence advective transport plays role to make steep rise of mass fraction and temperature in Figs. 2a, b. As injection and formation pressure gets in equilibrium then advection is not so extensive, hence after steep rise temperature and mass fraction versus time curve showing steady nature for both observation points. During this time transport by diffusion-dispersion is dominated over advection.

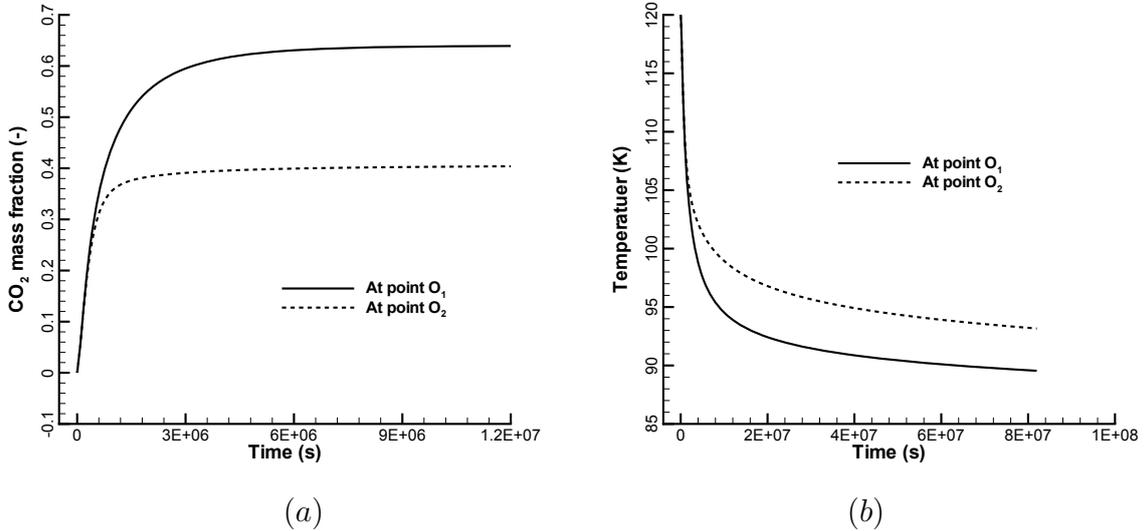


Figure 2: Time evolution of CO₂ mass fraction and temperature at points O_1 and O_2 located equidistantly (1 m) above and below to injection point, respectively

Figs. 3a, b show the distribution of carbon dioxide and methane mass-fractions for four different times. Multi-component transport in the gas mixture is mainly by advection and diffusion whereas dispersion is due to random motion of molecules. Advection is driven by pressure gradients as well as buoyancy forces (in the case when fluid density differences are significant), and diffusion depends on concentration gradients. From Fig. 3a it can be seen that near the injection point advective mass-fluxes are dominating over diffusive ones because of high pressure gradients for compressible flow. At the same time, due to the larger density of injected carbon dioxide than natural gas, gravity makes safe disposal of carbon dioxide in such reservoir caprock system.

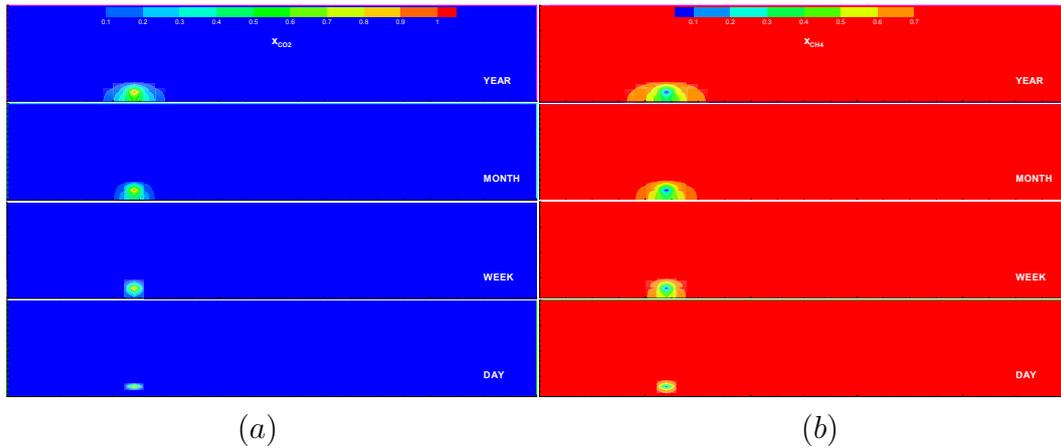


Figure 3: Mass fraction distribution of (a) CO_2 ; (b) CH_4 in the permeable layers of a gas reservoir

Figs. 4a, b depict the evolution of pressure and temperature distribution in the reservoir layer in the sequence of CO_2 injection through point I. As gas is compressible, transient flow is developing at the beginning; simulation shows a symmetry pattern of pressure and temperature distribution, respectively around injection point I (see Figs. 4a, b). Distribution pattern for long duration showing that gravity making the pressure distribution asymmetric and a positive pressure gradient is developed in negative z-direction. Therefore, mass and heat transport in bottom reservoir layer is dominated over the middle one due to gravity (see Figs. 3a, b). This supports the strategy to store carbon dioxide along the lower surface of the bottom reservoir layer and producing natural gases from above reservoir layers

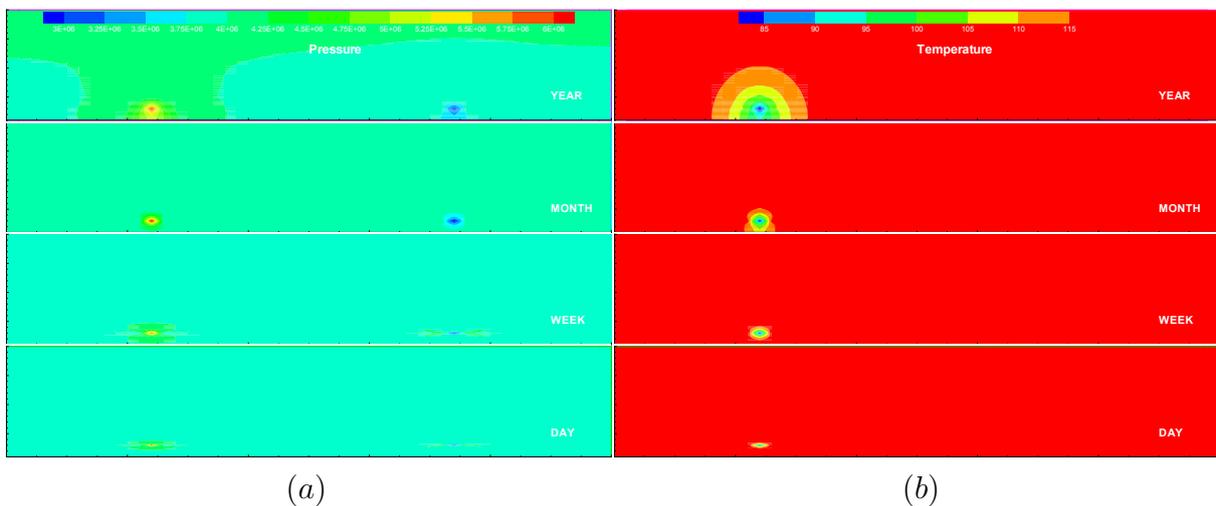


Figure 4: Evolution of (a) gas pressure and (b) temperature distribution through the permeable layers of a depleted gas reservoir

In Fig. 5 we present carbon dioxide mass fraction profile along vertical line through injection point I. Figure shows that at lower surface of the bottom reservoir layers (where injection is performed) most of carbon dioxide is disposed. This figure shows some sharp discontinuity along interface of reservoir layers Ω_0 and Ω_1 . This is due to lower permeability of layer Ω_1 , and carbon dioxide can not penetrate into it easily. We have found that in the caprock layer (least permeable), i.e. Ω_2 , after one year maximum carbon dioxide mass fraction is 0.001 and in the reservoir layer Ω_0 (most permeable) is 0.366. This shows present caprock-reservoir system is safe for carbon dioxide disposal.

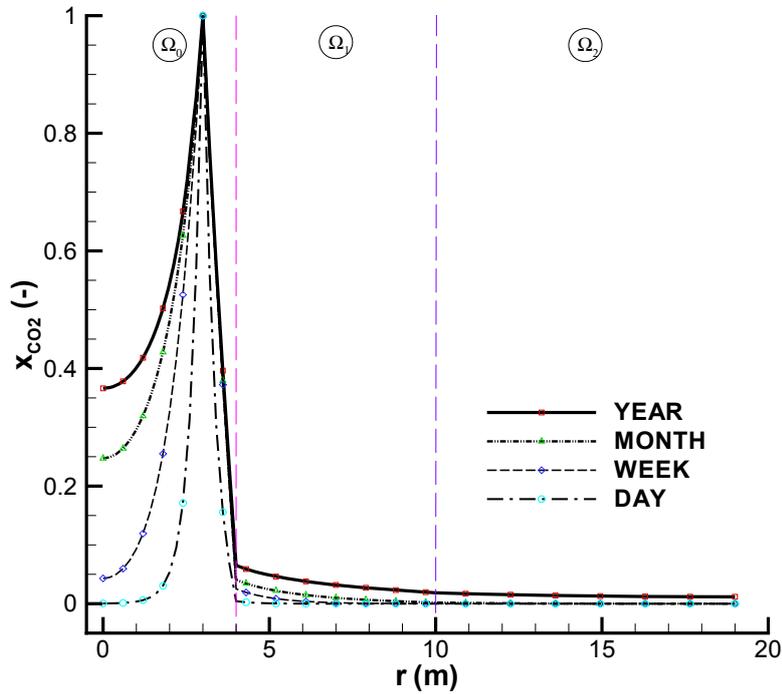


Figure 5: Profile of the CO₂ mass fraction along the vertical line passing through injection point

6 CONCLUSIONS

In this paper, we have demonstrated that the benchmark was carefully defined to analyze carbon dioxide disposal coupled with enhanced natural gas recovery into depleted gas reservoir system. We provide the following conclusions.

- The governing equations for the compositional gas flow model presented here include the mass balance Eq. (3) for the gaseous mixture along with heat and mass transport

equation according to Eq. (5) and Eq. (7), respectively. Equation of state for considered mixture of nonpolar gases at high pressure has been used.

- We used a combined monolithic / staggered coupling scheme, i.e. monolithic for the gas flow and staggered for transport processes.
- Non-isothermal compositional gas flow model developed here works for a pT range up to 7.0×10^6 Pa and 393.15 K. This allows model applications for the calculation of gaseous mixture properties in the vicinity of heat loss due to expansion of injected carbon dioxide.
- Properties of CO_2 and CH_4 are favorable for repressurization without extensive mixing over time scales of practical interest. Simulations of the process of CO_2 injection, into a depleted gas reservoir carried out with OGS confirm the plausibility of CSEGR as a way to CO_2 sequestration while enhancing natural gas recovery.

7 ACKNOWLEDGEMENTS

The authors acknowledge the funding by the German Federal Ministry of Education and Research (BMBF) within the framework of joint projects as parts of the Special Programme GEOTECHNOLOGIEN.

REFERENCES

- [1] Pusch, G. Feasibility Study on the Potential of CO_2 Storage for Enhanced Gas Recovery in Mature German Gas Reservoirs (CSEGR). *Final report of BMBF project* (2008) **03G0627A**:3323–3341.
- [2] Oldenburg, C.M. Carbon dioxide as cushion gas for natural gas storage, *Energy Fuels* (2003) **17**:240–246.
- [3] Holloway, S., Pearcea, J.M., Hardsa, V.L., Ohsumib, T. and Galec, J. Natural emissions of CO_2 from the geosphere and their bearing on the geological storage of carbon dioxide. *Energy* (2007) **32**:1194–1201.
- [4] Khadse, A., Qayyumi, M., Mahajani, S. and Aghalayam, P. Underground coal gasification: A new clean coal utilization technique for India. *Energy* (2007) **32**:2061–2071.
- [5] Garg, A. and Shukla, P.R. Coal and energy security for India: Role of carbon dioxide (CO_2) capture and storage (CCS). *Energy* (2009) **34**:1032–1041.
- [6] Tore, T.A. and Gale, J. Demonstrating storage of CO_2 in geological reservoirs: The Sleipner and SACS projects. *Energy* (2004) **29**:1361–1369.
- [7] Peng, D.Y. and Robinson, D.B. A new two-constant equation of state. *Ind. Eng. Chem. Fundam.* (1976) **15**:59–64.

- [8] Kolditz, O. and Bauer, S. A process-oriented approach to computing multi-field problems in porous media. *J. Hydroinf.* (2004) **6**:225–244.
- [9] Wang, W. and Kolditz, O. Object-oriented finite element analysis of thermo-hydro-mechanical (THM) problems in porous media. *Int. J. Numer. Methods Eng.* (2007) **69**:162–201.
- [10] Chung, T.H., Ajlan, M., Lee, L.L. and Starling, K.E. Generalized multiparameter correlation for nonpolar and polar fluid transport properties . *Ind. Eng. Chem. Res.* (1988) **27**:671–679.

NUMERICAL ANALYSIS OF CO₂ INJECTION INTO DEFORMABLE SALINE RESERVOIRS: BENCHMARKING AND INITIAL OBSERVATIONS

JOSHUA TARON^{*}, CHAN-HEE PARK[†], UWE-JENS GÖRKE^{*}, WENQING WANG^{*},
OLAF KOLDITZ^{*,‡}

^{*} Department of Environmental Informatics
Helmholtz Center for Environmental Research - UFZ
Permoserstr. 15, D-04318 Leipzig, Germany
e-mail: joshua.taron@ufz.de, www.ufz.de

[†] Geothermal Resources Department
Korea Institute of Geoscience and Mineral Resources – KIGAM
92 Gwahang-no, Yuseong-gu, Daejeon 305-350, Korea

[‡] University of Technology Dresden, Germany

Key words: Multiphase flow, Deformation, CO₂ Reservoir, Carbon Capture and Storage, Coupled Problems, Hydro-mechanical.

Abstract. A numerical scheme is presented for the solution of coupled multiphase hydro-mechanical problems in deformable porous media. Model verification is conducted against analytical solutions for multiphase flow with capillarity and coupled multiphase hydro-mechanical consolidation. A hybrid monolithic(flow)-staggered(mechanical) numerical solution scheme is verified to be stable for real materials, provided proper error control is placed on the hydraulic to mechanical iteration and the time-stepping scheme. Initial results of CO₂ injection into an aquifer-caprock system do not show significant differences in CO₂ migration rate between flow-only and hydro-mechanical simulations for conservative injection scenarios. However, the results highlight important regions in the reservoir with regard to potential mechanical failure and caprock integrity and suggest the need for further analysis.

1 INTRODUCTION

Long-term storage of CO₂ within deep geological reservoirs, saline aquifers or otherwise, is increasingly cited as a promising method for the global reduction of greenhouse gas emissions. However, the behavior of such reservoirs during CO₂ injection is not well understood in all of the relevant physico-chemical aspects necessary to ensure safe disposal. In addition to complexities associated with modeling individual processes, such as multi-phase hydraulic transport, chemical dissolution and transport, or mechanical deformation, the interaction between such processes is often complex and non-linear.

Being less dense than water, supercritical CO₂ requires an impermeable seal (caprock) above the target reservoir to prevent escape to the surface. Over the vast areas that injected CO₂ will migrate during long term injection, potentially greater than 100 km² for a 1000-MW coal-fired power plant injecting over 30 years into a 100m thick reservoir [1], the caprock is extremely unlikely to exhibit homogeneity or a lack of potential escape conduits such as faults or large fracture zones.

Beyond uncertainty in initial state, the aquifer-caprock system will be prone to dynamic alteration in permeability and porosity resulting from hydraulic and thermal stimulation. Mechanical failure will be an important consideration where elevated shear stresses are likely, such as at boundaries of materials with strongly different properties (such as the reservoir caprock interface), and such failure has direct consequences on the integrity of the system.

This paper will be one in a series of which examine development and application of a large scale numerical model (OpenGeoSys, [2]) for analysis of these situations. To have sufficient confidence in the ability of a numerical model to answer these very complex questions, it is first necessary to build confidence in the model in more fundamental respects. In this paper we focus on model verification for coupled multi-phase hydromechanical (H2M) problems. Several benchmarks test the individual and combined processes against analytical accuracy. Stability is discussed in relation to time-stepping schemes and inter-process coupling tolerance. A simple reservoir-caprock system is examined to highlight some important differences between typical multi-phase flow simulations and those that include contribution from mechanical deformation and also to examine mechanical integrity of the storage system.

2 GOVERNING EQUATIONS

The governing field equations for modeling multiphase flow in deformable porous media are formulated based on local mass and momentum balance relations of the constituents. The general mass balance of a component, κ , may be written (neglecting diffusive/dispersive flux),

$$\frac{\partial(\phi S_{\kappa} \rho_{\kappa})}{\partial t} + \nabla \cdot (\phi S_{\kappa} \rho_{\kappa} \mathbf{v}_{\kappa}) = 0, \quad (1)$$

in terms of the porosity, ϕ , saturation, S_{κ} , density, ρ_{κ} , and phase velocity, \mathbf{v}_{κ} . Expanding the time derivative term and utilizing the material time derivative of a component relative to the motion of the deformable solid,

$$\frac{d}{d^s t} = \frac{\partial}{\partial t} + \mathbf{v}_s \cdot \nabla, \quad (2)$$

introduces the relationship,

$$\phi \frac{d(S_{\kappa} \rho_{\kappa})}{d^s t} + S_{\kappa} \rho_{\kappa} \frac{d\phi}{d^s t} + \phi S_{\kappa} \rho_{\kappa} \nabla \cdot \mathbf{v}_s + \nabla \cdot q_{\kappa}^r = 0, \quad (3)$$

where the flux, q_{κ}^r , is relative to solid motion. The Lagrangian form (utilizing Eq. (2)) of solid mass balance is,

$$\frac{d(1-\phi)\rho_s}{d^s t} + (1-\phi)\rho_s \nabla \cdot \mathbf{v}_s = 0, \quad (4)$$

from which the expanded storage term yields the porosity derivative [3],

$$\frac{d\phi}{d^s t} = \frac{(1-\phi)}{\rho_s} \frac{d\rho_s}{d^s t} + (1-\phi) \nabla \cdot \mathbf{v}_s. \quad (5)$$

Substituting this into Eq. (3) yields,

$$\phi \frac{d(S_\kappa \rho_\kappa)}{d^s t} + S_\kappa \rho_\kappa \left[\nabla \cdot \mathbf{v}_s + \frac{(1-\phi)}{\rho_s} \frac{d\rho_s}{d^s t} \right] = -\nabla \cdot \mathbf{q}_\kappa^r, \quad (6)$$

where an assumption of small strain negates the resulting two terms: $\mathbf{v}_s \cdot \nabla S_\kappa \rho_\kappa$ and $\mathbf{v}_s \cdot \nabla \rho_s$. A Biot formulation is utilized to represent the solid density time derivative [cf. 4] and Eq. (6) is divided into two fluids; a wetting fluid (subscript w) and a non-wetting fluid (subscript nw). Algebraic manipulations target three primary variables for the numerical solution; wetting fluid pressure, P_w , non-wetting fluid saturation, S_{nw} , and the solid displacement vector, \mathbf{u} :

$$S_w \left[\frac{\phi}{K_w} + \frac{\alpha - \phi}{K_g} \right] \frac{dP_w}{d^s t} - \phi \frac{dS_{nw}}{d^s t} + \alpha S_w \nabla \cdot \frac{d\mathbf{u}}{d^s t} = -\nabla \cdot \left[\frac{\mathbf{k}k_w^r}{\mu_w} (-\nabla P_w + \rho_w \mathbf{g}) \right] \quad (7)$$

and

$$S_{nw} \left[\frac{\phi}{K_{nw}} + \frac{\alpha - \phi}{K_g} \right] \frac{dP_w}{d^s t} + \phi \left[1 - \frac{S_{nw}}{K_{nw}} \frac{dP_c}{dS_w} \right] \frac{dS_{nw}}{d^s t} + \alpha S_{nw} \nabla \cdot \frac{d\mathbf{u}}{d^s t} = -\nabla \cdot \left[\frac{\mathbf{k}k_{nw}^r}{\mu_{nw}} \left(-\nabla P_w - \frac{dP_c}{dS_w} \nabla S_{nw} + \rho_{nw} \mathbf{g} \right) \right] \quad (8)$$

Relative fluid flux (the right hand side term) is obtained from Darcy's law. The solid displacement derivative is $d\mathbf{u}/d^s t = \mathbf{v}_s$, and variables are the intrinsic permeability tensor, \mathbf{k} , viscosity, μ_κ , fluid density, ρ_κ , gravity vector, \mathbf{g} , solid *grain* modulus, K_g , and Biot's alpha $\alpha \approx 1 - K/K_g$, where K is the solid bulk modulus. Fluid pressure is related to the capillary pressure, P_c , as $P_w + P_c = P_{nw}$. The bulk modulus, K_κ , is, by definition,

$$1/K_\kappa \equiv \frac{1}{\rho_\kappa} \frac{\partial \rho_\kappa}{\partial p_\kappa} \bigg|_{\partial T / \partial t = 0}. \quad (9)$$

The third governing equation (for the H2M problem) is given by linear momentum balance on the solid mixture (stress equilibrium equation),

$$\nabla \cdot (\boldsymbol{\sigma} - \bar{P}\mathbf{1}) + \rho_m \mathbf{g} = 0, \quad (10)$$

for the total stress tensor, $\boldsymbol{\sigma}$, with an appropriate stress/strain constitutive relationship, $\boldsymbol{\sigma} = \mathbf{D}\boldsymbol{\varepsilon}$, in terms of the strain tensor $\boldsymbol{\varepsilon} = \frac{1}{2}(\nabla\mathbf{u} + (\nabla\mathbf{u})^T)$. Effective fluid pressure is defined as $\bar{P} = S_w P_w + S_{nw} P_{nw}$ and the mixture density is $\rho_m = \phi(S_w \rho_w + S_{nw} \rho_{nw}) + (1 - \phi) \rho_s$. Equations (7) and (8) are solved globally, with an iterative coupling to Eq. (10). Alternate assumptions are sometimes used for the mean fluid pressure, such as the perfect wetting fluid assumption (all grains are at all times perfectly wetted by the wetting fluid), $\bar{P} = P_w$.

2.1 Saturation equations

Relative permeability and capillary pressure are defined utilizing the Brooks-Corey relationship. For the effective saturation, $S_e = (S_w - S_w^r) / (1.0 - S_w^r - S_{nw}^r)$, these relationships are:

$$\begin{aligned} k_w^r &= S_e^{3+2/m} \\ k_{nw}^r &= (1 - S_e)^2 (1 - S_e^{1+2/m}), \\ P_c &= P_b S_e^{-1/m} \end{aligned} \quad (11)$$

for the entry pressure, P_b , and the residual saturations, S_{κ}^r .

3 NUMERICAL METHOD

Simulations utilize the open source numerical simulator OpenGeoSys (OGS) [2], in continued development by the authors. Weak formulations of the above governing equations (Eqs. (7), (8), and (10)) are derived using the method of weighted residuals. A standard Galerkin procedure is followed, multiplying the equations by arbitrary test functions and integrating over the domain of interest. Time discretization is designed as a generalized first order difference scheme, but all simulations in this paper utilize a fully implicit Euler scheme. The fluid equations (7 and 8) are solved globally in a single matrix, for a resulting non-linear system that is iterated with a Picard linearization. Coupling to the solid equation (and vice versa) is performed with an iterative linking of this global equation (7 and 8) to that of the solid (10). Iteration is performed until a tolerance is met that defines stabilization of error between the solid and fluid system. Such an error tolerance is also a reasonable foundation to build an adaptive time stepping scheme. This is briefly discussed in the following section.

3.1 Tolerance and stability

Convergence of the iteration between the solid and fluid system can depend on error reduction in any of the three primary variables, and potentially a third: the capillary pressure. As the fluid scheme utilizes saturation as a primary variable, the 0 to 1 bounds introduce potential instability in the system. As wetting saturation falls to near the wetting residual saturation, the non-linear relationship to P_c generates a rapid increase in P_c , and thus a rapid mean pressure response. Therefore, iteration and/or time stepping based on P_w and S_{nw} are inadequate in this case, as they will not recognize the rapid system change. We find the most stable method to define error and time stability in the fully H2M system to be that of the mean

pressure. Therefore, both H2 \leftrightarrow M coupling and adaptive time stepping can be stably confined by controlling equilibration of the mean pressure, $\bar{P} = P_w S_w + P_{nw} S_{nw} = P_w + P_c S_{nw}$.

The semi-staggered solution also has potential stability issues dependent upon material properties. When the fluid becomes highly incompressible relative to the solid (stiff system) the solution will fail. This is best defined relative to the Skempton coefficient, B (Table 1), a lower value of B ensures greater stability and we have observed that this criterion is generally independent of loading rate. The Skempton coefficient, $B = -d\bar{P} / d\sigma_m|_{\zeta=0}$ is, in fact, a perfect criterion, as it is a direct measure of strength of coupling between the solid and fluid system. In fact, in undrained coupling methodologies (cf. [5]), this is the coupling linkage between Eqs. (7-8) and (10). In 1-D, as in the analytical solution below, it is the 1-D Skempton coefficient, B_v , that measures strength of coupling and thus stability.

For real systems, where fluids and solid grains are compressible, we have experienced no trouble. It is the introduction of incompressible fluids that tend to cause instability. It is none-the-less important for a given problem and set of solid/fluid properties to examine stability with appropriate benchmarks (as provided below) before extending to the full system.

4 MODEL VERIFICATION

4.1 Two-phase flow

Validation of the flow component is provided by a 1-D solution to the incompressible fluid mass balance equation [6, 7],

$$\phi(1 - S_w^r - S_{nw}^r) \frac{\partial S_\kappa}{\partial t} = -q_\kappa(t, 0) \frac{\partial f_\kappa(S_w)}{\partial x} + \frac{\partial}{\partial x} \left(D_\kappa(S_w) \frac{\partial S_\kappa}{\partial x} \right) \quad (12)$$

where f_κ is a fractional flow function,

$$f_\kappa(S_w) = \frac{\lambda_\kappa(S_w)}{\lambda_w(S_w) + \lambda_{nw}(S_w)} \quad (13)$$

for the mobility, $\lambda_\kappa = k_\kappa^r / \mu_\kappa$. And D_κ is a diffusivity function ($D_{nw} = -D_w$),

$$D_w(S_w) = \frac{\lambda_w(S_w) \lambda_{nw}(S_w)}{\lambda_w(S_w) + \lambda_{nw}(S_w)} \frac{dP_c}{dS_w}. \quad (14)$$

and where the source term is required to take the form $q_\kappa(t, 0) = At^{-1/2}$ for $A > 0$. McWhorter and Sunada [7] provided the first analytical solution to this relationship, and a nicely modified improved form was presented by [6]. It is this modified form that we utilize in our comparison. See [6] for the semi-analytical methodology required to generate the solution. A 1-D flow problem is set up with an initially high wetting saturation prescribed in the domain. A source injection of non-wetting fluid, $q_\kappa(t, 0)$, is applied to the leftmost inlet and saturation monitored in time. The results are provided in Figure 1. It is quite difficult to represent the sharp front dictated by the analytical solution, which would require a very tight spatial discretization. None-the-less, the solution is quite agreeable.

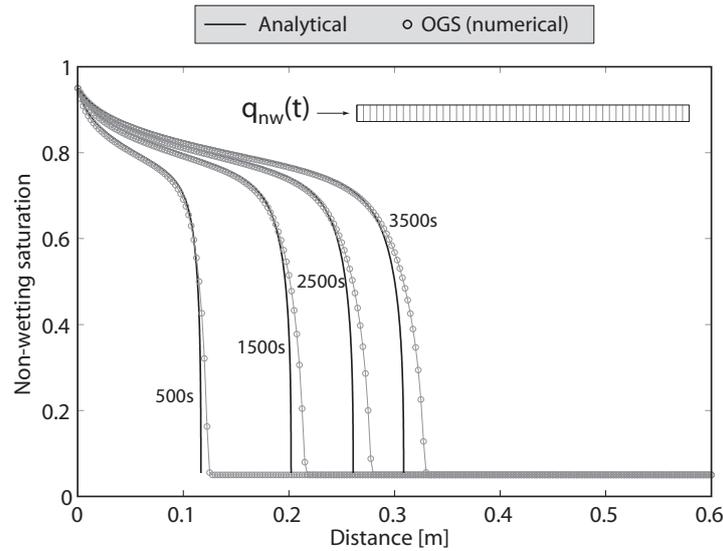


Figure 1. Saturation profiles for the numerical and analytical solutions of the 1-D two-phase flow problem.

Note: A simpler analytical solution is available if $P_c = 0 = \text{constant}$ is prescribed (Buckley-Leverett equation), but a solution that considers P_c variation is strongly desirable for code validation. Furthermore, because the dP_c/dS_w term vanishes in Eq. (8), and this term represents a stabilizing diffusive term in the relationship (see Eq. (14)), we cannot achieve a stable solution to the $P_c = 0$ problem without an upwinding scheme. An upwinding scheme was implemented for this purpose and achieved accurate reproduction of the $P_c = 0$ analytical solution, but those results are not presented here. All results shown in this paper do *not* utilize an upwinding scheme.

4.2 H2M coupling

In this section we test the fundamental premise and validity of a coupling between fluid flow and mechanical deformation. Mechanical compression generates a fluid pressure response, while pressure storage and dissipation affects the mechanical condition via the effective stress. Terzaghi has provided the framework to test such a problem.

This problem tests the fundamental linkage within a hydro-mechanical coupling. It is a convenient test of both the deformation and flow modules but most importantly guarantees that the coupling is correct between them. Without it, H2M (or HM) coupling does not exist. It is necessary for this benchmark to define the composite fluid bulk modulus,

$$\frac{1}{K_f} = \frac{S_w}{K_w} + \frac{S_{nw}}{K_{nw}}. \quad (15)$$

This relationship is accurate for immiscible (or slowly miscible) fluids without penetrating bubbles and allows us provide an analytical solution for H2M where none would otherwise be available.

Table 1. Poroelastic quantities (see [8, 9]).

Parameter	Description	Equation
B	Skempton coefficient	$\frac{\alpha}{[\alpha - \phi(1 - \alpha)] + \phi \frac{K}{K_f}}$
K^u	Undrained bulk modulus	$\frac{K}{1.0 - \alpha B}$
G	Shear modulus	$3K \frac{1 - 2\nu}{2 + 2\nu}$
ν^u	Undrained Poisson's ratio	$\frac{3K^u - 2G}{2(3K^u + G)}$
B_v	Uniaxial Skempton coefficient	$\frac{B(1 + \nu_u)}{3(1 - \nu_u)}$
K_v	Uniaxial bulk modulus	$3K \frac{(1 - \nu)}{1 + \nu}$
K_v^u	Uniaxial undrained bulk modulus	$3K^u \frac{(1 - \nu^u)}{1 + \nu^u}$
S_v	Uniaxial storage	$\frac{\alpha}{K_v B_v}$

4.2.1 Analytical solution

For a single fluid phase, the analytical solution for pressure dissipation and solid deformation in time are available. The analytical solution to this problem has been utilized a number of times for this very purpose. Beginning from the 1-D fluid diffusion equation of hydrogeology,

$$\frac{\partial \bar{P}}{\partial t} - c \frac{\partial^2 \bar{P}}{\partial z^2} = 0 \quad (16)$$

where c is 1-D fluid diffusivity. The pore pressure response to a vertical load, σ_z , applied linearly over time ($\sigma_z^{t=0^-} = 0$) to the top of the column at a rate, $\dot{\sigma}_z = d\sigma_z / dt$, is, [10, Eq. 6.50],

$$\bar{P}(z, t) = P_0 \left\{ 1.0 - \left(\frac{L - z}{L} \right)^2 - \frac{32}{\pi^3} \left[\sum_{m=0}^{\infty} \frac{(-1)^m}{(2m + 1)^3} \exp[-\lambda^2 ct] \cos[\lambda(L - z)] \right] \right\}, \quad (17)$$

where the total pressure generation is

$$P_0 = \frac{L^2}{2c} (B_v \dot{\sigma}_z), \quad (18)$$

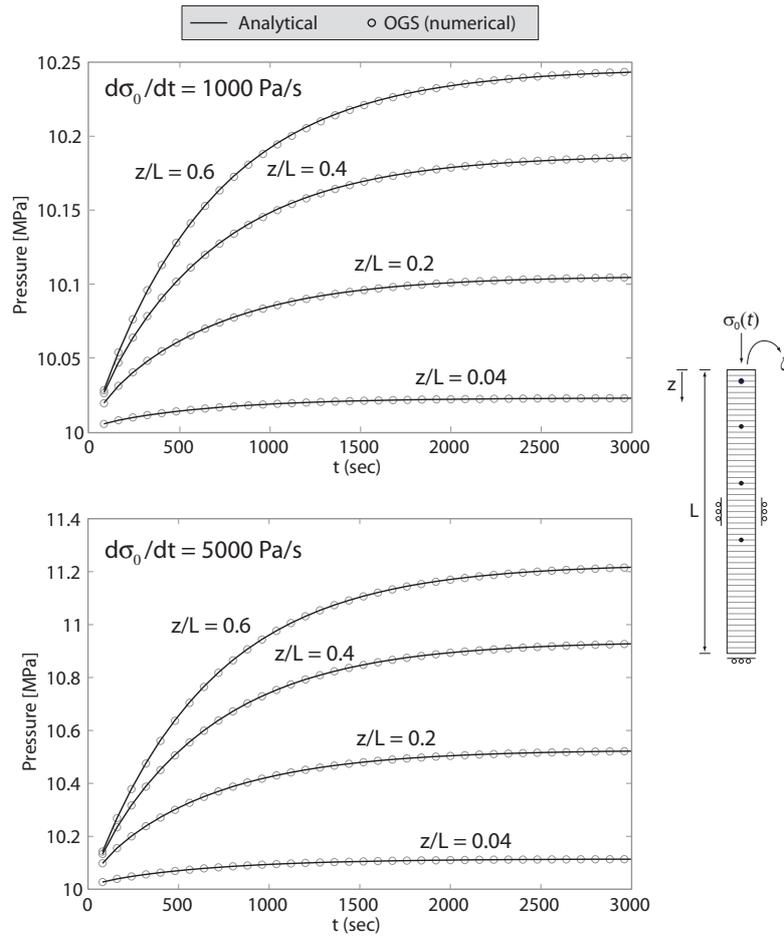


Figure 2. Numerical and analytical solutions of the coupled H2M Terzaghi consolidation problem with variable loading rate.

for the factor, $\lambda = (2m + 1)\pi / (2L)$, the total column length, L , and the location in the column (downward from the applied stress), z . The 1-D Skempton coefficient,

$$B_v = - \left. \frac{d\bar{P}}{d\sigma_{zz}} \right|_{\epsilon_{xx}=\epsilon_{yy}=\zeta=0} = \frac{\alpha}{K_v S_v}, \quad (19)$$

is given purely by micromechanical, poroelastic considerations from the uniaxial drained bulk modulus, K_v , and the 1-D specific storage, S_v (Table 1). The 1-D diffusivity is also a derivative of the 1-D storage:

$$c = \frac{k}{\mu S_v}, \quad (20)$$

If utilizing an applied step load at time $t = 0^+$ an analytical solution is available for pressure and displacement. For this validation, only the linear loading rate solution is examined. Because displacement is the primary variable in the FEM formulation, the

displacement must be accurate in order to generate the correct pressure response: we find no need to reproduce the results of a step load analysis here.

The domain is given the initial properties, $S_w = 0.8$, $k_w^r = k_{nw}^r = 0.5$, and $K_w = 2.93$ GPa and $K_{nw} = 1.19$ GPa. Capillary pressure is set constant at zero and relative permeability is constant. Utilizing Eq. (15) the appropriate coefficient (Eq. (19)) for calculation of the analytical solution (Eq. (17)) can be obtained. Two loading rates ($\dot{\sigma}_z = d\sigma_z / dt$) are examined in Figure 2 and adaptive time stepping and iteration tolerance control is utilized as discussed in section 3.1.

5 TEST CASE: AQUIFER CAPROCK SYSTEM

This section presents a few initial results of H2M flow in a simple aquifer-caprock system. Figure 3 illustrates the geometry and boundary conditions. A depth of 2000m is chosen, and fluid properties of water and CO₂ assigned appropriately. A vertical stress of 44MPa is applied at the upper boundary and an initial fluid pressure assigned to the total domain at 20MPa. CO₂ is injected centrally to the reservoir, at the maximum non-wetting saturation and injection pressure is constant, conservatively, at 22MPa. East and west boundaries constrain zero horizontal displacement, and so the vertical and horizontal stress state is allowed to develop naturally. Real reservoirs will exhibit a greater degree of initial instability due to tectonic stress states.

The aquifer is given mechanical parameters of generic sandstone, and the caprock modified slightly from this in terms of saturation dependent properties. Mechanical properties are not altered from reservoir values for the caprock, which generates a more stable physical situation. Introducing a boundary of differing mechanical properties would decrease the mechanical stability beyond that observed in Figure 4. Property values are shown in Table 2. Stability of the reservoir is indexed to a “factor of safety (f_s^s)” defined by a Mohr-Coulomb failure criterion. In 2-D, the criterion states that mechanical failure is favorable when the current maximum shear stress, $\tau_m = 1/2(\sigma_1 - \sigma_3)$, becomes [11],

$$|\tau_m| \geq C_h \cos \varphi + (\sigma_m - \bar{P}) \sin \varphi, \quad (21)$$

where σ_1 and σ_3 are the maximum and minimum principle stresses, respectively, and $\sigma_m = 1/2(\sigma_1 + \sigma_3)$ is the mean normal stress. A relationship is presented based on Eq. (21) that defines a “factor of safety (f_s^s)” for shear slip/failure along an optimally oriented failure plane,

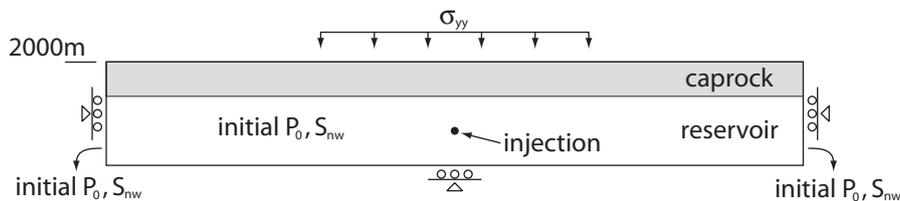


Figure 3. Geometry of reservoir system.

Table 2. Parameter values for reservoir simulation.

Parameter	Reservoir	Caprock
S_{lr}	0.30	0.30
S_{gr}	0.02	0.02
P_b	19.6kPa	3100kPa
\mathbf{k}	$1.9 \times 10^{-13} \text{ m}^2$	$1.9 \times 10^{-17} \text{ m}^2$
ϕ	0.19	0.02
Young's Modulus	14.4 GPa	14.4 GPa
Poisson Ratio	0.2	0.2
α	0.8	0.8

$$f_s^s = \frac{|\tau_m|}{C_h \cos \varphi + (\sigma_m - \bar{P}) \sin \varphi}, \quad (22)$$

for the cohesion, $C_h = 50\text{MPa}$, and friction coefficient, $\tan \varphi = 0.6$. Values of $f_s^s \geq 1$ imply incipient failure. Results of this criterion are presented in Figure 4 (left) for two values of cohesion. The lower value (5MPa) represents the case where optimally oriented, pre-existing fractures are present. The factor of safety is not violated in either case. However, the physical description utilized here is extremely conservative with respect to initial reservoir stability, discontinuity of mechanical properties across layers, and the injection pressure. Regardless, prime locations for reservoir instability are easily discernable. With lower cohesion, a greater fraction of the caprock is in the higher range of f_s^s and the magnitude of the instability is also higher (the latter is of course expected). It is clear that the highest risk of failure is at the reservoir/caprock interface near the injection region (a result in agreement with the study of [12]), not necessarily good news given the importance of caprock integrity. Additionally, although not strongly observable in the figures due to the domination of the material interface, a region of increased instability can be observed to follow the advance of the CO_2 migration front. Whether or not this will be an important observation is the focus of future study.

Figure 4 (right) illustrates the difference in CO_2 migration between H2M and H2 simulations. While some small difference is visible, it is not significant in this case. However, as porosity and permeability are not allowed to change in these simulations, the difference was not expected to be large. Permeability was shown to increase by nearly 50% in the CO_2 injection study of [12]. Furthermore, the introduction of thermal effects, shown to have dire consequences on permeability in geothermal reservoirs ([13]), may be important in this regard. This, also, will be included in future work.

6 CONCLUSIONS

A numerical scheme has been implemented for the coupled solution of multiphase flow and mechanical deformation during the injection of CO_2 into geological reservoirs. Benchmarking of the multiphase hydraulic and also the multiphase hydro-mechanical coupling has been conducted against analytical solutions. Results are accurate for the fundamental coupling methodology.

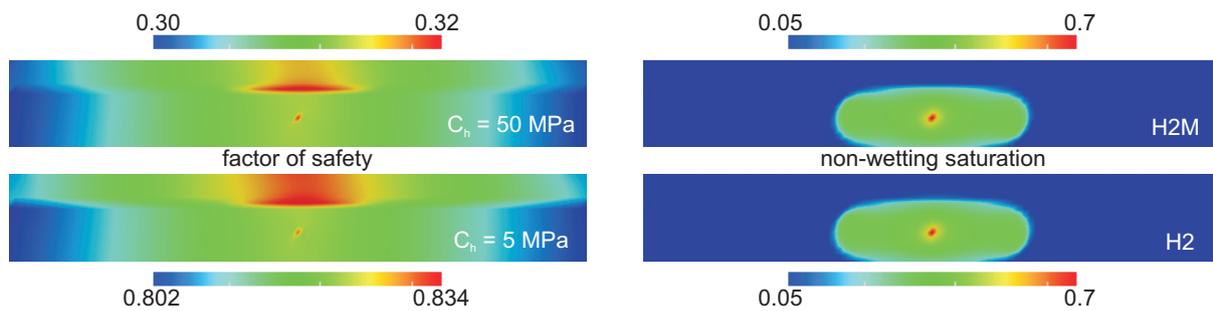


Figure 4. Left: Factor of safety (Top: $C_h = 50$ MPa Bottom: $C_h = 5$ MPa) and Right: CO₂ saturation (Top: H2M simulation Bottom: H2 simulation) for alternate test conditions.

Initial results for a simplified aquifer-caprock system do not show significant differences between H2 and H2M simulations for the very conservative injection scenario utilized here. More importantly, however, the coupled problem allows examination of reservoir integrity and the potential for breaching of an intact caprock. The interface between zones of different material properties (such as between the reservoir and caprock) are particular targets for mechanical failure, and improper selection of injections rates could easily lead to a breach of trapping integrity. More complex stress states will be important for a more detailed analysis in addition to sensitivity analysis of injection scenarios.

Future work will seek greater complexity in geometric representation of the reservoir system. Principally, the introduction of high permeability zones within the caprock and extension of the geometry to include additional geological layers. Greater complexity will be sought with regard to constitutive relationships defining how porosity and permeability will be altered dynamically within the system.

6 ACKNOWLEDGEMENTS

This study is the result of partial funding from the German Federal Ministry of Education and Research (BMBF) as part of joint projects in the framework of the Special Programme GEOTECHNOLOGIEN and from the Helmholtz Association within the Research Programme “Renewable Energies – Geothermal Technologies”.

REFERENCES

- [1] Pruess K, Xu T, Apps J, et al. Numerical modeling of aquifer disposal of CO₂. Society of Petroleum Engineers (SPE), Paper number 66537 2001.
- [2] Wang W, Kosakowski G, and Kolditz O. A parallel finite element scheme for thermo-hydro-mechanical (THM) coupled problems in porous media. *Computational Geosciences* 2009;35:1631-1641.
- [3] Bear J and Bachmat Y. Introduction to modeling of transport phenomena in porous media. Netherlands: Kluwer Academic; 1991,553.
- [4] Lewis RW and Schrefler BA. The finite element method in the static and dynamic deformation and consolidation of porous media, 2nd ed. West Sussex, England: John Wiley & Sons; 1998,492.

- [5] Taron J, Elsworth D, and Min KB. Numerical simulation of thermal-hydrologic-mechanical-chemical processes in deformable, fractured porous media. *Int J Rock Mech Min Sci* 2009;46(5):842-854.
- [6] Fucik R, Mikyska J, Benes M, et al. An improved semi-analytical solution for verification of numerical models of two-phase flow in porous media. *Vadose Zone Journal* 2007;6:93-104.
- [7] McWhorter DB and Sunada DK. Exact integral solutions for two-phase flow. *Water Resour Res* 1990;26(3):399-413.
- [8] Detournay E and Cheng AH-D, Fundamentals of poroelasticity, In: J.A. Hudson, Editor. *Comprehensive rock engineering*. New York: Pergamon; 1993.
- [9] Wang HF. *Theory of linear poroelasticity*. Princeton: Princeton University Press; 2000.
- [10] Wang HF. *Theory of linear poroelasticity*. Princeton: Princeton University Press; 2000,287.
- [11] Jaeger JC, Cook NGW, and Zimmerman RW. *Fundamentals of rock mechanics*, 4th ed. Malden, MA: Blackwell Publishing; 2007.
- [12] Rutqvist J and Tsang C-F. A study of caprock hydromechanical changes associated with CO₂ injection into a brine formation. *Environ Geol* 2002;42:296-305.
- [13] Taron J and Elsworth D. Thermal-hydrologic-mechanical-chemical processes in the evolution of engineered geothermal reservoirs. *Int J Rock Mech Min Sci* 2009;46(5):855-864.

NUMERICAL STUDY OF THE CHEMO-PORO-MECHANICAL BEHAVIOUR OF THE CEMENT SHEATH DURING CO₂ INJECTION

VALERIE VALLIN*, JEAN-MICHEL PEREIRA*, ANTONIN FABBRI†,
HENRY WONG‡ AND NICOLAS JACQUEMET†

*Université Paris-Est, Laboratoire Navier, École des Ponts ParisTech
6-8 av. Blaise Pascal, Cité Descartes, Champs-sur-Marne
77455 Marne-la-Vallée Cedex 2, France
e-mail: valerie.vallin@enpc.fr, jeanmichel.pereira@enpc.fr, navier.enpc.fr

†BRGM
3 av. Claude Guillemin, 45060 Orléans Cedex 2, France
e-mail: a.fabbri@brgm.fr, www.brgm.fr

‡Département Génie Civil et Bâtiment, Université de Lyon, ENTPE
rue Maurice Audin, 69518 Vaulx-en-Velin cedex, France
e-mail: henry.wong@entpe.fr, www.entpe.fr

Key words: CO₂, Geological storage, Cement carbonation, Chemo-poromechanics

Abstract. This paper aims at describing the influence of carbonation on the poromechanical behaviour of a wellbore cement. Attention is paid to the estimation of the damage risk of the cement sheath in the context of CO₂ geological storage. A chemo-poromechanical model is presented. It accounts for the main chemical reactions occurring within the cement matrix and their consequences in terms of matrix dissolution and carbonates precipitation. These porosity changes are accounted for by introducing a chemical porosity associated to the cement matrix dissolution and the porous volume occupied by carbonate precipitates. The model has been implemented in a finite volume code. An axisymmetrical configuration is considered. Evolutions of transport characteristics and mechanical behaviour during carbonation process are predicted. The damage risk of the cement sheath is finally estimated through the calculation of the elastic energy stored within the solid matrix.

1 INTRODUCTION

A possible solution to mitigate the impact of greenhouse gases emissions into the atmosphere relies on the geosequestration of carbon dioxide. This technology consists in

injecting CO_2 into a relatively deep geological formation through an injection well. Selected sequestration sites must ensure the perenity of the storage. To this end, a favourable geological configuration has to include a highly porous reservoir rock covered by a caprock characterised by low permeability and porosity.

The study presented in this paper deals with the chemo-poromechanical couplings occurring in wellbore cements in the context of CO_2 geosequestration. The main objective is to follow the evolution of chemically induced mechanical alterations of the cement and to estimate the damage risk of the cement well. Indeed, the presence of CO_2 within the injection site may affect the poromechanical behaviour of the cement and the rocks. This work concerns exclusively the behaviour of a cement well. The matrix of this latter is composed of several hydrated minerals such as lime hydrates (mainly Portlandite and calcium silicates). In presence of CO_2 , cement will be submitted to dissolution and precipitation reactions, which will heavily affect its stiffness and strength. These chemical reactions can present a serious threat to the perenity of CO_2 geosequestration.

A cylindrical cement well-bore submitted to the injection of CO_2 is considered in this study. This boundary value problem is solved by considering axisymmetric conditions and assuming infinitesimal transformations. At any time, the porous medium is saturated by a single liquid phase. CO_2 injection effects are simulated by the presence of CO_2 -rich brine put in contact with the porous system.

In the first part of the paper, the chemo-poromechanical model is presented before describing its implementation within a finite volume code. Then, the simulated boundary value problem, corresponding to a possible leakage of CO_2 between the caprock and the cement sheath, will be presented and numerically solved. Evolutions of pore overpressure, porosity, permeability and solid matrix moduli during cement carbonation process will be predicted. Finally, the damage risk of the cement sheath will also be estimated through the calculation of the elastic energy stored within the solid matrix.

2 CHEMO-POROMECHANICAL MODEL FOR CEMENT CARBONATION

In all the study, we assume isothermal conditions and infinitesimal transformations. We consider a two-phase porous medium, composed of a solid phase and a fluid phase. Each of these phases is composed of several species. The solid phase, denoted by S , contains not only the solid cement matrix itself composed by several minerals M_i but also carbonate crystals C . The in-pore fluid phase F is made of water, the solvent, denoted w , and dissolved species denoted α which partly come from the dissolution of the solid phase and CO_2 . In presence of CO_2 the main chemical alteration to which the cement is submitted corresponds to a carbonation process leading its matrix leaching and to the formation of carbonate crystals. Furthermore, it has been assumed that the carbonate crystals precipitate exclusively as calcite, given the relative stability of this polymorph as compared to other forms.

2.1 Saturated poroelasticity

Mechanical behaviour of saturated porous solids can be efficiently accounted for using a sound thermodynamical framework. Such an approach, as demonstrated by Coussy [1], [2], is relevant because it allows the extension to porous media of the balance equations of continuum mechanics relative to mass, momentum, energy and entropy. Key points of this approach are recalled in the following.

In this work, in-pore fluid solution is assumed to be a mixture composed of several species: the water being the solvent and several solutes such as the species coming from the dissolved matrix, dissolved carbonates and dissolved CO_2 . Let $n_i \times \Omega_0$ be the number of moles of the species i present in a unit volume of the porous medium Ω_0 , so that n_i is the apparent molar density of species i . The isothermal Gibbs-Duhem equality assuming chemical equilibrium can be written relatively to the in-pore solution and to the carbonate crystal so that:

$$-\delta_C \frac{dp_C}{dt} + n_C^S \frac{d\mu_C^S}{dt} = 0 \quad (1)$$

$$-\phi_F \frac{dp_F}{dt} + \sum_{i=\alpha,w} n_i^F \frac{d\mu_i^F}{dt} = 0 \quad (2)$$

where p_C and p_F are the pressures of the carbonate crystal and the in-pore fluid respectively and μ_C^S and μ_α^F stand for the molar chemical potentials of calcite and aqueous species α respectively. In the equation 2, the energy transmitted through the shear stress is assumed to be neglected in front of that transmitted through the normal stress between carbonate crystals and solid matrix.

Let ϕ_0 be the initial porosity and Ω_0 be the initial volume of the infinitesimal representative element of the porous continuum so that the initial porous volume of the medium is $\phi_0 \times \Omega_0$. A difference is made between the effective porosity of the porous medium and the porosity of the solid matrix. The former, denoted ϕ_F , corresponds to the porous volume occupied by the in-pore fluid per unit of porous medium. The latter denoted ϕ corresponds to the space per unit of volume of porous medium which is not occupied by the cement phase. The difference between these two porosities is denoted δ_C . By definition, the volume $\delta_C \times \Omega_0$ equals the volume occupied by carbonate crystals. This distinction is relevant to follow precisely the processes of dissolution/precipitations of the cement matrix and the calcite. These porosities are linked through the following relation:

$$\phi = \phi_F + \delta_C \quad (3)$$

When a cementitious material is submitted to the presence of CO_2 , it undergoes several variations on porosities. Some are due to chemical reactions, denoted ϕ_L for leaching of cement matrix and ϕ_P for carbonates precipitation. Others are related to the deformation of the porous medium. Let φ_F and φ_C be the deformation of the porosity filled respectively

by the fluid phase and by the calcite phase. The porosities involved in Eq. (3) can thus be rewritten as follows:

$$\phi_F = \phi_0 + \phi_L - \phi_P + \varphi_F \quad (4)$$

$$\delta_C = \phi_P + \varphi_C \quad (5)$$

Let now ϵ be the overall infinitesimal strain tensor of Ω_0 and σ the overall stress tensor to which the considered system is subjected. Ω_0 being an open thermodynamic system, Ω_0 will exchange during its evolution moles of species α with its surroundings. Finally, let introduce Ψ the free energy of the system. The first and the second laws of thermodynamics combine to provide the isothermal Clausius-Duhem inequality related to the system in the form (see [1]):

$$\sigma : \dot{\epsilon} - \sum_{i=\alpha,w} \mu_i^F \operatorname{div} \mathbf{w}_i^F - \frac{d\Psi}{dt} \geq 0 \quad (6)$$

where \mathbf{w}_i^F is the vector of molar transport of the aqueous species i . Thus $-\operatorname{div} \mathbf{w}_i^F$ is the rate of moles of aqueous species i externally supplied to the infinitesimal porous element Ω_0 by its contiguous elements. Considering Fick's diffusion law and Darcy's law for the diffusive and advective transport of species i , the vector \mathbf{w}_i^F can be expressed as:

$$\mathbf{w}_i^F = -d_{\text{eff}} \operatorname{grad} C_i^F - C_i^F \frac{\kappa}{\eta_{\text{vis}}} \operatorname{grad} p_F \quad (7)$$

with d_{eff} the effective diffusion coefficient, κ the intrinsic permeability, η_{vis} the dynamic viscosity of the fluid phase, C_i^F the molar concentration of the species i and p_F the fluid pressure. From Gibbs-Duhem equalities (1) and (2) and Clausius-Duhem inequality (6), assuming isothermal conditions, the constitutive equations of isotropic linear poroelasticity can be deduced (see [1] for more details):

$$\sigma - \sigma_0 = \left(K - \frac{2}{3}G \right) (\varepsilon - \varepsilon_0) \mathbf{1} + 2G (\epsilon - \epsilon_0) - \sum_{K=F,C} b_K (p_K - p_{K,0}) \mathbf{1} \quad (8)$$

$$\varphi_J - \varphi_{J,0} = b_J (\varepsilon - \varepsilon_0) + \sum_{K=F,C} \frac{p_K - p_{K,0}}{N_{JK}} ; \quad J = F, C \quad (9)$$

where $\varepsilon = \operatorname{tr}(\epsilon)$ is the volumetric deformation. K and G respectively are the bulk modulus and the shear modulus of the empty porous solid and b_J and N_{JK} respectively are the generalised Biot coefficient and the generalised poroelastic coupling moduli.

In order to account for the effects of porosity variations and others parameters of the solid matrix on the mechanical moduli, the three-phase self-consistent micromechanical model under the assumption of local isotropy [5] is used. The effective bulk modulus of the porous medium thus writes:

$$K = \frac{4G_m K_m (1 - \phi)}{4G_m + 3K_m \phi} \quad (10)$$

where K_m and G_m are the bulk and shear moduli of the solid matrix. The effective shear modulus G can be obtained thanks to G_m , the effective porosity and Poisson's ratio ν of the porous medium. Its expression is detailed in [5].

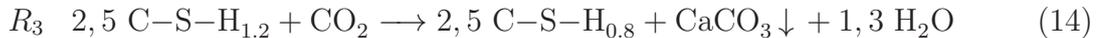
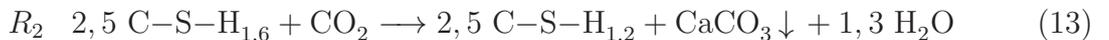
Finally, thanks to the linearity of poroelastic constitutive equations (8), the elastic energy W stored within the solid matrix simply writes in the following form:

$$W = \boldsymbol{\sigma}_0 : \boldsymbol{\epsilon}_0 + \frac{\left(\sigma - \sigma_0 + \sum_{J=F,C} b_J (p_J - p_{J,0}) \right)^2}{2K} + \sum_{I,J=F,C} \frac{(p_I - p_{I,0})(p_J - p_{J,0})}{2N_{IJ}} + \frac{(\mathbf{s} - \mathbf{s}_0) : (\mathbf{s} - \mathbf{s}_0)}{4G} \quad (11)$$

where $\sigma = \frac{1}{3} \text{tr}(\boldsymbol{\sigma})$ denotes the mean stress and $\mathbf{s} = \boldsymbol{\sigma} - \sigma \mathbf{1}$ the deviatoric stress. The calculation of this energy W is relevant when evaluating the risk of damage to the solid matrix.

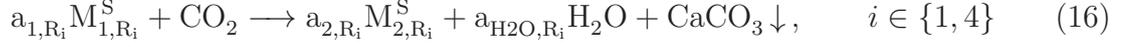
2.2 Chemical reactions

In this work, we will only consider two hydrates of the cement : Portlandite ($\text{Ca}(\text{OH})_2$), noted CH in cement industry, and calcium silicate hydrates, noted C–S–H. Indeed, following [8], predominant mechanisms of the carbonation of cement are those related to these last two hydrates. The chemical mechanism associated to the carbonation of CH is simple whereas the evolution of the C–S–H during carbonation corresponds to a progressive loss of density linked to a decalcification of the structure until the formation of amorphous silica SiO_2 . In order to identify the reactions pathways of carbonation, a reactive transport simulation has been carried out using the reactive transport code TOUGHREACT, a software well suited for modelling multiphase geochemical reactive transport in porous and fractured media (cf [9]). The main chemical reactions which have been identified during these simulations are the following:



The first reaction R_1 corresponds to the carbonation of CH, whereas the others describe the carbonation of C–S–H by successive decalcifications, from C–S–H_{1.6} to SiO_2 . The subscripts 1.6, 1.2, and 0.8 reflect the loss of density of the C–S–H and correspond to the C/S density ratio. Based on the simulations performed with TOUGHREACT, we will assume that all the reactions R_i are instantaneous, complete and successive. In other

words, reaction kinetics are neglected. We can write these four reactions in the following synthetic form:



where a_{j,R_i} is the stoichiometric coefficient of the mineral M_{j,R_i}^S in the reaction R_i , with $j = 1$ for a reactant and $j = 2$ for a product in the same reaction.

Chemical equilibrium conditions between fluid, crystals and cement matrix components are now recalled. The reaction of precipitation of calcium carbonates is assumed to be instantaneous. This assumption can be made because the kinetics of dissolution/precipitation is really faster than diffusive transport of CO_2 in the liquid phase (cf [7]). Doing so, the thermodynamical equilibrium between dissolved species in the fluid and precipitated carbonate crystals is true at any time and in each infinitesimal representative volume. This amounts to say that the chemical potential of carbonate crystals is in equilibrium with the one of dissolved crystal:

$$\mu_C = \mu_{\text{Ca}^{2+}} + \mu_{\text{CO}_3^{2-}} \quad (17)$$

Finally, during the molar transport of CO_2 and during chemical reactions, each molar quantity of species, denoted $\frac{dn_i}{dt}$ is not only due to exchanges of matter with surrounding elementary volumes but also to chemical reactions. Assuming that molar transport of solid minerals is not significant, conservation equations can be expressed such as:

$$\frac{dn_{M_i,C}^S}{dt} = \dot{n}_{M_i,C}^S \quad (18)$$

$$\frac{dn_{\alpha,w}^F}{dt} = -\text{div} \mathbf{w}_{\alpha,w}^F + \dot{n}_{\alpha,w}^F \quad (19)$$

with $\dot{n}_{M_i,C}^S$ and $\dot{n}_{\alpha,w}^F$ the molar variations of solid and dissolved species due to the chemical reactions and $\mathbf{w}_{\alpha,w}^F$ the molar transport given by equation (7).

2.3 Chemo-poromechanical behaviour

To conclude on the coupling between the chemical reactions and the poromechanical behaviour, we express the overall mass conservation of the in-pore fluid:

$$\frac{1}{\rho_F} \frac{d(\phi_F \rho_F)}{dt} = \frac{1}{\rho_F} \sum_{i=\alpha,w} \mathcal{M}_i \frac{dn_i}{dt} \quad (20)$$

By introducing the bulk modulus of the in-pore fluid, K_F , the variation of the density of the fluid in isothermal conditions can be written in the form:

$$\frac{1}{\rho_F} \frac{d\rho_F}{dt} = \frac{1}{K_F} \frac{dp_F}{dt} + \sum_{i=\alpha,w} \gamma_i^F \frac{dn_i^F}{dt} \quad \text{with} \quad \gamma_i^F = \left[\rho_F \frac{d(1/\rho_F)}{dn_i} \right]_{p_F, n_j, j \neq i} \quad (21)$$

where γ_i^F is the variation of the density of the fluid which takes into account the effects of chemical reactions on the porous medium behaviour due to the modification of the chemical composition of the in-pore fluid. Moreover, with the assumption of no transmission of shear between calcite and cement matrix (see section 2.1), the mechanical behaviour of the crystal can be written as:

$$\frac{1}{\rho_C} \frac{d\rho_C}{dt} = \frac{1}{K_C} \frac{dp_C}{dt} \quad (22)$$

where K_C is the bulk modulus of the carbonate crystal, ρ_C its density and p_C its pressure. Thanks to equations (4), (5), (20) and (21) and by assuming a dilute in-pore fluid (that is to say that water is in excess as compared to dissolved species), the conservation of fluid mass can be expressed as follows:

$$\frac{\phi_F}{K_F} \frac{dp_F}{dt} + \frac{d\phi_F}{dt} = (\nu_C^S - \nu_{C,dis}^F) \dot{n}_C^S - \sum_{M_i} ((\nu_{M_i}^S - \nu_{M_i,dis}^F) \dot{n}_{M_i}^S) + \text{div} \left(\frac{\kappa}{\eta_{vis}} \mathbf{grad} p_F \right) \quad (23)$$

According to the assumptions related to reactions kinetics previously introduced, the advancement rate $\dot{\xi}_{R_i}$ of each reaction R_i can directly be estimated from the quantity of CO_2 supplied to the system by diffusion and advection:

$$\dot{\xi}_{R_i} = \begin{cases} \text{div} \left(d_{\text{eff}} \mathbf{grad} \left(\frac{n_{\text{CO}_2}}{\phi_F} \right) + \left(\frac{n_{\text{CO}_2}}{\phi_F} \right) \frac{\kappa}{\eta_{vis}} \mathbf{grad} p_F \right) & \text{if } M_{1,R_i} \neq 0 \text{ and } M_{1,R_{i-1}} = 0 \\ 0 & \text{otherwise} \end{cases} \quad (24)$$

Finally, the variation of the molar quantity of CO_2 corresponds to the gas quantity brought by advection and diffusion minus the quantity consumed by chemical reactions:

$$\frac{dn_{\text{CO}_2}}{dt} = \text{div} \left(d_{\text{eff}} \mathbf{grad} \left(\frac{n_{\text{CO}_2}}{\phi_F} \right) + \left(\frac{n_{\text{CO}_2}}{\phi_F} \right) \frac{\kappa}{\eta_{vis}} \mathbf{grad} p_F \right) - \sum_{R_i} \dot{\xi}_{R_i} \quad (25)$$

Equation (8) submitted to the momentum balance condition and equations (23), (24) and (25) define the chemo-poromechanical model that has to be solved. Its implementation into a finite volume code is described in the following section.

3 COMPUTER CODE

The chemo-poromechanical model introduced so far is now implemented into a numerical code in order to solve a realistic boundary value problem, corresponding to the modelling of wellbore cement subjected to CO_2 injection under deep reservoir conditions. As a preliminary study, a simplified geometry is considered by assuming axial symmetry and plane strain conditions. In these conditions, the different unknowns only depend on the radial space variable (distance to the well). Assuming also that the material is isotropic, the strain and stress tensors are diagonal. The cement studied here corresponds

ϕ_0	κ_0	η_{vis} (m ² /s)	K_m (MPa)	G_m (MPa)	K_F (MPa)
30%	$1 \cdot 10^{-19}$	$1.79 \cdot 10^{-9}$	17500	10575	2200
ν_{CH}^S	$\nu_{CSH1.6}^S$	$\nu_{CSH1.2}^S$	$\nu_{CSH0.8}^S$	ν_{SiO2}^S	ν_C^S
33.1	84.7	72	59.3	29	36.9

Table 1: Table of parameters (with molar volumes in cm³/mol)

to a class G cement. According to [6], for such cements, the permeability and the diffusivity coefficients depend on porosity changes according to:

$$\kappa = 1, 2 \left(\frac{\phi_F}{0, 26} \right)^{11} \cdot \kappa_0 \quad [m^2] \quad (26)$$

$$d_{\text{eff}} = 100 \exp(9.95\phi_F - 29.08) \quad [m^2/s] \quad (27)$$

where κ_0 is a material constant. A table of the parameters is presented in the table 1.

In order to solve numerically the problem, we implement the previous equations into a finite volume code. This code has been initially developed by M. Mainguy [7] and extended by A. Fabbri [3] to account for crystallisation phenomena. As seen from carbonation experiments on cement samples (see [4] for instance), a quite sharp separation exists between the carbonated and healthy parts of the samples. Furthermore, this front remains sharp during the carbonation process. As a consequence, it can reasonably be assumed that a sharp carbonation front exists at any time in the sample.

Based on the assumptions made in this study, it appears that the use of the finite volume discretization is pertinent for the numerical resolution since it prevents from the occurrence of numerical instabilities or divergence problems because of this sharp carbonation front.

In the finite volume method, we discretize the modelled domain in a given number of control volumes associated by central nodes and interfaces. We choose to keep a constant radius step between nodes of volumes, and thus between interfaces. The unknowns of the problem, as fluid pressure, advancement rate of reaction, CO₂ concentration and displacements, are localized on the nodes (and thus constant per control volume). To fix the boundary conditions, two half-control volumes are added at the boundaries of the structure. In our program, an implicit Euler scheme is used for the temporal discretization. Then, the resolution of our problem is ensured by the Newton-Raphson method.

The cement is composed of 50% of Portlandite and 50% of C–S–H in volume. With the hypothesis of the sharp carbonation front, we consider that for a given control volume, CO₂ reacts with all the reactants in presence before entering into the next control volume. Finally, given that the water bulk modulus K_F is sensibly lower than the bulk moduli of the cement matrix K and the carbonate crystal (70 GPa), we neglect the term $\frac{d\phi_F}{dt}$ in equation (23). To simplify the problem, calcite is assumed to be equilibrated with the fluid ($p_C = p_F$).

4 RESULTS AND DISCUSSION

4.1 Statement of the problem

Before presenting the results of the simulations, we will make explicit initial and boundary conditions used.

We consider a cylindrical sample of cement with an internal radius of 89 mm, an external radius of 108 mm and an unit height. Since we are interested in the simulation of the leakage of gas dissolved in water through the interface between the caprock and the well cement seen as a preferential path, CO_2 will be assumed to reach the cement through its external surface. We consider a uniform fluid pressure of 13 MPa at the beginning, which roughly corresponds to the injection pressure of CO_2 at around 1000m depth. Moreover, the initial state being considered as a reference, null displacement ($u_r = 0$) and deformation ($\epsilon_r = \epsilon_\theta = 0$) are supposed in the well. The in situ stress field is uniform and isotropic. The mean total stress is 25 MPa. In addition, the cement is healthy and without any carbonate crystal before the injection of CO_2 . The initial porosity is 30%, a realistic value encountered in class G cements. The in-pore water does not contain dissolved CO_2 at the initial state.

At the beginning of the injection stage ($t = 0s$), the fluid pressure of 13 MPa is kept constant on the outside face of the cylinder and, with the arrival of CO_2 -rich fluid, a CO_2 concentration corresponding to the saturated concentration of CO_2 in water is fixed on this surface. On the inside face, because of the presence of the metallic casing, an impermeable boundary is assumed. Finally, concerning the mechanical problem, we choose to apply a radial stress of compression of 25 MPa on the external surface and of 13 MPa on the internal surface of the well, which corresponds to the stress induced by injection pressure.

4.2 Results

The results are now presented. We choose to simulate the injection of CO_2 into a cement sample during 180 days. First, we can follow the evolutions of the chemical reactions as a function of space at several times by looking at the evolutions of the molar quantities of the CH, C–S–H and CaCO_3 in figure 1.

We recall that the incoming of CO_2 is imposed by the external surface of the sample (at $r = r_{ext} = 108$ mm). As we can see on figure 1(a), after 30 days of exposure, the zone of the sample between $r=100\text{mm}$ and $r=108$ mm does not contain CH and C–S–H. So that, the first 8mm are totally carbonated. The limit between the part containing no CaCO_3 and the one containing only CaCO_3 (apart from SiO_2 coming from C–S–H carbonation) corresponds to the carbonation front. Figure 1(b) shows that all the sample is fully carbonated after 161 days of exposure to CO_2 . After this time, all the CH and C–S–H have been consumed, letting place to the carbonate crystals. As seen previously, these modifications of constitution of the cement matrix involve evolutions of the porosity of the porous media. Figure 2(a) shows how the chemical reactions affect the fluid porosity ϕ_F of the sample during the 180 days of exposure.

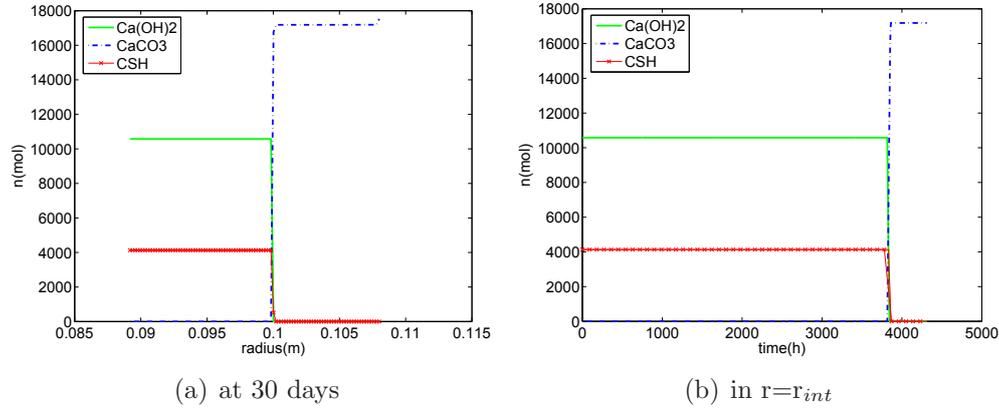


Figure 1: Evolutions of molar quantities of the CH, C–S–H and CaCO_3 under CO_2 injection according to the distance from the cylinder axis and the time

When the cement sample is healthy (that is before being carbonated), its porosity is 30%, whereas in the part completely carbonated, its porosity decreases down to 24.4%. This important loss of porosity involves, as seen from equations (26) and (27), a real alteration of transport characteristics of the cement such as a real decrease of the permeability which is divided per 10 and of the diffusivity coefficient of the CO_2 in water which is divided by 2 in altered zones.

Moreover, as a consequence of the decrease of the fluid porosity when the carbonate crystals are forming, the in-pore fluid pressure increases and can not be immediately evacuated because of the low permeability. So, from the graph 2(c), we can see a peak of in-pore fluid pressure localized at the carbonation front and which creates a water flow directed towards the external surface of the cement sample. However, even though the fluid porosity decreases with the carbonation, the dissolution of the cement matrix due to its leaching leads to an increase of the effective porosity of the matrix ϕ (see figure 2(b)). The porosity ϕ (that is excluding carbonates) of the cement matrix varies between 30% and 82% in our case. Consequently and according to Eq.(10), the cement matrix stiffness is highly degraded: the bulk modulus is divided by 9 and the shear modulus per 3 in the carbonated zones. This indicates a potential risk for the durability of CO_2 storage. The risk of damage of the cement can be evaluated by the estimation of the matrix elastic free energy stored in the cement matrix. As seen in Figure 2(d), the elastic free energy increases and damage could be expected. But this risk is really increased when the permeability decreases. Indeed, by comparing figures 3(a) and 3(b), the elastic free energy of the matrix reaches almost 600 MPa for $\kappa = 6 \cdot 10^{-22} \text{ m}^2$ whereas when $\kappa = 6 \cdot 10^{-21} \text{ m}^2$, W does not exceed 3 MPa. Moreover, we can highlight that the carbonation front induces a peak of the stored elastic free energy. The carbonation process could thus create an important risk of damage into the cement matrix.

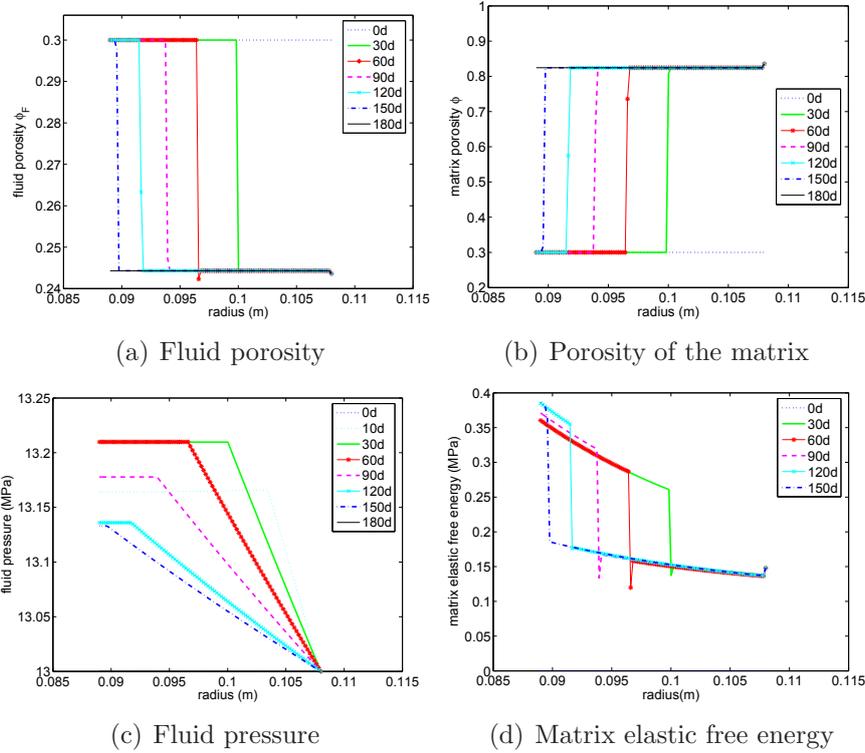


Figure 2: Evolutions of the porosities, the fluid pressure and the matrix elastic free energy in the cement sample during 180 days of exposure to the CO_2

5 CONCLUSIONS

Thanks to the development of a chemo-poromechanical model implemented into a finite volume code adapted to an axisymmetrical problem, we have simulated the behaviour of a cylindrical cement wellbore during the injection of CO_2 . Taking into account the chemical reactions related to the cement carbonation enables to predict the modifications of the transport properties and the mechanical behaviour of the cement. Given that the leaching of the cement matrix and the precipitation of carbonates lead to important variations of the fluid and matrix porosities, permeability, diffusivity and elastic moduli are highly impacted by the carbonation process. Obtained results show a significant risk of damage in the cement that is mainly the result of the important excess pore pressures generated at the carbonation front.

ACKNOWLEDGEMENTS

This work has been supported by French Research National Agency (ANR) through Captage et stockage du CO_2 program (project INTERFACE n°ANR-08-PCO2-006)

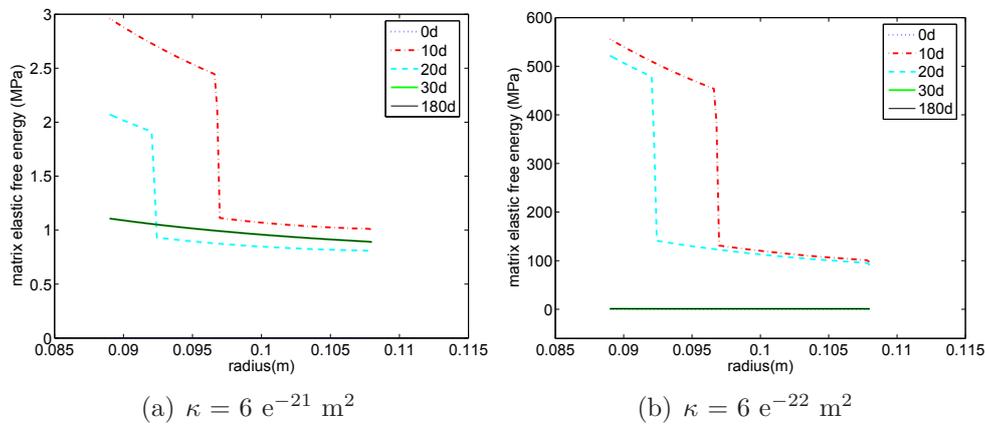


Figure 3: Evolution of the matrix elastic free energy in the cement sample during 180 days of exposure to the CO_2 when the permeability is equal to $6 \cdot 10^{-21} \text{ m}^2$ and to $6 \cdot 10^{-22} \text{ m}^2$

REFERENCES

- [1] O. Coussy. *Poromechanics*. John Wiley & Sons Inc, 2004.
- [2] O. Coussy. *Mechanics and physics of porous solids*. Wiley, 2010.
- [3] A. Fabbri. *Physico-mécanique des matériaux cimentaires soumis au gel-dégel*. PhD thesis, Thèse de doctorat, Université de Marne-La-Vallée, 2006.
- [4] A. Fabbri, J. Corvisier, A. Schubnel, F. Brunet, B. Goff, G. Rimmelé, and V. Barlet-Goudard. Effect of carbonation on the hydro-mechanical properties of portland cements. *Cement and concrete research*, 2009.
- [5] T. Fen-Chong. *Analyse micro-mécanique des variations dimensionnelles de matériaux alvéolaires*. PhD thesis, Thèse de doctorat, École Polytechnique, Paris, 1998.
- [6] S. Ghabezloo. *Comportement thermo-poro-mécanique d'un ciment pétrolier*. PhD thesis, Thèse de doctorat, École Nationale des Ponts et Chaussées, Paris, 2008.
- [7] M. Mainguy. *Modèles de diffusion non linéaires en milieux poreux. Application à la dissolution et au séchage des matériaux cimentaires*. PhD thesis, Thèse de doctorat, École Nationale des Ponts et Chaussées, Paris, 1999.
- [8] M. Thiery. *Modélisation de la carbonatation atmosphérique des bétons*. PhD thesis, Thèse de doctorat, École Nationale des Ponts et Chaussées, Paris, 2005.
- [9] T. Xu, E. Sonnenthal, N. Spycher, and K. Pruess. Toughreact—a simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: Applications to geothermal injectivity and CO_2 geological sequestration. *Computers & Geosciences*, 32(2):145 – 165, 2006.

COMPARISON OF DATA TRANSFER METHODS BETWEEN TWO DIFFERENT MESHES.

P. Bussetta* and J-P. Ponthot*

* LTAS MN²L, Aerospace & Mechanical Engineering Department, B52/3,
Université de Liège
Chemin des Chevreuils, 1; B-4000 Liege, Belgium

Key words: Data Transfer Method, Weak Form, Mortar Element, Finite Element, Finite Volume

Abstract. Many problems solved with the Finite Element Method require more than one mesh (i.e. one specific mesh for each Physic or a remeshing is needed). The Data Transfer Method used, has a great importance in the capacity to solve the problem and in the reliability of the solution. In general, the data is composed of two kinds of fields (defined thanks to the nodal values or at the integration points). In this paper, the more used Data Transfer Method is compared with the Data Transfer Methods based on a Weak Form (using Mortar Element or Finite Volume).

1 Introduction

The numerical resolution with the Finite Elements Method of many multi-physical problems require more than one mesh. On the one hand, some problems require one specific mesh for each Physic (i.e. one mesh for the mechanical part and another one for the thermal part). On the other hand, in some cases, during the computation a remeshing is needed. In all these cases, the Data Transfer Method used to transfer information from one mesh to another has a great importance in the capacity to solve the problem and in the reliability of the solution. In general, two kinds of fields have to be transferred: the first is defined thanks to the nodal values (primary field) and the second one is defined at the integration points (secondary field). Currently, despite the research effort, no Transfer Method has been recognized as the best. Each method has important disadvantages.

This paper compares the more used Data Transfer Method (Element Transfer Method [1, 2]) with the Data Transfer Methods based on the Weak Form (using Mortar Element [3, 4] or Finite Volume [5, 6, 7]).

2 Definition of the problem

In this paper a problem is computed with the Finite Element Method. During the computation, a Data Transfer is needed from the mesh call old mesh to the one call new mesh. Some primary and secondary fields known on the old mesh are needed on the new mesh. The field P is known on the old mesh thanks to the primary field P^{old} . The old mesh is composed of n_e^{old} elements, n_n^{old} nodes and the nodal values of P^{old} is noted P_{\bullet}^{old} . The value of P^{old} on each element e^{old} is written as:

$$P^{old} = \sum_{j=1}^{n_n^{elem}} N_j \cdot P_j^{old} \quad (1)$$

where N_j is the shape function of the node j in the element e^{old} and n_n^{elem} the number of node of the element. The field S is estimated on the old mesh by the secondary field S^{old} . The value of S^{old} on each element e^{old} is defined by means of the value at the integration points (noted $\bullet S^{old}$). The fields P and S are evaluated on the new mesh thanks to the primary field P^{new} and the secondary field S^{new} , respectively. The aim of the Data Transfer Method is to define on the new mesh, the nodal values of P^{new} (P_{\bullet}^{new}) and the value at the integration points of S^{new} ($\bullet S^{new}$). The properties of the Data Transfer Method should be:

- weak numerical diffusion,
- conservation of the extrema,
- easily treatment of the boundaries.

3 Transfer Methods

The Data Transfer Method makes the link between the two discretisations. The reliability of the field on the new mesh is directly linked with the Data Transfer Method used.

3.1 Element Transfer Method

The Element Transfer Method (ETM) is the most commonly used. The computation of the field on the new mesh is done in two steps:

- Firstly, for each characteristic point (node or integration point) of the new mesh a search is done to find the element of the old mesh e^{old} in which the characteristic point lies inside.
- For the secondary field S^{old} , the nodal values of the element e^{old} (${}^e S_{\bullet}^{old}$) are computed by extrapolation of the values at the integration points of this element ($\bullet S^{old}$). So, in the general case, these nodal values are different for each element. Then, the

value of the field on this characteristic point is computed by interpolation of the nodal values of the element e^{old} ($[1, 2]$), as:

$$P_{\bullet}^{new} = \sum_{j=1}^{n_n^{elem}} N_j \cdot P_j^{old} = P^{old}(x) \quad (2)$$

or :

$$\bullet S^{new} = \sum_{j=1}^{n_n^{elem}} N_j \cdot {}^e S_j^{old} \quad (3)$$

This method does not deal with the elements of the new mesh. These elements have no influence on the value of the field. The transfer is done from the old mesh to a characteristic point of the new mesh. On the one hand, for the secondary field, this method does not conserve the extrema because of the extrapolation on the nodal values. On the other hand, the extrema are conserved for the primary field, but due to geometrical approximations, some nodes on the boundary can be outside of the old mesh. So, a special treatment is required for the boundaries.

3.2 Mortar Element Transfer Method

The Mortar Element Transfer Method (METM) is based on a weak conservation form of the field (using Mortar Element [3]). The field on the new mesh are not directly computed at the characteristic points. But, it is evaluated considering that the integral of the difference between the value of the fields on the new mesh and the value on the old mesh is null ([3, 4, 8]). This integral is done over the new mesh. To compute this integral, for each element, the nodal values of the secondary field (S^{old} or S^{new}) are defined (in function of the values at the integration points of the element). The nodal values are noted ${}^e S_{\bullet}^{old}$ for an element of the old mesh and ${}^e S_{\bullet}^{new}$ for an element of the new mesh. The value of these fields on each element is defined like the primary field (see equation 1, but in the general case, these nodal values are different for each element). The value of the primary field P^{new} on the new mesh is done thanks to the relation:

$$\sum_{e^{new}=1}^{n_e^{new}} \int_{e^{new}} (P^{new} - P^{old}) f \, de = 0 \quad (4)$$

and the relation between the secondary field S^{new} and S^{old} is defined for each element e^{new} of the new mesh by:

$$\int_{e^{new}} (S^{new} - S^{old}) f \, de = 0 \quad (5)$$

where f is a weighting function defined on each element e^{new} (like a primary field). The nodal value of the function f , f_{\bullet} can take any value.

The transfer relation of the secondary field (equation (5)) can be written like:

$$\sum_{A=1}^{n_n^{new}} \left(\sum_{B=1}^{n_n^{new}} N_{AB}^1(e^{new}) S_B^{new} - N_A^2(e^{new}, S^{old}) \right) = 0. \quad (6)$$

Where $N_{AB}^1(e^{new})$ and $N_{AC}^2(e^{new}, S^{old})$ are the mortar elements linked with the element e^{new} defined as:

$$N_{AB}^1(e^{new}) = \int_{e^{new}} N_A N_B de \quad N_{AC}^2(e^{new}, S^{old}) = \int_{e^{new}} N_A S^{old} de \quad (7)$$

where N_i is the shape function of node i (A or B) in the element e^{new} . N_{AC}^2 is the coupling term.

The transfer relation of the primary field (equation (4)) can be written like:

$$\sum_{A=1}^{n_n^{new}} \left(\sum_{B=1}^{n_n^{new}} N_{AB}^1 P_B^{new} - \sum_{C=1}^{n_n^{old}} N_{AC}^2 P_C^{old} \right) f_A = 0. \quad (8)$$

Where N_{AB}^1 and N_{AC}^2 are the mortar elements defined as:

$$N_{AB}^1 = \sum_{e^{new}=1}^{n_e^{new}} N_{AB}^1(e^{new}) \quad N_{AC}^2 = \sum_{e^{new}=1}^{n_e^{new}} N_A^2(e^{new}, N_C) \quad (9)$$

where N_C is the shape functions of the node C in the corresponding element e^{old} .

3.2.1 Computation of mortar element

The first Mortar Element (N_{AB}^1) is computed by numerical integration over the element of the new mesh e^{new} (because is a product of two shape functions of this element). The evaluation of the coupling term (the second Mortar Element, $N_A^2(e^{new}, \bullet)$) is more complex. Because, in the general case, the field S^{old} is not continue on each element of the new mesh. In addition, the sum of the shape functions N_C^{old} on each element of the old mesh is not a polynomial function on each element of the new mesh (e^{new}). A numerical and an exact integration are used to compute this Mortar Element.

Numerical integration The mortar element is computed by numerical integration over each element of the new mesh. For the element of the new mesh e^{new} , the computation is done with n^{ip} integration points. The numerical integration supposes that the value of the field on the old mesh can be evaluated by a polynomial function on each element of the new mesh.

Exact integration Each element e^{new} of the new mesh is divided in n_e^{sub} elements, as each sub-element (e^{sub}) is only over one element of the old mesh. So, on each element e^{sub} , N_C^{old} is a polynomial function. Finally, the mortar element is computed exactly by numerical integration over each sub-element e^{sub} (because is an integration of polynomial function). The exact integration of mortar element considers all intersections between the element of the new mesh and the elements of the old mesh. However, with the numerical integration, the mortar elements are evaluated in function of the intersection in which the integration point lies inside. So, the intersections than are smaller than the influence area of the integration point can be ignored.

3.2.2 Evaluation of the field on the new mesh

Global solving (GS) The relation between the nodal value of the field S^{new} on each element of the new mesh and the value of the field S^{old} (equation (6)) can be written as:

$$\begin{bmatrix} N_{11}^1(e^{new}) & \cdots & N_{1n_n^{elem}}^1(e^{new}) \\ \vdots & \ddots & \vdots \\ N_{n_n^{elem}1}^1(e^{new}) & \cdots & N_{n_n^{elem}n_n^{elem}}^1(e^{new}) \end{bmatrix} \begin{bmatrix} S_1^{new} \\ \vdots \\ S_{n_n^{elem}}^{new} \end{bmatrix} = \begin{bmatrix} N_1^2(e^{new}, S^{old}) \\ \vdots \\ N_{n_n^{elem}}^2(e^{new}, S^{old}) \end{bmatrix} \quad (10)$$

The size of this equation is equal to the number of node of the element of the new mesh. The value at each integration point is equal to the interpolation of the nodal values. This equation can be solved for each element of the new mesh.

For the primary field, the equation (8) can be written as:

$$\begin{bmatrix} N_{11}^1 & \cdots & N_{1n_n^{new}}^1 \\ \vdots & \ddots & \vdots \\ N_{n_n^{new}1}^1 & \cdots & N_{n_n^{new}n_n^{new}}^1 \end{bmatrix} \begin{bmatrix} T_1^{new} \\ \vdots \\ T_{n_n^{new}}^{new} \end{bmatrix} = \begin{bmatrix} \sum_{C=1}^{n^{old}} N_{1C}^2 T_C^{old} \\ \vdots \\ \sum_{C=1}^{n_n^{old}} N_{n_n^{new}C}^2 T_C^{old} \end{bmatrix} \quad (11)$$

The size of this equation is equal to the number of node of the new mesh. In addition, the solution of this equation cannot certify the conservation of the extrema.

Local solving (LS) To obtain a local system, a diagonal matrix is used. The value of the diagonal term is equal to the sum of the line (or the column, because the matrix is symmetric). This is totally equivalent to the row-sum technique used to lump mass matrix in explicit time integration method.

So the value of the field S on each node of the element e^{new} is done by:

$$S_A^{new} = \frac{N_A^2(e^{new}, S^{old})}{\sum_{B=1}^{n_n^{elem}} N_{AB}^1(e^{new})} \quad (12)$$

The value at each integration point of the element e^{new} is computed by interpolation of the nodal values.

And, the value of the field P on each node is done by:

$$P_A^{new} = \frac{\sum_{C=1}^{n_n^{old}} N_{AC}^2 P_C^{old}}{\sum_{B=1}^{n_n^{new}} N_{AB}^1} \quad (13)$$

With this method, the weak conservation of the field is done in a cell composed of the elements of the new mesh including the node A . This technique increases the area of computation of nodal value and in the same time the numerical diffusion. But, in opposition of the global solving, the local solving conserves the extrema.

To sum up, the evaluation of the field is done at the node of the new mesh on a function of the shape function of the elements of the new mesh and the value of the field on the elements of the old mesh.

3.3 Finite Volume Transfer Method

With the Finite Volume Transfer Method (FVTM), each field is rebuild thanks to a finite volume mesh (called old finite volume mesh on the old finite element mesh and new finite volume mesh on the new finite element mesh). The finite volumes are called cells. For the primary field, the cells are based on the node of the finite elements (see figure 1). On the finite volume mesh build for a secondary field, the cells are based on the integration points of the elements of the finite element mesh (see figure 1). So, each field is transferred from one old finite volume mesh to a new finite volume mesh. The same procedure is used to transfer the primary and the secondary field (to more information see [5]).

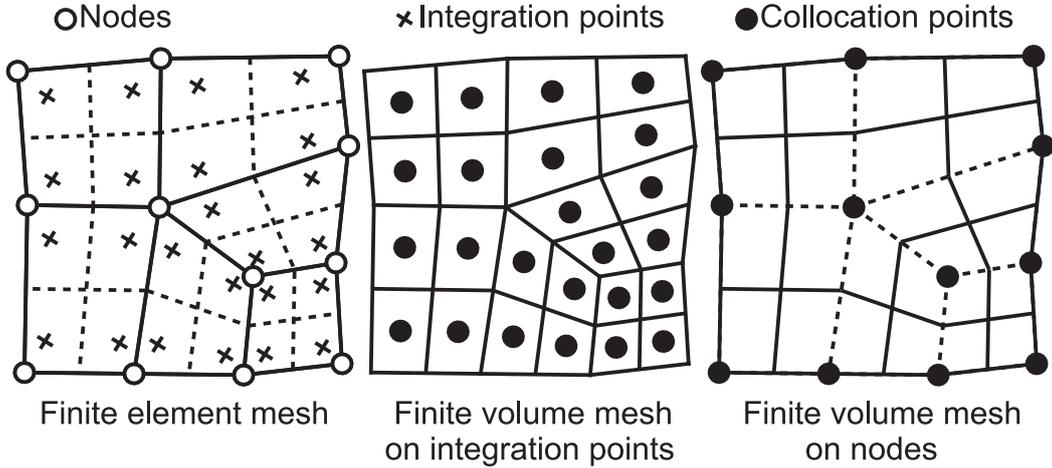


Figure 1: Finite element mesh and finite volume mesh based on integration points and on the nodes

The value of the field φ^{new} (primary or secondary) on one cell is equal to the value of the field at the corresponding characteristic point (node or integration point) of the finite

element mesh. The value of the of φ^{new} on one cell c^{new} of the new finite volume mesh (φ_{\bullet}^{new}) is done by:

$$\varphi_{c^{new}}^{new} = \frac{\int_{c^{new}} \varphi^{old} dc}{V_c^{new}}. \quad (14)$$

Where φ^{old} is the value of the field on the old finite volume mesh. Like with the mortar element, a numerical or an exact integration can be used to evaluate the coupling between the cells of the old and the new mesh.

Numerical integration The value of the field φ^{new} on the cell c^{new} is defined by numerical integration over this cell. This computation supposes that the field on the old finite volume mesh can be evaluated by a polynomial function on this cell.

Exact integration A super-mesh is built, each cell c^{new} of the new mesh is divided in n_{sub}^c cells, like each of them corresponds only to one cell of the old mesh. For each cell c^{new} the value of $\varphi_{c^{new}}^{new}$ is done by:

$$\varphi_{c^{new}}^{new} = \frac{\sum_{c^{sub}=1}^{n_{sub}^c} V_{c^{sub}} \times \varphi_{c^{old}}^{old}}{V_c^{new}}. \quad (15)$$

Where $\varphi_{c^{old}}^{old}$ is the value of the cell of the old mesh corresponding to the sub-cell c^{sub} . $V_{c^{sub}}$ is the value of the volume (the surface in two dimensions) of the sub-cell c^{sub} . The exact integration of coupling considers all intersections between the cell of the new mesh and the cells of the old mesh.

4 Examples

The difference between these Data Transfer Methods is shown on two dimensional academic examples. These examples expose the numerical diffusion and the evaluation of the data on the boundaries. The meshes are composed of quadrilateral elements. The evaluation of the Mortar Element N_{AB}^1 is done by numerical integration using two Gauss points in each direction. For the numerical integration, the coupling elements (Mortar Element N_{AC}^2 or coupling between cells) are evaluated with five Gauss points in each direction. For the exact integration, the evaluation of the coupling elements is done using six Gauss points on each triangle of the subdivision (to exact integration of quadratic function). These two examples are be used to compare the transfer of a primary field with the Transfer Element Method and the Transfer with Mortar Element in [9].

4.1 Numerical diffusion

The numerical diffusion lies to the Data Transfer Operator is studied with this example. A primary and a secondary field are known on the square. The square's sides are meshed by 30 elements. A rotation of $\pi/8$ is apply on this square. The Data is transferred from

the moved mesh to the initial mesh. The initial value of the field is of a hundred inside a circle and null outside. The centre of the circle is identical to the centre of the square and the radius is the half of the side's square.

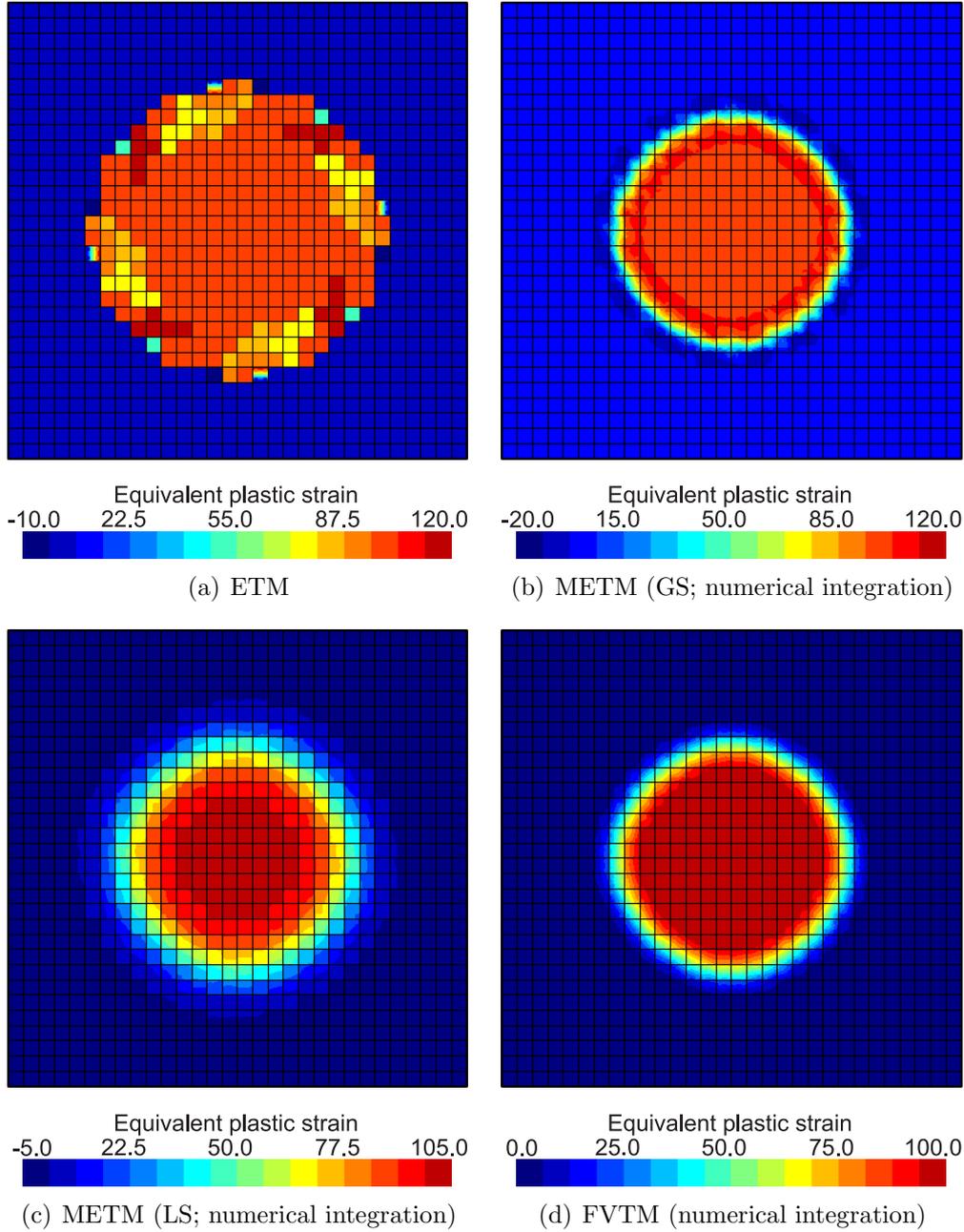


Figure 2: Numerical diffusion after twenty transfers of the secondary field (4 Gauss points)

The comparison of the numerical diffusion is done after twenty transfers (see figures 2(a), 2(b), 2(c) and 2(d)). These figures show that the numerical diffusion is minimized with the METM with global solving and the FVTM (the both with numerical computation of coupling terms). But the METM with global solving does not conserve the extrema, this Transfer Method introduces oscillations around a steep variation (see figure 2(b)). In this case, the difference between the numerical and the exact integration of coupling terms is not significant.

4.2 Influence of boundaries

To study the influence of the boundaries, a mesh of a disc with a hole is used. Like the first problem, the transfer is done after rotation of $\pi/8$ from the moved mesh to the initial mesh. The half circle is divided in fifteen elements and the radius in ten. The exact value of the field is a linear function of the abscissa (from zero to hundred). The difficulty is that in the general case, the nodes on the boundaries of the new mesh are not inside an element of the old mesh.

Figures 3 show the value of the field after sixteen transfers (one revolution). This problem proves that the ETM applied to a primary field requires a special technique to deal with the boundaries. The nodes located on the boundaries of the new mesh do not lie inside any element of the old mesh. This problem does not appear with the transfer of a secondary field (see figures 3), because the computation of the field is done on the integration points of the new mesh and these points generally lie inside of an element of the old mesh. In addition, the computation of the coupling terms by numerical integration does not consider the part of the elements that is outside of the other mesh. This explains that the METM with global solving does not introduce any error after the transfer (see figure 3(b)). The error after the transfer with the other Transfer Methods is a numerical diffusion and not a wrong evaluation of the field on the boundaries (see figures 3(c), 3(a) and 3(d)). The exact integration of the coupling terms introduces an error of space discretisation of the boundaries. The parts of the element of the old mesh that do not lie inside any element of the new mesh are not considered on the Mortar Elements. In the same time, the parts of the element of the new mesh that do not lie inside any element of the old mesh are considered and the value of the field inside is null. So, the integral of the field over the new mesh is not equal to the integral over the old mesh. This error impairs the quality of the solution.

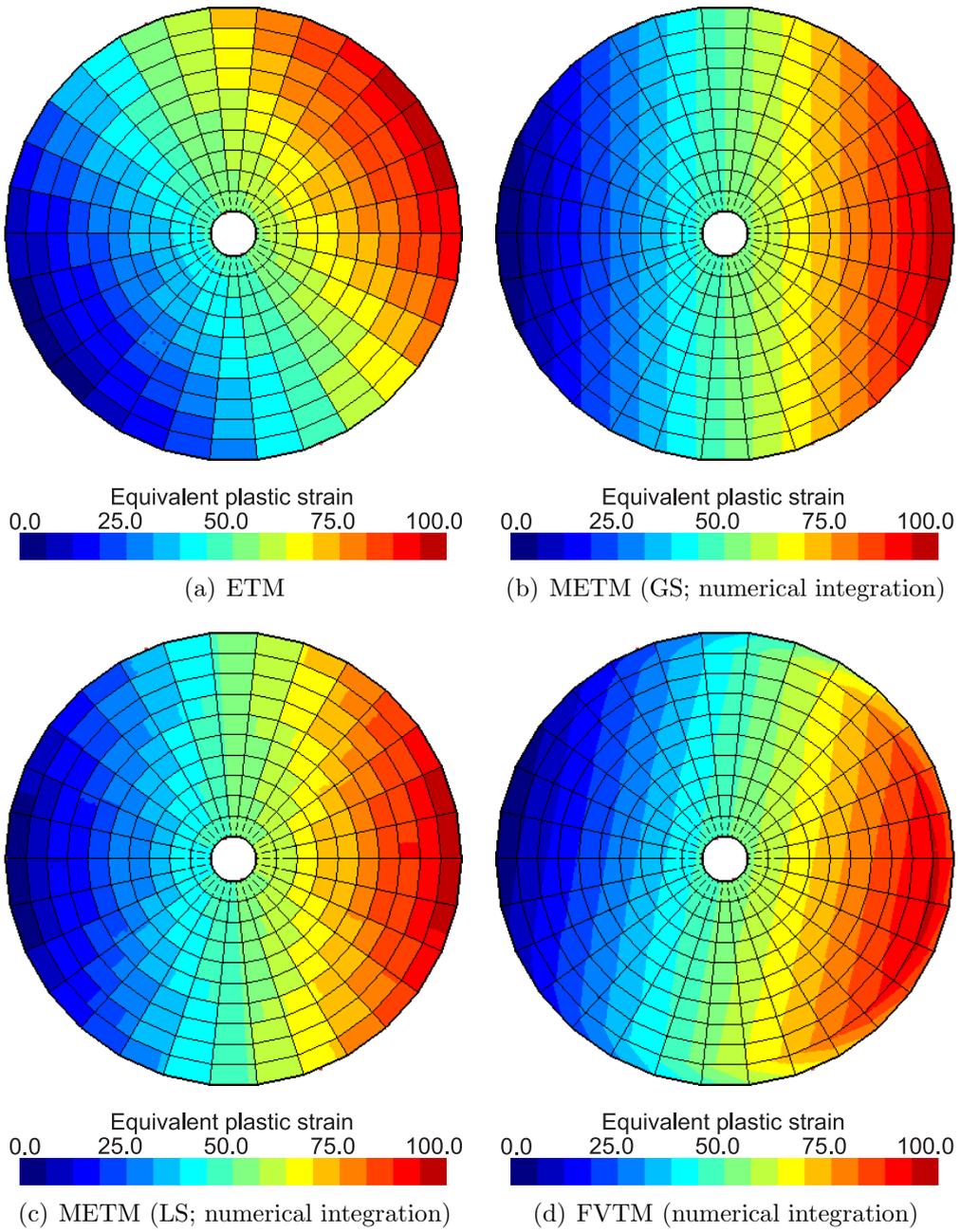


Figure 3: Numerical diffusion after sixteen transfers of the secondary field (4 Gauss points)

5 Conclusion and future works

In conclusion, this paper presents a comparison of Data Transfer Methods between two different meshes. The Data Transfer Methods based on the Weak Form (the Mortar

Element Transfer Method and the Finite Volume Transfer Method) are compared to the more used, the Element Transfer Method. With these methods (METM and FVTM) the value of the field at one characteristic point (a node or an integration point) of the new mesh is a function of the value of field on the old mesh and the elements of the new mesh. This paper shows that with the numerical integration of the coupling terms, these methods deal with complex boundaries without any specific procedure. In addition, the METM with global solving minimizes the numerical diffusion, but the global computation can introduce oscillations around steep variations of the field. So, this method cannot conserve the extrema. On the other hand, the local computation increases the smoothing of the field.

These Data Transfer Methods are applied to couple the simulation of the friction stir welding and the numerical simulation after the end of the welding. A remeshing is done after the end of the welding to continue the computation on a more appropriate mesh.

Acknowledgements

The authors wish to acknowledge the Walloon Region for its financial support to the STIRHETAL project (WINNOMAT program, convention number 0716690) in the context of which this work was performed.

REFERENCES

- [1] P.H. Saksono and D. Perić. On finite element modelling of surface tension. *Comput. Mech.*, 38:251–263, 2006.
- [2] M. Ortiz and J.J. Quigley. Adaptive mesh refinement in strain localization problems. *Computer Methods in Applied Mechanics and Engineering*, 90:781–804, 1991.
- [3] D. Dureisseix and H. Bavestrello. Information transfer between incompatible finite element meshes: Application to coupled thermo-viscoelasticity. *Comput. Methods Appl. Mech. Engrg.*, 195:6523–6541, 2006.
- [4] A. Orlando. *Analysis of adaptative finite element solutions in elastoplasticity with reference to transfer operation techniques*. PhD thesis, University of Wales, 2002.
- [5] R. Boman. *Développement d'un formalisme Arbitraire Lagrangien Eulérien tridimensionnel en dynamique implicite. Application aux opérations de mise forme*. PhD thesis, University of Liege, 2010.
- [6] A. Orlando and D. Perić. Analysis of transfer procedures in elastoplasticity based on the error in the constitutive equations: Theory and numerical illustration. *Int. J. Numer. Meth. Engrg.*, 60:1595–1631, 2004.

- [7] F. Alauzet and G. Olivier. An l^∞ - l^p space-time anisotropic mesh adaptation strategy for time-dependent problems. In *Proceedings of the IV European Conference on Computational Mechanics*, Paris, France, 16-21 may 2010.
- [8] P.E. Farrell, M.D. Piggott, C.C. Pain, G.J. Gorman, and C.R. Wilson. Conservative interpolation between unstructured meshes via supermesh construction. *Comput. Methods Appl. Mech. Engrg.*, 198:2632–2642, 2009.
- [9] P. Bussetta and J.-P. Ponthot. Comparison of field transfer methods between two meshes. In *Proceedings of the IV European Conference on Computational Mechanics*, Paris, France, 16-21 may 2010.

Thermomechanical simulation of roll forming process based on the coupling of two independent solvers

Y. Carretta*, R. Boman*, A. Stephany*, T. Bouache^ψ, R. Canivenc^ψ, P. Montmitonnet^ψ,
N. Legrand[†], M. Laugier[†] and J.-P Ponthot*

* LTAS-MN²L – Computational Mechanics,
University of Liège 1, Chemin des Chevreuils Bât. B52/3, B-4000 Liège 1, Belgium,
phone : +32-43/66.93.10 e-mail : jp.ponthot@ulg.ac.be – Web page : <http://metafor.ltas.ulg.ac.be/>

^ψ CEMEF – Centre for Material Forming Mines-ParisTech
Rue Claude Daunesse – BP 207 – 06904 Sophia-Antipolis CEDEX – FRANCE
e-mail : Pierre.Montmitonnet@mines-paristech.fr – Web page : <http://www.cemef.mines-paristech.fr/>

[†]Downstream Processes dept., ArcelorMittal Maizières R&D Industrials operations
Voie Romaine BP 30320, F-57283, Maizières-lès-Metz CEDEX France
e-mail : {nicolas.legrand, maxime.laugier}@arcelormittal.com

Key words: cold rolling, mixed lubrication, thermomechanical coupling

Abstract. This paper deals with an industrial application of a coupling procedure involving two different computer programs : MetaLub and ThermRoll. The first one models the mechanical phenomena occurring in the roll-bite during cold rolling such as elasto-plastic strip strains, elastic-roll deformations, asperity flattening and lubricant flow. The second one uses the previous results to compute the steady-state temperature field of the work-roll and the strip. These two codes are briefly described and a model of an industrial stand demonstrates the importance of the thermomechanical coupling in order to get a better understanding of the process.

1 INTRODUCTION

The general framework of this paper is in the field of numerical simulation of lubrication in cold strip rolling. This process implies many complex mechanisms at different scale levels that prevail in the roll bite which is working in mixed lubrication regime. As this mixed type of contact condition is not yet fully understood from the physics point of view, numerical models are as essential as ever to get a better understanding of this industrial process. In this paper, the coupling of two independent computer programs, MetaLub and ThermRoll, devoted to strip rolling simulations, is presented.

MetaLub [1-10] is based on the slab method. It iteratively solves the one-dimensional equations resulting from the longitudinal discretisation of the strip and from a coupled model of mixed lubrication at the interface. This lubrication model takes into account the evolution of the oil film thickness as well as the asperity crushing along the roll bite. An adiabatic model can be used to compute the temperature rise in the strip due to its deformation. However, the thermal conduction as well as heat transfer between the roll and the strip cannot

be taken into account. Nevertheless, in some cases, these heat transfers may have a huge impact on the rolling conditions. To model heat transfer during strip rolling, a software named ThermRoll, has been developed at CEMEF [11]. It predicts steady-state temperature field in the work-rolls and in the strip for several cooling configurations. ThermRoll uses the results of MetaLub to compute plastic and frictional heat generations and then solves the heat transfer equations for a given mechanical state. ThermRoll provides two temperatures profiles : one for the strip and one for the rolls. These profiles are then used by MetaLub for further iterations. An automatic external iterative procedure has been implemented to fully automate this coupling between the two codes. This method implies an exchange of result files between MetaLub and ThermRoll during the iterations. These operations are greatly simplified by thanks to the Python language [12].

The coupling between these two computer programs was used to model the behaviour of the last stand of an ArcelorMittal rolling mill. The following sections briefly describe the features of the mechanical and thermal models and the assumptions made in order to be as close as possible of the real rolling conditions encountered in the stand. Then, the results obtained with these two codes are discussed and compared to measurements.

2 METALUB MODEL

In MetaLub, the strip equilibrium is solved by the classical slab method [13-14] which computes a mean value, through the thickness of the strip, of stresses, strains and velocity at each point along the roll bite. For the purpose of this application, the mechanical behaviour of the steel strip is described by an elastic perfectly-plastic law with a yield limit σ_0 of 778 MPa. Tension and strip thickness at the entry and the exit of the roll-bite are shown in Table 1.

Table 1: Tension and thickness in the strip at the entry and the exit

	Tension [MPa]	Thickness [mm]
Entry	155	0.349
Exit	110	0.210

The lubricant viscosity η is modelled using the Williams-Landel-Ferry law [15]. The numerical values corresponding to the various parameters of the WLF law are listed in Table 2. $T_g(p_l)$ is the glass transition temperature at the fluid pressure p_l .

$$\log \eta(T, p_l) = \log \eta_g - \frac{C_1 [T - T_g(p_l)] F(p_l)}{C_2 + [T - T_g(p_l)] F(p_l)} \quad (1)$$

$$\begin{cases} T_g(p_l) = T_g(0) + A_1 \ln(1 + A_2 p_l) \\ F(p_l) = 1 - B_1 \ln(1 + B_2 p_l) \\ \eta_g = 10^{12} Pa.s \end{cases} \quad (2)$$

Table 2: Parameters of the WLF law used to model the viscosity of the lubricant η

A ₁ [°C]	A ₂ [Pa ⁻¹]	B ₁ [-]	B ₂ [Pa ⁻¹]	C ₁ [-]	C ₂ [°C]	T _g (0) [°C]
3,65	4,64.10 ⁻⁹	0,24381	1,55.10 ⁻⁸	15,91	27,868	-108,18

The local average friction stress is computed using the following sharing law

$$\tau(x) = A\tau_a + (1 - A)\tau_b \quad (3)$$

where A is the relative contact area between the strip and the work-roll, τ_a is the shear stress due to solid-to-solid contact, and τ_b is due to friction between the lubricant trapped between the strip asperities and the work-roll surface. The value of τ_a is computed using a Coulomb law with a limiting stress :

$$\tau_a = \min\left(\mu_{Coulomb} \cdot \sigma_n, \mu_{Tresca} \frac{\sigma_0}{\sqrt{3}}\right) \quad (4)$$

The parameter $\mu_{Coulomb}$ is set to match the experimental results (forward-slip i.e. the relative velocity difference between the strip and the work-roll at the exit of the roll-bite) and it is assumed that $\mu_{Tresca} = 2.6 \mu_{Coulomb}$. τ_b depends on the lubricant viscosity, the lubricant film thickness, the relative contact area and the relative velocity between the roll and the strip.

The relative contact area A is computed using Wilson and Sheu's analytical crushing law [16]. A Christensen's profile [17] is used to model the distribution of the asperities heights . This is an approximation of a Gaussian profile, more realistic since it eliminates points at infinite distance from the centreline. The arithmetic roughness of the work-rolls at stand 4 and 5 are $R_{a,rolls,4} = 0.65 \mu\text{m}$ and $R_{a,rolls,5} = 0.5 \mu\text{m}$ respectively. The roughness of the strip between stand 4 and 5 is assumed equal to the roll roughness of stand 4. The corresponding composite roughness R_q , which is a quadratic mean of the work-roll and strip quadratic roughness, is $1.03 \mu\text{m}$.

The rolls diameter is 515 mm and the rolling velocity is 1000 meters per minute. Work-rolls deformations are taken into account since, during the simulations, they appeared to be of primary importance in the heat generation. Indeed, rolls' deformations modify, among others, $Q_{Plastic}$ and $Q_{Friction}$ profiles. Jortner's [18,19] formulation is chosen to model this deformation.

3 THERMROLL MODEL

ThermRoll is a 2D thermal numerical model based on the Finite Volume method, which computes the temperature fields in the roll and the strip in the steady-state rolling regime.

The plastic heat generated in the strip $Q_{Plastic}$ as well as the frictional heat at strip/roll contact $Q_{Friction}$ are computed from the results of the mechanical simulation performed with MetaLub. Frictional heat at strip/roll contact is equally divided between the roll and the strip.

All boundary conditions are linear, with the heat flux ϕ given by:

$$\phi = H \cdot (T - T_{ref}) \quad (5)$$

T_{ref} is the temperature of the local counterpart, solid or fluid. The heat transfer coefficients H are defined separately for the roll/air contact, roll cooling water contact, strip/roll contact, work-roll/back-up roll contact, strip/strip coolant jet contact... In each zone, a constant value, or a position-dependent function for a maximal flexibility, can be given. The solid contact $H_{contact}$ can be made pressure- and temperature-dependent, using for instance Tseng's correlation [20]:

$$H_{contact} = 3800 \cdot \frac{2\lambda_{Roll}\lambda_{Strip}}{\lambda_{Roll} + \lambda_{Strip}} \cdot R_q^{-0.257} \left[\frac{P}{P + 3 \cdot \sigma_0} \right]^{0.94} + \left(\frac{\lambda_{Lub}}{h} \right) (1 - A) \quad (6)$$

λ 's are the conductivities of the roll and strip metals and of the lubricant, R_q is the RMS roughness, σ_0 the strip yield stress. The H s values and the reference temperatures modelling the thermal behaviour of stand 5 are schematically represented in Figure 1 and listed in Table 3.

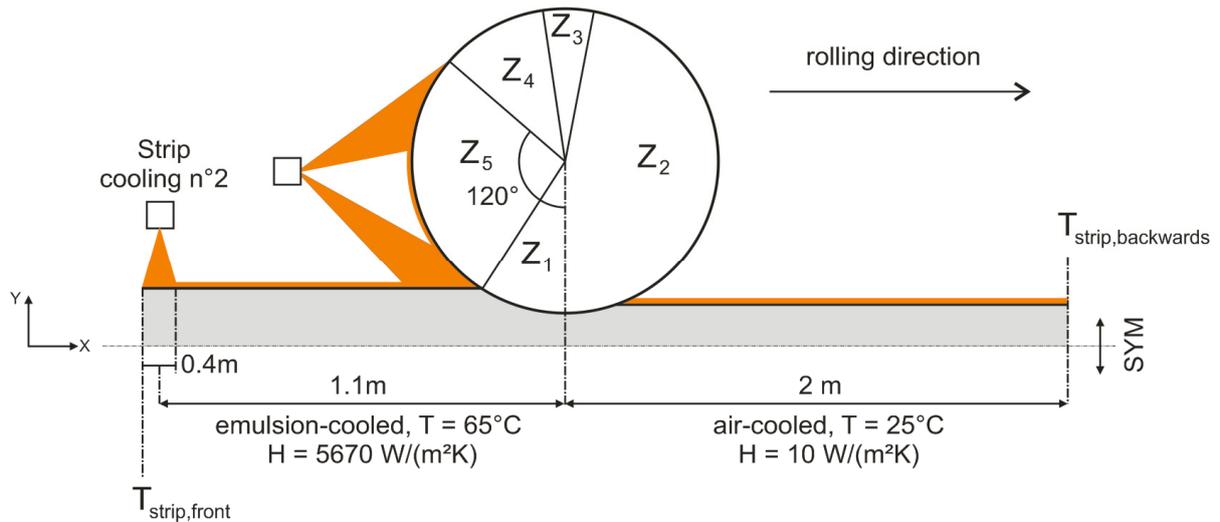


Figure 1: Schematic illustration of the thermal model of the fifth stand

Table 3: Inputs of the work-roll thermal model of the fifth stand

	Z_1	Z_2	Z_3	Z_4	Z_5
H ($\text{Wm}^{-2}\text{K}^{-1}$)	2.10^6	10	2.10^5	10	18000
T_{ref} ($^{\circ}\text{C}$)	-	25	40	25	65
	Roll-bite	Air-cooled	Contact with back-up roll	Air-cooled	Emulsion cooled

4 COUPLING PROCEDURE BETWEEN MECHANICAL AND THERMAL COMPUTATIONS

The iterative scheme involving MetaLub and ThermRoll can be seen in Figure 2. A first simulation is launched with MetaLub. Once MetaLub simulation is completed, ThermRoll comes into play using Q_{Plastic} and Q_{Friction} obtained with the mechanical model. ThermRoll computes a two-dimensional temperature field in the strip and the work-roll as well as the corresponding mean profiles; $T_{\text{strip}}(x)$ and $T_{\text{roll}}(\theta)$ along the roll-bite and the work-roll circumference respectively. These profiles are then employed in the next mechanical iteration. This procedure stops once the maximum difference between two successive strip temperature profiles $T_{\text{strip}}(x)$ is less than a tolerance specified by the user (0.5°C).

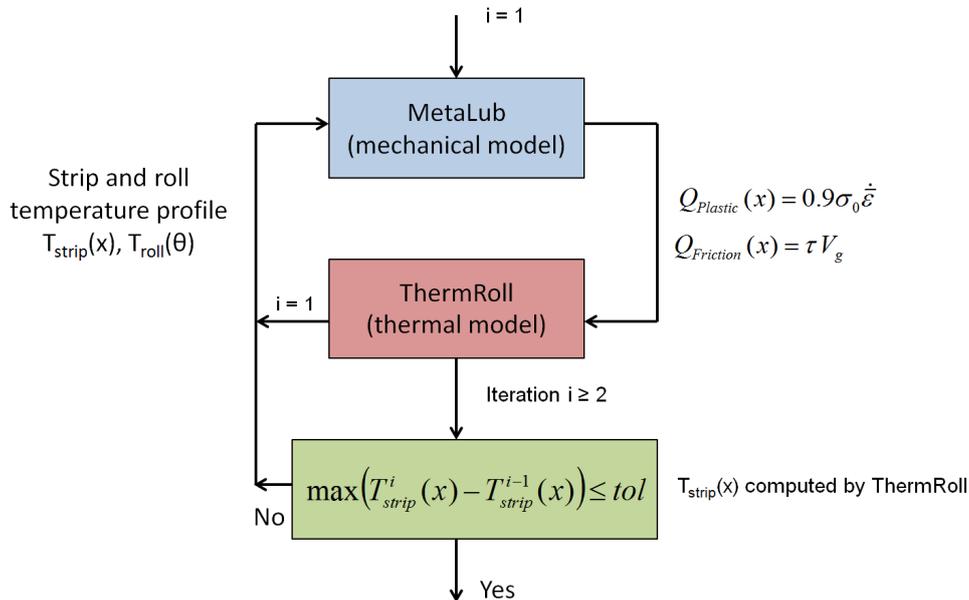


Figure 2: Staggered thermomechanical scheme

MetaLub and ThermRoll are two independent software. Moreover, the first one is written in C++ while the second one is coded in Fortran90. They essentially communicate by exchanging files; once a mechanical or thermal simulation is completed, the following is started using the results of the previous one.

5.2 Fully-lubricated configuration : spray 1 ON and spray 2 ON

In the first experiment, the strip cooling sprays 1 and 2 were enabled. In this case, fully lubricated (non-starved) conditions, which means the roll-bite absorbs as much oil as it can, were assumed in MetaLub. Two numerical parameters were set in order to match the experimental conditions. The first one is the friction coefficient $\mu_{Coulomb}$ of the mechanical model which allows us to get close to the forward slip experimental value. The second one, in the thermal model, concerns the temperature ($T_{strip,front}$) about 1.3 m prior to the roll-bite entry which is set to approach the exit temperature ($T_{strip,backwards}$) measured 2 m behind stand 5 exit. Numerical results obtained with $\mu_{Coulomb} = 0.04$, $T_{strip,front} = 118^\circ\text{C}$ are shown in Table 4. As one can see, the fully-lubricated simulations match experimental values of the forward slip and the exit temperature for an initial temperature of 118°C .

Table 4: Comparison between experimental measurement and numerical results obtained with $\mu_{Coulomb} = 0.04$ and $T_{strip,front} = 118^\circ\text{C}$ in fully lubricated conditions

	Experiment	Numerical results
Forward slip S_F	3 [%]	3.3 [%]
$T_{strip,backwards}$	153 [$^\circ\text{C}$]	152.4 [$^\circ\text{C}$]

5.3 Sub-lubricated configuration : spray 1 OFF and spray 2 ON

Turning OFF the spray 1 has two main consequences. The first is the decrease of the amount of lubricant available at the inlet of the roll-bite in stand 5. The second effect is an increase of the strip temperature before stand 5. These two effects have been separately studied in the numerical model by modifying $T_{strip,front}$ and h_0 .

In the first case, the initial temperature was kept constant ($T_{strip,front} = 118^\circ\text{C}$) and the inlet lubricant film thickness (h_0) was progressively decreased. These results correspond to the green curve in Figures 4 and 5. One can see in Figure 4 that h_0 higher than $10\ \mu\text{m}$ gives similar results as the ones obtained in a fully lubricated case. For lower initial lubricant film thickness starvation occurs. Indeed, reducing h_0 from $10\ \mu\text{m}$ to $5\ \mu\text{m}$ induces a slight rise of the forward slip of about 2.7%. Then a significant increase is observed until a film thickness of $0.1\ \mu\text{m}$ is reached. Finally for a lubricant film thinner than $0.1\ \mu\text{m}$, the forward slip is more or less constant. In Figure 5, it is observed that the value of the inlet oil film thickness does not change the exit temperature much.

In Figure 4, numerical results, corresponding to an inlet film thickness of $3.48\ \mu\text{m}$, match the forward slip experimental value measured after having turned OFF the spray 1. Nevertheless the exit temperature obtained in this case is 5°C lower than the experimental measurement.

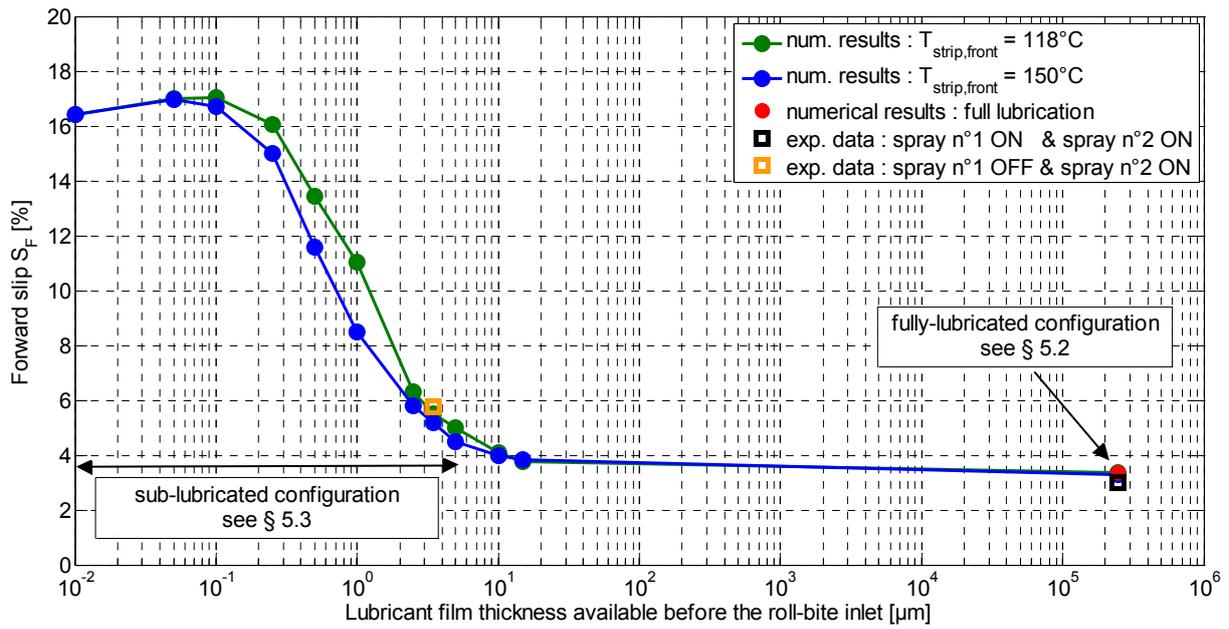


Figure 4: Forward slip corresponding to two entry temperatures ($T_{strip,front} \sim 1.3$ m ahead of the roll-bite entry) for various lubricant film thicknesses available at the inlet of the roll-bite

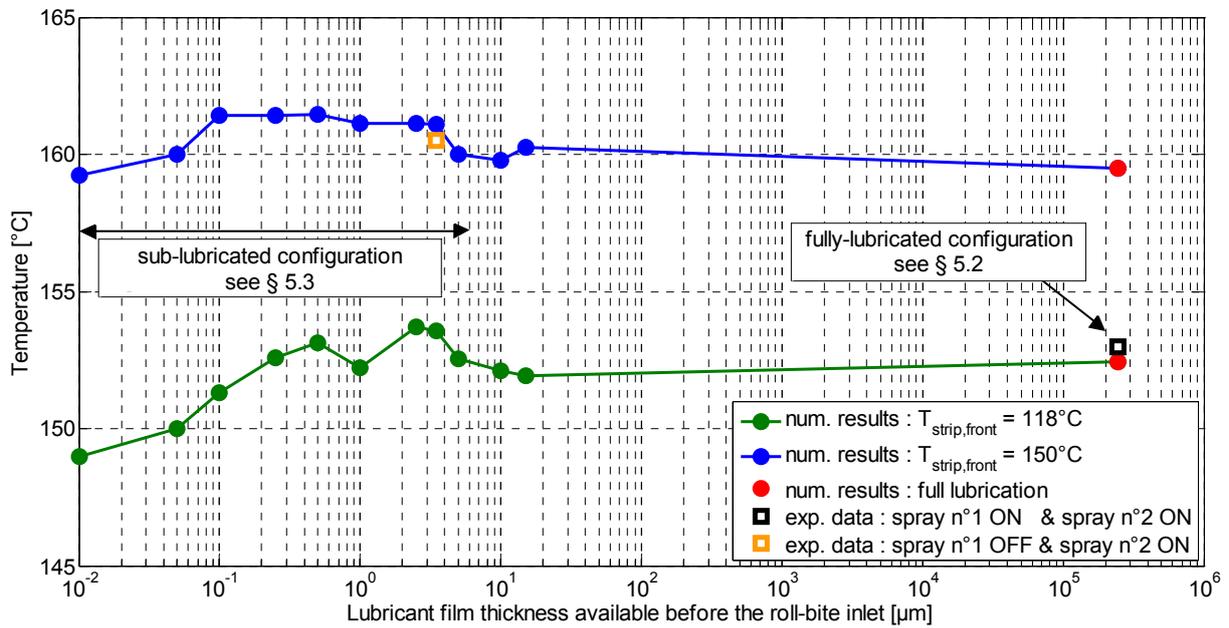


Figure 5: Exit temperature ($T_{strip,backwards}$, 2m after the roll-bite exit) corresponding to two entry temperatures ($T_{strip,front} \sim 1.3$ m ahead of the roll-bite entry) for various lubricant film thickness available at the entry of the roll-bite

In the second case, simulations were performed with $T_{strip,front} = 150^{\circ}\text{C}$. The forward slip values are close to the one obtained with $T_{strip,front} = 118^{\circ}\text{C}$. The maximum absolute difference between these two curves is less than 2%. The temperature difference at the entry of the roll-bite, after the emulsion cooling, is 10 °C. Figure 4 shows that this temperature difference, for a given inlet lubricant film thickness, does not have a huge impact on the forward slip. As one can see in Figures 4 and 5, $T_{strip,front} = 150^{\circ}\text{C}$ and a 3.48 μm initial lubricant film thickness match the forward slip and the exit temperature measured when the spray 1 is OFF.

From the analysis of Figures 4 and 5, it can be concluded that reducing the amount of lubricant at the roll-bite entry essentially affects the forward slip while the exit temperature does not vary much. On the contrary, moderate temperature change at the entry of the roll-bite does not affect much the forward slip while it has a more important impact on the exit temperature.

6 CONCLUSION

A coupled mechanical / thermal / tribological model has been created to examine the behaviour of the stand 5 of an industrial mill. A parametric study on lubrication and thermal conditions has been conducted. These numerical results have been compared to the experimental ones. This comparison gives a better understanding of the stand behaviour. Indeed, experimental results revealed that turning OFF a lubricant spray on the strip had two impacts: an increase of the forward slip, and a temperature rise at the exit of the roll-bite. The thermomechanical model highlights that these two effects are due to two different reasons: the exit temperature rise is due to an increase of the entry temperature while the increase of the forward slip is due to lower lubricant film thickness at the entry of the roll-bite.

In the two lubrication configurations investigated here, the obtained numerical results are close to experimental ones.

To sum it up, the coupling of MetaLub-ThermRoll has been proven to be a very helpful tool to better understand some complicated rolling mill configurations that can occur when the mixed lubrication regime as well as the strip temperature play a critical role on the behaviour of the process. This thermomechanical model can be a solid support for industrial cold rolling process optimisation.

7 ACKNOWLEDGEMENT

Yves Carretta would like to thank the FRIA (Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agriculture) which provides a financial support to this research project.

The Walloon region and the European Social Fund are also gratefully acknowledged for their support of Antoine Stephany through the grant FIRST-EUROPE/METALUB, convention no. 3310300R0172/215052.

REFERENCES

- [1] A. Stephany. *Contribution à l'étude numérique de la lubrification en régime mixte en laminage à froid*. PhD dissertation (in French), Université de Liège (2008).
- [2] Y. Carretta, R. Boman, A. Stephany, N. Legrand, M. Laugier and J.-P. Ponthot. *MetaLub - A Slab Method Software for the Numerical Simulation of Mixed Lubrication Regime in Cold Strip Rolling*. Proceedings of the Institution of Mechanical Engineers, Part J: Journal of Engineering Tribology, 2010, in Press.
- [3] Stephany A., Ponthot J.-P., Collette C. and Schellings, J. *Efficient algorithmic approach for mixed-lubrication in cold rolling*. Journal of Materials Processing Technology, 2004, 153-154, 307-313.
- [4] Stephany A., Ponthot J.-P., Collette C. and Gratacos P. *Nouveau concept algorithmique pour un modèle de laminage à froid en régime mixte*. Proceedings of the 6^{ème} colloque national en calcul des structures, Giens, France, 2003, Vol. III, 157-164.
- [5] Stephany A., Ponthot J.-P., Collette C. and Schellings J. *Efficient algorithmic approach for mixed lubrication in cold rolling*. Proceedings of Advances in Materials and Processing Technologies, 2003, 1085-1088.
- [6] Stephany A., Ponthot J.-P., Collette C. and Schellings J. *New algorithm concept for a model of mixed lubrication in cold rolling*. Proceedings of METEC 2003: European Rolling Conference, Düsseldorf, Germany, 2003, 411-416.
- [7] Stephany A., and Ponthot J.-P. *Improvements of an iterative and staggered Approach for a Model of Mixed Lubrication in Cold Rolling*. Proceedings of the 8th International Conference on Numerical Methods in Industrial Forming Processes, Columbus, Ohio, 2004, 406-411.
- [8] Stephany A., and Ponthot J.-P. *Numerical simulation of mixed lubrication in cold rolling using an object-oriented philosophy*. Proceeding of the Third International Conference on Advanced Computational Methods in Engineering, Ghent, Belgium, 2005.
- [9] Stephany A., Ponthot J.-P. and Legrand N. *Numerical simulation in cold rolling using adiabatic strip model: application to large roll deformation*. Proceedings of AUSTRIB 06, Brisbane, 2006.
- [10] Montmitonnet P., Stephany A., Cassarini S., Ponthot J.-P., Laugier M. and Legrand N. *Modelling of metal forming lubrication by O/W emulsions*. Proceedings of the International Conference on Tribology in Manufacturing Processes, Yokohama, Japan, 2007, 85-90.
- [11] T. Bouache, N. Legrand, P. Montmitonnet. *A numerical heat transfer analysis in mixed-film lubrication for cold strip rolling*. Proceedings of the 5th European Rolling Conference, 2009.
- [12] H.P. Langtangen. *Python scripting for computational science*. 3rd edition, Springer 2008
- [13] P. Cosse and M. Econopoulos. *Mathematical study of cold rolling*. C.N.R.M., 1968, 17, 15-32.
- [14] N. Marsault. *Modélisation du Régime de Lubrification Mixte en Laminage à Froid*. PhD dissertation (in French) Ecole Nationale Supérieure des Mines de Paris (1998)
- [15] L. Williams, R. Landel et J. Ferry. *The Temperature Dependence of Relaxation Mechanisms in Amorphous Polymers and Other Glass-forming Liquids*. J. Am. Chem. Soc., 77(14) :3701-3707, 1955.

- [16] W.R.D. Wilson and S. Sheu *Real area of contact and boundary friction in metal forming*. International Journal of Mechanical Science, 1988, 30, 475-489.
- [17] H. Christensen *Stochastic models for hydrodynamic lubrication of rough surfaces*. Proceedings of the Institution of Mechanical Engineers, 1970, 184, 1013-1022.
- [18] A. Stephany, J.-P. Ponthot and N. Legrand. *Numerical simulation in cold rolling using adiabatic strip model: application to large roll deformation*. Proceedings of AUSTRIB 06, Brisbane, 2006.
- [19] D. Jortner, D. Osterle and C. Zorowski. *An analysis of cold rolling*. International Journal of Mechanical Sciences, 1960, 2, 179-194..
- [20] A.A. Tseng. *Thermal Modeling of Roll and Strip Interface in Rolling Processes: Part 2 – Simulation*. Numer. Heat Transfer, Part A, 35 (1999) 135-154.

THIXOFORGING TOOLS MATERIALS: DETERMINATION OF APPROPRIATE FEATURES AND EXPERIMENTAL EVALUATION

AHMED RASSILI* AND JEAN-CHRISTOPHE. PIERRET*

*Thixo Unit ULg, Bd de Colonster, 4
B-4000 Liège, Belgium
e-mail: a.rassili@ulg.ac.be

Key words: Thixoforging, steel, high melting point, tool, material.

Abstract. Whereas thixoforming of low melting point alloys as aluminium or magnesium is now an industrial reality, thixoforming of high melting point alloys, as steel, is still at the research level. High working temperature, die wearing and production rate are problems that must be solved and are under investigation. The aim of this work is to evaluate the thermal and mechanical loadings applied to the tools during the steel thixoforging process in order to determine if classical hot-work tool steel could be an appropriate tool material. This evaluation has been realized thanks to experimental trials and to simulations on the finite elements code Forge2009©. The effect of the loadings on the tool's failure modes are highlighted and compared to the ones observed in classical hot forging. Beyond this, the failure modes of hot-work tool steel, the X38CrMoV5 or H11, is presented.

1 INTRODUCTION

Due to high slug temperature (usually higher than 1350°C), tools surfaces reach very high temperature. In hot forging, this temperature could already reach 500°C [1-3]; in thixoforging, it could be higher than 700°C. Such a temperature is higher than classical tool steels annealing temperature and could leads to a fall of the mechanical properties. In order to minimize the thermal shocks, dies are usually pre-heated from 40 to 350°C in hot forging.

Thixoforging process, as hot forging is composed of three sequential steps:

- Brutal contact of high temperature slug on the tool. If needed, tool closing could be done before or after this step.
- Forming step during which mechanical constraints are applied to the tool.
- Part ejection and tool cooling.

In production, these steps are repeated in a cycle. Tool damaging could be due to different mechanisms: fatigue cracking following thermomechanical loading cycle, microstructure evolution or scaling due to hot working, geometrical modification generated by wearing or plastic deformation. The common failure modes observed in hot forging are shown on Figure 1.

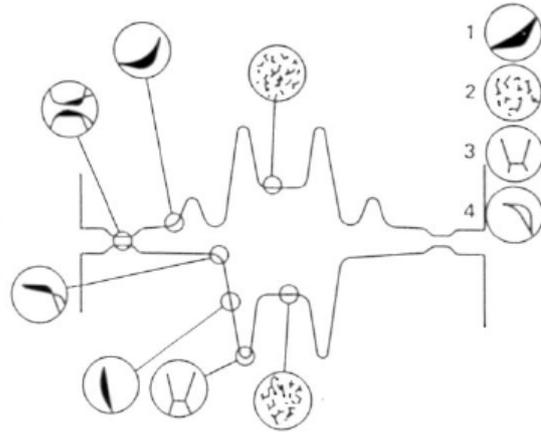


Figure 1: Most common hot forging tools failure modes and their localizations: (1) abrasive wearing, (2) thermal fatigue, (3) mechanical fatigue, (4) plastic deformation [4]

In thixoforging, thermomechanical loadings are quite different as forming loads are lower but thermal loads are higher. The failure modes could be different too. Their determination is the first aim of this paper. The resistance of classical hot forging tool steel (X38CrMoV5) to these loadings is also studied.

2 EXPERIMENTAL

2.1 Tool

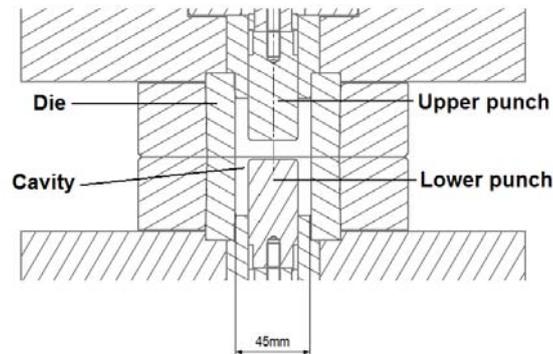


Figure 2: “H” or double-cup tool

The tool used during this work (Figure 2) is made in H11 hot work steel and forms a double-cup part (or axisymmetric H). The deformation is a compression followed by an important reverse extrusion. Due to small thickness of the walls, this geometry highlights the thermal effects occurring during forming. The dies and the punches are instrumented by thermocouples in order to measure their inner thermal fields. At the beginning of cycle, the tool is open and the punches are out of the dies. When heating is done, robot puts the slug in the lower die and moves back. Then, the upper part of the tool moves down to close it and the two punches form the part. It is also possible to form it with the upper punch alone if the lower one is already inside the die at the beginning of cycle. This possibility was used in a former work to determine friction parameters [5].

2.2 Tool material

Tool has been made of X38CrMoV5 hot working tool steel. It has a good thermal shocks resistance thanks to the presence of chrome, molybdenum and vanadium. It is commonly used as die material in hot forging [6]. The X38CrMoV5 composition is given on Table 1.

Chrome, molybdenum and vanadium make carbides which increase wearing resistance. Chrome and molybdenum delay the softening due to annealing. Chrome and vanadium inhibit the grains coarsening during austenitizing and chrome and silicon increase scaling resistance.

Table 1: Mass composition of X38CrMoV5 hot working tool steel [6]

	C	Cr	Mn	V	Ni	Mo	Si
[%]	0.40	5.05	0.49	0.47	0.20	1.25	0.92

Nevertheless, this steel grade loses a part of its mechanical properties at high temperature. Table 2 gives the mechanical properties of X38CrMoV5 for four working temperatures for a material previously oil-quenched at 1040°C after two tempering at 640°C. At 600°C, resistance is nearly divided by two. Extrapolated until 800°C, Rp0.2 falls to 400MPa, so lower than the locking force applied on the dies. Moreover, the austenitizing beginning temperature (830°C) is closed to the working one [7].

Table 2: Mechanical properties of X38CrMoV5 at different working temperatures [7]

Temperature [°C]	Rm [MPa]	Rp0.2 [MPa]	A [%]
20	1400	1170	12
400	1170	1020	13
500	1050	900	18
600	810	700	25

2.3 Modelling

The Finite Elements code Forge2009© was used for the simulations. The constitutive law used in this work is quite simple and mainly driven by the liquid fraction, and so the temperature. Thus, the structure of the raw material and its evolution are not explicitly represented. Even if this is a limitation of the calculation results, the error on the flow behavior is small for high solid fraction. Thermal exchanges are already taken into account by the FE code.

The constitutive law is a classical Spittel one (which is the default law used by the solver) when material temperature is lower than solidus and a modification of this Spittel equation when the material temperature is higher than solidus. The modification induces a linear decrease of the consistency by multiplying it by a factor going from one to zero between the solidus and the liquidus. There is then a smooth transition between semi-solid and solid behavior during cooling.

The constitutive law is

$$\sigma = A e^{m_1 T} \varepsilon^{m_2} e^{\frac{m_4}{\dot{\varepsilon}}} \dot{\varepsilon}^{m_3} \quad (1)$$

for $T < T_{\text{solidus}}$ and

$$\sigma = A \left(\frac{T_{\text{liq}} - T}{T_{\text{liq}} - T_{\text{sol}}} \right) e^{m_1 T} \varepsilon^{m_2} e^{\frac{m_4}{\dot{\varepsilon}}} \dot{\varepsilon}^{m_3} \quad (2)$$

for $T_{\text{solidus}} < T < T_{\text{liquidus}}$

In these equations, σ is the stress, ε is the strain, $\dot{\varepsilon}$ is the strain rate, T is the temperature, T_{liq} is the liquidus temperature, T_{sol} is the solidus temperature and A , m_1 , m_2 , m_3 and m_4 are constants depending on the steel grade. For 100Cr6 steel, the values of the constant parameters are given in Table 3. The values of A and m_1 to m_4 come from the database of Forge2009© and the values of T_{liq} and T_{sol} have been obtained by Differential Scanning Calorimetry (DSC) [8].

Table 3: Values of the constants used in equations (1) and (2)

Parameter	Value
A	2707.108
m1	-0.00325
m2	-0.00325
m3	0.1529
m4	-0.05494
T _{sol}	1315°C
T _{liq}	1480°C

3 RESULTS AND DISCUSSION

3.1 Mechanical loading

In the case of thixoforming, mechanical loadings are about ten to twenty times lower than in hot forging [9]. Figure 3 shows the Von Mises equivalent stresses, calculated by the Forge2009© software, at the end of forming inside the lower part of the tool. The simulated forming is a 100Cr6 steel slug symmetrically deformed with a tool speed of 170mm/s. It appears that maximum stress, for the areas in contact with the semi-solid material, is around 170MPa. This maximum stress is located in the centre of the punch top surface. If the punch temperature reaches a value for which its material yield stress is lower than 170MPa, there would be a deformation of this punch.

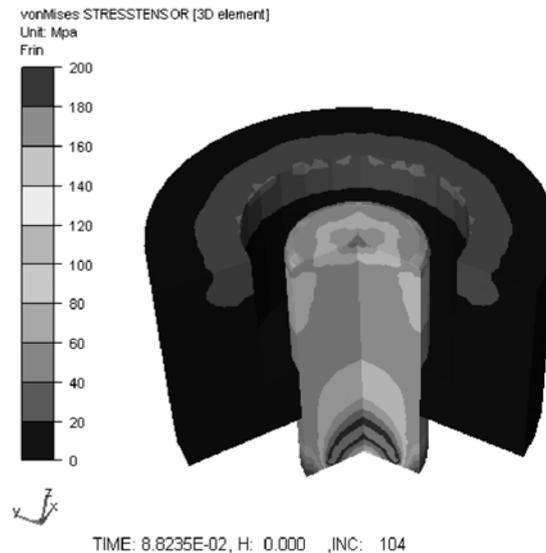


Figure 3: Von Mises equivalent stresses inside the tool at the end of forming

The accordance of the simulation total load to the recorded forming load for the upper punch is shown on Figure 4.

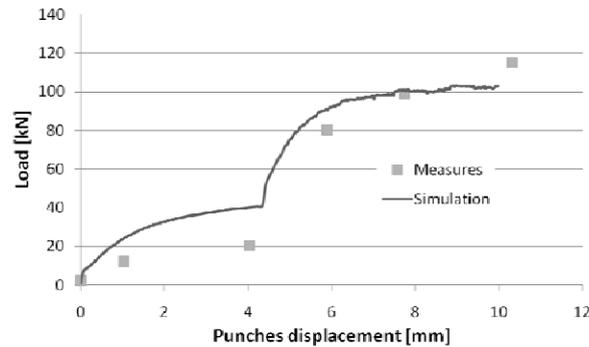


Figure 4: Comparison of the forming load measured during test and calculated by Forge2009©

The simulation did not take into account the locking force applied to the dies to keep the tool closed. In the present case, this force is 2000kN. On a surface of about 3200mm², the pressure is close to 630MPa.

3.2 Thermal loading

In hot forging, slugs are usually heated at a temperature higher than 1000°C. Their contact with the dies could heat these ones up to 500°C. In thixoforming, the working temperatures are still higher, until more than 1400°C. Tools surfaces are then subjected to very high temperature. The double-cup tool has been designed in order to be instrumented by K-type thermocouples. The measures of these thermocouples allowed validating the temperature fields calculated by simulation, as shown on Figure 5.

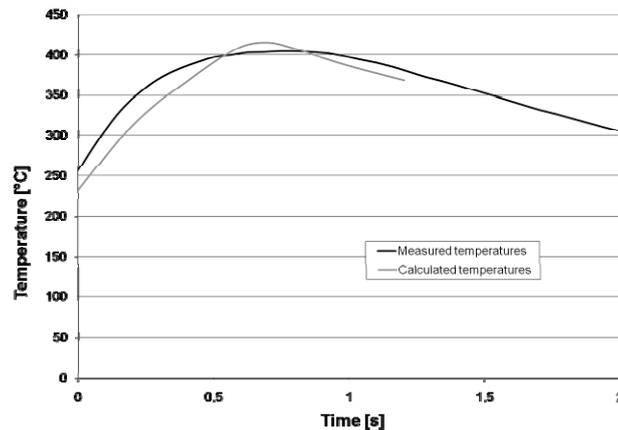


Figure 5: Comparison of the temperatures inside the tool measured by thermocouples and calculated by Forge2009©

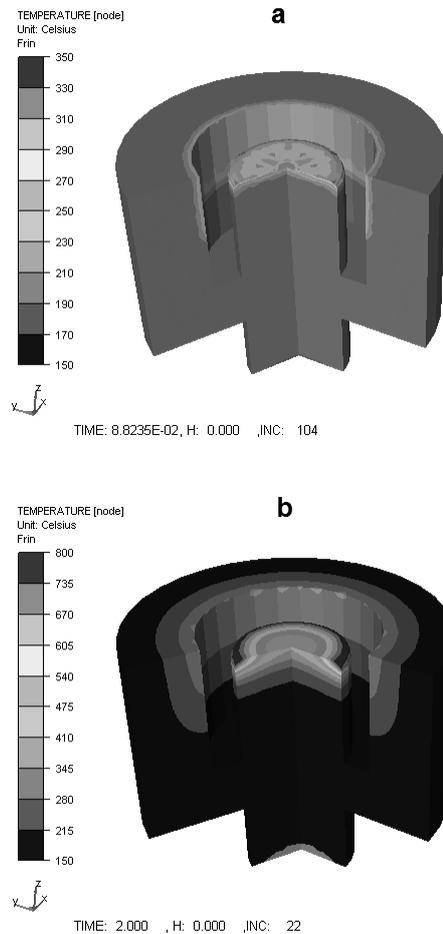


Figure 6: Temperature fields inside the tool at the end of forming (a) and at ejection time (b)

Figure 6 shows the temperature fields inside the lower part of the tool at the end of the forming and two seconds after, when the tool opens and the part is ejected. Simulation shows that, at the moment of the ejection, the surface temperature could reach 800°C. In this case, if the stresses are important, because of galling during the ejection for example the tool could easily be damaged.

Simulation has been run for hot work tool steel. In the case of another tool material, and thus another thermal conductivity, the surface temperature should be different. In the case of a lower thermal conductivity, the surface temperature would be higher, which will be interesting from the forming point of view as the flowing material temperature would stay higher during a longer time and thus, the forming load would be lower. At the opposite, from the tool point of view, this higher surface temperature would increase the risk to overrun the tool material yield stress and to damage the tool. Thermal stresses, coming from thermal gradients, depend of these gradients value and of the thermal dilatation coefficient.

3.3 Wearing

Figure 7 shows the area of maximum wearing. As in hot forging, they are located where sliding speeds are the higher, thus mainly at the punch edge. On the die, there is not any wearing at the joining plane level as the tangential speed is null on this area in the case of a symmetric deformation. As the working temperature is higher in thixoforming, the tool wearing resistance is lower than in hot forging.

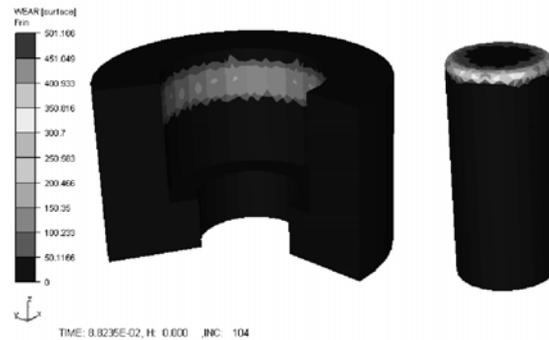


Figure 7: Wearing areas inside the tool

3.4 Hardness

Hardness of the tooling's lower punch has been measured after 50 cycles of forming. Figure 8 shows the hardness values measured on the punch and the line along which the measured have been made.

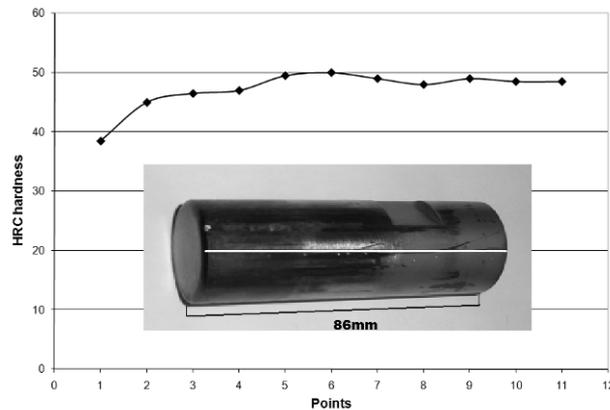


Figure 8: Hardness measured along the lower punch after 50 parts forming

The graph shows that the lower part of the punch has kept its original hardness, around 48-50 HRC. But the hardness of the last three centimetres has noticeably decreased. The closest point to the plane surface has a hardness of only 38.5 HRC. This softening is due to annealing occurring at high temperature. This means that the punch is more easily deformable. Moreover, some marks are visible on the punch surface, due to galling and abrasive wearing.

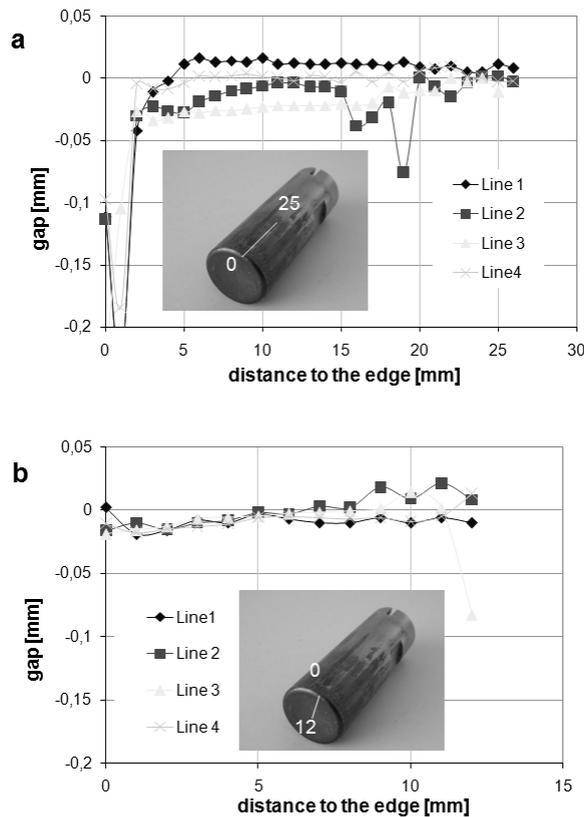


Figure 9 Profile measures on the punch after 50 formings along the lateral (a) and top (b) surfaces

3.5 Mechanical resistance

Figs. 9 and 10 give profile measures of the upper punch (Figure 9) and die (Figure 10). These profiles have been measured on four different lines in order to limit the impact of local damages.

The profile of the punch lateral surface (Figure 9a) does not show significant modification of this surface. On the top surface (Figure 9b), a small modification could be noted, as the point zero, corresponding to the punch edge, is few hundredths of millimetres lower than point 12, corresponding to the surface centre. This slight deformation is due to the

punch diving inside the semi-solid steel. It seems thus that mechanical resistance of the punch is high enough to avoid great plastic deformation.

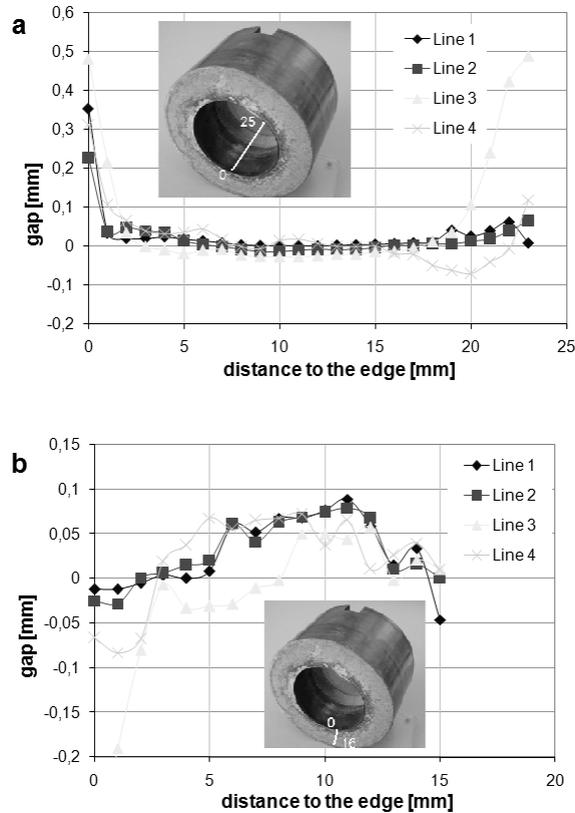


Figure 10 Profile measures on the die after 50 parts forming along the lateral inside (a) and top (b) surfaces

The top surface of the die (Figure 10b) does not show any major modification, except the usual one due to scum (residual lubricant, scaling...) crushing between the dies bearing areas during the locking. However, the inner lateral surface (Figure 10a) shows a great deformation (2 to 5 tenth of millimetre) at the joining plane level. This is due to the dies locking which induces important stresses (about 630MPa). Around this joining plane, the temperature could reach 500°C (Figure 6). In this case, the X30CrMoV5 tool steel Rp0.2 is 900MPa (Table 2), so clearly higher than the locking force. Nevertheless, one must take into account the locking force regulation. As it is not perfect, the locking load often reaches a transient value higher than 2500kN for a setting of 2000kN. In this case, the stress could reach 800MPa, what is close to the X38CrMoV5 steel Rp0.2 value. It is sufficient to create a deformation, really small but increasing from trial to trial. The deformation is thus a kind of fold of the die toward the inside at the joining plane level. This is an important issue, as the deformation acts as a reverse draft during the ejection, locking the part inside the die. At the opposite of the punch case, the hot mechanical resistance of the X38CrMoV5 hot work tool steel does not seem sufficient in this case.

An other super-thickness is visible on Figure 10a close to point 25, but this is only due to lubricant waste accumulation. This is not a die deformation.

4 CONCLUSIONS

Up to now, thixoforming tool lifetime is still the main lock to the technology industrialization. Due to high working temperature, mechanical features of the hot work tool steels classically used in hot forging strongly decrease. In particular, hardness and yield stress are too low to guarantee the tooling integrity.

Hardness could be increased by surface treatment, as nitriding, glazing [10, 11 and 12] or oxidation. The problem is to keep a good adherence of the coating or a good stability of the structural treatment.

Plastic deformation is the main issue. It is due to mechanical and thermal stresses. Compared to hot forging, mechanical stresses are clearly lower but thermal stresses are higher. However, at industrial production rate (6-12 parts per minute), the working temperature should be higher but temperature variation would be lower, so the thermal fatigue should be lower than in the case of laboratory study. An important point is also to minimize the contact time between tool and semi-solid steel in order to minimize the tool temperature. Parts ejection must then be as fast as possible to decrease thermal loading.

The X38CrMoV5 hot working tool steel could be used only for low stresses forming. Its hardness is not sufficient to avoid abrasive wearing at high working temperature.

REFERENCES

- [1] Andreis G., Fuchs KD., Schruff I., The wear behaviour of hot-work tool steels used in forging process. Proceedings of the 5th International Conference on Tooling, 1999.
- [2] Walter S., Haferkamp H., Niemeyer M., Bach FW., Henze A, Material failure mechanisms of forging dies. Proceedings of the 5th international Conference on Tooling, 1999.
- [3] Kircher D., Michaud H., Bogard V., Analyse des dégradations d'outillages de forge à chaud à l'aide de la simulation numérique. *Matériaux et Technologies*, 1-2, 1999.
- [4] ASM Handbook, Forming and Forging. ASM, 1998.
- [5] Pierret JC., Rassili A., Vaneetveld G., Bigot R., Lecomte-Beckers J., Friction coefficients evaluation for steel thixoforming. *Int J Mater Form* (2010) Vol. 3 Suppl 1:763 – 766
- [6] Barreau O., Etude du frottement et de l'usure d'aciers à outil de travail à chaud. Phd tesis, Institut National Polytechnique de Grenoble, 2004.
- [7] Leveque R., Aciers à outils, données numériques non normalisées. www.techniques-ingenieur.fr, 2004.
- [8] Lecomte-Beckers J., Rassili A., Robelet M., Poncin C., Koeune R., Study of the liquid fraction and thermophysical properties of semi-solid steels and application to the simulation of inductive heating for thixoforming. *Advanced Methods in Material Forming*, 2007.

- [9] Bigot R., Favier V., Rouff C., Characterisation of semi-solid material mechanical behaviour by indentation test. *Journal of Material Processing Technologies*, 160, 2004, 43-53.
- [10] Brabazon D., Naher S., Biggs P., Laser surface modification of tool steel for semi-solid steel forming. *Solid State Phenomena* 141-143, 2008, 255-260.
- [11] Brabazon D., Naher S., Biggs P., Glazing of tool dies for semi-solid steel forming. *International Journal on Material Forming* sup. 1, 2008, 985-988.
- [12] *Thixoforming Steel*: Editors Helen Atkinson and Ahmed Rassili, ISBN 978-3-8322-9133-4, Shaker Verlag Publications (2010).

FATIGUE ANALYSIS OF THE STRUCTURAL COMPONENTS OF A MECHANICAL PRESS CUTTING HIGH-STRENGTH STEELS

BERND-ARNO BEHRENS, RICHARD KRIMM AND CHRISTIAN WAGER

Institute of Metal Forming and Metal-Forming Machines (IFUM)
Leibniz University Hanover
An der Universität 2, 30823 Garbsen, Germany
e-mail: wager@ifum.uni-hannover.de, www.ifum.uni-hannover.de

Key words: Durability Analysis, Fatigue Analysis, Multi-Body Simulation, Coupled Simulation, Cutting Press, High-Strength Steel.

Abstract.

In this paper the fatigue analysis of the structural components of a high-speed press is presented. The fatigue analysis is based on the results of a hybrid multi-body simulation. During the hybrid multi-body simulation the press is virtually loaded by the forces occurring while cutting high-strength steel. The hybrid multi-body simulation was coupled with a durability software package for the fatigue analysis. A high cycle stress approach was applied for the fatigue analysis of the press frame, ram, con rods and eccentric shaft. In addition, critically dimensioned test structures were designed, manufactured and assembled to a high-speed press at the Institute of Metal Forming and Metal-Forming Machines to validate the fatigue analysis of the components of the press. The test structures were loaded due to the cutting of high-strength steel until crack initiation was detected by means of the dye penetrant inspection. The predicted and experimental determined fatigue life correlated. The simulation is based on design data and no calibration of the model was performed.

1 INTRODUCTION

Recently press manufacturers reported increasing risks of early crack initiation at the structural components of presses due to cutting high-strength steels. Crack growth at structural components of presses, for instance at the press frame, probably causes a sudden breakdown of the production line as well as high repair costs for the press manufacturer. Thus, an approach to determine the fatigue strength of structural components of presses is required that takes into account the specific demands while cutting high-strength steels.

A shock occurs due to the sudden drop of the cutting force when cutting high-strength steels. This shock initiates the press structure as well as assembled parts to vibrate [1]. The occurring load amplitudes and load changes are much higher than the values previously known from the cutting of conventional deep drawing materials. Therefore, the interaction between the cutting process and the cutting press must be considered for the development of future presses to realise the required fatigue strength to process high-strength steels.

A hybrid multi-body simulation (MBS) model based on design data of a high-speed press was established at the Institute of Metal Forming and Metal-Forming Machines (IFUM). Due

to the replacement of the rigid bodies of the press representing the press frame, ram, con rods and eccentric shaft by means of finite element (FE) meshes the elasticities of these components were taken into consideration during the simulation [2, 3]. In this paper the setup and validation of a fatigue analysis based on the results of the hybrid MBS are shown. The durability analyses of the structural components of the press are based on a high cycle stress approach and were carried out by means of commercial durability software package that was coupled with the MBS. It is shown that a coupled simulation entitles product engineers to determine the fatigue strength of structural components of presses within the design process.

2 MULTI-BODY SIMULATION OF THE HIGH-SPEED PRESS

The static and dynamic properties of a high-speed press were investigated by means of a hybrid MBS at the IFUM [2, 3]. Therein, the elasticities of the press frame, ram, con rods and eccentric shaft were taken into consideration during the MBS. This was realised by applying the Craig-Bampton method [4, 5] which is generally established [6] and implemented in most commercial MBS software packages [7-9]. By means of the Craig-Bampton method static and dynamic modes of the bodies that should be considered elastic during the MBS were calculated. The static modes depend on the type of joints acting on the bodies due to the connection to other bodies of the MBS. The dynamic modes are obtained by means of a modal analysis of the bodies themselves. Therefore, the degrees of freedom of the FE-models were reduced to a set of static and dynamic modes for each body. The actual deformations of the bodies during the MBS were determined by superimposing a combination of the static and dynamic modes. By this approach a good correlation between the simulated and experimentally determined properties of the press was realised. For instance the oscillation of the ram due to the drop of the cutting force was virtually determined as shown in figure 1. This oscillation mainly depends on the elasticities of the above mentioned components of the press and their mass.

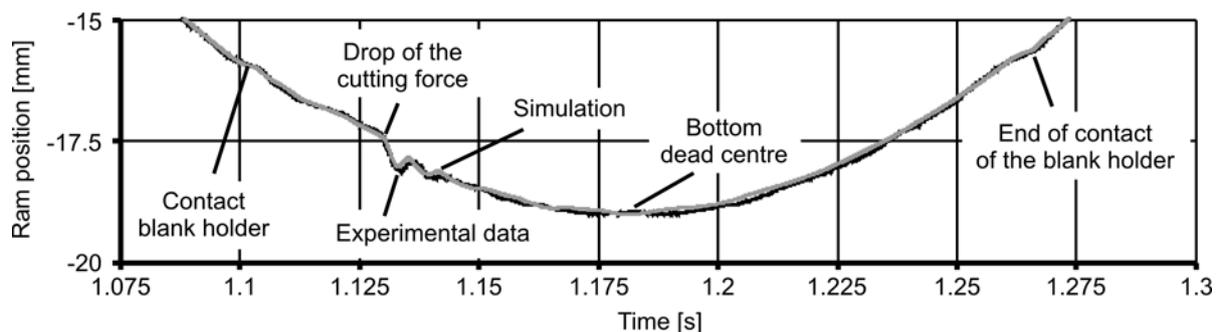


Figure 1: Simulated and experimentally determined oscillation of the ram due to the drop of the cutting force

The hybrid MBS of the high-speed press was validated by means of further numerical and experimental investigations as shown in [2, 3]. Due to the reduction of the degrees of freedom of the FE-models to a small number of static and dynamic modes the above described approach allows the simulation of the dynamics of a mechanical press with accuracy as well

as with considerable low computational requirements.

3 FATIGUE ANALYSIS OF THE COMPONENTS OF THE PRESS

The following approach of a fatigue simulation allows the analysis of each component that is considered to be elastic within a MBS. As described above, for each component considered elastic during the MBS a set of static and dynamic modes was determined by means of the finite element method (FEM) following the Craig-Bampton method. In addition, for each static and dynamic mode of deformation of a given FE-model, a corresponding stress mode can be determined by means of a FE-solver. These stress modes are also known as modal stresses [6, 10-12]. As the deformation of the entire body can be determined by superimposing the individual modes of deformation, the corresponding stress state of the entire body can be determined by superimposing the modal stresses. The modal stresses of the different elastic components of the press were determined by means of the FEM and imported into a durability software package. In addition, the durability software package was coupled with the MBS. The transient deformations of the elastic bodies were transferred from the MBS to the coupled durability software package. Within the durability software package the transient stress states of the flexible components were determined based on deformation of the bodies given by the coupled MBS and the modal stresses of the bodies imported from the FEM. Therefore, the transient stress state of any component that is considered to be elastic within the MBS can be analysed by means of the durability software package.

Within the fatigue analysis the virtual determined transient stresses are compared with the material related stress-number curves (S/N curves). The press frame, ram and con rods of the analysed high-speed press are made of ductile cast iron type EN-GJS-500-7. The eccentric shaft of the analysed press is made of a quenched and tempered steel type 30CrNiMo8. The S/N curves of these two materials are given by the Guideline of the Forschungskuratorium Maschinenbau "Analytical Strength Assessment" (FKM-Guideline) [13]. According to the FKM-Guideline the S/N curves are defined by the fatigue limits for completely reversed normal stresses $\sigma_{W,zd}$ and the tensile strength R_m . Hence, these values were applied to define the S/N curves required for the fatigue analysis of the structural components of the press by means of the durability software package.

The structural components of the press are not loaded to pure cyclic loads. In fact mean stresses lead to an almost zero-tension loading. Therefore, an additional approach was required to take the influence of the mean stresses into consideration. As such an approach the mean stress correction according to Goodman [14, 15] is known. The Goodman model requires a so called mean stress sensitivity M_σ which describes the influence of the mean stress. The mean stress sensitivity M_σ was determined according to the FKM-Guideline.

During the fatigue analysis of the structural components of the press it was distinguished whether a component is exposed to proportional stresses or non-proportional stresses. Proportional stresses occur if the loads acting on a body are applied proportional to each other. For instance the con rods of the press are exposed to proportional stresses. The con rods have only two points where they are connected to other components of the press. These are the connections to the eccentric shaft and to the ram. If a given load F_R is applied to a con rod by the ram this load is transferred to the eccentric shaft, leading to a reaction force of the

eccentric shaft F_E at the second joint of the con rod. The reaction force due to the eccentric shaft has the same amplitude as the force applied by the ram. Thus, the two loads acting on the con rods are always proportional to each other and therefore the stress state of the con rod is also proportional. On the other hand the eccentric shaft is loaded highly non-proportional. In figure 2 the different forces and momentums acting at the eccentric shaft are illustrated. The force F_F is required to support the mass of the fly wheel and acts continuously with a constant amplitude. The momentum M_F caused by the fly wheel depends on the sum of the other forces acting at the eccentric shaft. These are for instance the forces due to the con rods F_{C1} and F_{C2} which drive the ram. F_{C1} and F_{C2} are dynamic loads as they depend on the force required to drive the ram as well as on the process force applied to the ram. Similar to the forces F_{C1} and F_{C2} the force F_M is a dynamic load. F_M is required to drive the mass of the mass balancing unit. In addition, a second momentum M_{Fe} acts on the eccentric shaft which is required to drive the feeder attached to the press. Finally, the forces acting on the eccentric shaft due to the bearings F_{B1} to F_{B4} depend on all the other forces described above. All these forces acting on the eccentric shaft are not proportional to each other. Therefore, the eccentric shaft is non-proportional loaded leading to a non-proportional stress state at the eccentric shaft.

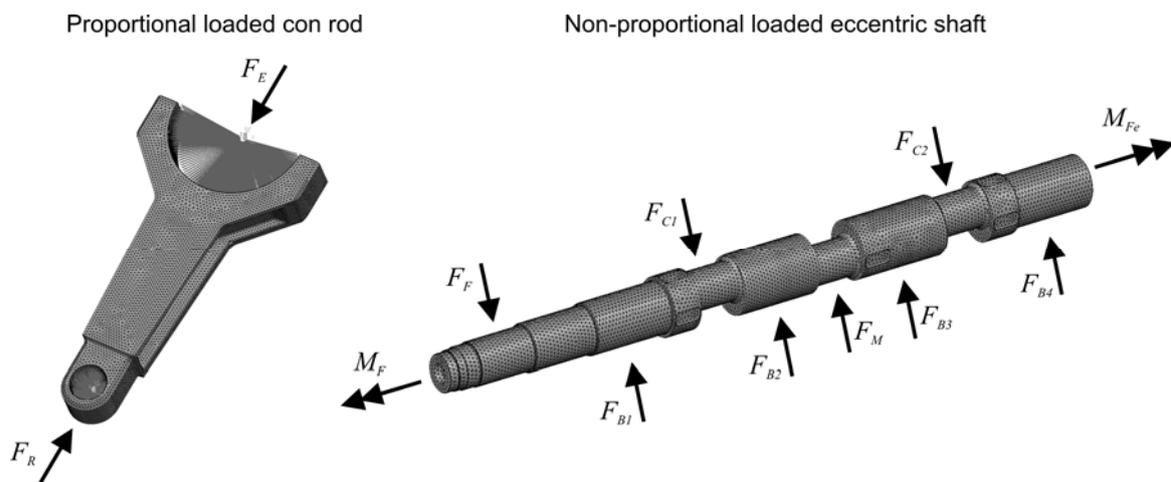


Figure 2: Loads acting on the con rod and eccentric shaft

For a proportional loaded component it is not required to take into consideration on which plane of the component the maximum stress occurs. The maximum stresses occur always on the same plane and thus the fatigue damage of each loading cycle can be added to the fatigue damage caused by the previous loading cycle. However, if the stress states of a component are non-proportional it must be distinguished on which plane the different transient stress amplitudes act. Therefore, in a non-proportional stress approach each stress amplitude is associated with a plane. The fatigue solver identifies the plane with the highest fatigue damage and defines this plane as the critical plane. On the critical plane crack initiation occurs first. Hence, depending on the stress states of the components of the press different fatigue approaches were applied to analyse the fatigue strength of the components of the press

(Figure 3). The con rods were analysed by means of a signed von Mises approach as they are loaded proportional. The other components of the press were analysed by means of a critical plane approach as they are loaded non-proportional.

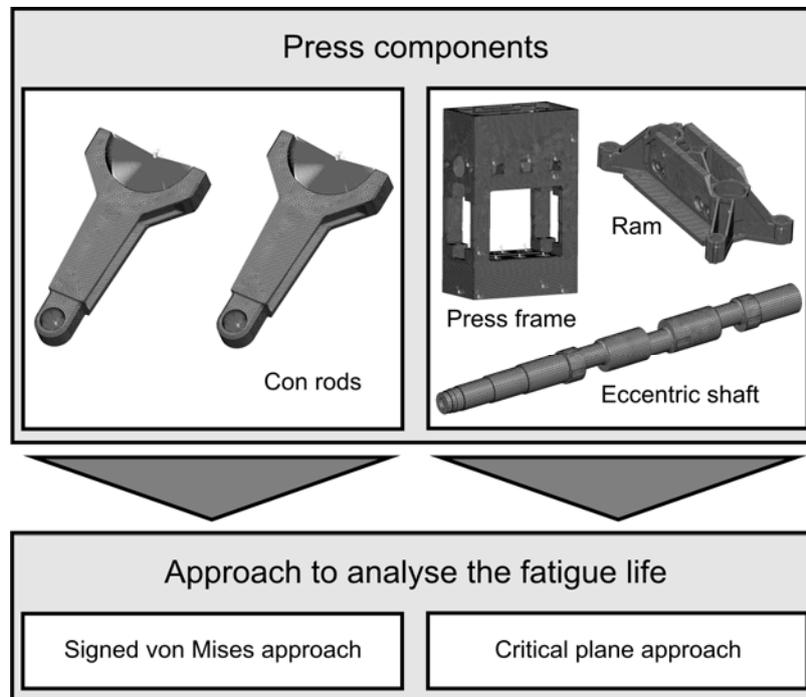


Figure 3: Approaches for the fatigue analysis of the structural components of the press

In figure 4 the results of the fatigue analysis for the con rods of the press are shown. At the left con rod of the press a maximum stress of 46 N/mm^2 was identified. According to the simulation the maximum occurring stress of the right con rod (44 N/mm^2) is slightly lower than the maximum stress occurring on the left con rod. The results are based on the loads acting on the press due to cutting a disc of 35 mm diameter made of high-strength steel type MS-W 1200. The determined fatigue damage is shown on the right hand side of figure 4. For both con rods of the press no fatigue damage was detected. Therefore, the con rods of the press are considered non-critical regarding crack initiation for the selected loading of the press.

The maximum stresses at the press frame and the ram were of the same range as the maximum stresses determined for the con rods. Therefore, crack initiation at the press frame and the ram is unlikely for the selected loading of the press. The maximum stress determined for the eccentric shaft (148 N/mm^2) was considerably higher. However, according to the FKM-Guideline the material of the eccentric shaft (30CrNiMo8) is also capable to carry higher loads. Therefore, crack initiation at the eccentric shaft either due to normal stresses or shear stresses is also unlikely for the selected loading of the press.

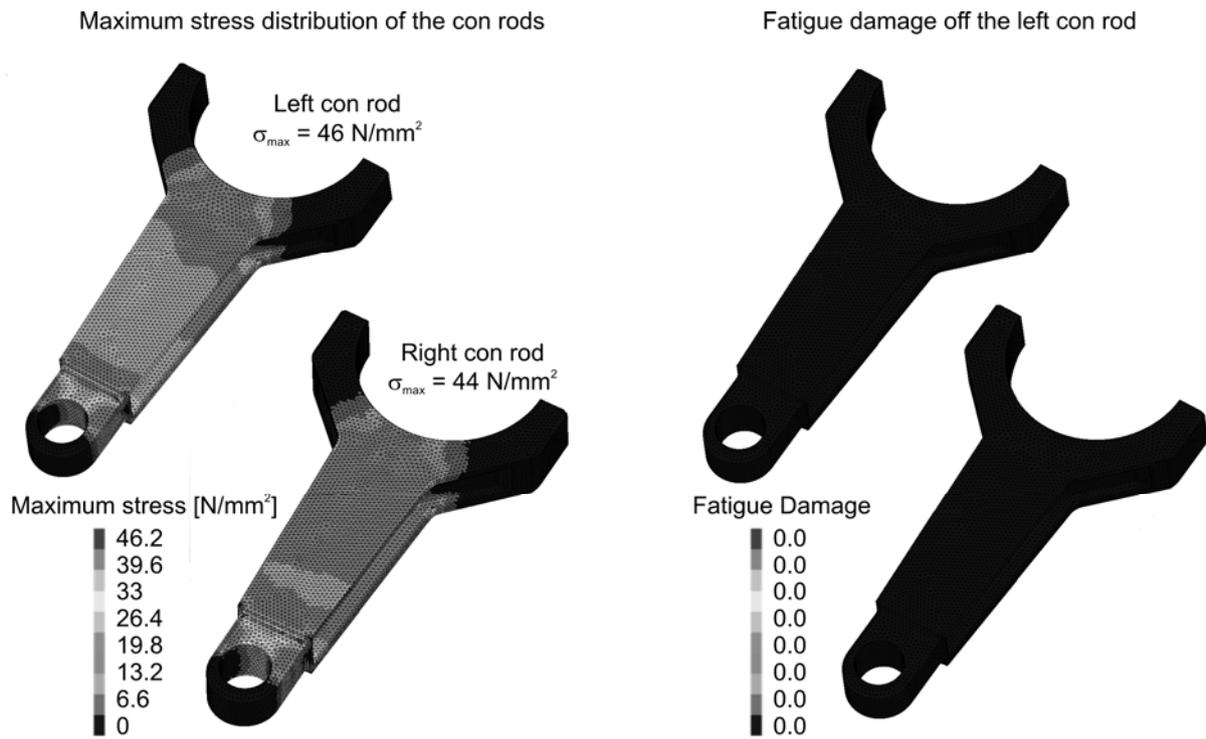


Figure 4: Stress distribution and fatigue damage determined at the con rods of the high-speed press

A validation of the fatigue analyses of the components of the press is not possible as no crack initiation was predicted by means of the coupled simulation. Therefore, it would be only possible to validate durability. An increase of the loadings of the press would eventually lead to a predicted crack initiation after a given number of strokes; however the validation of such a scenario would certainly damage the press. Therefore, it was decided to validate the selected approach of the fatigue analysis of the components of the press by means of specimens loaded with a fatigue testing machine as well as by means of test structures assembled to the press.

4 FATIGUE ANALYSIS OF NOTCHED SPECIMENS LOADED BY MEANS OF A FATIGUE TESTING MACHINE

As it is not practicable to load the press available at the IFUM until crack initiation occurs the validation of the selected approach for the fatigue analysis of the components of the press was realised by means of notched specimens. The notched specimens were loaded by means of a fatigue testing machine until crack initiation occurred. Figure 5 shows a notched specimen as it was clamped during the test setup. The material of the components of the press (ductile cast iron) is not available as a standardised profile bar. Therefore, S235JR was selected to manufacture the notched specimens of a standardised profile bar. This material is applied by many press manufactures to construct welded press structures such as press frames, rams and the like. On the market different profile bars as well as different sheet sizes are available in the quality of S235JR and therefore welded structures can be easily manufactured applying such standardised sizes. Thus, the notched specimens were also

manufactured applying a standardised profile bar of the material S235JR. The loading of the notched specimen was simulated by means of a hybrid MBS which was coupled with a durability software package as it was previously done for the structural components of the press. The notched specimen was considered elastic during the MBS. The clamps of the fatigue testing machine were modelled rigid. The coupled simulation was used to determine the forces which were applied to the notched specimen by means of the real fatigue testing machines. It was determined that a force amplitude of 5 kN leads to a stress amplitude of 220 N/mm² at the root of the notches. The stress distribution of the loaded specimen and the region of the maximum stresses at the root of the notches are shown in figure 5. Five notched specimens were loaded with a force amplitude of 5 kN until crack initiation was detected or 10 million load cycles were reached. On two out of the five notched specimens crack initiation was detected. The numbers of load cycles at which crack initiation was detected are marked in the graph of figure 5 showing the S/N curve of the material S235JR. As crack initiation was not detected on all of the specimens it was decided to increase the force applied by means of the fatigue testing machine. According to the coupled simulation a force amplitude of 6 kN leads to a maximum stress amplitude of 264 N/mm². Therefore, five notched specimens were loaded with a force amplitude of 6 kN. On all five specimen loaded to a force amplitude of 6 kN crack initiation was detected. The numbers of cycles at which crack initiation was detected on a maximum stress amplitude of 264 N/mm² (force amplitude 6 kN) are also marked in the graph of figure 5. The graph in figure 5 shows that no crack initiation was detected at a number of cycles that is lower than the number of cycles given for a certain stress amplitude by the S/N curve. Thus, no crack initiation occurred at a lower number of cycles than predicted by means of the coupled simulation.

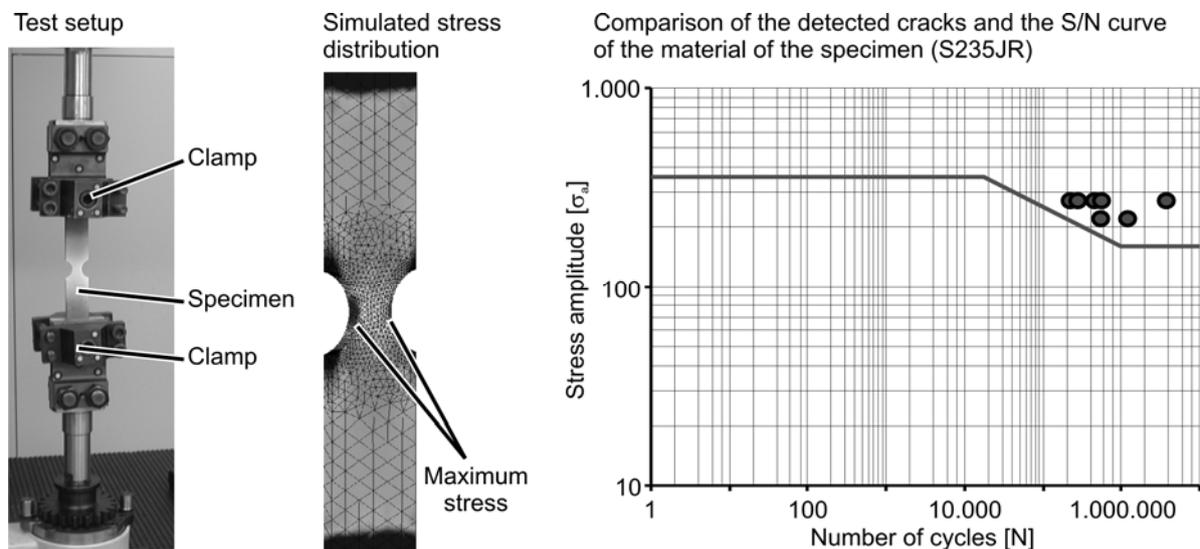


Figure 5: Validation of the fatigue analyses by means of a zero-tensile loaded specimen

Having planned to load test structures assembled to the real press to bending as it is shown in chapter 5 an additional notched specimen was loaded to bending by means of the fatigue

testing machine. Thus, a test rig was designed and manufactured to apply bending loads. In figure 6 the test setup to apply bending loads is shown. The loading of the notched specimen exposed to bending forces was also simulated by means of the coupled simulation. During the coupled simulation the specimen was considered as elastic whereas all other components of the test rig were modelled rigid. Ten specimens were loaded to different loading levels. The loading levels were determined by means of the coupled simulation. The numbers of loading cycles at which crack initiation was detected are marked in the graph of figure 6. The graph shows that each crack initiation was detected at a higher number of cycles than the number of cycles given by the S/N curve for a certain stress amplitude. In addition, it can be seen that at the highest stress level selected crack initiation occurred at 200,000 to 300,000 cycles.

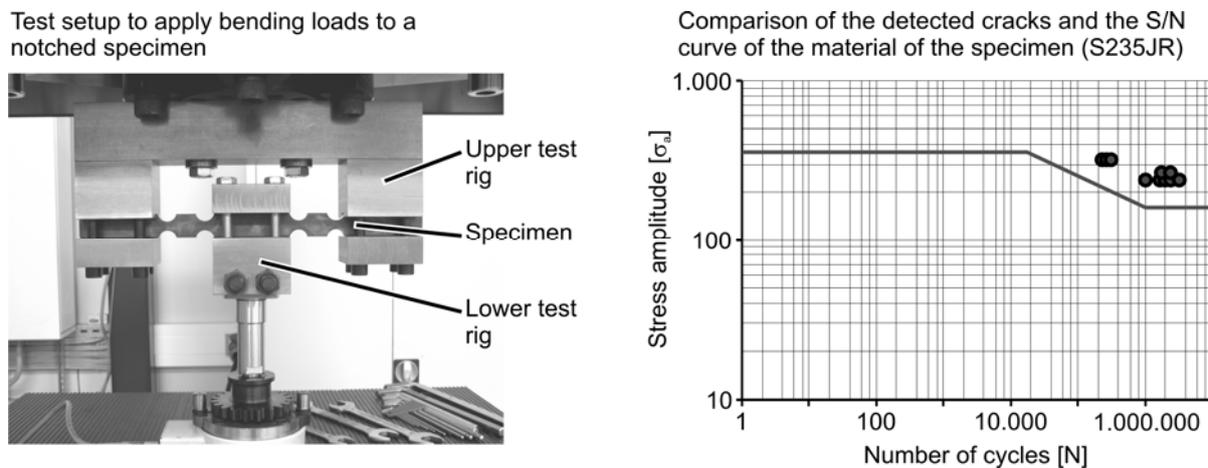


Figure 6: Validation of the fatigue analyses by means of a bending loaded specimen

The load levels (stress amplitudes) determined by means of the experiments described above were used to design critical dimensioned test structures. This critically dimensioned test structures were assembled to the high-speed press at the IFUM to validate the fatigue analysis of the components of the press by means of the coupled simulation as described in the following chapter.

5 VALIDATION OF THE FATIGUE ANALYSES BY MEANS OF TEST STRUCTURES ASSEMBLED TO THE PRESS

Based on the results of chapter 4 critically dimensioned test structures were designed, manufactured and assembled to the press. The test structures were loaded due to cutting high-strength steel. The fatigue life of the test structures was virtually determined by means of the coupled simulation. Therefore, the hybrid MBS model of the high-speed press was extended. Figure 7 shows the extended hybrid MBS model. It can be seen that test structures are virtually assembled between the ram of the press and the top of the cutting tool. Therefore, the test structures are assembled in the flux of the process forces. According to figure 7 the test structures are loaded to bending similar to the specimens which were loaded to bending by

means of the fatigue testing machine (figure 6). The coupled simulation was applied to determine the size of the notches. Therefore, it could be assured that at the root of the notches similar stresses occur while cutting high-strength steel as at the specimens loaded to bending which had a fatigue life of 200,000 to 300,000 cycles. Figure 7 shows also the real test structures assembled to the high-speed press.

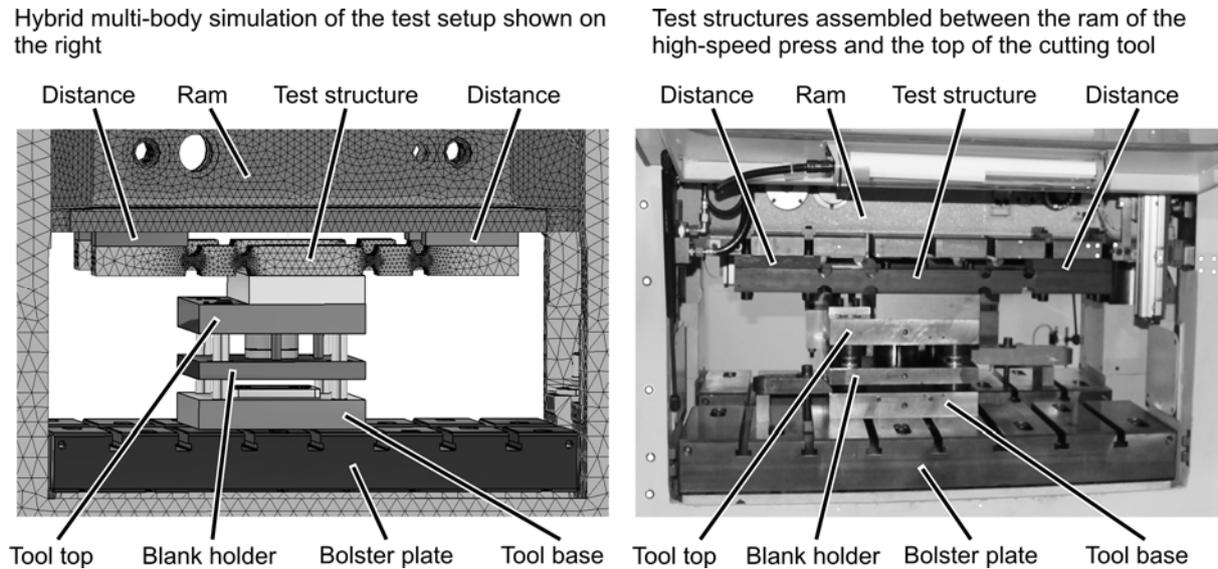


Figure 7: Critically dimensioned test structure in the flux of the process force

After approximately 200,000 cycles of cutting high-strength steel with a stroke rate of 200 strokes per minute crack initiation was detected by means of the dye penetrant inspection [17]. The detected cracks are shown in figure 8. Crack initiation occurred on both notched specimen at the predicted position at the root of the notches.

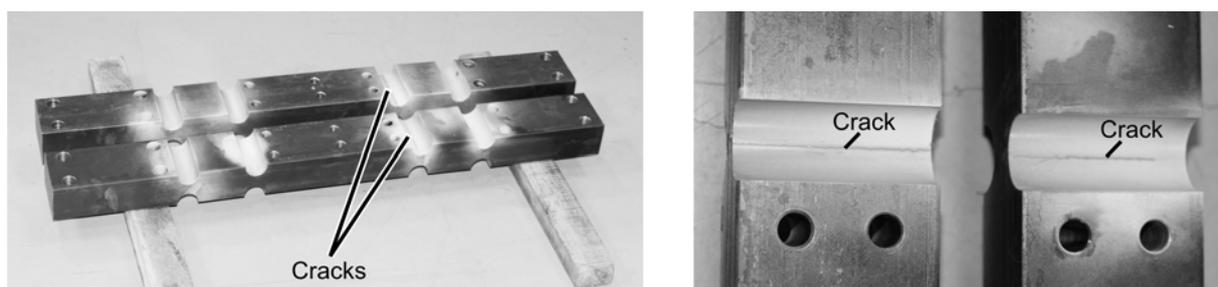


Figure 8: Cracks at the test structures in the flux of the process force

Taking into consideration that the simulation was based on design data and no calibration of the model was carried out the simulation represents the experiment well. Furthermore, it has to be considered that according to figure 5 and figure 6, the fatigue lifes of the specimens made of S235JR spread largely. Hence, no precise prediction of the fatigue life could have

been expected. Within the coming months further tests will be carried out at the IFUM to confirm the result of the fatigue analysis of the critically dimensioned test structure assembled to the press.

In addition to the test structure assembled to the press in the flux of the process force a second test structure was assembled to the bolster plate by means of a test rig (figure 9). This test structure is however not in the flux of the process force. On both sides of the test structure a mass is mounted. The vibration of the bolster plate due to the drop of the cutting force is transferred to the test structure via the test rig. The vibration of the test structure initiates the mounted mass to oscillation. Due to resonance, the oscillation of the mass increases at critical stroke rates. The critical stroke rate determined by means of the MBS is 220 strokes per minute. During the cutting of MS-W 1200 the critical stroke rate of the real test structures assembled to the test rig was found to be 216 strokes per minute. According to the simulation the amplitude of the oscillating mass (movement) is approximately 0.4 mm whereas the measured amplitude is approximately 0.6 mm (figure 9). The notched test structure is loaded to bending due to the oscillation of the mass. During each stroke of the press six load changes occur at the test structure. The test structure is designed for a fatigue life of approximately 2 million load changes. So far no crack initiation was detected by means of the dye penetrant inspection. However, within the coming months approximately 2.5 million strokes will be carried out at the IFUM to validate the results of the fatigue analysis while cutting high-strength steel.

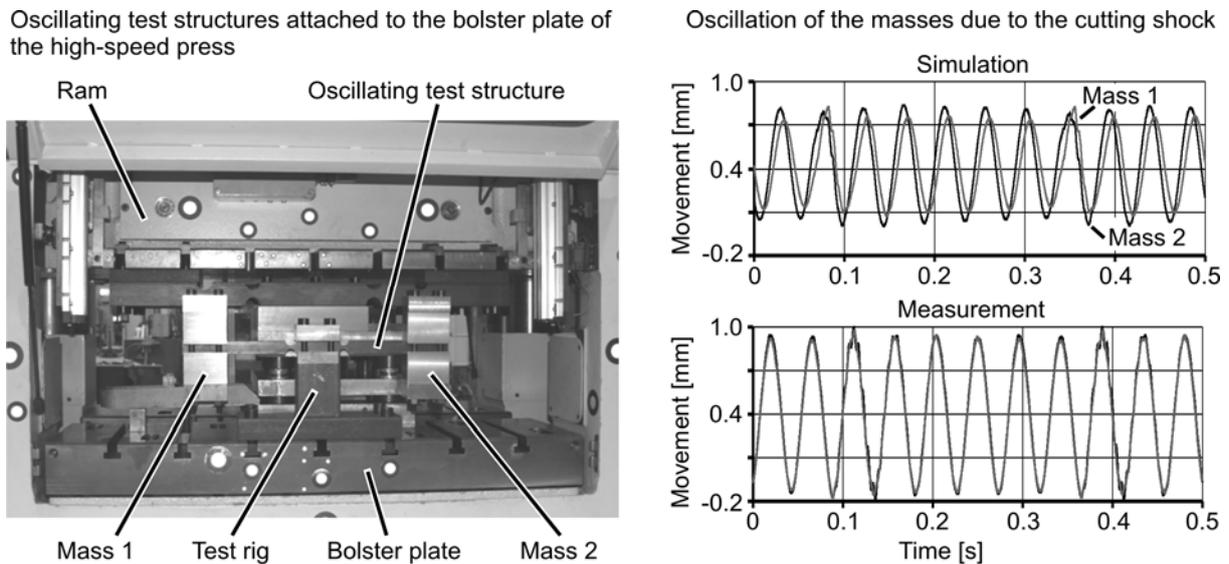


Figure 9: Oscillating test structure

6 CONCLUSION

In this paper an approach was shown in which a hybrid multi-body simulation model of a high-speed press was coupled with a commercial durability software package. Thus, a fatigue analysis of the components of the press considered as elastic during the hybrid multi-body simulation was realised. Within the project presented in this paper the press frame, ram, con

rods and eccentric shaft were considered elastic during the multi-body simulation. The selected approach of the fatigue analysis of the components of the press was a stress based high-cycle fatigue approach. During the fatigue analysis of the different components of the press it was distinguished whether a component was loaded by forces acting proportional to each other or by forces acting non-proportional to each other.

Furthermore, intensive tests in which notched specimens were loaded by means of a fatigue testing machine until crack initiation occurred were shown. The loading of the notched specimens and the corresponding fatigue life was analysed by means of a coupled simulation. Thereby, the same fatigue approach was applied as for the components of the press. Hence, the fatigue approach applied for the components of the press was validated by means of the test carried out with the notched specimen.

Finally, critically dimensioned test structures were designed, manufactured and assembled to the real press at the Institute of Metal Forming and Metal-Forming Machines. These test structures were also analysed by means of the coupled simulation regarding their fatigue life. The critically dimensioned test structures were loaded due to cutting high-strength steel until crack initiation occurred. The coupled simulation was validated as the experimentally determined and simulated fatigue life of the test structures correlated.

With the above described approach for a fatigue analysis of the structural components of a high-speed press it is possible to determine their fatigue life. The described approach has a high potential to improve the design of presses as it is based on design data and no calibration of the model is required. Therefore, the approach can be applied during the design process of presses without the use of further experimental tests.

ACKNOWLEDGEMENTS

The presented results have been obtained within the research project “Analysis of the press demands while cutting high-strength and super high-strength steel materials” (AiF 15680 N/1) at the Institute of Metal Forming and Metal-Forming Machines (IFUM) of the Leibniz Universität Hannover (LUH). The authors owe thanks to the German Machine Tool Builders' Association (VDW) and to the German Federation of Industrial Research Associations (AiF) for the financial support of this project.

REFERENCES

- [1] Schuler AG, *Harte Werkstoffe sanft schneiden*, Blech Rohre Profile, Vol. 5 (2005), pp. 18-23.
- [2] Behrens, B.-A., Wager, C., Krimm, R., Kammler, M. and Schrödter, J. *Investigation of the Elastic Deformation of a Mechanical Press due to Cutting High-Strength Steels*, Proc. International Conference on Innovative Technology 2010 (IN-TECH 2010), Prague, Czech Republic, September 14-16, 2010, pp. 425-428.
- [3] Behrens, B.-A., El-Galy, I., Wager, C., Schrödter, J., Olle, P., Betancur Escobar, S., Weigel, N., Bouguecha, A., *Aktuelle Forschungsergebnisse am Institut für Umformtechnik und Umformmaschinen*, 20. Umformtechnisches Kolloquium Hannover 2011, Hannover, Germany, February 23-24, 2011, pp. 35-51.

- [4] Craig, R., Bampton, M. *Coupling of Substructures for Dynamic Analysis*, AIAA Journal 6, Nr. 7 (1968), pp. 1313-1319.
- [5] Craig, R. *Structural dynamics – An introduction to computer methods*. John Wiley & Sons, New York, USA, 1981.
- [6] Fischer, P., Witteveen, W. *Fatigue Analysis of FE-MBS Hybride Structures Incorporating Modal Stresses / Dynamics and Vibration Effects*, 9. Aachener Kolloquium Fahrzeug- und Motorentechnik, Aachen, Germany, October 4-6, 2000, pp. 411-440.
- [7] MSC.Software Corporation, *Theory of Flexible Bodies in MSC.ADAMS - Theoretical Background*, Santa Ana, California, USA, 2003.
- [8] Meywerk, M. *CAE-Methoden in der Fahrzeugtechnik*, Springer-Verlag, Berlin, Germany, 2007.
- [9] LMS International, *LMS Virtual.Lab Motion – Optimizing real-life performance of mechanical systems*, Leuven, Belgium, 2010.
- [10] Dannbauer, H., Gaier, C. and Steinbatz, M. *A Statistical Measure of the Non-Proportionality of Stresses – Investigations and Applications*, 7th International Conference on Biaxial/Multiaxial Fatigue & Fracture (ICBMFF), Berlin, Germany, June 28-July 1, 2004, pp. 94-98.
- [11] Riener, H., Mayr, A. and Steinbatz, M. *Improvement Potential in Time and Money using Integration of a „Virtual Test Bench“ in the Simulation and Testing Process*, DVM Conference on Suspension Systems, Osnabruck, Germany, 2002.
- [12] Cali, M., Oliveri, S. M. and Sequenzia, G. *Geometric modeling and modal stress formulation of flexible multi-body dynamic analysis of crankshaft*, Proceedings of IMAC 25th, Society of Experimental Mechanics Inc, Bethel, Connecticut, USA, 2007, Paper Nr. 227.
- [13] Haibach, E. *Analytical Strength Assessment of Components in Mechanical Engineering*, 5th revised edition, Forschungskuroatorium Maschinenbau (FKM), Frankfurt, Germany, 2003.
- [14] Dowling, N. E. *Mean Stress Effects in Stress-Life and Strain-Life Fatigue*, Second SAE Brasil International Conference on Fatigue, Sao Paulo, Brasil, 2004, SAE Paper No. 2004-01-2227, Society of Automotive Engineers, June 2004.
- [15] Pook, L. P. *Metal fatigue: what it is, why it matters*. Springer-Verlag, Berlin, Germany, 2007.
- [16] EN 571-1: Non-destructive testing - Penetrant testing - Part 1: General principles. 1997.

FINITE VOLUME AND FINITE ELEMENT SCHEMES FOR THE EULER EQUATIONS IN CYLINDRICAL AND SPHERICAL COORDINATES

DANTE DE SANTIS*, GIANLUCA GERACI* AND ALBERTO
GUARDONE†

*INRIA Bordeaux–Sud-Ouest, équipe-projet Bacchus
Cours de la Libération, 33405 Talence, France
e-mail: dante.de_santis@inria.fr, gianluca.geraci@inria.fr

†Dipartimento di Ingegneria Aerospaziale
Politecnico di Milano
Via La Masa 34, 20156 Milano, Italy
e-mail: alberto.guardone@polimi.it

Key words: Compressible flows, Cylindrical/Spherical Coordinates, Explosion problem, Finite Element/Volume Methods.

Abstract. A numerical scheme is presented for the solution of the compressible Euler equations over unstructured grids in cylindrical and spherical coordinates. The proposed scheme is based on a mixed finite volume / finite element approach. Numerical simulations are presented for the explosion problem in two spatial dimensions in cylindrical and spherical coordinates, and the numerical results are compared with the one-dimensional simulation for cylindrically and spherically symmetric explosions.

1 INTRODUCTION

Some gasdynamics problems exhibit such strong symmetries that could be more convenient to describe these problems in a curvilinear reference, i.e. cylindrical or spherical, rather than in a Cartesian one. These are, e.g., astrophysical flows, Inertial Confinement Fusion (ICF) applications, sonoluminescence phenomena and nuclear explosions [1].

To compute the numerical solution of the compressible flow equations for these kind of flows, an interesting possibility is provided by the use of a mixed finite volume (FV) / finite element (FE) approach [2]. For example, in viscous flows, it is possible to use the FV and the FE to compute the advection and dissipation terms, respectively, within the same algorithm.

The combined use of these two different techniques is made possible by the introduction of suitable equivalence conditions that relate the FV metrics, i.e. cell volumes and inte-

grated normals, to the FE integrals. Equivalence conditions relating FV and FE schemes have been derived for Cartesian coordinates in two and three spatial dimensions [3],[4] and for cylindrical coordinates in axially symmetric two-dimensional problems [6]. In both cited references, equivalence conditions are obtained by neglecting higher order FE contributions. Subsequently in [7], equivalence conditions for the cylindrical coordinates have been derived for the first time without introducing any approximation into the FE discrete expression of the divergence operator, and in [8] the differences between the consistent scheme and an alternative one violating the the equivalence conditions have been quantified for the case of one-dimensional problems in cylindrical and spherical coordinates. In the present paper the consistent formulation is presented in the cylindrical and spherical reference systems. Numerical results for the explosion problem are also provided to demonstrate the correctness of the proposed approach.

The present paper is structured as follows. In section 2, the scalar equations in cylindrical and spherical coordinates are introduced. In sections 3 and 4 the spatial discretizations of the scalar equation by FE and FV are presented in cylindrical and spherical coordinates respectively. Equivalence conditions between FE and FV metrics are also shown. In section 3, numerical simulations are presented for the explosion problem in the cylindrical coordinates on the R - θ plane and in the the spherical coordinates on the r - ϕ plane. Numerical results are also compared to one-dimensional simulations.

2 SCALAR EQUATIONS IN CYLINDRICAL AND SPHERICAL COORDINATES

For simplicity we consider first a multidimensional scalar conservation law. The model equation in a three-dimensional cylindrical reference reads

$$\frac{\partial u}{\partial t} + \frac{\partial f_Z}{\partial Z} + \frac{1}{R} \frac{\partial}{\partial R}(R f_R) + \frac{1}{R} \frac{\partial f_\theta}{\partial \theta} = 0, \quad (1)$$

where t is the time, Z , R and θ are the axial, radial and azimuthal coordinates, respectively, $u = u(Z, R, \theta, t)$ is the scalar unknown and $\mathbf{f}^\circ(u) = (f_Z, f_R, f_\theta)$ is the so-called flux function. A more compact form of the above equation is obtained by introducing the divergence operator in three-dimensional cylindrical coordinates $\nabla^\circ \cdot (\cdot)$ as follows

$$\frac{\partial(u)}{\partial t} + \nabla^\circ \cdot \mathbf{f}^\circ(u) = 0. \quad (2)$$

The model equation in a three-dimensional spherical reference reads

$$\frac{\partial u}{\partial t} + \frac{1}{r^2} \frac{\partial}{\partial r}(r^2 f_r) + \frac{1}{r \sin \theta} \frac{\partial}{\partial \theta}(\sin \theta f_\theta) + \frac{1}{r \sin \theta} \frac{\partial f_\phi}{\partial \phi} = 0,$$

where r , θ and ϕ are the radial, polar and azimuthal coordinates, respectively, $u = u(r, \theta, \phi, t)$ is the scalar unknown and $\mathbf{f}^\circ(u) = (f_r, f_\theta, f_\phi)$ is the flux function. In a

compact form, using the divergence operator in three-dimensional spherical coordinates $\nabla^\circ \cdot (\cdot)$, it can be rearranged as

$$\frac{\partial(u)}{\partial t} + \nabla^\circ \cdot \mathbf{f}^\circ(u) = 0. \quad (3)$$

3 FINITE VOLUME/ELEMENT METHOD IN CYLINDRICAL COORDINATES

3.1 Node-pair finite element discretization

The scalar conservation law (2) is now written in a weak form by multiplying it by the radial coordinate R and by a suitable Lagrangian test function $\phi_i \in V_h \subset H^1(\Omega)$. Integrating over the support Ω_i of ϕ_i gives

$$\int_{\Omega_i} R\phi_i \frac{\partial u}{\partial t} d\Omega_i + \int_{\Omega_i} R\phi_i \nabla^\circ \cdot \mathbf{f}^\circ(u) d\Omega_i = 0, \quad \forall i \in \mathcal{N}, \quad (4)$$

where \mathcal{N} is the set of all nodes of the triangulation. Note that by multiplying by R , the numerical singularity of the cylindrical reference system is formally removed [6]. In the following, to simplify the notation, the infinitesimal volume $d\Omega = RdRd\theta dZ$ is not indicated in the integrals. Integrating by parts immediately gives

$$\int_{\Omega_i} R\phi_i \frac{\partial u}{\partial t} = \int_{\Omega_i} R\mathbf{f}^\circ \cdot \nabla^\circ \phi_i + \int_{\Omega_i} \phi_i \mathbf{f}^\circ \cdot \nabla^\circ R - \int_{\partial\Omega_i^\circ} R\phi_i \mathbf{f}^\circ \cdot \mathbf{n}_i^\circ \quad (5)$$

where $\partial\Omega_i^\circ = \partial\Omega_i \cap \partial\Omega$, with $\partial\Omega_i$ and $\partial\Omega$ are the boundary of Ω_i and of the computational domain Ω , respectively, and where $\mathbf{n}_i^\circ = n_Z \hat{\mathbf{Z}} + n_R \hat{\mathbf{R}} + n_\theta \hat{\boldsymbol{\theta}}$ is the outward normal versor to Ω_i . The scalar unknown is now interpolated as

$$u(Z, R, \theta, t) \simeq u_h(Z, R, \theta, t) = \sum_{k \in \mathcal{N}} u_k(t) \varphi_k(Z, R, \theta),$$

and the flux function $\mathbf{f}^\circ(u_h)$ is now expanded using the same shape functions $\phi_h \in V_h$ [9] as follows

$$\mathbf{f}^\circ(u_h) = \mathbf{f}^\circ \left(\sum_{k \in \mathcal{N}} u_k(t) \varphi_k(Z, R, \theta) \right) \simeq \sum_{k \in \mathcal{N}} \mathbf{f}_k^\circ(t) \varphi_k(Z, R, \theta),$$

to obtain

$$\begin{aligned} \sum_{k \in \mathcal{N}_i} M_{ik}^\circ \frac{du_k}{dt} &= \sum_{k \in \mathcal{N}_i} \mathbf{f}_k^\circ(t) \cdot \int_{\Omega_{ik}} R\varphi_k \nabla^\circ \phi_i + \sum_{k \in \mathcal{N}_i} \mathbf{f}_k^\circ(t) \cdot \int_{\Omega_{ik}} \phi_i \varphi_k \nabla^\circ R \\ &\quad - \sum_{k \in \mathcal{N}_i^\circ} \mathbf{f}_k^\circ(t) \cdot \int_{\partial\Omega_{ik}^\circ} R\varphi_i \varphi_k \mathbf{n}_i^\circ, \end{aligned}$$

where \mathcal{N}_i is the set of shape functions φ_k whose support Ω_k overlap Ω_i of φ_i , \mathcal{N}_i^∂ is the set of all boundary nodes of Ω_i , $\Omega_{ik} = \Omega_i \cap \Omega_k$ and $\partial\Omega_{ik}^\partial = \partial\Omega_i \cap \partial\Omega_k \cap \partial\Omega$. The previous equation simplifies to [6]

$$\begin{aligned} \sum_{k \in \mathcal{N}_i} M_{ik}^\varnothing \frac{du_k}{dt} = & - \sum_{k \in \mathcal{N}_{i,\neq}} \left(\frac{\mathbf{f}_k^\varnothing + \mathbf{f}_i^\varnothing}{2} \cdot \boldsymbol{\eta}_{ik}^\varnothing - \frac{\mathbf{f}_k^\varnothing - \mathbf{f}_i^\varnothing}{2} \cdot \boldsymbol{\zeta}_{ik}^\varnothing \right) \\ & + \mathbf{f}_i^\varnothing \cdot \widehat{\mathbf{L}}_i^\varnothing - \sum_{k \in \mathcal{N}_{i,\neq}^\partial} \frac{\mathbf{f}_k^\varnothing - \mathbf{f}_i^\varnothing}{2} \cdot \boldsymbol{\chi}_{ik}^\varnothing - \mathbf{f}_i^\varnothing \cdot \boldsymbol{\xi}_i^\varnothing, \end{aligned} \quad (6)$$

where $\mathcal{N}_{i,\neq} = \mathcal{N}_i \setminus \{i\}$ and $\mathcal{N}_{i,\neq}^\partial = \mathcal{N}_i^\partial \setminus \{i\}$, and the following FE metric quantities have been introduced

$$\begin{aligned} M_{ik}^\varnothing &\stackrel{\text{def}}{=} \int_{\Omega_{ik}} R \varphi_i \varphi_k, & M_{ik} &\stackrel{\text{def}}{=} \int_{\Omega_{ik}} \varphi_i \varphi_k, & \boldsymbol{\zeta}_{ik}^\varnothing &\stackrel{\text{def}}{=} \int_{\Omega_{ik}} \varphi_i \varphi_k \widehat{\mathbf{R}}, & \widehat{\mathbf{L}}_i^\varnothing &\stackrel{\text{def}}{=} \sum_{k \in \mathcal{N}_i} \boldsymbol{\zeta}_{ik}^\varnothing = \int_{\Omega_i} \varphi_i \widehat{\mathbf{R}}, \\ \boldsymbol{\eta}_{ik}^\varnothing &\stackrel{\text{def}}{=} \int_{\Omega_{ik}} R (\varphi_i \nabla^\varnothing \varphi_k - \varphi_k \nabla^\varnothing \varphi_i), & \boldsymbol{\chi}_{ik}^\varnothing &\stackrel{\text{def}}{=} \int_{\partial\Omega_{ik}^\partial} R \varphi_i \varphi_k, & \boldsymbol{\xi}_i^\varnothing &\stackrel{\text{def}}{=} \int_{\partial\Omega_i^\partial} R \varphi_i. \end{aligned}$$

In the next section, the corresponding FV metrics are derived.

3.2 Edge-based finite volume discretization

The spatially discrete form of the scalar conservation law (2) is now obtained according to the node-centred finite volume approach [10]. To this purpose, the integral form of (2) times the radial coordinate R is enforced over a finite number of non-overlapping finite volumes \mathcal{C}_i , with boundary $\partial\mathcal{C}_i$. Over each control volume \mathcal{C}_i the cell-averaged unknown is introduced as follows

$$u(Z, R, \theta, t) \simeq u_i(t) = \frac{1}{V_i} \int_{\mathcal{C}_i} u(Z, R, \theta, t),$$

where V_i is the volume of the i -th cell. Splitting the boundary integral into interface and edge contributions

$$V_i^\varnothing \frac{du_i}{dt} = - \sum_{k \in \mathcal{N}_{i,\neq}} \int_{\partial\mathcal{C}_{ik}} R \mathbf{f}^\varnothing \cdot \mathbf{n}_i^\varnothing - \int_{\partial\mathcal{C}_i^\partial} R \mathbf{f}^\varnothing \cdot \mathbf{n}_i^\varnothing + \int_{\mathcal{C}_i} \mathbf{f}^\varnothing \cdot \widehat{\mathbf{R}}, \quad (7)$$

where $\mathcal{N}_{i,\neq}$ is the set of the finite volume \mathcal{C}_k sharing a boundary with \mathcal{C}_i , excluding \mathcal{C}_i and where $\partial\mathcal{C}_{ik} = \partial\mathcal{C}_i \cap \partial\mathcal{C}_k \neq \emptyset, k \neq i$, is the so-called cell interface. As it is standard practice, the flux vector is assumed to be constant over each cell interface. Under this assumption, the domain and boundary contributions read

$$\begin{aligned} \int_{\partial\mathcal{C}_{ik}} R \mathbf{f}^\varnothing \cdot \mathbf{n}_i^\varnothing &\simeq \mathbf{f}_{ik}^\varnothing \cdot \int_{\partial\mathcal{C}_{ik}} R \mathbf{n}_i^\varnothing = \mathbf{f}_{ik}^\varnothing \cdot \boldsymbol{\nu}_{ik}^\varnothing & \text{with} & \quad \boldsymbol{\nu}_{ik}^\varnothing \stackrel{\text{def}}{=} \int_{\partial\mathcal{C}_{ik}} R \mathbf{n}_i^\varnothing & \text{and} \\ \int_{\partial\mathcal{C}_i^\partial} R \mathbf{f}^\varnothing \cdot \mathbf{n}_i^\varnothing &\simeq \mathbf{f}_i^\varnothing \cdot \int_{\partial\mathcal{C}_i^\partial} R \mathbf{n}_i^\varnothing = \mathbf{f}_i^\varnothing \cdot \boldsymbol{\nu}_i^\varnothing & \text{with} & \quad \boldsymbol{\nu}_i^\varnothing \stackrel{\text{def}}{=} \int_{\partial\mathcal{C}_i^\partial} R \mathbf{n}_i^\varnothing, \end{aligned}$$

respectively. If a second-order centred approximation of the fluxes is considered, namely, $\mathbf{f}_{ik} = (\mathbf{f}_i^\varnothing + \mathbf{f}_k^\varnothing)/2$, the final form of the finite volume approximation of (2) reads,

$$V_i^\varnothing \frac{du_i}{dt} = - \sum_{k \in \mathcal{N}_{i,\neq}} \frac{\mathbf{f}_i^\varnothing + \mathbf{f}_k^\varnothing}{2} \cdot \boldsymbol{\nu}_{ik}^\varnothing + \mathbf{f}_i^\varnothing \cdot \widehat{\mathbf{V}}_i^\varnothing - \mathbf{f}_i^\varnothing \cdot \boldsymbol{\nu}_i^\varnothing \quad \text{with } \widehat{\mathbf{V}}_i^\varnothing \stackrel{\text{def}}{=} \int_{\mathcal{C}_i} \widehat{\mathbf{R}}, \quad (8)$$

to be compared with the corresponding FE discretization (6).

3.3 Finite Element/Volume equivalence

The equivalence conditions relating the above FV metric quantities and the FE ones defined in the previous section are now derived. To this purpose, relevant properties of the FE and FV discretizations are briefly recalled.

Considering FE metric quantities first, from its definition the vector $\boldsymbol{\eta}_{ik}^\varnothing$ is asymmetric, namely, $\boldsymbol{\eta}_{ik}^\varnothing = -\boldsymbol{\eta}_{ki}^\varnothing$ which will be referred in the following as property FEM-a. Property FEM-b is obtained by noting that $\sum_{k \in \mathcal{N}_i} (\boldsymbol{\eta}_{ik}^\varnothing - \boldsymbol{\zeta}_{ik}^\varnothing) + \boldsymbol{\xi}_i^\varnothing = \mathbf{0}$ which gives immediately

$$\widehat{\mathbf{L}}_i^\varnothing = \sum_{k \in \mathcal{N}_{i,\neq}} \boldsymbol{\eta}_{ik}^\varnothing + \boldsymbol{\xi}_i^\varnothing. \quad (9)$$

Property FEM-c stems from the following identity

$$3L_i^\varnothing = \int_{\Omega_i} R \varphi_i \nabla^\varnothing \cdot \mathbf{p}^\varnothing = \int_{\partial\Omega_i^\varnothing} R \varphi_i \mathbf{p}^\varnothing \cdot \mathbf{n}_i^\varnothing - \int_{\Omega_i} \varphi_i \mathbf{p}^\varnothing \cdot \widehat{\mathbf{R}} - \int_{\Omega_i} R \mathbf{p}^\varnothing \cdot \nabla^\varnothing \varphi_i. \quad (10)$$

By substituting the exact expansion $\mathbf{p}^\varnothing = \sum_{k \in \mathcal{N}_i} \mathbf{p}_k^\varnothing \varphi_k$, and applying node-pair representation described in section 4.1, Eq. (10) can be written as

$$\begin{aligned} 3L_i^\varnothing &= \sum_{k \in \mathcal{N}_{i,\neq}} \left[\frac{\mathbf{p}_k^\varnothing + \mathbf{p}_i^\varnothing}{2} \cdot \boldsymbol{\eta}_{ik}^\varnothing - \frac{\mathbf{p}_k^\varnothing - \mathbf{p}_i^\varnothing}{2} \cdot \boldsymbol{\zeta}_{ik}^\varnothing \right] \\ &\quad - \widehat{\mathbf{L}}_i^\varnothing \cdot \mathbf{p}_i^\varnothing + \sum_{k \in \mathcal{N}_{i,\neq}^\varnothing} \frac{\mathbf{p}_k^\varnothing - \mathbf{p}_i^\varnothing}{2} \cdot \boldsymbol{\chi}_{ik}^\varnothing + \mathbf{p}_i^\varnothing \cdot \boldsymbol{\xi}_i^\varnothing. \end{aligned}$$

By substituting property FEM-b in the above identity, one finally obtain property FEM-c as

$$3L_i^\varnothing = \sum_{k \in \mathcal{N}_{i,\neq}} \left[\frac{\mathbf{p}_k^\varnothing + \mathbf{p}_i^\varnothing}{2} \cdot \boldsymbol{\eta}_{ik}^\varnothing - \frac{\mathbf{p}_k^\varnothing - \mathbf{p}_i^\varnothing}{2} \cdot \boldsymbol{\zeta}_{ik}^\varnothing \right] + \sum_{k \in \mathcal{N}_{i,\neq}^\varnothing} \frac{\mathbf{p}_k^\varnothing - \mathbf{p}_i^\varnothing}{2} \cdot \boldsymbol{\chi}_{ik}^\varnothing. \quad (11)$$

Considering now FV metric quantities, from the fact that $\mathbf{n}_i^\varnothing = -\mathbf{n}_k^\varnothing$ over $\partial\mathcal{C}_{ik}$, property FVM-a reads $\boldsymbol{\nu}_{ik}^\varnothing = -\boldsymbol{\nu}_{ki}^\varnothing$, which corresponds to the conservation property of the scheme. From the Gauss theorem, one also has

$$\int_{\mathcal{C}_i} \nabla^\varnothing R = \oint_{\partial\mathcal{C}_i} R \mathbf{n}_i^\varnothing,$$

which, from the definition of FV metric quantities, gives property FVM-b as

$$\widehat{\mathbf{V}}_i^\varnothing = \sum_{k \in \mathcal{N}_{i,\neq}} \boldsymbol{\nu}_{ik}^\varnothing + \boldsymbol{\nu}_i^\varnothing. \quad (12)$$

Property FVM-c is obtained by noting that

$$3V_i^\varnothing = \int_{C_i} R \nabla^\varnothing \cdot \mathbf{p}^\varnothing = \oint_{\partial C_i} R \mathbf{p}^\varnothing \cdot \mathbf{n}_i^\varnothing - \int_{C_i} \widehat{\mathbf{R}} \cdot \mathbf{p}^\varnothing. \quad (13)$$

The right hand side of (13) is now computed by means of the FV discretization described in section 3.2 as

$$3V_i^\varnothing = \sum_{k \in \mathcal{N}_{i,\neq}} \frac{\mathbf{p}_k^\varnothing + \mathbf{p}_i^\varnothing}{2} \cdot \boldsymbol{\nu}_{ik}^\varnothing - \mathbf{p}_i^\varnothing \cdot \widehat{\mathbf{L}}_i^\varnothing + \mathbf{p}_i^\varnothing \cdot \boldsymbol{\nu}_i^\varnothing.$$

which from property FVM-b becomes

$$3V_i^\varnothing = \sum_{k \in \mathcal{N}_{i,\neq}} \frac{\mathbf{p}_k^\varnothing + \mathbf{p}_i^\varnothing}{2} \cdot \boldsymbol{\nu}_{ik}^\varnothing. \quad (14)$$

Therefore, a FV approximation can be formally obtained from FE metric quantities defined over the same grid points by setting (see properties FEM/FVM-a and -b)

$$\boldsymbol{\nu}_{ik}^\varnothing = \boldsymbol{\eta}_{ik}^\varnothing, \quad \boldsymbol{\nu}_i^\varnothing = \boldsymbol{\xi}_i^\varnothing, \quad \widehat{\mathbf{V}}_i^\varnothing = \widehat{\mathbf{L}}_i^\varnothing.$$

Note that the mass lumping approximation,

$$\sum_{k \in \mathcal{N}_i} M_{ik}^\varnothing \frac{du_k}{dt} \simeq L_i^\varnothing \frac{du_i}{dt}$$

has to be introduced in the Eq. (6) for the equivalence conditions to be applicable. By subtracting Eq. (11) to Eq. (14), one finally has

$$V_i^\varnothing = L_i^\varnothing + \sum_{k \in \mathcal{N}_{i,\neq}} \frac{\mathbf{p}_k^\varnothing - \mathbf{p}_i^\varnothing}{6} \cdot \boldsymbol{\zeta}_{ik}^\varnothing - \sum_{k \in \mathcal{N}_{i,\neq}^\partial} \frac{\mathbf{p}_k^\varnothing - \mathbf{p}_i^\varnothing}{6} \cdot \boldsymbol{\chi}_{ik}^\varnothing. \quad (15)$$

It is remarkable that, differently from the Cartesian case [3, 4], in the cylindrical reference the FV cell is not coincident with the FE lumped mass matrix. Moreover, the shape of the FV cells that guarantees equivalence with FE discretization still remains to be determined.

4 FINITE VOLUME/ELEMENT METHOD IN SPHERICAL COORDINATES

4.1 Node-pair finite element discretization

In the present section the spatial discretization of conservation law in a spherical reference system is presented.

The scalar conservation law (3) is multiplied by the quantity $r \sin \theta$ to formally remove the singularity of the spherical coordinate system along the zenith direction. The FE node-pair discretization is obtained as done in the cylindrical case, namely

$$\begin{aligned}
 M_{ik}^{\circ} \frac{du_i}{dt} = & - \sum_{k \in \mathcal{N}_{i,\neq}} \left(\frac{\mathbf{f}_k^{\circ} + \mathbf{f}_i^{\circ}}{2} \cdot \boldsymbol{\eta}_{ik}^{\circ} - \frac{\mathbf{f}_k^{\circ} - \mathbf{f}_i^{\circ}}{2} \cdot \boldsymbol{\zeta}_{ik}^{\circ} \right) + \mathbf{f}_i^{\circ} \cdot \widehat{\mathbf{L}}_i^{\circ} \\
 & - \sum_{k \in \mathcal{N}_{i,\neq}^{\partial}} \frac{\mathbf{f}_k^{\circ} - \mathbf{f}_i^{\circ}}{2} \cdot \boldsymbol{\chi}_{ik}^{\circ} - \mathbf{f}_i^{\circ} \cdot \boldsymbol{\xi}_i^{\circ},
 \end{aligned} \tag{16}$$

where $\mathcal{N}_{i,\neq} = \mathcal{N} \setminus \{i\}$ and $\mathcal{N}_{i,\neq}^{\partial} = \mathcal{N}^{\partial} \setminus \{i\}$ and the following metric quantities have been introduced

$$\begin{aligned}
 M_{ik}^{\circ} &\stackrel{\text{def}}{=} \int_{\Omega_{ik}} r \sin \theta \varphi_i \varphi_k d\Omega^{\circ}, & \boldsymbol{\eta}_{ik}^{\circ} &\stackrel{\text{def}}{=} \int_{\Omega_{ik}} r \sin \theta (\varphi_i \nabla^{\circ} \varphi_k - \varphi_k \nabla^{\circ} \varphi_i), \\
 \boldsymbol{\zeta}_{ik}^{\circ} &\stackrel{\text{def}}{=} \int_{\Omega_{ik}} \varphi_i \varphi_k \widehat{\mathbf{r}}, & \widehat{\mathbf{L}}_i^{\circ} &\stackrel{\text{def}}{=} \sum_{k \in \mathcal{N}_i} \boldsymbol{\zeta}_{ik}^{\circ}, \\
 \boldsymbol{\chi}_{ik}^{\circ} &\stackrel{\text{def}}{=} \int_{\partial\Omega_{ik}^{\partial}} r \sin \theta \varphi_i \varphi_k, & \boldsymbol{\xi}_i^{\circ} &\stackrel{\text{def}}{=} \int_{\partial\Omega_i^{\partial}} r \sin \theta \varphi_i.
 \end{aligned}$$

4.2 Edge-based finite volume discretization

The FV spatially discrete form of the scalar conservation law (3) multiplied by the quantity $r \sin \theta$ reads

$$V_i^{\circ} \frac{du_i}{dt} = - \sum_{k \in \mathcal{N}_{i,\neq}} \frac{\mathbf{f}_i^{\circ} + \mathbf{f}_k^{\circ}}{2} \cdot \boldsymbol{\nu}_{ik}^{\circ} + \mathbf{f}_i^{\circ} \cdot \widehat{\mathbf{V}}_i^{\circ} - \mathbf{f}_i^{\circ} \cdot \boldsymbol{\nu}_i^{\circ} \tag{17}$$

where

$$\widehat{\mathbf{V}}_i^{\circ} \stackrel{\text{def}}{=} \int_{\partial\mathcal{C}_i^{\partial}} \widehat{\mathbf{r}}, \quad \boldsymbol{\nu}_{ik}^{\circ} \stackrel{\text{def}}{=} \int_{\partial\mathcal{C}_{ik}} r \sin \theta \mathbf{n}_i^{\circ} \quad \text{and} \quad \boldsymbol{\nu}_i^{\circ} \stackrel{\text{def}}{=} \int_{\partial\mathcal{C}_i^{\partial}} r \sin \theta \mathbf{n}_i^{\circ}$$

4.3 Finite Element/Volume equivalence

The equivalence conditions relating the FV and FE scheme are obtained like in the cylindrical case. The following properties are introduced for the FE metric quantities

$$\begin{aligned} \boldsymbol{\eta}_{ik}^\circ &= -\boldsymbol{\eta}_{ki}^\circ \quad (\text{FEM - a}), & \widehat{\mathbf{L}}_i^\circ &= \sum_{k \in \mathcal{N}_{i, \neq}} \boldsymbol{\eta}_{ik}^\circ + \boldsymbol{\xi}_i^\circ \quad (\text{FEM - b}), \\ 3L_i^\circ &= \int_{\partial\Omega_i^\partial} r \sin \theta \varphi_i \mathbf{x}^\circ \cdot \mathbf{n}_i^\circ - \int_{\Omega_i} \varphi_i \mathbf{x}^\circ \cdot \hat{\mathbf{r}} - \int_{\Omega_i} r \sin \theta \mathbf{x}^\circ \cdot \nabla^\circ \varphi_i, \quad (\text{FEM - c}) \end{aligned} \quad (18)$$

and for the FV metric quantities

$$\begin{aligned} \mathbf{n}_i^\circ &= -\mathbf{n}_k^\circ \quad (\text{FVM - a}), & \widehat{\mathbf{V}}_i^\circ &= \sum_{k \in \mathcal{N}_{i, \neq}} \boldsymbol{\nu}_{ik}^\circ + \boldsymbol{\nu}_i^\circ \quad (\text{FVM - b}), \\ 3V_i^\circ &= \sum_{k \in \mathcal{N}_{i, \neq}} \frac{\mathbf{x}_i^\circ + \mathbf{x}_k^\circ}{2} \cdot \boldsymbol{\nu}_{ik}^\circ. \quad (\text{FVM - c}) \end{aligned} \quad (19)$$

Therefore, a FV approximation can be formally obtained from FE metric quantities defined over the same grid points by setting (see properties FEM/FVM-a and -b)

$$\boldsymbol{\nu}_{ik}^\circ = \boldsymbol{\eta}_{ik}^\circ, \quad \boldsymbol{\nu}_i^\circ = \boldsymbol{\xi}_i^\circ, \quad \widehat{\mathbf{V}}_i^\circ = \widehat{\mathbf{L}}_i^\circ.$$

By subtracting (FEM-c) to (FVM-c), one finally has

$$V_i^\circ = L_i^\circ + \sum_{k \in \mathcal{N}_{i, \neq}} \frac{\mathbf{x}_k^\circ - \mathbf{x}_i^\circ}{6} \cdot \boldsymbol{\zeta}_{ik}^\circ - \sum_{k \in \mathcal{N}_{i, \neq}} \frac{\mathbf{x}_k^\circ - \mathbf{x}_i^\circ}{6} \cdot \boldsymbol{\chi}_{ik}^\circ, \quad (20)$$

where, like the cylindrical case, the mass lumping approximation has been introduced. It is clear that also in this case the FV cell is not coincident with the FE lumped mass matrix and the shape of the FV cells that guarantees equivalence with FE discretization still remains undetermined.

5 FULLY DISCRETE FORM OF THE EULER EQUATIONS IN CYLINDRICAL AND SPHERICAL COORDINATES

The Euler equations in cylindrical and spherical coordinates for compressible inviscid flows are now briefly recalled. The differential form reads

$$\frac{\partial \mathbf{u}^\circ}{\partial t} + \nabla^\circ \cdot \mathbf{f}^\circ = \frac{1}{R} \mathbf{s}^\circ \quad \text{and} \quad \frac{\partial \mathbf{u}^\circ}{\partial t} + \nabla^\circ \cdot \mathbf{f}^\circ = \frac{1}{r \sin \theta} \mathbf{s}^\circ,$$

where $\mathbf{u}^\circ(Z, R, \theta, t) = (\rho, \mathbf{m}^\circ, E^t)^\top$ and $\mathbf{u}^\circ(r, \theta, \phi, t) = (\rho, \mathbf{m}^\circ, E^t)^\top$, with ρ density, $\mathbf{m}^\circ = (m_Z, m_R, m_\theta)^\top$, $\mathbf{m}^\circ = (m_r, m_\theta, m_\phi)^\top$ momentum and E^t total energy per unit

volume. The flux function and the source term are defined as follow in the cylindrical case

$$\mathbf{f}^\varnothing = \begin{pmatrix} m_Z & m_R & m_\theta \\ \frac{m_Z^2}{\rho} + \Pi & \frac{m_R m_Z}{\rho} & \frac{m_\theta m_Z}{\rho} \\ \frac{m_Z m_R}{\rho} & \frac{m_R^2}{\rho} + \Pi & \frac{m_\theta m_R}{\rho} \\ \frac{m_Z m_\theta}{\rho} & \frac{m_R m_\theta}{\rho} & \frac{m_\theta^2}{\rho} + \Pi \\ \frac{m_Z}{\rho} (E^t + \Pi) & \frac{m_R}{\rho} (E^t + \Pi) & \frac{m_\theta}{\rho} (E^t + \Pi) \end{pmatrix}, \quad \mathbf{s}^\varnothing = \begin{pmatrix} 0 \\ 0 \\ \frac{m_\theta^2}{\rho} + \Pi \\ \frac{m_\theta m_R}{\rho} \\ 0 \end{pmatrix},$$

and in the spherical one

$$\mathbf{f}^\circ(\mathbf{u}^\circ) = \begin{pmatrix} m_r & m_\theta & m_\phi \\ \frac{m_r^2}{\rho} + \Pi & \frac{m_\theta m_r}{\rho} & \frac{m_\phi m_r}{\rho} \\ \frac{m_r m_\theta}{\rho} & \frac{m_\theta^2}{\rho} + \Pi & \frac{m_\phi m_\theta}{\rho} \\ \frac{m_r m_\phi}{\rho} & \frac{m_\theta m_\phi}{\rho} & \frac{m_\phi^2}{\rho} + \Pi \\ \frac{m_r}{\rho} (E^t + \Pi) & \frac{m_\theta}{\rho} (E^t + \Pi) & \frac{m_\phi}{\rho} (E^t + \Pi) \end{pmatrix},$$

$$\mathbf{s}^\circ(\mathbf{u}^\circ) = \underbrace{\begin{pmatrix} 0 \\ \left(\frac{m_\theta^2}{\rho} + \Pi\right) + \left(\frac{m_\phi^2}{\rho} + \Pi\right) \\ -\frac{m_\theta m_r}{\rho} \\ -\frac{m_\phi m_r}{\rho} \\ 0 \end{pmatrix}}_{\check{\mathbf{s}}^\circ} \sin \theta + \underbrace{\begin{pmatrix} 0 \\ 0 \\ \frac{m_\phi^2}{\rho} + \Pi \\ -\frac{m_\phi m_\theta}{\rho} \\ 0 \end{pmatrix}}_{\bar{\mathbf{s}}^\circ} \cos \theta.$$

The FV spatially-discrete form of the Euler equations reads

$$V_i^\varnothing \frac{d\mathbf{u}_i^\varnothing}{dt} = - \sum_{k \in \mathcal{N}_{i,\neq}} \frac{\mathbf{f}_k^\varnothing + \mathbf{f}_i^\varnothing}{2} \cdot \boldsymbol{\eta}_{ik}^\varnothing + L_i \mathbf{s}_i^\varnothing - \mathbf{f}_i^\varnothing \cdot \boldsymbol{\xi}_i^\varnothing, \quad (21)$$

$$V_i^\circ \frac{d\mathbf{u}_i^\circ}{dt} = - \sum_{k \in \mathcal{N}_{i,\neq}} \frac{\mathbf{f}_i^\circ + \mathbf{f}_k^\circ}{2} \cdot \boldsymbol{\eta}_{ik}^\circ + \mathbf{f}_i^\circ \cdot \widehat{\mathbf{L}}_i^\circ - \mathbf{f}_i^\circ \cdot \boldsymbol{\xi}_i^\circ + \check{\mathbf{s}}_i^\circ \widehat{L}_i + \bar{\mathbf{s}}_i^\circ \check{L}_i,$$

where the following notation has been used $\mathbf{s}_i^\varnothing \stackrel{\text{def}}{=} \mathbf{s}^\varnothing(\mathbf{u}_i^\varnothing)$, $\check{\mathbf{s}}_i^\circ \stackrel{\text{def}}{=} \check{\mathbf{s}}^\circ(\mathbf{u}_i^\circ)$, $\bar{\mathbf{s}}_i^\circ \stackrel{\text{def}}{=} \bar{\mathbf{s}}^\circ(\mathbf{u}_i^\circ)$. The terms V_i^\varnothing and V_i° are computed from the equivalence condition in cylindrical and

spherical case, respectively. In the computation, a TVD [11] numerical flux is used, with the van Leer limiter [12]. The fully discrete form of the Euler system is obtained by a two-step Backward Differencing Formulæ. At each time level, a dual time-stepping technique is used to solve the time-implicit problem [13].

6 NUMERICAL RESULTS

In the present section, numerical results for diverging cylindrical and spherical shock waves are presented in the two-dimensional cases (R - θ and r - ϕ planes) and the one-dimensional case.

Initial conditions are as follows. The velocity is assumed to be zero everywhere; the density is uniform and equal to 1, whereas the pressure is uniform and equal to 10 in a circular region centered at the origin with radius $r = 0.5$. In the remaining portion of the domain, the pressure is uniform and equal to 1. In all computations, the ideal gas model for nitrogen ($\gamma = c_P/c_v = 1.39$) is used. The thermodynamics variables are made dimensionless by the corresponding value at rest (outer region), the spatial dimensions are made dimensionless by the radius of the domain. Velocity and time are made dimensionless by the square root of the reference pressure divided by the reference density; the reference time is the reference length divided by the reference velocity.

Numerical results are shown in figure 1, for the cylindrical and for the spherical cases, where the density isolines at different time levels are shown together with the pressure profiles along the axis $y = 0$. A shock wave propagates towards the outer boundary of the computational domain; the shock wave is followed by a contact discontinuity. A rarefaction wave propagates towards the origin and is then reflected outward. The grid is the made with 39 153 nodes 77 587 triangles and the time step is 1.5×10^{-4} . The pressure profiles are compared against reference one-dimensional results for the three different time levels. One-dimensional computations were performed over a evenly-spaced grid made of 2001 nodes, which corresponds to an element spacing of 5×10^{-4} .

7 CONCLUSIONS

A novel unstructured-grid hybrid finite element/volume method for a orthogonal curvilinear coordinates was presented. The method moves from suitable equivalence conditions linking finite element integrals to the corresponding finite volume metrics, such as the cell volume or the integrated normals. The equivalence conditions were derived here without introducing any approximation and allow to determine all needed finite volume metric quantities from finite element ones. Numerical results are presented for two-dimensional compressible flows: these consist in the numerical simulation of the explosion problem, in which an initial discontinuity in pressure results in the formation of a diverging shock. The computed pressure and density profile agree fairly well with one-dimensional simulation in cylindrical and spherical symmetry over a very fine grid.

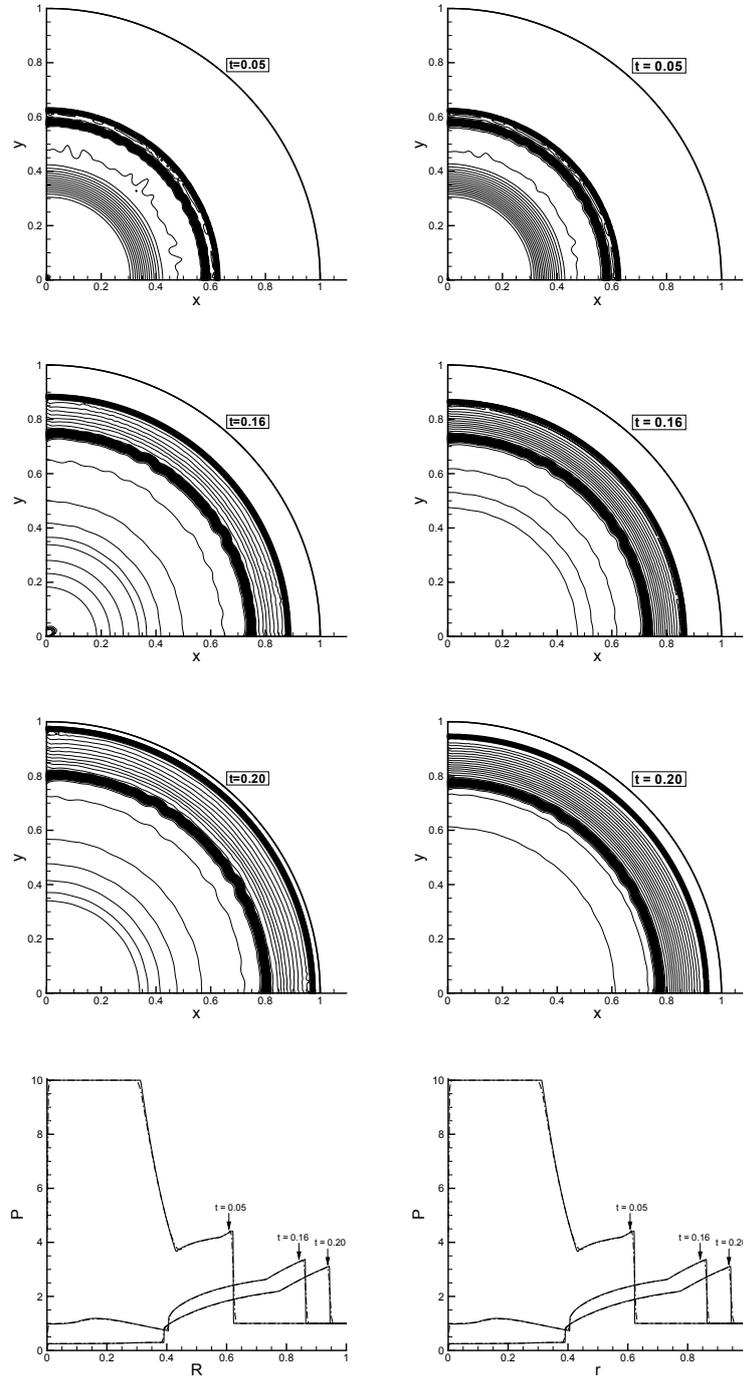


Figure 1: Density isolines for the explosion problem: left $R-\theta$ and right $r-\phi$ plane (Each isoline corresponds to a density difference of $\Delta\rho/\rho_{ref} = 0.03$). The pressure signal is along the axis $y = 0$. The solid line is the 1D reference solution, the dash-dotted line is the 2D solution.

REFERENCES

- [1] L. I. Sedov., *Similarity and dimensional methods in mechanics.*, Academic Press, 1959.
- [2] L. Fezou and B. Stoufflet., *A class of implicit upwind schemes for Euler simulations with unstructured meshes.*, J. Comput. Phys., 84(1):174–206, 1989.
- [3] V. Selmin., *The node-centred finite volume approach: bridge between finite differences and finite elements.*, Comp. Meth. Appl. Mech. Engng., 102:107–138, 1993.
- [4] V. Selmin. and L. Formaggia., *Unified construction of finite element and finite volume discretizations for compressible flows.*, Int. J. Numer. Meth. Eng., 39:1–32, 1996.
- [5] H. B. Chen, L. Zhang, and E. Panarella., *Stability of imploding spherical shock waves.*, Journal of Fusion Energy, 14:389–392, 1995.
- [6] A. Guardone and L. Vigevano., *Finite element/volume solution to axisymmetric conservation laws.*, J. Comput. Phys., 224:489–518, 2007.
- [7] D. De Santis, G. Geraci, and A. Guardone, *Equivalence conditions for finite volume/ element discretizations in cylindrical coordinates.*, V European Conference on Computational Fluid Dynamics ECCOMAS CFD 2010, June 2010.
- [8] A. Guardone, D. De Santis, G. Geraci, M. Pasta, *On the relation between finite element and finite volume schemes for compressible flows with cylindrical and spherical symmetry.*, J. Comput. Phys., 230:680–694, 2011.
- [9] J. Donea and A. Huerta., *Finite element methods for flow problems.*, Wiley, 2002.
- [10] R. J. LeVeque., *Finite volume methods for conservation laws and hyperbolic systems.*, Cambridge University Press, 2002.
- [11] A. Harten and J. M. Hyman., *Self adjusting grid methods for one-dimensional hyperbolic conservation laws.*, J. Comput. Phys., 50:253–269, 1983.
- [12] B. van Leer., *Towards the ultimate conservative difference scheme II. Monotonicity and conservation combined in a second order scheme.*, J. Comput. Phys., 14:361–370, 1974
- [13] V. Venkatakrishnan and D. J. Mavriplis., *Implicit method for the computation of unsteady flows on unstructured grids.*, J. Comput. Phys., 127:380–397, 1996.

ITERATIVE COUPLING OF BOUNDARY AND DISCRETE ELEMENT METHODS USING AN OVERLAPPING FEM ZONE

L. Malinowski^{*†}, G.F. Karlis[†], G. Beer[†] and J. Rojek[‡]

^{*†}Institute for Structural Analysis
Technical University Graz
Lessingstrasse 25, 8010 Graz, Austria
e-mail: lmalinowski@tugraz.at

[‡]Institute of Fundamental Technological Research
Polish Academy of Sciences
Pawińskiego 5B, 02-106 Warsaw, Poland

Key words: Coupled Problems, Geotechnical Engineering, Tunnelling, Boundary Element Method, Discrete Element Method, Iterative Coupling

Abstract. One of the characteristics of the numerical simulation in geotechnical engineering is that non-linear/discontinuous behaviour is concentrated on small portions of the total domain. A representative example is tunnel excavation, where high stress concentration is observed only near the excavation area and not over the whole rock mass. The Discrete Element Method (DEM) is ideal for handling discontinuous behaviour. However, employing DEM over the whole domain would be prohibitive, due to the large computational effort it requires. On the other hand, the Boundary Element Method (BEM) can effectively treat infinite and semi-infinite domains. The aim of the current work is to develop a simulation methodology that can iteratively couple Boundary Element and Discrete Element Methods in order to treat discontinuous behavior that arises in small portions of the total domain. To simplify the meshing procedure, an overlapping FEM zone is introduced on the DEM side of the common boundary.

1 INTRODUCTION

The Boundary Element Method (BEM) is an established numerical method that performs particularly well in problems where infinite domains are involved. Such problems are the norm in geomechanics, where the soil and the underground rock masses can be modelled as infinite or semi-infinite domains. The BEM has been successfully applied to a variety of problems related to underground structures [1, 2, 3, 4, 5, 6, 7].

However, the existence of non-linear behavior that can appear in problems like tunnel excavation, gives rise to additional volume integrals in BEM, that require the use of internal cells. Then the main advantage of the BEM, i.e. the discretization exclusively of the domain boundary, is partially lost, because of additional calculations that require integrations over the domain volume. In addition, discontinuous behavior can not be modelled.

In order to deal with these problems, the Discrete Element Method (DEM) is employed. DEM is a dynamic method, initially proposed by Cundall and Strack [8] and Cundall [9] and is used to investigate the behavior of geomaterials. A significant advantage of the DEM is its ability to model large displacements and discontinuous and non-linear behavior. In view of these advantages the DEM has been used to model the brittle behavior of rocks and to simulate crack initiation and propagation [16, 11, 12, 13]. However, these benefits come at a high computational cost, that prohibits the use of DEM in large domains.

The aim of the current work is to combine these two methods, by utilizing each, where they work best. It is common in geotechnical engineering problems, that discontinuous/non-linear behavior is confined to only small portions of the total domain. A representative example is tunnel excavation, where high stress concentration is observed and confined near the excavation zone and not over the whole rock mass. In the present work, an attempt is made to develop a simulation methodology that can iteratively couple the two aforementioned numerical methods in order to effectively treat non-linear/discontinuous behavior in problems related to tunnel excavation.

This paper is organized as follows: sections two and three briefly describe the BEM and DEM formulations that are used in the analysis. Section four describes the iterative coupling algorithm that is used and in section five the numerical results for a two-dimensional problem are obtained and discussed. Finally, section six contains the conclusions of this study and the plans for future work on that topic.

2 BEM FORMULATION

The boundary element method used in the present work is a displacement-based, collocation method that is typically derived from Betti's reciprocal theorem [14]. Based on that and Green's third identity, the following integral equation, for a linear elastic domain Ω can be derived.

$$\mathbf{c}(\mathbf{P}) \cdot \mathbf{u}(\mathbf{P}) = \int_{\partial\Omega} \mathbf{U}(\mathbf{P}; \mathbf{Q}) \cdot \mathbf{t}(\mathbf{Q}) dS - \int_{\partial\Omega} \mathbf{T}(\mathbf{P}; \mathbf{Q}) \cdot \mathbf{u}(\mathbf{Q}) dS \quad (1)$$

with $\mathbf{u}(\mathbf{P})$, $\mathbf{t}(\mathbf{P})$ being the displacement and traction vectors at point \mathbf{P} , point \mathbf{Q} running over the domain boundary ($\partial\Omega$) and $\mathbf{U}(\mathbf{P}; \mathbf{Q})$, $\mathbf{T}(\mathbf{P}; \mathbf{Q})$ being the fundamental solutions for the displacements and tractions respectively. Furthermore, the tensor $\mathbf{c}(\mathbf{P})$ is equal to the unit tensor \mathbf{I} , if \mathbf{P} lies inside the domain, $1/2\mathbf{I}$ if \mathbf{P} lies on the domain's smooth boundary and $\mathbf{0}$ if \mathbf{P} lies outside of the domain.

Discretizing the domain boundary into nodes and elements and writing the above equation for every node of the boundary of the domain, the following integral representation for node \mathbf{x}^k is obtained.

$$\mathbf{c}(\mathbf{x}^k) \cdot \mathbf{u}^k = \sum_{e=1}^E \sum_{i=1}^{A(e)} \Delta \mathbf{U}_{ei}^k \cdot \mathbf{t}^k - \sum_{e=1}^E \sum_{i=1}^{A(e)} \Delta \mathbf{T}_{ei}^k \cdot \mathbf{u}^k \quad (2)$$

with \mathbf{u}^k , \mathbf{t}^k being the displacements and tractions of node \mathbf{x}^k , E being the total number of elements and $A(e)$ the total number of nodes for element e . Moreover, $\Delta \mathbf{U}_{ei}^k$ and $\Delta \mathbf{T}_{ei}^k$ are matrices that contain integrals over the element e , of the fundamental solutions U and T respectively, multiplied by the i -th interpolation function and Jacobian of element e , that are typically calculated using Gauss quadrature.

Collocating eq (2) for all boundary nodes and utilizing the boundary conditions of the problem, results to the final linear system for the BEM, which is of the form

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \quad (3)$$

3 DEM FORMULATION

The Discrete Element Method is a dynamic method, where the material is represented as a collection of particles interacting with each other. The formulation utilized in the present work is the one proposed by Oñate and Rojek [15] and Labra et al.[17].

Assuming N particles in total and writing the equations of motion for translations and rotations for all of them results in

$$\mathbf{M}_D \cdot \{\ddot{\mathbf{u}}\} = \{\mathbf{F}\} \quad (4)$$

$$\mathbf{J}_D \cdot \{\dot{\boldsymbol{\omega}}\} = \{\mathbf{T}\} \quad (5)$$

with $\{\ddot{\mathbf{u}}\}$ and $\{\dot{\boldsymbol{\omega}}\}$ containing the components of acceleration and angular acceleration respectively. Furthermore, matrices \mathbf{M}_D and \mathbf{J}_D are block diagonal matrices that involve the particle masses m_i and moments of inertial J_i and are of the form

$$\mathbf{M}_D = \begin{bmatrix} m_1 \cdot \mathbf{I} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & m_N \cdot \mathbf{I} \end{bmatrix} \quad \text{and} \quad \mathbf{J}_D = \begin{bmatrix} J_1 \cdot \mathbf{I} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & J_N \cdot \mathbf{I} \end{bmatrix} \quad (6)$$

with \mathbf{I} being the 2×2 and 3×3 identity matrix for 2D and 3D respectively. Finally vectors $\{\mathbf{F}\}$ and $\{\mathbf{T}\}$ contain the total force and total moment for each degree of freedom of each particle, respectively. It is worth noting that the total force applied to a particle is described as the sum of the external forces, the contact forces from the adjacent particles and the forces resulting from the imposed external damping. Similarly, the total moment is given by the sum of the external moment due to the external load, the rotational

component of the contact forces between adjacent particles and the moment resulting from the imposed external damping.

Since DEM is a dynamic method, eqs (4) and (5) have to be integrated in time. This is done by employing a central difference scheme. The calculation of the displacements and the incremental rotations for step $n + 1$ is done in terms of the displacement and rotation and their time derivatives over the previous time step, utilizing also an intermediate, auxiliary time step, indicated as $n + 1/2$.

$$\{\ddot{\mathbf{u}}\}^n = \mathbf{M}_D^{-1} \{\mathbf{F}\}^n \quad (7)$$

$$\{\dot{\mathbf{u}}\}^{n+1/2} = \{\dot{\mathbf{u}}\}^{n-1/2} + \{\ddot{\mathbf{u}}\}^n \Delta t \quad (8)$$

$$\{\mathbf{u}\}^{n+1} = \{\mathbf{u}\}^n + \{\dot{\mathbf{u}}\}^{n+1/2} \Delta t \quad (9)$$

$$\{\dot{\boldsymbol{\omega}}\}^n = \mathbf{J}_D^{-1} \{\mathbf{T}\}^n \quad (10)$$

$$\{\boldsymbol{\omega}\}^{n+1/2} = \{\boldsymbol{\omega}\}^{n-1/2} + \{\dot{\boldsymbol{\omega}}\}^n \Delta t \quad (11)$$

$$\Delta \{\boldsymbol{\theta}\}^{n+1} = \{\boldsymbol{\omega}\}^{n+1/2} \Delta t \quad (12)$$

with $\Delta \{\boldsymbol{\theta}\}$ being the incremental rotations.

4 ITERATIVE COUPLING

In order to couple the BEM and DEM subdomains, an iterative coupling scheme has been adopted. An important advantage of iterative coupling is that it allows solving each region separately, without having to assemble a global linear system. The interaction between the adjacent regions during the iteration steps is achieved with the exchange of boundary conditions on their common boundary. From now on, the term interface will be used to describe the common boundary of the two subdomains. The proposed scheme is a strong coupling scheme, i.e. there must be a “1-1” correspondence of the nodes of the two adjacent subdomains on their common boundary.

A basic requirement for the coupling scheme to work, is that the same fields must be calculated on all interfaces, regardless the method that is implemented in each subdomain. Otherwise, the exchange of boundary conditions will not be possible. To that end, the BEM formulation described in Section 2 was modified to calculate or accept as boundary conditions the nodal point forces instead of the element-based tractions on the interface.

4.1 BEM - Nodal Force Calculation on the Interface

As explained by Beer et al.[14], the nodal tractions on the interface of a BEM subdomain can be expressed as

$$\{\mathbf{t}\}_c = \{\mathbf{t}\}_{c0} + \mathbf{K}_{BE} \{\mathbf{u}\}_c \quad (13)$$

where the vector $\{\mathbf{t}\}_c$ contains the tractions on the interface nodes, $\{\mathbf{t}\}_{c0}$ contains the tractions on the interface nodes due to the applied loads on the rest of the boundary, $\{\mathbf{u}\}_c$

contains the displacements on the interface nodes and \mathbf{K}_{BE} is the so called pseudo-stiffness matrix, calculated for the interface nodes.

To calculate $\{\mathbf{t}\}_{c0}$, one has to solve the BEM subdomain using the linear system (3) with a constrained interface ($\{\mathbf{u}\}_c = \mathbf{0}$) and the boundary conditions on the rest of the domain boundary as specified by the problem.

For calculating the pseudo-stiffness matrix \mathbf{K}_{BE} , the linear system (3) is utilized again. The collocation procedure is conducted for all nodes of the domain, in order to assemble the system matrix and all nodes, including those of the interface, are considered to have unknown tractions. Then the linear system is solved (number of interface nodes) \times (number of dofs per node) times. Each time, a unit displacement is applied to one degree of freedom of one node of the interface, whereas all other displacements are set to zero. This results to the following traction vector:

$$\{\bar{\mathbf{t}}\}_i = \begin{Bmatrix} \{\mathbf{t}\}_i^b \\ \{\mathbf{t}\}_i^c \end{Bmatrix} \quad (14)$$

with i ranging from 1 to (number of interface nodes) \times (number of dofs per node), $\{\mathbf{t}\}_i^c$ indicating the interface tractions and $\{\mathbf{t}\}_i^b$ indicating the tractions on the rest of the boundary of the subdomain. Using vectors $\{\bar{\mathbf{t}}_i\}$ as columns, a matrix \mathbf{B} can be constructed, with dimensions [(number of nodes) \times (number of dofs per node)] \times [(number of interface nodes) \times (number of dofs per node)].

$$\mathbf{B} = \begin{bmatrix} \{\mathbf{t}\}_1^b & \{\mathbf{t}\}_2^b & \dots \\ \{\mathbf{t}\}_1^c & \{\mathbf{t}\}_2^c & \dots \end{bmatrix} \quad (15)$$

Isolating the part of \mathbf{B} that corresponds to the tractions on the interface, yields the pseudo-stiffness matrix \mathbf{K}_{BE} .

$$\mathbf{K}_{BE} = [\{\mathbf{t}\}_1^c \quad \{\mathbf{t}\}_2^c \quad \dots] \quad (16)$$

After \mathbf{K}_{BE} and the applied load contribution $\{\mathbf{t}\}_{c0}$ have been calculated, eq (13) can be solved either for displacements or for tractions on the interface nodes. However, it requires further processing in order to utilize nodal forces instead of tractions.

To obtain the x -component of the equivalent nodal point force for a node k , we apply a unit virtual displacement to that node in the x direction. Then, the work done by the tractions at that point must be equal to the work done by the equivalent nodal point force.

$$F_x^k \cdot 1 = \int_S t_x \delta u_x dS \quad (17)$$

with S being the boundary of the domain (including the interface). By discretizing eq (17)

and introducing the displacement and traction interpolation functions

$$\begin{aligned} t_x &= \sum_{n=1}^{N_e} N_n t_{xe}^n \\ \delta u_x &= N_j \delta u_{xj} = N_j \cdot 1, j = \text{local node numbering of node } k \end{aligned}$$

we end up with the following expression:

$$F_x^k = \sum_{e=1}^E \int_{S_e} \sum_{n=1}^{N_e} N_n t_{xn}^e \quad (18)$$

with E being the total number of elements that node k belongs to and N_e being the number of nodes of element e . The same procedure can be followed for the y -component of the nodal force and the total equivalent nodal force at node k can be expressed using vector notation as follows:

$$\mathbf{F}_k = \sum_{e=1}^E \sum_{n=1}^{N_e} \mathbf{N}_{jn}^e t_n^e \quad (19)$$

with E, N_e being the same as in eq (18) and \mathbf{N}_{jn}^e being a matrix of the form

$$\mathbf{N}_{jn}^e = N_{jn}^e \cdot \mathbf{I}, \text{ with } N_{jn}^e = \int_{S_e} N_j N_n dS_e \quad (20)$$

where N_j, N_n are the displacement and traction interpolation functions respectively.

Repeating the same procedure for every node of the interface (i.e. where we need the equivalent nodal point forces) we obtain

$$\{\mathbf{F}\}_c = \mathbf{N} \{\mathbf{t}\}_c = \mathbf{N} \{\mathbf{t}\}_{c0} + \mathbf{N} \mathbf{K}_{BE} \{\mathbf{u}\}_c \quad (21)$$

with matrix \mathbf{N} containing the contributions of \mathbf{N}_{jn}^e . The above procedure is explained in further detail in [14] and results to

$$\{\mathbf{F}\}_c = \{\mathbf{F}\}_{c0} + \mathbf{K} \{\mathbf{u}\}_c \quad (22)$$

where $\{\mathbf{F}\}_c$ and $\{\mathbf{u}\}_c$ are the nodal point forces and the displacements on the interface nodes respectively, $\{\mathbf{F}\}_{c0}$ is the contribution of the external load on the interface nodes and \mathbf{K} is the stiffness matrix of the interface.

Now using eq (22) the displacements or the forces on the interface can be calculated, when the forces or displacements are provided from the adjacent region respectively. The results of the calculation can then be transferred back to the adjacent subdomain. Note that during the iterative procedure only the interface has to be solved.

From the above, it is easy to see that in order to calculate the solution at the subdomain interface, only the calculation of the matrix \mathbf{K} and of the forces on the interface nodes $\{\mathbf{F}\}_{c0}$ is required and is done only once. This means that the fields on the boundary of the subdomain do not have to be calculated in every iteration step, but only in the end, after the iterative procedure has converged.

4.2 DEM - Overlapping FEM Zone

Directly coupling DEM and BEM using a strong coupling scheme would impose additional requirements on the DEM mesh, such as gradually increasing the particle size to match that of the BEM element size on the interface. In order to avoid this problem, a FEM zone was introduced between the BEM and DEM subdomains, that partially overlaps with the DEM subdomain, as described by Rojek and Oñate [16] and Labra et al.[17].

In short, the methodology described in Section 3 is enhanced with the so called explicit dynamic FEM formulation over the FEM zone.

$$\mathbf{M}_F \cdot \{\ddot{\mathbf{u}}\}_F = \{\mathbf{F}\}_F^{ext} - \{\mathbf{F}\}_F^{int} \quad (23)$$

with \mathbf{M}_F being the mass matrix and \mathbf{u}_F , \mathbf{F}_F^{ext} and \mathbf{F}_F^{int} being the displacements, external and internal forces respectively. The same central difference scheme as before has been utilized for the time integration.

Over the overlapping zone, where DEM and FEM coexist, additional kinematic constraints have to be applied to the DEM particles in order to constrain their displacement field by that of the underlying finite elements. The above constraints are imposed by means of a penalty function which forms the final linear system for the DEM-FEM subdomain along with the equations of motion derived from the DEM and FEM formulations. This formulation is described in great detail in [17].

4.3 Coupling Algorithm

A common problem in iterative coupling is avoiding floating domains. In that context, each static subdomain should either have sufficiently prescribed displacements or be of infinite extent, to avoid being subjected to rigid body motion. In the case neither of the above is true, the required displacements should be provided by an adjacent region in the form of appropriate boundary conditions. To fulfill this requirement in a general manner, an algorithm has been developed to determine the sequence in which the subdomains will be solved during each iteration step and what kind of boundary conditions should be imposed on their interfaces with the adjacent regions. This algorithm is invoked only once, before starting the iterative procedure and results in an array, referred to as the solving sequence array, that contains all the subdomains of the problem in the reverse sequence in which they should be solved. Furthermore, the boundary conditions that will be used on the interfaces are determined for each subdomain. The algorithm consists of the following steps:

- While there are still unhandled subdomains
 - If current subdomain has sufficient displacement boundary conditions (bcs) or is infinite
 - ◊ Set Neumann type bcs on its interfaces with other subdomains

- ◊ For all adjacent subdomains set Dirichlet type bcs on their interfaces with the current subdomain
- ◊ Add current subdomain to the solving sequence array

After the solving sequence and the types of boundary conditions have been determined on all interfaces, the iterative procedure takes place:

- While convergence or maximum number of iterations have not been reached
 - Run through the solving sequence array from the back to the front and back again
 - ◊ For the current subdomain, obtain boundary conditions from adjacent subdomains
 - ◊ Solve current subdomain
 - ◊ Check for convergence

The convergence criterion used, was the one proposed by Elleithy et al.[18]. Namely, the displacements of the BEM subdomain on the interface $\{\mathbf{u}_{BEM}\}_I$ are checked over two consequent iteration steps.

$$\frac{\|\{\mathbf{u}_{BEM}\}_I^{n+1} - \{\mathbf{u}_{BEM}\}_I^n\|}{\|\{\mathbf{u}_{BEM}\}_I^{n+1}\|} < \epsilon \quad (24)$$

with n and $n + 1$ being consequent iteration steps and ϵ being the predefined tolerance.

To illustrate the procedure described above, consider two subdomains, Ω_1 and Ω_2 that have a common boundary, as shown in Figure 1. The elements of Ω_1 that lie on the common boundary with Ω_2 form the interface I_{12} , whereas the elements of Ω_2 that lie on the common boundary with Ω_1 form the interface that is indicated as I_{21} . The displacements and forces on I_{12} , at the n -th iteration step, are denoted as \mathbf{u}_{12}^n and \mathbf{f}_{12}^n respectively. Similarly, \mathbf{u}_{21}^n and \mathbf{f}_{21}^n are used to indicate the displacements and forces on the interface I_{21} .

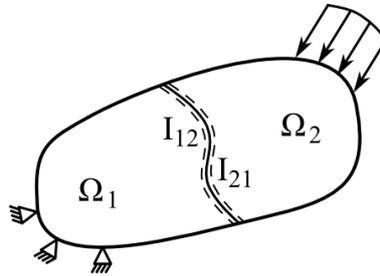


Figure 1: A problem with two subdomains sharing a common boundary

Now assume that the subdomain Ω_1 is fixed and that a constant load is applied to Ω_2 . In that case, the algorithm sequence determination algorithm would check Ω_1 , and since it is fixed, it would apply Neumann boundary conditions to I_{12} , Dirichlet boundary conditions on I_{21} and would store Ω_1 in the solving sequence array. Afterwards, the algorithm would check subdomain Ω_2 , which now has prescribed Dirichlet boundary conditions (on I_{21}). It has no interfaces with subdomains other than Ω_1 , which has already been addressed. The subdomain Ω_2 is added to the solving sequence array and the algorithm ends. At this point both subdomains have been assigned with a boundary condition on their interface.

The next step is to proceed to the iterative procedure itself. The first subdomain to be solved is the last added to the solving sequence array, i.e. Ω_2 . In order to solve Ω_2 , the displacements have to be specified on the interface I_{21} . For the first iteration, zero displacements are assumed, whereas for the next iterations, the displacements are provided by Ω_1 .

To ensure convergence for the aforementioned iterative procedure, a relaxation scheme had to be adopted. A wide range of relaxation algorithms exist in the literature [19]. For the present work, the same relaxation scheme was applied on all interfaces, regardless the type of the imposed boundary conditions, as shown in eqs (25) and (26).

$$\mathbf{u}_{12}^{n+1} = \alpha \mathbf{u}_{21}^{n+1} + (1 - \alpha) \mathbf{u}_{12}^n \quad (25)$$

$$\mathbf{f}_{12}^{n+1} = \beta \mathbf{f}_{21}^{n+1} + (1 - \beta) \mathbf{f}_{12}^n \quad (26)$$

The values of the relaxation parameters α and β are problem dependent and can affect not only the convergence speed of the iterative procedure, but also the convergence itself.

5 NUMERICAL EXAMPLE

In order to demonstrate the aforementioned coupling strategy, a two dimensional, linear elastic compression test is solved. The problem domain consists of an $\alpha \times 2\alpha$ rectangular plate, with $\alpha = 0.5$ meters.

The upper half of the plate is modelled as a linear elastic BEM region with Young's modulus $E = 100GPa$ and Poisson's ratio $\nu = 0$. The lower half of the plate was modelled as a DEM region with a partially overlapping FEM zone, as shown in Figure 2(a). Note that the DEM and the FEM regions are treated as one subdomain, as described in Section 4.2. For the FEM zone, the same material parameters are used, whereas for the DEM the particle radius is $r = 0.025\alpha$ and the normal stiffness was set to $k_n = 94.6GPa$ in order to obtain the same effective Young's modulus as in the BEM subdomain.

A uniform load $L = 0.05GPa$ is applied to the top of the plate, whereas its base is constrained with zero displacement in all directions. During the iterative procedure, the BEM subdomain is solved first with zero displacement prescribed on its interface with the adjacent subdomain. The forces calculated on the interface are applied to the DEM as boundary condition. Since the DEM is dynamic, the boundary condition was applied incrementally starting from zero force at time $t = 0$ and increasing linearly until the

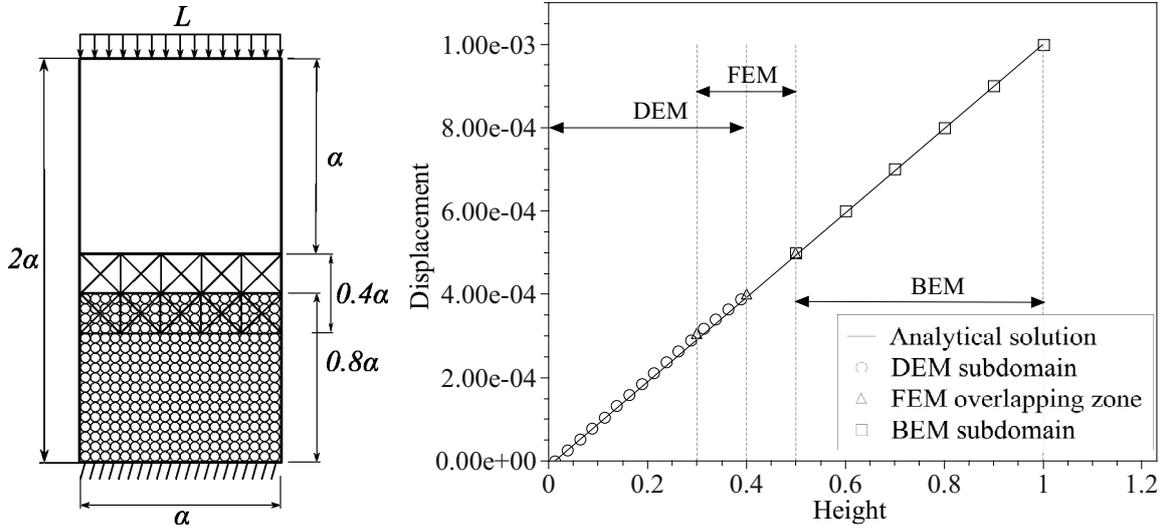


Figure 2: (a) The problem domain and (b) the displacements along the vertical axis.

value provided by the BEM is reached at time $t = 0.01\text{sec}$ and remains constant until the total time $T = 0.02\text{sec}$ is reached. Then, the obtained results on the interface are transferred to the BEM subdomain allowing the iterative process to continue.

The convergence criterion was set to $\epsilon = 1e - 4$. After the iterative procedure has converged, the displacements along the vertical direction were compared to the analytical solution of the problem (Figure 2(b)). The mean error obtained was 0.02%.

6 CONCLUSIONS AND FUTURE WORK

In the present work a DEM-BEM iterative coupling scheme has been presented. In order to simplify the coupling procedure, a FEM zone has been introduced between the two subdomains, partially overlapping with the DEM region. The DEM with the overlapping FEM zone and the BEM subdomains were solved independently from one another during the coupling procedure and the interaction between them was achieved by exchanging boundary conditions on their common boundary. The type of boundary conditions to be used on the interface was automatically determined, taking into account the existing boundary conditions of each subdomain, in order to avoid floating domains.

A numerical example has been solved to demonstrate the aforementioned algorithm. The convergence of the coupling scheme depends on the values of the relaxation parameters. The optimal values of these parameters are problem specific and to calculate them efficiently, an automatic relaxation algorithm should be used, as the ones proposed by Yan et al.[20] and Lin et al.[21].

In order to be able to model non-linear effects on the DEM subdomain with the proposed scheme, additional feedback from the BEM subdomain is required, during the

solution of the DEM region. Since DEM is a dynamic method, the strategy to follow would be to pause the DEM calculation at specific time steps, couple iteratively with the BEM region and after the iterative procedure has converged, proceed to the next time step. However, this case requires further investigation and is the subject of future work.

REFERENCES

- [1] Beer, G. and Watson, O.W. and Swoboda, G. Three-dimensional analysis of tunnels using infinite Boundary Elements. *Computers and Geotechnics* (1987) **3(1)**:37–58.
- [2] Pande G.N. and Beer, G. and Williams, J.R. Numerical Methods in Rock Mechanics. J.Wiley, (1987).
- [3] Beer, G. Numerical Simulation in Tunneling. Springer, (2003).
- [4] Beer, G. Technology Innovation in Underground Construction. CRC Press, (2009).
- [5] Moser, W. and Duenser, Ch. and Beer G. Mapped infinite elements for three-dimensional multi-region boundary element analysis. *International Journal for Numerical Methods in Engineering* (2004) **61(3)**:314–328.
- [6] von DEstorff, O. and Stamos, A.A. and Beskos, D.E. and Antes H. Dynamic interaction effects in underground traffic systems. *Engineering Analysis with Boundary Elements* (1991) **8(4)**:167–175.
- [7] Pöttler, R. and Swoboda, G.A. Coupled Beam-Boundary-element model (BE-BEM) for analysis of underground openings. *Computers and Geotechnics* (1986) **2(4)**:239–256.
- [8] Cundall, P.A. and Strack, O.D.L. A discrete numerical method for granular assemblies. *Geotechnique* (1979) **29**:47–65.
- [9] Cundall, P.A. Formulation of a three dimensional distinct element model-Part I. A scheme to detect and represent contacts in a system of many polyhedral blocks. *International Journal of Rock Mechanics, Mining Sciences & Geotechnics Abstracts* (1988) **25(3)**:107–116.
- [10] Rojek, J. Discrete Element Modelling of Rock Cutting. *Computer Methods in Material Science* (2007) **7(2)**:224–230.
- [11] Potyondy, D. and Cundall, P.A. A bonded-particle model for rock. *International Journal of Rock Mechanics and Mining Sciences* (2004) **41**:1329–1364.
- [12] Shiu, W. and Donzé, F.V. and Daudeville, L. Compaction process in concrete during missile impact: a DEM analysis. *Computers and Concrete* (2008) **5(4)**:329–342.

- [13] Moon, T. and Nakagawa, M. and Berger, J. Measuring of Fracture Toughness using the Distinct Element Method. *International Journal of Rock Mechanics and Mining Sciences* (2006) **44(3)**:449–456.
- [14] Beer, G. and Smith, I. and Duenser, Ch. The Boundary Element Method in Programming for Engineers and Scientists. Springer, (2008).
- [15] Oñate, E. and Rojek, J. Combination of discrete and finite element methods for dynamic analysis of geomechanics problems. *Computer Methods in Applied Mechanics and Engineering* (2004) **193(27-29)**:3087–3128.
- [16] Rojek, J. and Oñate, E. Multiscale analysis using a coupled discrete/finite element model. *Interaction and Multiscale Mechanics* (2007) **1(1)**:1–31.
- [17] Labra, C. and Rojek, J. and Oñate, E. and Zarate, F. Advances in discrete element modelling of underground excavations. *Acta Geotechnica* (2008) **3(4)**:317–322.
- [18] Elleithy, W.M. and Al-Gahtani, H.J. and El-Gebeily, M. Iterative coupling of BE and FE methods in elastostatics. *Engineering Analysis with Boundary Elements* (2001) **25**:685–695.
- [19] Elleithy, W.M. and Tanaka, M. Interface relaxation algorithms for BEM-BEM coupling and FEM-BEM coupling. *Computer Methods in Applied Mechanics and Engineering* (2003) **192**:2977–2992.
- [20] Yan, B. and Du, J. and Hu, Ning. and Sekine, H. A domain decomposition algorithm with finite element - boundary element coupling. *Applied Mathematics and Mechanics (English Edition)* (2006) **27(4)**:519–525.
- [21] Lin, C.-C. and Lawton, E.C. and Caliendo, J.A. and Anderson, L.R. An iterative finite element - boundary element algorithm. *Computers & Structures* (1996) **59(5)**:899–909.

A HARDWARE-BASED MULTI-DISCIPLINARY DESIGN OPTIMISATION METHOD FOR AERONAUTICAL APPLICATIONS

J. KOK*, F. GONZALEZ*, N. KELSON† AND T. GURNETT†

*Australian Research Centre for Aerospace Automation (ARCAA)
Queensland University of Technology (QUT)
Eagle Farm, 4009 Queensland, Australia
e-mail: info@arcaa.aero, www.arcaa.aero/

†High Performance Computing (HPC) and Research Support Group
Queensland University of Technology (QUT)
Brisbane, 4000 Queensland, Australia
e-mail: qut.itshpc@qut.edu.au, www.itservices.qut.edu.au/hpc/

Key words: Aeronautics, Field Programmable Gate Array, Multi-disciplinary Design Optimisation, Unmanned Aerial Vehicle

Abstract. There are many applications in aeronautics where there exist strong couplings between disciplines. One practical example is within the context of Unmanned Aerial Vehicle (UAV) automation where there exists strong coupling between operation constraints, aerodynamics, vehicle dynamics, mission and path planning. UAV path planning can be done either online or offline. The current state of path planning optimisation online UAVs with high performance computation is not at the same level as its ground-based offline optimizer's counterpart, this is mainly due to the volume, power and weight limitations on the UAV; some small UAVs do not have the computational power needed for some optimisation and path planning task. In this paper, we describe an optimisation method which can be applied to Multi-disciplinary Design Optimisation problems and UAV path planning problems. Hardware-based design optimisation techniques are used. The power and physical limitations of UAV, which may not be a problem in PC-based solutions, can be approached by utilizing a Field Programmable Gate Array (FPGA) as an algorithm accelerator. The inevitable latency produced by the iterative process of an Evolutionary Algorithm (EA) is concealed by exploiting the parallelism component within the dataflow paradigm of the EA on an FPGA architecture. Results compare software PC-based solutions and the hardware-based solutions for benchmark mathematical problems as well as a simple real world engineering problem. Results also indicate the practicality of the method which can be used for more complex single and multi-objective coupled problems in aeronautical applications.

1 INTRODUCTION

Multi-disciplinary design optimisation (MDO) has been actively applied across engineering disciplines over a wide range of problems, the majority of which lies within the field of aeronautics [1]. The complexity of modern systems has directed research focus towards improving MDO search algorithms and analysis tools [2–4]. One method to speed up the runtime of an MDO search algorithm is to implement its features in hardware, where concurrent data processing is possible. MDO search algorithms that take several hours to run could be executed in fractions of a second, impacting significantly on the design time of a project.

One such hardware which has advanced extensively in technology is the Field Programmable Gate Array (FPGA). An FPGA implementation in the field of aeronautics will also contribute to the Unmanned Aerial Vehicle (UAV) automation community where it is necessary to have a high level of integration within the capabilities of a vehicle. One practical example is within the context of Unmanned Aerial Vehicle (UAV) automation where there exists strong coupling between operation constraints, aerodynamics, vehicle dynamics, mission and path planning. UAV path planning can be done either online or offline. The current state of path planning optimisation online UAVs with high performance computation is not at the same level as its ground-based offline optimizer’s counterpart, this is mainly due to the volume, power and weight limitations on the UAV; some small UAVs do not have the computational power needed for some optimisation and path planning task. In this paper, we describe an optimisation method which can be applied to Multi-disciplinary Design Optimisation problems and UAV path planning problems. Hardware-based design optimisation techniques are used. The power and physical limitations of UAV, which may not be a problem in PC-based solutions, can be approached by utilizing a Field Programmable Gate Array (FPGA) as an algorithm accelerator. The inevitable latency produced by the iterative process of an Evolutionary Algorithm (EA) is concealed by exploiting the parallelism component within the dataflow paradigm of the EA on an FPGA architecture.

In this paper, an FPGA implementation of the search algorithm of an MDO is proposed. Our proposed design inherits the diversity and Pareto optimality features necessary for an effective MDO search algorithm. Simulation results showing speed up of 43 times for discrete and discontinuous test problems, and 244 times for the constrained test problem.

2 BACKGROUND

Previous work incorporating search algorithm design onto FPGA [5–10] have ascertained this concept. One of the limitations was the application to single objective search algorithms only, which is not realistically applicable to real-world scenarios where conflicting objectives often arise. An important point to note is that the paradigm behind an MDO search algorithm design is very different from that of a single objective optimisation [11]; the MDO search algorithm is aimed to find multiple trade-off optimal solutions with a wide range of

values across the objectives, whereas a single objective optimisation search algorithm is directed towards one optimal solution. Bonissone and Subbu [12] presented an Evolutionary Multi-objective Optimisation on an FPGA but were restricted to the basic filtering of Pareto fronts and not addressing the core features such as solution diversity and Pareto optimality.

3 METHODOLOGY

The methodology applied to this research is as depicted in Figure 1 below. Coinciding with the methodology, the rest of the paper is organized as follows. Section 4 summarises the features of EAs which contribute to a robust MDO search algorithm. Section 5 presents the proposed FPGA-based MDO implementation. Section 6 reports on the test problems conducted to verify the functionality of the FPGA-based MDO and demonstrate its effectiveness. Section 7 concludes with a brief summary and highlighting future work.

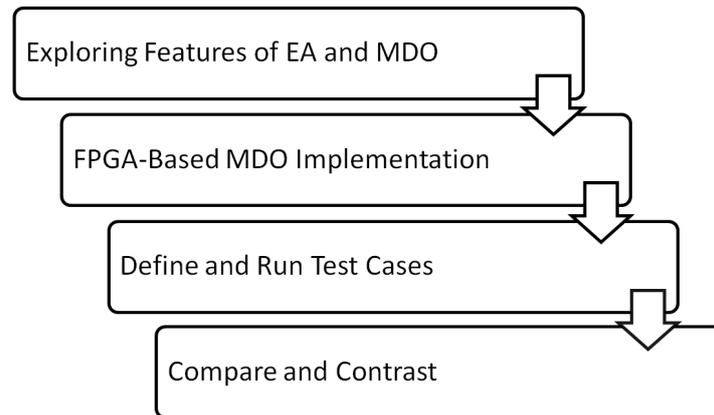


Figure 1: Methodology for developing an FPGA-based MDO

4 MULTI-DISCIPLINARY DESIGN OPTIMISATION

The two main objectives in multi-objective optimisation and MDO is to find a set of solutions which are as diverse as possible and as close as possible to the set of optimal solutions within the feasible search space, known as the Pareto-optimal front [11]. There are a number of search methods used in MDO, such as NSGA-II [13], DPGA [14], SPEA [15], TDGA [16] and PAES [17]. One class of the techniques actively applied on MDO is Evolutionary Algorithms (EAs) [18,19], which is a Metaheuristic optimisation algorithm inspired by the theory of natural evolution. The attractiveness of EAs lies in its population-based characteristic, where manipulating multiple solutions leads to the simultaneous discovery of a Pareto set [4]. There are a number of EAs and MOEAs with different features. In this paper we use and explore the features of a well known multi-objective EA, NSGA-II [13]. Even though NSGA-II has some drawback [13], it also has very good

features such as diversity maintenance technique using a crowding distance assignment and Pareto ranking for preserving elitists.

5 PROPOSED FPGA-BASED MDO

5.1 Algorithm Overview

The architecture of the FPGA-based MDO is depicted below in Figure 2. The algorithm incorporates the key features, fixed-point representation, random number generator, crossover, mutation, Pareto front ranking, crowding distance assignment, selection and evaluation, which are needed for securing diverse Pareto-optimal fronts. The population, consisting of candidate solutions, is stored on the block RAM onboard the FPGA. Tournament selection is randomly carried out across the population, determining better candidate solutions to be genetically altered, which produces the offspring for that current generation. After the offspring have been crossed over, mutated and evaluated, they are concatenated with the parent population to undergo Pareto front ranking and crowding distance assignment. The higher ranking and wider spread solutions are updated back into the population block RAM. It is stressed that the data flow arrow in Figure 2 denotes parallel processing. The following subsections will describe the functionality of the basic features and modules of the proposed FPGA-based MDO.

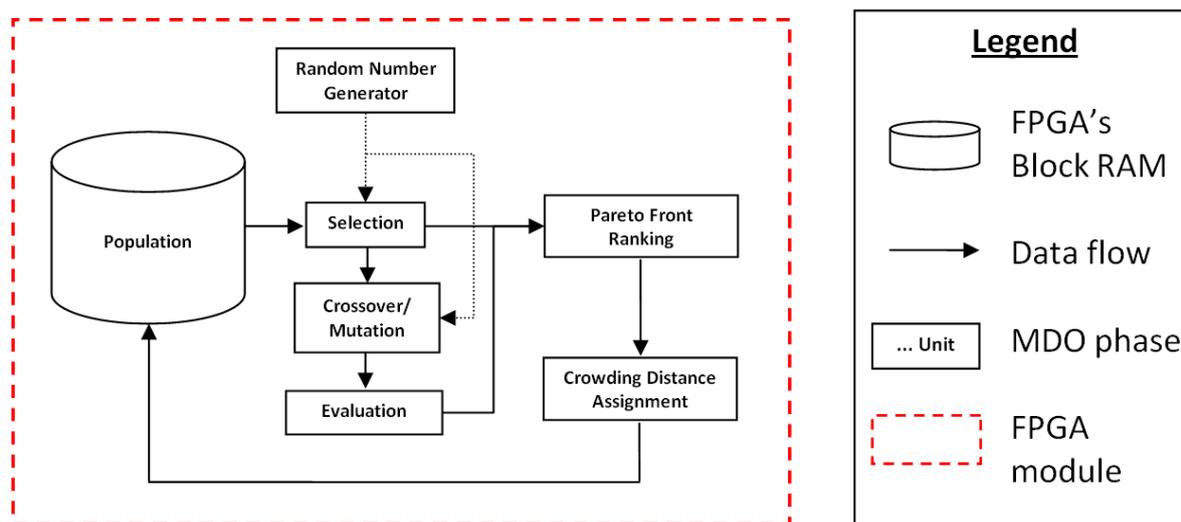


Figure 2: Architecture of the FPGA-based MDO search algorithm.

5.2 Representation

One of the first design decisions is determining the representation of candidate solutions. These could be coded either in real-coded binary, floating-point or fixed-point numbers. Michalewicz [19] experimented and concluded that floating-point and fixed-point represen-

tation is faster, more consistent and higher in precision than real-coded representation. Since the complexity of floating-point arithmetic consumes a larger logic footprint and is not as efficient as fixed-point arithmetic, therefore the candidate parameters are encoded in fixed-point representation.

5.3 Random Number Generator

When deciding on a logic level Random Number Generator (RNG) for an FPGA implementation, two main factors are taken into consideration. They are the randomness and its period. Matsumoto and Nishimura [20] proposed a pseudo RNG called Mersenne Twister, which was argued to be as fast and random as the standard ANSI-C "rand()". The Mersenne Twister is essentially a uniformly distributed pseudo RNG based on a matrix linear recurrence over a large finite binary field. Another advantage of FPGA implementation of a Mersenne Twister is its low resource consumption and its capability to generate new random sequences every clock cycle.

5.4 Crossover/Mutation

A simulated binary crossover and a polynomial mutation proposed by Deb and Agarwal [21] are implemented. The intention of a crossover operation is to exchange useful information between two candidate solutions, whereas a mutation operation is aimed to slightly alter a candidate solution. Thus, crossover and mutation can be seen as exploitation and exploration respectively. A balance between them is applied to guide a search algorithm.

5.5 Pareto Front Ranking

Pareto front ranking is based on the non-dominance feature of a candidate solution [11]. Solution A is said to dominate solution B if all of solution A's fitness is better than solution B's, else solution B is non-dominated. Non-dominated solutions are allocated higher rank than dominated ones, hence ensuring the preservation of Pareto fronts.

5.6 Crowding Distance Assignment

A technique proposed by Deb(2001) [11] known as crowding distance assignment is implemented to maintain the diversity of the Pareto fronts. The advantage of this technique lies in the nature by which it operates, whereby it does not require any performance dependent parameter.

5.7 Selection

Tournament selection based on the Pareto front rank and crowding distance is used as a competition winning criteria for the next generation of offspring. Higher ranking solutions wins over lower ranking solutions. If two solutions are of the same rank, the solution with higher crowding distance wins. This method ensures the survival of the fittest.

5.8 Evaluation

The evaluation module is the only module that is application dependent. It would contain the necessary objective functions to be optimized.

6 EXPERIMENTS AND RESULTS

6.1 Test problems

Three different types of test problems were used as a test bench. These three test problems were specifically chosen to analyze the functionality of the algorithm when applied on discrete, discontinuous and constrained problems. The first test problem is SCH1 [22]:

$$\begin{aligned} \text{Minimize } f_1(x) &= x^2, \\ \text{Minimize } f_2(x) &= (x - 2)^2. \end{aligned} \tag{1}$$

This problem has a convex set of Pareto-optimal solution. It is aimed to highlight the diversity of solutions and whether or not they lie on the convex Pareto-optimal front.

The second test problem is SCH2 [22] :

$$\begin{aligned} \text{Minimize } f_1(x) &= \begin{cases} -x & \text{if } x < 1, \\ x - 2 & \text{if } 1 < x \leq 3 \\ 4 - x & \text{if } 3 < x \leq 4 \\ x - 4 & \text{if } x > 4 \end{cases} \\ \text{Minimize } f_2(x) &= (x - 5)^2. \end{aligned} \tag{2}$$

This problem has a Pareto-optimal front consisting of two discontinuous regions. The aim for this test problem is to verify if the algorithm will capture the true Pareto-optimal front and not get trapped in a local front.

The third test problem is a well-studied constrained welded beam design problem [23]:

$$\begin{aligned} \text{Minimize } f_1(x) &= 1.10471h^2l + 0.04811tb(14 + l), \\ \text{Minimize } f_2(x) &= \frac{2.1952}{t^3b}. \end{aligned} \tag{3}$$

$$\begin{aligned} \text{Subjected to } g_1(x) &\equiv 13,600 - \tau(x) \geq 0, \\ g_2(x) &\equiv 30,000 - \sigma(x) \geq 0, \\ g_3(x) &\equiv b - h \geq 0, \\ g_4(x) &\equiv P_c(x) - 6,000 \geq 0, \\ &0.125 \leq h, b \leq 5, \\ &0.1 \leq l, t \leq 10. \end{aligned} \tag{4}$$

where:

$$\begin{aligned}
 \tau(x) &= \sqrt{(\tau')^2 + (\tau'')^2 + \frac{l\tau'\tau''}{\sqrt{0.25(l^2 + (h+t)^2)}}, \\
 \tau' &= \frac{6,000}{\sqrt{2hl}}, \\
 \tau'' &= \frac{6,000(14 + 0.5l)\sqrt{0.25(l^2 + (h+t)^2)}}{2(0.707hl(\frac{l^2}{12} + 0.25(h+t)^2))}, \\
 \sigma(x) &= \frac{504,000}{t^2b}, \\
 P_c(x) &= 64,746.022(1 - 0.0282346t)tb^3.
 \end{aligned} \tag{5}$$

This optimisation problem consists of four design variables and five inequality constraints. It is aimed to test the search algorithm functionality under constraints.

6.2 Experiment Setup

The PC-based implementation for comparison is the NSGA-II [13], which is fundamentally based on an elitist genetic algorithm. NSGA-II was simulated through MATLAB on a Intel(R) Core(TM)2 Duo CPU E8600 @ 3.33GHz, 3.49 GB of RAM. The proposed FPGA-based MDO was simulated on a Xilinx Virtex 4 (xc4vlx200-11ff1513). This section illustrates the FPGA-based MDO and NSGA-II results on the three tests problems. The parameters used for both algorithms and all the test problems are as follows:

- population size = 20,
- number of generations = 100,
- crossover rate = 0.9,
- mutation rate = 0.1.

6.3 Discussion of the Results

Figure 3 and Figure 4 shows the Pareto fronts captured for SCH1 and SCH2 respectively. The proposed FPGA-based MDO was demonstrated to be able to converge to the true Pareto-optimal front and not being trapped in a local front for the discontinuous problem (SCH2). In both problems the FPGA-based MDO found diverse set of solutions, about of the same quality as those found by the NSGA-II.

Figure 5 shows the performance of the algorithms of the constrained welded beam problem. Both algorithms were not able to converge to the true Pareto-optimal front after a low specified number of generations; the true Pareto front is found after 10000 generations of the PC-based NSGA-II. The FPGA-based MDO is able to find 4 solutions closer and better than NSGA-II after 100 generations. The fronts captured by both algorithms are also diverse; no clustering occurred.

Table 1 shows the computation time results for the test problems. The computation time was significantly enhanced, in particular for the welded beam problem where there was a speed improvement of 244 times. The NSGA-II, which is processing sequentially, is problem complexity dependent; as the objectives becomes more complex, the algorithm takes a longer time to process the information. Whereas the FPGA-based MDO search algorithm have no correlation with the problem definition because of its parallel processing, which accounts to the significant speed improvement when faced with the more complex welded beam problem.

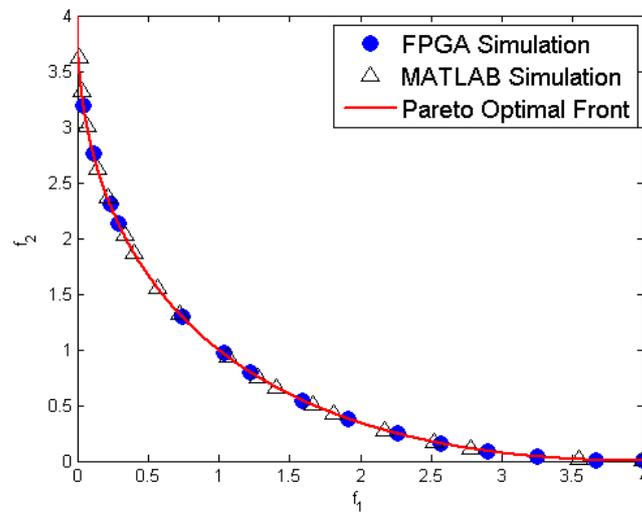


Figure 3: Pareto fronts obtained using MATLAB and FPGA for the SCH1 problem showing both algorithms converging to the true Pareto-optimal front.

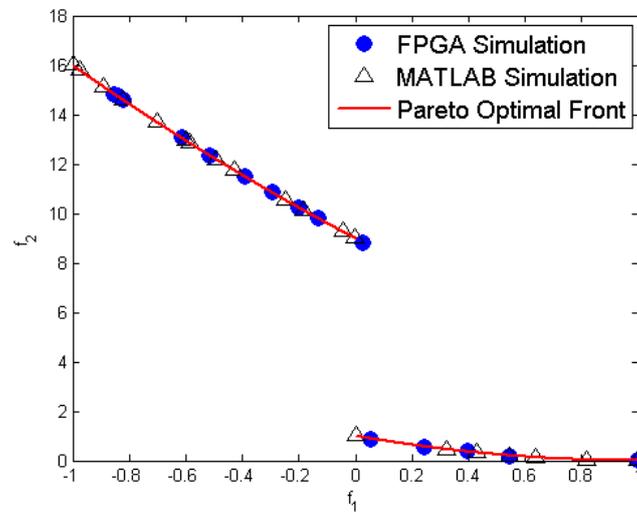


Figure 4: Pareto fronts obtained using MATLAB and FPGA for the SCH2 problem showing both algorithms converging to the true Pareto-optimal front.

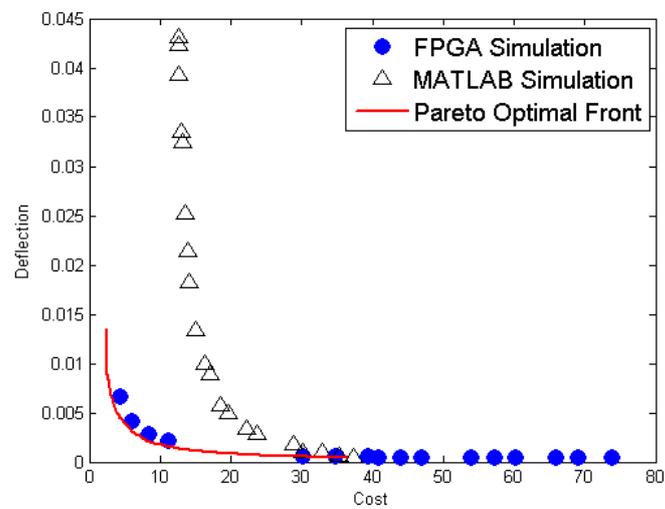


Figure 5: Pareto fronts obtained using MATLAB and FPGA for the welded beam problem showing both algorithms not converging to the true Pareto-optimal front.

Table 1: Computation time results showing the speed improvements

Test Problem	Computation Time (ms)		Speed Improvement (X)
	NSGA-II	FPGA-Based MDO	
SCH1	232	7.542	30
SCH2	354	8.128	43
Welded Beam Problem	1805	7.392	244

7 CONCLUSIONS

We have proposed a computationally fast and effective search method to address the increasingly complex field of Multi-disciplinary Design Optimisation (MDO). On three different types of test problems, which pertained to discrete, discontinuous and constrained problems, the proposed FPGA-based MDO was able perform computationally better than NSGA-II, with speed improvements of over 30 times, and producing results of about similar quality. However, with the welded beam problem, the FPGA-based MDO with its data represented in predefined 32-bit fixed-point widths was not able to attain a wider Pareto-front. In the aeronautics community, this optimisation method can be used on the ground for design problems (aerodynamics, structural etc.) and also practicable for onboard aeronautics application, where equipment size and power consumption are limiting current onboard capabilities.

Future work will look into improving the solution quality by analysing how fixed-point representation and resolution affects the Pareto-front found, especially when objective space of the problem is unknown beforehand. One possible research development is to dynamically alter the width of the fixed-point representation based on the search progress. Additionally, the computation time can be drastically improved by pipelining the flow of FPGA modules.

ACKNOWLEDGEMENTS

Computational resources and services used in this work were provided by the Australian Research Centre for Aerospace Automation (ARCAA) and the High Performance Computing and Research Group of the Queensland University of Technology (QUT), Brisbane, Australia.

REFERENCES

- [1] J. Sobieszczanski-Sobieski and R. T. Haftka, “Multidisciplinary aerospace design optimization: Survey of recent developments,” in *Structural and Multidisciplinary Optimization*, vol. 14, no. 1, 1997, pp. 1–23.
- [2] J. Sobieszczanski-Sobieski, “Multidisciplinary design optimization: An emerging new engineering discipline,” Rio de Janeiro, Brazil, 1993.

- [3] S. Venkataraman and R. T. Haftka, “Structural optimization complexity: What has Moore’s Law done for us?” in *Structural and Multidisciplinary Optimization*, vol. 28, no. 6, 2004, pp. 375–387.
- [4] D. S. Lee, L. F. Gonzalez, J. Periaux, and K. Srinivas, “Robust design optimisation using multi-objective evolutionary algorithms,” *Computers and Fluids*, vol. 37, no. 5, pp. 565–583, June 2008.
- [5] S. D. Scott, A. Samal, and S. Seth, “HGA: A hardware-based genetic algorithm,” in *Proceedings of the 1995 ACM/SIGDA Third International Symposium on Field-Programmable Gate Array*, 1995, pp. 53–59.
- [6] C. Apornthewan and P. Chongstitvatana, “A hardware implementation of the compact genetic algorithm,” in *Proceedings of the 2001 Congress on Evolutionary Computation*, vol. 1, 2001, pp. 624–629.
- [7] B. Shackelford, G. Snider, R. J. Carter, E. Okushi, M. Yasuda, K. Seo, and H. Yasuura, “A high-performance, pipelined, FPGA-based genetic algorithm machine,” *Genetic Programming and Evolvable Machines*, vol. 2, no. 1, pp. 33–60, March 2001.
- [8] W. Tang and L. Yip, “Hardware implementation of genetic algorithms using FPGA,” in *Proceedings of the 2004 47th Midwest Symposium on Circuits and Systems*, 2004, pp. 549–552.
- [9] P. R. Fernando, S. Katkoori, D. Keymeulen, R. Zebulum, and A. Stoica, “Customizable FPGA IP core implementation of a general-purpose genetic algorithm engine,” *IEEE Transactions on Evolutionary Computation*, vol. 14, no. 1, pp. 133–149, 2010.
- [10] J. Kok, L. F. Gonzalez, R. A. Walker, T. Gurnett, and N. A. Kelson, “A synthesizable hardware evolutionary algorithm design for unmanned aerial system real-time path planning,” in *Proceedings of the 2010 Australasian Conference on Robotics and Automation*, Brisbane, Australia, December 2010.
- [11] K. Deb, *Multi-objective optimization using evolutionary algorithms*, 1st ed., ser. Wiley-Interscience series in systems and optimization. John Wiley and Sons, 2001.
- [12] S. Bonissone and R. Subbu, “Evolutionary multiobjective optimization on a chip,” in *Proceedings of the 2007 Workshop on Evolvable and Adaptive Hardware*, Honolulu, Hawaii, September 2007, pp. 61–66.
- [13] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: NSGA-II,” *IEEE Transaction on Evolutionary Computation*, vol. 6, no. 2, April 2002.

- [14] A. Osyczka and S. Kundu, "A new method to solve generalized multicriteria optimization problems using the simple genetic algorithm," *Structural Optimization*, vol. 10, no. 2, pp. 94–99, 1995.
- [15] E. Zitzler and L. Thiele, "An evolutionary algorithm for multiobjective optimization: The strength pareto approach," Zurich, Switzerland, 1998.
- [16] H. Kita, Y. Yabumoto, N. Mori, and Y. Nishikawa, "Multi-objective optimization by means of thermodynamical genetic algorithm," in *Proceedings of Parallel Problem Solving from Nature IV (PPSN-IV)*, 1996, pp. 504–512.
- [17] A. the non-dominated front using the Pareto archived evolution strategy, "Evolutionary computation journal," *Computers and Fluids*, vol. 8, no. 2, pp. 149–172, 2000.
- [18] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989.
- [19] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, 1992.
- [20] M. Matsumoto and T. Nishimura, "Mersenne twister: A 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Transactions on Modeling and Computer Simulation*, vol. 8, no. 1, pp. 3–30, January 1998.
- [21] K. Deb and R. B. Agarwal, "Simulated binary crossover for continuous search space," *Complex Systems*, vol. 9, pp. 115–148, April 1995.
- [22] J. D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms," in *Proceedings of the First International Conference on Genetic Algorithms*, 1985, pp. 93–100.
- [23] G. V. Reklaitis, A. Ravindran, and K. M. Ragsdell, *Engineering Optimization Methods and Applications*. New York: Wiley, 1983.

EFFICIENT FORCE FIELD CALCULATION IN ARTICULATED MULTISCALE MOLECULAR SIMULATIONS

MOHAMMAD POURSIINA* AND KURT S. ANDERSON*

*Computational Dynamics Lab (CDL)
Department of Mechanical, Aerospace and Nuclear Engineering
Rensselaer Polytechnic Institute
Troy, NY 12180, USA
e-mail: poursm@rpi.edu, anderk5@rpi.edu

Key words: Force Field, Pseudo-center, Pseudo-inertia Tensor, Multiscale Problems

Abstract. Multiscale simulations of molecular systems such as proteins, DNAs, and RNAs are implemented using models with different resolutions ranging from a fully atomistic model to coarse-grained molecules, up to a continuum level system descriptions. For such simulations, pairwise force calculation is a serious bottleneck which can impose a prohibitive amount of computational load on the simulation if not performed wisely. Herein, a novel method is presented to approximate the resultant force and the associated moment for long-range particle-body and body-body interactions applicable to multiresolution coarse-grained simulation of biopolymers. The resultant moment is due to the fact that the net force does not necessarily act through the center of mass of the body. This moment is neglected in bead-based coarse models which use particle dynamics to form the equations of motion of each large spherical pseudo-atom. The presented method significantly reduces the cost of the force field calculations specially in multiscale models which contain rigid subdomains.

1 INTRODUCTION

Function of most biopolymers such as DNA and RNA is highly related to their structure. As such, modeling and simulation of such systems are performed to predict the associated structure accurately in a timely manner. Such systems with large number of particles ($n \approx 10^6$ [1]) suffer from the cumbersome pairwise force field calculations with the computational complexity of $O(n^2)$ at each time step. Therefore, different algorithms have been developed to efficiently calculate far-field interactions. Ewald summation method applicable to periodic systems [2, 3] is the approach which calculates the long-range field in Fourier space. For non-periodic systems, particle-mesh (PM) and particle-particle particle-mesh (PPPM) techniques perform satisfactorily when the required precision is low and the particles are distributed more or less uniformly [4]-[6]. As the next generation of the fast algorithms, the first tree methods

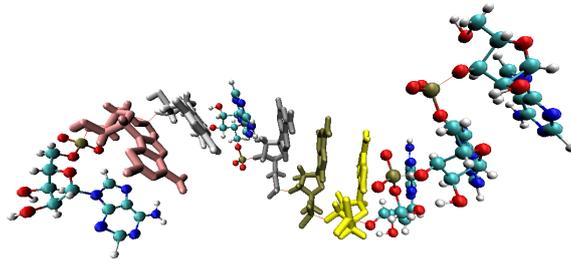


Figure 1: Multiscale modeling of an RNA

[4] presented in [7] and [8] provide the monopole approximation of the long-range field in the system. Using higher order multipole expansion, Fast Multipole Method (FMM) is presented in [9] to approximate the electrostatic field of a group of particles with a computational complexity of $O(n)$. This method is later extended and used to approximate different potential fields for different systems [10]-[13].

Eliminating high frequency modes of motion by considering groups of atoms as rigid or flexible bodies, coarse-grained simulations of biopolymers improve the integration time step. Additionally, they reduce the computational cost associated with pairwise force calculations by ignoring the interactions among particles embedded in rigid regions of the system. This technique may be implemented by replacing different subdomains of the system with rigid or flexible bodies. Using internal coordinates, these regions are connected together via kinematic joints [14]-[16]. Alternatively, coarse-grained modeling may be implemented by considering a group of atoms as a spherical bead, and applying particle dynamics to formulate the equations of motion of each large pseudo-atom [17, 18]. Such systems may also be modeled in the multiscale and/or adaptive [16] framework to capture the dynamics of the system more accurately. These models may contain multiple resolutions ranging from fine scale atomistic domains, coarse-grained macromolecules to continuum level system descriptions.

Herein, we extend the approach used to approximate long-range gravitational force and the associated moment in spacecraft dynamics [19] to the pairwise forces in articulated multiscale simulation of molecular systems. We provide an approximation of the resultant force and the associated moment applied from a particle to a body (pseudo-atom) consisting of many particles. The resultant moment approximated here is due to the fact that the net force does not necessarily act through the center of mass of the body (pseudo-atom). This moment should be considered when the equations of motion are formed in an articulated multibody framework. However, it is neglected in most coarse-grained bead models which use particle dynamics. These approximations can be used to find the long-range interactions between the atoms and rigidified nucleotides of an RNA shown in Fig. 1. We also derive the approximation for the spatial force due to the interactions between particles embedded in two pseudo-atoms (bodies). For the same system shown in Fig. 1, this approximation is applied to find the far-field interactions among the rigidified nucleotides. Herein, multipole and Taylor series expansion coefficients are elaborated and as such, expressed using terms with physical meanings. For instance, we define pseudo-

center and pseudo-inertia tensor for each body. Then, we express the approximations in terms of the location of the pseudo-center and specific elements of the pseudo-inertia tensor of each body. This tensor is calculated for each *rigid* subdomain of the system before starting the simulation. We also show that if the pseudo-center is defined for the subdomain and is selected as the origin of the body-fixed frame, the first (dipole) moment disappears in the force approximation.

2 FORCE ON A SMALL BODY FROM A PARTICLE

Consider particle \bar{P} , and body B (not necessarily a rigid body) containing N particles shown in Fig. 2. Any arbitrary particle P_i belonging to B experiences a pairwise force due to its interaction with \bar{P} . Herein, we assume that this force is expressed by the following relation

$$\mathbf{F}_{\bar{P}P_i} = \frac{\kappa \bar{\lambda} \lambda_i}{(|\mathbf{r}'_i|)^s} \mathbf{e}_{\mathbf{r}'_i} = \kappa \bar{\lambda} \lambda_i \mathbf{r}'_i (\mathbf{r}'_i^2)^{-\frac{s+1}{2}}, \quad (1)$$

where, λ_i and $\bar{\lambda}$ are the quantities corresponding to the force field, κ is the constant associated with the force field of interest, and s is an integer. In this formula, $\mathbf{e}_{\mathbf{r}'_i}$ is the unit vector from particle \bar{P} to P_i . This general formulation may be used to address the gravitational, Coulombic or London forces. For instance, if one is interested in pairwise forces due to the Coulomb's law, κ is replaced by the Coulomb force constant, λ_i represents the charge of the particle, and s becomes 2. For each body, we define the lumped quantity corresponding to the quantity of interest λ as

$$\Lambda \triangleq \sum_{i=1}^N \lambda_i. \quad (2)$$

We also define the ‘‘pseudo-center’’ of the body denoted by C_λ . Position vector of this point with respect to the center of mass (i.e., B^*) is determined by the following relation

$$\mathbf{R}_\lambda = \frac{\sum_{i=1}^N \mathbf{R}_i \lambda_i}{\Lambda}, \quad (3)$$

provided that $\Lambda \neq 0$. In the above relation, \mathbf{R}_i is the position vector of the particle (atom) P_i measured from the center of mass of the body. In the following derivations, it is assumed that the origin of the body-fixed frame coincides with the pseudo-center of the body.

Using the following relation based on Fig.2,

$$\mathbf{r}'_i = \mathbf{R} + \mathbf{r}_i, \quad (4)$$

Eq. (1) is rewritten as

$$\mathbf{F}_{\bar{P}P_i} = \kappa \bar{\lambda} \lambda_i (\mathbf{R} + \mathbf{r}_i) (\mathbf{R}^2 + \mathbf{r}_i^2 + 2\mathbf{R} \cdot \mathbf{r}_i)^{-\frac{s+1}{2}}. \quad (5)$$

Introducing the intermediate vector \mathbf{q}_i as

$$\mathbf{q}_i = \frac{\mathbf{r}_i}{R}, \quad (6)$$

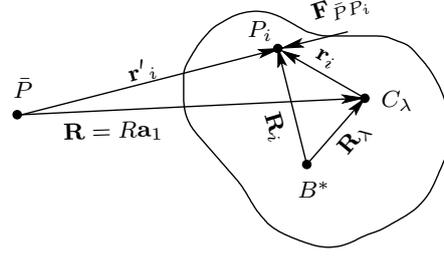


Figure 2: Particle \bar{P} applies the force $\mathbf{F}_{\bar{P}P_i}$ to the particle P_i on body B .

and expressing vector \mathbf{R} as

$$\mathbf{R} = R\mathbf{a}_1, \quad (7)$$

one can rewrite Eq. (5) in the following format

$$\mathbf{F}_{\bar{P}P_i} = \frac{\kappa\bar{\lambda}\lambda_i}{R^s}(\mathbf{a}_1 + \mathbf{q}_i)(1 + \mathbf{q}_i^2 + 2\mathbf{a}_1 \cdot \mathbf{q}_i)^{-\frac{s+1}{2}}. \quad (8)$$

Recalling the binomial series expansion,

$$(1 + x)^n = 1 + nx + \frac{n(n-1)}{2}x^2 + \frac{n(n-1)(n-2)}{3!}x^3 + \dots \quad (|x| < 1), \quad (9)$$

the term $(1 + \mathbf{q}_i^2 + 2\mathbf{a}_1 \cdot \mathbf{q}_i)^{-\frac{s+1}{2}}$ in Eq. (8) is expanded as follows if $|\mathbf{q}_i^2 + 2\mathbf{a}_1 \cdot \mathbf{q}_i| < 1$

$$\begin{aligned} & (1 + \mathbf{q}_i^2 + 2\mathbf{a}_1 \cdot \mathbf{q}_i)^{-\frac{s+1}{2}} \\ &= 1 - \frac{s+1}{2}(\mathbf{q}_i^2 + 2\mathbf{a}_1 \cdot \mathbf{q}_i) + \frac{(s+1)(s+3)}{8}(\mathbf{q}_i^2 + 2\mathbf{a}_1 \cdot \mathbf{q}_i)^2 \\ & - \frac{(s+1)(s+3)(s+5)}{48}(\mathbf{q}_i^2 + 2\mathbf{a}_1 \cdot \mathbf{q}_i)^3 + \dots \end{aligned} \quad (10)$$

Using the expression provided in Eq. (10), the total force experienced by body B due to the pairwise interactions between its own particles and \bar{P} is rewritten as

$$\begin{aligned} \mathbf{F}_{\bar{P}B} &= \frac{\kappa\bar{\lambda}}{R^s} \left[\sum_{i=1}^N \lambda_i \mathbf{a}_1 - \sum_{i=1}^N (s+1)\lambda_i (\mathbf{a}_1 \cdot \mathbf{q}_i) \mathbf{a}_1 + \sum_{i=1}^N \lambda_i \mathbf{q}_i \right. \\ & - \sum_{i=1}^N \frac{(s+1)}{2} \lambda_i \mathbf{q}_i^2 \mathbf{a}_1 + \sum_{i=1}^N \frac{(s+1)(s+3)}{2} \lambda_i (\mathbf{a}_1 \cdot \mathbf{q}_i)^2 \mathbf{a}_1 \\ & \left. - \sum_{i=1}^N (s+1)\lambda_i (\mathbf{a}_1 \cdot \mathbf{q}_i) \mathbf{q}_i + \dots \right]. \end{aligned} \quad (11)$$

We elaborate on different terms in the above expression based on their orders with respect to $\frac{|\mathbf{r}_i|}{R}$.

- $O\left(\frac{|\mathbf{r}_i|}{R}\right)^0$:

$$\frac{\kappa\bar{\lambda}}{R^s} \sum_{i=1}^N \lambda_i \mathbf{a}_1 = \frac{\kappa\bar{\lambda}}{R^s} \mathbf{a}_1 \sum_{i=1}^N \lambda_i \stackrel{(2)}{=} \frac{\kappa\bar{\lambda}\Lambda}{R^s} \mathbf{a}_1 \quad (12)$$

- $O\left(\frac{|\mathbf{r}_i|}{R}\right)$:

$$- \frac{\kappa\bar{\lambda}(s+1)}{R^{(s+1)}} \mathbf{a}_1 \mathbf{a}_1 \cdot \sum_{i=1}^N \lambda_i \mathbf{r}_i + \frac{\kappa\bar{\lambda}}{R^{(s+1)}} \sum_{i=1}^N \lambda_i \mathbf{r}_i \quad (13)$$

Although it is mentioned in [4] that the dipole moment measured from the center of charge (pseudo-center in this paper) vanishes if “the charges are all positive” [4] ($\lambda_i > 0$), in the following we show that the dipole (first) moment measured from the pseudo-center (if defined) becomes zero independent of the sign of λ_i

$$\sum_{i=1}^N \lambda_i \mathbf{r}_i = \sum_{i=1}^N \lambda_i (\mathbf{R}_i - \mathbf{R}_\lambda) = \sum_{i=1}^N \lambda_i \mathbf{R}_i - \mathbf{R}_\lambda \sum_{i=1}^N \lambda_i \stackrel{(2)}{=} \Lambda \mathbf{R}_\lambda - \Lambda \mathbf{R}_\lambda = \mathbf{0}. \quad (14)$$

Consequently, the first order terms disappear if the pseudo-center is defined as the origin of the body-fixed frame.

- $O\left(\frac{|\mathbf{r}_i|}{R}\right)^2$:

$$\begin{aligned} \frac{\kappa\bar{\lambda}}{R^{(s+2)}} \left[-\frac{(s+1)}{2} \mathbf{a}_1 \sum_{i=1}^N \lambda_i \mathbf{r}_i^2 + \frac{(s+1)(s+3)}{2} \mathbf{a}_1 \mathbf{a}_1 \cdot \sum_{i=1}^N \lambda_i \mathbf{r}_i \mathbf{r}_i \cdot \mathbf{a}_1 \right. \\ \left. - (s+1) \mathbf{a}_1 \cdot \sum_{i=1}^N \lambda_i \mathbf{r}_i \mathbf{r}_i \right]. \quad (15) \end{aligned}$$

To simplify the above expression, we define the “pseudo-inertia” tensor for body B associated with the quantity λ with respect to its own pseudo-center as

$$\mathcal{I}_\lambda^{B/C_\lambda} \triangleq \sum_{i=1}^N (\mathcal{U} \mathbf{r}_i^2 - \mathbf{r}_i \mathbf{r}_i) \lambda_i, \quad (16)$$

where \mathcal{U} denotes the identity tensor. This tensor represents the tensor of the moment of inertia (second moment) of the body if one studies the gravitational force [19]. For the Coulombic force, this tensor turns out to be the quadrupole moment tensor [20]. Defining the trace of the pseudo-inertia tensor as

$$\text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}) = 2 \sum_{i=1}^N \lambda_i \mathbf{r}_i^2 \implies \sum_{i=1}^N \lambda_i \mathbf{r}_i^2 = \frac{1}{2} \text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}), \quad (17)$$

and using the following property

$$\sum_{i=1}^N \lambda_i \mathbf{r}_i \mathbf{r}_i = \frac{\text{tr}(\mathcal{I}_\lambda^{B/C_\lambda})}{2} \mathcal{U} - \mathcal{I}_\lambda^{B/C_\lambda}, \quad (18)$$

the second order terms are simplified as

$$\begin{aligned} & \frac{\kappa \bar{\lambda}}{R^{(s+2)}} \left[-\frac{(s+1)}{4} \text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}) \mathbf{a}_1 + \frac{(s+1)(s+3)}{2} \mathbf{a}_1 \mathbf{a}_1 \cdot \left(\frac{\text{tr}(\mathcal{I}_\lambda^{B/C_\lambda})}{2} \mathcal{U} - \mathcal{I}_\lambda^{B/C_\lambda} \right) \cdot \mathbf{a}_1 \right. \\ & \left. - (s+1) \mathbf{a}_1 \cdot \left(\frac{\text{tr}(\mathcal{I}_\lambda^{B/C_\lambda})}{2} \mathcal{U} - \mathcal{I}_\lambda^{B/C_\lambda} \right) \right]. \end{aligned} \quad (19)$$

Substituting Eqns. (12), (13), and (19) into Eq. (11), the net force applied to the body is expressed as

$$\begin{aligned} \mathbf{F}_{\bar{P}B} &= \frac{\kappa \bar{\lambda}}{R^s} \left\{ \Lambda \mathbf{a}_1 + \left[\frac{s(s+1)}{4R^2} \text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}) - \frac{(s+1)(s+3)}{2R^2} \mathbf{a}_1 \cdot \mathcal{I}_\lambda^{B/C_\lambda} \cdot \mathbf{a}_1 \right] \mathbf{a}_1 \right. \\ & \left. + \frac{s+1}{R^2} \mathbf{a}_1 \cdot \mathcal{I}_\lambda^{B/C_\lambda} + O\left(\frac{\mathbf{r}}{R}\right)^3 \right\}, \end{aligned} \quad (20)$$

where \mathbf{r} is the position vector of a generic point on B with respect to the pseudo-center of the body. If the particle \bar{P} is far enough from the pseudo-center of the body such that the largest distance from the pseudo-center to the particles embedded in body B is smaller than the distance between \bar{P} and C_λ , i.e., $\max_{i \in B} |\mathbf{r}_i| \ll R$, one can ignore the third and higher order terms in Eq. (20). Additionally, introducing a dextral, orthogonal set of unit vectors, \mathbf{a}_1 , \mathbf{a}_2 , and \mathbf{a}_3 , and defining the elements of the pseudo-inertia matrix in \mathbf{a} -basis as

$$I_{ij} = \mathbf{a}_i \cdot \mathcal{I}_\lambda^{B/C_\lambda} \cdot \mathbf{a}_j \quad (i, j = 1, 2, 3), \quad (21)$$

one arrives at the following approximation for the resultant force

$$\begin{aligned} \tilde{\mathbf{F}}_{\bar{P}B} &= \frac{\kappa \bar{\lambda}}{R^s} \left\{ \Lambda \mathbf{a}_1 + \left[\frac{s(s+1)}{4R^2} \text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}) - \frac{(s+1)(s+3)}{2R^2} I_{11} \right] \mathbf{a}_1 \right. \\ & \left. + \frac{s+1}{R^2} (I_{11} \mathbf{a}_1 + I_{21} \mathbf{a}_2 + I_{31} \mathbf{a}_3) \right\}. \end{aligned} \quad (22)$$

The above equation represents the second order multipole approximation of the force applied from particle \bar{P} to body B . The first term shows the interaction between the particle and the body in which the whole body is treated as a particle located at its pseudo-center with the lumped quantity Λ . Since the origin of the body-fixed frame is located at its pseudo-center, the first moment does not appear in the approximate force. Further, the trace of the pseudo-inertia tensor which appears in this expression is an invariant quantity for *rigid* superatoms. In other words, it does not change when the orientation and location of the rigid body changes within the

course of the simulation. As such, once this term is calculated, there is no need to update it. The pseudo-inertia tensor is also a constant matrix for the rigid body if expressed in its body-basis. Therefore, if this tensor is calculated for a rigid subdomain of the system at some time either before or during the simulation, there is no additional cost associated with forming or using this dyadic during the course of the simulation. It is only necessary to monitor the superatom location and orientation to update this tensor at each time step.

Since the pseudo-center is not defined (see Eq. (3)) for a body with zero lumped quantity Λ , the center of mass of the body (subdomain) is considered as the origin of the body-fixed frame to derive the above equations. As such, the pseudo-inertia tensor used in Eq. (22) is defined about B^* . Since the origin of the body-fixed frame is located at the center of mass of the body rather than the pseudo-center, unlike the previous derivations, the first moment is not zero. For such a system, using the relation $\sum_{\substack{i=1 \\ i \neq j}}^N \lambda_i = -\lambda_j$ from Eq. (2), one can rewrite the first moment measured from the center of mass of the body as

$$\sum_{i=1}^N \lambda_i \mathbf{R}_i = \sum_{\substack{i=1 \\ i \neq j}}^N (\lambda_i \mathbf{R}_i) + \lambda_j \mathbf{R}_j = \sum_{\substack{i=1 \\ i \neq j}}^N \lambda_i \mathbf{R}_i - \mathbf{R}_j \sum_{\substack{i=1 \\ i \neq j}}^N \lambda_i = \sum_{\substack{i=1 \\ i \neq j}}^N \lambda_i (\mathbf{R}_i - \mathbf{R}_j), \quad (23)$$

which is effectively the first moment measured from the reference point j . Since j is an arbitrary point of the body, this relation demonstrates that if the pseudo-center is not defined for the subdomain, the first moment is constant regardless of the choice of the origin. Consequently, when the lumped quantity of the pseudo-atom is zero, it is necessary to consider its origin-independent first moment in Eq. (22). Further, this term becomes time-invariant for rigid subdomains of the system.

3 FORCE ON A SMALL BODY FROM A SMALL BODY

Consider bodies B and \bar{B} shown in Fig. 3 containing N and \bar{N} particles, respectively. Each body experiences a resultant force due to the pairwise interactions between particles $\{P_i\}_{i=1}^N$ belonging to B , and $\{\bar{P}_j\}_{j=1}^{\bar{N}}$ embedded in \bar{B} . Summing over the low order approximation of the resultant forces applied to B by particles \bar{P}_j on \bar{B} from Eq. (22), one arrives at the approximate net force applied to body B from \bar{B} as

$$\begin{aligned} \tilde{\mathbf{F}}_{\bar{B}B} &= \sum_{j=1}^{\bar{N}} \frac{\kappa \Lambda \lambda_j}{\bar{R}_j^s} \bar{\mathbf{a}}_1^j + \frac{\kappa \lambda_j}{\bar{R}_j^s} \sum_{j=1}^{\bar{N}} \left\{ \left[\frac{s(s+1)}{4\bar{R}_j^2} \text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}) \right. \right. \\ &\quad \left. \left. - \frac{(s+1)(s+3)}{2\bar{R}_j^2} \bar{\mathbf{a}}_1^j \cdot \mathcal{I}_\lambda^{B/C_\lambda} \cdot \bar{\mathbf{a}}_1^j \right] \bar{\mathbf{a}}_1^j + \frac{s+1}{\bar{R}_j^2} \bar{\mathbf{a}}_1^j \cdot \mathcal{I}_\lambda^{B/C_\lambda} \right\}, \quad (24) \end{aligned}$$

where \bar{R}_j denotes the distance from \bar{P}_j to the pseudo-center of body B , and $\bar{\mathbf{a}}_1^j$ is the corresponding unit vector.

We initially elaborate on the first summation in the above equation. Based on the geometry

~

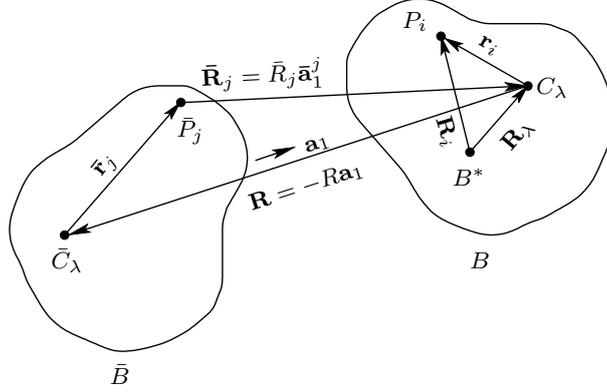


Figure 3: Body \bar{B} applies a resultant force to body B due to the pairwise interaction between the particles belonging to different bodies.

shown in Fig. 3, the term $\frac{\bar{\mathbf{a}}_1^j}{\bar{R}_j^s}$ can be replaced by the following relation

$$\frac{\bar{\mathbf{a}}_1^j}{\bar{R}_j^s} = \bar{\mathbf{R}}_j (\bar{\mathbf{R}}_j^2)^{-\frac{s+1}{2}} = -(\mathbf{R} + \bar{\mathbf{r}}_j) (\mathbf{R}^2 + \bar{\mathbf{r}}_j^2 + 2\mathbf{R} \cdot \bar{\mathbf{r}}_j)^{-\frac{s+1}{2}}, \quad (25)$$

where \mathbf{R} is the position vector of the pseudo-center of body \bar{B} relative to the pseudo-center of body B which is written as

$$\mathbf{R} = -R\mathbf{a}_1. \quad (26)$$

Introducing the intermediate vector $\bar{\mathbf{q}}_j = -\frac{\bar{\mathbf{r}}_j}{R}$, Eq. (25) becomes

$$\frac{\bar{\mathbf{a}}_1^j}{\bar{R}_j^s} = \frac{1}{R^s} (\mathbf{a}_1 + \bar{\mathbf{q}}_j) (1 + \bar{\mathbf{q}}_j^2 + 2\mathbf{a}_1 \cdot \bar{\mathbf{q}}_j)^{-\frac{s+1}{2}}. \quad (27)$$

We replace the first summation of Eq. (24) with the above expression and expand the term $(1 + \bar{\mathbf{q}}_j^2 + 2\mathbf{a}_1 \cdot \bar{\mathbf{q}}_j)^{-\frac{s+1}{2}}$ using the binomial expansion described in section 2. The zeroth order terms in $\frac{\bar{\mathbf{r}}_j}{R}$ are treated the same as those provided in Eq. (12). Similar to the results provided for Eq. (13), the first order terms in $\frac{\bar{\mathbf{r}}_j}{R}$ disappear since the pseudo-center coincides with the origin of the body-fixed frame. The second order terms are even functions of $\bar{\mathbf{q}}_j = -\frac{\bar{\mathbf{r}}_j}{R}$. As such, the negative sign used in the definition of $\bar{\mathbf{q}}_j$ disappears. Therefore, this vector in the second order terms is treated exactly the same as \mathbf{q}_i in Eq. (15). Finally, the first term of Eq. (24) is approximated by

$$\begin{aligned} \sum_{j=1}^{\bar{N}} \frac{\kappa\Lambda\lambda_j}{\bar{R}_j^s} \bar{\mathbf{a}}_1^j &\approx \frac{\kappa\Lambda\bar{\Lambda}}{R^s} \mathbf{a}_1 + \frac{\kappa\Lambda}{R^s} \left\{ \left[\frac{s(s+1)}{4R^2} \text{tr}(\bar{\mathcal{I}}_{\lambda}^{\bar{B}/\bar{C}_{\lambda}}) \right. \right. \\ &\quad \left. \left. - \frac{(s+1)(s+3)}{2R^2} \mathbf{a}_1 \cdot \bar{\mathcal{I}}_{\lambda}^{\bar{B}/\bar{C}_{\lambda}} \cdot \mathbf{a}_1 \right] \mathbf{a}_1 + \frac{s+1}{R^2} \mathbf{a}_1 \cdot \bar{\mathcal{I}}_{\lambda}^{\bar{B}/\bar{C}_{\lambda}} \right\}, \end{aligned} \quad (28)$$

where $\bar{\mathcal{I}}_\lambda^{\bar{B}/\bar{C}_\lambda}$ is the pseudo-inertia tensor of body \bar{B} about its pseudo-center, and the lumped quantity of body \bar{B} is defined as $\bar{\Lambda} \triangleq \sum_{j=1}^N \bar{\lambda}_j$.

In the second summation of Eq. (24), \bar{R} and $\bar{\mathbf{a}}_1^j$ may be replaced by R and \mathbf{a}_1 , respectively, since this summation involves the terms of second or higher degree in $|\mathbf{r}_i|$, so that no terms of interest for the purpose at hand are lost through this replacement. Defining the elements of the pseudo-inertia tensors of the bodies in a-basis, this relation is expressed in terms of the elements of the pseudo-inertia tensor as

$$\begin{aligned} \tilde{\mathbf{F}}_{\bar{B}B} &= \frac{\kappa\Lambda\bar{\Lambda}}{R^s}\mathbf{a}_1 + \frac{\kappa\Lambda}{R^s}\left\{\left[\frac{s(s+1)}{4R^2}\text{tr}(\bar{\mathcal{I}}_\lambda^{\bar{B}/\bar{C}_\lambda}) - \frac{(s+1)(s+3)}{2R^2}\bar{I}_{11}\right]\mathbf{a}_1\right. \\ &+ \left.\frac{s+1}{R^2}(\bar{I}_{11}\mathbf{a}_1 + \bar{I}_{21}\mathbf{a}_2 + \bar{I}_{31}\mathbf{a}_3)\right\} + \frac{\kappa\bar{\Lambda}}{R^s}\left\{\left[\frac{s(s+1)}{4R^2}\text{tr}(\mathcal{I}_\lambda^{B/C_\lambda})\right.\right. \\ &\left.\left.- \frac{(s+1)(s+3)}{2R^2}I_{11}\right]\mathbf{a}_1 + \frac{s+1}{R^2}(I_{11}\mathbf{a}_1 + I_{21}\mathbf{a}_2 + I_{31}\mathbf{a}_3)\right\}. \end{aligned} \quad (29)$$

The approximation provided in Eq. (29) is valid if the distance between the pseudo-centers of bodies B and \bar{B} is greater than the distance between the pseudo-center of each body and the corresponding farthest particle, i.e., $\max_{i \in B} |\mathbf{r}_i| \ll R$ and $\max_{j \in \bar{B}} |\bar{\mathbf{r}}_j| \ll R$. The symmetry observed in the net force properly implies that the provided approximation does not violate Newton's third law of motion.

4 MOMENT ON A SMALL BODY FROM A PARTICLE

Since the resultant force calculated in section 2 does not generally act through the center of mass of the body, it creates a moment about B^* . Based on the geometry shown in Fig. 2, the resultant moment about B^* due to the interaction between \bar{P} and P_i is expressed by the following cross product

$$\mathbf{M}_{\bar{P}P_i}^{B^*} = \mathbf{R}_i \times \mathbf{F}_{\bar{P}P_i}. \quad (30)$$

According to Fig. 2, and knowing that both \mathbf{r}'_i and $\mathbf{F}_{\bar{P}P_i}$ are collinear vectors, this moment becomes

$$\mathbf{M}_{\bar{P}P_i}^{B^*} = (\mathbf{R}_\lambda - \mathbf{R}) \times \mathbf{F}_{\bar{P}P_i} + \underbrace{\mathbf{r}'_i \times \mathbf{F}_{\bar{P}P_i}}_{\mathbf{0}}. \quad (31)$$

As such, body B experiences the following moment about its center of mass due to the interactions between its own particles and \bar{P}

$$\mathbf{M}_{\bar{P}B}^{B^*} = \sum_{i=1}^N (\mathbf{R}_\lambda - \mathbf{R}) \times \mathbf{F}_{\bar{P}P_i} = (\mathbf{R}_\lambda - \mathbf{R}) \times \sum_{i=1}^N \mathbf{F}_{\bar{P}P_i} = (\mathbf{R}_\lambda - \mathbf{R}) \times \mathbf{F}_{\bar{P}B}. \quad (32)$$

Substituting \mathbf{R} from Eq. (7) into the above equation, using the approximate force from Eq. (20), and defining the elements of the pseudo-inertia tensor in a-basis, the approximate moment about

the center of mass of the body is expressed as

$$\begin{aligned}\tilde{\mathbf{M}}_{PB}^{B*} &= \frac{\kappa\bar{\lambda}}{R^s}\mathbf{R}_\lambda \times \left\{ \Lambda\mathbf{a}_1 + \left[\frac{s(s+1)}{4R^2}\text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}) - \frac{(s+1)(s+3)}{2R^2}I_{11} \right] \mathbf{a}_1 \right. \\ &\quad \left. + \frac{s+1}{R^2}(I_{11}\mathbf{a}_1 + I_{21}\mathbf{a}_2 + I_{31}\mathbf{a}_3) \right\} - \frac{\kappa\bar{\lambda}(s+1)}{R^{(s+1)}}(I_{21}\mathbf{a}_3 - I_{31}\mathbf{a}_2).\end{aligned}\quad (33)$$

In the above equation, if the pseudo-center and the center of mass of the body coincide, i.e., $\mathbf{R}_\lambda = \mathbf{0}$, the last term only contributes to the moment applied to body B . In this case, if \mathbf{a}_1 is aligned with one of the principal axes of the pseudo-inertia tensor of body B , the approximate moment formulation provides a zero value. As such, the analyst may use the exact calculations or higher order approximation to find the moment.

5 MOMENT ON A SMALL BODY FROM A SMALL BODY

Since the resultant force applied to body B from body \bar{B} does not necessarily pass through the center of mass of body B , it creates a moment about B^* . As such, the resultant moment can be approximated as a summation over the approximate moments applied by each particle \bar{P}_j on \bar{B} to body B as

$$\begin{aligned}\tilde{\mathbf{M}}_{BB}^{B*} &= \sum_{j=1}^{\bar{N}} \frac{\kappa\lambda_j\Lambda}{R_j^s}\mathbf{R}_\lambda \times \bar{\mathbf{a}}_1^j + \sum_{j=1}^{\bar{N}} \frac{\kappa\lambda_j}{R_j^s}\mathbf{R}_\lambda \times \left\{ \left[\frac{s(s+1)}{4R_j^2}\text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}) \right. \right. \\ &\quad \left. \left. - \frac{(s+1)(s+3)}{2R_j^2}\bar{\mathbf{a}}_1^j \cdot \mathcal{I}_\lambda^{B/C_\lambda} \cdot \bar{\mathbf{a}}_1^j \right] \bar{\mathbf{a}}_1^j + \frac{s+1}{R_j^2}\bar{\mathbf{a}}_1^j \cdot \mathcal{I}_\lambda^{B/C_\lambda} \right\} \\ &\quad - \sum_{j=1}^{\bar{N}} \frac{\kappa\lambda_j(s+1)}{R_j^{(s+1)}}\bar{\mathbf{a}}_1^j \times (\bar{\mathbf{a}}_1^j \cdot \mathcal{I}_\lambda^{B/C_\lambda}).\end{aligned}\quad (34)$$

The first summation can be expanded exactly the same as the manipulation conducted in section 3. Similarly, since the terms in the second and third summations are second or higher orders in $|\bar{\mathbf{r}}_j|$, we may replace \bar{R} and $\bar{\mathbf{a}}_1^j$ by R and \mathbf{a}_1 , respectively. Defining the elements of the pseudo-inertia tensors of the bodies in \mathbf{a} -basis, the approximate moment about the center of mass of body B from \bar{B} due to the pairwise interactions between the particles embedded in B and \bar{B} is expressed as

$$\begin{aligned}\tilde{\mathbf{M}}_{BB}^{B*} &= \frac{\kappa\Lambda\bar{\Lambda}}{R^s}\mathbf{R}_\lambda \times \mathbf{a}_1 \\ &\quad + \frac{\kappa\Lambda}{R^s}\mathbf{R}_\lambda \times \left\{ \left[\frac{s(s+1)}{4R^2}\text{tr}(\bar{\mathcal{I}}_\lambda^{\bar{B}/\bar{C}_\lambda}) - \frac{(s+1)(s+3)}{2R^2}\bar{I}_{11} \right] \mathbf{a}_1 + \frac{s+1}{R^2}(\bar{I}_{11}\mathbf{a}_1 + \bar{I}_{21}\mathbf{a}_2 + \bar{I}_{31}\mathbf{a}_3) \right\} \\ &\quad + \frac{\kappa\bar{\Lambda}}{R^s}\mathbf{R}_\lambda \times \left\{ \left[\frac{s(s+1)}{4R^2}\text{tr}(\mathcal{I}_\lambda^{B/C_\lambda}) - \frac{(s+1)(s+3)}{2R^2}I_{11} \right] \mathbf{a}_1 + \frac{s+1}{R^2}(I_{11}\mathbf{a}_1 + I_{21}\mathbf{a}_2 + I_{31}\mathbf{a}_3) \right\} \\ &\quad - \frac{\kappa\bar{\Lambda}(s+1)}{R^{(s+1)}}(I_{21}\mathbf{a}_3 - I_{31}\mathbf{a}_2).\end{aligned}\quad (35)$$

In the above equation, if the pseudo-center and center of mass of body B coincide, only the last term survives and contributes to the moment. In this case, if \mathbf{a}_1 is aligned with one of the principal axes of the pseudo-inertia tensor of body B , the moment becomes zero. Consequently, either the exact moment calculation may be used to find the resultant moment or the moment is approximated considering higher order terms.

6 CONCLUSIONS

A novel method to approximate the resultant force and moment due to the long-range particle-body and body-body interactions in Cartesian coordinates has been presented. The effective moment is due to the fact the resultant force does not necessarily act through the center of mass of the body. In this low order multipole approximation, for each body of the system, a pseudo-center and pseudo-inertia tensor have been defined. Using the pseudo-center of the body (if defined) as the origin of the body-fixed frame, the first (dipole) moment disappears in the force approximation. In this derivation, the quadrupole terms have been elaborated and expressed in terms of the elements of the pseudo-inertia tensor (quadrupole moment tensor). This matrix is constant for each rigid body (subdomain) if expressed in its own body basis. As such, it is formed in the preprocessing step, and there is no additional cost associated with forming this tensor during the course of the simulation.

REFERENCES

- [1] Ding, H.Q., Karasawa, N., and Goddard III, W.A. Atomic level simulation on a million particles: The cell multipole method for Coulomb and London nonbond interactions. *J. Chem. Phys.* (1992) **97**: 4309–4315.
- [2] Ewald, P. Evaluation of optical and electrostatic lattice potentials. *Annual Physics* (1921) **64**: 253–287.
- [3] de Leeuw, S., Perram, J., and Smith, E. Simulation of electrostatic systems in periodic boundary conditions. I. lattice sums and dielectric constants. *Proceeding of Royal Society London A* (1980) **373**: 27–56.
- [4] Esselink, K. A comparison of algorithms for long-range interactions. *Comput. Phys. Commun.* (1995) **87**: 375–395.
- [5] Hockney, R.W., and Eastwood, J.W. *Computer simulation using particles*. McGraw Hill, New York (1981).
- [6] Sadus, R.J. *Molecular simulation of fluids*. Elsevier Science, first edition (1999).
- [7] Appel, A.W. An efficient program for many-body simulation. *SIAM J. Sci. Stat. Comp.* (1985) **6**: 85–103.

- [8] Barns, J., and Hut, P. A hierarchical $O(N \log N)$ force-calculation algorithm. *Letters to Nature* (1986) **324**: 446–449.
- [9] Greengard, L., and Rokhlin, V. A fast algorithm for particle simulations. *J. Comput. Phys.* (1997) **135**: 280–292.
- [10] Dehnen, W. A hierarchical $O(N)$ force calculation algorithm. *J. Comput. Phys.* (2002) **179**: 27–42.
- [11] Mathiowetz, A.M., Jain, A., Karasawa, N., and Goddard III, W.A. Protein simulations using techniques suitable for very large systems: the cell multipole method for nonbond interactions and the newton-euler inverse mass operator method for internal coordinate dynamics. *Proteins: Structures, Function, and Genetics* (1994) **20**: 227–247.
- [12] Boschitsch, A.H., Fenley, M.O., and Olson, W.K. A fast adaptive multipole algorithm for calculating screened Coulomb Yukawa interactions. *J. Comput. Phys.* (1999) **151**: 212–241.
- [13] Fong, W., and Darve, E. The black-box fast multipole method. *J. Comput. Phys.* (2009) **228**: 8712–8725.
- [14] Chun, H.M., Padilla, C.E., Chin, D.N., Watanabe, M., Karlov, V.I., Alper, H.E., Soosaar, K., Blair, K.B., Becker, O.M., Caves, L.S.D., Nagle, R., Haney, D.N., and Farmer, B.L. MBO(N)D: A multibody method for long-time molecular dynamics simulations. *J. Comput. Chem.* (2000) **21**: 159–184.
- [15] Jain, A., Vaidehi, N., and Rodriguez, G. A fast recursive algorithm for molecular dynamics simulation. *J. Comput. Phys.* (1993) **106**: 258–268.
- [16] Poursina, M., Bhalerao, K.D., Flores, S., Anderson, K.S., and Laederach, A. Strategies for articulated multibody-based adaptive coarse grain simulation of RNA. *Method. Enzymol.* (2011) **487**: 73–98.
- [17] Praprotnik, M., Site, L., and Kremer, K. Adaptive resolution molecular-dynamics simulation: Changing the degrees of freedom on the fly. *J. Chem. Phys.* (2005) **123**: 224106–224114.
- [18] Tozzini, V. Coarse-grained models for proteins. *Curr. Opin. Struct. Biol.* (2005) **15**: 144–150.
- [19] Kane, T.R., Likins, P.W., and Levinson, D.A. *Spacecraft Dynamics*. McGraw-Hill, NY (1983).
- [20] Leach, A.R. *Molecular Modelling Principles and Applications*. Prentice Hall, second edition (2001).

INFLUENCE OF MECHANICAL STRAINS ON ELECTROMAGNETIC SIGNALS OF A MICROSTRIP ANTENNA. FEM/BIM MODEL

N. ADNET^{*}, F. PABLO^{*}, I. BRUANT^{*} AND L. PROSLIER^{*}

^{*} Laboratoire Energétique Mécanique Electromagnétisme EA 4416 (LEME)
Université Paris Ouest Nanterre La Défense
50 rue de Sèvres, 92410 Ville d'Avray (France)
e-mail: {nicolas.adnet, frederic.pablo, isabelle.bruant, laurent.proslier}@u-paris10.fr
web page: <http://leme.u-paris10.fr>

Key words: mechanical/electromagnetic coupling, conformal cavity patch antenna, radar cross section, vector finite element, nodal finite element, boundary integral method, finite element method, Maxwell's equations.

Abstract. A hybrid numerical technique is proposed for a characterization of the radar cross section of a microstrip patch antenna residing in a dielectric filled cavity which is loaded by a sinusoidal mechanical pressure. A new 3D hexahedral finite element is developed in order to take into account the deformed shape of the antenna within the electromagnetic computations. The numerical tool combines the finite element and boundary integral methods to formulate a system for the solution of the fields at the aperture and those inside the cavity. In this work, numerical examples are presented for demonstrating the ability and the validity of the hexahedral element.

1 INTRODUCTION

The framework of the present study is the development of metamaterial strip antennas for aeronautical applications. This work is part of the MSIE project (Intelligent Materials and Structures for Electromagnetism), launched in 2008 by the competitiveness French cluster AStech. The aim of this project is the reduction of antennas on aircrafts using conformable microstrips made of metamaterials.

Such antennas, which are thin and conformable to planar or warped surfaces, can easily be embedded on any aircraft surfaces. Since these surfaces are subjected to aerodynamic and thermal loads, the influences of geometry change and mechanical strains on the electromagnetic signals of the antenna have to be studied. Available commercial predictive analysis tools appear not to be well suited for structural systems involving multiphysical couplings, and in particular electromagnetism.

The main difficulty in the solution of a problem involving deformable antennas is the accurate evaluation of the integrals over the unbounded domain of the electromagnetic radiation. Classical numerical methods for structural mechanics, such as Finite Element Method (FEM), require the discretisation of the domain far from the source region, which

leads to an excessive computational effort. An efficient solution consists in the combined Finite Element-Boundary Integral (FE-BI) Method [1], first developed in mechanical engineering and later introduced in electromagnetism.

This method consists in introducing a fictitious boundary that encloses the structures to be studied. Classical FEM is then used to approximate the fields in the interior domain, whereas the fields in the exterior region are evaluated by the boundary integral method. These fields are coupled at the fictitious boundary via the field continuity conditions. This leads to a coupled system of the interior and boundary fields to be solved.

The new numerical tool here proposed is based on a weak coupling between electromagnetism and mechanical behaviour. The FEM/BIM approach is used with a new dedicated 3D hexahedral finite element. The mechanical field is approximated through a classical nodal formulation while the electromagnetic one is expressed by an edge formulation [2]. This completely new software has been implemented from scratch and it has been validated through various benchmarks found in the literature [4]. This contribution presents an extension of the FEM/BIM technique applied to numerical simulations for the characterization of a patch antenna subjected to bending loads.

2 FORMULATION OF THE COUPLED PROBLEM

2.1 Geometric description

In the present work, the considered antennal structure is illustrated in **Figure 1**. It consists in a rectangular patch antenna residing on a cavity [1] recessed in an infinite ground plane. The cavity is filled with a homogeneous dielectric material having a relative permittivity ϵ_r and a relative permeability μ_r . The antenna thickness is assumed to be smaller than the substrate height. Therefore, it will be neglected in the following developments. Furthermore, this structure is supposed to be integrated on composite aircraft panels.

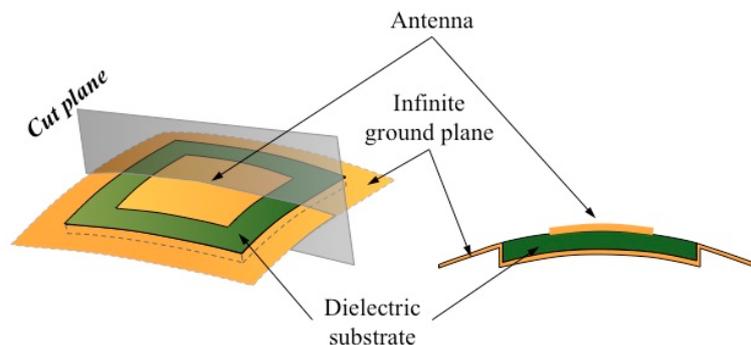


Figure 1: Studied one patch antenna

In the following sections, the free space region above the cavity and the ground plane, and below these ones will be respectively denoted by V_∞ and V symbols (see **Figure 2**).

The antenna is studied in a receiving run mode. Thus, the cavity is illuminated by electromagnetic waves emitted by a source point of the region V_∞ . Considering the Cartesian coordinate system presented in the **Figure 3**, the incident wave, emitted from a source point M'

and received by an observation point M , is respectively located by the following position vectors $\vec{r}' = \overline{OM'}$ and $\vec{r} = \overline{OM}$. The related Euclidian norms are defined by $r' = \|\vec{r}'\| = \sqrt{x'^2 + y'^2 + z'^2}$ and $r = \|\vec{r}\| = \sqrt{x^2 + y^2 + z^2}$.

At last, the cavity is assumed to be subjected to some mechanical loads such as pressures. The governing equations and variational formulations are presented in the next sections.

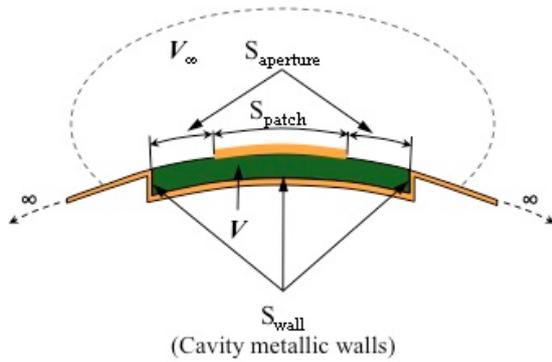


Figure 2: The cavity model and its surrounding regions

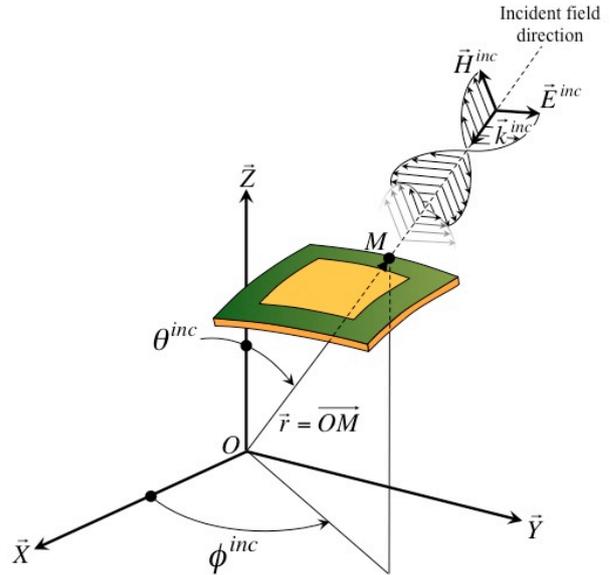


Figure 3: The incident field and the Cartesian coordinate system

2.2 Mechanical governing equations and weak form

Assuming the antenna is embedded on a composite panel, the cavity volume V will be subjected to mechanical distortions of its support. Taking into account strains in the electromagnetic computations is then essential.

Considering a static approach, the mechanical governing equations consists in the equilibrium equation, the constitutive equation in V and the strain-displacement relationship [3] :

$$\overline{\text{div}} T + \vec{f} = \vec{0} \quad (1)$$

$$T = C S(\vec{u}) \quad (2)$$

$$S(\vec{u}) = \frac{1}{2} (\text{grad } \vec{u} + \text{grad}^T \vec{u}) \quad (3)$$

where T , $S(\vec{u})$, \vec{f} and C are respectively the stress tensor, the strain tensor, the prescribed body forces vector applied to V and the elastic stiffness tensor. In addition to these equations, one has to take into account the mechanical boundary conditions : a first one where mechanical vector displacement \vec{u} is imposed, and the second one where a surface force vector \vec{F} is applied on the boundary ∂V_F .

The weak form of the mechanical boundary value problem is then given by :

find $\bar{u} \in U_{ad}$ such that : (4)

$$\int_V S^*(\bar{u}^*) : T(\bar{u}) dV = \int_V \bar{u}^* \vec{f} dV + \int_{\partial V_f} \bar{u}^* \vec{F} dS$$

for any admissible virtual displacement $\bar{u}^* \in U_{ad}^*$. This variational principle is the usual starting point for any finite element approximation in mechanical modelling.

2.3 Electromagnetic governing equations and weak form

Here, the FEM/BIM technique [4] is used. The proposed method involves the equivalence principle to subdivide the original problem into two equivalent problems. They are then coupled by enforcing field continuity. The fields in each region are subsequently formulated via a variational or integral equation approach leading to a coupled set of integral equations solved via the finite element method.

2.3.1 The governing equations

The Maxwell's equations (in the usual time harmonic convention $e^{j\omega t}$) and the constitutive relations are :

$$\overrightarrow{rot} \vec{E} = -j\omega \mu_r \mu_0 \vec{H} \quad (5)$$

$$\overrightarrow{rot} \vec{H} = j\omega \epsilon_r \epsilon_0 \vec{E} \quad (6)$$

$$div(\vec{D}) = \rho \quad (7)$$

$$div(\vec{B}) = 0 \quad (8)$$

$$\vec{D} = \epsilon_0 \epsilon_r \vec{E}, \vec{B} = \mu_0 \mu_r \vec{H} \quad (9)$$

where \vec{D} , \vec{B} , ϵ_0 , μ_0 and ρ respectively are the electric flux density related to the \vec{E} field, the magnetic flux density related to the \vec{H} field, the permittivity and the permeability in the free space region and the electric charge density.

2.3.2 Boundary conditions and loads

- *The excitation vector* : Considering the antenna runs in a receiving mode, the excitation vector is defined by the incident plane wave model [4]. The propagation direction of the travelling wave is defined by \vec{k}^{inc} . Thus, the excitation vector received by an observation point located by \vec{r} is given as a function of the electric field polarisation \vec{E}_0 and the wavenumber k_0 by :

$$\vec{E}^{inc} = \vec{E}_0 e^{-j k_0 \vec{k}^{inc} \cdot \vec{r}} \quad (10)$$

- *The perfect electric conductor* : the boundary conditions associated to the cavity model are those of perfect electric conductors prescribed on the cavity walls indicated in the **Figure 2**. They enforce no tangential electric fields on the metallic regions such as the

patch and the ground plane :

$$\vec{n} \wedge \vec{E} = \vec{0} \quad (11)$$

- *The integral equation on the aperture surface S* : in accordance with the equivalence principle [4], the fields existing in the free space region (above the ground plane) are then due to the radiation caused by the equivalent magnetic currents $\vec{M} = \vec{E} \wedge \vec{n}$ residing on the ground plane. Accordingly, the magnetic field \vec{H}^{ext} created by \vec{M} in the external region is defined by :

$$\vec{H}^{ext} = 2\vec{H}^{inc} - 2j\frac{k}{Z_0} \iint_S \overline{\overline{G_0}}(\vec{r}, \vec{r}') (\vec{E} \wedge \vec{n}) dS' \quad (12)$$

where \vec{H}^{inc} denotes the magnetic field due to an incident wave, Z_0 is the intrinsic impedance in free space and $\overline{\overline{G_0}}$ is the free space dyadic Green's function [5]. Enforcing continuity of the tangential electric fields across the interface S , the fields in both the internal and external regions are then coupled :

$$\vec{n} \wedge \begin{cases} \vec{E}^{int} \\ \vec{H}^{int} \end{cases} = \vec{n} \wedge \begin{cases} \vec{E}^{ext} \\ \vec{H}^{ext} \end{cases} \quad (13)$$

where the superscripts *int* and *ext* respectively denote the fields created in the internal region V and the unbounded external region V_∞ .

2.3.3 Weak form formulation

Considering the equations (5) to (13), the variational principle leads to the weak form of the electromagnetic problem :

$$\text{find } \vec{E} \in E_{ad} \text{ such that :} \quad (14)$$

$$b(\vec{E}, \vec{E}^*) = c(\vec{E}^*)$$

for all admissible virtual electric field $\vec{E}^* \in E_{ad}^*$, with :

$$\begin{aligned} b(\vec{E}, \vec{E}^*) = & \int_V \overrightarrow{rot} \vec{E}^* \mu_r^{-1} \overrightarrow{rot} \vec{E} dV - k_0^2 \int_V \vec{E}^* \epsilon_r \vec{E} dV \\ & + k_0^2 \int_S \int_S \vec{E}_S^* (\vec{n} \wedge \overline{\overline{G_0}}(\vec{r}, \vec{r}') \wedge \vec{n}') \vec{E}'_S dS' dS \end{aligned} \quad (15)$$

$$c(\vec{E}^*) = -2jk_0 \int_S (\vec{n} \wedge \vec{E}_S^*) (\vec{k}^{inc} \wedge \vec{E}^{inc}) dS \quad (16)$$

where $S = S_{aperture}$ represents the aperture zone of the cavity that is not covered by the metallic patch as illustrated in **Figure 2**. The \vec{n} and \vec{n}' denote the normals associated to the observation point M and the source point M' lying on the aperture area. The fields \vec{E} and \vec{E}_S are respectively the internal and aperture electric fields.

A difficulty in the evaluation of this quadruple surface integral within (15) is the usual singularity associated with the derivatives of the free space Green's function. This can be

numerically computed by considering two different Gauss quadratures, each is related to the observation surface (S) and the source surface (S').

3 SOLVING THE COUPLED PROBLEM

3.1 Weak form of the coupling

The mechanical behaviour of the antenna is assumed to be non sensitive to the electromagnetic phenomena. Therefore, a weak coupling between the electromagnetic part and the mechanical one is considered. The problem can be studied in two successive steps :

- firstly, the mechanical problem (4) is solved in order to compute the deformed shape of the antenna. The displacements are interpolated at the element nodes.
- secondly, the computation of the electromagnetic problem (14) is done by taking into account the deformed mesh.

3.2 A 3D hexahedral vector/node-based element

The antenna is meshed using 3D hexahedral finite elements (see **Figure 4**). The same element is used to solve either the mechanical or the electromagnetic problem. In order to ensure the compatibility between the two physical domains, the mechanical displacements are linearly interpolated at the 8 nodes of the element :

$$\vec{u} = \sum_{n=1}^8 N_n(\xi, \eta, \zeta) \vec{u}_n \quad (17)$$

whereas the electric fields are approximated through the 12 edges of those :

$$\vec{E} = \sum_{n=1}^{12} \vec{W}_n(\xi, \eta, \zeta) E_n \quad (18)$$

The N_n are the classical nodal shape functions of a hexahedra while \vec{W}_n are the vector shape functions [1] related to this kind of finite element. In (19), the quantities E_n represents the electric field magnitude lying on the hexahedra edges.

Through this way, the hexahedral elements are able to take into account the deformed shape information when performing the electromagnetic computation. By virtue of the finite element method, the technique is then pertinent to study a patch antenna residing on a dielectric substrate that is mechanically loaded.

Introducing the field approximations in equations (4) and (14), the mechanical discretized problem can be written as the following form :

$$K_{FEM} \{q_U\} = \{F_{ext}\}_{mech} \quad (19)$$

where K_{FEM} , $\{q_U\}$ and $\{F_{ext}\}_{mech}$, are respectively the stiffness matrix, the unknown displacements and the mechanical excitation vector. For the electromagnetic problem, one obtains :

$$\left(Y_{FEM} - k_0^2(L_{FEM} - L_{BIM})\right)\{q_E\} = \{F_{ext}\}_{em} \quad (20)$$

where Y_{FEM} , L_{FEM} , L_{BIM} , $\{q_E\}$ and $\{F_{ext}\}_{em}$ represent respectively the admittance matrix, the inductance matrices related to the interior and the exterior fields, the unknown electric fields and the electromagnetic excitation vector. The subscripts FEM and BIM denote the matrices that are resulting from the discretization of the interior volume V and the aperture region S .

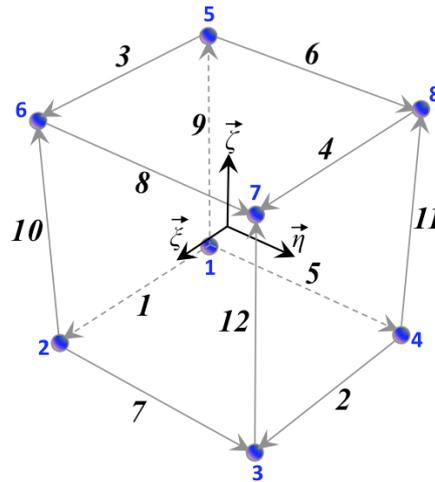


Figure 4: The 3D hexahedral vector/node-based finite element

4 NUMERICAL VALIDATION

In order to validate the coupling formulation, a numerical test presented in [4] is considered. It consists in computing the radar cross section (RCS) of a patch antenna as a function of its curvature. In the test, the deformed shape of the antenna is just controlled by the curvature radius. In this section, the RCS of the same antenna is computed with a deformed mesh resulting from a mechanical pressure load.

The radar cross section σ is computed from the scattered magnetic fields \vec{H}^{far} and is defined by :

$$\sigma = \lim_{r \rightarrow \infty} 4\pi r^2 \frac{|\vec{H}^{far}|^2}{|\vec{H}^{inc}|^2} \quad (21)$$

where $|\vec{H}^{inc}|^2$ is equal to unity. The RCS is usually normalized to λ^2 , the squared wavelength.

The scattered magnetic fields result from the integration of the surface magnetic currents emitted by a test point M' lying on the aperture surface S , and received by an observation point M located at the infinity space :

$$\vec{H}^{far} = j \frac{k_0}{Z_0} \int_S \vec{G}_0(r, r') (\vec{n} \wedge \vec{E}'_S) dS' \quad (22)$$

4.1 Geometric description of the benchmark

The structure consists in a rectangular patch residing in a dielectric filled cavity, which the relative permittivity is $\epsilon_r = 2.17$, the relative permeability is $\mu_r = 1.00$ and the Young modulus is about $E = 1\,000$ MPa. The patch dimensions are 20 mm x 30 mm and the cavity size is 50 mm x 60 mm x 0.7874 mm. The antenna is simply supported along the y direction at $x = 0$ and is mechanically loaded by a sinusoidal pressure :

$$p = p_0 \sin\left(\pi \frac{x}{L}\right) \tag{23}$$

where $L = 50$ mm here.

Considering a receiving antenna, the electromagnetic excitation is a vertical incident field and set for a transversal magnetic polarisation ($\alpha = 0^\circ$, see Figure 5).

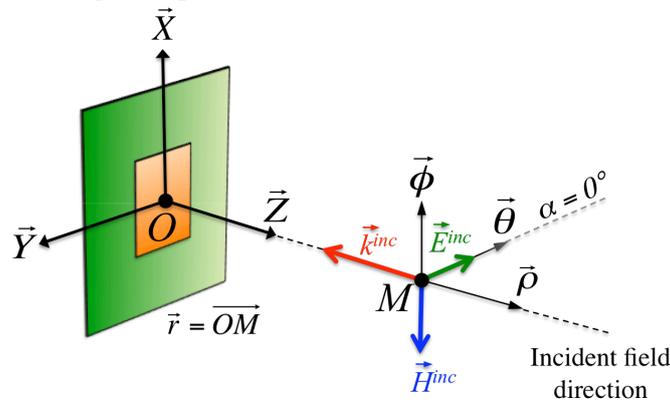
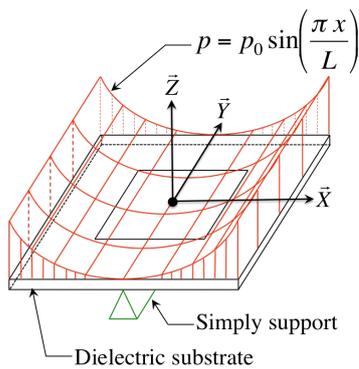


Figure 5: Incident field and polarisation setup

The sensitivity of the RCS to the curvature is studied with respect to various pressure magnitudes p_0 and with respect to a frequency range going from 4.0 GHz to 6.0 GHz. The magnitudes p_0 listed in Table 1 were chosen in order to match two curvatures studied by [4] for the considered antenna. Figure 6 presents two configurations of the strained antenna resulting from the cylindrical pressure p .



Pressure magnitude p_0	Associated curvature radius R
0.54×10^{-3} MPa	2 000 mm
7.21×10^{-3} MPa	150 mm

Table 1: Curvature radii as a function of the mechanical pressure p loading the antenna

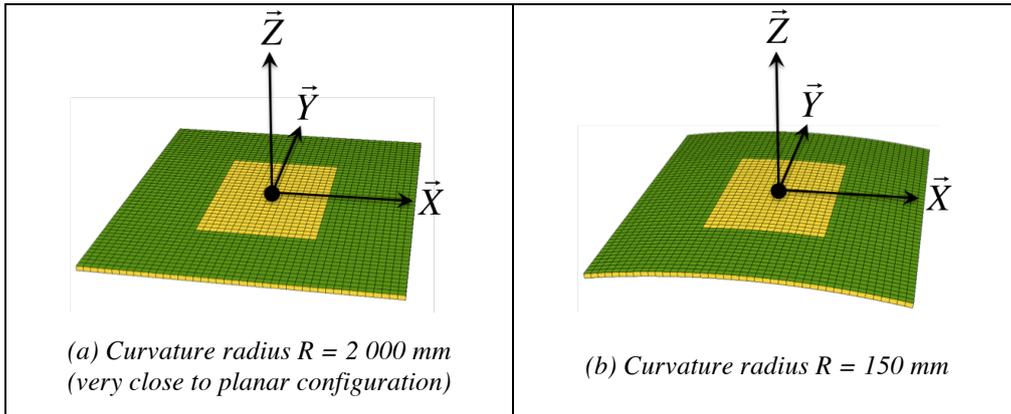


Figure 6: Configurations of the strained patch antenna loaded by a sinusoidal pressure

4.2 RCS as a function of the frequency

The backscattered RCS of the antenna are shown in **Figure 7**. They are related to the planar and the curved cases which the curvature radii $R = 150$ mm and $R = 2\,000$ mm. The results performed with the hexahedral elements are in good agreement with the brick elements simulations done in [4]. The resonance frequencies and the dB levels are similar with both the hexahedra and brick finite elements. As expected for a very large radius (2 000 mm), it can be seen that the planar and the curved RCS are nearly merged. A 19% RCS shift is noticed for the 150 mm curved configuration. Therefore, the planar configuration represents an envelope for infinite radii of curvature and bounds lower ones.

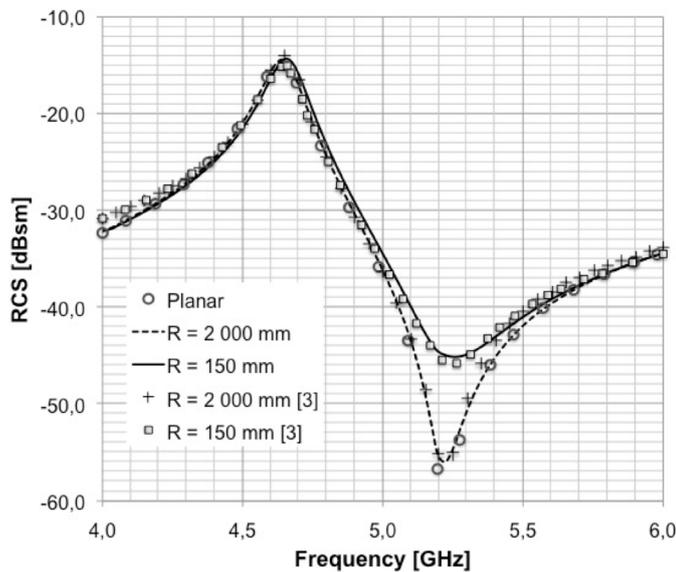


Figure 7: RCS of a patch antenna related to a planar and two cylindrical structural configurations

Conversely, an important point has to be underlined concerning the numerical treatment of the BIM part. Our solutions are computed with two different Gauss quadratures whereas an

exact form of the planar Green's function contribution is considered in the computational process used in [4]. Moreover, the ability of hexahedral elements to take into account each node coordinates allows both planar and conformal configurations to be computed through the same element. For brick elements, treating planar and conformal shape needs to implement a *planar* and a *curved* brick element respectively. In the future, the hexahedral elements will enable us to treat more arbitrary strained structures than using bricks and furthermore, to investigate the effects of singly or doubly curvatures on the RCS.

5 CONCLUSIONS

The presented work concerns the computation of the RCS property of a patch antenna loaded by a sinusoidal mechanical pressure. A weak coupling between the electromagnetic and the mechanical domains was assumed. The simulations were performed through the FEM/BIM method considering a new 3D hexahedral finite element that combines the interpolation of the electromagnetic fields (at the element edges) and the mechanical fields (at the nodes of the hexahedra). Then, the main feature of the hexahedral element is the ability to take into account the deformed shape of the antenna within the electromagnetic computations.

Moreover, the FEM/BIM technique allows the computational domain to be only reduced to the structure volume since the unbounded external region is modelled through the BIM and then avoid the free space to be meshed.

The good agreement between the numerical results and the literature ones confirms the validity of the hexahedral element and the ability of the method to be applied to simulate the mechanic/electromagnetic coupled problem for a cylindrical conforming context. In order to enlarge the use of this work, more a kind of mechanical loads imposed on the antenna has to be considered such as twisting loads.

REFERENCES

- [1] Jin, J.M. *The finite element method in Electromagnetics*. Wiley-Interscience. 2nd edition (2002).
- [2] Nédélec, J. C. *A new family of mixed elements in R^3* . Numerical Mathematics, Vol. **50**, (1986) pp. 57-81.
- [3] Fung, Y.C. *Foundations of Solid Mechanics*. Prentice-Hall, Inc (1965).
- [4] Volakis, J.L. Chatterjee, A. and Kempel L.C. *Finite Element Method for Electromagnetics; Antennas, Microwave circuits and scattering applications*. IEEE Press (1998).
- [5] Jin, J.M. and Volakis, J.L. *A hybrid finite element method for scattering and radiation by microstrip patch antennas and arrays residing in a cavity*. IEEE Trans. Ant. Propagat., (1991) 39(11):1598-1604.
- [6] Idelsohn, S.R. and Oñate, E. Finite element and finite volumes. Two good friends. *Int. J. Num. Meth. Engng* (1994) **37**:3323-3341.

ELECTROMAGNETIC PLASMA MODELING IN CIRCUIT BREAKER WITHIN THE FINITE VOLUME METHOD

L. RONDOT^{*}, J.-P. GONNET[†] AND V. MAZAURIC[♠]

^{*} CEDRAT, 15 chemin de Malacher, 38246 – Meylan, France
e-mail: loic.rondot@cedrat.com, web page: <http://www.magsoft-flux.com>

[†] Schneider Electric, Power Business, 38EQI, 38050 – Grenoble Cedex, France
e-mail: jean-paul.gonnet@schneider-electric.com, web page: <http://www.schneider-electric.com>

[♠] Schneider Electric, Strategy & Technology., 38TEC, 38050 – Grenoble Cedex, France
e-mail: vincent.mazauric@schneider-electric.com, web page: <http://www.schneider-electric.com>

Key words: Multiphysics Problems, Finite Volume Method, Plasma, Circuit Breaker.

Abstract. In order to ensure the galvanic isolation of an electrical system following a manual operation or a default strike, current limitation properties of the electric arc are used, forcing a fast decrease to zero current. Modeling this process reveals complex, since it involves a large amount of physical phenomena (radiation, phase transitions, electromagnetism, fluid dynamics, plasma physics). In order to get a robust solving, enhancing strongly coupled resolution and time constants compatibility, the Finite Volume Method has been chosen. This method was first implemented on intrinsic electromagnetism problems (current flow, magnetostatics including non-linear materials, and magnetodynamics). Once validated, the models have been successfully used in the Schneider's current-interruption dedicated software, thus allowing a significantly improved simulation of Schneider Electric circuit breakers.

1 INTRODUCTION

Various multiphysic modeling, and more particularly the arc interruption modeling, requires fluid dynamics and electromagnetic models [1]. Whereas conventional Finite Volume Method (FVM) usually dedicated to Computational Fluid Dynamics enforces the local conservation of mass, momentum and energy [2], low-frequency electromagnetic software minimizes global energy functionals within the Finite Elements Method (FEM) [3]. Hence, magnetohydrodynamics problems are currently resolved by using either:

- a FE-CFD code, which is not suitable for circuit breaker applications because the method is unable to model the shock waves during the interruption process; or
- an hybrid method combining a FVM and a FEM, thereby sacrificing the high level of integration and the accuracy achieved with a single mesh [4].

Thus, the search for a common, effective and integrated model and important hydrodynamic constraints call for a single numerical method for the two phenomena. In this work, an electromagnetic model based on the FVM is adopted.

2 ELECTROMAGNETISM WITHIN FINITE VOLUME METHOD

2.1 Finite Volume Method (FVM)

The Finite Volume Method is based on the local conservation (Fig. 1). Expressing the balance in any elementary volume between the dissipative term (RHS) and the variation over time and the flux exchanged at its boundary (LHS), it reads:

$$\underbrace{\frac{\partial \gamma}{\partial t}}_{\text{Transient term}} + \underbrace{\text{div}(-k \mathbf{grad} \gamma + \mathbf{v} \gamma)}_{\text{diffusive and convective flux term}} = \underbrace{\sum}_{\text{source term}} \quad (1)$$

where γ denotes the transported physical quantity, k is the diffusion coefficient of the transported quantity and \mathbf{v} its velocity.

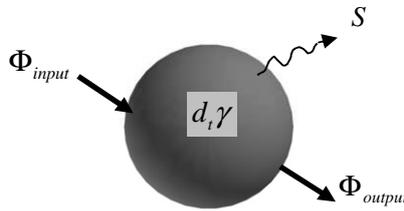


Figure 1: Local balance of the flow Φ : $d_t \gamma$ is the local variation and S denotes the source term.

In FEM the unknowns are associated to the nodes or edges of the mesh and the integral approximation is made by an approximation of the solution. In contrast, the unknowns are on the cell-centroid and the integral approximation is made by an operator approximation in FVM. Hence, the diffusive flux D_f , across a face f of a scalar V is given by:

$$D_f = \iiint_{\Omega} \text{div}(-k \mathbf{grad} V) d\Omega = \underbrace{-k \frac{V_1 - V_0}{ds} \frac{\mathbf{A} \cdot \mathbf{A}}{\mathbf{A} \cdot \mathbf{es}}}_{\text{Primary flux (implicit term)}} + \underbrace{k \mathbf{grad} V \cdot \left(\mathbf{A} - \mathbf{es} \frac{\mathbf{A} \cdot \mathbf{A}}{\mathbf{A} \cdot \mathbf{es}} \right)}_{\text{Secondary flux (explicit term)}} \quad (2)$$

where k is the diffusion coefficient at the face, V_0 and V_1 the scalar value in the cell c_0 and c_1 , \mathbf{A} the area normal vector of face directed from cell c_0 to c_1 , ds the distance between the cell centroids, and \mathbf{es} the unit normal vector in this direction.

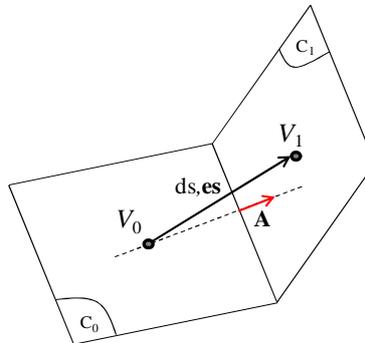


Figure 2: Diffusive flux approximation for non-orthogonal meshes

While the first term of D_f on the right hand side represents the primary gradient directed along the vector \mathbf{es} , the second term represents the cross diffusion term directed along the vector \mathbf{es}^\perp . Notice that the former is an implicit term whereas the latter is explicit and allows correcting fluxes for non-orthogonal meshes (Fig. 2).

The previous description was adapted successfully to solve linear magnetostatic problem within the so-called \mathbf{T}_0 - ϕ formulation [5]. This formulation is known to require a lower number of degree of freedom than the magnetic potential vector formulation [6]. Unfortunately, no extension to non-linear case was given, preventing to compute the inclusion of the electric arc in the splitter plates. In order to make feasible the whole arc interruption process – including its vanishing by the inclusion in the splitter plates –, a specific formulation should be developed to take non-linear magnetostatic properties into account.

2.2 Non linear Magnetostatics (3D)

Within the 3D magnetostatic \mathbf{T}_0 - ϕ formulation, the magnetic field reads $\mathbf{H} = \mathbf{T}_0 - \mathbf{grad}\phi$. Hence, the magnetic flux density divergence-free is expressed as a diffusion equation with a source term resulting from the field \mathbf{T}_0 obtained in vacuum:

$$\underbrace{\operatorname{div}(-\mu_r \nabla \phi)}_{\text{diffusive term}} = \underbrace{\operatorname{div}(-\mu_r \mathbf{T}_0)}_{\text{source term}} \quad (4)$$

where μ_r is the non-linear magnetic permeability. Notice that $-\mathbf{grad}\phi$ is nothing but the demagnetizing field if \mathbf{T}_0 is given by the Biot and Savart's law.

Among the various choices to describe the non-linearity [7], a two-parameter \tan^{-1} law is used to fit the anhysteretic curve. After an initial magnetostatic resolution in vacuum to compute \mathbf{T}_0 and the set-up of the relative magnetic permeability μ_r – typically the half-value of the relative magnetic permeability at the origin μ_{r_init} –, the computation is performed iteratively thanks to an update of μ_r . To avoid the oscillation around the solution, an over-relaxation on μ_r is used. The inspection of the flux density conservation provides a criterion to check the convergence (Fig. 3).

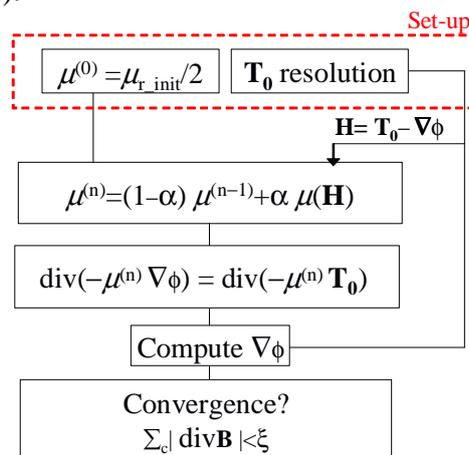


Figure 3: Flow chart algorithm for the resolution of a non-linear magnetostatic problem.

The previous developments are used in a 2D case where two steady and opposite currents flow in circular conductors to magnetize a non-linear ferromagnetic plate (Fig. 4).

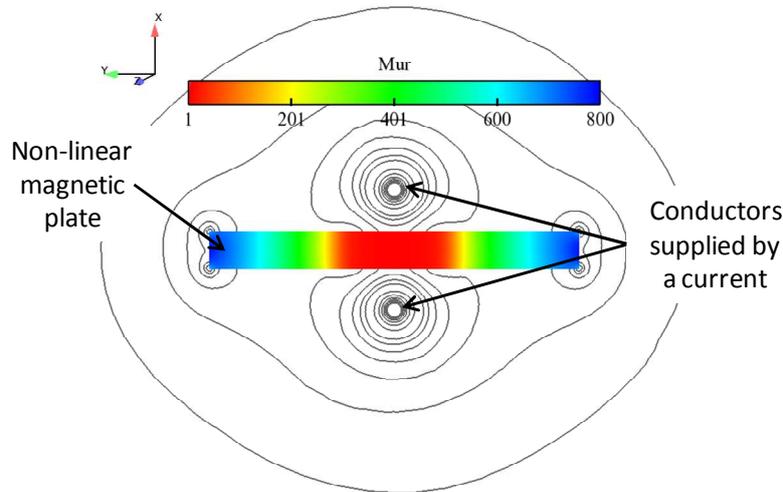


Figure 4: Refraction of the H-lines around a non-linear ferromagnetic plate magnetized by two conductors supplied by opposite currents: The level of saturation of the plate may be followed by the permeability value μ_r .

The saturation is observed qualitatively and the numerical comparison with finite element computations – achieved with Flux2D[®] software – show that the relative error in energy deviates less than 5% (Fig. 5). Hence, the extension of the $\mathbf{T}-\phi$ formulation in 3D is quite straightforward [8].

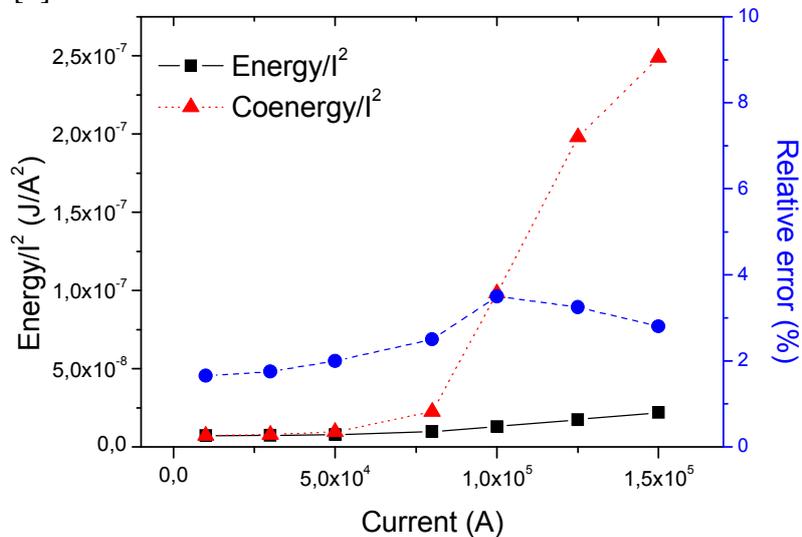


Figure 5: Energy and co-energy curves reduced by squared current (left) and the maximum relative error computed with FEM (right) vs. supplied current.

3 CIRCUIT BREAKER MODELING

In the interruption process, the magneto-dynamic effects (eddy currents) can be neglected both in feeders and splitter plates (often called arc chutes) [9]. Therefore it is possible to model the process within the previous FV electromagnetic developments.

3.1 The framework of the resolution procedure

The resolution procedure follows the chart represented in Fig.6 where one iteration is represented. The previous formulation was implemented within the plasma physics-dedicated Schneider Electric software. This code already included:

- a real gas model,
- a radiation model,
- a table of electrical properties,
- an arc root model.

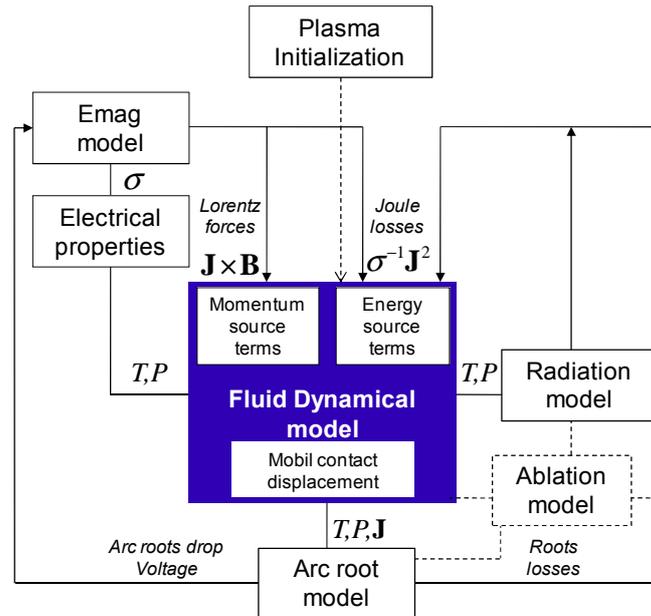


Figure 6: Resolution procedure chart for the interruption process modeling (T is the temperature, P the pressure, and σ denotes the electrical conductivity).

All the models (real gas, radiation, root model and electromagnetism) are driven by the fluid dynamical based core [10]. These models are the inputs of source terms in energy and momentum. In the case of electromagnetism, the Lorentz forces and the Joule losses are introduced in fluid dynamical solver (Fig. 5).

This resolution code is achieved within the CFD Fluent[®] code thanks to the explicit solver, which use a Gauss-Seidel method with a multi-grid resolution.

2.2 Circuit breaker modeling

The Fig. 7 shows a LV arc chamber currently designed in Schneider Electric. After modeling of the whole interruption process occurring therein thanks to the non-linear developments derived above, comparison with experiments is provided in Fig. 8. Agreement is quite encouraging on both current and voltage drop.



Figure 7: Typical circuit breaker arc chamber: While two electric arcs are initiated by the rotation of the moving conductor (center), their displacement towards the splitter plates (right and left) is enhanced both by the loop effect (provided by the feeders shape) and the switching reluctance effect (provided by the ferromagnetic parts).

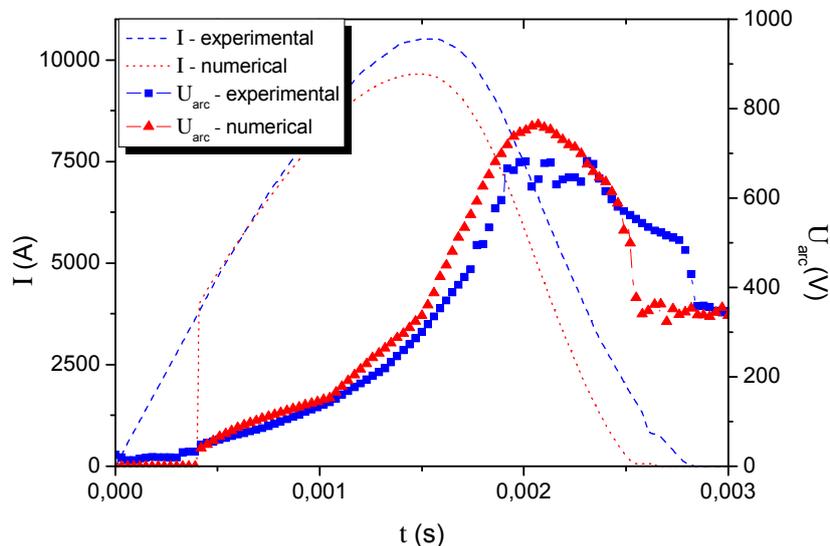


Figure 8: Comparison of numerical and experimental results on current and voltage drop during an interruption process.

Hence, further investigations on local behaviors occurring during the interruption process could be considered to enhance the arc chamber design. This kind of investigation was performed on an experimental mock-up, composed of two feeders, three ferromagnetic splitter plates, and a far pressure outlet (Fig. 9). The mesh has 400,000 cells and the unknown solving are $(\rho, \mathbf{v}, H, V, \mathbf{T}_0, \phi)$, respectively density, velocity, Gibbs' energy, electrical potential, field in the vacuum and magnetic scalar potential. The computational time for an arc interruption process modeling is about 3 weeks on Pentium Xeon single core 2 GHz -2Gb RAM, each time step running for 10ns.

The Fig. 9 also provides three iso-values of the current density during the whole arc interruption process, and especially when the arc is getting into the ferromagnetic splitter plates (also called arc chutes). As a result, the saturation of the splitter plates is effective

during the breaking process, showing that a non-linear treatment of the field is required to avoid an over-estimation of the driving force acting on the electric arc [6].

Furthermore, such investigations are necessary to enhance current limitation and subsequently thermal and electro-dynamical damages of the electrical installation.

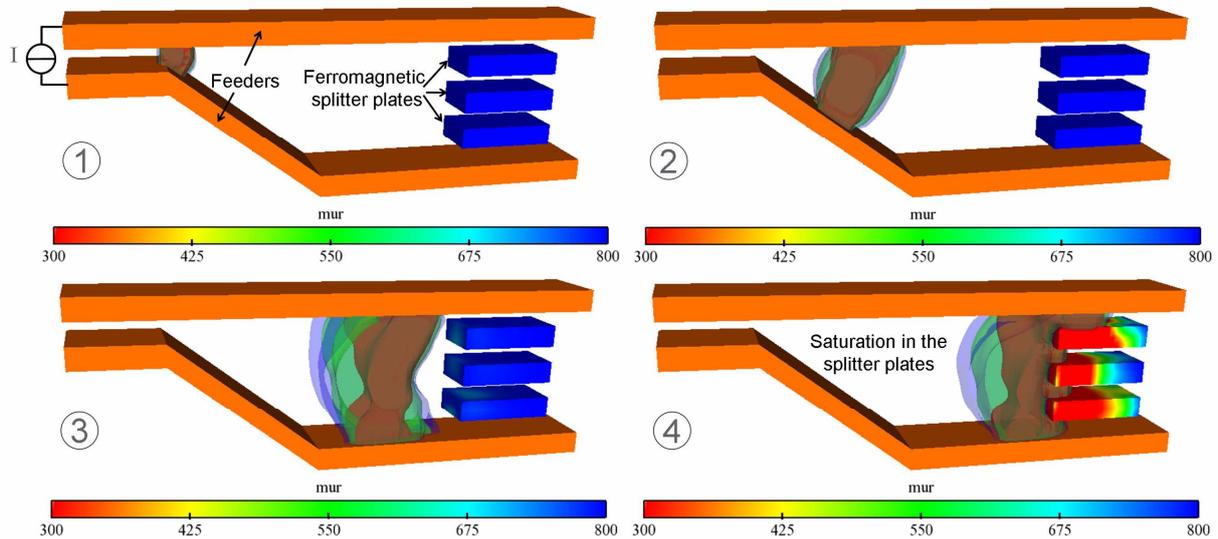


Figure 9: Modeling of the arc interruption process: While the electric arc is displayed with three iso-values of the current density ($1.5 \cdot 10^7$; $8.0 \cdot 10^6$; $5 \cdot 10^6$ A·m⁻²), the saturation of the splitter plates is effective and represented with the relative magnetic permeability μ_r .

CONCLUSION

Within sight of the results obtained thanks to non-linear magnetostatic developments, the Finite Volume Method seems suitable to model efficiently electromagnetism. This method does not have any ambition to compete with Finite Element Method but simply to allow a stronger coupling between the fluid dynamical and electromagnetism in the specific case of arc interruption. These electromagnetic developments allow already modeling the interruption process with ferromagnetic materials. The comparison with the experimental data is not easy but is under progress. Nevertheless, although non-linear ferromagnetic material modeling is already efficient, CPU-time remains too huge to be used for the modeling of the whole arc interruption process at a design center level. Subsequent improvements and productivity in the design of arc interruption products (Medium- and Low-Voltage circuit-breaker) are expected.

REFERENCES

- [1] C.A. Borghi and A. Cristofolini, "A Hybrid Implicit Numerical Method for the Analysis of the Magneto-Plasmadynamics in a Gas Discharge," *IEEE Transactions on Magnetics*, vol. 37, pp. 3401-3404, 2001.
- [2] S.V. Patankar, Numerical Heat Transfer and Fluid Flow in *Computational Methods in Mechanics and Thermal Sciences*, Hemisphere Publishing Corporation: New York U.S.A.
- [3] V. Mazauric, "From thermostatics to Maxwell's equations: A variational approach of electromagnetism," *IEEE Transactions on Magnetics*, vol. 40, pp. 945-948, 2004.
- [4] I.J. Lytle, Use of MPCCI to Perform Multidisciplinary Analyses for Electrical Distribution Equipment, *American Institute of Aeronautics and Astronautics*, paper 0122, 2005.
- [5] L. Rondot, V. Mazauric, Y. Delannoy, and G. Meunier, "Dedicating Finite Volume Method to electromagnetic plasma modeling: Circuit breaker application," *International Journal in Applied Electromagnetics and Mechanics*, vol. 28, pp. 3-9, 2008.
- [6] M. Springstube and M. Lindmayer, "Three-Dimensional-Simulation of Arc Motion Between Arc Runners Including the Influence of Ferromagnetic Material," *IEEE Transactions on Component and Packaging Technologies*, vol. 25, pp. 409-414, 2002.
- [7] W. Thompson and P. Munk, "Mathematical model of nonlinear magnetic cores", *IEEE Transactions on Magnetics*, vol. 6, p 523, 1970.
- [8] O. Biro, K. Preis, G. Vrisk, K. R. Richter, and I. Ticar, "Computation of 3D magnetostatic fields using a reduced scalar potential," *IEEE Transaction on Magnetics*, vol. 29, pp. 1329-1332, 1993.
- [9] O. Chadebec, G. Meunier, V.G. Mazauric, Y. Le Floch, and P. Labie, "Eddy-current effects in circuit breakers during arc displacement phase," *IEEE Transactions on Magnetics*, vol. 40, pp. 1358-1361, 2004.
- [10] F. Karetta and M. Lindmayer, "Simulation of the gasdynamic and electromagnetic processes in low voltage switching arcs", *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 21, pp. 96-103, 1998.

INVESTIGATION OF THE PURELY HYPERBOLIC MAXWELL SYSTEM FOR DIVERGENCE CLEANING IN DISCONTINUOUS GALERKIN BASED PARTICLE-IN-CELL METHODS

Andreas Stock*, Jonathan Neudorfer[†], Rudolf Schneider[‡],
Christoph Altmann[†] and Claus-Dieter Munz[†]

* [†]Institute for Aerodynamics and Gasdynamics (IAG)
University of Stuttgart
Pfaffenwaldring 21, 70569 Stuttgart, Germany
e-mail: stock@iag.uni-stuttgart.de, www.iag.uni-stuttgart.de

[‡]Karlsruhe Institute of Technology
Institute for Pulsed Power and Microwave Technology,
PF 3640, 76021 Karlsruhe, Germany
e-mail: rudolf.schneider@kit.edu, www.ihm.kit.edu

Key words: Particle-In-Cell (PIC) Method, Computational Plasma Physics, Hyperbolic Divergence Cleaning, Discontinuous Galerkin (DG) Method

Abstract. For the Particle-In-Cell (PIC) method which is used to simulate non-neutral plasma the charge conservation is violated. To solve this problem several approaches have been suggested. In this paper we shall investigate the hyperbolic divergence cleaning method. It has been proposed in two different manners, one enforcing only Gauss' law and another considering both Gauss' law and the magnetic monopole divergence constraint. We shall investigate the differences between the two approaches with numerical simulations of a quasi 2D diode with the PIC method and a pure Maxwell equations example satisfying the charge conservation in its initial conditions. We use a discontinuous Galerkin FEM-type space discretization method to solve the Maxwell equations. The results of the numerical investigations are presented and the advantages or disadvantages of both approaches w.r.t. performance and quality are discussed.

1 Introduction

For given charge and current density, ρ and \mathbf{j} , the evolution of the electromagnetic field is given by the Maxwell equations,

$$\frac{\partial \mathbf{E}}{\partial t} = c^2 \nabla \times \mathbf{B} - \frac{\mathbf{j}}{\varepsilon_0}, \quad (1)$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \quad (2)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (3)$$

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}, \quad (4)$$

where \mathbf{E} and \mathbf{B} denote the electric field and the magnetic induction field, respectively. c is the speed of light and ε_0 is the electric permittivity.

When ρ and \mathbf{j} satisfy the charge conservation equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{j} = 0, \quad (5)$$

for all times $t \geq 0$, and when the initial \mathbf{E} and \mathbf{B} satisfy (4) and (3), respectively, the Maxwell system of hyperbolic evolution equations (1) and (2) has a unique solution.

In the discrete case, i.e. using a numerical method such as the Finite-Element or Finite-Volume method, approximation errors of the fields (\mathbf{E}, \mathbf{B}) and the sources (ρ, \mathbf{j}) lead to small inconsistencies accumulating to large errors for long simulation times. Thus, the charge conservation (5) is not satisfied by the discrete approximation any more. In the scope of the Maxwell equations this means that Gauss' law (4) is violated.

Another situation where the charge conservation is not satisfied occurs in the Particle-In-Cell (PIC) framework where particles and fields are simulated in a self-consistent manner on discrete level [3, 4, 9]. In the PIC technique the particles are directly used as sources that are assigned to the grid via special deposition methods, e.g. nearest-grid-point, volume-weighting or shape-function [3, 9]. The deposition leads to errors in Gauss' law (4) which can become very large in the course of time.

To resolve this problem mainly two approaches have been proposed, i.e. the projection method [3] and a set of methods relying on a generalized Lagrange multiplier (GLM) approach for the Maxwell equations. The latter has three different formulations, the hyperbolic-elliptic constrained formulation proposed by Assous et al. [2], the hyperbolic-parabolic form introduced by Marder [11] and the purely hyperbolic Maxwell (PHM) system introduced by Munz et al. [12, 13], known as hyperbolic divergence cleaning.

In the frame of the standard projection approach a Poisson equation has to be solved which requires large computational effort and is cumbersome for straightforward implementation of the scheme on parallel platforms.

The hyperbolic divergence cleaning technique seemed to be the best choice w.r.t. numerical complexity, locality and parallelization since it can be solved explicitly and, beyond that, yields qualitatively similar results to the projection method [9]. It is easy to

implement in an explicit discontinuous Galerkin method, which will be used within this work.

To satisfy charge conservation (5) the divergence constraint of Gauss' law (4) has to be coupled with the evolution equation (1). This is done by a generalized Lagrange multiplier Φ . For the hyperbolic divergence cleaning Munz et al. proposed in [12] to incorporate the Lagrange multiplier in the Maxwell equations as follows

$$\frac{\partial \mathbf{E}}{\partial t} = c^2 \nabla \times \mathbf{B} - c^2 \nabla \Phi - \frac{\mathbf{j}}{\varepsilon_0}, \quad (6a)$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}, \quad (6b)$$

$$\frac{\partial \Phi}{\partial t} = -\chi^2 \nabla \cdot \mathbf{E} + \chi^2 \frac{\rho}{\varepsilon_0}, \quad (6c)$$

where χ is a positive dimensionless parameter. The new variable Φ is an additional degree of freedom, leading to a seven variable system in three space dimensions, named PHM7 in the following. The physical interpretation of the new variable is a correction potential that transports the divergence errors of the electrical field out of the domain with the speed $\chi \cdot c$. The higher χ is the better (4) is satisfied. Obviously the monopole constraint (3) for the magnetic fields is left out of consideration in the PHM7 system. None the less it was recognized that violations of the magnetic divergence constraint also appear, similar as in the context of Magneto-Hydro-Dynamics (MHD) [5, 1]. Thus in the same year Munz et al. presented in [13] a system involving the magnetic divergence constraint, yielding

$$\frac{\partial \mathbf{E}}{\partial t} = c^2 \nabla \times \mathbf{B} - \chi c^2 \nabla \Phi - \frac{\mathbf{j}}{\varepsilon_0}, \quad (7a)$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} - \gamma \nabla \psi, \quad (7b)$$

$$\frac{\partial \Psi}{\partial t} = -\gamma c^2 \nabla \cdot \mathbf{B}, \quad (7c)$$

$$\frac{\partial \Phi}{\partial t} = -\chi \nabla \cdot \mathbf{E} + \chi \frac{\rho}{\varepsilon_0}, \quad (7d)$$

where γ has the same function as χ . This system couples the originally elliptic divergence constraint (3) with the hyperbolic evolution equation (2) through the additional variable Ψ , yielding eight variables in total. Thus we call it PHM8 in the following. In this paper we shall investigate the differences between the PHM7 and PHM8 system w.r.t. the quality of the divergence cleaning and the performance with the help of a pure Maxwell equation and a PIC problem.

1.1 Numerical Method

For the spatial discretization we use a mixed nodal and modal FEM-type discontinuous Galerkin (DG) method [6, 7] and a low storage fourth-order explicit Runge-Kutta (LSERK4) scheme [10] for the temporal time discretization.

The computational domain Ω is divided into K non-overlapping hexahedral or tetrahedral elements; pyramids and prisms are also possible. In each element locally the state vector $\mathbf{u}^k(\mathbf{x}, t)$ is approximated by a polynomial $\mathbf{u}_h^k(\mathbf{x}, t)$ of order p ,

$$\mathbf{u}^k(\mathbf{x}, t) \approx \mathbf{u}_h^k(\mathbf{x}, t) = \sum_{i=1}^M \tilde{\mathbf{u}}^k(\mathbf{x}_i, t) L_i(\mathbf{x}) = \sum_{n=1}^N \hat{\mathbf{u}}_n^k(t) \phi_n(\mathbf{x}) \quad (8)$$

where $\{L_i\}_{i=1,\dots,M}$ is the multivariate Lagrange basis, $\tilde{\mathbf{u}}(\mathbf{x}_i, t)$ is the state evaluated at M nodal interpolation points (IPs) \mathbf{x}_i , $\{\phi_n\}_{n=1,\dots,N}$ is the modal basis and $\hat{\mathbf{u}}(t)$ the modal coefficients. In three dimensions each element needs $N = (p + 1)(p + 2)(p + 3)/6$ modal degrees of freedom (DOF). The choice of M and the location of the IPs \mathbf{x}_i is discussed in [6]. M only equals N in case of linear tetrahedrons; for all other types of elements $M > N$. Important for the spatial resolution of the electromagnetic fields is the number of IPs in different element types. This has to be taken into account later when comparing the Cartesian mesh build with hexahedral elements and the unstructured mesh build with tetrahedral elements. The modal polynomial is used to achieve a quadrature free method and the nodal polynomial is used to determine the physical fluxes as well as to incorporate the sources, i.e. the particles in the PIC method. For further details on the DG method we refer to [7].

2 Numerical Experiments

In order to quantify errors arising from the approximation of the magnetic divergence constraint (3) and Gauss' law (4), mainly two quantities will be investigated: The divergence of the magnetic field component, $\nabla \cdot \mathbf{B}$, and the error in Gauss' law, $\nabla \cdot \mathbf{E} - \rho/\varepsilon_0$, which indirectly also measures the violation of the charge conservation (5). We consider the L_2 -norm of the time derivative of the Lagrange multiplier. For instance, in the case of the PHM8 system we obtain

$$\frac{1}{\chi} \left\| \frac{\partial \Phi}{\partial t} \right\|_{L_2} := \frac{1}{\chi} \sqrt{\int_{\Omega} \left| \frac{\partial \Phi}{\partial t} \right|^2 dV} = \sqrt{\int_{\Omega} \left| \nabla \cdot \mathbf{E} - \frac{\rho}{\varepsilon_0} \right|^2 dV}. \quad (9)$$

Here we evaluate the right expression because in this case, even for $\chi = 0$, the norm can be computed. In the following we will call (9) the divergence error of the charge conservation. The L_2 -norm of the magnetic field divergence is computed in a similarly straightforward way.

For the PHM8 system we always set $\chi = \gamma$. A different value can be chosen if it is necessary to clean the divergence of one of the two electromagnetic field components stronger than the other.

2.1 Example - 3D Issautier Example

The following example points out the significant differences between the PHM8 and PHM7 systems. In PIC codes the charge conservation (5) usually is violated due to the deposition method. In a pure Maxwell equation setup without particles but analytically

initiated sources we can force charge conservation. This has been done by Issautier et al. for a two-dimensional transverse electric (TE) Maxwell system [8]. We extended their idea to a three-dimensional fully electromagnetic setup with the sources

$$\begin{aligned} \rho(\mathbf{x}, t) &= \sin(\omega t) [\sin(\pi y) + \sin(\pi x) + \sin(\pi z)], \\ j_x(\mathbf{x}, t) &= (\cos(\omega t) - 1) [\pi \cos(\pi x) + \pi^2 x \sin(\pi z)] - \cos(\omega t) x \sin(\pi z), \end{aligned} \quad (10a)$$

$$j_y(\mathbf{x}, t) = (\cos(\omega t) - 1) [\pi \cos(\pi y) + \pi^2 y \sin(\pi x)] - \cos(\omega t) y \sin(\pi x), \quad (10b)$$

$$j_z(\mathbf{x}, t) = (\cos(\omega t) - 1) [\pi \cos(\pi z) + \pi^2 z \sin(\pi y)] - \cos(\omega t) z \sin(\pi y). \quad (10c)$$

The divergence of the current density is

$$\nabla \cdot \mathbf{j} = -\cos(\omega t) [\sin(\pi y) + \sin(\pi x) + \sin(\pi z)], \quad (11)$$

and the derivative of the charge density yields

$$\frac{\partial \rho}{\partial t} = \cos(\omega t) [\sin(\pi y) + \sin(\pi x) + \sin(\pi z)], \quad (12)$$

satisfying the charge conservation equation (5). The spatial setup is given as a unity cube by $(1 \times 1 \times 1 \text{ [m}^3\text{)})$ centered in the origin with irradiation boundaries (see [12]) and $\omega = 10^{-9} \text{ [1/s]}$. For the Cartesian mesh we used $10 \times 10 \times 10$ hexahedrons and for the unstructured mesh roughly 2400 tetrahedrons. Note that a hexahedron can be filled with six tetrahedrons of the same edge length. The edge length does not necessarily determine the resolution since in our DG implementation a fourth-order hexahedron has 56 IPs and a tetrahedron has 20 IPs, yielding 56000 IPs and 48000 IPs for the Cartesian unstructured mesh, respectively. Therefore we can assume a nearly equal resolution even if the number of tetrahedrons does not multiply by a factor of six.

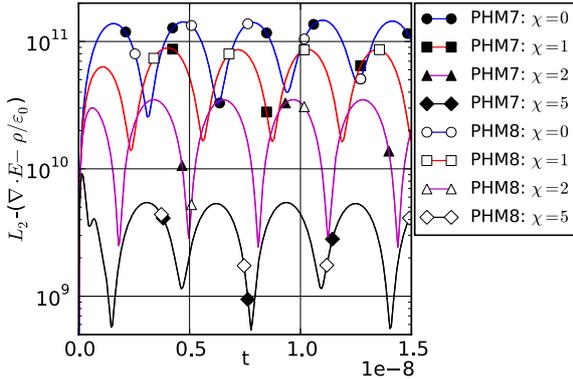


Figure 1: L_2 -norm of charge conservation error for the Cartesian mesh.

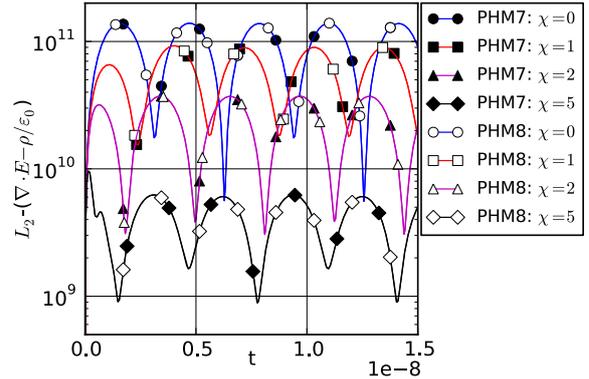


Figure 2: L_2 -norm of charge conservation error for the unstructured mesh.

The computations have been performed with $\chi = [0, 1, 2, 5]$ and a fourth-order DG method. Figure 1 and 2 show the L_2 -norm (9) for the Cartesian and the unstructured mesh computed with the PHM system. The results of both PHM systems are equal.

There is no difference between the PHM7 and PHM8 results. This is reasonable since the compared L_2 -norm (9) is based in both cases on the same equations, i.e. (1) and (4) with the specific Lagrange multiplier Φ . The influence of the magnetic field component in this early phase of the simulation is still small enough and does not influence the electric field. In a later phase when the magnetic field component is stronger developed it will also affect the electrical field and a difference between the PHM7 and PHM8 system will occur. In good comparison with results by [12, 13] the error is damped more strongly with increasing χ .

Not very well visible, the Cartesian mesh evolves slightly lower errors than the unstructured mesh, especially for $\chi = 5$. It is known that unstructured grids evolve higher divergence errors [12, 13]. None the less, an optimal mesh requires a detailed analysis of the wave propagation directions. A spherical structured mesh might be an even better choice for this problem.

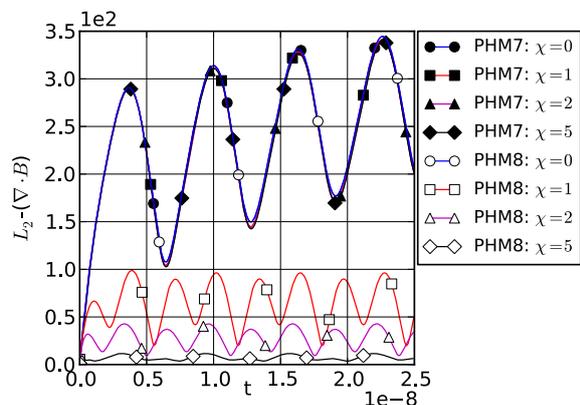


Figure 3: L_2 -norm of $\nabla \cdot \mathbf{B}$ for the Cartesian mesh.

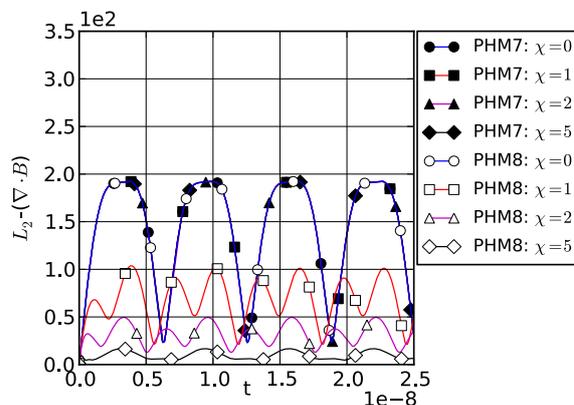


Figure 4: L_2 -norm of $\nabla \cdot \mathbf{B}$ for the unstructured mesh.

Figure 3 and 4 show the divergence of the magnetic field for the Cartesian and the unstructured mesh. With the PHM7 system and the Cartesian mesh, with any χ the magnetic divergence is equal and growing. For the unstructured mesh it is continuously oscillating without increase. In both cases the divergence of the magnetic field is much larger than for the PHM8 system. This is obvious since (3) is not enforced by the PHM7 system. In contrast to that, for the PHM8 system with $\chi = \gamma = [1, 2, 5]$ the divergence is damped to a lower and not growing value. For the PHM8 system with $\chi = \gamma = 0$ the divergence evolves equal to the PHM7 system since the constraint (3) is not enforced either. In agreement to the previous results for the charge conservation, also the Cartesian mesh shows better results for the damping of the magnetic divergence.

The influence of the magnetic field on the electric field is described by (1). There it is added with the factor c^2 to the derivative of the electric field. Thus, even a small divergence error in the magnetic fields affects the electric field. The growing values of the magnetic divergence for the PHM7 system with the structured mesh and also the non-

growing values for the unstructured mesh might become a problem for longer simulation times because the Poynting vector $\sim \mathbf{E} \times \mathbf{B}$ becomes erroneous. These results clearly demonstrate the necessity of enforcing also the magnetic divergence constraint, as in the PHM8 system.

Concerning the computational performance the expectation is that the PHM7 system is faster than the PHM8 system since one variable less has to be computed. In an optimal sense this would mean a reduction of the computational time of 12.5 %. Of course this is not realistic since a certain overhead (data-structure, I/O, initialization, etc.) exists. None the less a significant reduction of about 5 to 10 % can be expected.

Table 1 lists up the computational time t_{Tot} for this example. The initialization time is below 1 % of t_{Tot} to guarantee that mainly the performance of the DG scheme is measured. The speed-up S is given as $1 - t_{Tot,PHM7}/t_{Tot,PHM8}$.

Contrary to the expectation it seems as if the PHM7 system is slower than the PHM8 system. In total, we performed 11 runs leading to a mean value of -0.025 for the speed up and a relatively large standard deviation of 0.032, which means that in some runs we also achieved a positive speed up. Obviously for this test case the PHM7 system is even slower than the PHM8 system. None the less, w.r.t. the standard deviation we can identify a certain scattering of the results, which lies within the range of scattering that arises by performing the same computation again and again.

For the memory consumption we found a reduction of ≈ 1.5 % for the PHM7 system.

	$\chi = 0$		$\chi = 1$		$\chi = 2$		$\chi = 5$	
	PHM7	PHM8	PHM7	PHM8	PHM7	PHM8	PHM7	PHM8
t_{Tot} [s]	792.38	753.82	3111.71	3101.47	1492.81	1524.40	3840.32	3824.26
S [-]	-0.051		-0.003		0.021		-0.004	

Table 1: Computational time t_{Tot} and speed-up S for the Issautier example with different χ . Note, that the simulation time for each run was different.

2.2 Example - 2D Diode

To show the impact of the hyperbolic divergence cleaning, we chose a simple but very typical situation occurring in the context of PIC simulations. The domain ($0.1 \times 0.1 \times 0.0015$ [m³]) is considered to be a plane diode with cathode and anode situated on the left and on the right sides, respectively. Figures 5 and 6 show the domain in the x-y-plane for a Cartesian and an unstructured mesh which have been used in the following computations. The spatial setup is a three-dimensional grid trying to imitate a plane two-dimensional diode. In x-direction we have implemented perfectly conducting, in y-direction open and in z-direction periodic boundaries.

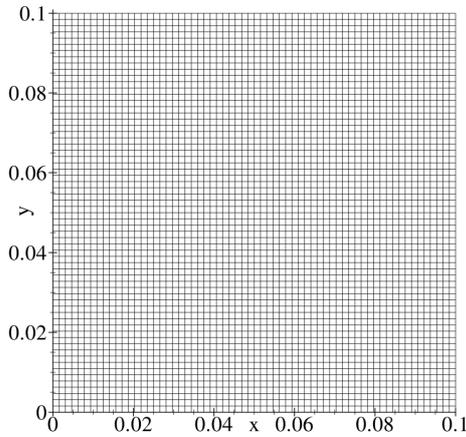


Figure 5: Cartesian grid of $64 \times 64 \times 1$ second-order hexahedrons with ≈ 33000 IPs.

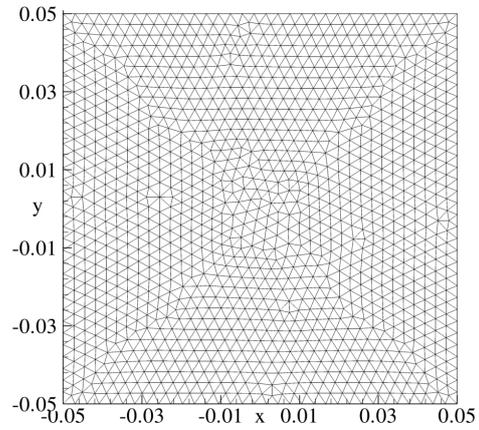


Figure 6: Unstructured grid of 7452 second-order tetrahedrons (edge length 0.0001) with ≈ 30000 IPs.

In the center of the cathode we inject an electron beam imitating a current density that has 75% of the space charge limiting current density. The length of the emitter is 37.5 [mm], the voltage is 100 [kV] and the current density is ≈ 5536 [Am^{-2}]. The current density is simulated by an electron beam where 9000 charges/ns are injected with a macro particle factor¹ of $2.180367 \cdot 10^5$, which are accelerated in x-direction by an external electric field of -10^6 [Vm^{-1}]. In the following we will only consider the PHM8 system since no significant qualitative and quantitative differences compared to the PHM7 system have been found.

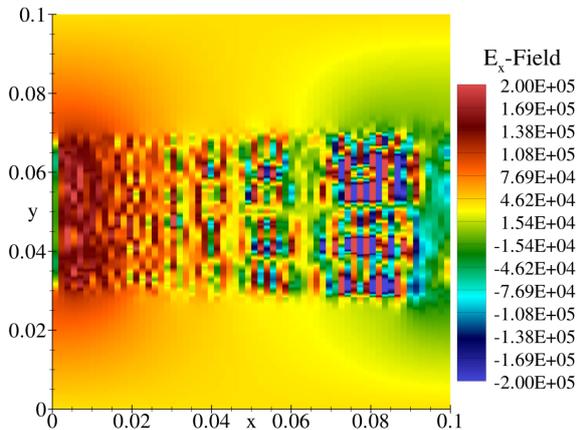


Figure 7: E_x -field at 50 [ns] of y-x-plane slice at 50% z-thickness for $\chi = 0$ with Cartesian mesh.

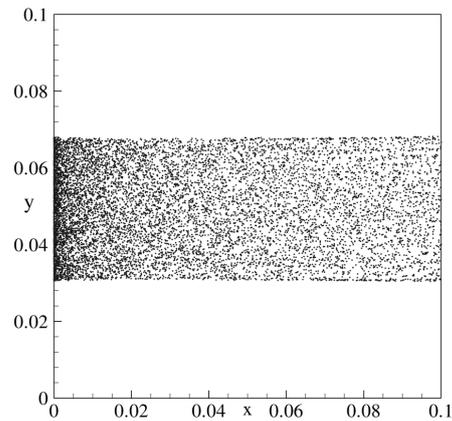


Figure 8: Particles for Cartesian mesh at 50 [ns] with $\chi = 0$.

Figure 7 shows the situation for the Cartesian mesh at simulation time $t = 50$ [ns] without divergence cleaning, only using the hyperbolic evolution equations (1) and (2).

¹One macro particle imitates the presence of a certain number of real particles. This number is the macro particle factor.

Strong gradients of the E_x -field in the region of the electron beam indicate large charge conservation errors. Figure 8 shows the corresponding position of the particles. A weak bunching of the beam in the left part of the figure can be observed. Filaments of particles, specially on the edges of the beam, can be found in the right part of the figure. These filaments are a typical sign for inconsistencies between the electrical field and the charge density, i.e. a violation of Gauss' law (4).

In Figure 9 and 10 we find the situation for $t = 50$ [ns] with $\chi = 1$. Here we can see a smooth electrical field without strong gradients and no significant divergence errors. The particle beam is diverging and filaments, especially on the edges, are not occurring. Moreover, a more homogeneous particle distribution can be found.

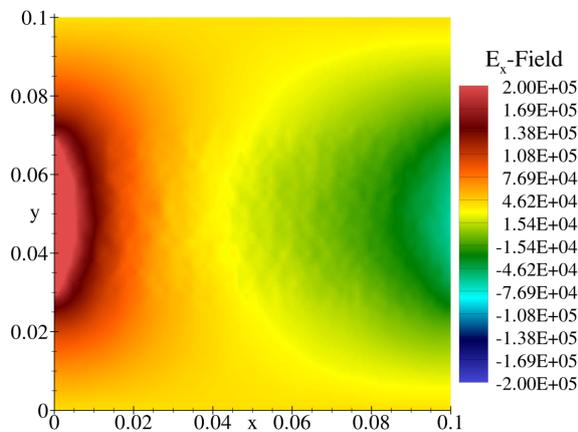


Figure 9: E_x -field at 50 [ns] of y-x-plane slice at 50% z-thickness for $\chi = 1$ with Cartesian mesh.

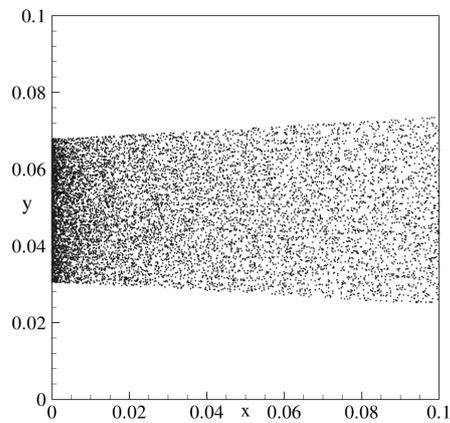


Figure 10: Particles for Cartesian mesh at 50 [ns] with $\chi = 1$.

The results for the unstructured mesh without divergence cleaning can be found in Figure 11 showing the E_x field and Figure 12 showing the corresponding particles at $t = 50$ [ns]. Strong scattering in the E_x -field indicates large charge conservation errors. The particles are totally scattered in a strong diverging fashion along the diode, which can be explained by the violation of the charge conservation. For $\chi = 1$ the results for the unstructured mesh are very similar to the results of the Cartesian mesh shown in Figure 9 and 10.

Figure 13 shows the charge conservation error for $\chi = [0, 1]$ in logarithmic scale due to a strong growth in the error with $\chi = 0$. We can find that for the unstructured grid the error grows significantly stronger than for the Cartesian grid. This can be explained by the propagation direction of the particles and the electromagnetic waves. While the Cartesian mesh is aligned with the propagation direction the unstructured mesh is not aligned. This leads to stronger errors.

Figure 14 shows the charge conservation error for the structured and unstructured grid with $\chi = [1, 2, 5]$. The errors evolve about linearly in the beginning phase and remain constant after 1.5 [ns]. Electrons are initially emitted from the left traveling about 1 [ns]

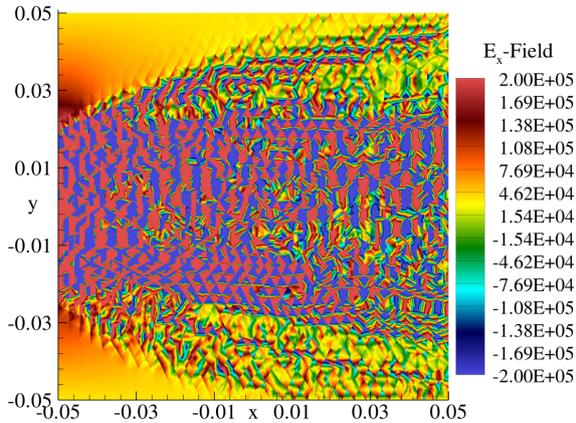


Figure 11: E_x -field at 50[ns] of y-x-plane slice at 50% z-thickness for $\chi = 0$ with unstructured mesh.

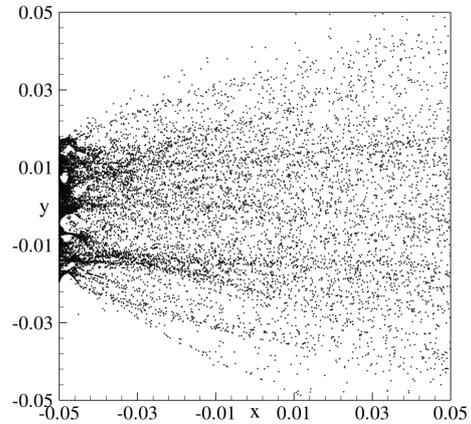


Figure 12: Particles for unstructured mesh at 50[ns] with $\chi = 0$.

to the right. After the initial phase and another 0.5 [ns] the current is fully developed and the errors stay constant. For increasing χ the error decreases as expected. Again the unstructured mesh has larger errors than the Cartesian mesh, which can be explained as mentioned above.

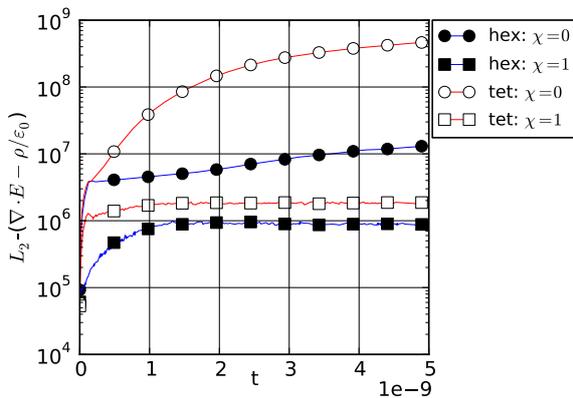


Figure 13: L_2 -norm of the charge conservation error for the Cartesian and unstructured mesh using hexahedral and tetrahedral elements, respectively. Computed with the PHM8 system for $\chi = [0, 1]$.

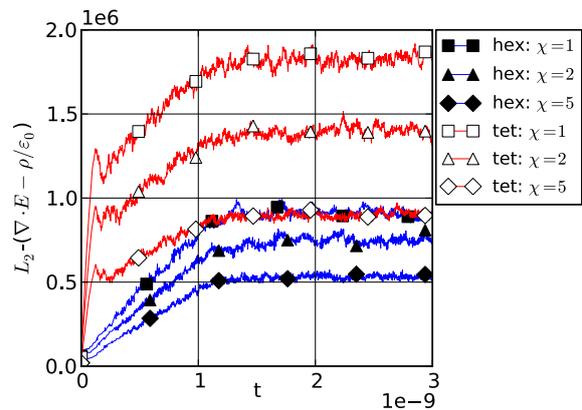


Figure 14: L_2 -norm of the charge conservation error for the Cartesian and unstructured mesh using hexahedral and tetrahedral elements, respectively. Computed with the PHM8 system for $\chi = [1, 2, 5]$.

Performance tests with seven runs revealed a small speed up of the PHM7 system with a mean value of 0.021 and a standard deviation of 0.035. This is not the expected speed up of $\approx 10\%$. Apparently, for a complex three-dimensional DG code on a modern cache-based CPU, it does not matter whether we use a seven or eight variable system. Latencies, caching and prefetching of data from the memory seem to have a greater effect than the computational overhead of one additional variable.

3 Conclusions

The necessity for divergence cleaning of the Maxwell equations has been demonstrated. Both test cases revealed no significant differences in the charge conservation error for both the PHM7 and PHM8 system. A significant difference exists for the enforcing of the magnetic divergence constraint. For the PHM7 system the magnetic divergence is increasing and influencing the results of electromagnetic wave propagation on a long time scale. On the short scale these errors do not influence the overall errors as shown in the pure Maxwell equations example. For an example where magnetic fields play an important role, such as the Weibel instability, this issue has to be taken into account.

A major influence on the errors is caused by the grid. In both examples the unstructured grid led to larger errors than the Cartesian grid. While for the PIC example significantly larger errors were observed, for the pure Maxwell equation example only slightly bigger errors appeared. When the grid is uniform with the wave and source propagation directions, errors are smaller than for a non-uniform grid. With $\chi = 1$ satisfactory results are achieved for both Cartesian and unstructured grids. None the less, a proper study of the wave propagation directions seems to be an important issue. A grid designed for these specific demands can improve the charge conservation and reduce magnetic divergence errors drastically and is better than using a high χ . Using high values of χ additionally is not advisable since the time step decreases due to the CFL condition leading to longer computation times.

Performance tests revealed no significant speed up for the PHM7 system compared to the PHM8 system. But memory reductions of about 1.5 % for the PHM7 system were measured. Apparently, for a complex three-dimensional DG code on a modern cache-based CPU, it does not matter whether we use a seven or eight variable system. Latencies, caching and prefetching of data from the memory seem to have a greater effect than the computational overhead of one additional variable. Thus performance considerations are not relevant for the choice of the system since the differences between them w.r.t. this issue are negligible.

In general we suggest to choose the PHM8 system since on long time scales not only the violation of charge conservation has to be considered, but also the harm of the magnetic divergence constraint, which affects the Poynting vector as well. While the PHM7 system neglects the magnetic divergence constraint the PHM8 system enforces this constraint.

4 Acknowledgement

We gratefully acknowledge the Deutsche Forschungsgemeinschaft (DFG) for funding this research within the project “Numerical Modeling and Simulation of Highly Rarefied Plasma Flows“.

REFERENCES

- [1] C. Altmann, G. Gassner, F. Lörcher and C.-D. Munz, *A Space-Time Expansion Discontinuous Galerkin Scheme With Local Time Stepping for the Ideal and Viscous*

- MHD Equations*, IEEE Transactions on Plasma Science, 2009, Vol. 37, 513 - 519.
- [2] F. Assous and P. Degond and E. Heintze and P.-A. Raviart and J. Segre, *On a finite-element method for solving the three-dimensional Maxwell equations*, J. Comput. Phys., 1993, Vol. 109, 222 - 237.
- [3] Birdsall, C.K. and Langdon, A.B., *Plasma Physics via Computer Simulation*, Adam Hilger, Bristol, Philadelphia, New York, 1991.
- [4] Boris, J. P., *Relativistic Plasma Simulations – Optimization of a Hybrid Code*, Proc. 4th Conf. on Num. Sim. of Plasmas, 1970, NRL Washington, Wash. DC, p. 3 - 67.
- [5] A. Dedner and F. Kemm and D. Kröner and C.-D. Munz and T. Schnitzer and M. Wesenberg, *Hyperbolic divergence cleaning for the MHD equations*, J. Comput. Phys., 2001, Vol. 175, p. 645 - 673.
- [6] G. Gassner and F. Lörcher and C.-D. Munz and J. S. Hesthaven, *Polymorphic Nodal Elements and their Application in Discontinuous Galerkin Methods*, J. Comput. Phys., 2009, Vol. 228, p. 1573 - 1590.
- [7] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods*, 2008, Springer, New York.
- [8] Issautier, D. and Poupaud, F. and Cioni, J.P. and Frezoui, L., *A 2D Vlasov-Maxwell solver on unstructured meshes*, in: Proc. 3rd Intl. Conf. on Math. & Num. Aspects on Wave Prop.; G. Cohen, P.Joly, E. Becache, J.E. Roberts (eds), 1995, p. 355 - 371.
- [9] Jacobs, G.B. and Hesthaven, J.S., *High-order nodal discontinuous Galerkin particle-in-cell method on unstructured grids*, J. Comput. Phys., 2006, Vol. 214, p. 96- 121.
- [10] Christopher A. Kennedy and Mark H. Carpenter and R. Michael Lewis, *Low-storage, explicit Runge-Kutta schemes for the compressible Navier-Stokes equations*, Appl. Num. Math., 2000, Vol. 35, p. 177 - 219.
- [11] B. Marder, *A method for incorporating Gauss' law into electromagnetic PIC codes*, J. Comput. Phys., 1987, Vol. 68, p. 48 - 55.
- [12] Munz, C.-D. and Omnes, P. and Schneider, R. and Sonnendrücker, E. and Voß, U., *Divergence correction techniques for Maxwell solvers based on a hyperbolic model*. J. Comput. Phys., 2000, Vol. 161, p. 484 - 511.
- [13] Munz, C.-D. and Omnes, P. and Schneider, R., *A three-dimensional finite-volume solver for the Maxwell equations with divergence cleaning on unstructured meshes*. Computer Physics Communications, 2000, Vol. 130, p. 83 - 117.

NUMERICAL SIMULATION FRAMEWORK FOR WEAKLY COUPLED MULTIPHYSICAL PROBLEMS IN ELECTRICAL ENGINEERING

PIERGIORGIO ALOTTO*, MICHAEL JAINDL†, RALPH KUTSCHERA†,
CHRISTIAN MAGELE† and ALICE KÖSTINGER†

*Dipartimento di Ingegneria Elettrica (DIE)
Universit degli Studi di Padova
Via Gradenigo 6/A, 35131 Padova, Italy
e-mail: alotto@die.unipd.it - Web page: <http://www.die.unipd.it>

† Institute for Fundamentals and Theory in Electrical Engineering (IGTE)
Graz University of Technology
Kopernikusgasse 23/III, 8010 Graz, Austria
Email: christian.magele@tugraz.at - Web page: <http://www.igte.tugraz.at>

Key words: Numerical Optimization, Simulation, Multiphysics, Weak Coupling.

Abstract. Every engineering discipline faces the fact of ever-shortening time-to-market windows and development cycles. In order to counteract these, virtual prototyping, simulation and problem optimization are employed in a rapidly increasing number of cases. Yet, the key to efficient problem formulation by professionals still lies in the use of sophisticated simulation software capable of processing numerous diverse design and optimization tasks in a versatile way.

More often than not, different tools for different workflows need to be coordinated and interdepend on each others data in the design process chain. When toolchains need to be run multiple times, as it is typically the case in numerical optimization, the lineup overhead tends to be tedious to both man and machine.

This paper describes different aspects concerning the design of a software and data framework which tackles the problem of lining up software tools that may be incoherent in terms of data exchange and control mode. The resulting system covers all parts of multiphysical simulation problems that may arise in electrical engineering and its adjoining disciplines as an application of the finite element method.

1 INTRODUCTION

A typical optimization cycle making use of the finite element method (FEM) [1] covers four domains: problem setup, model creation and discretization, solving the equation

system, and interpretation of the results to feed the next iteration, if necessary. Well-defined interfaces are mandatory for obtaining best results within such a cycle. Internally, one physics model's calculation relies on the results obtained by another one and vice versa, making accessible data structures mandatory. Externally, interfaces with optimization toolboxes or equation system solvers need to be found. These data structures form the major part of the research reported in this paper.

The central objectives aimed for consist of:

- The facilitation of solving forward problems because of direct, thus fast data access
- Dynamically adjustable model discretizations in both granularity and geometry
- Option of different models and discretizations for different physics
- Encapsulation of solution data, enabling access control for certain user groups.

Four classes of actors can be defined in a framework adhering to the design objectives mentioned above. The Controller takes the input as an XML string [5] and other models' postprocessing results, before each single model is set up using these data. Consequently, model solving is started. It also decides if a given stopping criterion is fulfilled and halts or re-starts the process accordingly.

Depending on the nature of a given problem, different Solver architectures can be employed, i.e. FEM, circuit analysis, or analytical solving (Python). Postprocessors enable access between different models' solutions and setups. The controlling functionality is provided by Matlab, interfacing with external structures (such as Python [6], FieldView [4]) is also provided. The Evaluator, in the end, computes output XML from all solutions accrued through direct access to these, including all intermediate optimization cycles' solutions. It also interfaces with a Python interpreter, thus enabling complex query constructions the formulation of which would be complicated or limited in predefined syntax.

When numerical optimization of multiphysical problems is to be addressed, three different issues arise. Each of these accounts for several dimensions of complexity and, therefore, different approaches towards solving have been researched. As a result, numerical simulation, optimization and coupling frameworks are in use in areas as different as natural science, engineering, medicine, and social sciences.

Common to all of these is that the numerical simulation of large-scale, complex systems requires consideration of a number of physical components, such as electromagnetics, solid mechanics, heat transfer, or fluid dynamics and the like [7]. This multidisciplinary nature of research requires simulation code development to be partitioned among different research groups, yielding independently developed software modules. Integration of these modules into a coupled system is usually done by integration frameworks [8].

2 FRAMEWORK COMPONENTS

In order to create a software tool capable of serving the above mentioned needs, a set of functionality requirements for the implementation of such a tool can be compiled as follows.

Sought for was the creation of a

- coupled
 - transient
 - multi-conductor
 - multi-objective
- optimization system
 - following a modular software architecture
 - allowing for multiple solver kernels to be triggered
 - in a client-server environment
 - * with definable user access rights
 - * offering output functionality suitable for web-browsers and
 - * accepting input via some special client or web application

For the framework that has been developed, the system boundary was set at taking over commands for solving a certain coupled problem, and at the delivery of the current solution to the optimization control again, thus offering full connectivity to an optimization toolbox. Fig. 1 shows an exemplary block diagram, containing both actors (coloured rectangles) and data to be exchanged between these (rounded rectangles). On the contrary, the optimization control functionality is denoted as an orange pentacle.

As the input for the problem setup should be any given problem that may be subject to optimization, data interoperability and data exchange were seen to be the main principles to be followed. Therefore, as a format for exchanging and storing data, eXtensible Markup Language XML was chosen [5].

In the engineering world, MATLAB is an omnipresent player in any profession and is used for a vast multiplicity of purposes. Therefore, a wide number of so-called toolboxes with standardized interfaces exist. The combination of this fact with offering a fully integrated development environment (IDE) for technical algorithm design, even allowing for object oriented design, led to the decision of selecting MATLAB for prototype development.

Depending on the application area, data representation follows one of three approaches. The finite element analysis makes use of certain parts that can be seen as objects, such as matrices, vectors, meshes and elements. Therefore, concepts of object oriented design

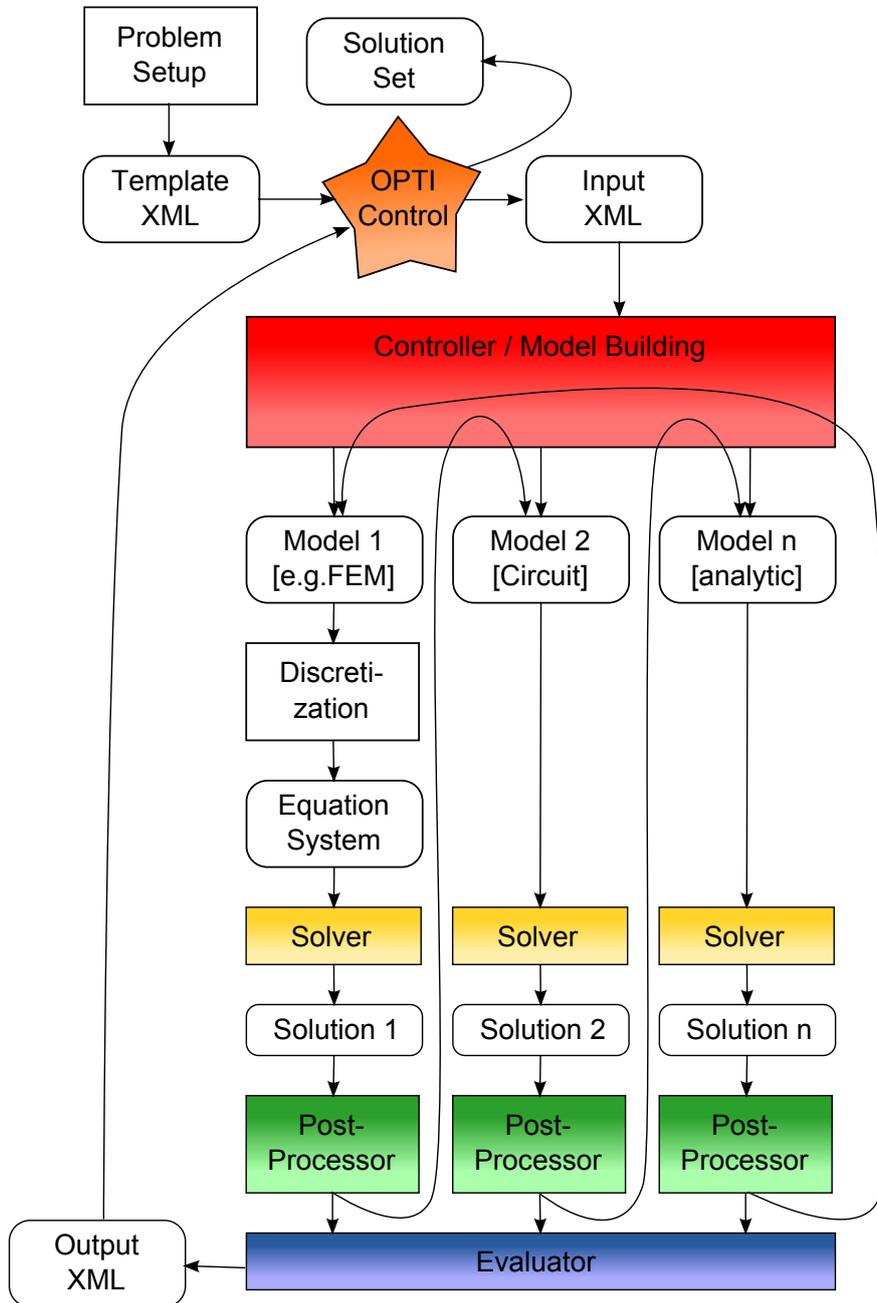


Figure 1: Block diagram of the ideal optimization framework

and programming are employed. That means for example that a computation's results are kept in a computer's memory represented by objects.

Persistent yet accessible storage of large data sets, on the other hand, requires employment of appropriate strategies. The Hierarchical Data Format (HDF) has proven to largely comply with these demands as well as to be robust even when holding huge amounts of data [2]. Therefore, interfaces with HDF have been created. Together with XML officiating as the data exchange backbone between the framework and the outside world as well as between actors within the framework, expedient data representation models have been identified and implemented for each use case.

2.1 Controller

The Controller offers the main interface to the outside world. It is also the main instance that governs the handling of the full process between input and output data as depicted in figure 1.

All other actors within that process are fed by data provided through the controller, and are also started and queried for results by this command structure accordingly. It therefore makes use of the overall input as an XML string and other models' simulation results as described in chapter 2.5. The input is split up into single models' data which is, where necessary, supplemented by data retrieved from other models' results in order to consequently yield self-contained and fully-fledged model data that can be forwarded to model solving functionality of whatever nature.

This linkage of different models defines the weak coupling process mentioned before. Most prominently, this will be the coupling of two finite element method's models, but the current functionality is not limited to these. For the simulation of electrical circuits, for example, active parts are typically presented as equivalent circuits. As such equivalent circuits often do not provide sufficient information about the behaviour of the replaced system, the field equations are solved numerically and then coupled with the circuit equations [9]. In order to fulfill these demands, the circuit can either be modelled in external applications, or through inclusion of the respective logics using the Python interface which has been created for this purpose.

2.2 FEM Solver

As a part of research conducted on the topic of creating a software framework for coupled simulation problems, a novel Finite Element Method solver was created. FEMtastic, as it was called by the developers, in its current state of development allows obtaining solutions to electrostatic, thermal, static current flow and magnetostatic problems alike, offering support to problems that are either of two-dimensional nature or can be modelled by a twodimensional setup. This can, for example, be achieved for rotational symmetric problems; cross sections through a problem geometry with an assumedly infinite third dimension would also qualify.

2.3 Analytical Solvers - Matlab and Python

Usually, computer applications are distributed in some sort of self-contained way, i.e. they offer encapsulated functionality aimed at solving a given problem or fulfilling a given task. Some programmes provide extensions or even extendability with limited complexity; the potential sphere of operations is limited in both cases.

Therefore, a completely open approach was selected for the work under scrutiny, open in a sense of not binding the user to a newly created set of directives and therefore to predefined functionality. Rather, a scripting framework was added, which would allow fast prototyping of new functions or even applications, thus postponing code specialization as much as possible [10].

For achieving such a goal, interpreted languages were chosen because of their characteristics of flexibility, interactivity and extensibility. Matlab, as forming the backbone programming language of the whole project, lends itself to being employed automatically. Apart from that, Python was selected for a number of reasons. It offers concise and almost pseudocodelike syntax, modularity, object oriented design principles, and the availability of numeric extensions that allow for efficient handling of large volumes of numerical data [10].

Within the simulation framework under scrutiny, an interface for carrying out Python as well as Matlab code was created. The latter one emerges straightforward from the design decision of Matlab being the development environment for most of the project, Python on the other hand offers the above mentioned assets. Therefore, models that do not have to undergo FEM simulation can be solved using deterministic code using the interfaces to these two environments.

2.4 Evaluator

The Evaluators main responsibility is to generate output values on the basis of design parameters. These serve, for example, as objective function values for evaluation of the obtained result by superimposed control structures.

Its functionality is more or less confined to executing the output queries that are extracted by the controller from the input XML data. As it is necessary when dealing with the interface to analytic solvers the Evaluator, in the end, offering similar functionality these query statements need to contain a returnValue array that is filled during the course of execution. The contents of this data vector are then inserted into the output XML structure.

2.5 Different Models' Data Interfaces: Implementing Weak Coupling

The import of values from one model to another is the process that is pursued with respect to weakly coupling them. Therefore, it plays an eminent role within the whole simulation workflow. Importing values is currently allowed for models that are to be solved analytically, and for certain characteristics of FEM models. For these, imported

values would qualify to be applied as sources or material parameters.

The *Import* - directive within the data model offers two ways of deriving values that are to be forwarded to the next model's setup. Figure 2 shows the according data schema diagram.



Figure 2: XML schema for importing

On the one hand, a number of readily available functions for common coupling problems exists. For these, the *type* - attribute of the *Import* element may be employed. Currently, the following directives are available:

- **max(ν)** gathers the maximal value of a quantity ν in the domain of the current macro element (where applicable). ν depends on the nature of the problem to be solved. Apart from that, also certain derivatives from the addressed result set may be queried. For electrostatic and current flow models, for example, ν may take the following values:
 - **U** delivers the maximal value of the electric potential Φ (denoted as **U**, to simplify matters)
 - **E** delivers the maximal value of the electric field strength depending on the potential distribution within the given domain
 - **J** delivers the maximal value of the current density, depending on electric field and material properties within the given domain
- **source** retrieves integral quantities that serve as a source with respect to the given simulation model's nature. For magnetostatic problems, this would be the total current I within a given domain.
- **energy** yields the total energy stored in a magnetic or electric field within a particular domain.

Then again, more complex dependencies have to be modelled using very specific functions that are not easily mapped using said features. For example, the temperature dependency of a metal's specific conductivity follows a curve that is best implemented using explicit formulae. In such cases, special code of any necessary complexity may

be implemented using *CouplingCode*. Its attribute *interpreter* denotes which code interpretation base should be deployed, e.g. Matlab directly, or Python via an implemented interface.

Common to both interface varieties is the necessity to determine where to take information from. *fromModel* denotes exactly that linkage, and refers to a model's name attribute. In that respect, the weak coupling scheme pursued within this work is of information pulling nature; a single model and its solving architecture does not have to be aware of being coupled with others, it is only queried for certain results when necessary.

When two or more models are not coupled in one direction, but information is also gathered from the last model in the chain to feed the first one again, the resulting workflow may easily end in infinite round trips. This may be desirable when particular results have to converge on account of changing external conditions determined by another simulation.

Nevertheless, such a process needs to be brought to a stop once stability is reached. This is usually done by comparing results of two succeeding simulation runs for their difference, and halt the round-trip accordingly as soon as a certain *threshold* is under-run. Both absolute and relative values are accepted for the respective attribute, with 1% being the default value that is employed automatically in order to avoid infinite loops.

3 SAMPLE APPLICATION: ELECTROTHERMAL HARDENING

Hardening of steel components is commonly done using the method of annealing. This typically requires controlled heating of the workpieces to exact temperatures following a just as exact time schedule in order to effectuate the desirable material properties.

Therefore, the design of an appropriate heating device would include the determination of the dimensions a heating plate should have, given the fact that the device is supplied with a fixed voltage. The optimization strategy and the determination of fitness with respect to the optimization objectives is left to an optimization toolbox not within the scope of this paper.

The two-dimensional cross-section model of such an appliance made from alloy is shown in Fig. 3 together with its cupreous feed connections shown on either side of the heating plate. An arbitrary workpiece of steel is placed on top of it. All measurements are given in meters, the third dimension is assumed to be infinite in length, which eases the model to being of two-dimensional nature only.

The electric potential distribution that arises drives current through the heating plate (as depicted in figure 4). As a result of the power dissipation due to ohmic losses therein, it will heat up and thus serve as a source for the thermal model simulation. Within this model, the feed connections can be neglected, but as, of course, the workpiece to be hardened has to be taken into account, a problem space discretization entirely different to the current flow model's is the result. Figure 5 shows the according model space and its discretization as well as the resulting simulated thermal distribution.

Now, taking advantage of the finite elements being realized as software objects, the input values for source elements of the thermal model can easily be determined for each

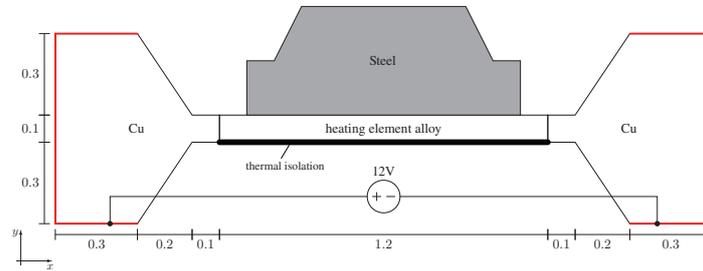


Figure 3: Heating device - problem geometry

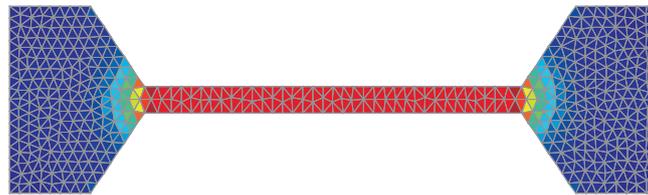


Figure 4: Current density causing the losses that serve as thermal source

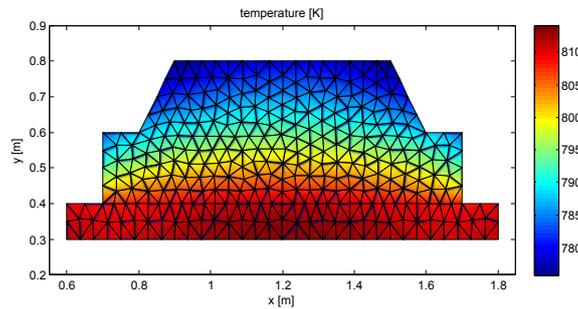


Figure 5: Thermal distribution within the modelled domain

finite element's integration points. As each FE object of the current flow problem offers its own postprocessing functionality, exact input values - the power dissipation in each integration point as a result of the current density \vec{J} - can be retrieved and fed into the new thermal source elements.

Also, the temperature that arises within the heating plate can influence its material properties, e.g. specific conductivity. Therefore it may be advisable to re-feed this information into the electric model and simulate it again, before starting the thermal simulation using the now freshly calculated current flow. This round-trip can be repeated until the difference between two iterations falls below a certain threshold.

Details to early stages of reasearch on this topic are summarized and published in [3]. The computation of a full optimization cycle for the given problem setup requires, as

outlined in above, an intermediate result to be computed. In the case of the first model – the static current flow model – that would be the distribution of the current density \mathbf{J} . In order to gain the final result, the thermal model has to be computed subsequently in order to receive the temperature distribution or the scalar temperature field T of the steel workpiece.

For the reason of having two coupled models originating from different physical domains this approach is called *multiphysics*. After computing the current flow within the heating plate, the respective power losses can be calculated. In the course of calculating the second thermal model, these power losses will act as energy sources for heat.

Figure 5 outlines the thermal subtask. Only the heating plate and the steel workpiece have to be modelled; thermal effects within the electrodes can safely be neglected in just the same manner as the influences on the current flow in the workpiece are disregarded. The heating plate is then constructed of an alloy with a thermal conductivity of $k = 85 \text{ W K}^{-1} \text{ m}^{-1}$, the steel workpiece on the other hand shows a thermal conductivity of $k = 42 \text{ W K}^{-1} \text{ m}^{-1}$.

The framework supports weakly coupled problems through the `<Import>` expression in the input XML file. It advises the controller to access the model that is identified by the therein contained number. It is then queried for its postprocessing features to return certain computed values. The controller, in the end, forwards these values to the actual model for assembling it into its calculations.

The current implementation necessitates parts analysed in both models to be located at the same positions in their respective geometry. For the given example this would mean that the heating plate features the same coordinates in both the current flow and the thermal model. Different setups of macroelements or a different meshing and triangulation would, on the contrary, be of no concern.

$$p = \frac{1}{\sigma} \mathbf{J}^T \mathbf{J} \quad (1)$$

The below mentioned listing shows how the import functionality for a given model is activated, in this actual case for a current flow - thermal coupling. Equation 1 states the relation between the current flow \mathbf{J} and the power loss p in $\frac{\text{W}}{\text{m}^3}$, with the latter being the source of the thermal problem. This relation is specified through the value of the `<Type>` directive being set to `current-thermal`. The value for `<Threshold>` can be set to any arbitrary value, when no roundtrip coupling is applied.

```
<Source>
  <Import>
    <FromModel>1</FromModel>
    <Type>current-thermal</Type>
    <Threshold>1</Threshold>
  </Import>
</Source>
```

In case the influence of the heating plate's temperature should also be fed back to its material properties, an import of the respective value for this would have to be put into execution before another calculation cycle is started as this may influence the current flow and the temperature distribution accordingly. In order to find a stopping criterion, either value being imported can be monitored by the controller. Should the difference of imported values between two cycles fall below the given threshold in the `<Import>` directive, the calculation cycle is halted. `<Threshold>` accepts both percentages and absolute values, and is – as mentioned above – ignored when the models to be calculated are not coupled in both ways.

The given setup was evaluated including such a feedback loop. That means that the temperature that arises within the heating plate has to be obtained and forwarded as input to calculate the heating plate's temperature dependent specific conductivity $\sigma_{(T)}$. Engineering reference compendia list a temperature coefficient α for the electrical resistivity ρ . Therefore, the respective value of $\alpha = 0,4\%K^{-1}$ for the alloy material chosen for the heating plate has to be applied to obtain $\sigma_{(T)}$ as in (3).

$$\rho_{(T)} = \rho_{(T_0)} \cdot (1 + \alpha\Delta T) \quad (2)$$

$$\sigma_{(T)} = \sigma_{(T_0)} \cdot \frac{1}{1 + \alpha\Delta T} \quad (3)$$

The first simulation run per default uses T_0 for determination of $\sigma_{(T)}$. After that, the maximum temperature within the heating plate is obtained and stored for later evaluation. As soon as the difference between the latest two values determined by this drops below 1%, the stopping criterion flag is set which brings the evaluation to an end.

4 RETROSPECT AND OUTLOOK

For the overall software framework design, MATLAB has proven to be accurate and flexible in general. It allows a clear differentiation between software modules and components, uniting them into common data structures, and thus making them accessible for the computation of coupled problems.

Also, the use of XML as common data interchange format has proven to cater to a number of different problem setups. Nevertheless, when pushing forward the software functionalities capabilities, the notation of problem geometry in XML may turn out to be inefficient, maybe even not sufficient to cover all needs.

IGES, the Initial Graphics Exchange Specification, should be mentioned as one of the resources worth investigating. It is a free, well established, manufacturer independent data format that is widely used for data interchange among computer-aided-design products.

The incorporation of a data format like this would also overcome two drawbacks the current framework exhibits. Firstly, it would be possible to also work with complex threedimensional topographic layouts. Secondly, through incorporating a prevalent data format also for internal purposes, interfaces to external programme modules would be

more consistent, and the complexity of and efforts to create wrappers could therefore be reduced.

In addition to that, modern computing equipment with cheap multi-core or multi-processor infrastructure allows to ventilate parallelization of the computational load. The implementation of parallelized algorithms usually contributes significantly to speeding up a numerical simulation and optimization process. Therefore, implementation of appropriate mechanisms ranges among the most desirable goals to be scheduled for future development.

REFERENCES

- [1] B.Cranganu-Cretu, M.Jaindl, A.Köstinger, C.Magele, W.Renhart and J.Smajic, *Multi-Objective Optimization of Shielding Devices for Eddy-Currents using Niching Evolution Strategies*. Proceedings of the 10th OIPE Workshop, Ilmenau, 2008.
- [2] P.Alotto, *Data Structures for mesh-based electromagnetic simulation codes*. Proceedings of the 8th Int. Symposium on Electric and Magnetic Fields, Mondovi, 2009.
- [3] M.Jaindl, A.Reinbacher-Köstinger, R.Kutschera and C.Magele, *Numerical Optimization Framework for Weakly Coupled Multiphysical Problems*. Proceedings of the 14th International IGTE Symposium on Numerical Field Calculation in Electrical Engineering, Graz, 2010.
- [4] *Computer program package EleFAnT2D*. IGTE TUGraz, <http://www.igte.tugraz.at>
- [5] T.Bray, J.Paoli, C.M.Sperberg-McQueen, E.Maler and F.Yergeau, *Extensible Markup Language (XML) 1.0 (Fourth Edition)*, W3 Consortium, <http://www.w3.org/TR/2006/REC-xml-20060816/>
- [6] M.Lutz, *Programming Python*, ISBN 0596009259, O'Reilly Media, Inc., 2006
- [7] X.Jiao, M.T.Campbell and M.T.Heath, *Rocom: an object-oriented, data-centric software integration framework for multiphysics simulations*. Proceedings of the 17th Int. Conference on Supercomputing ICS '03, pp. 358–368, San Francisco 2003.
- [8] P.Zeng, Y.Hao, W.Shao and Y.Liu, *Towards a Software Integration Framework in Product Collaborative Design Environment*. Computer Science and Software Engineering, 2008 International Conference on, pp. 527 -530, 2008.
- [9] A.Nicolet, F.Delilnce, N.Bamps, A.Genon and W.Legros, *A coupling between electric circuits and 2D magnetic field modeling*. Magnetics, IEEE Transactions on, Vol.29/2, pp. 1697-1700, 1993.
- [10] M.F.Sanner, *Python: A Programming Language for Software Integration and Development*. J. Mol. Graphics Mod., Vol.17, pp.5761, 1999.

NUMERICAL MODELLING OF COUPLED ELECTRO-MECHANICAL PROBLEMS FOR THE STATE SPACE CONTROLLER DESIGN

TAMARA N. NESTOROVIC^{*} AND MIROSLAV M. TRAJKOV[†]

^{*} Mechanics of Adaptive Systems, Institute of Computational Engineering
Ruhr-Universität Bochum
Universitätsstraße 150, Geb. IA-01/128, D-44801 Bochum, Germany
e-mail: ^{*}tamara.nestorovic@rub.de, [†]miroslav.trajkov@rub.de

Key words: Coupled electro-mechanical behavior, State space models, Model reduction for the controller design.

Abstract. Model development for coupled electro-mechanical problems in light-weight smart structure design is the subject of this paper. The paper addresses development of reliable models for the controller design of piezoelectric smart structures and systems, within an overall design procedure. Model development is based on the finite element (FE) approach, with application of modal reduction techniques for obtaining the state space models convenient for the controller design. Modal truncation and balanced modal reduction are considered as modal reduction techniques, with regard to controllability and observability issues. From the model optimization and verification point of view the experimental modal analysis and identification issues are addressed as well. Examples of model application to controller design document the feasibility of the technique.

1 INTRODUCTION

Modelling of light-weight smart structures has been a research challenge over the past years. As a part of the overall design of smart structures it represents an important phase in development procedure, which supports all subsequent steps including simulation, controller design, testing and implementation. Smart structures and systems in general comprise integrity (structural and functional) of a structure or a system, multifunctional materials, actuators/sensors and appropriate control in order to achieve desired performances under varying environmental conditions. This paper considers a special type of active multifunctional materials – piezoelectric materials, which belong to the group of ferroelectrics. Due to special properties of the piezoelectric materials, which qualify them for the implementation as actuators and sensors in smart structures, their modelling deserves a special attention. Numerical modelling of smart light-weight structures, which include piezoelectric actuators and sensors, will be treated as a coupled electro-mechanical problem, which integrates the primary mechanical structural behaviour and piezoelectric influence of the active material.

Development of reliable coupled electro-mechanical models for the controller design of piezoelectric smart structures and systems, as a part of an overall design procedure, is the subject of this paper. Special attention is paid to development of state space models, since

they represent the basis for many controller applications [1]. Model development is based on the finite element (FE) approach, and appropriate formulation of the coupled electro-mechanical problem [2], with application of modal reduction techniques for obtaining the state space models convenient for the controller design. Development of reduced state space models based on the FE approach has mostly been treated in literature by applying pure model truncation techniques, which assume retaining only the eigenmodes of interest, assessed by the designer without considering or adopting criteria for the mode selection. This approach does not consider controllability and observability issues within the model reduction algorithm, which represents its lack. The model reduction technique proposed in this paper overcomes the mentioned drawback of the pure model truncation, by introducing weighting of the eigenmodes of interest and consideration of the controllability and observability issues, via appropriate controllability and observability Gramians. In this way a consistent conclusion about the influence of the modes with respect to highest controllability and observability Gramians can be adopted as a criterion for the mode selection. An algorithm for such balanced modal reduction is developed and implemented in this paper for development of reduced order state space models, which are reliable and appropriate for the controller design. As an alternative approach, the subspace-based state space model identification is also presented. The implementation of the techniques is documented by an example of state space model development for coupled electro-mechanical behavior applied to vibration control of a piezoelectric light-weight structure.

2 FINITE ELEMENT BASED REDUCED MODELS

The finite element (FE) based modeling of piezoelectric smart systems and structures represents a suitable basis for the overall simulation and design procedure. This approach results in models which are convenient for the controller design [3-5] and for the analysis of appropriate actuator/sensor placement [6].

The FE analysis is based on the finite element semi-discrete form of the equations of motion of a piezoelectric smart system, which describe its electro-mechanical behavior. These equations can be derived using the established approximation method of displacements and electric potential and the standard finite element procedure [6]. Here the coupled electro-mechanical behavior of smart structures will be considered.

Constitutive equations in the stress-charge form (1) are used for the development of the equations of motion for a smart structure:

$$\boldsymbol{\sigma} = \mathbf{C}\boldsymbol{\varepsilon} - \mathbf{e}\mathbf{E}, \quad \mathbf{D} = \mathbf{e}^T\boldsymbol{\varepsilon} + \boldsymbol{\kappa}\mathbf{E} \quad (1)$$

with following notations:

$\boldsymbol{\sigma}^T = [\sigma_{11} \ \sigma_{22} \ \sigma_{33} \ \sigma_{12} \ \sigma_{23} \ \sigma_{31}]$ mechanical stress vector, $\mathbf{C}_{(6 \times 6)}$ symmetric elasticity matrix, $\boldsymbol{\varepsilon}^T = [\varepsilon_{11} \ \varepsilon_{22} \ \varepsilon_{33} \ 2\varepsilon_{12} \ 2\varepsilon_{23} \ 2\varepsilon_{31}]$ strain vector, $\mathbf{E}^T = [E_1 \ E_2 \ E_3]$ electric field vector, $\mathbf{e}_{(6 \times 3)}$ piezoelectric matrix, $\mathbf{D}^T = [D_1 \ D_2 \ D_3]$ vector of electrical displacement and $\boldsymbol{\kappa}_{(3 \times 3)}$ symmetric dielectric matrix.

The system of equations which describe electro-mechanical behavior consists of the constitutive equations (1) together with the mechanical equilibrium and electric equilibrium

(charge equation of electrostatics resulting from the 4th Maxwell equation):

$$\mathbf{D}_u^T \boldsymbol{\sigma} + \mathbf{P} - \rho \mathbf{v} = \mathbf{0}, \quad \mathbf{D}_\phi^T \mathbf{D} = \mathbf{0} \quad (2)$$

where $\mathbf{P}^T = [P_1 \ P_2 \ P_3]$ represents the body force vector, $\mathbf{v}^T = [v_1 \ v_2 \ v_3]$ is the vector of mechanical displacements, ρ is the mass density and \mathbf{D}_u and \mathbf{D}_ϕ are differentiation matrices:

$$\mathbf{D}_u^T = \begin{bmatrix} \frac{\partial}{\partial x_1} & 0 & 0 & \frac{\partial}{\partial x_2} & 0 & \frac{\partial}{\partial x_3} \\ 0 & \frac{\partial}{\partial x_2} & 0 & \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_3} & 0 \\ 0 & 0 & \frac{\partial}{\partial x_3} & 0 & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_1} \end{bmatrix}, \quad \mathbf{D}_\phi^T = \begin{bmatrix} \frac{\partial}{\partial x_1} & \frac{\partial}{\partial x_2} & \frac{\partial}{\partial x_3} \end{bmatrix}. \quad (3)$$

Variational statement of the governing equations for the coupled electro-mechanical problem derived from the Hamilton's principle represents the basis for development of the FE model [2]. It is obtained in the form:

$$\begin{aligned} & - \int_V (\rho \delta \mathbf{v}^T \ddot{\mathbf{v}} - \delta \boldsymbol{\varepsilon}^T \mathbf{C} \boldsymbol{\varepsilon} + \delta \boldsymbol{\varepsilon}^T \mathbf{e}^T \mathbf{E}) dV + \int_V (\delta \mathbf{E}^T \mathbf{e} \boldsymbol{\varepsilon} + \delta \mathbf{E}^T \boldsymbol{\kappa} \mathbf{E} + \delta \mathbf{v}^T \mathbf{F}_v) dV \\ & + \int_{\Omega_1} \delta \mathbf{v}^T \mathbf{F}_\Omega d\Omega + \delta \mathbf{v}^T \mathbf{F}_p - \int_{\Omega_2} \delta \phi q d\Omega - \delta \phi Q = 0 \end{aligned} \quad (4)$$

where \mathbf{F}_Ω represents the surface applied forces (defined on surface Ω_1), \mathbf{F}_p the point loads, ϕ the electric potential, q the surface charge brought on surface Ω_2 and Q the applied concentrated electric charges. Applying the approximation of displacements and electric potential with the shape functions over an element, representing the structure by a finite number of elements and adding up all elements contributions, the finite element semi-discrete form of the equations of motion is obtained:

$$\mathbf{M} \ddot{\mathbf{q}} + \mathbf{D}_d \dot{\mathbf{q}} + \mathbf{K} \mathbf{q} = \bar{\mathbf{E}} \mathbf{f}(t) + \bar{\mathbf{B}} \mathbf{u}(t) \quad (5)$$

where vector \mathbf{q} represents the vector of generalized displacements including mechanical displacements and electric potential and contains all degrees of freedom. Matrices \mathbf{M} , \mathbf{D}_d and \mathbf{K} are the mass matrix, the damping matrix and the stiffness matrix, respectively. The total load vector is divided into the vector of the external forces $\mathbf{F}_E = \bar{\mathbf{E}} \mathbf{f}(t)$ and the vector of the control forces $\mathbf{F}_C = \bar{\mathbf{B}} \mathbf{u}(t)$, where the forces are generalized quantities including also electric charges. Vector $\mathbf{f}(t)$ represents the vector of external disturbances, and $\mathbf{u}(t)$ is the vector of the controller influence on the structure. Matrices $\bar{\mathbf{E}}$ and $\bar{\mathbf{B}}$ describe the positions of the forces and the control parameters in the finite element structure, respectively.

Model (5) represents a standard FE formulation of equations of motion for a piezoelectric structure. Since FE models in this form often may have more thousands of degrees of freedom (depending on the meshing, i.e. number of finite elements), they are often not suitable for subsequent investigation phases, especially not for the controller design, and therefore a model reduction is required. A technique often used for the model reduction of large flexible vibrating structures is modal truncation, which will be briefly represented in the following.

2.1 Modal truncation

This model reduction technique is based on the modal analysis, which enables determination of structural eigenmodes and eigenfrequencies. Eigenmodes, i.e. eigenvectors, form the modal matrix Φ_m . Modal coordinates \mathbf{z} can be introduced in the following way:

$$\mathbf{q}(t) = \Phi_m \mathbf{z}(t). \quad (6)$$

If (6) is substituted into equation of motion (5), then the following ortho-normalization relations can be applied $\Phi_m^T \mathbf{M} \Phi_m = \mathbf{I}$, $\Phi_m^T \mathbf{K} \Phi_m = \Omega$, $\Delta = \Phi_m^T \mathbf{D}_d \Phi_m$, where Ω represents the spectral matrix and Δ the modal damping matrix. In modal truncation only a limited number of eigenmodes of interest is taken into account. The remaining modes are truncated, which enables model reduction. Introducing the state space vector:

$$\mathbf{x}(t) = [\mathbf{z} \quad \dot{\mathbf{z}}]^T \quad (7)$$

after performing appropriate transformations, the state equation of the modally reduced state space model can be obtained in the form:

$$\dot{\mathbf{x}} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\Omega & -\Delta \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{0} \\ \Phi_m^T \bar{\mathbf{B}} \end{bmatrix} \mathbf{u}(t) + \begin{bmatrix} \mathbf{0} \\ \Phi_m^T \bar{\mathbf{E}} \end{bmatrix} \mathbf{f}(t). \quad (8)$$

Equation (8) corresponds to the state equation in the state space form of a structural model (9), which is convenient for the controller design.

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{E}\mathbf{f}(t) \\ \mathbf{y} &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{F}\mathbf{f}(t) \end{aligned} \quad (9)$$

Controllability and observability issues play an important role in controller design and its application. For state space models of a smart structure, controllability and observability can be proven using appropriate numeric criteria. In literature different criteria can be found, see e.g. [7]. Controllability/observability of single selected modes of interest cannot be strictly assessed for truncated models using standard controllability/observability criteria based on the rank of the controllability and observability matrices. Balanced modal reduction overcomes this drawback.

2.2 Balanced modal reduction

Ranks of the controllability and observability matrices, although relatively simple criteria, provide only an answer to the controllability/observability question in terms “yes” or “no”. As very well known, if the rank of the controllability/observability matrices is equal to the number of states, the model i.e. the realization is controllable/observable. This approach gives good results only for lower system orders, otherwise numerical difficulties may be encountered. Controllability and observability properties of the state space systems can be qualitatively expressed in terms of controllability (\mathbf{P}) and observability (\mathbf{Q}) Gramians, defined in the following way:

$$\mathbf{P} = \int_0^{\infty} e^{\mathbf{A}t} \mathbf{B} \mathbf{B}^T e^{\mathbf{A}^T t} dt, \quad \mathbf{Q} = \int_0^{\infty} e^{\mathbf{A}t} \mathbf{C}^T \mathbf{C} e^{\mathbf{A}^T t} dt. \quad (10)$$

\mathbf{P} and \mathbf{Q} satisfy algebraic Lyapunov linear matrix equations:

$$\mathbf{A}\mathbf{P} + \mathbf{P}\mathbf{A}^T = -\mathbf{B}\mathbf{B}^T, \quad \mathbf{A}^T\mathbf{Q} + \mathbf{Q}\mathbf{A} = -\mathbf{C}^T\mathbf{C}. \quad (11)$$

For an arbitrary transformation of the states by some transformation matrix, appropriate Gramians are obtained, with the property that the eigenvalues of the controllability and observability Gramians products remain invariant. This invariants are the Hankel singular values of the system, and they represent the basis of the balanced model reduction. In balanced realization each state (mode) is equally controllable and observable and the reduced order model is obtained by truncating the least controllable and observable modes. The task of the balanced reduction is actually to find such state transformation, which provides equal controllability and observability of retained modes. In other words, the controllability and observability Gramians of the retained modes are diagonal and equal, and based on this criteria balanced model reduction can be performed.

3 MODAL REDUCTION FOR A SMART PIEZOELECTRIC BEAM

For a piezoelectric beam represented in Figure 1 the finite element model was obtained using the standard FE procedure for modeling piezoelectric materials and mechanical structures.

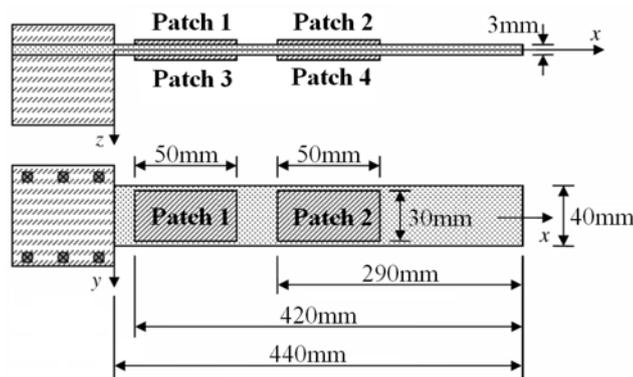


Figure 1: Geometry and layout of the smart structure

The smart structure consists of a cantilever aluminum beam (Young's modulus 70 GPa and density 2.7 g/cm³) and four piezoelectric patches (DuraAct™ P-876.A15), which are attached to the beam, two on each side of the beam. For control purposes these four patches are used as actuators to enable active vibration control of the beam. A scanning digital laser Doppler vibrometer (VH-1000-D), which acts as a sensor, is used to measure the velocity of the bending vibration at the tip of the beam. The sensor provides the feedback signal in the active control algorithm. Among investigations of different MIMO and SISO models, some selected results for the model with one input (piezoelectric actuator excitation) and one output (measured velocity at the tip of the beam) are presented here. The model was obtained using the balanced reduction procedure previously described.

Balanced realization with 20 states was obtained from the original FE model which was reduced by selecting 10 eigenmodes of the beam. A diagram of controllability and observability Gramians for the balanced realization is represented in Figure 2, showing

diagonal elements with highest magnitudes for the states which are most controllable /observable.

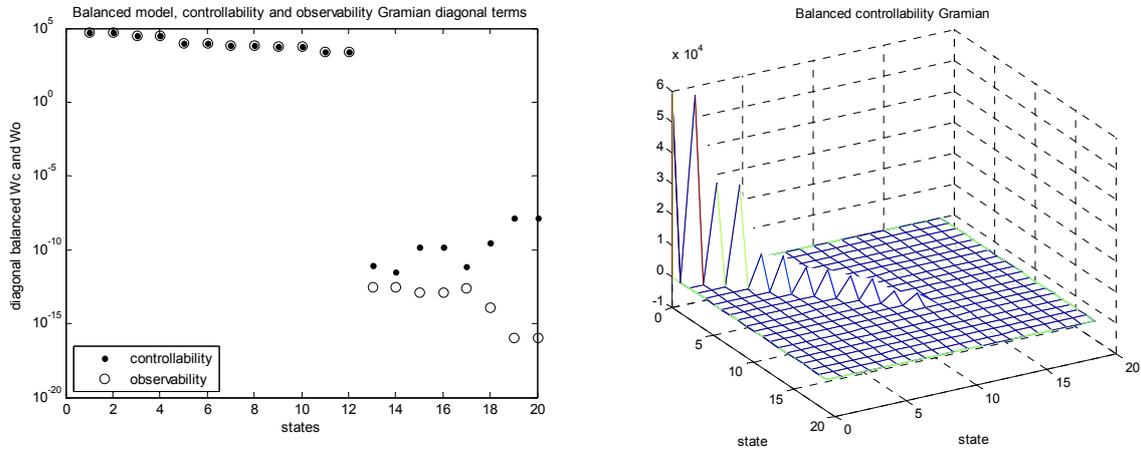


Figure 2: Balanced controllability/observability Gramians

Based on this diagram, selection of the states, which are most controllable and observable at the same time, can be performed. In this case those are obviously first twelve states. Furthermore, the number of the states which should be retained can be selected as well. In this case four eigenmodes are selected. Balanced reduction determines, which of the modes should be retained based on the previously stated controllability and observability criteria for the states. The order of the states in the diagram of grammians does not necessarily correspond to the order of the structural eigenmodes.

Analysis of the balanced reduced order model and its comparison with the original unreduced FE model shows an excellent agreement of frequency responses for the retained modes (Figure 3). It can be clearly seen that the fourth and the sixth modes of the original model are truncated, which was determined based on the balanced reduction and maximal controllability/observability influence of the states corresponding to appropriate structural eigenmodes.

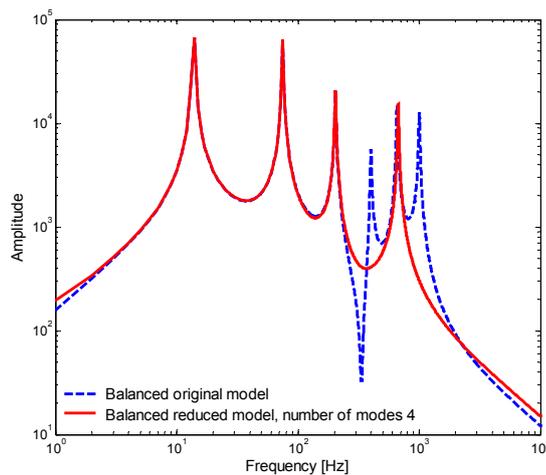


Figure 3: Frequency responses of the original and reduced model

4 MODEL DEVELOPMENT USING SYSTEM IDENTIFICATION

Another possibility for development of reliable smart structural models is system identification. The premise for this approach is the existence of a real structure, which should undergo experimental identification procedure. *Subspace based system identification* was proven to be a very efficient means for the identification of smart structural models, [3]. This procedure is suitable for the identification of discrete-time state space models in the form:

$$\mathbf{x}[k+1] = \mathbf{\Phi}\mathbf{x}[k] + \mathbf{\Gamma}\mathbf{u}[k], \quad \mathbf{y}[k] = \mathbf{C}\mathbf{x}[k] + \mathbf{D}\mathbf{u}[k]. \quad (12)$$

The task of the subspace identification is to determine the order n of the unknown system as well as the system matrices $\mathbf{\Phi} \in \mathbb{R}^{n \times n}$, $\mathbf{\Gamma} \in \mathbb{R}^{n \times m}$, $\mathbf{C} \in \mathbb{R}^{l \times n}$, $\mathbf{D} \in \mathbb{R}^{l \times m}$. Input and output measurement data are organized in the form of the following input-output relation:

$$\mathbf{Y}[k] = \mathbf{\Gamma}_\alpha \mathbf{x}[k] + \mathbf{\Phi}_\alpha \mathbf{U}[k], \quad (13)$$

where $\mathbf{\Gamma}_\alpha$ represents the observability matrix for the system (13), and $\mathbf{\Phi}_\alpha$ is the Toeplitz matrix of impulse responses from \mathbf{u} to \mathbf{y} . A detailed subspace identification procedure is described in [1,3]. Comparison of the frequency responses obtained for identified models using subspace identification method with different model orders is represented in Figure 4.

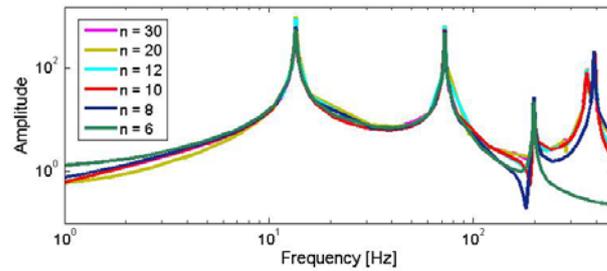


Figure 4: Frequency responses of identified models with different orders

5 CONTROLLER IMPLEMENTATION RESULTS

State space models obtained by previously described procedures can be successfully implemented in the controller design for vibration suppression. Development of reliable models, which enable an overall design of smart structures including controller design, plays a very important role for further application steps.

Active control of the piezoelectric beam represented in Figure 1 was performed by applying a discrete-time negative feedback control in the form:

$$\mathbf{u}[k] = -\mathbf{L}\mathbf{x}[k] \quad (14)$$

where \mathbf{u} represents control voltages applied to the piezoelectric patches, and the state variables \mathbf{x} are estimated using an observer (Kalman filter), [8]. The feedback gain matrix \mathbf{L} is determined for an optimal linear quadratic regulator (LQR). The controller design task is to determine the control law which minimizes the performance index:

$$J = \sum_{k=0}^{\infty} (\mathbf{x}[k]^T \mathbf{Q}\mathbf{x}[k] + \mathbf{u}[k]^T \mathbf{R}\mathbf{u}[k]) \quad (15)$$

where the matrices \mathbf{Q} and \mathbf{R} are the designer specified symmetric positive definite weighting matrices.

Implementation of the controllers, designed based on developed state space structural models with the purpose of vibration suppression is illustrated in the following figures. Figure 5 represents the sensor signal (velocity at the tip of the beam) in the presence of a random noise; the controller is switched on after 2 seconds. Active control of the piezoelectric beam results in significant reduction of the vibration magnitudes in comparison with uncontrolled case.

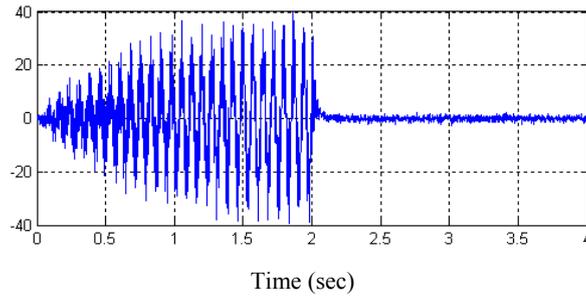


Figure 5: Velocity of a point at the tip of the beam (uncontrolled and controlled using piezoelectric patches)

Successful performance of the controlled system is also demonstrated for the case of the initial displacement disturbance. Free vibrations of the beam caused by an initial displacement applied to the tip of the beam are comparable with impulse disturbance vibrations.

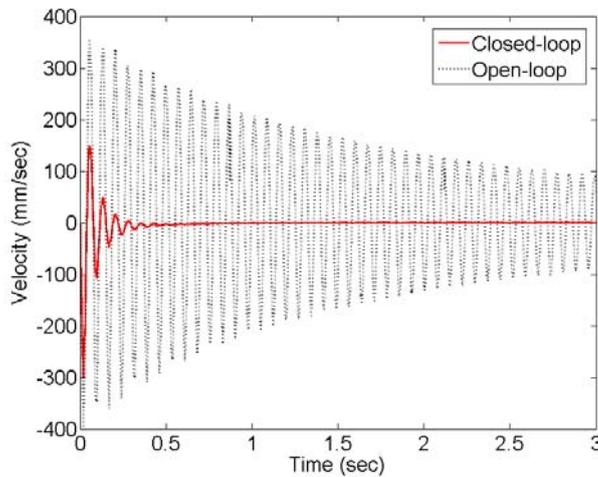


Figure 6: Free vibration response (velocity) of the controlled and uncontrolled system

The free vibration response (velocity) of the open-loop and closed-loop system subjected to an initial displacement of is represented in Figure 6. Designed controller attenuates significantly the magnitudes of the free end displacement [8].

In both cases controller development and application were performed owing to reliable reduced state space models.

6 CONCLUSIONS

In this paper the modeling of coupled electro-mechanical behavior of piezoelectric light-weight structures is considered. Numerical modeling based on finite element procedure is used to obtain the reduced state space models, which are convenient for the controller design. For the model reduction the balanced modal reduction is employed, which is superior in comparison with the pure modal truncation, since it provides necessary information on model controllability and observability issues. Besides, the numerical modeling using the subspace based model identification is also presented, as a technique for reliable modeling in case when a real structure or a prototype is available. State space models obtained applying proposed techniques provide a suitable basis for all subsequent steps in the overall design and implementation of actively controlled smart structures. The feasibility of the models for the controller design was demonstrated on example of a flexible cantilever beam with piezoelectric patches. Controllers designed based on the developed state space models perform significant vibration reduction in comparison with the uncontrolled cases.

ACKNOWLEDGEMENT

Parts of the research presented in the paper are supported by the German Research Foundation Project DFG-837. This support is highly appreciated.

REFERENCES

- [1] Nestorović, T. *Controller Design for the Vibration Suppression of Smart Structures*. Fortschritt-Berichte VDI, Vol. 8, Nr. 1071, Düsseldorf: VDI Verlag (2005)
- [2] Tiersten H.F. Hamilton's principle for linear piezoelectric media. *Proc. IEEE* (1967) 1523–1524.
- [3] Nestorović-Trajkov, T. and Gabbert, U. Active control of a piezoelectric funnel-shaped structure based on subspace identification. *Structural Control and Health Monitoring* (2006) **13** (6): 1068–1079.
- [4] Nestorović-Trajkov, T. Köppe, H. and Gabbert, U. Active Vibration Control Using Optimal LQ Tracking System with Additional Dynamics. *Int. J Control* (2005) **78** (15): 1182–1197.
- [5] Nestorović, T. and Trajkov, M. Development of Reliable Models for the Controller Design of Smart Structures and Systems, *The International Conference "Mechanical Engineering in XXI Century" – Proceedings* University of Niš, Faculty of Mechanical Engineering, Niš, Serbia, (2010) pp. 263-266, ISBN: 978-86-6055-008-0
- [6] Gabbert, U., Lefèvre, J., Laugwitz, F. and Nestorović, T. Modelling and analysis of piezoelectric smart structures for vibration and noise control, *International Journal of Applied Electromagnetics and Mechanics* (2009) **31** (1): 29-39
- [7] Kailath, T. *Linear Systems*. Prentice Hall (1980)
- [8] Durrani, N. and Nestorović, T. Active Vibration Control of Piezoelectric Structures. *DAGA 2011 (37th annual convention for acoustics)* 21-24 March 2011, Düsseldorf (submitted, Paper Nr. DAGA2011/14)

PRECISE ELECTROMAGNETIC-THERMOELASTIC ACTUATOR FOR LASER TECHNOLOGIES

IVO DOLEŽEL[†], VÁCLAV KOTLAN^{*} AND BOHUŠ ULRYCH^{*}

[†] Institute of Thermomechanics, v. v. i.
Academy of Sciences of the Czech Republic
Dolejskova 5, 182 00 Praha 8
email: dolezel@it.cas.cz

^{*} Faculty of Electrical Engineering
University of West Bohemia
Univerzitní 16, 306 14 Plzeň
e-mail: {vkotlan, ulrych}@kte.zcu.cz

Key words: Thermoelastic Actuator, Nonlinear Coupled Problem, Numerical Analysis, Electromagnetic Field, Temperature Field, Field of Thermoelastic Displacements.

Abstract. A new type of a combined electromagnetic-thermoelastic actuator is presented. The device is intended for accurate setting of position. Its mathematical model represents a nonlinear coupled problem characterized by the interaction of electromagnetic field, temperature field and field of thermoelastic displacements. Its numerical solution is illustrated on a typical example whose results are discussed.

1 INTRODUCTION

Accurate setting of position belongs to standard requirements in numerous laboratory and industrial processes and practically is realized on the basis of several physical (for example, mechanical, hydraulic, pneumatic, electromagnetic, and other) principles. But even when such systems are robust and reliable, they always exhibit one disadvantage consisting in the presence of movable parts (micrometric screws, pistons, plungers, etc.), whose movement usually negatively affects the available accuracy.

This is also the case of classical electromagnetic actuators [1–3]. They usually work with the movement of a plunger that is controlled magnetically, by the field current. Their accuracy does not exceed 10^{-4} – 10^{-3} m. Several years ago, the authors proposed a thermoelastic actuator [4–5] working with extremely small shifts (10^{-6} – 10^{-3} m) of its dilatation element. Its accuracy is very high, but on the other hand, there is no possibility of any additional movement, so that the range of possible positions is equal to the mentioned thermoelastic shift. A combination of both above systems presented in this paper results in a device providing a possibility of accurate wide-range setting of position.

2 DESCRIPTION OF THE DEVICE AND ITS OPERATION

The basic arrangement of the combined actuator is depicted in Fig. 1. Its working regime consists of two steps: electromagnetic operation and consequent thermoelastic operation.

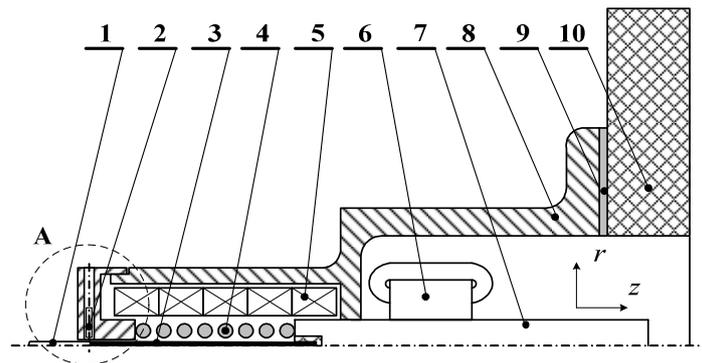


Figure 1. A combined electromagnetic-thermoelastic actuator:

- 1–ceramic part of the plunger, 2–electromagnetic clutch I, 3–nonferromagnetic metal part of the plunger, 4–return spring, 5–field coil consisting of several sections, 6–electromagnetic clutch II, 7–ferromagnetic part of the plunger, 8–ferromagnetic shell, 9–thermal insulation, 10–stiff frame or wall

In the course of the *electromagnetic operation*, the field coil **5** consisting of generally more sections carries direct current $I_{\text{ext,DC}}$ of density $J_{\text{ext,DC}}$, that generates in the system quasistationary magnetic field B_{DC} . This field produces magnetic force F_m acting on the ferromagnetic part **7** of the plunger and pulling it into the field coil. In order that the plunger starts moving, the force F_m must be higher than the sum of the force F_s produced by the return spring **5** and possible external force F_{ext} . After reaching the desired shift ζ_d (which, however, is only approximate), the electromagnetic friction clutch **6** switches on and the field coil **5** is disconnected from the DC source. Now the rough position of the nonferromagnetic part of the plunger **2** is fixed, but with a certain error.

During the thermoelastic operation, selected sections of the coil **5** start carrying harmonic current of density $J_{\text{ext,AC}}$ and frequency f . This current generates in the device a periodic magnetic field B_{AC} . This field induces in the plunger (mostly in its nonferromagnetic part **3**) eddy currents of density J_{eddy} that cause its heating and consequent dilatation Δu with respect to its fixed part specified by the friction surface of the clutch **6**. This shift may also be fixed by another friction clutch **2**.

The aim of the paper is to present the complete mathematical model of the device (a triply coupled task characterized by the interaction of electromagnetic field, temperature field and field of thermoelastic displacements), numerical solution of its computer model and to discuss the results.

3 CONTINUOUS MATHEMATICAL MODEL OF THE SYSTEM

The arrangement of the device may be (with a small error given by the presence of the friction clutch) considered axisymmetric. Its complete model consists of several partial models providing the results for both above regimes.

For the electromagnetic operation we need to know the static and dynamic characteristics of the device. The static characteristic is represented by the dependence of the electromagnetic force F_m on the shift ζ of the plunger, see Fig. 2).

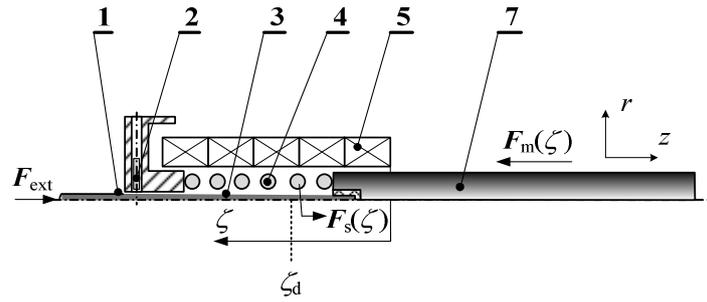


Figure 2. To the computation of the static characteristic $F_m = F_m(\zeta)$ of the device (the structural parts correspond to the parts in Fig. 1, symbol ζ_d denotes the desired position)

For the thermoelastic operation we have to determine the distribution of periodical magnetic field in the system, temperature rise of the plunger (particularly its nonferromagnetic part **3**) and its dilatation.

3.1 Electromagnetic operation

Magnetic field in the system (for any position ζ of the plunger) may be described by the partial differential equation (PDE) for the magnetic vector potential A in the form [6]

$$\operatorname{curl} \left(\frac{1}{\mu} \operatorname{curl} A \right) = \mathbf{J}_{\text{ext,DC}}, \quad (1)$$

where μ is the magnetic permeability. The magnetic force F_m acting on the plunger is then expressed by the integral

$$\mathbf{F}_m = \frac{1}{2} \oint_{S_p} [\mathbf{H}(\mathbf{n} \cdot \mathbf{B}) + \mathbf{B}(\mathbf{n} \cdot \mathbf{H}) - \mathbf{n}(\mathbf{H} \cdot \mathbf{B})] dS, \quad (2)$$

where \mathbf{H} and \mathbf{B} are vectors of the magnetic field, \mathbf{n} denotes the unit vector of the outward normal, and finally, S_p is the whole surface of the plunger (consisting of parts **3** and **7**). The force F_m depends on the instantaneous position of the plunger and can be controlled by the external field current $I_{\text{ext,DC}}$. During the movement up to the desired position ζ_d there must hold $|F_m(\zeta)| \geq |F_s(\zeta) + F_{\text{ext}}|$, where $F_s(\zeta)$ depends on the actual compression of the return spring **4**.

The dynamic characteristic of the plunger $\zeta = \zeta(t)$ follows from the solution of two ordinary nonlinear differential equations

$$\begin{aligned} m \frac{dv}{dt} &= F_m(\zeta) - F_s(\zeta) - F_{\text{ext}} - F_f(\zeta), & \frac{d\zeta}{dt} &= v, \\ v(0) &= 0, & \zeta(0) &= 0, \end{aligned} \quad (3)$$

where $F_f(\zeta)$ is a possible friction force.

As the above mechanical transient is very short (on the order of 10^{-2} – 10^{-1} s), the temperature rise of the system due to the Joule losses in the field coil is very low and in practical computations can be neglected.

3.2 Thermoelastic operation

Its model represents a triply coupled task characteristic by an interaction of the electromagnetic field, temperature field and field of thermoelastic displacements.

Distribution of the *periodical electromagnetic field* in the actuator is given by the solution of the well-known parabolic equation for magnetic vector potential \mathbf{A} in the form [6]

$$\operatorname{curl}\left(\frac{1}{\mu}\operatorname{curl}\mathbf{A}\right)+\gamma\frac{\partial\mathbf{A}}{\partial t}=\mathbf{J}_{\text{ext,AC}}, \quad (4)$$

where μ denotes the magnetic permeability, γ the electric conductivity and \mathbf{J}_{ext} the vector of the external harmonic current density in the field coil. But solution to (4) would require a lot of time because the quasi-steady (periodical) state is reached only after about 10 periods of the field current (which would represent at least 10^3 time steps). That is why the model was somewhat simplified using the assumption that the magnetic field is harmonic. In such a case it can be described by the Helmholtz equation for the phasor $\underline{\mathbf{A}}$ of the magnetic vector potential \mathbf{A}

$$\operatorname{curl}\operatorname{curl}\underline{\mathbf{A}}+\mathbf{j}\cdot\omega\gamma\mu\underline{\mathbf{A}}=\mu\underline{\mathbf{J}}_{\text{ext,AC}}, \quad (5)$$

where ω is the angular frequency. But the magnetic permeability μ of ferromagnetic parts is not supposed constant; it is always assigned to the local value of magnetic flux density \mathbf{B} (that is, due to the previous assumption, harmonic as well). Its computation is based on an iterative procedure. The conditions along the axis of the device and artificial boundary placed at a sufficient distance from the system are of the Dirichlet type ($\underline{\mathbf{A}}=\mathbf{0}$).

As the device may be considered axisymmetric, the magnetic vector potential \mathbf{A} has only the circumferential component A_α , while the magnetic flux density \mathbf{B} has two components B_r and B_z .

The *nonstationary temperature field* in the system is described by the heat transfer equation [7]

$$\operatorname{div}(\lambda\cdot\operatorname{grad}T)=\rho c\frac{\partial T}{\partial t}-p \quad (6)$$

where λ is the thermal conductivity, ρ denotes the mass density, and c stands for the specific heat (all of these parameters are generally temperature-dependent functions). Finally, symbol p denotes the time average internal volume sources of heat that generally consist of the volumetric Joule losses p_J due to the induced eddy currents, and volumetric magnetization losses p_m . Thus,

$$p=p_J+p_m, \quad (7)$$

where

$$p_J = |\underline{\mathbf{J}}_{\text{eddy}}|^2 / \gamma, \quad \underline{\mathbf{J}}_{\text{eddy}} = \mathbf{j} \cdot \omega \gamma \underline{\mathbf{A}}, \quad (8)$$

while p_m are determined from the known measured loss dependence $p_m = p_m(|\underline{\mathbf{B}}|)$ for the used material. The boundary condition generally respects both convection and radiation, but its particular application depends on the case solved.

The solution of the *thermoelastic problem* may be performed in several manners. After some research in the domain we decided to use the Lamé equation for the vector of displacements \mathbf{u} that reads [8]

$$(\varphi + \psi) \cdot \text{grad}(\text{div} \mathbf{u}) + \psi \cdot \Delta \mathbf{u} - (3\varphi + 2\psi) \cdot \alpha_T \cdot \text{grad} T + \mathbf{f} = \mathbf{0}, \quad (9)$$

where $\varphi > 0$, $\psi > 0$ are coefficients associated with material parameters by the relations

$$\varphi = \frac{\nu \cdot E}{(1 + \nu)(1 - 2\nu)}, \quad \psi = \frac{E}{2(1 + \nu)}.$$

Here E denotes the modulus of elasticity and ν the Poisson coefficient of the transverse contraction. Finally, $\mathbf{u} = (u_r, u_\varphi, u_z)$ represents the displacement vector, α_T the coefficient of the linear thermal dilatability of material and \mathbf{f} the vector of the internal volumetric forces. The boundary conditions depend on the arrangement solved.

4 NUMERICAL COMPUTATION

The mathematical model presented in the previous paragraph was solved by the finite element method in the quasi-coupled formulation using the codes QuickField [9] (magnetic field) and COMSOL Multiphysics [10] (temperature field and field of thermoelastic displacements) supplemented with a lot of own procedures and scripts. The last two fields are supposed not to affect the primary magnetic field. Particular attention was paid to the convergence of results (in the case of the thermoelastic operation) in the dependence of the position of the artificial boundary (periodical magnetic field) and density of the discretization mesh (all three physical fields). In order to obtain values with three valid digits for the magnetic field, the corresponding mesh must consist of more than 150 000 quadratic elements; for the two remaining fields this number could be substantially lower (the corresponding definition areas are smaller and both fields are relatively smooth) – about 60 000 elements.

5 ILLUSTRATIVE EXAMPLE

For an illustration, we will present the most important results for the device whose principal dimensions are depicted in Fig. 3. This device does not contain the magnetic clutches **2**, **6** (but their presence influences the main results by not more than about 2–3 %), and ceramic part of the plunger, so that the arrangement is fully axisymmetric.

The nonferromagnetic part of the plunger **1** is made of zinc, the return spring **2** is made of spring steel DIN 17 221, the field coil **3** is made of Cu, the shell **4** and ferromagnetic part **6** of the plunger are made of carbon steel CSN 12 040, and thermal insulation **5** is made of Teflon.

The basic physical parameters of the mentioned materials including their temperature de-

dependencies were found in [11]. For example, Fig. 4 depicts the magnetization characteristic of steel 12 040 (left part) and temperature dependence of its thermal conductivity (right part).

The air inside the device (representing the ambient medium) is supposed unmoving, its thermal conductivity being 0.03 W/mK. The external and friction forces (F_{ext} and F_f from (3)) are equal to zero.

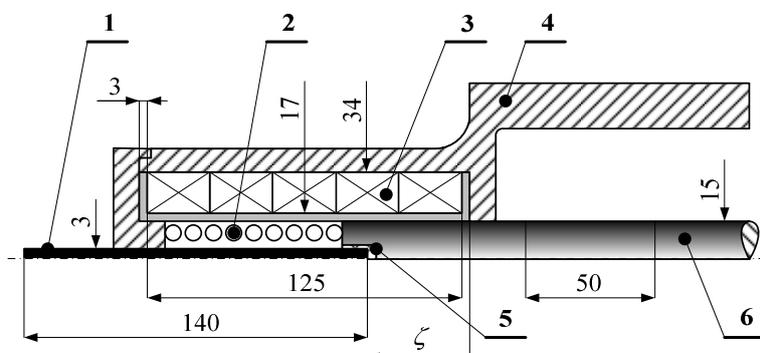


Figure 3. The main part of the investigated device with the principal dimensions in mm: 1–nonferromagnetic part of the plunger, 2–return spring, 3–field coil of the actuator, 4–ferromagnetic shell, 5–connecting Teflon matrix, 6–ferromagnetic part of the plunger

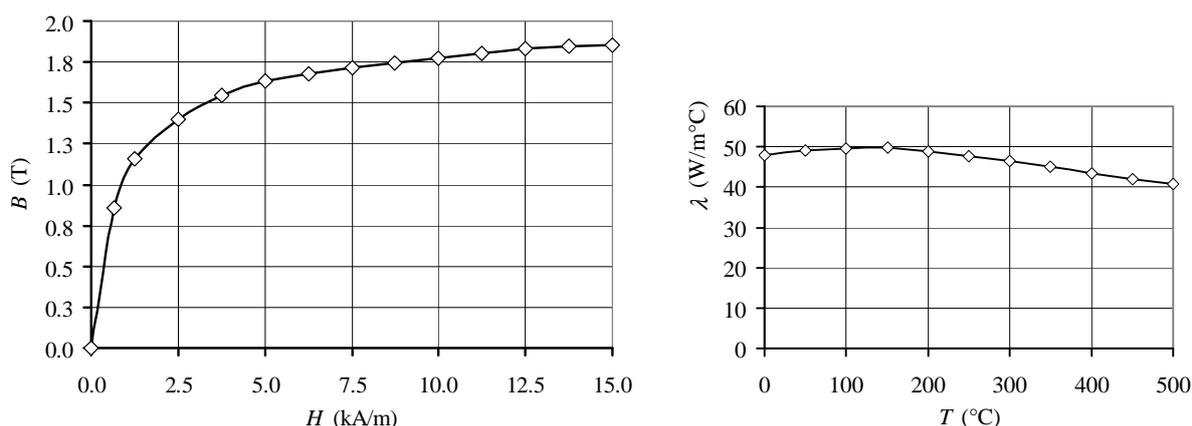


Figure 4. Magnetization characteristic of carbon steel 12 040 (left) and temperature dependence of its thermal conductivity λ (right)

The computations provided a lot of qualitative and quantitative results. The qualitative results are represented by maps of all three fields, but they are rather of low importance. Much more important are the qualitative results depicted in the following figures.

Figure 5 shows the dependencies of the magnetic force F_m on shift ζ for two different values of DC field current density $J_{\text{ext,DC}}$ and a similar dependence of the counterforce $F_s(\zeta)$ exerted by the return spring. It is clear that for the selected return spring the current density $J_{\text{ext,DC}} = 10^6 \text{ A/m}^2$ is too low to overcome the counterforce $F_s(\zeta)$ (the plunger would not move at all). But DC field current density $J_{\text{ext,DC}} = 2 \times 10^6 \text{ A/m}^2$ is already enough for the movement of the plunger in any range of shifts ($F_m > F_s$ everywhere).

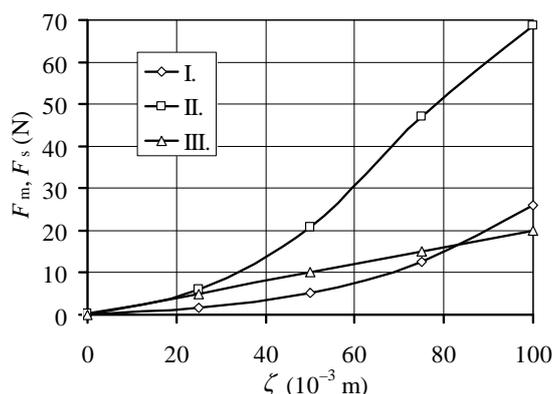


Figure 5. Dependencies of the magnetic force F_m and counterforce F_s exerted by the return spring on shift ζ

I - $F_m(\zeta)$ for DC field current density $J_{\text{ext,DC}} = 10^6$ A/m²

II - $F_m(\zeta)$ for DC field current density $J_{\text{ext,DC}} = 2 \times 10^6$ A/m²

III - $F_s(\zeta)$ for the constant of elasticity of the spring $k = 0.2$ N/m

The same conclusion can be seen in Fig. 6 that depicts the resultant static characteristics of the electromagnetic operation (defined as $F_{\text{tot}} = F_m - F_s$) for three different DC field current densities and the same elasticity constant of the return spring. For field current density - $J_{\text{ext,DC}} = 5 \times 10^6$ A/m² we obtain an unreal characteristic, because one its part lies below zero.

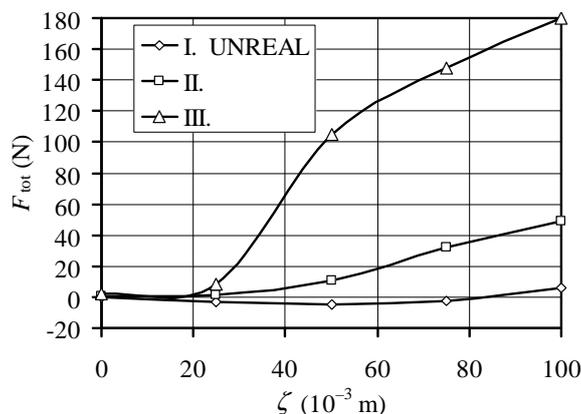


Figure 6. Static characteristics of the device for

I - $J_{\text{ext,DC}} = 10^6$ A/m², II - $J_{\text{ext,DC}} = 2 \times 10^6$ A/m², III - $J_{\text{ext,DC}} = 5 \times 10^6$ A/m²

Figure 7 shows the dependence of the total thermoelastic shift u_z of the leftmost point of the zinc part of the plunger on its average temperature T_{av} , the previous electromagnetic shift being $\zeta = 25$ mm. The dependence is almost linear. A similar graph is shown in Fig. 8 depicting the dependence of the average temperature rise ΔT of the zinc part on the shift u_z of its leftmost end (the previous electromagnetic shift ζ being 25 mm again). This dependence may be very useful for an accurate determination of this shift (because the average temperature rise of the zinc part may be well measured either indirectly or by thermocouples).

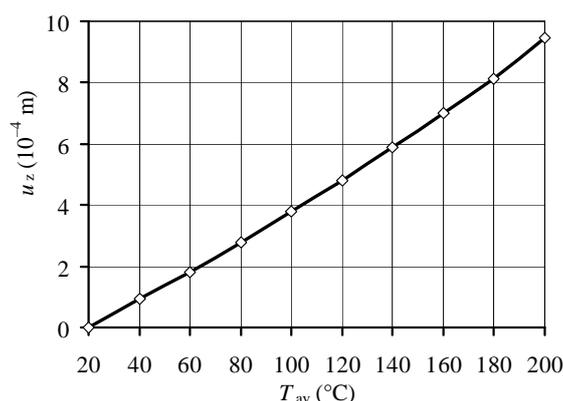


Figure 7. Dependence of the total thermoelastic shift u_z of the leftmost end of the zinc part on its average temperature T_{av} for $\zeta = 25$ mm

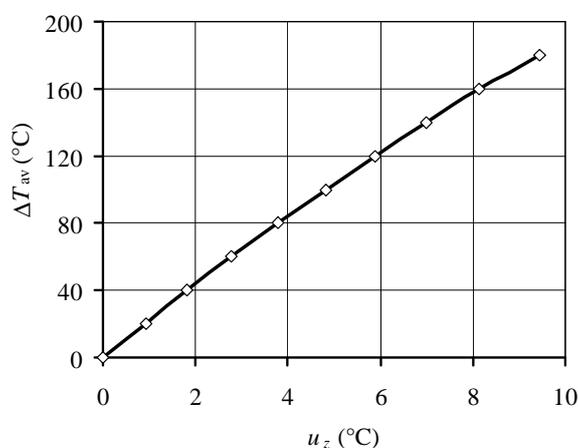


Figure 8. Dependence of the average temperature rise ΔT_{av} of the zinc part on shift u_z of the leftmost end of this zinc part for $\zeta = 25$ mm

Of great importance are also the dynamic characteristics of the system, because the thermoelastic operation can take tens of seconds or even minutes. Figure 9 shows the time evolution of the average temperature T_{av} of the zinc part of the plunger for three different field current densities $J_{ext,AC}$ of frequency $f = 5$ kHz, shift ζ being 25 mm. And Fig. 10 shows the corresponding time evolutions of the shift u_z of the leftmost end of the zinc part of the plunger, the other parameters remaining the same.

Of course, the temperature in the system must not exceed 200 °C due to thermal endurance of used insulation elements. This is also marked in both Figs. 9 and 10 by the unreal zones.

But even these results are far from being complete, because the figures show the graphs just for several selected parameters. Some more results obtained for other parameters are listed in Tab. 1.

6 CONCLUSION

The device exhibits a lot of possibilities of practical applications in optics (setting of focal distances of lens systems), laser technologies (setting of position of the laser head or focusing

of the laser beam) or microscope technologies (setting of positions of specimens with respect to the focus of light or electron beams). Next work in the domain will be aimed at further acceleration of computations and detailed analysis of behavior of other parts of the device (mainly both clutches).

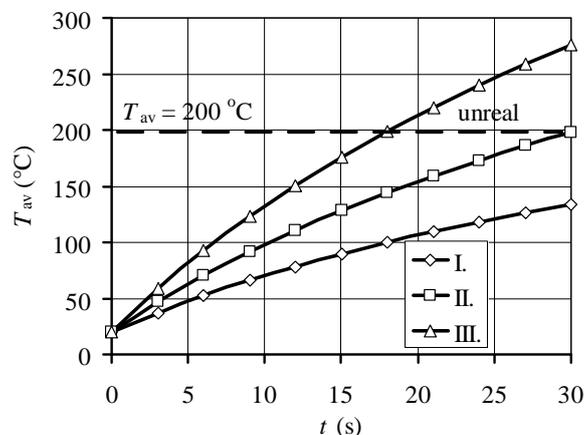


Figure 9. Time evolution of the average temperature T_{av} of the zinc part of the plunger for three different field current densities ($\zeta = 25$ mm, $f = 5$ kHz):

$$\text{I} - J_{\text{ext,AC}} = 4 \times 10^6 \text{ A/m}^2, \text{ II} - J_{\text{ext,AC}} = 5 \times 10^6 \text{ A/m}^2, \text{ III} - J_{\text{ext,AC}} = 6 \times 10^6 \text{ A/m}^2$$

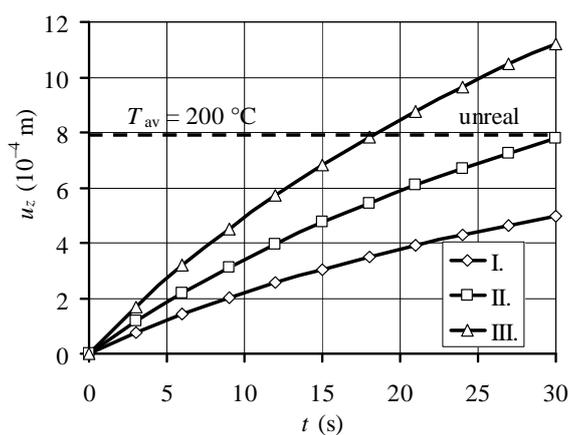


Figure 10. Time evolution of the shift u_z of the leftmost part of the zinc part of the plunger for three different field current densities ($\zeta = 25$ mm, $f = 5$ kHz):

$$\text{I} - J_{\text{ext,AC}} = 4 \times 10^6 \text{ A/m}^2, \text{ II} - J_{\text{ext,AC}} = 5 \times 10^6 \text{ A/m}^2, \text{ III} - J_{\text{ext,AC}} = 6 \times 10^6 \text{ A/m}^2$$

7 ACKNOWLEDGMENT

This work was supported by the European Regional Development Fund and Ministry of Education, Youth and Sports of the Czech Republic under the project No. CZ.1.05/2.1.00/03.0094: Regional Innovation Centre for Electrical Engineering (RICE), by Grant project GA CR 102/09/1305 and Research Plan MSM 6840770017.

Table 1. Results for other parameters (symbol p denoting the volumetric losses in the zinc part of the plunger)

f (Hz)	P (W/m ³)	T_{av} (°C)	$u_{z,max}$ (m)
$J_{ext,AC} = 2 \times 10^6 \text{ A/m}^2$			
50	1.088×10^3	29.60	4.20×10^{-5}
500	1.062×10^5	32.55	5.40×10^{-5}
5000	4.977×10^6	115.50	4.14×10^{-4}
$J_{ext,AC} = 5 \times 10^6 \text{ A/m}^2$			
50	6.800×10^3	76.86	2.47×10^{-4}
500	6.637×10^5	98.33	3.40×10^{-4}
5000	3.110×10^7	870.00	2.59×10^{-3}

The last row in the table is unreal because of unacceptably high temperature.

REFERENCES

- [1] Janocha, H. *Actuators, basics & applications*. Springer, New York, 2004.
- [2] Busch-Vishniac, I. J. *Electromechanical sensors and actuators*. Springer, Berlin, 1998.
- [3] Brauer, J. R. *Magnetic actuators and sensors*. Wiley, New York, 2006.
- [4] Dolezel, I., Karban, P., Ulrych, B., Pantelyat, M., Matyukhin, Y. and Gontarowskiy, P. Computer model of thermoelastic actuator solved as coupled contact problem. *COMPEL* (2007) **26**, No. 4: pp. 1063–1072.
- [5] Dolezel, I., Karban, P., Ulrych, B., Pantelyat, M., Matyukhin, Y., Gontarowskiy, P. and Shulzhenko, N. Limit operation regimes of actuators working on principle of thermoelasticity. *IEEE Trans. Magn.* (2008) **44**, No. 6: pp. 810–813.
- [6] Stratton, A. J. *Electromagnetic theory*. McGraw Hill, NY, 2007.
- [7] Holman, J. P. *Heat transfer*. McGraw Hill, NY, 2002.
- [8] Boley, B. and, Wiener, J. *Theory of Thermal Stresses*, Wiley, New York, NY, 1960.
- [9] www.quickfield.com.
- [10] www.mathworks.com.
- [11] Database of material parameters www.jahm.com.

NUMERICAL ANALYSIS OF DIELECTRIC MICRO-PARTICLE MOTION IN A FLUID AND ELECTRIC FIELD

MYUNG KI BAEK^{*}, YOUNG SUN KIM[†] AND IL HAN PARK^{*}

^{*} School of Information and Communication Engineering
Sungkyunkwan University
Suwon 440-746, Korea
e-mail: myungki79@naver.com, laem.skku.ac.kr

[†] Department of Electrical Engineering and Computer Science,
Massachusetts Institute of Technology,
Cambridge, MA, 02139, USA
email: youngsun@mit.edu

Key words: DEP force, Micro-Particle, Coupled Problem, Numerical Analysis, Particle Separator.

Abstract. We present numerical analysis of a coupled problem composed of fluidics, electromagnetic and particle dynamics. The forces acting on the dielectric micro-particle consist of a dielectrophoretic(DEP) force, drag force and gravitational force in the proposed analysis model. DEP force and drag force are calculated using the distribution of the electric field and fluid velocity field to analyze the characteristic of the micro-particle motion. The forces exerted by each field are driving terms in the Newton's equation for particle motion. The designed particle separating device, which has the one inlet and the two outlets, is simulated to validate proposed numerical scheme. The analysis results show the trace of the micro-particles can be analyzed using the proposed numerical approach.

1 INTRODUCTION

The electric force acting on dielectric material under non-uniform electric field is called dielectrophoretic(DEP) force. Recently, dielectric micro-particles, whose motions are controlled by the DEP force, play an increasing role in various areas, ranging from environmental engineering to biomedical fields [1]. For example, minute particles caused during a manufacturing process have a detrimental influence on people's health, and can damage modern equipment consisting of micro scale devices. These pollutants can be manipulated using the DEP force. Many particle control devices have been developed.

The dielectric micro-particles existing in the air experience many forces such as electric, drag, gravitational and buoyancy [2]. Among them, the dominant forces are electric force, drag force and gravity since buoyancy is very small relative to gravity. The electric force is related to both dielectric material properties and non-uniform electric field. Drag force is determined by fluid velocity. Last, gravity is associated with physical constants: volume, density and particle radius.

In this paper, we propose a numerical procedure to solve the coupled problem, and design a separating device. The micro-beads in air can be controlled by both fluid stream and electric force. That is, the motion of the particles is determined by the distributions of an electric field and fluid velocity. The characteristic of the particle motion is a form of coupled problem of electromagnetic, fluidics and particle dynamics [3].

The electric field is calculated using the finite element method for the electrostatics, and the fluid velocity field is calculated also using the finite element method for the potential flow to calculate numerically the DEP force and fluid drag force. The gravitational force can be obtained easily from material properties. The three forces are summed to be a driving force of Newton's equation to calculate particle motion. The Runge-Kutta method is used to numerically solve the coupled motional equation.

The micro-particles suspended in laminar flow experience positive DEP force in the designed device based on their relative permittivity and electric field intensity. The movement characteristics of the micro-particles are analyzed using the proposed numerical analysis procedure. Simulation results show validity of the proposed method and feasibility of the designed device.

2 MODELING OF PARTICLE MOTION

The force acting on dielectric material suspended in a non-uniform electric field is generated by interaction of imposed with the induced dipole moment. The DEP force on the dielectric particle, which is lossless in a DC electric field, can be obtained using the effective dipole moment that follows as:

$$\mathbf{p} = 4\pi\epsilon_1 R^3 K \mathbf{E}_0 \quad (1)$$

where ϵ_1 , R , and \mathbf{E}_0 are fluid permittivity, particle radius and applied electric field, respectively, and K , known as the Clausius-Mossotti(CM) factor, can be expressed as

$$K = \frac{\epsilon_2 - \epsilon_1}{\epsilon_2 + 2\epsilon_1} \quad (2)$$

where ϵ_2 is particle permittivity. Using the above equations, the well-known expression for the DEP force can be written as [4]:

$$\mathbf{F}_{\text{DEP}} = 2\pi\epsilon_1 R^3 K \nabla E_0^2 \quad (3)$$

According to (2), the CM factor, whose value range is in $-0.5 \leq K \leq 1.0$ range, provides a measure of the magnitude of the DEP force and the particle's direction of motion. If fluid permittivity is larger than that of the particle, then it is called the negative DEP force due to the CM factor.

The moving particles suspended in fluid experience the drag force as follows [2]:

$$\mathbf{F}_D = 6\pi\eta(\mathbf{u} - \mathbf{v})R \quad (4)$$

where η , \mathbf{u} and \mathbf{v} are viscosity, fluid velocity and particle velocity, respectively. The drag force can be calculated by the relative velocity between fluid and particle velocity.

These two forces on the particles are substituted into Newton's motional equation of motion to evaluate the movement characteristic of the particles:

$$\mathbf{F}_{DEP} + \mathbf{F}_D + \mathbf{F}_G = m \frac{d\mathbf{v}}{dt} \quad (5)$$

where m is particle mass, \mathbf{F}_G is gravity force. Equation (5) can be solved using the Runge-Kutta method.

3 ANALYSIS OF PARTICLE DYNAMICS

3.1 Analysis Model

The DEP force and the drag force are considered using three design variables the design the device. First, from the material standpoint, the analysis region is divided into two parts: fluid and micro-particles. The difference of permittivity between two materials affects the DEP force. Second, the drag force associated with the flow increases linearly with the fluid velocity because the radius of the particles used in analysis is constant. Third, applied voltage value is chosen to prevent its malfunction considering breakdown of air.

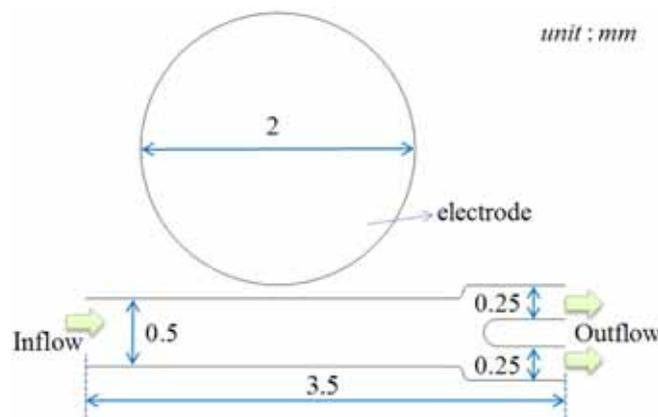


Figure 1: Analysis model

Fig.1 shows the designed device. It consists of the one electrode and the channel. The electrode, whose diameter is 2 mm, keeps 0.1 mm away from the channel. The maximum electric field intensity is about 2.4 MV/m when applied voltage is 5.5KV. This is under the breakdown field in the air. The channel has a one inlet and two outlets, and the micro beads are supplied with fluid velocity through the inlet. The fluid field with flow velocity, $\mathbf{u} = 0.5$ cm/s, is laminar flow, since the Reynolds number of the device and the particles is

much less than 2000 and 1 [5]: $Re_d \approx 0.16$ and $Re_p \approx 1.7 \times 10^{-3}$.

3.2 Analysis results

The drag force and DEP force can be calculated from each field distributions. Fig. 2 shows the fluid velocity field distribution obtained by the finite element method. The drag force can be obtained from equation (4) using the relative velocity between fluid and the particles at each point.

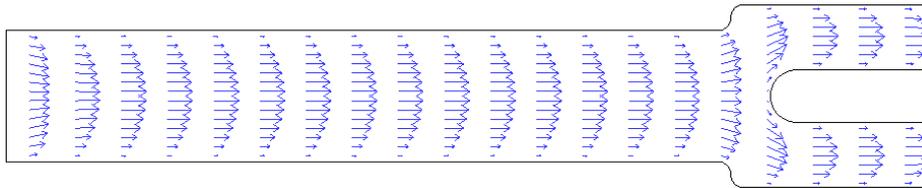


Figure 2: Flow velocity distribution

All particles experience the DEP force when the voltage is applied. Fig. 3(a) shows equi-potential distribution, and Fig. 3(b) depicts the gravitational force and the DEP force on the line that is near the bottom of the channel in Fig. 3(a). The DEP force between 0.3 mm and 2.6 mm is stronger than the gravitational force, which is about 0.8×10^{-12} [N], and the CM factor is a positive value. Therefore, in this region all particles move towards the electrode.

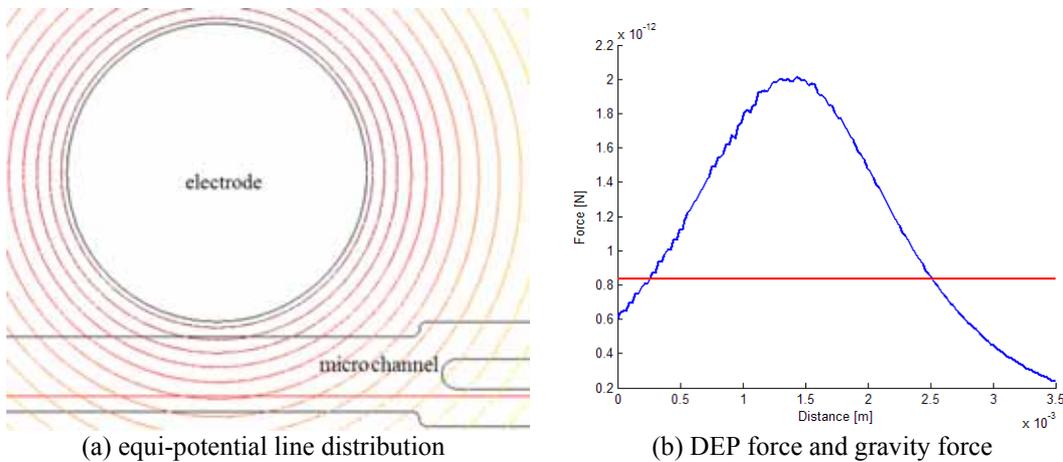


Figure 3: Comparing the forces near the bottom of the channel

The applied voltage to generate the same electric field becomes lower, as the electrode diameter becomes smaller. However, the DEP force acting on the particles at the same distance is weak in the case of a smaller electrode, because it is proportional to the electric field gradient. Fig. 4 shows the analytical result of the quarter of the electrode in Fig. 3. The region where the DEP force is dominant is smaller than one of the bigger electrode model.

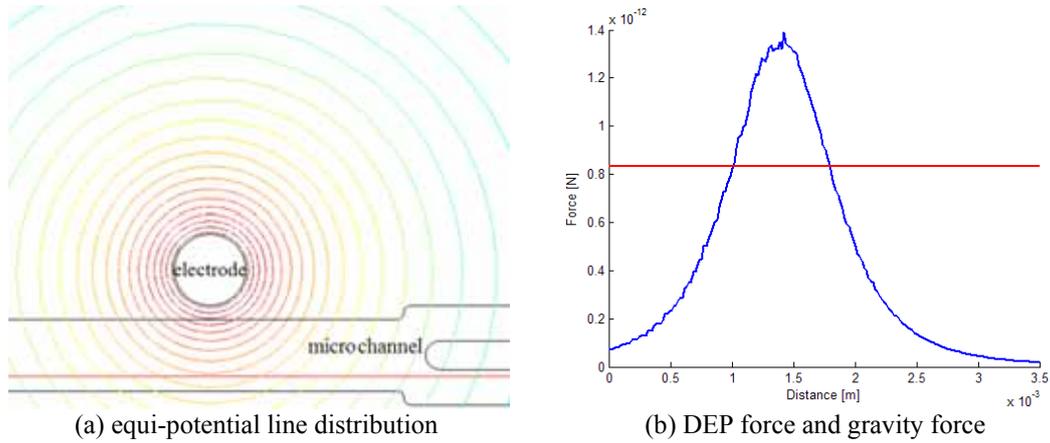


Figure 4: Comparing the forces near the bottom of the channel

The beads used in analysis are assumed to be carbon particles, instead of cells, to obtain the reliable analytical results using well-known material property. The diameter of the micro-particles is $5 \mu m$, and its mass density is assumed as $1.31 \text{ g} \cdot \text{cm}^{-3}$. Fig. 5 shows the particle motion in only the electric field.

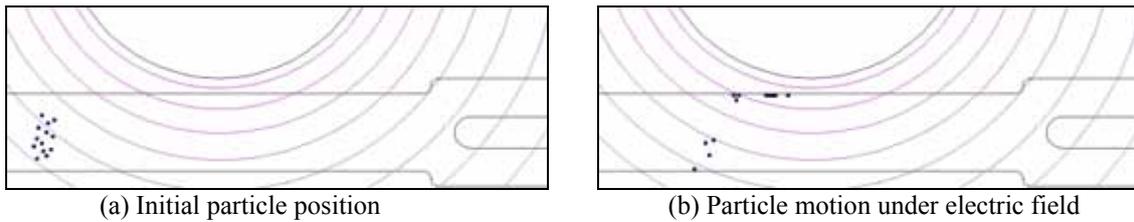


Figure 5: Electric field only exists

Fig. 6 shows the simulation results. The carbon micro-particles are supplied from the right side, and it can be manipulated by electric field. The injected beads move toward the lower channel when there is no electric field, since the forces acting on them are the drag and the gravitational forces. In contrast, if the voltage is applied, then all particles gather in the upper channel due to the DEP force.

4 CONCLUSION

We have proposed a numerical procedure that can solve the coupled problem to analyze the particle movement in fluid and electric fields. The total force acting on the beads was calculated from each field distributions by the finite element method, because the forces vary with its position. These values are substituted in Newton's equation to calculate the particle motion. The Runge-Kutta method is used to solve the coupled motional equation. The dynamic characteristics were estimated using the proposed algorithm in the micro-particle separating device designed.

Fluid velocity, material properties of a particle and an applied voltage are considered to design the separating device. The simulations showed the particles can be safely and rapidly manipulated using the DEP force, which is proportional to both the CM factor and electric field gradient. The usefulness of the proposed numerical scheme is validated.

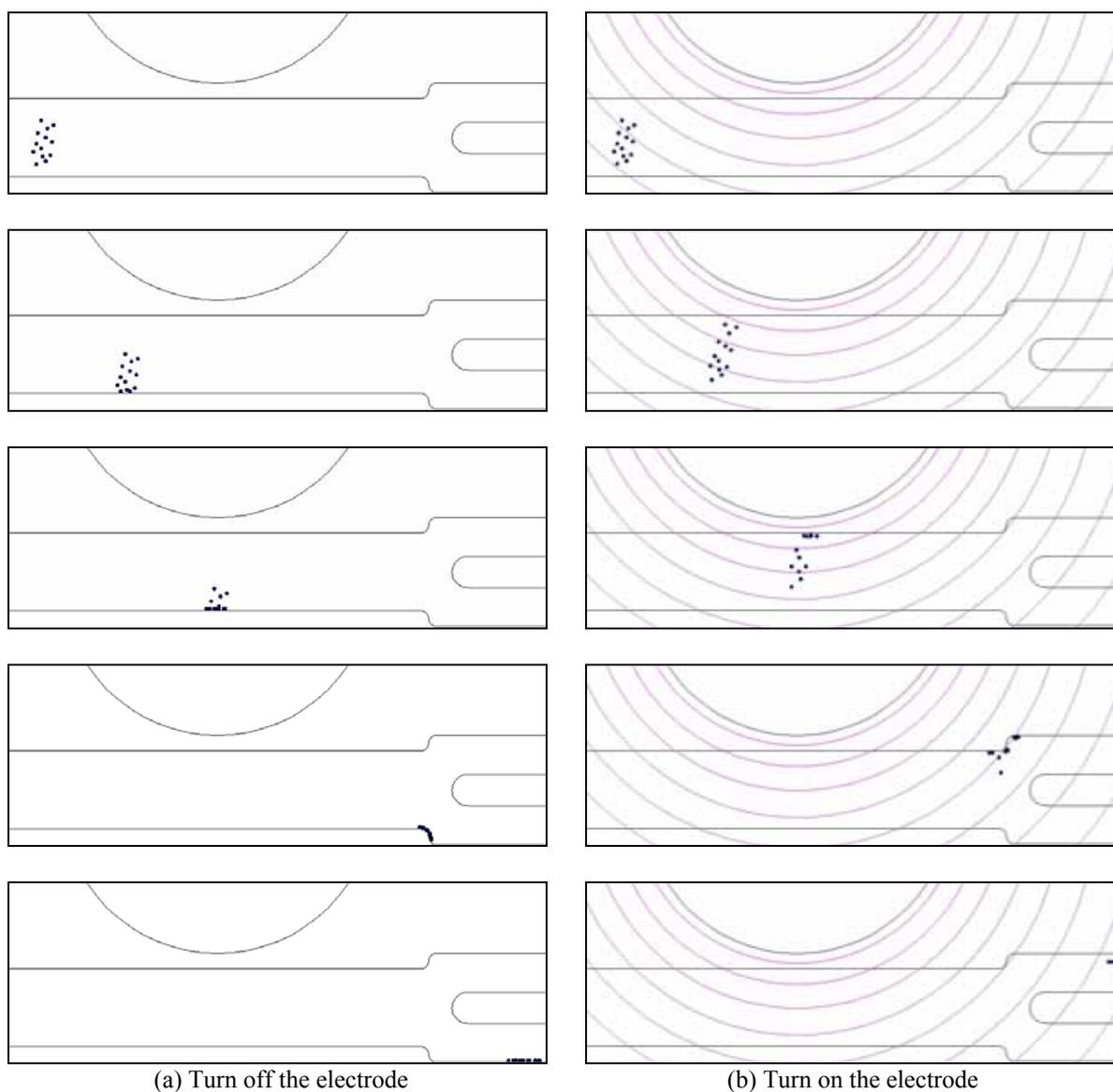


Figure 6: The micro-particle motion

REFERENCES

- [1] Thomas B. Jones, *Basic Theory of Dielectrophoresis and Electrorotation*, IEEE Engineering in Medicine and Biology Magazine(2003) 0739-5715.
- [2] William C. Hinds, *Aerosol Technology*, John Wiley & Sons, Inc., 2nd ed., (1999).
- [3] Joel Voldman, Rebecca A. Braff, Mehmet Toner, Martha L. Gray, and Martin A. Schmidt, *Holding Forces of Single-Particle Dielectrophoretic Traps*, Biophysical Journal (2001) 531-514.
- [4] Thomas B. Jones, *Electromechanics of Particles*, Cambridge University press, 1st ed., (1995).
- [5] Alexander J. Smits, *Fluid Mechanics*, John Wiley & Sons, Inc., 2nd ed., (2005).

NUMERICAL SIMULATION OF DROPLET SHAPES IN EXTERNAL ELECTRIC FIELD, GRAVITY AND SURFACE TENSION

TAN-IL SUNG^{*}, HONG-SOON CHOI[†], YOUNG-SUN KIM[‡] AND IL HAN PARK^{*}

^{*} School of Information and Communication Engineering
Sungkyunkwan University

Suwon 440-746, Korea

E-mail: ihpark@skku.ac.kr, Web page: <http://www.laem.skku.edu>

[†] School of Electrical Engineering

Kyungpook National University

Sangju 742-711, Korea

Email: tochs@knu.ac.kr, Web page: <http://www.knu.ac.kr>

[‡]Department of Electrical Engineering and Computer Science
MIT

Cambridge, MA 02139, USA

youngsun@mit.edu

Key words: Electrowetting, Electric body force, Surface tension, Shape deformation, Multiphysics Problems.

Abstract. The electrowetting is a phenomenon that the shape of droplet is deformed by external electric field. The most of electrowetting studies have been presented by researchers with mechanical viewpoint. In this paper, we present a numerical method to calculate the droplet shape by taking into account the effects of external electric field, surface tension and gravity. The numerical analysis for shape calculation is formulated by using the equilibrium condition of hydrostatical pressure in the coupled system of external electric field and surface tension in the presence of gravity. The model is numerically implemented and coupled using a standard finite element procedure. The proposed method is numerically tested and validated in a shaping problem of water droplet placed above a conductor coated by dielectric material in external electric field. The electrowetting phenomenon was successfully modeled and analyzed by the proposed approach.

1 INTRODUCTION

The electrowetting is a phenomenon that the shape of droplet is deformed by external electric field. Various applications, such as lap-on-a-chip, electronic display, and adjustable lenses are based on the electrowetting phenomenon. In recent papers, they assert that the shape of droplet is initially determined by the contact angle, and then the contact angle is changed by applying a voltage between two electrodes [1,2].

It is commonly understood that the contact angle, with no applied voltage, is only

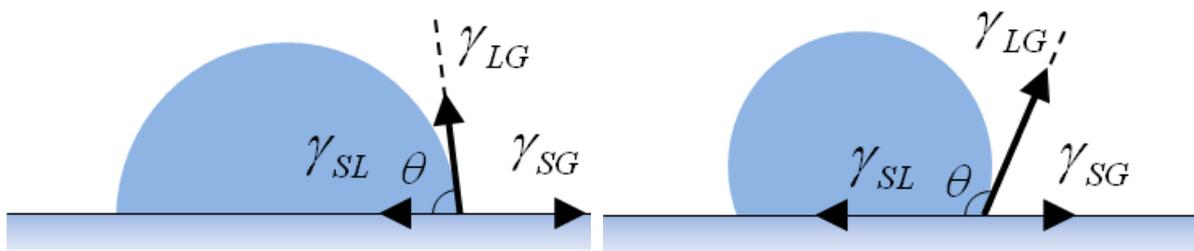
determined by the triple line as shown in Figure 1. The angle can be calculated using the following relation (1),

$$\gamma_{LG} \cos\theta = \gamma_{SG} - \gamma_{SL} \quad (1)$$

where γ is the surface tension coefficient (e.g., γ_{LG} at a liquid /gas interface) and θ is the contact angle. When a voltage is applied, the relation (1) is modified to (2) to consider the influence of the electric field [1],

$$\gamma_{LG} \cos\theta = \gamma_{SG} - \gamma_{SL} + \frac{1}{2}CV^2 \quad (2)$$

where C is the capacitance and V is the applied voltage. It was asserted that Maxwell stress can be used as the electric pressure [3].



(a) Hydrophilic contact, $\theta < 90^\circ$ (b) Hydrophobic contact, $\theta > 90^\circ$
Figure 1: Schematic view of the forces at the triple line (Solid, Liquid, Gas)

From the mechanical viewpoint [4]-[7], the shape of droplet is determined by the contact angle. And the change of contact angle by the voltage is explained to be due to the change of material property by the voltage. However, this paper aims to analyze the electrowetting in electrical viewpoint using the distributed electric force, which is generated by electric field distribution. That is, we couple the static fluid system with the electric system to form a coupled system equation for calculation of the droplet shape [8].

We present a numerical method to calculate the droplet shape by taking into account the effects of external electric field, surface tension and gravity. The different physical phenomena influence each other and they constitute a coupled system. So, their transient interaction between them is very complicated and difficult to analyze and calculate. Thus, in this paper, the numerical analysis for shape calculation is formulated by using the equilibrium condition of hydrostatical pressure in the coupled system of external electric field and surface tension in the presence of gravity.

The model is numerically implemented and coupled using a standard finite element procedure. The proposed method is numerically tested and validated in a shaping problem of water droplet placed above a conductor coated by dielectric material in external electric field.

2 PROPOSED APPROACH

From the hydrostatical force equilibrium, the pressure on the surface can be expressed as follows [8],

$$p_i = \int_{L_0}^{L_i} (\rho \mathbf{g} + \mathbf{f}_e) \cdot d\mathbf{l} + p_0 + p_{LS} \quad i = 1, 2, \dots, n \quad (3)$$

where p is the isotropic mechanical scalar pressure, i is the index for indicating different positions on the surface, n is the number of control positions, ρ is the mass density of the liquid, \mathbf{g} is the acceleration vector of gravity, \mathbf{f}_e is the Kelvin force density in the liquid, L_0 is an arbitrary fixed position on the free surface of the liquid, p_0 is the atmospheric pressure, L is a position on the droplet surface and p_{LS} is an additional pressure by surface curvature at position L .

The line integration path should go through the inside of the fluid to consider the internal body force. It is noted that the path can be arbitrary because the pressure p_i is independent of the integration path. In the state of hydrostatical equilibrium, all the p_i should be the atmospheric pressure, p_0 .

From [9], the formula of volume force density is written as

$$\mathbf{f}_e = -(\mathbf{P} \cdot \nabla) \nabla V = (\mathbf{P} \cdot \nabla) \mathbf{E} \quad (4)$$

where \mathbf{P} is polarization, \mathbf{E} is electric field intensity and V is the electric scalar potential.

The calculation technique of pressure using surface tension is shown in Figure 2(a). Also, the points in Figure 2(b) represent the control points to draw the surface of droplet. Pressure by surface tension is expressed as follows (5).

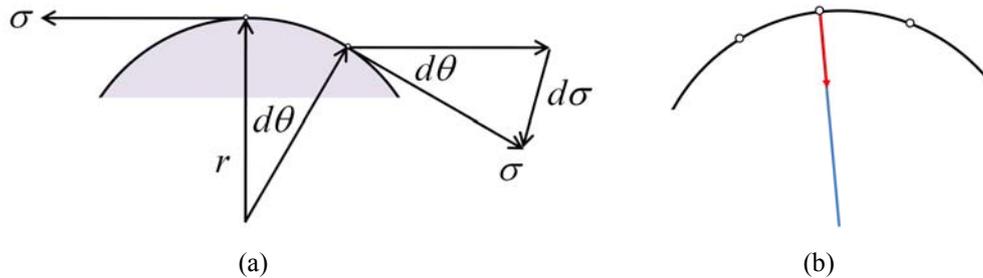


Figure 2: Schematic diagram to calculate (a) pressure by surface tension and (b) control points of droplet surface.

$$P_{st} = p_{LS} = \frac{2 \times \sigma d\theta}{r d\theta} = \frac{2\sigma}{r} \quad (N/m^2) \quad (5)$$

Where σ is the surface tension constant of water, r is the radius, determined by control points.

The model is numerically implemented and coupled using a standard finite element procedure. The proposed method is numerically tested and validated in a shaping problem of water droplet placed above a conductor coated by dielectric material in external electric field.

3 NUMERICAL ANALYSIS

From the stage of the experiment, Figure 3 shows the dimensions and properties for a numerical model. An axis-symmetric 3D formulation was derived and a finite element analysis was performed. From (4), the equation in cylindrical coordinates, multiplying it by r to avoid singularities at $r = 0$, the equation becomes

$$\mathbf{f}_e = -\frac{1}{r}(\mathbf{P} \cdot \nabla)r\mathbf{E}_0 \quad (6)$$

where r denotes the radial directional components and \mathbf{E}_0 is the external electric field intensity. \mathbf{E}_0 is expressed as follows [8],

$$\mathbf{E}_0 = \frac{1 + \varepsilon_r}{2}\mathbf{E} \quad (7)$$

The concept of an external field was first introduced by Kelvin, as follows in [9]. The electromagnetic external field can be also found in [10]-[12].

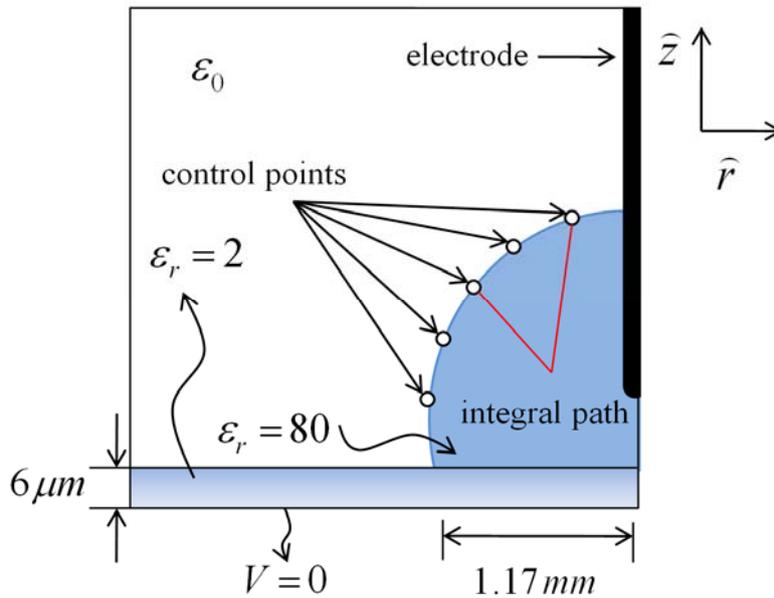


Figure 3: Numerical model from the experimental stage. Axis-symmetric 2D formulation and finite element analysis were performed. The line integration path should go through the inside of the fluid. The line integration path is used to calculate the pressure on the droplet surface. 30 points on the surface were used for the pressure calculating.

The droplet shape is detected with the pressure equilibrium condition at control points. The surface tension on the droplet through the radius and hysteresis phenomenon is ignored. The volume of the droplet is $5 \mu\text{l}$, and the Teflon coating thickness is $6 \mu\text{m}$. The integration path to the control points is shown in Figure 4.

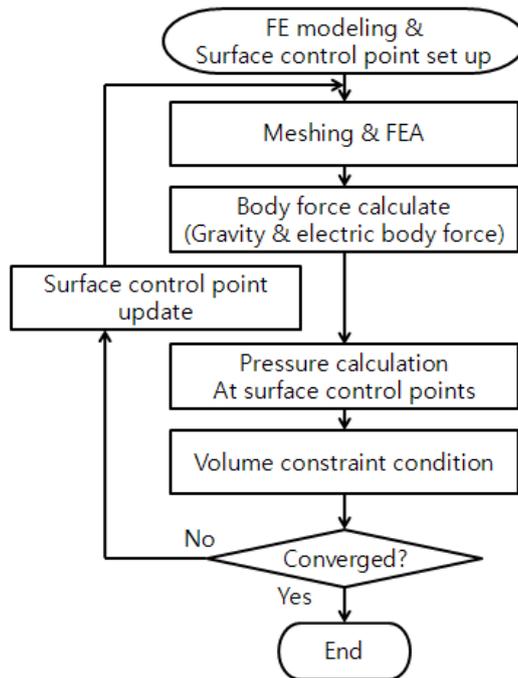


Figure 4: Flowchart to find surface control point by the hydrostatical equilibrium condition.

The free surface profile of the droplet can be obtained through numerical iteration, whose algorithm is based on the condition that the pressures on the free surface are the same. Figure 4 presents the overall flowchart of the procedure.

4 NUMERICAL RESULT

Figure 5 shows equi-potential lines passing through the surface of initial and final shape when 200ν is applied.

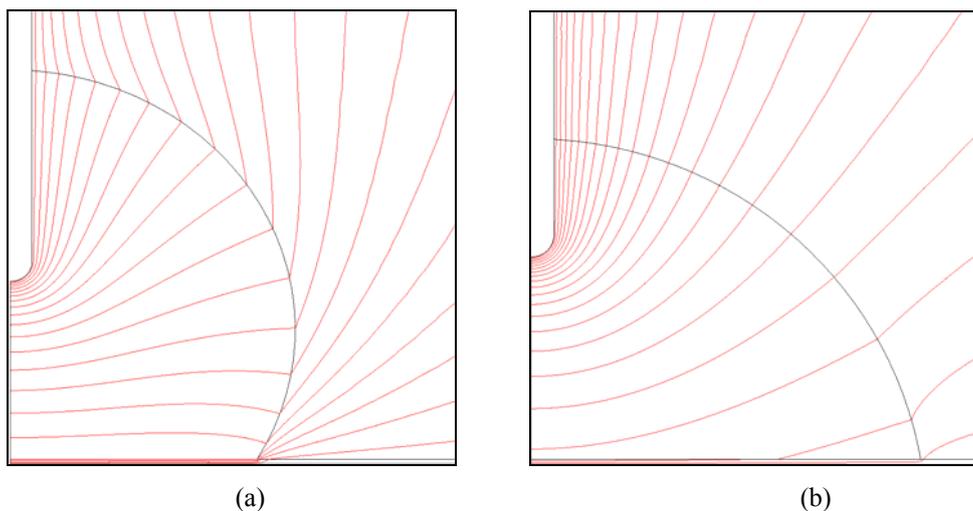


Figure 5: Body Surfaces of $1 \mu\text{l}$ droplet and flux distributions of (a) initial and (b) final shape.

Figure. 6 shows the electromagnetic body force density distribution near the tip of the electrode when the voltage is applied. The force density near the tip of the electrode is larger than the points far from the electrode.

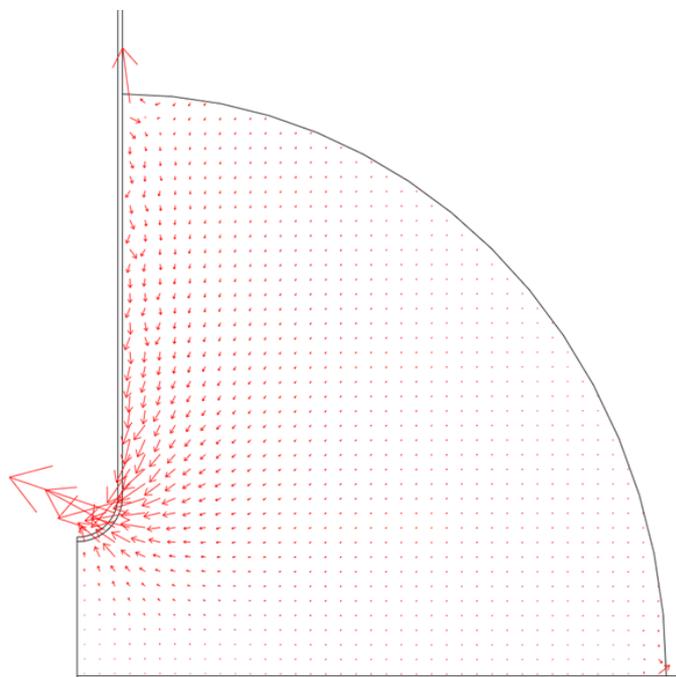


Figure 6: Body force density distribution. The arrows to visualize the body force density are drawn in each of the elements.

Figure 7 shows the trajectory of the droplet shape detected with the pressure equilibrium condition at control points.

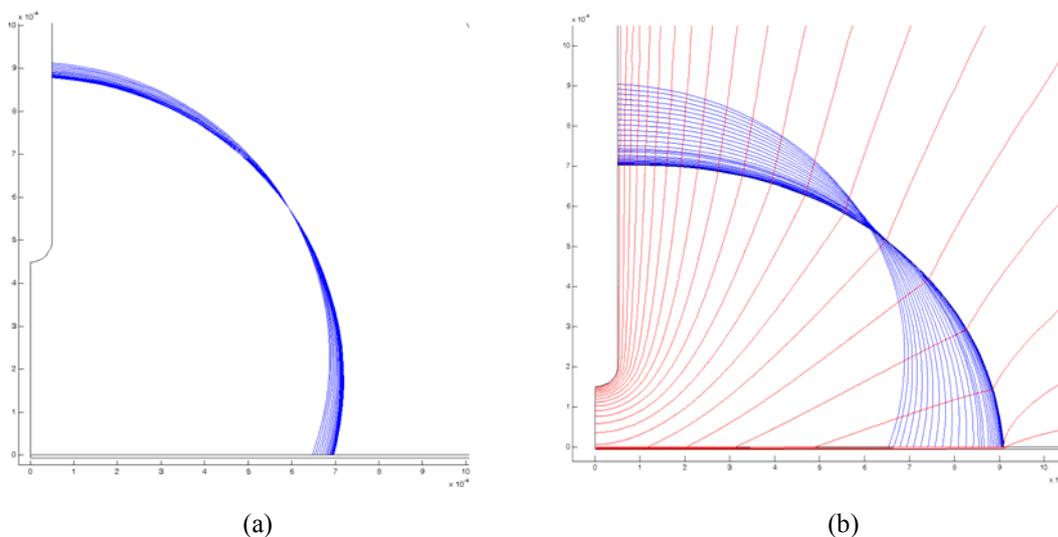


Figure 7: Numerical analysis results. Trajectories of the $1 \mu\text{l}$ droplet surface profile. Applied voltage : 50 V (a) and 200 V (b), The potential line for the last shape.

The final shapes of droplet according to the various applied voltage is shown in Figure 8. The volume constraint condition is adopted to determine the shape of droplet. The larger the voltage, the flatter the droplet shape. The result shows that the electrowetting phenomenon was successfully modeled and analyzed using the proposed approach.

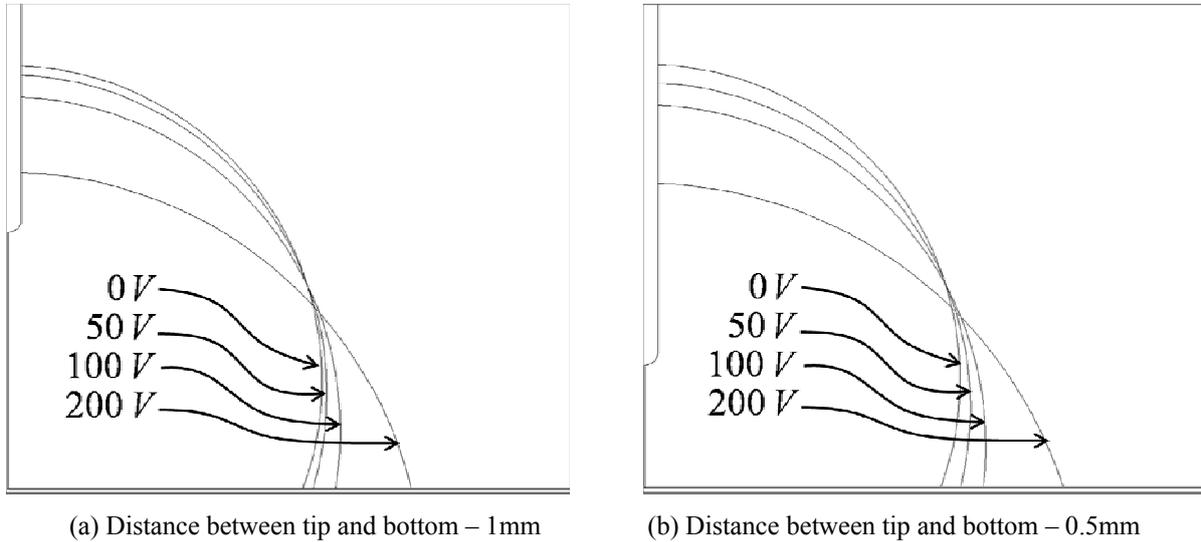


Figure 8: Comparison of droplet surface according to the applied voltage

5 EXPERIMENT

In Figure 9 presents photos of the experiment.

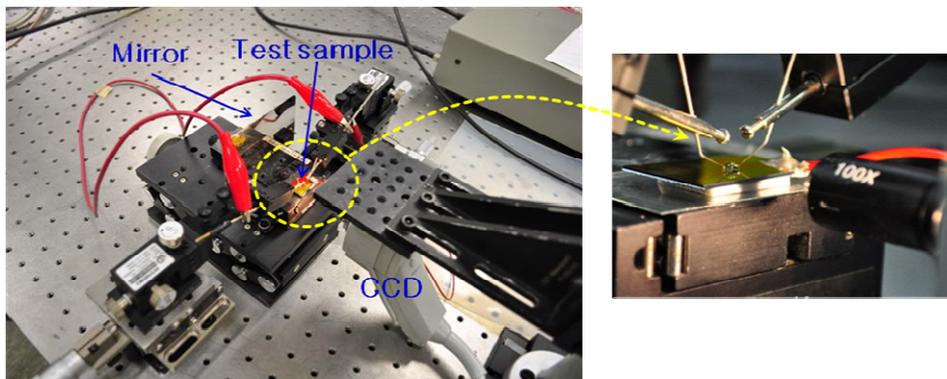


Figure 9: Experimental equipments used to observe electrowetting phenomena.

A detailed description of the measurement procedure follows. First, a $5 \mu\text{l}$ water droplet is placed on the grounded metal plate coated with Teflon. Second, voltage is applied to the electrode in the droplet. The distance between the tip of the electrode and coated metal plate was 1.5 mm and 0.5 mm . A picture of the droplet shape is taken using a CCD camera.

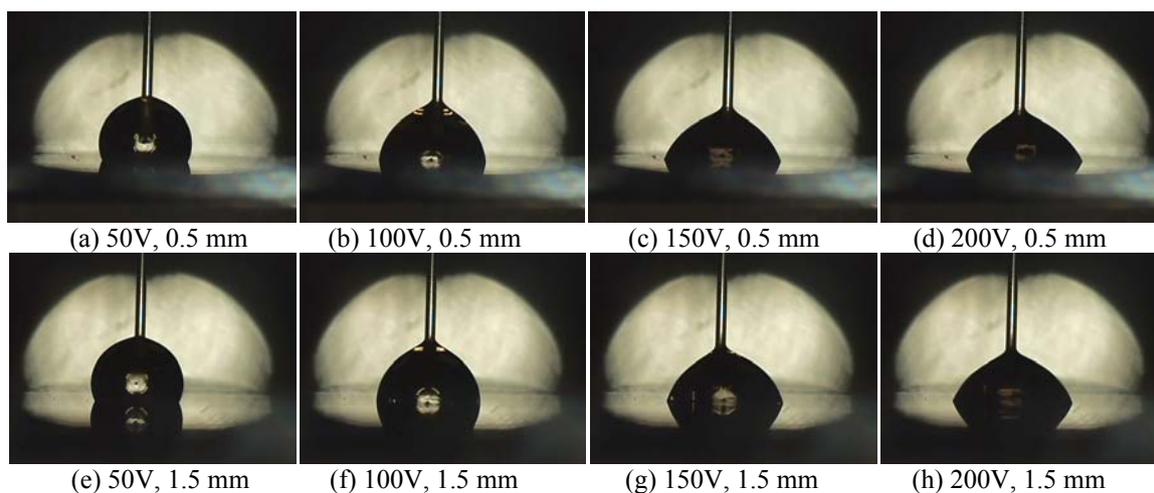


Figure 10: Photographs of shape deformation with applied voltage and depth of the electrode

5 CONCLUSION

In this paper, we present a numerical approach to represent the effect of electric field in electrowetting phenomenon. The electrowetting phenomenon can be regarded as a result of electromagnetic body force and surface tension. The proposed approach was successfully modeled and analyzed qualitatively for the electrowetting phenomenon. We have obtained a significant result although the experimental part permits only to analyze qualitatively the electrowetting phenomenon. We can get many advantages to analyze the electrowetting phenomenon from electrical viewpoint. Variations of voltage, the depth of electrode, the sort and thickness of the dielectric coating are easily treated in the electromagnetic viewpoint. The reason why contact angle is saturated, fingering effect near contact line, contact line instability radiation of light can be explained reliably, if electrowetting is continuously studied with electrical aspect.

ACKNOWLEDGMENT

This work was supported by the Human Resources Development of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Knowledge Economy. (No. 20104010100630-11-1-000)

REFERENCES

- [1] J. Berthier and P. Silberzan, *Microfluidics for Biotechnology*, Artech house, pp. 51-88, 2006.
- [2] Frieder Mugele and Jean-Christophe Baret, "Electrowetting: from basics to applications", *J. Phys: Condens. Matter.* 17(2005) R705-R774.
- [3] K. H. Kang, "How Electrostatic Fields Change Contact Angle in Electrowetting," *Langmuir*, 18, pp. 10318-10322, 2002.

- [4] Wim J. J. Welters and Lambertus G. J. Fokkink, "Fast Electrically Switchable Capillary Effects," *Langmuir*, 14, pp. 1535-1538, 1998.
- [5] T B Jones, "An electromechanical interpretation of electrowetting," *J. Micromech. Microeng.* 15(2005) 1184-1187.
- [6] Emilie Seyrat and Robert A. Hayes, "Amorphous fluoropolymers as insulators for reversible low-voltage electrowetting," *J. Appl. Phys.*, Vol. 90, No. 3, 1 August 2001,
- [7] E. J. De Souza, L. Gao, T. J. McCarthy, E. Arzt, and A. J. Crosby, "Effect of Contact Angle Hysteresis on the Measurement of Capillary Forces," *Langmuir*, 24, pp. 1391-1396, 2008.
- [8] H. S. Choi, Y. S. Kim, K. T. Kim, and I. H. Park, "Simulation of Hydrostatical Equilibrium of Ferrofluid Subjected to Magneto-static Field," *IEEE Trans. on Magn.*, Vol. 44, No.6 , pp.818-821, June 2008.
- [9] David J. Griffiths, *Introduction to Electrodynamics*, 3rd ed., Prentice-Hall, Inc. pp. 160-169, 2006.
- [10] J. J. Roche, "B and H, the intensity vectors of magnetism: A new approach to resolving a century-old controversy," *Am. J. Phys.* 68(5), pp. 163-168, May 2000.
- [11] C. Kittel, *Introduction to solid state physics*, John Wiley & Sons Inc., 2005, pp. 453-461
- [12] J. A. Stratton, *Electromagnetic Theory*, New York: McGraw Hill Book Co., 1941, pp. 258-261

NUMERICAL SIMULATIONS OF TUNGSTEN TARGETS HIT BY LHC PROTON BEAM

M. SCAPIN*, L. PERONI*, A. BERTARELLI†, AND A. DALLOCCHIO†

* Dipartimento di Meccanica (DIMEC)
Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
e-mail: martina.scapin@polito.it, web page: <http://www.polito.it>

† Mechanical and Materials Group, Engineering,
CERN,
CH-1211 Geneva 23, Switzerland
email: Alessandro.Dalocchio@cern.ch, web page: <http://www.cern.ch>

Key words: shock wave, hydrodynamic, EOS, LHC, high energy impact, thermo-mechanical

Abstract. The unprecedented energy intensities of modern hadron accelerators yield special problems with the materials that are placed close to or into the high intensity beams. The energy stored in a single beam of LHC particle accelerator is equivalent to about 80 kg of TNT explosive, stored in a transverse beam area with a typical value of 0.2 mm×0.2 mm. The materials placed close to the beam are used at, or even beyond, their damage limits. However, it is very difficult to predict structural efficiency and robustness accurately: beam-induced damage for high energy and high intensity occurs in a regime where practical experience does not exist. The interaction between high energy particle beams and metals induces a sudden non uniform temperature increase. This provokes a dynamic response of the structure entailing thermal stress waves and thermally induced vibrations or even the failure of the component. This study is performed in order to estimate the damage on a tungsten component due to the impact with a proton beam generated by LHC. The solved problems represent some accidental cases consequent to an abnormal release of the beam: the energy delivered on the components is calculated using the FLUKA code and then used as input in the numerical simulations, that are carried out via the FEM code LS-DYNA.

1 INTRODUCTION

The LHC [1,2] is a circular accelerator with a 26.659 km circumference situated at the border between Switzerland and France at an average depth of 100 m underground. This machine mainly provides the collision between two counter-circulating proton beams. At the design operating condition, each proton beam consists of 3×10^{14} protons at 7 TeV, so when the protons collide the collision energy is 14 TeV. The beam has 2808 bunches each having 1.11×10^{11} protons. The bunch length is 0.5 ns and the time between two successive bunches is 25 ns, so the duration of the entire beam is about 72 μ s.

The total energy stored in each beam at maximum energy is about 350 MJ: this is enough energy to melt 500 kg of copper. This large amount of energy is potentially destructive for

accelerator equipments having direct interaction with particles (e.g. the collimation system) in case of uncontrolled beam loss, so everything is done to ensure that this never happens. Besides, it is important to know what will be the damage in case of the LHC malfunction. It is in this perspective that a thermo-mechanical analysis becomes relevant. However, it is very difficult to predict structural efficiency and robustness accurately: beam-induced damage for high energy and high intensity occurs in a regime where, nowadays, practical experience does not exist. For now the importance of developing a reliable multidisciplinary methods (physics, hydrodynamics and structural engineering are involved) and accurate models that could be efficiently applied to estimate the damage occurring during an impact is therefore evident.

The interaction between high energy particle beams and solids can be considered from a structural point of view as an energy deposition inducing a sudden non uniform temperature increase. In function of which part of material is investigated the behaviour is different (Figure 1). In the material part closest to the beam, the pressure and temperature increase and the material could arrive at its melting temperature or vaporize. The material response in this condition is correctly described only using an equation of state that is able to describe the hydrodynamic behaviour, while in this portion of material the deviatoric stress is totally negligible. On the other hand, the remaining part of the material is characterized by high values of plastic strain, strain-rate and temperature, so the response is related with the strength material model used.

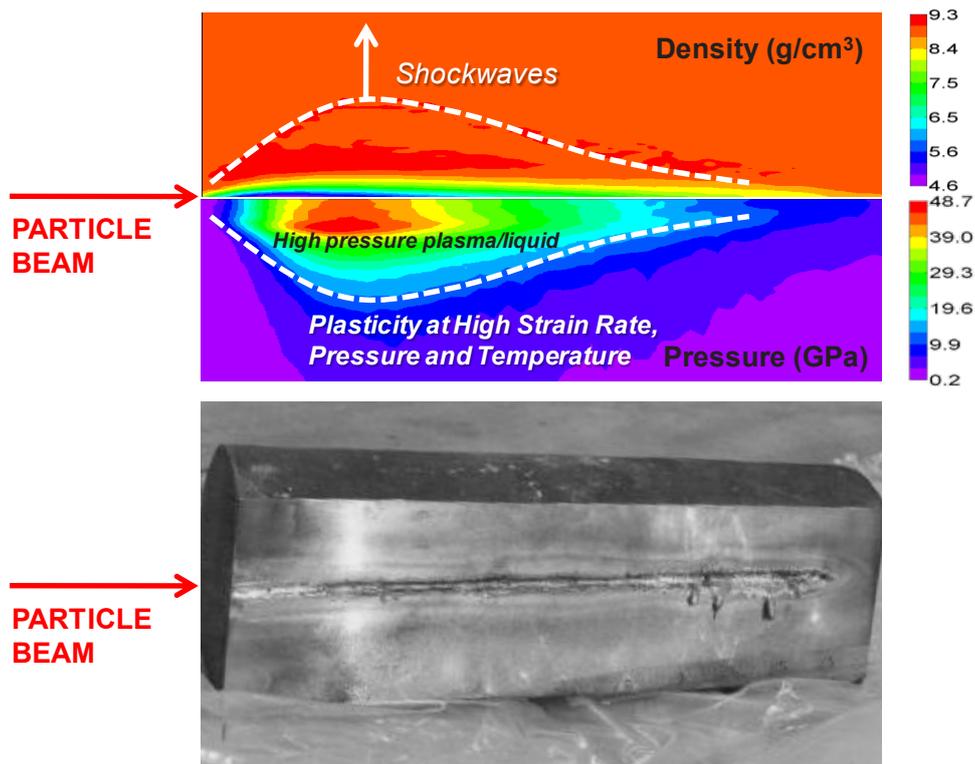


Figure 1: high energy particle beam impact effects on a metal structure

From these considerations it is clear what is the complexity of the problem: in order to correctly simulate the thermo-mechanical response of the hit material it is needed to take into

account both the hydrodynamic behaviour using a dedicated equation of state (EOS) and the deviatoric behaviour using a dedicated material model.

The numerical simulations are performed using the commercial FEM code LS-DYNA [3]. For the simulations the chosen equation of state is a polynomial EOS, in which the coefficients are obtained fitting a three-phase tabular equation of state, and the material model is the Johnson–Cook model.

The evaluation of thermal loads on the hit material is performed using a statistical code, called FLUKA [4], based on the Monte-Carlo method.

As mentioned before, the material involves in such high energy and high intensity impacts operates under extreme conditions, in which the possibility to perform experimental tests is limited. For this reason the importance of developing a reliable methods and accurate models that could be efficiently applied to estimate the damage occurring during an impact is therefore evident.

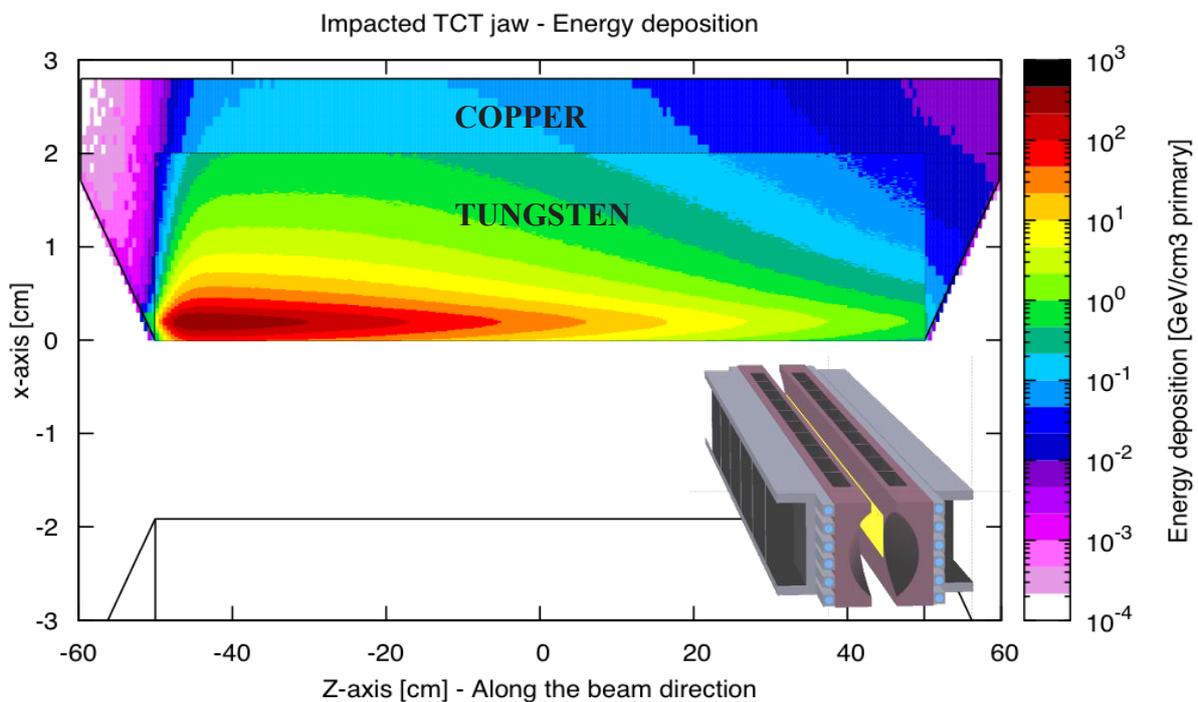


Figure 2: energy distribution (GeV/cm^3) on a tungsten/copper component (TCT collimators jaw) for a single proton at 5 TeV

In case of interaction between a high energy particle (protons) beam and a metallic material three main dynamic response regimes may occur, depending on the deposited energy, the energy density, the interaction duration and the material strength. The first possibility is the case in which the induced stress waves and the vibrations remain in the elastic domain. In this case the deposited energy density is low, the changes in density are negligible and the stress waves travel in the material at the sound speed. On the other hand, for medium energy levels, the stress waves are generated in the plastic domain. This implies the velocity of the waves is lower than the elastic domain speed and there are permanent deformations in the

component also once the load is over. The last case implies that a large amount of energy is delivered on the component. In the matter there is the dawn of shock waves, in which there is a nearly discontinuous change in the characteristic of the medium (pressure, temperature and density). The discontinuity moves with a supersonic velocity and this makes the mass transport phenomenon to be relevant.

In order to correctly simulate the effects of the high energy particle interaction with a solid target, it is necessary to calculate the energy deposition on the component delivered during the impact duration. In this work the calculation is done from the physicists at CERN via the statistical code FLUKA [4]. FLUKA is a fully integrated package for the calculation of particle transport and interaction with the matter, based on the Monte-Carlo analysis with many applications in high energy experimental physics and engineering. The FLUKA result (Figure 2) is the energy map on the same geometry on which the mechanical simulation is performed taking into account all the particles in the cascade generated by the interaction between the proton beam and the target. Finally the FLUKA results are used as input for thermo-structural studies.

2. HYDRODYNAMIC AND VISCOPLASTIC BEHAVIOUR

2.1 Equation of state

An equation of state (EOS) is a constitutive relation between state variables and describes the state of the matter. Usually it expresses a thermodynamic variable (such as pressure or internal energy) in function of two other independent state variables (such as density and temperature).

In this work a tabular multi-phase equation of state [5] is used. In particular, solid, liquid, vapour, gas and plasma region are taken into account. In Figure 3 the P - ρ (varying T) and P - E (varying ρ) planes for tungsten are depicted.

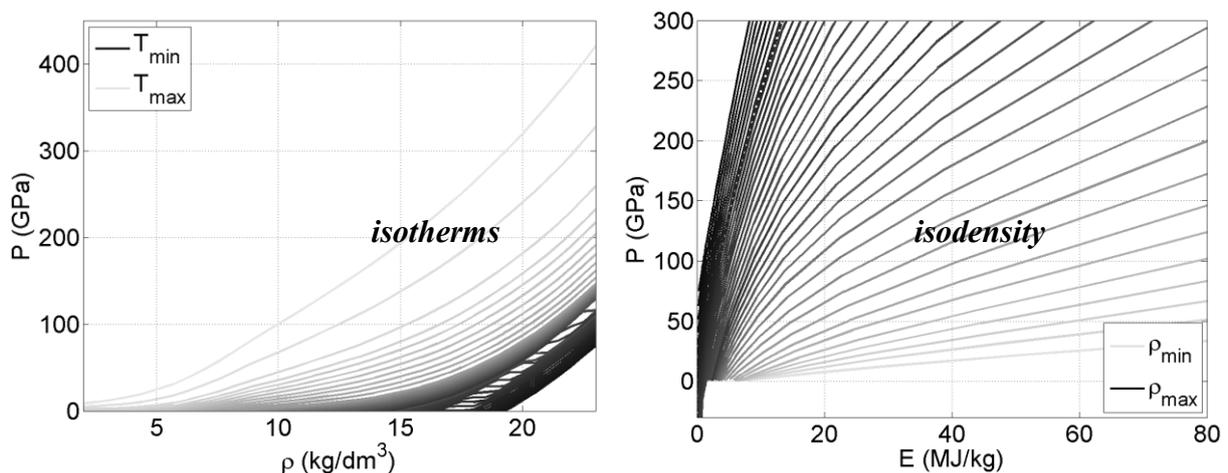


Figure 3: Equation of state for tungsten in the P - ρ (left) and P - E (right) planes for tungsten

Once the energy for the impact of a proton bunch is known, it is possible to understand the consequent material behaviour considering the EOS in the P - E plane varying the density (Figure 3). The arrival of the bunch pulse, obviously, modified the precedent matter condition: in particular, there is an increment in the energy of the medium. Since the impact is quite short, it can be considered as an isochoric transformation, so the jump in energy is done along the density curve characteristic of the previous state. This implies a new value in pressure univocally identified from the new energy and density conditions. The global result is that instantaneously the material reaches a different (greater) pressure. This, together with the fact that for this problem different parts of the component are not in the same condition, a shock front is generated: the elements with a high level of pressure rapidly expand into the elements with a low level.

2.2 Material model

In past decades a lot of material models for the description of the visco-plasticity in metals are proposed. The Johnson-Cook (J-C) model [6] is a purely empirical model; an example of semi-empirical model is the Steinberg-Cochran-Guinan-Lund model, which was first (S-G) developed for the description of high strain-rates behaviour [7], and after (S-L) extended to low strain-rates [8].

A first set of simulations is carried out with the J-C [6] model because also if it very simple it is able to predict the mechanical behaviour of the materials under different loading conditions. Besides, it is one of the most used material model, so it is implemented in many FEM codes and it is quite easy to find the values of J-C parameters for different materials. Nevertheless, the J-C model is quite inaccurate to describe the material behaviour in case of high pressure conditions and extremely high strain-rates. The J-C values are usually obtained in a range of strain-rate lower than 10^5 s^{-1} and could overestimate the strain-rate sensitivity for strain-rate $\geq 10^5 \text{ s}^{-1}$. As a matter of fact, it neglects the influence of the pressure and changes in volume on the flow stress and it considers the melting temperature as a constant, while the solid-liquid transition is influenced by the density.

In this work, a more suitable model is used: the S-G material model [7] in which the strength material characterization is controlled by the equations

$$\begin{aligned}
\sigma_y &= \sigma_A \frac{G}{G_0} \\
\sigma_A &= \sigma_0 \left[1 + \beta (\varepsilon_{pl,i} + \varepsilon_{pl}) \right]^n \leq \sigma_{\max} \\
\frac{G}{G_0} &= \left[1 + b P v^{1/3} - h (T - 300) \right] \\
T_m &= T_{m0} \exp[2a(1-v)] v^{-2(\gamma_0 - a - 1/3)}
\end{aligned} \tag{1}$$

in which both the flow stress and the shear modulus are function of temperature, strain and also pressure, v is the relative volume, b and n are the work-hardening parameters and $\varepsilon_{pl,i}$ is the initial equivalent plastic strain, normally equal to zero. The subscript 0 refers to the condition in which $T=300 \text{ K}$, $P=0$ and $\varepsilon=0$, the reference state. b and h are proportional to the derivative of the yield stress respect to the pressure and to the derivative of the shear modulus respect to the temperature, respectively. As in the J-C model, if the temperature overcomes

the melting temperature then the material strength goes to zero, but differently from the J-C model, the melting temperature is not constant but is a density function. The S-G model is strain-rate independent, according to the fact that for strain-rates greater than 10^5 s^{-1} the rapid decrease of rate dependent effects with increasing dynamic stress may be explained by the increase in temperature with increasing stress [7]. This assumption is reasonable in case of FCC material, but could be inaccurate in case of BCC materials [8].

4. BENCHMARKING

The numerical simulations are performed via the commercial code LS-DYNA. The validation of the numerical results is performed on a simplified geometry: a cylindrical bar (radius of 10 mm and length of 1 m) facially irradiated by 8 bunches at 7 TeV ($\sigma_r = 0.088 \text{ mm}$), each of them having 1.11×10^{11} protons. A lagrangian 2D axis-symmetric model is used with an explicit time integration (time step magnitude 0.01 ns). A non reflection boundary condition is applied at the external radius.

The response of a cylindrical structures subjected to a beam impact in the centre was investigated in different works for different materials [9,10]. For this reason, the results of similar simulations are easily comprehensible.

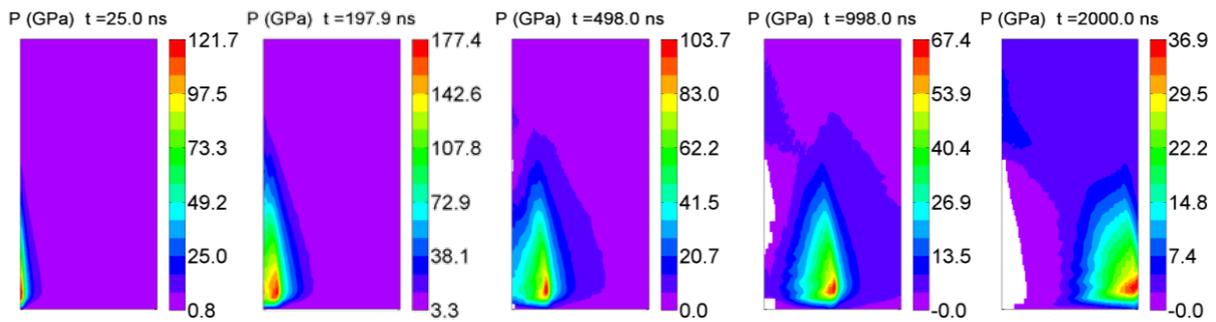


Figure 4: pressure evolution for tungsten components: radius (\rightarrow), length (\uparrow)

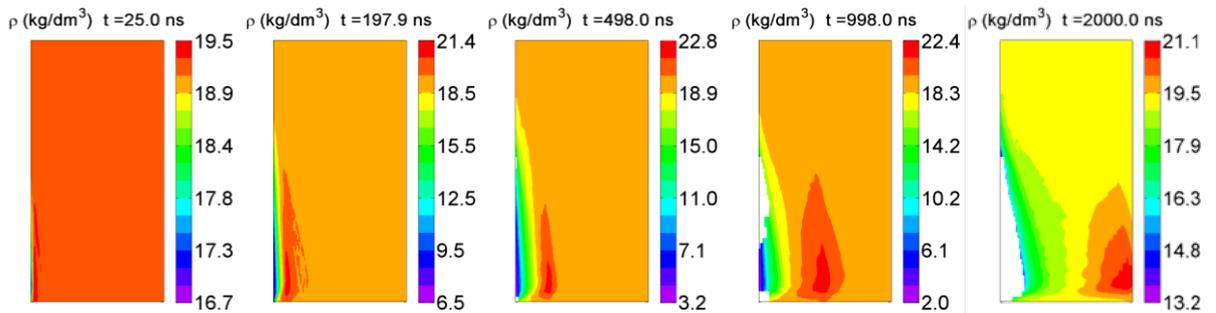


Figure 5: density evolution for tungsten component: radius (\rightarrow), length (\uparrow)

The results in terms of pressure, that are substantially independent from the chosen strength material model since the level of deposited energy is order of magnitude greater than plastic work, are reported in Figure 4.

A radial shock wave generated in the central/impact zone travels from the hit zone to the face. The levels of pressure reached at the end of the deposition ($\sim 200 \text{ ns}$) are very high. Since

the intensity level of the beam is the same used for the other materials [9,10], the difference in the pressure levels is essentially due to the difference in density: the greater the density the higher the energy absorption.

Figure 5, shows the density evolution: there is a rarefaction behind and a compression ahead the shock wave front. Due to the strong reduction in density, the material becomes gas and it is removed from the model.

In the centre the Von Mises stress (Figure 6) is zero because the temperature overcomes the melting value. In the remaining part of the component high levels of stress ($\sim 2 \sigma_0$) are reached.

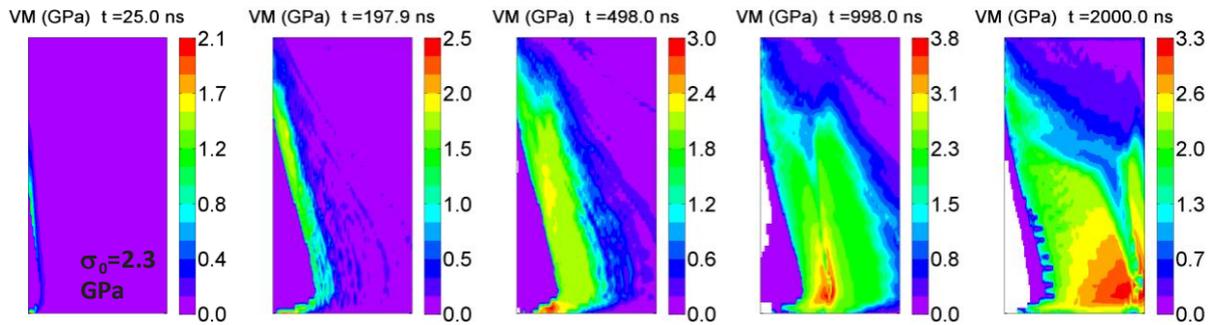


Figure 6: Von Mises stress evolution for tungsten component: radius (\rightarrow), length (\uparrow)

In Figure 7, an overview of the results is reported. The pressure and the density at the end of the deposition act as a pump (or similarly an explosive) which produces a dynamic deformation of the material ahead the shock front. In the remaining part of the component high levels of strain-rate (over 10^5 s^{-1}) and plastic strain (over 10%) are reached.

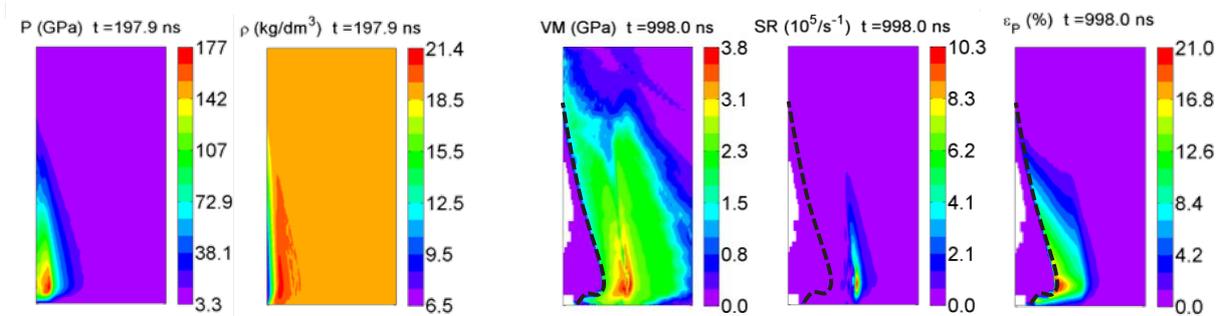


Figure 7: pressure and density distribution at the end of the deposition (left) and Von Mises, strain-rate and plastic strain at about $1 \mu\text{s}$ after the impact (right) for tungsten (bottom) component: radius (\rightarrow), length (\uparrow)

5. THREE-DIMENSIONAL REAL STRUCTURE

The method is finally applied to simulate the beam impact against a real complex geometry: the Target Collimator Tertiary (TCT) of LHC [11].

The lagrangian 3D numerical model (Figure 8) represents a simplified structure respect to the original one, but it considers the presence of two different material parts and the contact between the different model components. Substantially, the structure is a parallelepiped with

an internal part made by tungsten (the hit part) and an external C-section in copper. The simulation provides the impact of 8 bunches each having 1.13×10^{11} protons at 5 TeV ($\sigma_x = 0.3$ mm and $\sigma_y = 0.19$ mm).

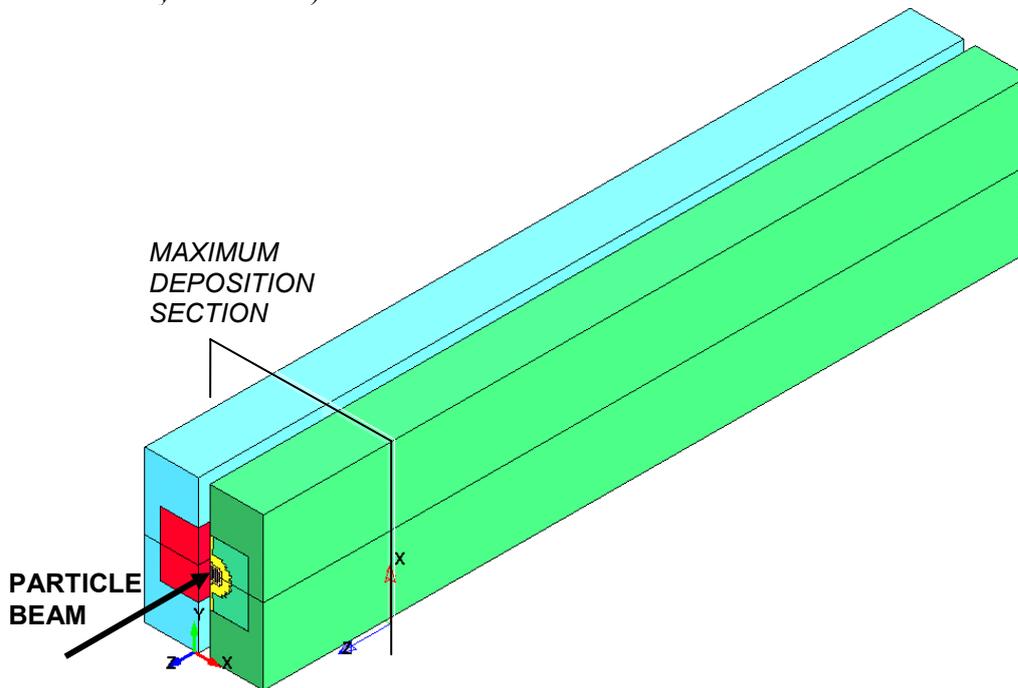


Figure 8. 3D view of the simplified TCT

In Figures 9-12 some simulation results on the model section of maximum deposition energy (Figure 8) are shown.

During the deposition phase (0÷182 ns) the pressure evolution is quite similar to the 2D cylindrical case. The reached level (156 GPa) is lower due to the different beam intensity and size. Once the deposition is finished a shock wave is generated and moves from the tungsten region to the copper part. Differently from 2D, in which the beam hits the centre of the structure, in this case the material is hit near a free surface. For this reason it can't create a wide zone of rarefaction because the pressure wave reaches in a short time the component free surface and is reflected: due to this the material is projected toward the other TCT jaw at high velocity (Figure 10) and consequently the other jaw may be damaged. The high level of particle speed in the material is a proof of shockwaves phenomena.

Since the significant difference in impedance between the two materials, the shock wave is in part reflected at the interface with the potential spallation of tungsten ($P_{\text{cutoff}} = 900$ MPa) and the reduction of the pressure level transmitted to copper.

In Figure 11 the results in terms of temperature are reported: the temperature is limited for a better understanding to the normal melting temperature (4520 K) since in the hit zone it could reach several thousand of K (plasma condition).

The level of stress is over 3 GPa in the tungsten part, that results heavy deformed after the beam impact. In the melted part of the component the Von Mises stress is zero and the behaviour is purely hydrostatic. On the other hand it reaches the maximum value behind the shock wave profile.

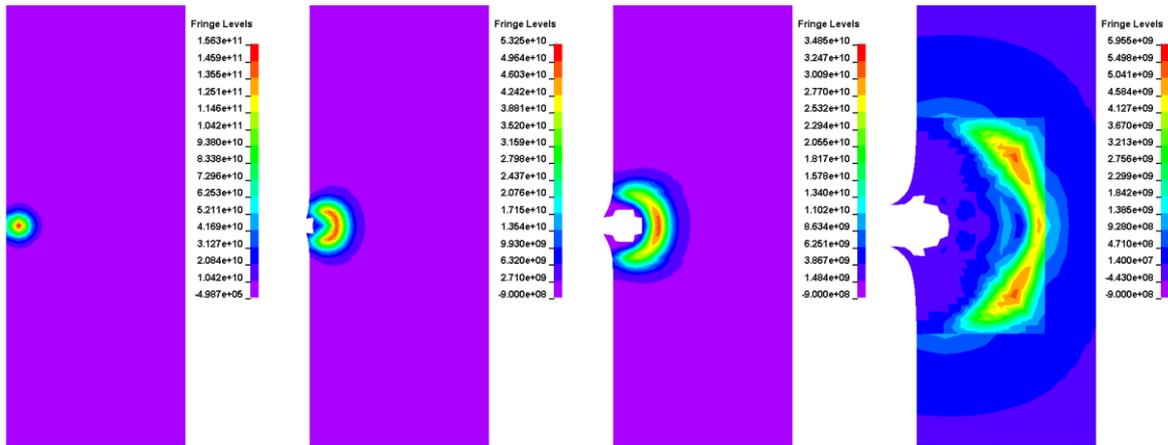


Figure 9. simulation results at: 182 ns (end of deposition), 600 ns, 1.2 μ s, 4 μ s, Pressure (Pa)

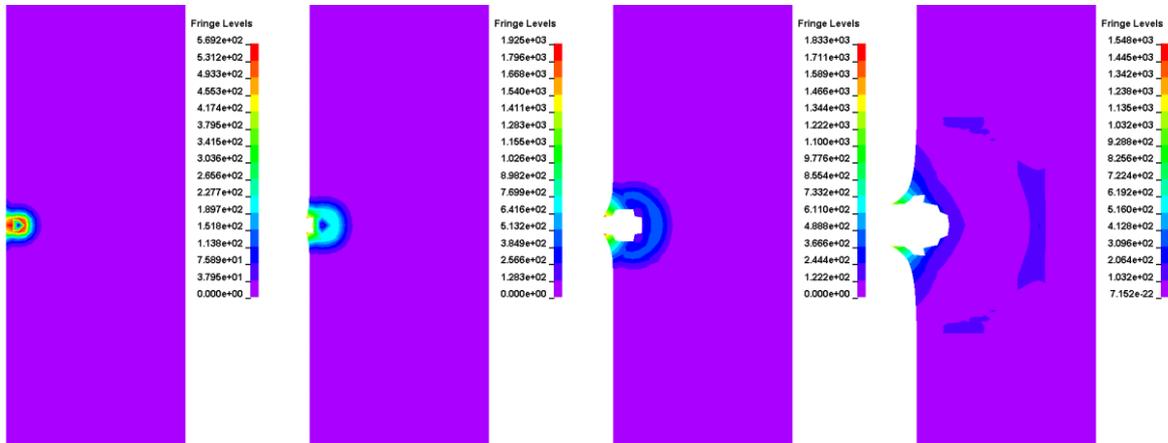


Figure 10. simulation results at: 182 ns (end of deposition), 600 ns, 1.2 μ s, 4 μ s, resultant velocity (m/s)

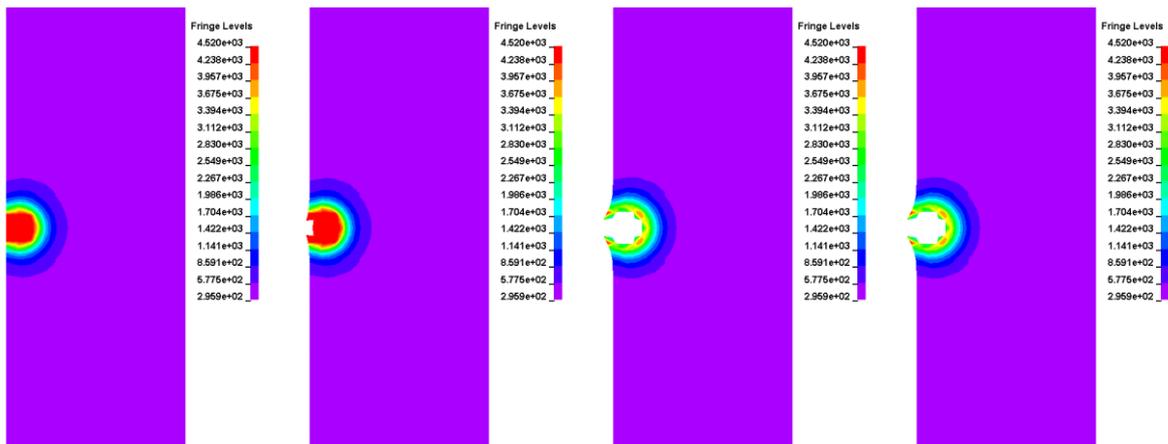


Figure 11. simulation results at: 182 ns (end of deposition), 600 ns, 1.2 μ s, 4 μ s, Temperature (K)

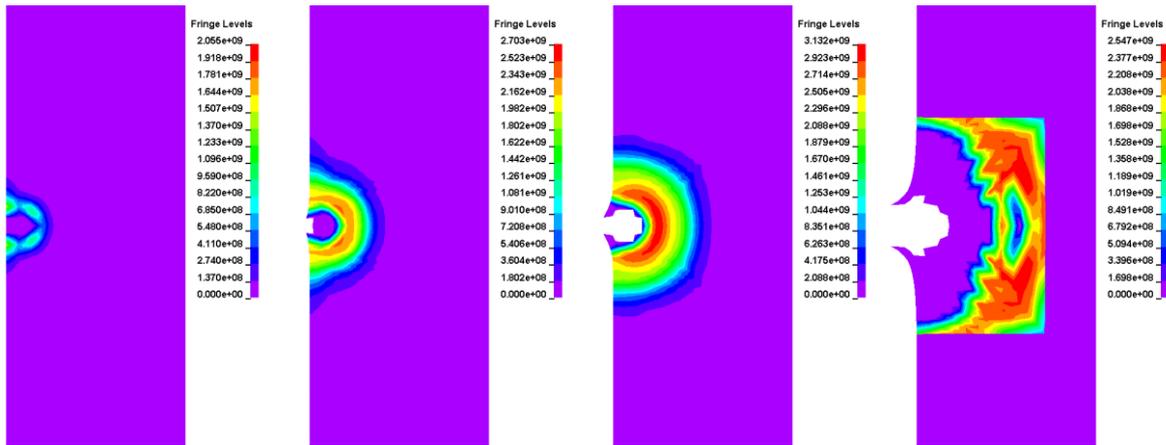


Figure 12. simulation results at: 182 ns (end of deposition), 600 ns, 1.2 μ s, 4 μ s, Von Mises Stress (Pa)

CONCLUSIONS

This study was performed in order to estimate the damage on a metal component due to the impact with a proton beam generated by the Large Hadron Collider (LHC) at CERN. The case study represents an accidental case consequent to an abnormal release of the beam, in which 8 bunches irradiate the target directly. The energy delivered on the component was calculated using the FLUKA code and then used as input in the numerical simulation, that were carried out via the FEM code LS-DYNA.

In this kind of problems both the hydrostatic and the deviatoric components are involved, but in different regions of the hit component. The hydrostatic behaviour of the target was described using a multi-phase equation of state, while the elasto-plasticity response was treated using Steinberg-Guinan material model.

In the first part, the validation of the numerical procedure on a simple geometry) and constitutive material models was presented. In particular, the high energy particle (7 TeV) impact was examined on a facially irradiated cylindrical bar: the beam hits the component directly on the centre of the basis. Then the final step was the study of the impact on a real structure with an energy beam of 5 TeV (the next target in the energy value of LHC beam).

REFERENCES

- [1] Wiedemann, H. *Particle Accelerator Physics I*. (1993) Berlin: Springer
- [2] Petterson, T.S., and Lefèvre, P. *The Large Hadron Collider: conceptual design*. (1995) Geneva: Desktop Publishing Service
- [3] Gladman, B. et al., *LS-DYNA® Keyword User's Manual – Volume I – Version 971*. (2007) LSTC
- [4] Fasso, A. et al., *FLUKA: A Multi-Particle Transport Code*. (2005)
- [5] Kerley, G.I., *Equation of State for Copper and Lead*. (2002) KTS report KTS02-1
- [6] Johnson, G. R. and Cook, W. A., A constitutive model and data for metals subjected to large strains, high strain rates and high temperatures, *7th International Symposium on Ballistic*, (1983) 541-547
- [7] Steinberg, D.J., Cochran, S.G. and Guinan, M.W., A constitutive model for metals

- applicable at high-strain rate. *Journal of Applied Physics*, (1980) **51**:1498
- [8] Steinberg, D.J. and Lund, C.M., A constitutive model for strain rates from 10^{-4} to 10^6 s⁻¹. *Journal de physique*, Symposium C3, (1988) **49**:433-440
- [9] Scapin, M., Peroni, L., and Dallochio, A. Thermo-mechanical modelling of high energy particle beam *Proc. 4th International Conference on Advanced Computational Engineering and Experimenting ACE-X 2010* 8-9 July 2010 Pari
- [10] Tahir, N.A., Shutov, A., Lomonosov, I.V., Dallochio, A., Bertarelli, A., and Piriz, A.R., Thermo-mechanical effects induced by beam impact on the LHC Phase II collimators: preliminary analysis using hydrodynamic approach
- [11] Redaelli, S., et al., Operational performance of the LHC collimation. *Proceedings of HB2010*
- [12] Le Blanc G Petit J Chanal P Y L'Eplattenier P and Avrillaud G 2008 Modelling the dynamic magneto-thermomechanical behavior of materials using a multi-phase EOS *Conf. 7th European LS-DYNA Conference* 14-15 May 2009 Salzburg
- [13] Steinberg D J 1996 Equation of State and Strength Properties of Selected Materials *Lawrence Livermore National Laboratory UCRL-MA-106439*

A COMPUTATIONAL STUDY OF EFFICIENCY OF CONTROLLABLE SEMIACTIVE MAGNETORHEOLOGICAL DAMPERS REDUCING LATERAL VIBRATION OF ROTORS WORKING IN CHAMBERS SUBMERGED IN LIQUIDS

JAROSLAV ZAPOMĚL^{*}, LIBOR ČERMÁK[†] AND PETR FERFECKI[#]

^{*} Department of Mechanics
VŠB-Technical University of Ostrava
17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic
e-mail: jaroslav.zapomel@vsb.cz

[†] Institute of Mathematics
Brno University of Technology
Technická 2, 616 69 Brno, Czech Republic
e-mail: cermak@fme.vutbr.cz

[#] Institute of Thermomechanics
Academy of Sciences of the Czech Republic, branch at VŠB-Technical University of Ostrava
17. listopadu 15, 708 33 Ostrava-Poruba, Czech Republic
e-mail: petr.ferfecki@vsb.cz

Key words: Rotor, Submerged Chamber, Semiactive Magnetorheological Squeeze Film Damper, Additional Masses and Damping, Coupled Physical Problems, Computational Methods.

Abstract. The design of rotating machines working in chambers submerged in liquids is a complicated technological problem. Vibration of the whole system produced by unbalance of the rotating parts can be significantly reduced if damping devices are inserted between the rotor and its stationary part. Due to variable operating conditions depending on the depth of the chamber submerging to achieve their optimum performance, the damping effect must be controllable. This is enabled by application of semiactive magnetorheological squeeze film dampers. Their damping is produced by squeezing a thin film of magnetorheological fluid between two rings and is controlled by the change of the magnetic flux passing through the lubricating layer. The liquid, in which the chamber is submerged, is considered as incompressible and acts on the chamber walls by its inertia and damping effects. The investigating behaviour of such systems requires to analyse several mutually coupled physical problems (mechanical, hydraulic, electrical and magnetic).

1 INTRODUCTION

The design of rotating machines working in chambers submerged in liquids belongs to complicated technological problems. Unbalance of the rotors always produces their lateral

vibration and forces transmitted through the coupling elements into the rotor housing and into the chamber walls. Their magnitude depends not only on the speed of the rotor rotation but also on the depth of submerging the chamber, the flow of the surrounding liquid and on position of the chamber in the tank or in the channel. The presence of the liquid shifts due to additional mass and damping the resonance frequencies and therefore also the critical revolutions of the rotor. Vibration of the whole system can be significantly reduced if damping devices are inserted between the rotor and its stationary part. Due to variable operating conditions, to achieve their optimum performance, their damping effect must be controllable. This is enabled by application of semiactive magnetorheological squeeze film dampers. Their damping is produced by squeezing a thin film of magnetorheological fluid filling the gap between the inner and outer damper rings. As resistance against the flow of magnetorheological liquid depends on magnetic induction, the damping effect can be controlled by the change of the current in the coils generating the magnetic field.

The regular experimental and theoretical research of squeeze film magnetorheological dampers started in the end of the 20th century. In the mathematical models the magnetorheological fluids are usually represented by Bingham or Bulkley-Herschel materials. Vibration and control of a flexible rotor attenuated by a magnetorheological damping device was studied by Wang et al. [1]. The mathematical models of a long squeeze film magnetorheological damper based on modification of the Reynolds equation and on the assumption that the rotor journal exhibits a circular orbit can be found in [2], [3] and [4]. These works were extended by Zapomel and Ferfecki [5], [6] who developed the mathematical models of short and long magnetorheological dampers applicable also in the cases when trajectory of the rotor journal centre in the damper is a curve of a general shape and consequently applied them for computational simulations investigating the transient response of a rigid rotor passing the critical speeds [7].

A computational procedure for analysis of lateral vibration of a rotor submerged in perfect liquid is reported by Zapomel in [8], [9]. The pressure distribution in the liquid is described by a Laplace equation that is obtained by simplification of the Navier-Stokes equations and the equation of continuity. Components of the resulting force acting on the disc of the rotor are obtained by integration of the pressure distribution around the circumference and along the width of the disc. These forces can be expressed as a linear combination of the components of the disc acceleration and therefore these coefficients of proportionality are considered as additional masses. To solve the Laplace equation a finite element method was adopted.

The procedure presented in this paper is intended for investigating the transient response of rigid rotors working in chambers submerged in liquids focusing on attenuation of the rotor vibration. Its application requires to solve several mutually coupled physical problems (mechanical, hydraulical, electrical, magnetic).

2 THE INVESTIGATED SYSTEM

The investigated rotor RT works in a chamber CH placed in a tank TA filled with liquid (water) LQ (Fig.1). The rotor consisting of a shaft and of one disc is supported by rolling element bearings whose outer races are coupled with the bearing housings by spring elements B1, B2. The chamber is coupled with the wall of the tank by spring elements C1 and C2

exhibiting also some damping. The rotor turns at constant angular speed and is loaded by its weight and by the disc unbalance. The springs supporting the rotor are prestressed to eliminate their deflection caused by the rotor weight. The chamber is cylindrical and is partly or fully submerged in the liquid. It is loaded by its weight and by the forces transmitted from the rotor through the bearings and the damping elements into the walls of the chamber faces.

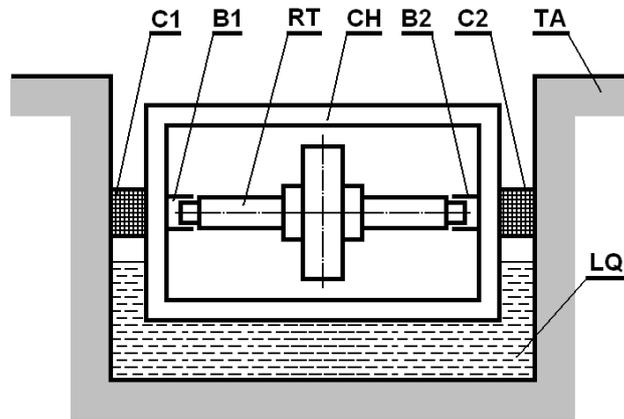


Figure 1: The investigated system

The task is to analyze influence of the change of damping of the magnetorheological dampers inserted between the rotor and its stationary part fixed with the chamber on attenuation of the system vibration at different operating conditions.

In the mathematical model the rotor, the chamber and the tank are considered as absolutely rigid bodies. The magnetorheological dampers placed between the rotor and its stationary part are represented by force couplings and the liquid in the tank is modelled as inviscid. Because of the design arrangement, the system is considered as symmetric relative to the middle plane of the disc and therefore the flow in the tank is two dimensional.

3 DETERMINATION OF THE DAMPING FORCE

The squeeze film magnetorheological damper consists of two rings between which there is a thin film of magnetorheological liquid (Fig.2). Both rings are coupled with the stationary part of the rotating machine, the outer ring directly, the inner one by a spring. The shaft is supported by a rolling element bearing whose outer race is coupled with the inner ring of the damper. Vibration of the inner ring relative to the outer one squeezes the film of magnetorheological liquid, which produces the damping force. In the body of the damper there is embedded an electric coil generating magnetic field. The magnetic flux passes through the layer of magnetorheological liquid and as resistance against its flow depends on magnetic induction, magnitude of the current can be used to control the damping effect.

The development of the mathematical model of the damping element is based on assumptions of the classical theory of lubrication with the exception of modelling the lubricating liquid. The fixed and movable rings are assumed to be absolutely rigid and their surfaces absolutely smooth, the width of the gap between the rings is very small relative to the rings radii, the magnetorheological liquid is represented by Bingham material, its yield shear

stress depends on magnetic induction, the flow in the lubricating layers is laminar and isothermal and the pressure in the radial direction remains constant and curvature of the oil film does not influence the lubricant flow.

The further attention will be focused only on the dampers whose geometry and design make possible to consider them as short [10]. Then the flow induced in the circumferential direction can be neglected.

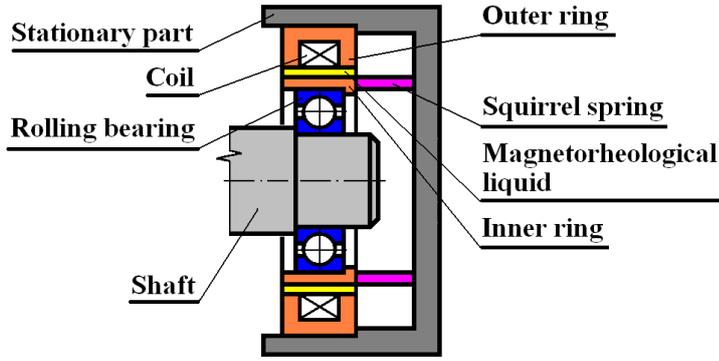


Figure 2: Scheme of the magnetorheological damper

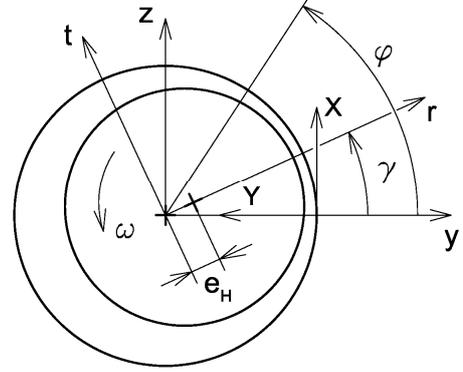


Figure 3: Coordinate system of the damper

The thickness of the lubricating film depends on position of the rotor journal centre relative to the system stationary part [10]

$$h = c - e_H \cos (\varphi - \gamma). \quad (1)$$

h denotes the thickness of the oil film, c represents the width of the gap between the fixed and movable rings, e_H is the rotor journal eccentricity, φ is the circumferential coordinate and γ denotes the position angle of the line of centres (Fig.3).

To describe the pressure field in the layer of magnetorheological fluid, the Reynolds equation was modified for the case of Bingham liquid [6]

$$h^3 p'^3 + 3(h^2 \tau_y - 4\eta_B \dot{h} Z)p'^2 - 4\tau_y^3 = 0 \quad \text{for} \quad p' < 0, \quad (2)$$

$$h^3 p'^3 - 3(h^2 \tau_y + 4\eta_B \dot{h} Z)p'^2 + 4\tau_y^3 = 0 \quad \text{for} \quad p' > 0. \quad (3)$$

Equations (2) and (3) represent cubic algebraic equations with respect to the pressure gradient in the axial direction p' . η_B denotes the Bingham dynamical viscosity, τ_y is the yield shear stress, Z is the axial coordinate and $(\dot{})$ denotes the first derivative with respect to time.

As follows from the derivation, the root that has the physical meaning must satisfy the following conditions:

- it must be real (not complex),
- the conditions of validity of equations (2) and (3) must be satisfied, this means that the real roots obtained from (2) must be negative and the real ones obtained from (3) must be positive,

$$\bullet \quad p' < -\frac{2\tau_y}{h} \quad \text{for} \quad p' < 0, \quad (4)$$

$$p' > \frac{2\tau_y}{h} \quad \text{for} \quad p' > 0. \quad (5)$$

The pressure profile p is given by integration of the pressure gradient

$$p = \int p' dZ. \quad (6)$$

Equations (2) and (3) are solved with the boundary condition expressing that pressure at the damper faces is equal to the pressure in the ambient space

$$p = p_A \quad \text{for} \quad Z = \pm \frac{L}{2}. \quad (7)$$

L is the length of the damper.

It is assumed that in the areas where the thickness of the lubricating film rises with time ($\dot{h} > 0$) a cavitation occurs. Pressure of the medium in cavitated regions remains constant and equal to the pressure in the ambient space. Then it holds for the pressure distribution in the film of the magnetorheological oil

$$p_D = p \quad \text{for} \quad p \geq p_A \quad \text{and} \quad p_D = p_A \quad \text{for} \quad p < p_A. \quad (8)$$

In the simplest design case of the damper, the rings, between which there is a layer of magnetorheological liquid, can be considered as a divided core of an electromagnet. Then the dependence of the yield shear stress on magnetic induction can be approximately expressed

$$\tau_y = k_y \left(\frac{NI}{2h} \right)^{n_y}. \quad (9)$$

k_y and n_y are the material constants of the magnetorheological liquid, N is the number the coil turns and I is the electric current.

y and z components of the damping force F_{MRy} , F_{MRz} are then given by integration of the pressure distribution around the circumference and along the length of the damping element. As it is assumed its symmetry relative to the plane perpendicular to the shaft centre line the resulting relations take the form

$$F_{MRy} = -2R \int_0^{\frac{L}{2}} \int_0^{2\pi} p_D \cos \varphi dZ d\varphi, \quad F_{MRz} = -2R \int_0^{\frac{L}{2}} \int_0^{2\pi} p_D \sin \varphi dZ d\varphi. \quad (10)$$

R is the radius of the gap between the damper rings.

4 THE FORCE OF THE LIQUID ACTING ON THE CHAMBER

To determine the force by which the liquid acts on the chamber it is assumed that the walls of the tank and of the chamber are absolutely rigid and smooth, the liquid is incompressible, newtonian, the flow is isothermal and due to the system symmetry it is two dimensional (2D).

Then the pressure and the velocity field is governed by the Navier-Stokes equations and by the equation of continuity

$$\frac{\partial p_F}{\partial x_1} + \rho_F \left(\frac{\partial v_1}{\partial t} + v_1 \frac{\partial v_1}{\partial x_1} + v_2 \frac{\partial v_1}{\partial x_2} \right) - \eta_F \left(\frac{\partial^2 v_1}{\partial x_1^2} + \frac{\partial^2 v_1}{\partial x_2^2} \right) = \rho_F g, \quad (11)$$

$$\frac{\partial p_F}{\partial x_2} + \rho_F \left(\frac{\partial v_2}{\partial t} + v_1 \frac{\partial v_2}{\partial x_1} + v_2 \frac{\partial v_2}{\partial x_2} \right) - \eta_F \left(\frac{\partial^2 v_2}{\partial x_1^2} + \frac{\partial^2 v_2}{\partial x_2^2} \right) = \rho_F g, \quad (12)$$

$$\frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} = 0. \quad (13)$$

x_1, x_2 are the y and z cartesian coordinates, v_1, v_2 are the velocity components of the flow in the y and z directions, p_F is the pressure, ρ_F, η_F are the density and dynamical viscosity of the liquid, g is the gravity acceleration and t is the time.

As the chamber performs oscillations only with small amplitude, the Reynolds number of the flow is low and the liquid inertia forces are dominant [11], [12]. On these conditions the viscous and convective terms in the Navier Stokes equations can be neglected and consequently (11) - (13) can be transformed into a Laplace equation describing the pressure distribution in the liquid [12]

$$\frac{\partial^2 p_F}{\partial x_1^2} + \frac{\partial^2 p_F}{\partial x_2^2} = 0. \quad (14)$$

Because of the performed simplifications, p_F in (14) covers only the pressure induced by vibration of the chamber but it does not include the hydrostatic component that is caused by the weight of the fluid and that produces the lifting force acting on the chamber.

The liquid particles touching the body surface are not constrained in the tangential direction and in the normal direction they move together with the chamber wall. This is expressed by the boundary condition needed for solving the Laplace equation (14)

$$\frac{\partial p_F}{\partial n} = -\rho_F a_n. \quad (15)$$

n is the coordinate in the direction the outer normal to the chamber surface and a_n is acceleration of the liquid particle on the boundary in the direction of the outer normal.

As the chamber exhibits only a sliding motion, the boundary condition (15) can be rewritten in the following form

$$\frac{\partial p_F}{\partial x_1} \cos \alpha_n + \frac{\partial p_F}{\partial x_2} \sin \alpha_n = -\rho_F (a_y \cos \alpha_n + a_z \sin \alpha_n). \quad (16)$$

a_y, a_z are the acceleration components of the chamber sliding motion in the y and z directions and α_n is the directional angle of the outer normal of the chamber surface.

The governing equation (14) together with the boundary condition (16) is solved by application of a finite element method. After discretization of the region filled with the liquid and after some manipulation described in details e.g. in [9], [12] one obtains at each point of

time a set of linear algebraic equations for unknown values of the pressure at all nodes.

It follows from the Laplace equation (14) and the boundary condition (16) that the pressure can be expressed at each moment of time as a sum of terms proportional to the acceleration components of the sliding motion of the chamber

$$p_F(x_1, x_2) = a_y p_y^*(x_1, x_2) + a_z p_z^*(x_1, x_2) + p_0. \quad (17)$$

p_y^* , p_z^* are the coefficients of proportionality that depend only on the border (the chamber wall) geometry and p_0 is the pressure component depending on the pressure at the free surface (atmospheric pressure). Then integration of the pressure distribution around the circumference and along the length of the submerged part of the chamber gives components of the hydraulic force acting on the chamber

$$F_{Fy} = -L_{CH} \int_{\Gamma_{CH}} p_F \cos \alpha_n ds, \quad F_{Fz} = -L_{CH} \int_{\Gamma_{CH}} p_F \sin \alpha_n ds. \quad (18)$$

L_{CH} is the length of the chamber and Γ_{CH} is the border of the region filled with the liquid corresponding to the chamber wall.

Taking into account relation (17), equations (18) can be rewritten into the form [9], [12]

$$F_{Fy} = -m_{Fyy} a_y - m_{Fyz} a_z + F_{Fy0}, \quad (19)$$

$$F_{Fz} = -m_{Fzy} a_y - m_{Fzz} a_z + F_{Fz0}. \quad (20)$$

Coefficients m_{Fyy} , m_{Fyz} , m_{Fzy} , m_{Fzz} are called additional masses and the force components F_{Fy0} , F_{Fz0} depend on the pressure on the free surface. The additional masses express inertia effects of the liquid by which the vibration of the chamber is influenced. They depend on the mutual position of the chamber relative to the tank and on geometric shape of their walls.

5 THE EQUATIONS OF MOTION OF THE INVESTIGATED SYSTEM

Lateral vibration of the rotor and of the chamber partly or fully submerged in water is described by four equations of motion. Taking into account the system symmetry, they have the form

$$m_R \ddot{y}_R + 2k_{CHR} y_R - 2k_{CHR} y_{CH} = m_R e_T \omega^2 \cos(\omega t + \psi_0) - 2F_{MRy}, \quad (21)$$

$$m_R \ddot{z}_R + 2k_{CHR} z_R - 2k_{CHR} z_{CH} = m_R e_T \omega^2 \sin(\omega t + \psi_0) - 2F_{MRz} - m_R g + 2F_{PS}, \quad (22)$$

$$m_{CH} \ddot{y}_{CH} + 2b_{CHT} \dot{y}_{CH} - 2k_{CHR} y_R + 2(k_{CHR} + k_{CHT}) y_{CH} = 2F_{MRy} + F_{Fy}, \quad (23)$$

$$m_{CH} \ddot{z}_{CH} + 2b_{CHT} \dot{z}_{CH} - 2k_{CHR} z_R + 2(k_{CHR} + k_{CHT}) z_{CH} = 2F_{MRz} + F_{Fz} - 2F_{PS} - m_{CH} g + F_L \quad (24)$$

where

$$2F_{PS} - m_R g = 0. \quad (25)$$

m_R , m_{CH} denote the mass of the rotor and of the chamber, k_{CHR} , k_{CHT} are stiffnesses of the coupling elements placed between the rotor and the chamber and the chamber and the tank, b_{CHT} is the coefficient of linear damping in the constraint between the chamber and the tank,

F_{MRy} , F_{MRz} are components of the damping force produced by the magnetorheological dampers in y and z directions, F_{Fy} , F_{Fz} are the y and z components of the hydraulic force by which the liquid acts on the chamber, F_{PS} is the prestress force, F_L is the lifting force, e_T is eccentricity of the rotor centre of gravity, ω is angular speed of the rotor rotation, ψ_0 is the phase lag, y_R , z_R , y_{CH} , z_{CH} are displacements of the rotor and chamber centres in the y and z directions respectively and $(\dot{})$, $(\ddot{})$ denote the first and second derivatives with respect to time.

After some manipulations the equations of motion can be rewritten into a matrix form

$$\mathbf{M} \ddot{\mathbf{x}} + \mathbf{B} \dot{\mathbf{x}} + \mathbf{K} \mathbf{x} = \mathbf{f}_{CO} \cos \omega t + \mathbf{f}_{SO} \sin \omega t + \mathbf{f}_{MR} + \mathbf{f}_F + \mathbf{f}_G + \mathbf{f}_L. \quad (26)$$

\mathbf{M} , \mathbf{B} , \mathbf{K} are the mass, damping and stiffness matrices, \mathbf{f}_{MR} is the vector of damping forces produced by the magnetorheological dampers, \mathbf{f}_F is the vector of hydraulic force by which the liquid acts on the chamber, \mathbf{f}_{CO} , \mathbf{f}_{SO} are the vectors of the unbalance force acting on the rotor, \mathbf{f}_G , \mathbf{f}_L are the vectors of gravitational and lifting forces and \mathbf{x} is the vector of displacements of the rotor and of chamber centres.

To solve the equation of motion (26) a modified Newmark method has been applied. To simplify the calculation procedure, the damping forces related to time $t+\Delta t$ are determined by means of their expansion into a Taylor series in the neighbourhood of time t and by neglecting the terms of the second and higher orders

$$\mathbf{f}_{MR,t+\Delta t} = \mathbf{f}_{MR,t} + \mathbf{D}_B (\dot{\mathbf{x}}_{t+\Delta t} - \dot{\mathbf{x}}_t) + \mathbf{D}_K (\mathbf{x}_{t+\Delta t} - \mathbf{x}_t) + \dots \quad (27)$$

\mathbf{D}_B and \mathbf{D}_K denote the Jacobi matrices of partial derivatives.

After these manipulations the equation of motion (26) referred to time $t+\Delta t$ is rewritten

$$\begin{aligned} & \mathbf{M} \ddot{\mathbf{x}}_{t+\Delta t} + (\mathbf{B} - \mathbf{D}_{B,t}) \dot{\mathbf{x}}_{t+\Delta t} + (\mathbf{K} - \mathbf{D}_{K,t}) \mathbf{x}_{t+\Delta t} = \\ & = \mathbf{f}_{CO} \cos \omega t + \mathbf{f}_{SO} \sin \omega t + \mathbf{f}_{MR,t} - \mathbf{D}_B \dot{\mathbf{x}}_t - \mathbf{D}_K \mathbf{x}_t + \mathbf{f}_F (\ddot{\mathbf{x}}_{t+\Delta t}) + \mathbf{f}_G + \mathbf{f}_L. \end{aligned} \quad (28)$$

According to (18) - (20) components of the hydraulic force F_{Fy} , F_{Fz} can be expressed by means of the chamber acceleration components

$$\mathbf{f}_{F,t+\Delta t} = -\mathbf{M}_F \ddot{\mathbf{x}}_{t+\Delta t} + \mathbf{f}_{P0}. \quad (29)$$

Consequently the equation of motion (26) related to time $t+\Delta t$ takes the form

$$\begin{aligned} & (\mathbf{M} + \mathbf{M}_F) \ddot{\mathbf{x}}_{t+\Delta t} + (\mathbf{B} - \mathbf{D}_{B,t}) \dot{\mathbf{x}}_{t+\Delta t} + (\mathbf{K} - \mathbf{D}_{K,t}) \mathbf{x}_{t+\Delta t} = \\ & = \mathbf{f}_{CO} \cos \omega t + \mathbf{f}_{SO} \sin \omega t + \mathbf{f}_{MR,t} - \mathbf{D}_B \dot{\mathbf{x}}_t - \mathbf{D}_K \mathbf{x}_t + \mathbf{f}_G + \mathbf{f}_L + \mathbf{f}_{P0}. \end{aligned} \quad (30)$$

\mathbf{M}_F is the matrix of additional mass and \mathbf{f}_{P0} is the vector of the hydraulic force depending on the pressure at the free surface. As the displacements are considered to be small, the change of the shape of the region filled with the water is negligible and coefficients of matrix \mathbf{M}_F remain constant. The resulting equation of motion (30) has the form required by the basic Newmark algorithm and its calculation arrives at each integration step at solving a set of linear algebraic equations. But the effective stiffness matrix must be repeatedly set up.

To perform solution of the equation of motion (30) it is assumed that at the beginning the system is in rest and takes the equilibrium position.

6 COMPUTATIONAL SIMULATIONS

The task was to investigate the influence of the damping effect of the magnetorheological dampers on the vibration amplitude of the rotor and of the chamber in the tank with water for two specified speeds of the rotor rotation (50 rad/s and 70 rad/s) and for the cases when the chamber is not submerged and is submerged to 1/3 of its radius.

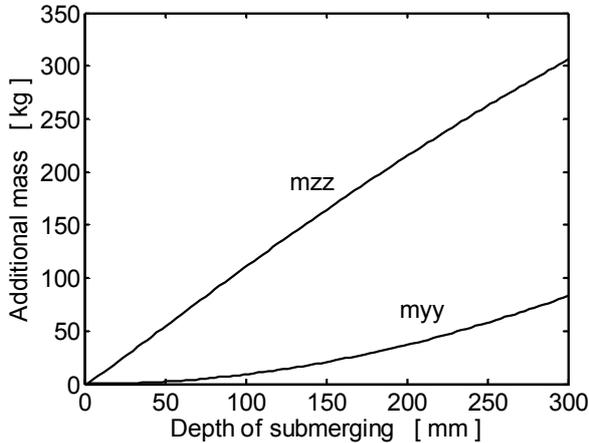


Figure 4: Additional masses

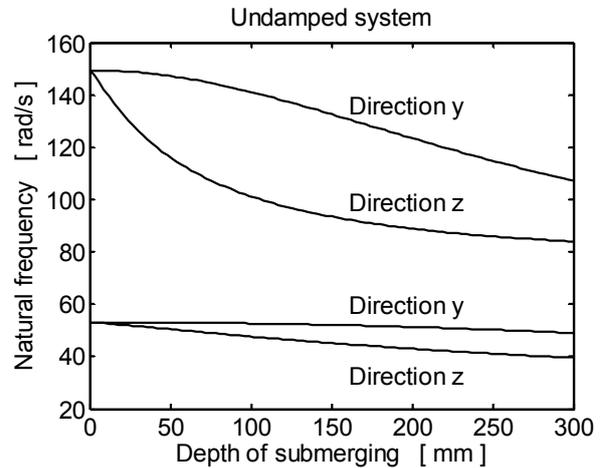


Figure 5: The submerged system natural frequencies

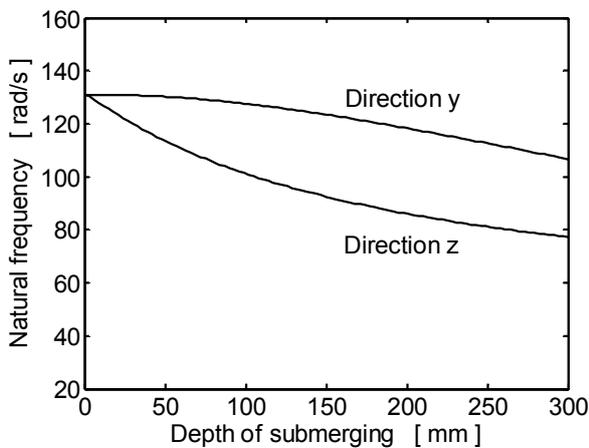


Figure 6: The overdamped system natural frequencies

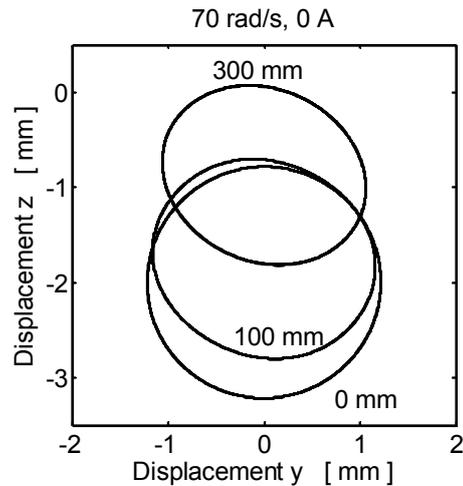


Figure 7: Positions of the chamber centre orbits

In Fig.4 one can see that the additional masses related to the horizontal (m_{yy}) and vertical (m_{zz}) directions grow up with increasing depth of submerging. Rising magnitudes of the additional masses arrives at the decrease of the system undamped natural frequencies (Fig.5). The splitting of the courses is caused by different magnitudes of the additional masses in the horizontal and vertical directions. If the coupling between the rotor and the chamber is strongly overdamped, then the rotor and the chamber behave as one rigid body. The dependence of its natural frequencies on the submerging depth is drawn in Fig.6. As evident

the change of damping in the magnetorheological dampers changes the system modal properties and shifts its natural frequencies. This can be utilized for attenuation of the oscillation amplitude.

In Fig.7 one can see the affect of the lifting force on the orbits of not submerged (0 mm) and partly submerged (100 mm, 300 mm) chamber. The trajectory centers are pushed upward to the free water surface.

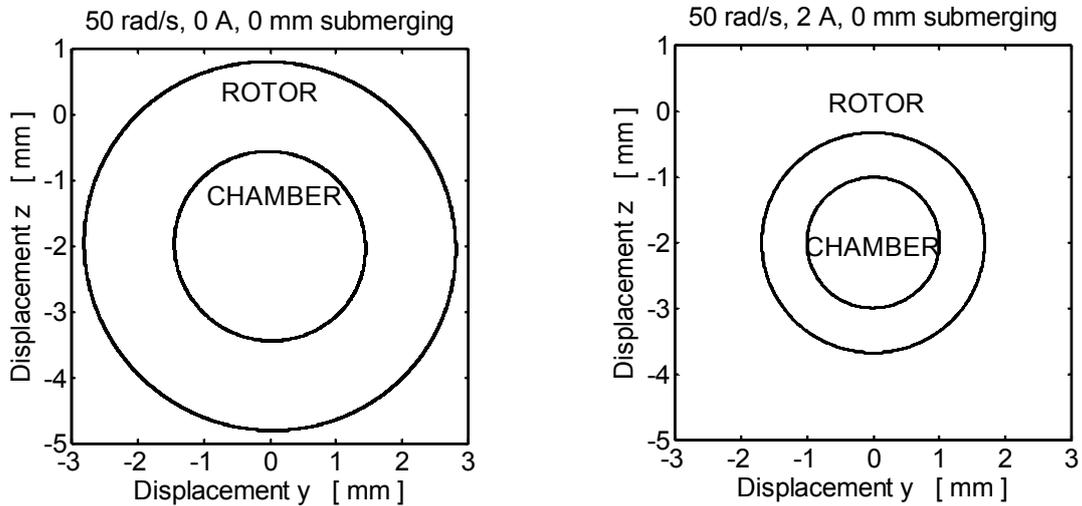


Figure 8: Orbits of not submerged chamber at speed of the rotor rotation 50 rad/s

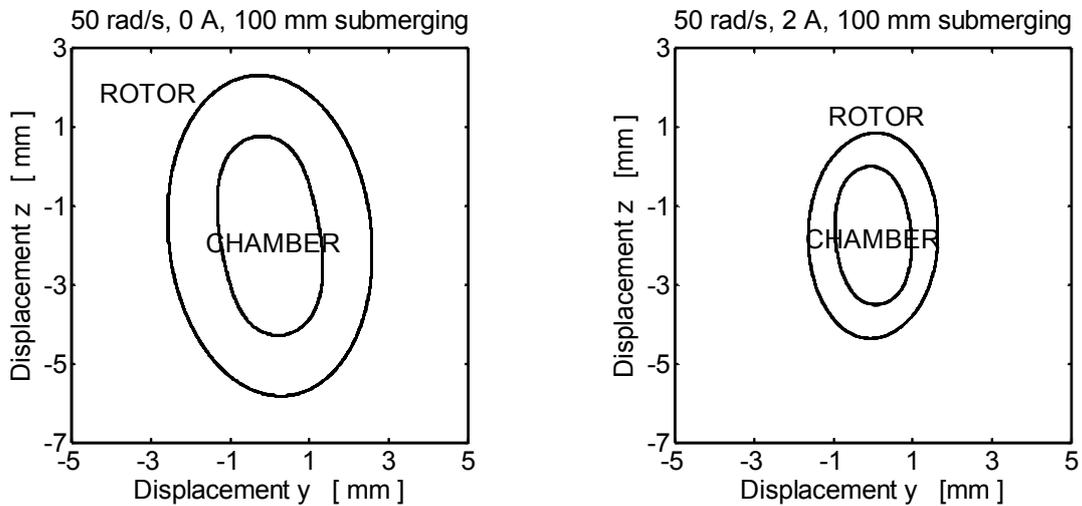


Figure 9: Orbits of submerged chamber at speed of the rotor rotation 50 rad/s

In Fig. 8 to 10 there are drawn orbits of the rotor and chamber centres for increasing magnitudes of the current controlling the damping effect of the magnetorheological dampers. Its evident that for the speed of the rotor rotation 50 rad/s rising damping (rising magnitude of the applied current) arrives at higher reduction of the oscillation amplitude. Comparing results

in Fig. 8 and 9, the displacements are larger if the chamber is partly submerged. Submerging the chamber also changes the shape of the orbits. Fig. 10 shows a different behaviour of the system. The rising current leads to increase of the size of orbits of the rotor and the chamber centres.

The observed phenomena can be explained by the change of two system parameters : amount of damping in the magnetorheological dampers and influence of the inertia effects of the liquid (additional mass of the water), in which the chamber is submerged, and by nonlinear character of the magnetorheological damping force. These parameters lead to changing the course of the resonance curves and to shifting the resonance peaks. The obtained results confirm that to reach the minimum amplitude of the system vibration the damping effect must be controllable to be possible to be adapted to the current operating conditions.

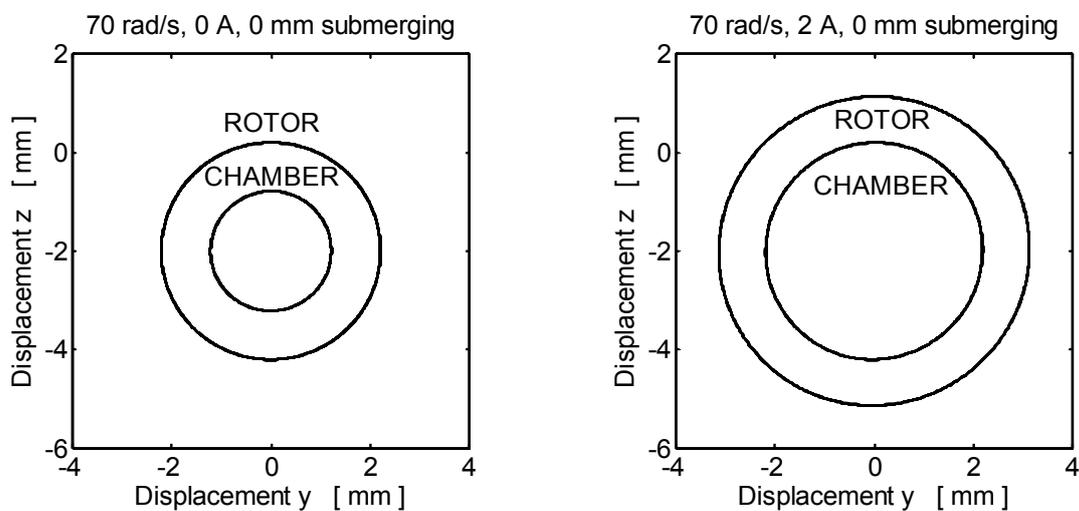


Figure 10: Orbits of not submerged chamber at speed of the rotor rotation 70 rad/s

7 CONCLUSIONS

The developed procedure represents a tool for analysis of the transient and steady state vibration of rigid rotors placed in chambers submerged in liquids. The variable depth of submerging changes the system parameters, which has the influence on the system resonance frequencies and critical speeds of the rotor rotation. Amplitude of the induced vibration can be significantly reduced if damping devices are inserted between the rotor and its stationary part fixed with the chamber. To achieve their optimum performance, the damping effect must be controllable. For this purpose application of squeeze film magnetorheological dampers were investigated. As the damping effect depends on magnitude of the magnetic flux passing through the layer of magnetorheological fluid, the change of the current generating the magnetic field can be used for its control. Results of the performed computational simulations confirm that the magnetorheological dampers make possible to adapt the damping effect to the current operating conditions and to minimize amplitude of the system vibration. Advantage of the proposed damping mechanism is that it always produces some amount of damping, which if needed, can be increased. In addition it does not require a complicated and

expensive control system and makes possible to achieve the optimum performance of the rotating machine by means of adapting the damping effect to the current operating conditions.

Investigating attenuation of vibration of rotating machines working in chambers submerged in liquids by means of application of magnetorheological dampers is a complicated technological problem. Its solving requires to analyze several mutually coupled physical problems (mechanical, hydraulic, electrical, magnetic).

ACKNOWLEDGEMENT

The work presented in this paper was supported by the research projects GA P101/10/0209, AVO Z20760514 and MSM 6198910027. The support is gratefully acknowledged.

REFERENCES

- [1] Wang, J. and Meng, G. and Hahn, E.-J. Experimental study on vibration properties and control of squeeze mode MR fluid damper-flexible rotor system. *Proceedings of the ninth Engineering Conference*, Chicago, (2003).
- [2] Forte, P. and Paterno, M. and Rustighi, E. A magnetorheological fluid damper for rotor applications. *International Journal of Rotating Machinery* (2004) **10**:175-182.
- [3] Forte, P. and Paterno, M. and Rustighi, E. A magnetorheological fluid damper for rotor applications. *Proceedings of the IFToMM Sixth International Conference on Rotor Dynamics*, Sydney, (2002).
- [4] Wang, G.-J. and Feng, N. and Meng, G. and Hahn, E.-J. Vibration control of a rotor by squeeze film damper with magnetorheological fluid. *Journal of Intelligent Material Systems and Structures* (2006) **17**:353-357.
- [5] Zapoměl, J. and Ferfecki, P. Mathematical modelling of a long squeeze film magnetorheological damper for rotor systems. *Modelling and Optimization of Physical Systems* (2010) **9**:97-102.
- [6] Zapoměl, J. and Ferfecki, P. Mathematical modelling of a short magnetorheological damper. *Transactions of the VŠB – Technical University of Ostrava, Mechanical Series* (2009) **LV**:289-294.
- [7] Zapoměl, J. and Ferfecki, P. A computational investigation of vibration attenuation of a rigid rotor turning at a variable speed by means of short magnetorheological dampers. *Applied and Computational Mechanics* (2009) **3**:411-422.
- [8] Zapoměl, J. The approach to computational analysis of vibration of vertical rotors supported by fluid film bearings and having a disc submerged in an inwettable liquid. *Proceedings of the 8th World Congress on Computational Mechanics*, Venice, (2008), CD 1-2.
- [9] Zapoměl, J. The procedure for numerical investigation of vibration of a rotor with fluid film bearings and a disc submerged in inwettable liquid. *Proceedings of the 9th International Conference on Vibrations in Rotating Machinery*, Exeter, (2008).
- [10] Krämer, E. *Dynamics of rotors and foundations*. Springer-Verlag, (1993).
- [11] Bathe, K.-J. *Finite Element Procedures in Engineering Analysis*. Prentice-Hall, (1982).
- [12] Levy, S. and Wilkinson, J.P.D. *The component element method in dynamics with application to earthquake and vehicle engineering*. McGraw-Hill, (1976).

DESIGN, SIMULATION AND PROTOTYPING OF POLYMERIC ORTHO-PLANAR MICROVALVES

EMANUELE BERTARELLI^{*,†}, ELENA BIANCHI^{*}, DIRK STROHMEIER[†],
JAN G. KORVINK^{†,§} AND GABRIELE DUBINI^{*}

^{*} Laboratory of Biological Structure Mechanics,
Department of Structural Engineering, Politecnico di Milano,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
e-mail: bertarelli@stru.polimi.it

[†] Laboratory for Simulation,
Department of Microsystems Engineering, University of Freiburg,
Georges-Koehler-Allee 103, 79110 Freiburg, Germany

[§] School of Soft Matter Research,
Freiburg Institute of Advanced Studies, University of Freiburg,
Albertstrasse 19, 79104 Freiburg, Germany

Key words: Microvalve, Rapid prototyping, Fluid-structure interaction.

Abstract. In this work, a passive ortho-planar microvalve design is investigated. The component is conceived to be integrated in micropumps and microfluidic systems. It is very compact and it can be realized through standard planar machining techniques, such as laser rapid prototyping or MEMS surface micromachining.

Analytical and Finite Element coupled approaches are used to describe the microvalve response to the fluid flow. A validation of the models is successfully carried out with respect to literature data for silicon microvalves. The proposed analytical model well describes the hydraulic resistance. Eventually, a calibration through a limited number of numerical simulations can be done. Subsequently, polymeric prototypes are realized and tested. The manufactured microvalves are effective in flow rectification and a power law $Q - \Delta p_V$ trend is evidenced. As expected, deviations are found with respect to the modeled ideal response. This seems to be due to tolerances in assembly and high structure compliance.

1 INTRODUCTION

Microvalves are fundamental components in microscale fluid manipulation systems. As an example, in reciprocating micropumps valves have to be used to convert the non-directional fluid motion in a directional flow [1].

In this research, a passive ortho-planar valve design is considered, suitable to be used for flow rectification in micropumping devices [2; 3]. The device fully develops in plane, while it works outside the plane through an orthogonal motion [4]. A major advantage of this type of structures is that they are very compact and they can be easily manufactured with planar techniques, also at the microscale. Moreover, there is no need to actively drive valve action. This results in low complexity in both manufacturing and control.

Basically, the investigated microvalve consists of a central disk suspended by compliant elements over a circular inlet hole. The working principle is shown in Figure 1. If the pressure in the upper chamber is higher than the pressure in the lower one, the disk is pushed to close the hole. A higher pressure in the lower chamber will result in a lift of the boss, allowing the fluid to flow.

Fundamental requirements are a low leakage (especially in the reverse direction) and a low pressure drop in the flow direction. As a consequence, spring stiffness and geometry should be carefully selected [5]. The most elementary option to realize ortho-planar springs is represented by straight beams with an out-of-plane action; more convoluted spring geometries can be considered, such as folded beams [4–6]. Valve designs characterized by one or two suspension arms exhibit large leakage, caused by kipping motion of the disc [6]. A three beams design is then selected here.

A wide variety of shapes and materials can be envisaged, as well as different fabrication methods. A number of Authors proposed and successfully realized micro check valves exploiting the ortho-planar working principle. Recently, SU-8 microvalves at the millimeter scale have been studied by Nguyen *et al.* especially for implementation into micropumps and microfluidic devices [6]. Small *et al.* [5] manufactured and characterized analogous systems, with an in-plane characteristic size in the range of hundred of microns and a $100\ \mu\text{m}$ layer thickness. Interestingly, micromachined in-silicon ortho-planar microvalves realized through conventional etching techniques were demonstrated too [7; 8]. The in-plane characteristic size was in the range of a hundred of microns; a thickness of tenth of microns was utilized to obtain structures with an adequate compliance. Oosterbroek and collaborators also reported a complete description of the proposed fabrication process, suitable to be embedded in a standard MEMS microfabrication workflow [7].

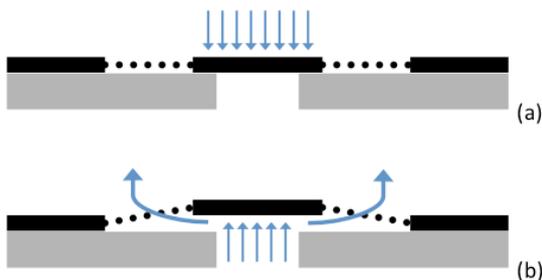


Figure 1: Schematic cross section of the microvalve in closed (a) and open state (b).

Although polymeric check valves have a much lower spring constant and a soft sealing surface that allows almost zero leakage [6], it has been shown that the micromachined in-silicon counterpart is characterized by a more predictable and reproducible behaviour (compare [5; 6], [7; 8]). Besides, the full compatibility with MEMS fabrication processes and the possibility to easily scale down the size of the device are important advantages.

In this work, analytical and Finite Element fluid-structure coupled approaches are adopted to describe microvalve response under the assumptions of rigid boss, linear elastic response of beam springs and zero leakage. A validation of the models is successfully carried out with respect to literature data for silicon microvalves [7]. Subsequently, polymeric prototypes are designed, realized through laser rapid prototyping, and tested.

2 ANALYTICAL MODEL

To perform the check valve design, a simple analytical model is defined first. The response of the structure under fluid load is described as follows. An ideally rigid disc with radius R_d is considered. Three equally spaced, straight radial beams with length L , width w and thickness t are designed. The fluid load is assumed to act on the boss only; the action of fluid flow on beams is neglected (Figure 2). This model exhibits radial symmetry.

A linear spring response of the beams is assumed. The force versus displacement relationship for the system is

$$F = k_V \delta \quad (1)$$

The total fluid load acting on the disc is assumed to be equally divided between each arm. Beams are clamped on the outer edge and subjected to a load $F_t = F/3$ at the tip, where orthogonal (out of plane) displacement only is allowed. For straight, slender beams ($L/t < 20$) in small deformations regime ($t/\delta < 10$), the displacement at the tip is

$$\delta = \frac{F_t L^3}{12EI} \quad (2)$$

where E is the Young's modulus of the material and $I = wh^3/12$. The equivalent spring constant become then

$$k_V = \frac{F}{\delta} = \frac{3Ewt^3}{L^3} \quad (3)$$

The expression of the force due to the fluid load, for a given pressure drop Δp_V , when

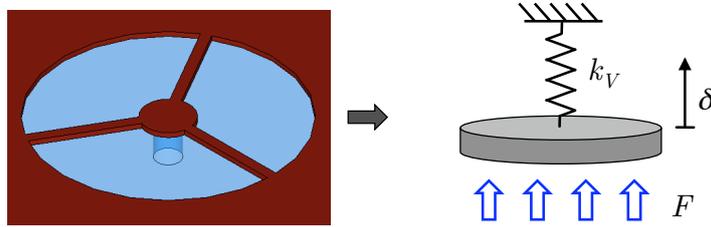


Figure 2: Representation of the model adopted to describe valve mechanical response: the displacement δ of the rigid boss under the fluid load F is described through an equivalent linear spring of constant k_V .

the internal radius is R_i and the external radius R_d , is [5]

$$F = \int_0^{2\pi} \int_0^{R_i} \Delta p_V r dr d\theta + \int_0^{2\pi} \int_{R_i}^{R_d} \frac{\Delta p_V}{R_d - R_i} (r - R_d) r dr d\theta = \frac{\pi}{3} \Delta p_V (R_d^2 + R_d R_i + R_i^2) \quad (4)$$

Finally, disc displacement δ can be expressed with respect to pressure drop:

$$\delta = \frac{\pi}{9} \frac{L^3}{Ewh^3} \Delta p_V (R_d^2 + R_d R_i + R_i^2) \quad (5)$$

At this point, the fluid flow through the microsystem is treated. Fluid is assumed to be water. In small scale systems, the gravity effect exhibits a negligible effect; thus, the correspondent body force is coherently neglected. The flow of an incompressible fluid in stationary conditions can be described by the generalized Bernoulli equation [9].

The fluid path through the valve can be conveniently split into two contributions [10]: the vertical inlet channel and the meatus with variable thickness δ between the disc and the base layer. Radial symmetry is considered. Along the inlet channel, the flow profile is not strictly unidimensional, and a rigorously correction on the kinetic energy terms have to be considered [11]. However, since the most relevant contribution to the overall pressure drop is expected in the meatus, the contribution of this part will be neglected.

In the meatus, laminar flow loss is considered, by applying a derivation of the Navier-Stokes formula [12]. For microchannels with a width d such that $d/\delta > 4.5$:

$$\Delta p = \frac{12\mu_w L Q}{d \delta^3} \quad (6)$$

where μ_w is water viscosity. The section is an annulus. Thus, the equivalent width can be evaluated by performing an integration as follows:

$$\Delta p = \int_{R_i}^{R_d} \frac{12\mu_w Q}{2\pi r \delta^3} dr = \frac{6\mu_w Q}{\pi \delta^3} \ln \left(\frac{R_d}{R_i} \right) \quad (7)$$

Now, by combining Equations 3, 5 and 7, the flow through the valve can be computed:

$$\begin{cases} \Delta p_V(\delta) = \frac{6\mu_w Q}{\pi \delta^3} \ln \left(\frac{R_d}{R_i} \right) \\ \delta(\Delta p_V) = \frac{\pi}{3} k_V \Delta p_V (R_d^2 + R_d R_i + R_i^2) \end{cases} \quad (8)$$

It is worth emphasizing that the solution above effectively considers the coupling between the solid domain and the fluid domain. The relation between pressure drop Δp_V and flow rate Q in Equation 8a depends on device geometry and valve opening $\delta(\Delta p_V)$, thus inherently on the structural response of the equivalent spring that represents the

beams. The deformation of the system computed in Equation 8b is determined by the fluid load, here explicitly represented in terms of the pressure drop Δp_V .

To summarize, the fluid flow Q can be expressed as

$$Q \propto k_V^{-3} \cdot \Delta p_V^4 \cdot \Gamma(R_i, R_d) \quad (9)$$

where $\Gamma(R_i, R_d)$ is a geometrical function. From Equation 9, it can be seen that the microvalve fluidic component is expected to be characterized by a fourth order power law $Q - \Delta p_V$ relationship. The effect of valve spring constant is relevant, since a third order relation with respect to k_V is found. Oosterbroek [7] developed an analogous analytical model and proposed an interesting adimensionalization of the full problem. Differently, the flow through valve gap was considered as a Poiseuille flow, but the same fourth order power law $Q - \Delta p_V$ and third order dependence on valve stiffness k_V were identified. The design of the compliant structure is a crucial aspect in ortho-planar microvalve development.

3 FINITE ELEMENT MODEL

With COMSOL Multiphysics 3.5a (COMSOL AB), a fully-coupled Fluid-Structure Interaction (FSI) Finite Element model is developed. A 2D model is defined (Figure 3), since a 3D model would be very expensive from a computational standpoint. Rigid boss assumption is in fact relieved, but the motion of the disc is constrained due to imposed radial symmetry.

A preliminary study with structural 3D models showed indeed that, for the proposed geometry, the assumption of radial symmetry during valve opening due to a uniform fluid load is not always respected. In particular, when the valve is realized with a material with $E \simeq 1 - 10$ GPa (here, $E = 3.5$ GPa), the boss deformation is relevant, leading to a wavy profile and a non-uniform meatus thickness. Thus, higher flow rate can be expected. When $E \simeq 100$ GPa or higher (e.g. for polysilicon, $E = 160$ GPa), the radial symmetry assumption is reasonably fulfilled. The 2D model here introduced do not aim therefore to provide a precise representation of polymeric valve response, but to give interesting information on valve functioning and to provide a comparison with the analytical approach above expounded.

In the FSI axisymmetric model, at inlet and outlet of the fluid domain a pressure boundary condition is imposed. Concerning the fluid-structure interaction, at the fluid/solid interface boundary velocity and fluid load are exchanged. The disc can move along the axial direction according to the response of the compliant elements. The elastic response of beams is assumed to be concentrated on a linear spring of constant k_V . For implementation purposes, spring action is distributed and applied as a pressure over the upper disc surface. In the present model, complete valve closure cannot be achieved. This would require the fluid domain in the gap to collapse on a line. This is not allowed in a standard Finite Element formulation. Thus, a manual remeshing is performed, when mesh deformation in the domain degenerates in element inversion.

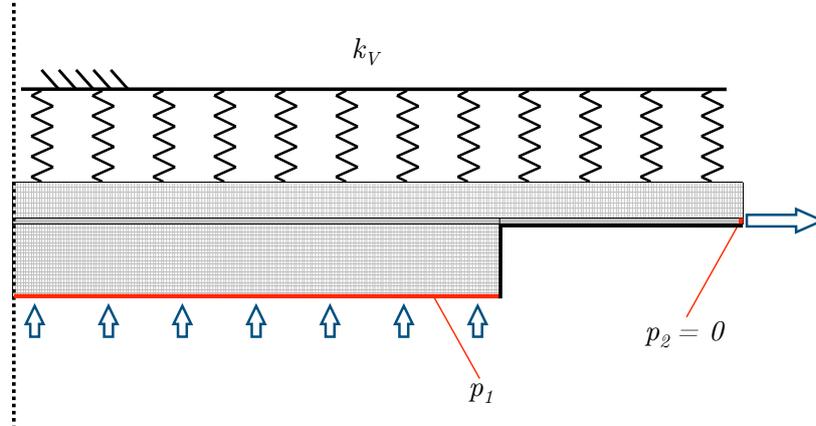


Figure 3: Axisymmetric Fluid-Structure Interaction Finite Element model of the microvalve.

Simulations are performed starting from an open valve configuration. The nodal force acting on the upper surface is evaluate as $\tilde{F}_{n,z} = k_v (w + \delta_0)$, where δ_0 is the valve gap at the beginning of the simulation, and w is the displacement. First, the system is initialized by means of the analytical model. Then, a portion of the $Q - \Delta p_V$ is swept. The solution scheme can be summarized as follows:

1. Initialization:

- For a trial pressure drop across the valve Δp_V^* (ideally, the upper or lower limit of the interval of interest), a guess for boss equilibrium position δ_0^* is obtained from Equation 5.
- The model geometry is drawn. Inlet pressure $p_1 = \Delta p_V^*$ is imposed.
- A value of the flow Q^* for the given configuration is obtained, for the fluid domain only (fixed disc). Field solution over the fluid domain is stored.
- The value of the flow Q is computed again, for full model, starting from stored solution for fluid domain. Field solution over whole model is stored.

2. Curve sweep:

- A small variation in the inlet pressure p_1 is imposed, and the new value of Q is computed. At each step, the field solution over whole model is stored.
- The procedure is repeated until excessive mesh deformation is found. Then, manual remeshing is performed, and the solution is restarted from the last step.

In this solution scheme, the field solution over a domain or the whole model is stored at each step. The solution is then used as initial condition for the further step of the analysis. This originates a relevant decrease of the computational cost of each iteration, since a guess, very close to the solution, is used.

4 PROTOTYPING

With the aim to realize a prototype of planar microvalves, laser rapid prototyping has been considered. This process can be adopted to easily manufacture a variety of planar valve designs, on various materials. The valves here reported are realized on Polyimide/Kapton film (DuPont). Machining is realized with a TruMark Series 6000 laser (Trumpf). Nominal dimensions $w = 125 \mu\text{m}$, $R_i = 250 \mu\text{m}$, $R_d = 1000 \mu\text{m}$, $L_i = 1500 \mu\text{m}$ are defined for all the valves, while different beam widths $w = 300 \mu\text{m}$, $400 \mu\text{m}$, $500 \mu\text{m}$ are selected.

After the valve layer is realized, PMMA top and bottom layers containing inlet and outlet channels are realized. Layers are finally stacked to obtain a sandwich – with the valve layer in the middle – by using a double-sided high performance acrylic adhesive tape (3M 467MP with 200MP adhesive). Assembly is done by means of alignment screws. Process is shown in Figure 4.

Finally, connection of microvalves to the external fluidic circuit is achieved by using a NanoPort connector (Upchurch Scientific) at the inlet. At the outlet a Polyimide conical connector is used; it is kept in place by a mechanical contrast, while an optimal sealing is achieved by realizing PDMS seals. The $Q - \Delta p_V$ characterization is performed in flow control. A flow rate is imposed by means of a syringe pump (Harvard PHD2000), and pressure drop across the valve is measured with a differential pressure transducer (Honeywell 142PC05D) read by a function implemented through LabView (National Instruments). To finally obtain the characteristic $Q - \Delta p_V$, a series of static measurements is performed varying the imposed flow rate.

5 RESULTS

5.1 Modelling

To perform a consistent validation of the proposed models, experimental data for silicon micromachined microvalve are considered [7]. This choice is made since this component during functioning fulfills the hypotheses of rigid boss and linear elastic response of beam springs. Comparison with experimental data is satisfactory (Figure 5).

The analytical formulation overestimates the flow. The valve characteristic obtained by the Fluid-Structure Interaction fully-coupled model very well represents the system response in the considered range. Interestingly, the discrepancy between the two proposed approaches is constant, and can be compensated by a correction factor on the flow rate. In this case, the ortho-planar circular microvalve response can be represented by an expression in the form

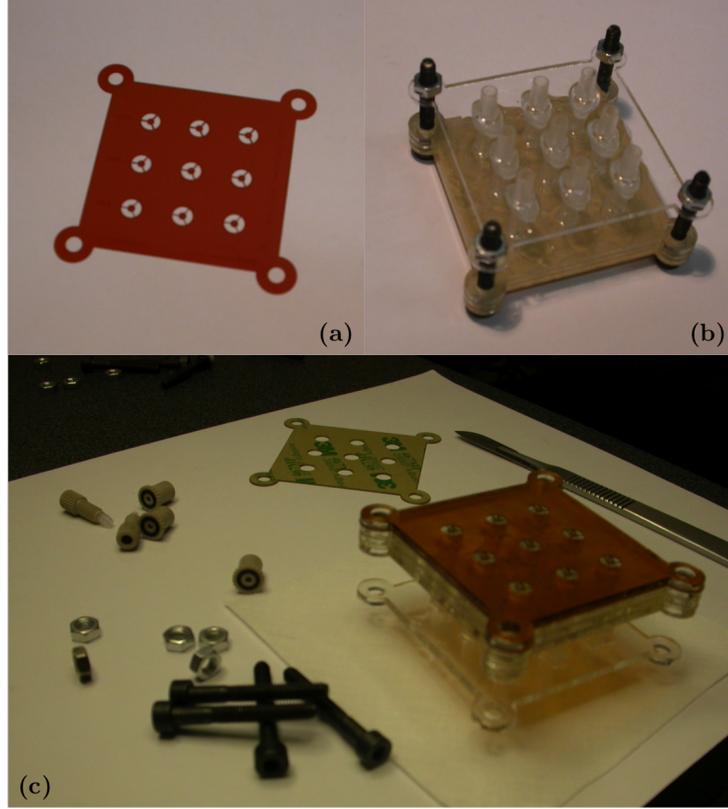


Figure 4: Microvalves prototypes: (a) Kapton valve layer; (b) outlet part, with conical connectors and contrast; (c) application of valve layer on the outlet part.

$$Q \simeq k_Q \cdot \frac{\pi^4}{162} \cdot \frac{\Delta p_V^4}{\mu_w k_V^3} \cdot \frac{(R_d^2 + R_d R_i + R_i^2)^3}{\ln\left(\frac{R_d}{R_i}\right)} \quad (10)$$

where $k_Q \simeq 0.765$ is a correction factor, which can be extracted by means of a small number of FSI simulations.

5.2 Prototype testing

In Figure 6a, results of microvalve characterization is reported. The tests evidenced the possibility to successfully obtain flow rectification. The three different geometries here proposed differ by spring stiffness: the measured flow rates coherently reflect this. While for a given pressure drop across the valve the flow rate is proportional to spring stiffness, the response cannot be in general identified as a power law as expected from models. Indeed, the more compliant valves ($w = 300 \mu\text{m}$ and $w = 400 \mu\text{m}$) exhibit a response which is somehow linear, similarly to polymeric microvalves reported by Small and co-workers [5].

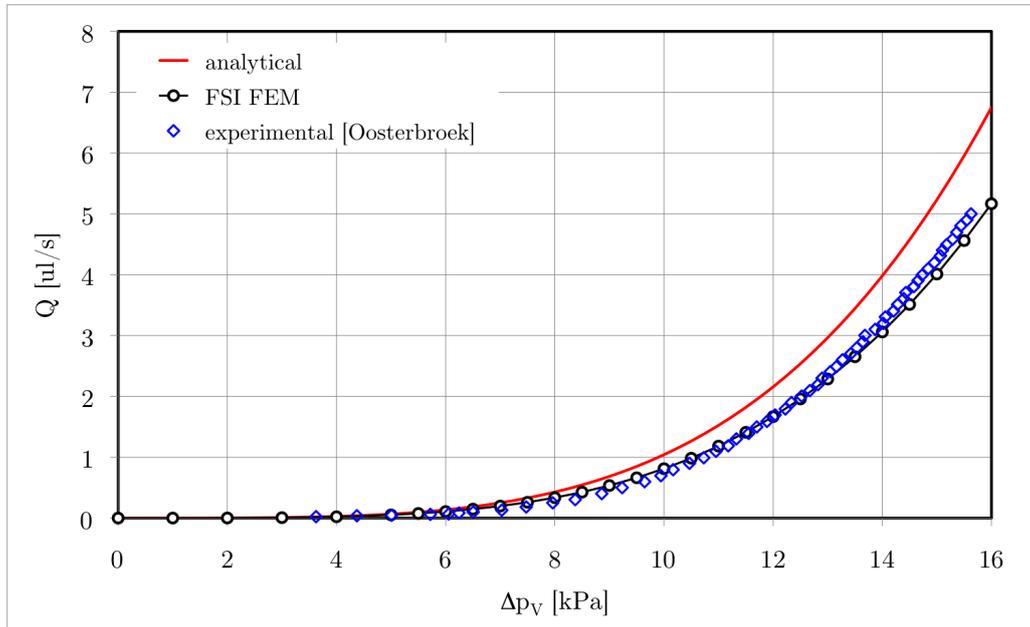


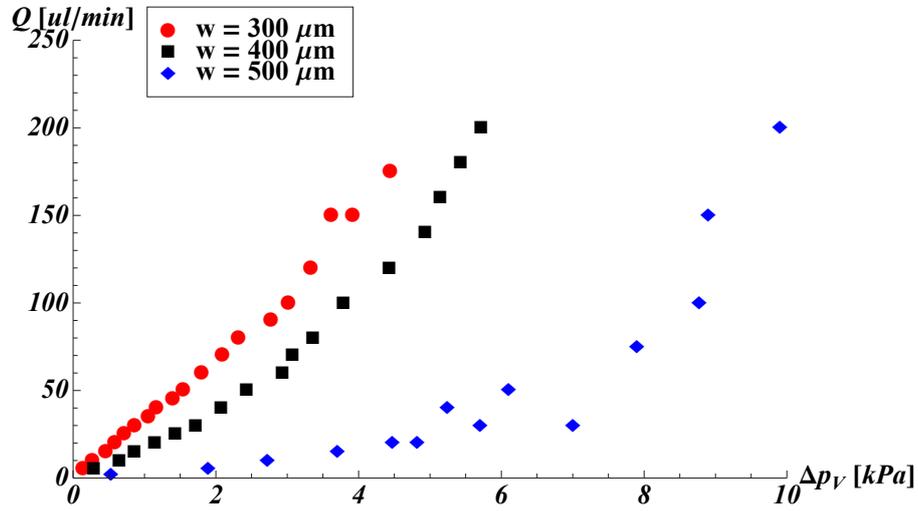
Figure 5: Microvalve flow curve: comparison of the proposed models with literature experimental data.

In fact, this behaviour can be explained by the limitations concerning the manufacturing process and the assembly. In particular, surface finish could be too rough and some asymmetry in the assembly could be present [5]. Moreover, melting due to laser cutting originates irregular edges and features size slightly differs from the nominal ones, as evidenced by optical microscope observations. These aspect, combined with a low spring stiffness, can generate a relevant leakage. In practice, the flow through the valve is basically determined by this leakage.

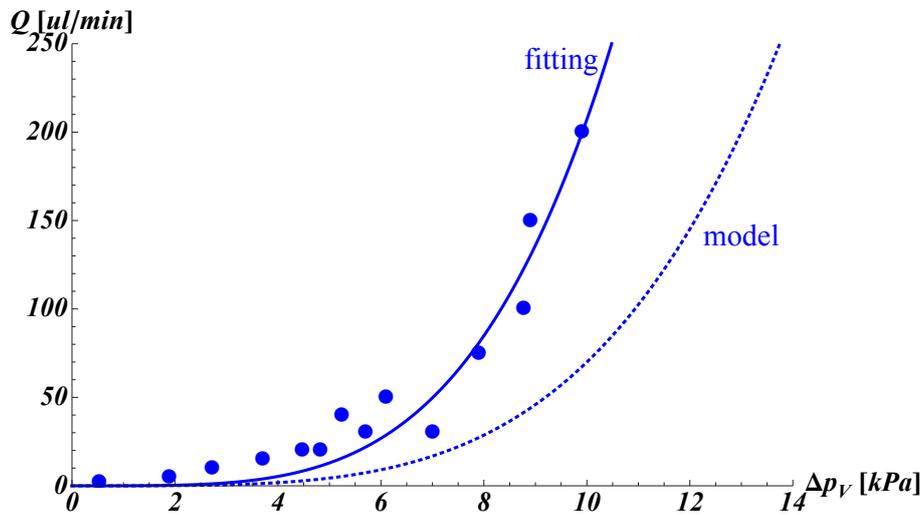
Concerning the stiffer valve ($w = 500 \mu\text{m}$), a fourth order power law response can be reasonably identified. However, as evidence in Figure 6b, from the analytical model a stiffer response was expected. From a qualitative point of view, it can be argued that the higher spring stiffness of this valve design is capable to partially overcome the leakage due to manufacturing issues. Besides, the non-uniform meatus thickness that comes from the high compliance of the boss – which could be appreciated through a 3D modelling of the device – can explain the higher measured flows.

6 CONCLUSIONS

In this work, the design of valves to be implemented in microsystems for microscale fluid transport is illustrated. The proposed planar shape is compact and valves are suitable to be realized by means of standard standard planar machining techniques, such as laser rapid prototyping or MEMS surface micromachining.



(a)



(b)

Figure 6: Characteristic flow rate versus pressure drop curve for the polymeric microvalves: (a) Comparison of experimental results for different valve geometry; (b) Fourth order power law data fitting of experimental results on $w = 500 \mu\text{m}$ microvalve, and prediction from the analytical model.

Analytical and Finite Element coupled approaches are defined to study the microvalve response to fluid flow, under the assumptions of rigid boss, linear elastic response of beam springs and zero leakage. A successful validation of both modelling strategies is achieved by comparison with experimental data available in the literature. It has been found that the proposed analytical model well describes the hydraulic resistance. Eventually, a calibration through a limited number of numerical simulations can be done.

Subsequently, polymeric prototypes are designed and realized through laser rapid prototyping. The manufactured microvalves are effective in flow rectification, while deviations are found with respect to the expected ideal response which are not captured by the proposed flow models. The deviations can be explained by limitations concerning the manufacturing process and the assembly, but also by considering that the proposed models are not suitable to describe in details the response of highly deformable microvalves. Indeed, a satisfactory agreement is found, while for a precise modeling 3D models able to capture structure non-symmetric deformation should be envisaged. It is worth noting that the modelling and simulation of polymeric microvalves is a challenging task. It is meaningful that, up to now, no effective simulation has been reported for this class of microvalves.

References

- [1] Iverson, B. and Garimella, S., Recent advances in microscale pumping technologies: a review and evaluation. *Microfluid. Nanofluid.* (2008) **5**:145–174.
- [2] Bertarelli, E., Ardito, R., Cioffi, M., Laganà, K., Procopio, F., Baldo, L., Corigliano, A., Contro, R. and Dubini, G., Design optimization of an electrostatic micropump: a multi-physics computational approach. In: *Proceedings of the 2nd South-East Conference on Computational Mechanics*, 181 & CD-ROM (2009).
- [3] Bertarelli, E., Ardito, R., Bianchi, E., Laganà, K., Procopio, F., Baldo, L., Corigliano, A., Dubini, G. and Contro, R., A computational study for design optimization of an electrostatic micropump in stable and pull-in regime. *AES Tech. Rev. B: IJAMAIM* (2010) **1**:19–15.
- [4] Parise, J., Howell, L. and Magleby, S., Ortho-planar linear-motion springs. *Mech. Mach. Theory* (2001) **36**:1281–1299.
- [5] Smal, O., Dehez, B., Raucent, B., Volder, M. D., Peirs, J., Reynaerts, D., Ceysens, F., Coosemans, J. and Puers, R., Modelling, characterization and testing of an ortho-planar micro-valve. *J. of Micro-Nano Mechatronics* (2008) **4**:131–143.
- [6] Nguyen, N.-T., Truong, T.-Q., Wong, K.-K., Ho, S.-S. and Low, C.-N., Micro check valves for integration into polymeric microfluidic devices. *J. Micromech. Microeng.* (2004) **14**:69–75.

- [7] Oosterbroek, R., *Modeling, design and realization of microfluidic components*. PhD thesis, Universiteit Twente (1999).
- [8] Oosterbroek, R., Berenschot, J., Schlautmann, S., Krijnen, G., Lammerink, T., Elwenspoek, M. and van den Berg, A., Designing, simulation and realization of in-plane operating micro valves, using new etching technique. *J. Micromech. Microeng.* (1999) **9**:194–198.
- [9] Truckenbrodt, E., *Fluidmechanik – Band 1: Grundlagen und Elementare Strömungsvergänge dichtebeständiger Fluide* (2008). Springer.
- [10] Ulrich, J. and Zengerle, R., Static and dynamic flow simulation of a KOH-etched microvalve using the finite-element method. *Sensor and Actuat, A* (1996) **53**:379–385.
- [11] White, F., *Fluid Mechanics*, 4th edition (1998). McGraw-Hill International Editions.
- [12] Bendib, S. and Français, O., Analytical study of microchannel and passive microvalve: application to micropump simulator. In: *Proceedings of SPIE - Design, Characterization, and Packaging for MEMS and Microelectronics II*, 283–291 (2001).

ELECTROSTATIC DIAPHRAGM MICROPUMP ELECTRO-FLUID-MECHANICAL SIMULATION

EMANUELE BERTARELLI, RAFFAELE ARDITO
AND ALBERTO CORIGLIANO

Department of Structural Engineering, Politecnico di Milano,
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
e-mail: bertarelli@stru.polimi.it

Key words: Electrostatic micropump, Multiphysics simulation, Electro-fluid-mechanical coupling, Biological micro-electro-mechanical systems (BioMEMS)

Abstract. In this work, a fully-silicon mechanical displacement micropump is proposed and investigated. Electrostatic actuation of a flexible diaphragm is used to generate the pressure difference required to transport the fluid at the microscale. The study is carried out by exploiting the Finite Element method in a multiphysics framework, considering simplified geometries and boundary conditions. These investigations suggest the possibility to adopt the proposed device for applications in biomedical and biological fields. Achievable stroke volumes and flow rates are computed: values are in line with those obtained for similar devices presented in the literature.

1 INTRODUCTION

Micropumping is emerging as a critical research area for many applications. Several extensive reviews concerning micropumps are available [1–3]; some of them have emphasized specific applications [4–6]. The motivations of this growing interest in micropumping can be found first in the need to develop miniaturized pumping mechanisms for applications in microfluidics. Indeed, a major research interest in this area is the development of autonomous platforms to perform a precise manipulation of biological samples or the administration of drugs and chemicals. It is worth mentioning that microfluidics is also an essential part of precision control systems for automotive, aerospace and machine tool industries [4]. Additionally, micropumps are being considered for applications in the thermal management of electronic components [6].

Nevertheless, the industrial transfer is still at the beginning. Some micropumps have already entered the marketplace, but to definitely obtain a broad diffusion, cost-optimised and fabrication-optimised designs and a good performance reproducibility are the next goals to be achieved [2]. The key step at this point is to perform technology-driven

studies, in which a micropumping principle is developed in the framework of a well-established fabrication technology. In particular, the perspective of taking advantage of MEMS Integrated Circuits batch-processing technology to obtain affordable and reliable devices is of great interest. Besides, a direct integration with electronic components can be achieved, which is a key feature for power supplying and control of the device, as well as for a straightforward embedding in the electronic devices to be cooled or the realization of devices for fluid actuation and manipulation along the same fabrication process [7]. Indeed, the possibility to realize effective microfluidic components with MEMS microfabrication technology has been proven [8].

In this work, a diaphragm displacement micropump is proposed. The motion of a flexible diaphragm through electrostatic actuation is used to generate a pressure difference and then to obtain the fluid transport. The electrostatic actuation is considered, since this choice shows some relevant advantages, such as low power consumption, fast response time and full MEMS compatibility [2]. Since no external actuation is required, the cost and complexity of the device is noticeably reduced with respect to other actuation solutions (e.g. piezoelectric). The main reference for device actuation design is the work by Zengerle *et al.* [9]. However, a different geometry is here proposed and the stable operation regime – avoiding actuator electrostatic collapse – is considered.

The micropump here presented is intended to be used for biomedical applications – such as drug delivery systems – or for microfluidics devices for biological fluid handling and analysis. Key requirements are high quality standards and low costs, since many drug delivery systems and biological analytical chips are disposable. In these contexts, the possibility to exploit microsystems fabrication to obtain cost-effective manufacturing in high volume, satisfying high requirements of accuracy and reliability, is a fundamental aspect and represents a research driving force. Biocompatibility issues have to be considered too, especially for implantable devices or components that are directly interfaced with the biological environment [10].

2 DEVICE DESIGN

The geometry of the in-silicon micropumping device under study is reported in Figure 1. A description of a micromachining process suitable to realize it is reported elsewhere [7]. A circular pumping chamber with $R = 2$ mm radius and $h_f = 10$ μm height is designed. The initial volume of the pumping chamber is $v_0 = \pi R^2 h_f \simeq 126$ nl. The diaphragm is realized in polysilicon, with a thickness h_m . A rigid ground electrode is placed above, with spacing h_c . The space between electrodes is assumed to be filled of a dielectric, namely air, with permittivity ε . The actuation voltage is such that $V_{max} = 60$ V.

Inlet and outlet ports are placed along the perimeter of the pumping chamber, in diametrically opposite positions. At this stage ideal valves are considered, with no pressure drop during fluid flow and no leakage if valve is closed. Herein, inlet pressure is referred as p_{in} and outlet pressure is referred as p_{out} . Inlet pressure can be regarded as the pressure of a reservoir which contains the fluid to be pumped, even pressurized to facilitate pumping.

The device functioning over a pumping cycle is illustrated in Figure 2. Starting from the rest position (a), when voltage is applied the diaphragm is loaded by an electrostatic pressure. This pressure is transferred to the fluid contained in the pumping chamber and generates a depression. When the chamber pressure is lower than the inlet pressure, the inlet valve opens and liquid flows into the expanding chamber (expansion stroke), while the diaphragm deflects as a consequence of the electrostatic force acting on it (b). When voltage is removed, the diaphragm bounces back. This increases the pressure in the pumping chamber (compression stroke). When the chamber pressure is higher than the outlet pressure, an amount of liquid is discharged through the outlet valve (c). In fact, the driving force in the emptying phase is given by the elastic recovery of the diaphragm.

For a generic pumping cycle, after fluid discharge the pressure in the pumping chamber will be equal to p_{out} . Indeed, the outlet valve closes when the pressure inside the pumping chamber is equal to the outlet pressure, since the system has reached an equilibrium condition. The pressure difference to overcome during the next filling is then $p_{out} - p_{in}$. Conversely, the pressure in the pumping chamber at the end of filling is p_{in} , then the pressure difference to overcome is then again $p_{out} - p_{in}$. Summarizing, since ideal valves are considered, the pressure drop to overcome during a pumping cycle, in both expansion and compression stroke, is given by the difference between the outlet pressure p_{out} and the inlet pressure p_{in} :

$$\Delta p = p_{out} - p_{in} \quad (1)$$

On the assumption that all the charges are localized within a thin layer below the electrode surface and that there is no bending of the diaphragm towards the counterelectrode, a lower limit for the electrostatic generated pressure is [9]

$$\tilde{p}_{el} = \frac{\varepsilon V^2}{2h_c^2} \quad (2)$$

Equation 2 represents the minimum value of the electrostatic force, calculated in rest position. Once the diaphragm deflects, the distance between the plates decreases and an higher (although non-uniform) pressure is exerted on the diaphragm on the new configuration [11].

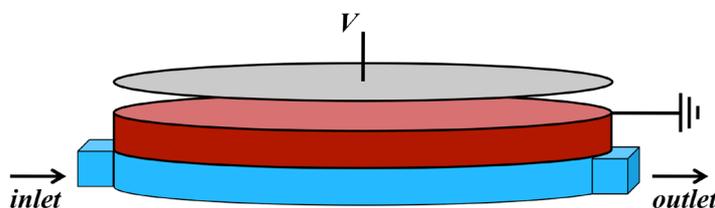


Figure 1: The proposed device design. Rigid electrode (upper gray plate), diaphragm (in red) and pumping chamber (in blue) are represented (not to scale).

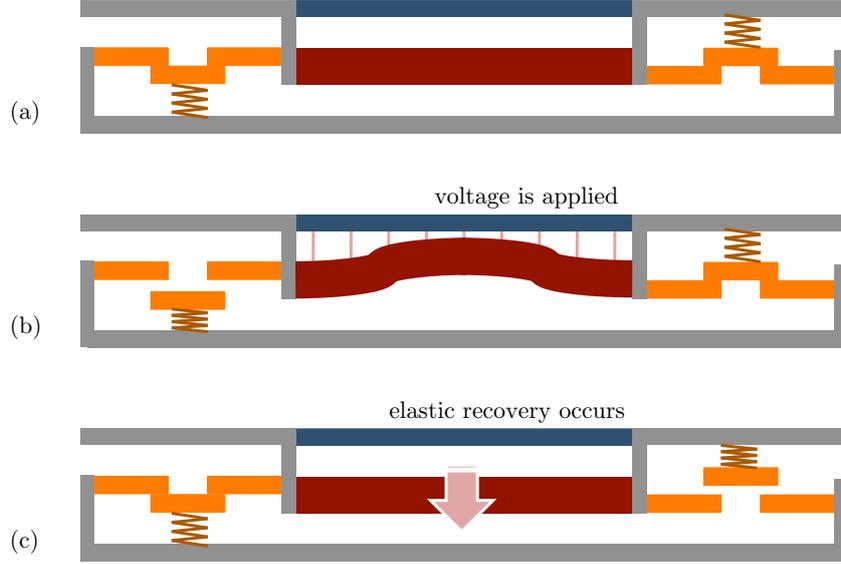


Figure 2: Device functioning over a pumping cycle: rest position (a), filling (b) and fluid discharge (c). Valve action is schematically represented.

To start a pumping cycle, the actuator electrostatic pressure in rest position must be higher than Δp . This is required to generate a sufficient chamber depression to open the inlet valve. This condition can be represented by the inequality

$$\frac{\varepsilon V^2}{2 h_c^2} > \Delta p \quad (3)$$

Equation 3 expresses a minimum requirement. If matched, the diaphragm starts to move and the electrostatic pressure to rise, up to the desired deflection (Figure 3). In this study, inlet and outlet pressure is the same: $\Delta p = 0$ [9].

Fluid discharge is obtained in the proposed device by elastic recovery only. In previous works on micropumps simulation [12; 13], when outlet and inlet pressure is considered to be the same, the pumping chamber variation in quasi-static conditons v_s was assumed as stroke volume Δv , considering that the diaphragm ideally bounces back to the undeformed configuration at each pumping cycle:

$$\Delta v \simeq v_s \quad (4)$$

Then, for a given pumping frequency f_p , the ideal flow should be given by

$$\mathcal{Q}_i = v_s \cdot f_p \quad (5)$$

If this approach is applied to the device here investigated, a theoretical flow transport capability of $\mathcal{Q}_i = 1-15 \mu\text{l}/\text{min}$ for an actuation frequency of 10-50 Hz is estimated [7; 11].

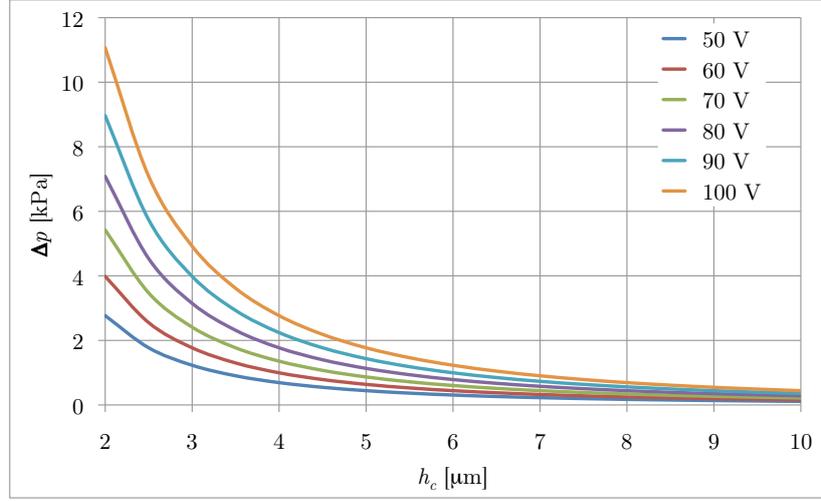


Figure 3: Limit value of the pressure drop $\Delta p = p_{out} - p_{in}$ that can be overcome, for different applied voltages and capacitor spacings.

Equation 4 is a reasonable approximation only if a negligible pressure drop is found in the pumping chamber and through inlet and outlet valves and channels. But the most relevant assumption is that the pump is capable to pump all the stroke at any frequency. In fact, both electrostatic active expansion and passive elastic compression require a finite time. As a consequence, increasing the frequency over a limit value (which is a characteristic of the system) leads to a decrease of the effective amount of fluid which is pumped at each stroke. Experimental measurements on micropumps prototypes evidenced that the effects here neglected might assume a relevant role and then influence in a considerable manner micropump performances [9; 14].

If the maximum stroke v_s is taken as design parameter, a pumping efficiency index can be defined as

$$\mathcal{E}(f_p) = \frac{\Delta v(f_p)}{v_s} \cdot f_p \quad (6)$$

where $\Delta v \leq v_s$ represents the stroke obtained for a given pumping frequency. Finally, for a given design and pumping frequency f_p , the effective flow is

$$Q(f_p) = v_s \cdot \mathcal{E}(f_p) \quad (7)$$

3 DEVICE MODELLING

In this work, dynamic electro-fluid-mechanical simulations are performed, with the aim to appreciate the effect of the fluid on device actuation. To the knowledge of the Authors, this is the first reported simulation of the fully coupled dynamic response of an electrostatic micropump.

An effective, sophisticated compact approach was proposed by Voigt *et al.* [15; 16]. To date, this is the only complete implemented model of the electrofluidic microsystem previously realized by Zengerle *et al.* [9]. Each component is a block, connected to the others in a Kirchoffian network, in which driving forces and resulting fluxes are treated with a lumped parameter approach. Each block is described by a physically based compact model extracted from continuum theory or from experimental data. Nevertheless, the lack of Finite Element simulations of the effective pumping mechanism does not enable a precise characterization of either the fluid dynamic field or the diaphragm action. Multiphysics Finite Element simulations can be used as a design tool not only to optimize the overall performance of the system, but also to reduce both the time and cost of development.

Due to the extreme aspect ratio of the device, a 3D multiphysics fully-coupled model has been found to be too expensive from the computational standpoint. In order to describe inlet and outlet while preserving flow directionality, an equivalent 2D plane strain model is defined which represents a diameter slice of the device. Parameters and constants adopted are reported in Table 1. Since the plane strain formulation exhibits a diaphragm flexural stiffness lower than the clamped circular plate, a scaling is necessary to obtain the same deflections as in the axisymmetric formulation. A straightforward solution is to scale Young's modulus, defining an equivalent modulus $E_{ps} = (8/3) E$ such that the maximum deflection obtained in plane strain simulations is the same that in the previously performed axisymmetric analysis [7]. With an out of plane equivalent chamber depth $d_f = \pi R/2$, the initial volume v_0 of the chamber is preserved. The volume $v(t)$ during actuation is in general larger, due to the different diaphragm deformed configuration, then it has to be scaled. The results can be considered a qualitative representation of the device dynamic behaviour. In fact, some relevant considerations can be drawn.

Second order quadrilateral elements are used for mechanical, electrostatic and fluid domains, with large displacements formulation and mesh movement described by the arbitrary Lagrangian-Eulerian scheme, with no remeshing. Characteristic length of the elements is $\ell \simeq h_m/10$. Suitable boundary conditions are defined along sub-domain boundaries, in terms of electrostatic forces and displacements (for the electro-mechanical interaction), and boundary velocities and fluid load (for the fluid-structure interaction).

E	polysilicon Young's modulus	160 GPa
ν	polysilicon Poisson's ratio	0.22
ρ	polysilicon density	2320 kg/m ³
ε	dielectric (air) permittivity	$8.854 \cdot 10^{-12}$ F/m
ρ_w	fluid (water) density	1000 kg/m ³
μ_w	fluid (water) viscosity	0.001 Pa · s

Table 1: Parameters and constants adopted in the Finite Element model.

4 RESULTS AND DISCUSSION

In Figure 4, square wave actuation is simulated at $f_p = 3$ Hz, for a device with $h_m = 30$ μm and $h_c = 6$ μm . Voltage is applied for half cycle, then switched off during the second half cycle. For this device geometry, this represents a limit frequency, beyond which the stroke volume can not be fully exploited. Indeed, if a higher actuation frequency is imposed, there is not enough time to reach equilibrium configuration at the end of neither expansion nor compression stroke. Increasing the frequency f_p , however, would improve the pumping performances: although a reduced stroke is exhibited, a higher number of pumping cycles is performed for a given time interval.

A set of simulations for different geometries, for actuation frequencies ranging from 1 Hz to 50 Hz with square wave actuation, is performed. This yield an estimate of the flow rate $\mathcal{Q}(f_p)$ in this simplified framework (Figure 5). As a first consideration, the micropump behaviour is well reproduced by the simplified 2D model here proposed, from a qualitative standpoint. In fact, it exhibits the same frequency trend found by the experimental results presented by Zengerle *et al.* [9] and Machauf *et al.* [14].

For low frequencies, here roughly below 1-5 Hz with respect to the device geometry, the full stroke is exploited. In this low-frequency regime, increasing the frequency of actuation corresponds to a proportional increase of flow rates. A further increase of the actuation frequency leads to a decrease of the stroke ($\Delta v \leq v_s$). This is initially overcompensated by faster actuation. When an optimal frequency is reached, the maximum of pumping effectiveness is achieved. A this point, a further increase of the actuation frequency f_p leads to a decrease of the generated fluid flow.

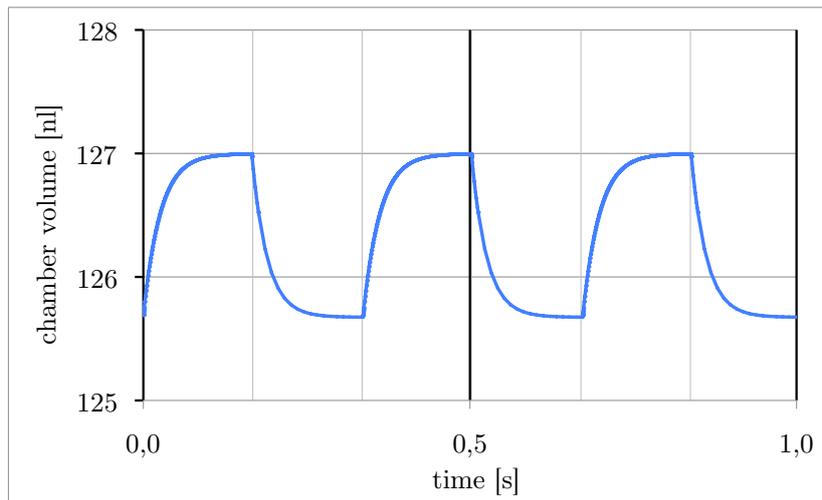
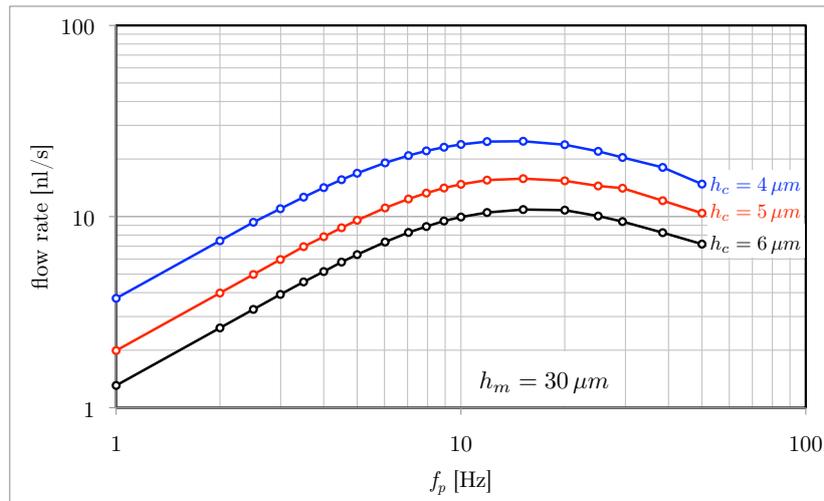
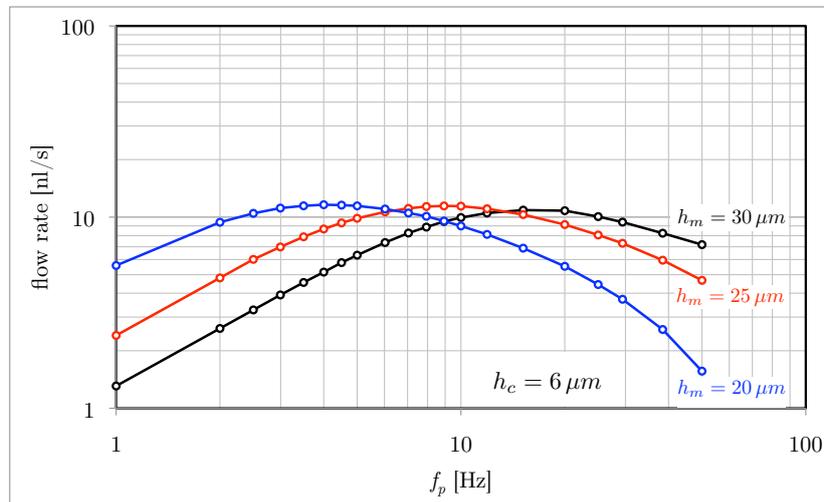


Figure 4: Chamber volume for square wave actuation at $f_p = 3$ Hz, for $h_m = 30$ μm and $h_c = 6$ μm .

From simulations, it appears that optimal pumping frequency is considerably influenced by membrane thickness, but not by capacitor spacing. For thinner diaphragms, a lower optimal actuation frequency is found. The obtained flow rate, for a given capacitor spacing, is the same in the investigated range of thickness. Although a higher stroke v_s is obtained with thinner diaphragms, this cannot be exploited at higher frequencies, due to pumped fluid inertia.



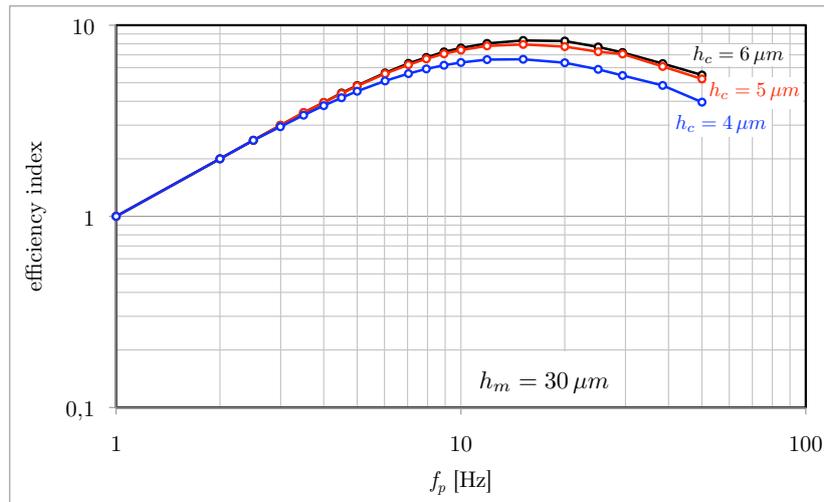
(a)



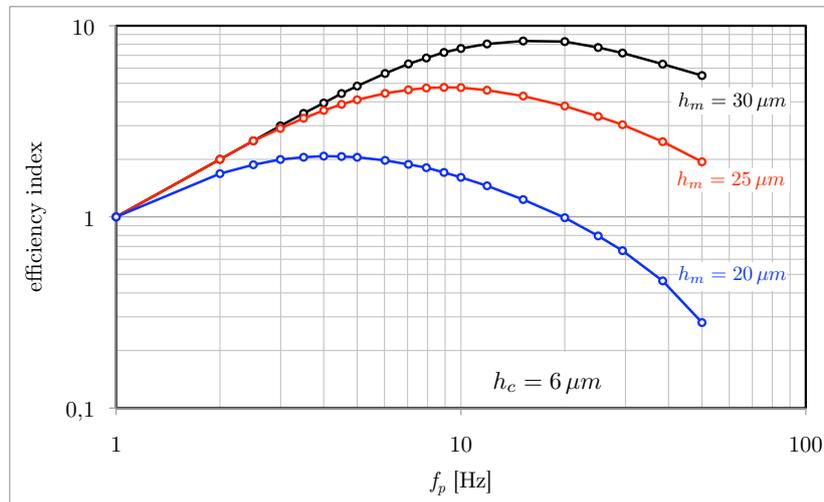
(b)

Figure 5: Flow rate $Q(f_p)$ calculated by means of the dynamic fully-coupled electro-fluid-mechanical model. The effect of capacitor spacing h_c (a) and membrane thickness h_m (b) is emphasized.

On the other hand, the fluid transport capability is greatly influenced by the capacitor spacing, not by membrane thickness. As expected, for smaller gaps a higher actuation pressure is exerted and more liquid can be displaced. Moreover, dynamic simulations evidenced that the higher stroke v_s achieved by decreasing capacitor spacing is better exploited, since the optimal actuation frequency is the same in the investigated range of capacitor spacings (for a given membrane thickness).



(a)



(b)

Figure 6: Efficiency index $\mathcal{E}(f_p)$ calculated by means of the dynamic fully-coupled electro-fluid-mechanical model. The effect of capacitor spacing h_c (a) and membrane thickness h_m (b) is emphasized.

Similar considerations can be drawn by considering the efficiency index $\mathcal{E}(f_p)$, plotted in Figure 6. The efficiency index, defined in Equation 6, is adopted here to quantify the advantage in improving the actuation frequency. It is clear that the micropump efficiency in the stable operation regime is greatly influenced by the membrane thickness h_m . Capacitor spacing has a reduced effect, that could become evident only for very small h_c . Considering dynamic effects and fluid inertia and referring to Equation 7, a flow rate Q of 1-2 $\mu\text{l}/\text{min}$ can be hypothesized for the device, considerably lower than Q_i computed with standard arguments from quasi-static simulations.

5 CONCLUSIONS

In this work, a fully-silicon mechanical displacement micropump is investigated. From a technological standpoint, the final goal of the research project is to take advantage of well-established MEMS technologies and realize cost-effective and reliable devices which can directly integrate with both electronic and microfluidic components.

An investigation by means of numerical methods is performed to provide an outlook on the device fluid transport capabilities. This preliminary work is intended to support the design process and the prototype realization, by providing suggestions on the development of the device, in terms of working principle, geometric parameters, actuation voltage. As expected, the results suggest the possibility to adopt this device design for a proper fluid actuation, in terms of actuation voltages and geometry. For the considered geometry, a flow rate of 1-2 $\mu\text{l}/\text{min}$ is expected for actuation frequencies from 10 to 30 Hz.

To take advantage of the full stroke, a maximum pumping frequency of 2-5 Hz can be hypothesized. A further increase of the actuation frequency leads to a decrease of the stroke ($\Delta v \leq v_s$), overcompensated by the faster actuation up to an optimal frequency, which corresponds to the maximum pumping effectiveness.

The optimal frequency identified for the current geometry is 10-30 Hz. A further increase of the actuation frequency f_p leads to a decrease of the generated fluid flow and a decay of performances. For a given actuation voltage in square wave, the optimal actuation frequency is greatly influenced by the membrane thickness h_m , but not by the capacitor spacing h_c . For thinner diaphragms, a lower optimal actuation frequency is found. By reducing the capacitor spacing h_c , higher flow rates can be obtained. The effect of membrane thickness h_m on the flow rate is negligible.

From a methodological point of view, it is worth to mention that the electro-fluid-mechanical dynamic simulations here presented return interesting results, although some rough simplifications are made to reduce the full problem to a 2D one. Indeed, to the knowledge of the Author, no detailed electro-fluid-mechanical fully-coupled models have been previously reported in the literature for electrostatic micropumps. The proposed simplified model is a step in this direction, and results are novel and encouraging.

To optimize the device and to improve pumping performances, the key component is the actuator, which deserves a particular attention. To increase the fluid transport capabilities dynamic drive could be exploited [17]. A viable alternative to increase the stroke volume avoiding pull-in is to apply charge control combined with tailored electrode shapes [18].

The device is here studied only in a stable range of actuation: a strong attention should be paid to pull-in phenomena, with the aim to control this feature and eventually to take advantage of it. The prototype developed by Zengerle and collaborators [9] works in the pull-in regime, developing a considerably higher stroke volume. In a similar circumstance, analyses should be performed to evaluate the adhesion (stiction) of the diaphragm and the its release after capacitor deactivation, also when the working fluid is present [19].

Acknowledgements

The Authors wish to acknowledge STMicroelectronics for the fruitful interaction in the development of this project.

References

- [1] Iverson, B. and Garimella, S., Recent advances in microscale pumping technologies: a review and evaluation. *Microfluid. Nanofluid.* (2008) **5**:145–174.
- [2] Woias, P., Micropumps – Past, progress and future prospects. *Sensor Actuat. B* (2005) **105**:28–38.
- [3] Laser, D. and Santiago, J., A review of micropumps. *J. Micromech. Microeng.* (2004) **14**:R35–R64.
- [4] Nisar, A., Afzulpurkar, N., Mahaisavariya, B. and Tuantranont, A., MEMS-based micropumps in drug delivery and biomedical applications. *Sensor Actuat. B* (2008) **130**:917–942.
- [5] Zhang, C., Xing, D. and Li, Y., Micropumps, microvalves, and micromixers within PCR microfluidic chips: advances and trends. *Biotechnol. Adv.* (2007) **25**:483–514.
- [6] Singhal, V., Garimella, S. and Raman, A., Microscale pumping technologies for microchannel cooling systems. *Appl. Mech. Rev.* (2004) **57**:191–221.
- [7] Bertarelli, E., Ardito, R., Cioffi, M., Laganà, K., Procopio, F., Baldo, L., Corigliano, A., Contro, R. and Dubini, G., Design optimization of an electrostatic micropump: a multi-physics computational approach. In: *Proceedings of the 2nd South-East Conference on Computational Mechanics*, 181 & CD-ROM (2009).
- [8] Oosterbroek, R., Schlautmann, S., Berenschot, J., Lammerink, T., van den Berg, A. and Elwenspoek, M., Modeling and validation of flow-structure interactions in

- passive micro valves. In: *Technical Proceedings of the 1998 International Conference on Modeling and Simulation of Microsystems*, 528–533 (1998).
- [9] Zengerle, R., Richter, A. and Sandmaier, H., A micro membrane pump with electrostatic actuation. In: *Proceedings of Micro Electro Mechanical Systems '92*, 19–24 (1992).
- [10] Kotzar, G., Freas, M., Abel, P., Fleischman, A., Roy, S., Zorman, C., Moran, J. and Melzak, J., Evaluation of MEMS materials of construction for implantable medical devices. *Biomaterials* (2002) **23**:2737–2750.
- [11] Bertarelli, E., Ardito, R., Bianchi, E., Laganà, K., Procopio, F., Baldo, L., Corigliano, A., Dubini, G. and Contro, R., A computational study for design optimization of an electrostatic micropump in stable and pull-in regime. *AES Tech. Rev. B: IJAMAIM* (2010) **1**:19–15.
- [12] Cao, L., Mantell, S. and Polla, D., Design and simulation of an implantable medical drug delivery system using microelectromechanical systems technology. *Sensor Actuat. A* (2001) **94**:117–125.
- [13] Teymoori, M. and Abbaspour-Sani, E., Design and simulation of a novel electrostatic peristaltic micromachined pump for drug delivery applications. *Sensor Actuat. A* (2005) **117**:222–229.
- [14] Machauf, A., Nemirovsky, Y. and Dinnar, U., A membrane micropump electrostatically actuated across the working fluid. *J. Micromech. Microeng.* (2005) **15**:2309–2316.
- [15] Voigt, P., Schrag, G. and Wachutka, G., Electrofluidic full-system modelling of a flap valve micropump based on Kirchhoffian network theory. *Sensor Actuat. A* (1998) **66**:9–14.
- [16] Voigt, P., Schrag, G. and Wachutka, G., Microfluidic system modeling using VHDL-AMS and circuit simulation. *Microelectr. J.* (1998) **29**:791–797.
- [17] Bertarelli, E., Ardito, R., Corigliano, A. and Contro, R., A plate model for the evaluation of pull-in instability occurrence in electrostatic micropump diaphragms. *Int. J. Appl. Mech.* (2011) **3**:1–19.
- [18] Bertarelli, E., Corigliano, A., Greiner A. and Korvink, J.G., Design of high stroke electrostatic micropumps: a charge control approach with ring electrodes. *Microsyst. Technol.* (2011) **17**:165–173 .
- [19] Ardito, R., Bertarelli, E., Contro, R. and Corigliano, A., Static and dynamic analyses of actuation devices in electrostatic micro-pumps. In: *Proceedings of the 7th International Conference on Engineering Computational Technology*, paper 121 (2010).

MULTIPHYSICAL COMPUTATIONS OF THE ELECTRICAL MACHINES USING FEM

A. STERMECKI ^{*†}, O. BÍRÓ ^{*†}, M. HETTEGGER ^{*†}, H. LANG [♦], G. OFNER [♦], S. RAINER [†] AND B. WEILHARTER ^{•†}

^{*}Institute for Fundamentals and Theory in Electrical Engineering (IGTE)
Graz University of Technology, Kopernikusgasse 24/3, 8010 Graz, Austria
e-mail: andrej.stermecki@tugraz.at, <http://www.igte.tugraz.at>

[•]Institute for Electrical Drives and Machines
Graz University of Technology, Inffeldgasse 18, 8010 Graz, Austria
e-mail: andrej.stermecki@tugraz.at, <http://www.igte.tugraz.at>

[†] Christian Doppler Laboratory for Multiphysical Simulation, Analysis and Design of Electrical Machines (MuSicEl)
Kopernikusgasse 24/3, 8010 Graz, Austria

[♦] ELIN Motoren GmbH
8160 Preding/Weiz, Austria
e-mail: georg.ofner@elinmotoren.at, web page: <http://www.elinmotoren.at>

Key words: Acoustic noise, Electric machines, Electromagnetic forces, Multiphysics problems, Thermal analysis, Vibrations.

Abstract. In this paper three problems representing the multiphysical aspect of electrical machine computation are addressed. The interaction between magnetic and structural mechanical systems is demonstrated by the finite element method (FEM) structural investigation of the turbo-generator end-winding deformations. Multiphysical simulation of the acoustical problem is presented by weakly coupled electromagnetic, structural dynamic and acoustic simulations. And finally, a procedure based on computational fluid dynamics (CFD) simulation for acquiring the convective heat transfer coefficients is proposed in order to improve the accuracy of the coupled electro-thermal FEM simulations.

1 INTRODUCTION

In order to comply with modern demands for better efficiency and more environment-friendly machine-drives, they should be considered as comprehensive coupled systems, where effects of different physics are interacting with each other. The numerical computation of these phenomena is rather pretentious and often requires a substantial enhancement of the state of the art numerical procedures and highly efficient computing powers. In 2007 a new laboratory titled Christian Doppler Laboratory for Multiphysical Simulation, Analysis and Design of Electrical Machines (MuSicEl) has been established in the scope of the Institute for Fundamentals and Theory in Electrical Engineering (IGTE) and the Institute for Electrical

Drives and Machines of the Graz University of Technology. The research work of the laboratory is focused particularly on the development of efficient computational methods required for the consideration of the multiphysical nature of the electrical machines. In this paper three such representative electrical machine multiphysical problems are addressed in order to present the research activities of the Christian Doppler Laboratory for Multiphysical Simulation, Analysis and Design of Electrical Machines. The descriptions of the presented computational methods are based on the work published by the authors in [1], [2] and [3].

The investigation of the turbo-generator end-winding deformations shall demonstrate the interaction between magnetic and structural mechanical systems. The numerical approach based on the sequential coupling between a current-flow, an electromagnetic and a mechanical model is presented. In order to take the excitation produced by stator windings into consideration, a procedure based on coupled current flow and magnetic field analyses using dissimilar finite element (FE) meshes and numerical calculation of the Biot-Savart field has been employed. The magnetic forces acting on current conducting elements have been calculated and have been applied to the structural model. Under design operating conditions, computed harmonic forces specified by amplitude and phase have been applied to the mechanical model in order to simulate the frequency response of the end-winding system.

For the numerical computation of audible noise generated by induction machines, weakly coupled electromagnetic, structural dynamic and acoustic simulations have been accomplished [2], [4]. In order to comply with the different demands of each physical domain, different numerical formulations have been used for the simulation models besides applying different space discretizations. Electromagnetic excitation forces have been estimated by solving a transient nonlinear eddy current problem formulated by FE methods. The resulting forces have been applied to a linear structural dynamic FE model in the frequency domain. Finally, the acoustic sound pressure level is determined from the structural surface vibrations using the boundary element method.

Knowing the thermal boundary conditions is crucial in order to carry out coupled electrical-thermal analyses. The accuracy of the whole thermal simulation depends on the precision of those boundary conditions. Especially for complex geometries like the end-windings, there is no exact analytical method for the calculation of the convective heat transfer coefficient. A procedure based on computational fluid dynamics (CFD) simulation for acquiring the convective heat transfer coefficients is proposed.

2 ANALYSIS OF SYNCHRONOUS GENERATOR END-WINDING DEFORMATIONS

The structural investigation of synchronous machine end-region phenomena based on a numerical approach is an especially ambitious task, since the coupling of two models, each fitting different physics, needs to be accomplished. The numerical approach based on the sequential coupling between a current-flow, an electromagnetic and a mechanical model is presented in this paper. An electromagnetic simulation has been carried out in the frequency domain using a T, Φ - Φ formulation [5]. The material nonlinearity has been taken into account by the employment of the effective B - H curve [6]. The magnetic forces acting on current conducting elements have been calculated in the frequency domain and their time-varying

components have been applied to the structural model. A different FE mesh discretization of the mechanical model has been required in order to take into account the exact structural characteristics of the winding system and the construction parts.

2.1 Stator current excitation

Owing to the employed \mathbf{T}, Φ - Φ formulation [5], the current density in the stator three-phase winding (\mathbf{J}) needs to be represented by an arbitrary current vector potential \mathbf{T}_0 satisfying the following condition:

$$\text{curl}(\mathbf{T}_0) = \mathbf{J}. \quad (1)$$

One possible way to find an adequate function \mathbf{T}_0 is to calculate the Biot-Savart field produced by the stator current excitation. Static current flow analysis has been carried out in order to obtain the current distribution in the stator winding. Knowing the current distribution, the Biot-Savart field has been numerically computed in the integration points of another FE mesh representing the motor, its housing and the surrounding air. As seen in Fig. 1, the meshes of the stator winding and of the motor are non-conforming.

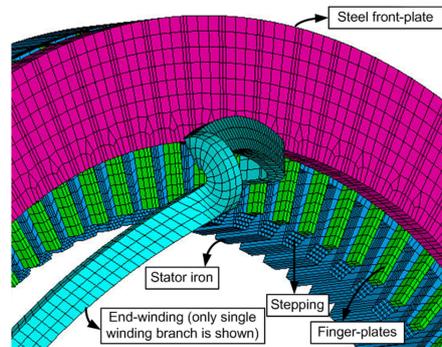


Figure 1. 3-D FE model of the motor and its end-winding.

The Biot-Savart field \mathbf{H}_s has been evaluated for each phase excitation separately and by summing up the three magnetic field contributions, the complex representation of the rotating magnetic field intensity is obtained [1]. Knowing the numerically calculated Biot-Savart field, the rotational part \mathbf{T}_0 of the magnetic field is represented with the aid of edge element shape functions in order to avoid cancellation errors:

$$\underline{t}_i = \int_{\text{Edge } i} \mathbf{H}_s dl; \quad \mathbf{T}_0 = \sum_{\text{Edge } i}^{n_e} \underline{t}_i N_i \quad (2)$$

where N_i stands for the vectorial edge element shape functions associated with the i -th edge and n_e for number of FE edges.

2.2 Rotor current excitation

The rotating magnetic field of the rotor is not generated by a 3-phase AC current excitation, but rather by DC currents and the mechanical rotation of the rotor making it impossible to be modeled in the same way as the stator excitation. In order to overcome this problem, a novel numerical method has been developed [1]. Using the \mathbf{T}, Φ - Φ formulation, the

rotor excitation has been defined by prescribing the rotating part \underline{T}_0 of the magnetic field in a thin finite element layer on the surface of the rotor [1]. It is sufficient to define the radial component of \underline{T}_0 only (Fig 3b), since this way the current density \underline{J} in axial and azimuthal directions is taken into consideration:

$$\text{curl}\underline{T}_0 = \frac{\partial T_{0,r}}{\partial z} \underline{e}_\phi - \frac{1}{r} \frac{\partial T_{0,r}}{\partial \phi} \underline{e}_z = \underline{J} \quad (3)$$

where $T_{0,r}$ denotes the radial component of the prescribed vector potential \underline{T}_0 , \underline{J} the current density in the rotor excitation winding and \underline{e}_ϕ and \underline{e}_z the unit vectors in z and ϕ directions in a polar coordinate system. This way, the current excitations in axial and azimuthal directions are modeled in correspondence with the actual rotor winding distribution. In order to define the adequate amplitude of the prescribed potential \underline{T}_0 (3), the well known analytical expressions for the rotor current excitation have been taken into consideration. The detailed description of this procedure can be find in [1]. Knowing the distribution of the magnetic field, the Lorentz forces acting on the stator end-winding have been calculated. Since the computation has been carried out in the frequency domain, the constant as well as the oscillating parts of the forces have been evaluated. The latter forces oscillating with double nominal frequency (100 Hz) are especially important for the structural investigations of the nominal operating conditions. Therefore, these forces have been used as an input for the structural mechanical analysis carried out in the next step.

2.3 Structural mechanical analysis

Another FE mesh has been used for the structural harmonic analysis. A detailed mechanical structure consisting of winding bars, poromat bricks, fabric tubes, cantilevers, finger-plates and the stator package has been modeled. Since non-conforming meshes have been used for all parts, contact elements have been employed to link the different parts. A separate pre-processor program has been developed in order to parametrically define the model's geometry and to obtain a finite element mesh composed of well shaped hexahedron elements. Finally, the investigation of the structural behavior has been carried out in the program package ANSYS [7].

The final goal of the presented mechanical investigations has been to determine deformations of the end-winding system caused by the calculated excitation. The extent of the resulting deformations can be attributed to individual deformation waves of different ordinal numbers. In order to avoid mechanical resonances, the lowest natural frequency of the end-winding system has to be located distant enough from the excitation frequency. It is safe to assume that, for the machine under research, all natural frequencies above 140 Hz are not going to be highly stimulated by the excitation frequency of 100 Hz. The natural frequencies have been ascertained by performing a modal analysis. Moreover, a harmonic analysis has been carried out as well in order to compute the exact deformations due to the force excitation.

2.4 Results

In the present work, a crucial impact of the boundary conditions on the calculated winding forces has been ascertained. In order to compute the magnetic field as accurately as possible,

the problem has been solved in the frequency domain and eddy currents induced in the housing, the bearing shield, the clamping-fingers and the clamping-plate have been taken into consideration. In Fig. 2a the eddy currents in the portion of the steel clamping-plate and the front clamping-fingers are shown.

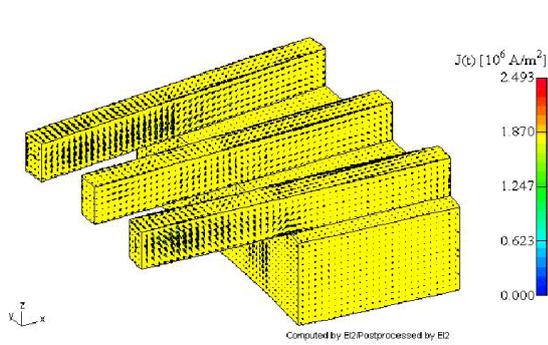


Figure. 2a. Induced eddy currents in the clamping -plate and -fingers.

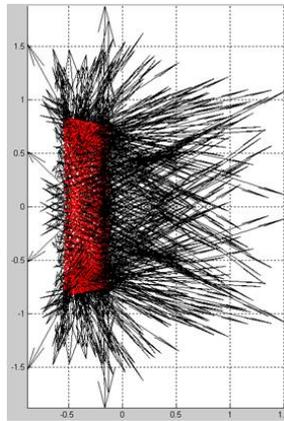


Figure. 2b. Side view of the forces acting on stator end-windings.

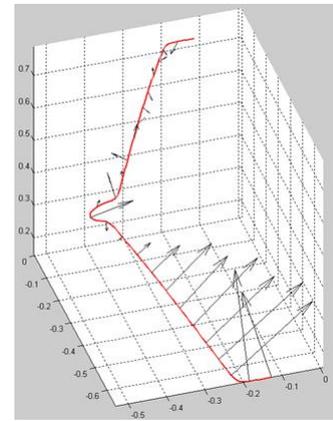


Figure. 2c. Forces acting on a single stator winding branch.

In Fig. 2b, the oscillating parts of the forces acting on the stator end-winding are shown. As seen from the Fig. 2b, the forces are symmetrically distributed over 72 stator slots and maximal force densities are obtained at the positions near the rotor package (Fig. 2c).

The results of the harmonic analysis are given in Fig. 3. Outcomes of two methods are presented and mutually compared. The deformations shown in Fig. 3a have been obtained by full stiffness and mass matrices. The modal superposition method taking into consideration the first twenty modes only has been used to obtain the deformations shown in Fig. 3b. Based on the presented results (Fig. 3), the conclusion can be drawn that all essential deformation characteristics are kept even when a less complex and less time-consuming modal superposition method is employed.

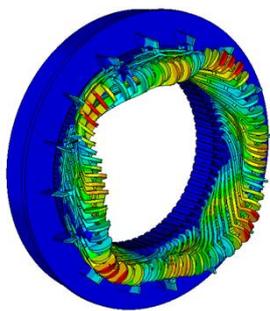


Figure. 3a. End-winding deformations (harmonic analysis using full mass and stiffness matrices).

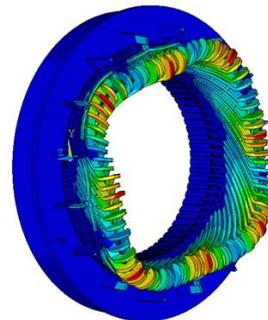


Figure. 3b. End-winding deformations (harmonic analysis using modal superposition of the first 20 modes).

The proposed method is very time-efficient, since a transient time-stepping analysis has

been avoided, which would otherwise be required to consider the rotor excitation. It has to be mentioned that the rotor slots harmonics have been neglected in this procedure, since only the fundamental spatial harmonic of the prescribed vector potential to the rotor has been taken into consideration. Nevertheless, a further improvement of this method is possible by imposing the excitation fields of higher-order harmonics to the FE layer on the rotor surface.

The proposed procedure has been employed to investigate the structural behavior of a synchronous generator. The findings of this study lead to a conclusion that, for the machine under research, the natural frequencies of the end-winding are not interfering with the excitation frequency. Consequently, the obtained deformations by harmonic analysis are also within the expected μm range and no negative influences on the end-winding life-cycle are to be expected. These simulation results have also been confirmed by the measurement results.

3 NUMERICAL COMPUTATION OF THE NOISE RADIATION OF AN INDUCTION MACHINE

The reasons for audible noise generated by electrical machines are, besides mechanical causes, as for example bearings, brush noise or rotor unbalance, and noise of aerodynamic origin, certain aspects in the electromagnetic field generation [8], [9]. Due to the permeance fluctuations, e.g. arising from the slotting or saturation effects, and the winding composition, the magnetic flux distribution in the airgap contains, apart from the fundamental wave, time-harmonics of higher order. Due to the electromagnetic field in the airgap, forces with according harmonics acting on the rotor and stator surface occur. Obviously, the electromagnetic forces acting on the stator lead to its deformation and thus also the deformation of the housing. The resulting vibrations result in pressure fluctuations of the enclosing air and thus audible noise is generated. Therefore, for computations of audible noise with electromagnetic origin, three different physics, namely electromagnetics, mechanics and acoustics, have to be considered and coupled accordingly. A comprehensive numerical approach shall be presented in this paper. A detailed version of this approach can be found in [2].

3.1 Electromagnetic Field Computation

The electromagnetic field in the airgap, with the magnetic flux density B , results in the mechanical stress σ , which can be computed in time domain according to [10] with the following relation

$$\sigma_n = \frac{1}{2\mu_0} (B_n^2 - B_t^2) \quad (4)$$

$$\sigma_t = \frac{1}{\mu_0} B_n B_t, \quad (5)$$

where μ_0 denotes the magnetic permeability for vacuum and the indices n and t the normal and tangential component of the magnetic flux density B . The determination of B is carried out numerically with a finite element multi-slice simulation [11], which also enables the consideration of skewing effects. The axial dimension is divided into five slices which are represented by two dimensional finite element models. These models, more precisely the

currents in the rotor bars, are then connected via an electric circuit. The $A, v - A$ formulation [12] is used to solve the nonlinear and transient eddy current problem. The solution yields results for the magnetic flux density B in the airgap at distinct axial positions. For the successive structural computation the stress distribution in the airgap along the axial dimension is needed. Hence the axial characteristic of B has to be reconstructed from the multi-slice simulation results. Therefore the time-harmonics and spatial harmonics of the magnetic flux density B , implying a 2D Fourier transformation, are needed. The spatial order in axial direction is computed analytically and a correlation with the harmonics of the multi-slice solution enables the reconstruction, see [13]. The mechanical stress acting on the teeth of the stator can now be computed according to (4) and (5) in time domain.

3.2 Structural Vibration Computation

To determine the structural vibrations the finite element method is used. A three-dimensional model of the investigated machine, comprising the main parts stator, rotor, windings and housing, has to be set up and discretized appropriately. As only the steady state solution is of interest, the simulation can be performed in frequency domain. According to [14], the dynamical behaviour of the structure for the excitation frequency Ω is therefore described by

$$(-\Omega^2 \mathbf{M} + \mathbf{K}) \hat{\mathbf{x}} = \hat{\mathbf{F}} \quad (6)$$

with the mass matrix \mathbf{M} and the stiffness matrix \mathbf{K} . $\hat{\mathbf{F}}$ and $\hat{\mathbf{x}}$ denote the complex nodal force amplitude and the complex deformation amplitude for the excitation frequency Ω . It is assumed that the deformation does not influence the magnetic flux distribution, which results in a weakly coupled problem. The mechanical stress σ has to be transformed into frequency domain, or directly computed in this domain. The nodal forces at the excitation frequencies can then be determined and impressed onto the structural model. Solving (6) then provides the deformation characteristic on the surface of the machine, which is of interest for the noise computation.

3.3 Acoustic Computation

To determine, in a last step, the noise radiation due to the machine vibrations and surface oscillations the homogeneous Helmholtz equation

$$\Delta p + k^2 p = 0 \quad (7)$$

with the wave number k has to be solved, regarding the sound pressure p , for an exterior radiation problem. Therefore the Sommerfeld radiation condition

$$\lim_{R \rightarrow \infty} \left[R \frac{\partial p}{\partial R} - jkp \right] = 0 \quad (8)$$

with R denoting the radial distance, has to be satisfied, meaning that any acoustic disturbance caused by the structure dies out at infinity.

To solve (7) the boundary element method [15], [16] is used. Only the surface of the investigated structure has to be discretized and furthermore (8) is fulfilled implicitly when using this method, which are great advantages compared to the finite element method. From the structural frequency domain solution for the deformation $\hat{\mathbf{x}}$ the velocity of the surface vibration can be computed. Again the influence of the sound pressure on the structural deformation is neglected and thus weak coupling is assumed. The normal components of the

velocity are applied to the surface elements as a boundary condition for the boundary element problem. The solution of the Neumann problem of (7) is the sound pressure on the surface of the investigated machine. To get the sound pressure distribution in the circumference of the investigated structure a field, where the boundary element solution is evaluated, discretized with appropriate elements, has to be set up.

3.4 Numerical Simulation Results

The numerical simulation has been carried out for an induction machine with a cylindrical cooling jacket. The rotor is skewed for one stator slot pitch and has a speed of 2991 rpm at nominal operating point. Acoustic measurement showed a noise peak at 1895.6 Hz of approximately 78 dB. The solution of the transient electromagnetic simulation shows a time harmonic of the surface stress occurring at this frequency. In Fig. 4a the computed real part of the radial component of the mechanical stress acting on the stator teeth is depicted.

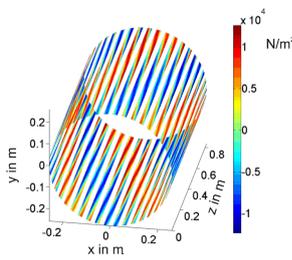


Figure. 4a. Teeth forces in radial direction

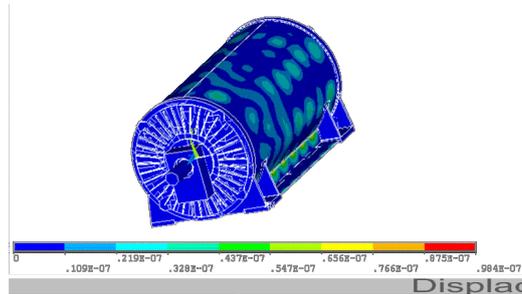


Figure. 4b. Real part displacement solution

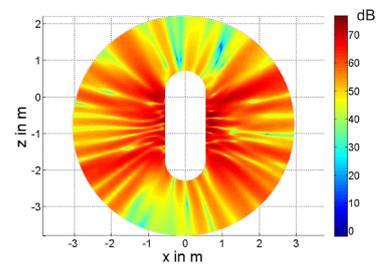


Figure. 4c. Sound pressure level on a plane field surrounding the machine

This resulting stress distribution is impressed on the mechanical finite element model and the structural deformation is computed according to (6). The solution is shown in Fig. 4b. It can be seen that the deformation characteristic is dominated by plate bending, which can be lead back to the eigenmodes near the excitation frequency. This plate bending causes the pressure fluctuations in the air and thus the radiating noise. The sound pressure level in a horizontal plane field, passing through the rotation axis of the rotor, is depicted in Fig. 4c. The main part of the noise is radiating from the cylindrical surfaces, as there the highest deformation amplitudes occur. The sound pressure level of more than 70 dB near the machine's surface is decreasing with the distance. At a surrounding sphere maximal sound pressure levels of approximately 70 dB can be detected. This however deviates from measurement results which yield approximately 78 dB at the excitation frequency and in 1 m distance for this induction machine. The deviations can be explained by the lack of validated material parameters for the structural model and by the possible influence of the forces acting on the rotor core stack.

4 DETERMINATION OF CONVECTIVE HEAT TRANSFER COEFFICIENT AT END WINDINGS FOR TEMPERATURE FIELD SIMULATIONS

The transport of thermal energy can be split into three different physical phenomena heat radiation, heat conductance and heat convection. All three of these phenomena can occur at the same time and can be cumulated into the total heat flux \dot{q}_{tot} .

Under common operating conditions of electrical machines the heat radiation transports a negligible portion of thermal energy. Due to its dependency on the fourth power of the temperature difference its influence on the total energy balance is rarely perceptible. For a simple test case of a radiating body in a free environment the heat radiation can be written as

$$\dot{q}_{rad} = \varepsilon\sigma \cdot (T^4 - T_\infty^4), \quad (9)$$

where ε is the emissivity, σ the Stefan-Boltzmann constant and T_∞ denotes the ambient temperature. Nevertheless it should not be discarded completely from the design process of electrical machines. In certain situations the heat radiation can be an observable fact, for instance for machine types, which allow higher temperatures at the housings surface. In this case the magnitude of heat flux by radiation can reach the value of heat flux by free convection, and must be considered [17].

The procedure of conductive heat flow simulation in solid materials is well known and reliable. The distribution of temperature inside the material can be solved for instance with finite element method (FEM) by solving the Laplace equation

$$\dot{q}_{cond} = -\lambda \cdot \nabla T \quad (10)$$

with the conductive heat flux \dot{q}_{cond} . A homogenisation of the material properties [17] can simplify the computation model [3] for components with non isotropic thermal conductivity λ like laminated iron sheets and insulated copper windings. At contact regions of different solid materials or a thin layers of a different substances a special treatment of the FEM model is necessary, in order to avoid a very fine resolution of the mesh. In these cases the application of certain element types can include these effects as proposed in [18]. Its accuracy satisfies most requests of engineering purposes if proper boundary conditions have been specified.

The boundary to the adjoining fluid coolant is specified by the convective heat flux

$$\dot{q}_{conv} = \alpha \cdot (T - T_\infty) \quad (11)$$

with the convective heat transfer coefficient α . This convective heat transfer coefficient depends basically on the fluid properties of the coolant and its flow conditions like velocity and turbulence. In general α can be defined for certain areas or locally α_x for certain points at the surface. If the heat transfer coefficient is defined locally one has to take under consideration the validity of its value. Due to its definition with the environmental temperature T_∞ it can be physically invalid if the difference between surface temperature and environmental temperature becomes zero but the convective heat flux \dot{q}_{conv} does not [19].

The convective heat transfer coefficient can be derived analytically only for very simple geometries in the case of laminar flow conditions. In the case of turbulent flow conditions at complex geometries with rough surfaces only empirical approaches could be used in the past. Due to the improvements of computational fluid dynamics (CFD) and the nowadays available

computational power, it is possible to derive the convective heat transfer coefficients of any geometry numerically. An exact derivation of the convective heat transfer can be done with direct numerical simulation (DNS) but it requires a mesh resolution based on the Kolmogorov length scale η

$$\eta = \left(\frac{\nu^3}{\varepsilon} \right)^{1/4}, \quad (12)$$

which is defined by the kinematic viscosity ν and the turbulence dissipation rate ε . By using this length scale the element size of the mesh has to be in the magnitude of several micrometers for an electric machine, which uses air as cooling fluid. Hence this dramatically enlarges the number of elements and degrees of freedom, thus DNS is not a suitable method for industrial applications. Models which are based on eddy viscosity are more appropriate for engineering purposes and should be used instead DNS. Hence these numerical simulation models are based on certain simplifications, they can not be used for all kind of flow conditions in the same way and can fail in calculating the heat transfer. Therefore it is still necessary to validate the derived results by measurements or other empirical approaches. Especially at low velocities and thus low Reynolds numbers certain turbulence models fail in predicting the heat flux at the surface. The occurrence of flow separation and reattachment of the fluid cannot always be predicted accurately but it makes a crucial difference in the derivation of the heat transfer.

According to the investigations in [20] and [21] the Shear Stress Transport (SST) turbulence model [22] and the Scale Adaptive Simulation (SAS) turbulence model [23] meets the requirements for the prediction of the heat transfer coefficient best.

In Fig. 5 the derived heat transfer coefficient is shown related to its maximum.

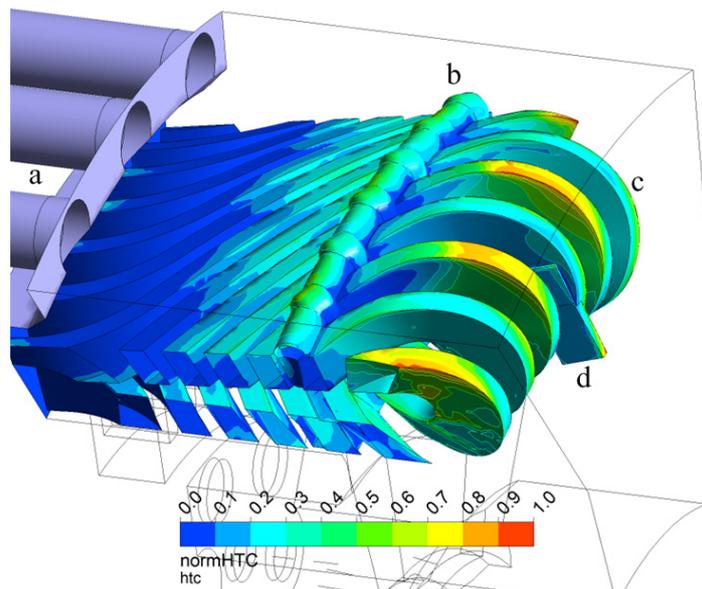


Figure 5: Norm values of the local heat transfer coefficient at the surface of the end-windings of an induction machine, with the cooling ducts (a), cord (b), end windings (c) and a heat flux sensor (d) which has been meshed with the end-windings' geometry in order to validate the simulation results.

As one can see, the heat flux reaches the maximal value at the outermost position of the end-windings, where the conductors are exposed directly to the air flow from the cooling ducts. The results shown in Fig. 5 have been derived by the use of the SAS turbulence model, using constant values for the temperature at the models' surfaces and temperature dependent fluid properties as used in [24].

These derived values for the local heat transfer coefficient can be used for the simulation of the temperature distribution inside the solid material of the electric machine. This should lead to more accurate results than using mean values for the convective heat transfer coefficient.

ACKNOWLEDGMENT

This work has been supported by the Christian Doppler Research Association (CDG), by the ELIN Motoren GmbH and Traktionssysteme Austria GmbH.

REFERENCES

- [1] A. Stermecki, O. Bíró, G. Ofner, and H. Lang, "Numerical Simulation of Electromagnetic Phenomena in Motor End-Regions", e & i Elektrotechnik und Informationstechnik, submitted for publication.
- [2] BG. Weilharter, O. Bíró, H. Lang and S. Rainer, "Computation of the Noise Radiation of an Induction Machine Using 3D FEM/BEM", COMPEL, submitted for publication.
- [3] M. Hettegger, O. Bíró, A. Stermecki, G. Ofner, "Temperature rise determination of an induction motor under blocked rotor conditions", PEMD 2010, Brighton 2010.
- [4] B. Weilharter, O. Bíró, H. Lang, G. Ofner and S. Rainer, "Validation of a Comprehensive Analytic Approach to Determine the Noise Behaviour of Induction Machines", ICEM 2010, Rome 2010.
- [5] O. Bíró, K. Preis, G. Vrisk, K. R. Richter and I. Tícar, "Computation of 3-D Magnetostatic Fields Using a Reduced Scalar Potential," IEEE Transaction on Magnetics, 29(2): 1329-1332, 1993.
- [6] G. Paoli, O. Biro, and G. Buchgraber, "Complex Representation in Nonlinear Time Harmonic Eddy Current Problems," IEEE Transaction on Magnetics, vol.34, no.5, pp.2625-2628, September 1998.
- [7] ANSYS 11.0 documentation (2007). SAS IP, Canonsburg.
- [8] P. L. Timar, *Noise and Vibration of Electrical Machines*, Elsevier, 1989.
- [9] H. Jordan, *Der geräuscharme Elektromotor*, W. Girardet, 1950.
- [10] J. Melcher, *Continuum Electromechanics*, MIT Press, Cambridge, MA, 1981.
- [11] P. Dziwniel, B. Boualem, F. Piriou, J.P. Ducreux and P. Thomas, "Comparison Between Two Approaches to Model Induction Machines with Skewed Slots", IEEE Transactions on Magnetics, vol. 36, no. 4, pp. 1453-1457, 2000.
- [12] O. Bíró, "Edge element formulation of eddy current problems", Comput. Methods Appl. Mesh. Engrg., vol 169, no. 3, pp 391-405, 1999.
- [13] B. Weilharter, O. Bíró and S. Rainer, "Computation of rotating force waves in induction machines using multi-slice models", IEEE Transactions on Magnetics, submitted for publication.

- [14] K. J. Bathe, *Finite Element Methods*, Springer.
- [15] R.D. Ciskowski and C. A. Brebbia, *Boundary Element Methods in Acoustics*, Computational Mechanics Publications, Elsevier Applied Science.
- [16] Y. J. Liu and F. J. Rizzo, "*Hypersingular boundary integral equations for radiation and scattering of elastic waves in three dimensions*", Computer Methods in Applied Mechanics and Engineering, vol. 107, pp. 131-144, 1993.
- [17] Richter, R. *Elektrische Maschinen* Springer, 1967, XVI, 691 Seiten.
- [18] Driesen, J.; Belmans, R. & Hameyer, K. *Finite-element modeling of thermal contact resistances and insulation layers in electrical machines Industry Applications*, IEEE Transactions on DOI - 10.1109/28.903121, Industry Applications, IEEE Transactions on, 2001, 37, 15-20.
- [19] Herwig, H. *Kritische Anmerkungen zu einem weitverbreiteten Konzept: der Wärmeübergangskoeffizient α* Forschung im Ingenieurwesen, Springer Berlin / Heidelberg, 1997, 63, 13-17.
- [20] Hettegger, M.; Streibl, B.; Biro, O. & Neudorfer, H. "*Identifying the heat transfer coefficients on the end-windings of an electrical machine by measurements and simulations*", Electrical Machines (ICEM), 2010 XIX International Conference on DOI - 10.1109/ICELMACH.2010.5608250, 2010, 1-5.
- [21] Hettegger, M.; Streibl, B.; Biro, O. & Neudorfer, H. "*Measurements and Simulations of the Heat Transfer on End Windings of an Induction Machine*", 14th International IGTE Symposium, 2010, 23 - 23.
- [22] Menter, F. R. "*Two-Equation Eddy-Viscosity Turbulence Models for Engineering Applications*", AIAA Journal, 1994, 32, 1598-1605.
- [23] Egorov, Y. & Menter, F. "*Development and Application of SST-SAS Turbulence model in the DESIDER Project*", CFX Technical Memorandum, 2007, CFX-VAL10/0602.
- [24] Hettegger, M.; Streibl, B.; Biro, O. & Neudorfer, H. "*Characterizing the heat transfer on the end-windings of an electrical machine for transient simulations*", MELECON 2010 - 2010 15th IEEE Mediterranean Electrotechnical Conference DOI - 10.1109/MELCON.2010.5476025, 2010, 581-586.

THE DIFFERENT LEVELS OF MAGNETO-MECHANICAL COUPLING IN ENERGY CONVERSION MACHINES AND DEVICES

A. BELAHCEN, R. KOUHIA AND K. FONTEYN

Aalto University
POBox 13000, FI-00076 Aalto, Finland
E-mail: anouar.belahcen@aalto.fi, reijo.kouhia@aalto.fi, katarzyna.fonteyn@tkk.fi
Web page: <http://sahkotekniikka.tkk.fi/en/>, <http://buildtech.tkk.fi/en/>

Key words: Electrical machines, Coupled Problems, Magnetic forces, Magnetostriction, Magneto-mechanics, Computing Methods.

Abstract. This paper reviews the methods for coupling the magnetic and mechanical problems in magnetic materials and their application to electrical machines. The reviewed methods include both the material models and the computing methods as well as the methods for computing the magnetic forces. The paper shows that there are different levels of coupling the magnetic system with the mechanical one and that the use of a method or another depends on the application and the level of accuracy aimed at. The paper also clarifies some terms and concepts related to the coupling terminology such as strong, weak, local, global, direct and indirect coupling and put these terms in a coherent context. Most of the examples are related to the two dimensional analysis but some three dimensional ones are also shown.

1 INTRODUCTION

Energy conversion devices refer to these devices that convert mechanical energy into electrical one or vice-versa by the media of a magnetic field. Such device could be an electric motor or generator but it could be also an actuator or in some cases even a sensor although there is very little energy conversion in this latter case.

In the case of a motor or generator, the energy conversion occurs so that the current flowing in the windings of the machine, connected to a voltage source, generates, in accordance with the Ampere law, a magnetic field in the core and the air gap of the machine that exerts magnetic forces and torque on different parts of the machine and thus produces motion of the rotor or any other moving part. Thus the electrical energy in the form of voltages and currents at the terminal of the machines is converted into mechanical energy in the form of motion and torque at its shaft or vice-versa. In the case of actuators the electrical energy is converted into mechanical one either through a rigid motion of a plunger or through an elastic deformation of the actuating part under the effect of the magnetic field e.g. magnetostrictive deformation. Fig. 1 shows a cross-section of an electrical motor and illustrates the power flow from the electric supply to the mechanical load. Fig. 2 schematically shows a simple electromagnet and illustrates the related power flow. Fig. 3 shows a schematic view of a magnetostrictive actuator and illustrates its operation and power flow. The energy conversion is always accompanied with power losses in different parts of the devices and only part of the input power is converted to the output.

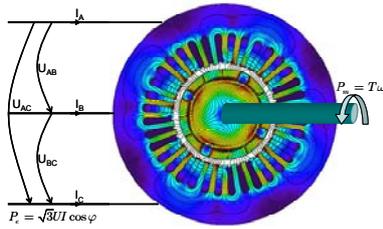


Figure 1: Illustration of the power flow in an electric motor. The electric power is converted into mechanical power through the magnetic field and the torque it produces. Part of this power is dissipated in the device.

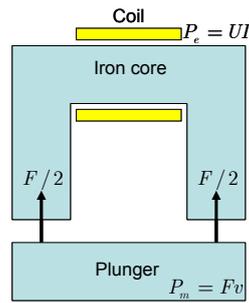


Figure 2: A schematic view of an electromagnet. The input electric power, which depends on the position of the plunger, is converted into mechanical power in the form of motion and force on this later one.

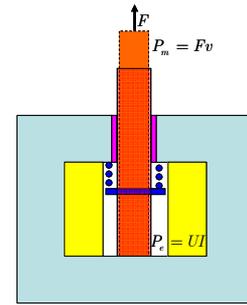


Figure 3: Schematic representation of a magnetostrictive actuator. The magnetic field in the actuating part causes its deformation, which is transmitted as mechanical output to the surrounding

Besides their operation as energy converters, the above devices are prone to parasitic phenomena such as vibrations and acoustic noise as well as wearing. The vibrations and noise occur due to the magnetic forces and the magnetostriction. Indeed, the time dependent magnetic field in the core and the air gap of the machine produce time dependent magnetic forces and induce magnetostrictive strains under the effect of which the structure of the device starts vibrating and emits acoustic noise. The acoustic noise is also produced by fluid flow in the machine and its surrounding as well as by pure mechanical effects such as contacts between parts or rolling of the bearing balls. In some cases, the acoustic noise is not generated in the structure of the machine but rather in some surrounding of it as the vibrations of the machine are transmitted to the surroundings.

On the other hand, the operation of the energy conversion device depends on the magnetic properties of the underlying material of its core, windings and permanent magnets. These properties such as the magnetization characteristics of the permanent magnets and the core material or the resistivity of the winding material or even the magnetostriction of the actuating parts strongly depend on the temperature or the mechanical stress level in the material or both. Fig. 4 shows e.g. the measured hysteretic magnetization curves of alloy steel under different levels of applied mechanical stresses at room temperature [1] and Fig. 5 the magnetostrictive strain of an electrical steel grade at different mechanical stresses and as function of the magnetic flux density [2]. Also, the rigid motion of different parts of an energy conversion device contributes and is a decisive factor in determining e.g. the current drawn from a voltage source by the device. This phenomenon can be appreciated when comparing the current of an electrical machine or an electromagnet in the case when the rotor or the plunger is under motion with the case of blocked rotor or plunger for example.

The design of efficient energy conversion devices and the prediction of their operation under different load and fault conditions as well as the minimization of the parasitic phenomena caused by these devices require not only simulation models from different fields of science vis. electromagnetism, mechanics etc. but also in most cases a coupling methodology between these models. In this paper we will focus on the magneto-mechanical or magneto-elastic models and the way they are coupled. In some examples the electromagnetic coupling is also handled.

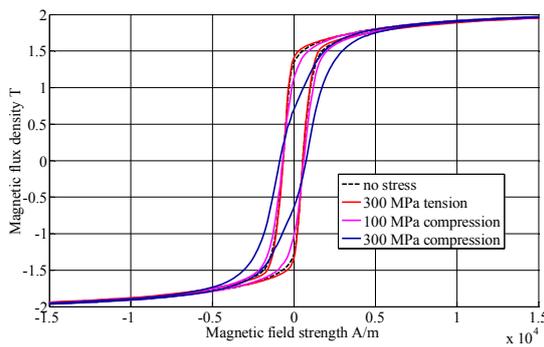


Figure 4: Measured quasi-static BH-loops of alloy steel at different applied stresses. The stress state in the sample is different from the applied one due to the magnetostriction. The frequency is 0.01 Hz.

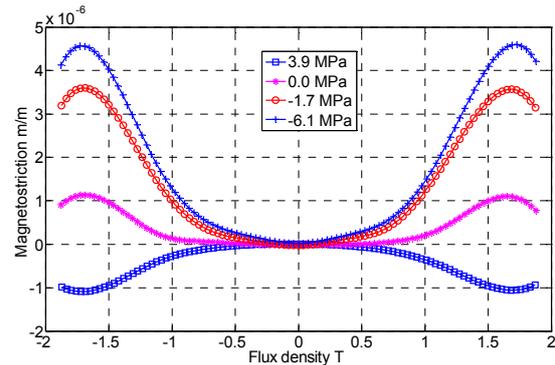


Figure 5: Measured magnetostrictive elongation of electrical steel under different applied stresses. The stress and the magnetic field are in the same direction and the measurement frequency is 5 Hz.

In Section 2 we will first establish a general context in which the coherent definition of different terms used to specify the nature of the coupling and its methodology is possible. In Section 3 we will review and explain the different levels of magneto-mechanical coupling and how they have been used in the recent literature. Section 4 clarifies the methods for magnetic force computation as well as the magnetostriction and the ways to deal with it. Last, in Section 5 we present some challenges and open issues in the magneto-mechanical modeling.

2 CLARIFICATION OF TERMS AND CONCEPTS

If one goes through the published literature dealing with the coupled problems in general and the electro-magneto-mechanical coupling in particular, he/she will find a multitude of terms used to specify the nature of the coupling methodology for solving a given problem. Such terms as direct or indirect coupling as well as strong and weak coupling are deliberately used and seldom explained. In this section we will first clarify these terms.

We propose a separation between the physical or phenomenological aspect of the coupling and its computational methodology or implementation aspects. In this respect, the physical coupling could be either strong or weak depending on the level of interaction between the fields or in other words how the change in one field say the magnetic field e.g. affects the change in the other field say displacement field e.g. It is obvious that this definition is a subjective one and depends on the accuracy at which we are aiming. Yet another terminology related to the phenomenological aspect is the concept of global and local coupling. By global coupling we mean an interaction between the field quantities without effects on the constitutive relations of the underlying materials, whereas the local coupling means the participation of the materials constitutive relations to this interaction. These concepts of weak vs. strong and local vs. global can be combined in pairs to specify the level of coupling both from its strength and its nature point of views. E.g. a problem can present strong global coupling and weak local coupling as is the case when modeling the operation of an electrical machine without interest in its vibrations. The strong global coupling here describes the coupling between the rotor motion and the magnetic field and the weak coupling describes the fact that in these conditions single-valued stress-independent magnetization properties are used to model the magnetic material and that the magnetostriction can be ignored.

The terms implicit, explicit, direct and indirect coupling should be reserved for the computational or implementation level. One speaks about explicit coupling when the governing equations for the quantities from different fields are written in a closed form, whereas the implicit coupling relates to governing equations that are written for each set of quantities separately but with the awareness that each set of equations includes parameters that have to be solved or updated from the other set. In both cases the equations can be formulated and solved either simultaneously and one speaks about a direct method or sequentially and we are speaking about an indirect method. Fig. 6 illustrates the intended use of these terms with some examples. It should be noted here that the explicit coupling is possible only in very special cases where some assumptions on the geometry of the problem or the properties of the materials are to be made. The explicit coupling also results in large system matrices with lower level of sparseness, which usually makes the solution of such problems slow. It is also to be noted that in most cases the solution of a coupled problem requires an iterative procedure regardless of the computational methodology used.

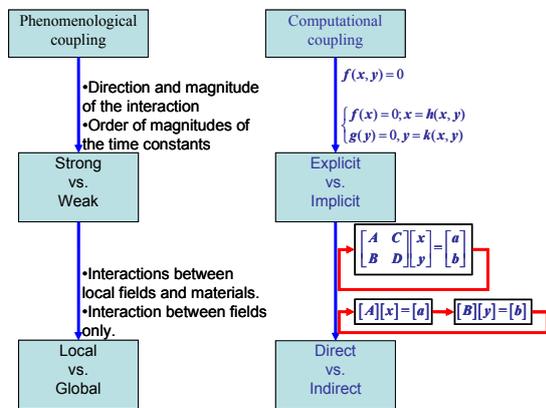


Figure 6: Proposed terminology for coupling methodology. x and y are hypothetic fields. f , g , h , and k are functions or operators defining the governing equations of the fields. A , B , C and D are matrices or sub-matrices, which in conjunction with the load vectors a and b define the algebraic equations to be solved. These vectors and matrices are in general depending on the fields and the problem is generally nonlinear and requires an iterative solution procedure.

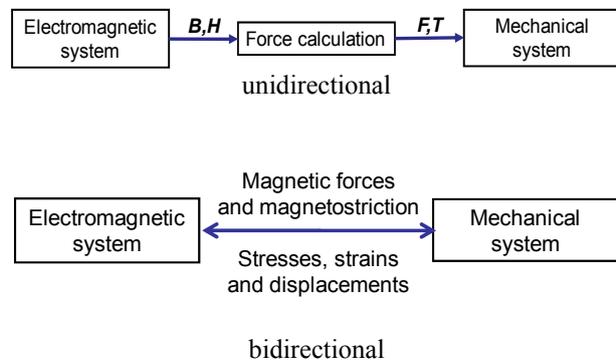


Figure 7: The Flow charts of the unidirectional and bidirectional magneto-mechanical coupling. B and H are the magnetic flux density and the Magnetic field strength. F and T are the magnetic forces and the torque. In some cases the magnetic vector potential A and the scalar electrical potential Φ are used instead of B and H . The unidirectional coupling takes only the effect of electromagnetic system on the mechanical one whereas the bidirectional coupling is simultaneous and all the phenomena happen at the same time in both systems.

3 THE DIFERENT LEVELS OF MAGNETO-MECHANICAL COUPLING

Now that the terminology is explained we will proceed with the different levels of coupling starting from the simplest one and in a bottom-up fashion evolving to the most general case and concentrating on the magneto-mechanical phenomena only.

3.1 Unidirectional coupling

The flow chart of the unidirectional magneto-mechanical coupling is shown in Fig. 7. From the phenomenological point of view such a coupling method suites a weak coupling, where there is no effect of the mechanical displacements on the magnetic field or the

underlying magnetic materials. In such a methodology the nonlinear Maxwell equations governing the magnetic field in the device are solved assuming that the magnetic material properties such as the permeability of iron or the relation between \mathbf{H} and \mathbf{B} , are not depending on the mechanical state of the material i.e. the mechanical stress. Such an assumption is usually possible due to the low level of stresses in the core of the machine and also due to the fact that the mechanical displacements are very small, except when a rigid motion is involved. The rigid motion is possible to handle separately. The one directional coupling is the most popular way of calculating the vibrations and noise from rotating electrical machines [3]-[6]. It allows for a complex and accurate electromagnetic modeling such as coupling the circuit equations of the machine windings and the electrical supply e.g. the frequency converter with the magnetic problem in the machine either in a 2D or 3D approach as well as for complex and accurate material modeling e.g. magnetic hysteresis [7]. It also allows for a complex modeling of the mechanical problem by the use of a detailed 3D geometry and updated FE model parameters [3], [5]. The coupling quantities in this approach are the magnetic forces that may or may not include the magnetostriction [2], [6], [8], [9]. The different methods for computing the magnetic forces and magnetostriction are explained later in Section 4. However, the most common ways to transfer the forces from the magnetic problem to the mechanical problem in this kind of coupling are either the so called teeth forces or the rotating stress waves in the air gap of the machine [2], [10]-[12]. Both methods are based on the Maxwell stress tensor as will be seen in Section 4. Fig. 8 shows a plot of the teeth forces for the stator of a 37 kW induction motor at a given time and Fig. 9 shows a spectral plot of the 2D Fourier decomposition of the Maxwell stress in the air gap of the same machine. Both forces are calculated from the 2D time stepping solution of the magnetic field in the cross section of the machine. The generalized nodal forces can also be used to couple the magnetic problem with the mechanical problem. The computation of these forces will be explained in Section 4.

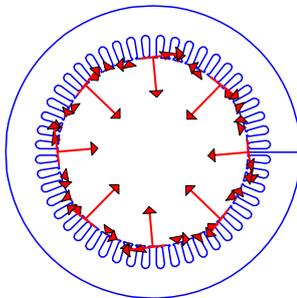


Figure 8: Normalized teeth forces of a 4 poles 37 kW induction machine at a given time. The direction and amplitude of the force vectors are time dependent.

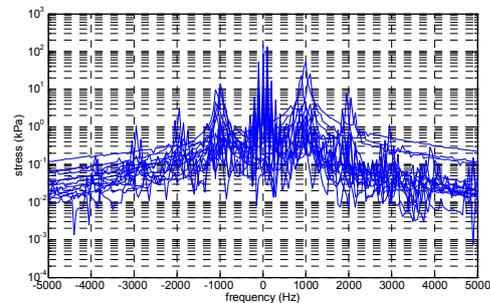


Figure 9: 2D spectral decomposition of the radial Maxwell stress in the air gap of the 37 kW machine. Each line corresponds to a spatial harmonic. Negative frequencies mean that the wave is rotating in opposite direction to that of the rotor.

3.2 Two directional coupling

In a two directional coupling it is possible to model both the effect of mechanics on magnetism and that of magnetism on mechanics. Such kind of coupling, suites best for phenomenological strongly coupled problems. This is the kind of coupling at which, most of the recent papers on the subject are aiming. Indeed, the solution of the magnetic problem requires knowledge of the material magnetic properties, which are stress dependent (see Fig.

4). On the other hand the solution of the elastic or mechanical problem requires knowledge of the load forces and the magnetostrictive strains, which depend on both the magnetic field and the stress of the material (see Fig. 5). Further, any change in the computed stresses and displacements will reflect on the computed magnetic forces through a change in the magnetic field and any change in the computed stresses and displacements or strains will reflect on the magnetic field through either the material properties or the geometry of the problem. These considerations are better understood from the flow chart presented in Fig. 7 above.

It should be understood that the coupling between the different phenomena is simultaneous and occurs in a continuous process rather than in a discrete or sequential one. This also means that the separation between the causes and consequences is rather artificial and should be used only for visualization purposes in view of a better understanding of the phenomena considered so that it could be used for the derivation of possible simplifications when needed.

In addition, the magneto-mechanical coupling is affected by the electric system, which is usually the source of the magnetic field. A load torque e.g. at the shaft of a machine connected to a voltage source will result in increased current withdrawn by the machine, which in turn will change the magnetic field configuration and further the force and the mechanical strains and displacements. Here again, the separation between the causes and the consequences is rather artificial and all these phenomena occur simultaneously. Reference [13] presents an example of this electro-magneto-mechanical coupling, where the coupling methodology seems to have some effect on the results of the computations. In the study, the motion of the plunger of an actuator, like the one in Fig. 2, has been modeled with time stepping and with its terminals connected to a voltage source. It was noticed that the current of the actuator was depending on whether the coupling was implemented as direct or indirect.

The magneto-mechanical coupling can be modeled in different manners depending on the nature of the coupling and the aim of the investigation. The vibrations of an electrical machine can be computed with reasonable accuracy with a unidirectional approach and without accounting for the effect of mechanical stresses or any other mechanical quantity on the material magnetic properties and the geometry of the machines. However, the computation of the hysteresis energy losses e.g. in the iron core of the machine requires knowledge of the stresses in the core of the machine as the losses are very much affected by these stresses [14]-[16]. The additional losses due to the stress affect the current withdrawn by the machine and the torque it produces and thus affects the stress state. The methodology to carry out such an analysis as well as the results of its application to an electrical machine has been presented in [17]. The methodology consists of presenting the reluctivity of electrical steel as a function of the magnetic flux density and the mechanical stress and solving the magneto-mechanical problem with an implicit approach and in an iterative fashion. In such analysis the magnetostriction of iron was ignored and the iron losses were computed in a posteriori manner which result in their effect on the current being ignored.

The magnetostriction of iron has been also modeled with equivalent magnetic forces acting on the structure of the machine [6], [9]. The computation method of these forces are presented later in section 4.3 and discussed altogether with other methods. It is worth notice that the approach adopted in the implementation of the problem is again implicit although the algebraic systems of equations have been solved simultaneously.

After time and space discretisation, the equations to be solved has the form

$$[\mathbf{P}] \begin{bmatrix} \Delta \mathbf{A}_{k+1}^n \\ \Delta \mathbf{u}_{k+1}^{r,n} \\ \Delta \mathbf{i}_{k+1}^{s,n} \\ \Delta \mathbf{u}_{k+1}^n \end{bmatrix} = [\mathbf{R}_{k+1}^n] \quad (1)$$

With \mathbf{P} the Jacobian matrix for the coupled magneto-elastic system, $\Delta \mathbf{A}_{k+1}^n$ the vector potential increment, $\Delta \mathbf{u}_{k+1}^{r,n}$ the rotor bar voltage increment, $\Delta \mathbf{i}_{k+1}^{s,n}$ the stator current increment, $\Delta \mathbf{u}_{k+1}^n$ the displacement increment and \mathbf{R}_{k+1}^n the residual at time step $k+1$ and iteration n [2]. As far as the phenomenological coupling is not very strong, the off-diagonal sub-matrices associated with the magneto-elastic coupling can be set to zero resulting in a sparse matrix but a lower convergence rate of the solution. When the coupling is strong due to material properties, the coupling terms could not be avoided.

3.3 Coupling through the material

The coupling procedures and methodologies presented in the previous section were implicit coupling in the sense that the equations for the magnetic and mechanical systems are written separately and the coupling is implemented through the magnetic and magnetostrictive forces. The equations of each system are then discretised and the solution is achieved through an iteration process that updates the material properties according to some rules or equations.

In [18], the coupling methodology is quite different. Here, the constitutive equations of the material are written in a coupled form, which is derived from energy considerations. Thus the coupling procedure is explicit and does not require separate modeling of the magnetostriction e.g. as a set of equivalent forces as will be discussed later. The explicit coupling, which requires coupled constitutive equation for the material, is the correct way to model magnetostrictive materials used in actuators. The explicit coupling presents also the possibility of modeling the so-called delta-E effect if the constitutive equations are well elaborated. However the explicit coupling required the energy functional to be parameterized and even though resulted in large system matrix with low level of sparseness as explained above. This is a natural property of the explicit coupling that requires coupling terms in the Jacobian matrix for the solution to converge.

The constitutive equations in [18] are written in terms of the stress tensor $\boldsymbol{\tau}$ and the magnetic field strength vector \mathbf{H} as functions of the strain tensor $\boldsymbol{\varepsilon}$ and the magnetic flux density vector \mathbf{B} . Additional parameters α_i are used to write the energy functional ψ of the magneto-mechanical system in terms of six invariants I_i , from which the constitutive equations are derived. Further, in defining the stress tensor it was assumed that the stress is the sum of the electromagnetic stress and an elastic stress $\boldsymbol{\tau} = \boldsymbol{\sigma}_{elastic} + \boldsymbol{\tau}_m$ where the electromagnetic stress tensor is defined in terms of the magnetic flux density \mathbf{B} and the magnetization \mathbf{M} vectors as

$$\boldsymbol{\tau}_m = \mu_0^{-1} \left(\mathbf{B} \otimes \mathbf{B} - \frac{1}{2} (\mathbf{B} \cdot \mathbf{B}) \mathbf{I} \right) + (\mathbf{M} \cdot \mathbf{B}) \mathbf{I} - \mathbf{B} \otimes \mathbf{M} \quad (2)$$

The space and time discretisation of the governing equations follows similar approach as in the implicit coupling and results in similar matrix equations.

4 METHODS FOR MAGNETIC FORCE COMPUTATION

The computation of magnetic forces from the solution of the magnetic field in a given geometry has been a subject of many discussions and publications. The main problem, was the fact that different computation methods were giving exactly the same total magnetic forces but the force distribution were different from one method to the other [19]. These differences were mainly present in magnetized media such as the iron core of an electrical machine. Lately, methods based on the principle of virtual work seem to gain confidence among the researchers and the application of these methods is somehow established especially in conjunction with the FEM.

In this section we present the main methods for force computation and explain how they can be used in coupling a magnetic system with a mechanical one.

4.1 Lorentz force

Consider a current carrying conductor with a constant permeability μ_0 . The force density within the coil is given by the classical Lorentz formula

$$\mathbf{f}_j = \mathbf{J} \times \mathbf{B} \quad (3)$$

Where $\mathbf{J} = \nabla \times \mathbf{H}$ is the current density in the conductor and $\mathbf{B} = \mu_0 \mathbf{H}$ is the magnetic flux density in the conductor. \mathbf{H} is the magnetic field strength. This formula can be used whenever the conductor has a constant permeability and is carrying a current. The resultant force acting on the conductor is naturally the integral of the force density over the volume of the conductor.

In magnetized media, the conduction current is usually zero and the Lorentz force equation results in null force density. However, the magnetized media can be represented at least in three different manners; surface magnetic pole distribution, surface current density, or a combination of both. In all cases, a local use of the Lorentz force distribution combined with the definition of magnetic moments, will result in a given force distribution either on the surface of the iron or inside it. A summary of these distributions is given all together with their discussion and consequences in [19].

4.2 Maxwell stress tensor

Starting from the Lorentz force formula, Maxwell derived a general purpose stress tensor that can be used to compute the force on any part on a device either magnetized or not

$$\boldsymbol{\sigma} = \frac{1}{\mu_0} \left(\mathbf{B} \otimes \mathbf{B} - \frac{1}{2} \mathbf{B} \cdot \mathbf{B} \mathbf{I} \right) \quad (4)$$

The force on any volume V bounded with the surface S is then calculated as

$$\mathbf{F} = \int_V \mathbf{f} dV = \int_V \nabla \cdot \boldsymbol{\sigma} dV = \oint_S \boldsymbol{\sigma} \cdot \mathbf{n} dS \quad (5)$$

Where \mathbf{n} is the normal vector to the surface element dS directed outwards of the volume element dV and $\mathbf{f} = \nabla \cdot \boldsymbol{\sigma}$ is interpreted as a force density, which result into

$$\mathbf{F} = \oint_S \left(\frac{1}{2\mu_0} (B_n^2 - B_t^2) \mathbf{n} + \frac{1}{\mu_0} B_n B_t \mathbf{t} \right) dS \quad (6)$$

Although no mathematical or theoretical evidence is given, the terms under the integral are

interpreted as a normal and tangential surface stresses or surface force densities. Such an interpretation has been the subject of a large number of publications.

If we accept this assumption (see Fig. 10 and 11), and in conjunction with rotating electrical machines, the stress can be developed into two dimensional Fourier series (space and time) and used as input or load for the mechanical system in the unidirectional coupling approach. A spectral plot of the radial component of the Maxwell stress in a 37 kW induction machine has been shown in Fig. 9. The usefulness of such a method is that even without making any mechanical analysis one can already “*guess*” what are the expected noise and vibration frequencies and modes that can be generated in the machine. Of course, all the frequencies and modes are not likely to be excited. The Maxwell stress or in general the electromagnetic stress tensor can also be used in the more general coupling method as explained in [18].

4.3 Method of virtual work

The method of virtual work for force calculation is not only one of the oldest methods to compute the magnetic forces but also is the one that have seen many developments in the last decades. The conventional virtual work method consisted of computing the magnetic field at two positions of a given moving part while the current or the flux are kept unchanged, computing the magnetic energy or co-energy at these positions and calculating the magnetic force as the ratio of the change in the magnetic energy or co-energy and the displacement of the part under investigation. Such a method is very heavy as it required two computations of the magnetic field and the displacement needed to be of adequate size for accuracy aspects. Reference [20] came with a method that best fits the FE computation and needs only one field solution. In this method the forces are computed from the derivative of the magnetic energy too but the derivation is made using the numerical and analytical properties of the FEM. However, the method was intended for the computation of total forces and torque of electrical machines and did not answer the critical question of force distribution. Reference [21] used the concept presented by [20] and applied it to the nodes of an FE mesh. In this way the author could resolve what he called “*generalized nodal forces*”. These forces do not represent the force distribution in the material. They rather give a combined method to transform the force distribution into a local total forces acting each on a given node of the mesh and representing the force on a volume around that node. The volume itself cannot be defined with this method neither the force distribution. However, the advantage of the generalized nodal force concept is that it allows for a coupling between the magnetic and the mechanical systems when they are treated with finite element method. They also allow for the use of the same FE mesh for the magnetic and elastic system and avoid the problem of projection from one mesh to the other. The computation routine for the nodal forces although originally present in terms of the magnetic vector potential has been extended to the general case of magnetic flux density and also could be derived from the Maxwell stress tensor as in [22]. These methods were using linear finite elements. Methods using higher order elements and methods to reconstruct the force distribution from the generalized nodal force have been presented in [23], [24]. Other projection methods related to the generalized nodal forces have been presented in [25]. A plot of the generalized nodal force in the stator of a 3 MVA synchronous machine is shown in Fig. 10. For comparison purpose a plot of the surface forces

from the Maxwell stress tensor is shown in Fig. 11. In both figures the field solution and the force computation have been carried out with linear triangular finite elements. Note that the scaling of the force vectors is not the same in the two figures but the shapes of the *force distributions* are similar and the absolute values are close to each other. This is due to the fact that the iron is saturated and the forces are concentrated on the surface of the iron [26].

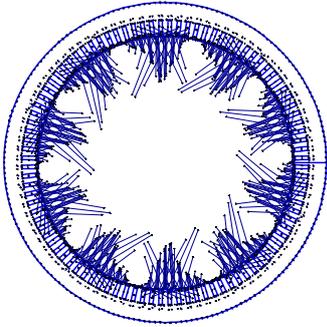


Figure 10: Generalized nodal forces in the stator of a 3 MVA synchronous machine.

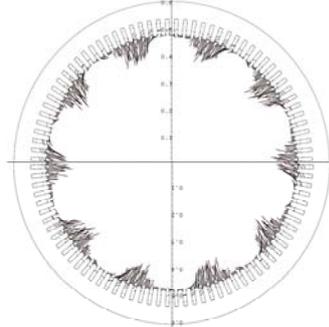


Figure 11: Maxwell force distribution on the stator of a 3 MVA synchronous machine.

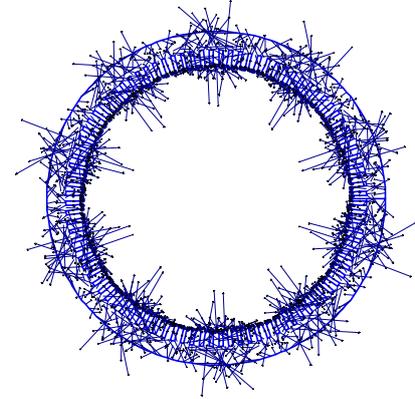


Figure 12: Magnetostrictive forces in the stator of a 3 MVA synchronous machine.

4.4 Magnetostriction

The magnetostriction is the phenomenon by which a sample of magnetic materials deforms under the effect of a magnetic field. Such a deformation is not due to the existence of external magnetic forces, it is rather the result of the internal stresses and strains in the material, which are due to the rearrangement of the magnetic domains and their interactions with the material lattices.

The magnetostriction has been traditionally accounted for through sets of equivalent forces, which are computed from either the magnetostrictive strain [6] or the magnetostrictive stress [8]. A plot of the magnetostrictive forces computed from the magnetostrictive stress in the stator of the 3 MVA synchronous machines is shown in Fig. 12.

However, the magnetostrictive stresses or strains are themselves dependent on the total stress in the material as shown in Fig. 5. Such a behavior could be simulated with stress dependents magnetostrictive forces [2] but it does not reproduce the actual stress behavior in the material when the materials boundary conditions are taken into account [2]. In this respect, the magnetostriction is better modeled within the material coupling methodology presented above [18]. In this methodology the magnetostriction is included in the magneto-elastic coupled constitutive equation through a relation between the stress, strain and magnetic flux density, which has been derived from energy considerations and measurements analysis. The fact that the magnetostriction is already in the constitutive equations makes it difficult to visualize it from complex computations which is the consequence of the fact that in real life the magnetic and mechanical coupling happen simultaneously and there is no way to separate them. Some attempts to find out how different phenomena in the magneto-mechanical coupling could be separated are still presented in [18].

5 OTHER CHALLENGES

In the previous sections, we handled the magneto-mechanical coupling as a separate problem from the electrical coupling and also from the mechanical load, which is usually connected to the shaft of electrical machines or the moving plunger of an actuator. We also assumed that the solution of the magneto-mechanical system is non-dissipative except for the dissipation in the resistive parts of the system. The other dissipations such as iron losses were estimated a posteriori.

The mechanical dissipation can be added to the system by additive mechanical damping [27] and the magnetic dissipation in the iron parts due to hysteresis and eddy currents can also be added through dynamic vector hysteresis models [7]. However, such additions to the different parts of the system result in non-coherent description of the energy, resulting in mathematically (and also physically) incorrect models. One challenge in the magneto-mechanical modeling is thus to include different dissipation phenomena at an early stage of the energy description, derive the dissipative constitutive equations for the material, and include these equations in the solution of complex systems such as electrical machines and actuators.

Most of the problems presented above have been implemented in two-dimensional analysis. This is due to the fact that the computational cost of these models is high and at the limit of what nowadays computation resources allow for. The computational resources are however developing very fast and already now one can carry out three-dimensional analysis of very complex systems within a reasonably short time. The development of the previous models and their adequacy for three-dimensional computations is then another challenge that has to be dealt with in the near future.

Last and not least, the characterization of the material and the identification of the models require experimental setups able to take measurements in three directions, whereas the existing setups are mainly designed for a single or at most two directional measurements.

REFERENCES

- [1] Belahcen, A. and Peussa, T. *Experimental and Numerical Investigations of the Inverse Magnetostriction-Based Mechanical Stress Sensing*. In Proc. COMPUMAG 2011, to be published.
- [2] Belahcen, A. *Magnetoelasticity magnetic forces and magnetostriction in electrical machines*. Dissertation, 2004, Helsinki University of Technology, Finland.
- [3] Roivainen, J. *Unit-Wave Response-Based Modeling of Electromechanical Noise and Vibration of Electrical Machines*. Dissertation, 2009, Helsinki University of Technology, Finland.
- [4] Carlson, R. Runcos, F. Oliveira, A. M. Kuo-Peng, P. Sadowski, N. *FEM Vibration Analysis of a Doubly-Fed Twin Stator Cage Induction Generator*. In Proc. COMPUMAG 2005, vol. 4, p. 222-223.
- [5] Schlensok, C. and Hameyer, K. *Body-Sound Analysis of a Power-Steering Drive Considering Manufacturing Faults*. IEEE trans. on vehicular technology, vol. 56, no. 4, July 2007, pp. 1553-1560.
- [6] Delaere, K. Heylen, W. Belmans, R. and Hameyer, K. *Comparison of Induction Machine Stator Vibration Spectra Induced by Reluctance Forces and Magnetostriction*. IEEE trans. on magn. vol. 38, no. 2, March 2002, pp. 969-972.
- [7] Dlala, E. Belahcen, A. Fonteyn, K. Belkasim, M. *Improving loss properties of the Mayergoyz vector hysteresis model*. IEEE trans. on magn. vol. 46, no. 3, 2010, pp.918-924.
- [8] Belahcen, A. *Magnetoelastic coupling in rotating electrical machines.*, IEEE trans. on magn., vol. 41, no. 5, 2005, pp. 1624-1627.
- [9] Belahcen, A. *Vibrations of rotating electrical machines due to magneto-mechanical coupling and magnetostriction*. IEEE trans. on magn., vol. 42, no. 4, 2006, pp. 971-974.

- [10] Salon, S. Chari, M. V. K. Sivasubramaniam, K. Kwon, O. M. and Selvaggi, J. *Computational Electromagnetics and the Search for Quiet Motors*. IEEE trans. on magn., vol. 45, no. 3, March 2009, pp. 1694-1699.
- [11] Henneberger, G. Sattler, Ph. K. Hadrys, W. Shen, D. *Procedure for the numerical computation of mechanical vibrations in electrical machines*. IEEE trans. on magn., vol. 28, no. 2, March 1992, pp. 1351-1354.
- [12] Belahcen, A. Arkkio, A. Klinge, P. Linjama, J. Voutilainen, V. Westerlund, J. *Radial forces calculation in a synchronous generator for noise analysis*. In Proc. CICEM 1999, vol. 1, pp. 119-122.
- [13] Ren, Z.; Razek, A. *Modelling of dynamical behaviours of electro-magneto-mechanical coupled systems*. In Proc. Second Intern. Conf. on Computation in Electromagnetics, 1994, pp. 20-23.
- [14] Basak, A. Moses, A.J. and Al-Bir, R. *Effect of clamping stress on power loss in power core strip and Si-Fe transformer cores*. IEEE trans. on magn. vol. 26, no. 5, September 1990, pp. 1999-2001.
- [15] Moses, A.J. and Rahmatizadeh, H. *Effects of stress on iron loss and flux distribution of an induction motor stator core*. IEEE trans. on magn., vol. 25, no. 5, September 1989, pp. 4003-4005.
- [16] Pitman, K.C., *The Influence of stress on ferromagnetic hysteresis*, IEEE trans. magn., 1990, vol. 26, no. 5, pp. 1978-1980.
- [17] Belahcen, A. and Arkkio, A. *Locally coupled magneto-mechanical model of electrical steel*. Intern. J. for Comp. and Math. in Elec. and Electronic Eng. COMPEL 2008, vol. 27, no. 6, pp.1451-1462.
- [18] Fonteyn, A. K. *Energy-based magneto-mechanical model for electrical steel sheets*. Dissertation, Aalto University 2010.
- [19] Carpenter, C. J. *Surface-integral methods of calculating forces on magnetized iron parts*. Proceedings of the IEE - Part C, vol. 107, no. 11, 1960, pp. 19-28.
- [20] Coulomb, J. L. *A methodology for the determination of global electromechanical quantities from a finite element analysis and its application to the evaluation of magnetic forces, torques and stiffness*. IEEE trans. on magn. vol. MAG-19, no. 6, November 1983, pp. 2514-2519.
- [21] Bossavit, A. *Edge-element computation of the force field in deformable bodies*. IEEE trans. on magn. vol. 28, no. 2, March 1992, pp. 1263-1266.
- [22] Kameari, A. *Local calculation of forces in 3D FEM with edge elements*. Intern. J. of Applied Electromagnetics in Materials, vol. 3, 1993, pp. 231-240.
- [23] Belahcen, A. *Nodal magnetic forces in non-linear triangular finite elements*. Intern. Sym. on Electric and Magnetic Fields EMF 2006, pp. 179-180.
- [24] Belahcen, A. *Computation of generalized nodal forces and force fields in electrical machines*. IEEE Conf. on Electromagnetic Field Computation CEFC 2008, 1p.
- [25] Parent, G. Dular, P. Ducreux, J.-P. and Piriou, F. *Using a Galerkin Projection Method for Coupled Problems*. IEEE trans. on magn. vol. 44, no. 6, June 2008, pp. 830- 833.
- [26] Belahcen, A. *Force calculation for vibration and noise analysis of electrical machines*. Intern. Conf. on Elec. Machines ICEM 2000, pp.1829-1833.
- [27] Belahcen, A. *Magnetoelastic coupling and rayleigh damping*. Intern. J. for Comp. and Mathematics in Elec. and Electronic Eng. COMPEL 2004, vol. 23, no. 3, pp.647-654.

COUPLING OF DAMAGE MECHANISMS IN THE PREDICTION OF CREEP FAILURE OF A WELDED BRANCHED HEADER PIPE

DAVID R. HAYHURST

School of Mechanical, Aerospace, & Civil Engineering,
The University of Manchester,
George Begg Building C-004,
MANCHESTER, M13 9PL, U.K.
e-mail: d.r.hayhurst@manchester.ac.uk

Key words: High Temperature Creep, Continuum Damage Mechanics, Coupled Creep-Damage, Constitutive Equations, Finite Element Method, Lifetime Prediction, Components.

Abstract. *The paper introduces the concept of high-temperature creep Continuum Damage Mechanics (CDM); and then presents the associated physics-based creep constitutive equations used to model CDM. The equations involve the coupling of inelastic creep deformation and damage mechanisms. The calibration of the constitutive equations from available creep data is addressed; and materials constants are presented for parent, weld and intermediate phase materials found in welds used in high-temperature pipework in the power generation industry.*

The paper continues with the introduction of a ferritic steel welded branched header pipe that is subjected to a constant internal pressure of 4 MPa at a temperature of 590 °C. The pressurised branch was tested to failure, and the associated test data is used as a benchmark against which lifetime and damage evolution predictions are assessed. The results of two Finite Element solutions are presented that have been obtained using the FRONTAL and the PGC equation solvers. Excellent correlation is shown between laboratory test results on the branch and the Finite element predictions.

1 INTRODUCTION

High temperature creep and rupture of materials is a major limiting factor in the design and operation of power generation plant and turbines. Lifetimes can be of the order of 35 years, and hence there is a need to predict deformation and the potential for failure over that time. Materials data is typically available for much shorter periods of time, and physics-based material models are needed that allow accurate data extrapolation. The problem is made more severe due to the non-linearity of creep in both stress and temperature. An example of a set of uni-axial creep curves, for Type IV weldment phase material at 640°C, is given in Fig.1 [1]. Creep curves have three stages: the initial or primary region, where strain rate decrease due to hardening; a secondary phase in which creep rates are approximately constant, where hardening and softening mechanisms balance; and a third or tertiary stage, where the internal softening mechanisms dominate, strain rates tend to infinity, and failure occurs. The paper concerns the identification of constitutive equations and the associated damage

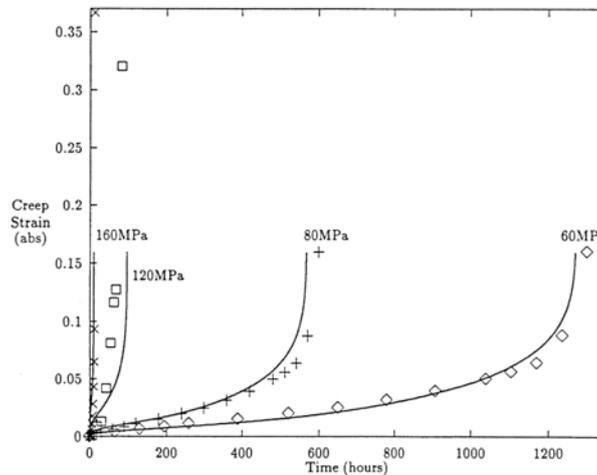


Figure 1: Comparison of experimental (symbols) and predicted (continuous) creep curves for simulated Type IV weld material at 640°C [1].

evolution equations that can be used to model their coupling, and to accurately predict creep strains and lifetimes of both testpieces and engineering components. To illustrate the process an internally pressurised branched steam pipe has been selected [2].

The calculation method utilises the concept of creep Continuum Damage Mechanics (CDM) [3,4] and this will be introduced first. Then the formulation of the creep damage evolution equations will be addressed, concentrating on the range of possible mechanism. Discussed next is the identification of the controlling damage mechanisms, and the calibration of the materials constants in the relevant damage and constitutive equations. The experimental study of a welded branched header pipe is then briefly outlined; and this is followed by the presentation of the results of full CDM analyses on the vessel.

2 CONCEPT OF CREEP CONTINUUM DAMAGE MECHANICS

Creep damage evolves during tertiary creep, and some mechanisms can be easily

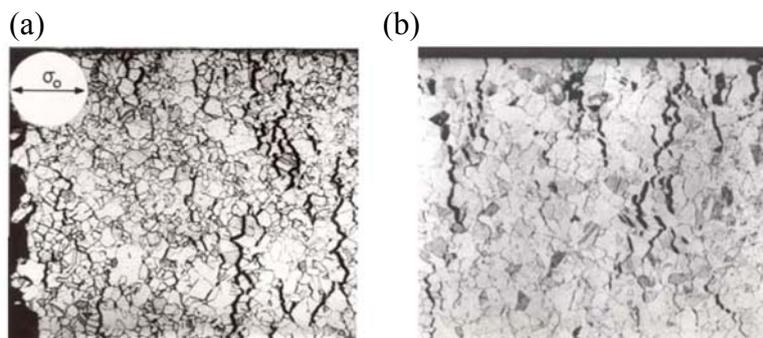


Figure 2: Micrographs taken from mid-thickness sections of 99% pure copper testpieces at 250°C after creep rupture: (a) bar subjected to a uni-axial stress σ_0 [5]; and, (b) the top, outer surface of a rectangular sectioned beam subjected to a uniform bending moment [6]. Failure times are of the order of 1500 h.

Observed using optical microscopy. This can be seen in Fig. 2 for 99% pure copper either at or close to failure [5,6]. Fig. 2a shows a uni-axial testpiece with its failure surface on the left hand edge. It can be seen that where the stress field is uniform then the damage field, shown as grain boundary cracks, is also uniform. Figure 2b shows a micrograph of the top, outer surface of a rectangular sectioned beam that has been subjected to a uniform bending moment [6]; the vertical failure surface is not shown on the figure. The stress on the upper surface of the beam was tensile, decreasing in magnitude as one moves down the figure, what is evident is that as the stress reduces, so does the intensity of damage. These two figures clearly demonstrate the existence of CDM.

3 FORMULATION OF COUPLED CREEP CONSTITUTIVE EQUATIONS

In this section a range of tertiary creep damage evolution mechanisms are presented, along with the equations that describe the controlling physics.

3.1 Hardening in primary creep

The parameter H models dislocation structures; during the process of hardening, dislocation entanglement, competes with softening, dislocation disentanglement, by thermal activation.

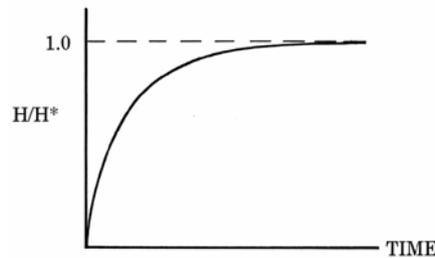


Figure 3: Change of Primary Creep State Variable H with time, [7].

Figure 3 shows this variation of H/H^* with time; the saturation limit of $H=H^*$ being achieved when secondary creep is attained. The governing equation is:

$$\dot{H} = (h\dot{\varepsilon}_e / \sigma_e)(1 - (H / H^*)), \quad (1)$$

where σ_e and ε_e are respectively the effective stress and strain and h is a material constant [7].

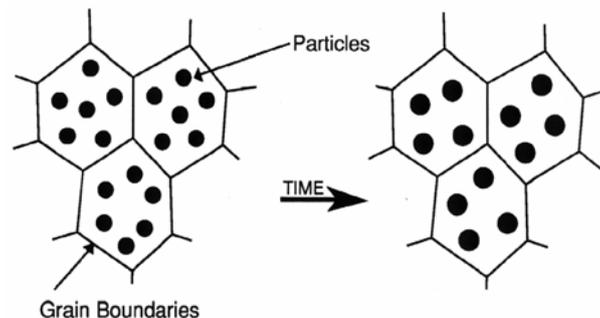


Figure 4: Softening due to stress independent ageing and particle coarsening [8].

3.2 Softening due to stress independent ageing and particle coarsening

The parameter Φ models the stress independent ageing due to the coarsening of particles, typically carbides contained within the grains, which are used to prevent dislocation motion [8]. This is shown schematically in Fig, 4. The parameter Φ increases monotonically from zero to unity. The controlling physical equation is:

$$\Phi = (K_c / 3)(1 - \Phi)^4, \quad (2)$$

where K_c is a material constant.

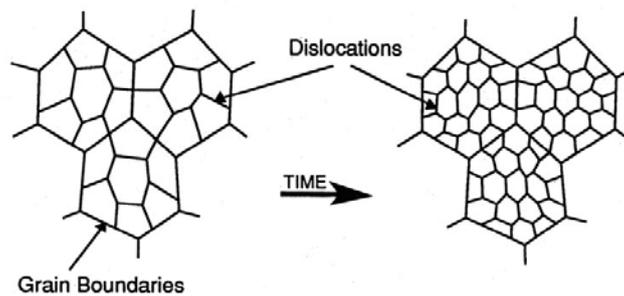


Figure 5: Softening due to multiplication of dislocation substructures [9].

3.3 Softening due to multiplication of mobile dislocations

A common softening mechanism is the multiplication of dislocation sub-structures as shown schematically in Fig.5 [9]. The evolution equation for the damage parameter ω_1 is:

$$\dot{\omega}_1 = C(1 - \omega_1)^2 \dot{\epsilon}_e \quad (3)$$

where C a material constant.

3.4 Softening due to void nucleation and growth

A dominant mechanism for low stress, long term, creep rupture is damage by creep

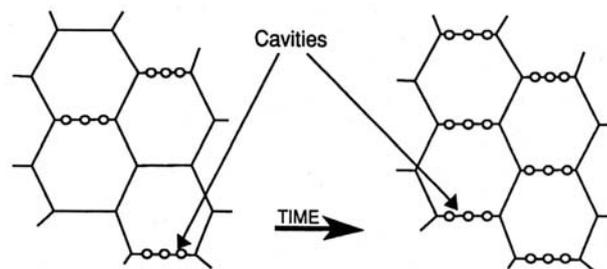


Figure 6: Softening due to grain boundary cavity nucleation and growth [9].

constrained cavitation, ω_2 , which is shown schematically in Fig. 6 [9]. The population of damaged grain boundary facets increases with time. The governing equation is:

$$\dot{\omega}_2 = DN \dot{\epsilon}_e \left(\sigma_1 / \sigma_e \right)^v, \quad (4)$$

where σ_1 is the maximum principal tension stress, D is a material constant and N is the Heavyside unit function, with $N=1$ for $\sigma_1 \geq 0$. The term in brackets reflects the multi-axial stress-state sensitivity of the damage process [5].

3.5 Softening due to continuum cavity growth

This mechanism dominates at higher stresses under conditions of tri-axial stress and the associated parameter is ω_3 ; it is shown schematically in Fig. 7 [8,10]. A nucleated void grows by power law creep of the material adjacent to a grain boundary and of that of the surrounding

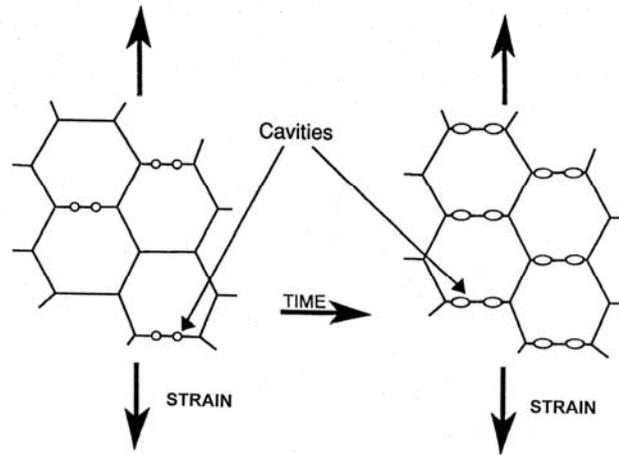


Figure 7: Softening due to continuum cavity growth [8].

grains. The controlling physical equation is:

$$\dot{\omega}_3 = \Theta \dot{\epsilon}_e \left\{ \frac{1}{(1-\omega_3)^n} - (1-\omega_3) \right\}, \quad (5)$$

where $\Theta = \tilde{H} \sinh \{ q(J_1 / \sigma_e) \}$, $J_1 = (\sigma_1 + \sigma_2 + \sigma_3)$ and σ_1 , σ_2 and σ_3 are principal stresses and n is the Norton's creep index.

4 CREEP DEFORMATION RATE EQUATION

It is widely accepted that the best equation that can be used to describe the creep strain rate process over a wide range of stress is the sinh function equation [9,10], shown here for primary-secondary creep:

$$\dot{\epsilon}_{ij} = A \left\{ 3s_{ij} / 2\sigma_e \right\} \sinh \left\{ B\sigma_e (1-H) \right\} \quad (6)$$

where A and B are material constants. The inclusion of the damage parameters, Θ , ω_1 , ω_2 , ω_3 , given by equations (2-5), is done by their inclusion in the argument of the sinh function and in the multiplier term [8,9].

4.1 Coupling with Elasticity

Elastic strains, at any instant, e_{ij} , are given by the relation: $\sigma_{ij} = C_{ijkl}e_{kl}$, where C_{ijkl} is the elasticity tensor. The total strains, ξ_{ij} , are assumed to be the sum of the elastic and creep strains, and given by $\xi_{ij} = e_{ij} + \varepsilon_{ij}$.

5 CALIBRATION OF CONSTITUTIVE MATERIALS CONSTANTS

The basic material model comprising one state damage evolution equation and a single creep constitutive equation may be calibrated against uni-axial data using most optimisation schemes [11]; and after the identification of an appropriate optimisational functional they converge well with good accuracy. But when equation (6) is coupled with the differential equations (1-5) they contains approximately twelve independent constants, and each has to be determined to provide a good fit to the experimental data. Use of schemes due to Dunne *et al* [11], for the basic equation set, has not proved successful for the multiple equation set. The difficulty has been overcome by Kowalewski *et al* [12] who developed a new approach; it involves taking each physical mechanism, or the equation that describes its kinetics i.e one of the equations (1-5), and isolating it from other mechanisms in such a way that it yields good estimates of the material constants for that mechanism. After the determination of good estimates for as many materials constants as possible, the values are used as initial values in the numerical minimisation of the functional Ψ :

$$\text{Minimum } \Psi = \sum_{i=1}^a \left[\left\{ \sum_{j=1}^b (\varepsilon_j^p - \varepsilon_j^{exp})^2 \right\}_i \right] + Z_i (t_i^p - t_i^{exp}) / t_i^{exp} \quad (7)$$

where a is the number of creep curves, $a = 4$ in Fig. 1, b is the number of data points per



Figure 7: Medium bore branch under test at a constant pressure of 4 MPa at 590°C.

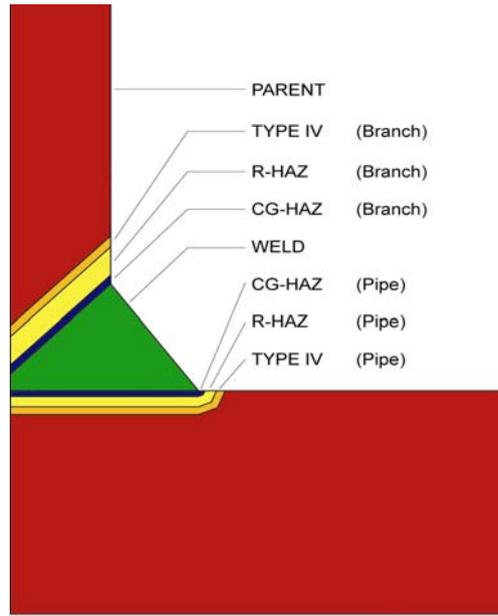


Figure 8: Five material zones of weld shown on crotch section

curve, ε^p and ε^{exp} are respectively the predicted and experimental values of strain, Z_i is an amplification constant, t^p and t^{exp} respectively denote predicted and experimental lifetimes. The second term in (7) is used only when $t^p > t^{exp}$. Progressing in this way, and by inclusion or exclusion of each of the equations (1-5), the quality of fit to the experimental data can be assessed by the magnitude of the Kernal functional Ψ . Hence, the dominant physical mechanisms can be identified. It is in this way that the appropriate mechanism-based constitutive equation can be identified from the available uni-axial materials data.

5.1 Constitutive parameters for the medium bore branched welded pipe.

The material of the branched welded pipe to be examined in this study is 0.5Cr, 0.5Mo, 0.25V and the weld material is 2.25Cr, 1Mo, c.f. Fig. 7. The two crotch positions are marked to be in-line with the axis of the main pipe. The welded pipe under consideration is composed of five different materials zones c.f. Fig. 8. For these materials, at 590°C, the procedure outlined above has been enacted and it has been found that $C = \Theta = 0$ in equations 3 and 5 respectively, and that the coupled set of differential equations is:

$$\begin{aligned}\dot{\varepsilon}_{ij} &= \frac{3s_{ij}}{2\sigma_e} A \sinh \left[\frac{B\sigma_e(1-H)}{(1-\Phi)(1-\omega)} \right], \\ \dot{H} &= (h\dot{\varepsilon}_e / \sigma_e)(1-(H/H^*)), \\ \Phi &= (K_c/3)(1-\Phi)^4, \\ \dot{\omega}_2 &= DN\dot{\varepsilon}_e(\sigma_1 / \sigma_e)^V,\end{aligned}\tag{8}$$

where $v = a \{ \exp(b\sigma_e / \sigma_o) \}$, and a and b are material constants. The materials constants determined in this way are given in Table 1.

Table 1: Ferritic steel constitutive parameters used in the medium bore branched pipe analyses [13]

590°C	Parent	Weld	Refined HAZ	Type IV	CG HAZ
A (h^{-1})	1.3913×10^{-9}	2.8898×10^{-8}	1.3913×10^{-9}	5.2879×10^{-9}	1.5844×10^{-10}
B (MPa^{-1})	1.4337×10^{-9}	1.1109×10^{-1}	1.4337×10^{-1}	1.1762×10^{-1}	1.9746×10^{-1}
C (-)	2.9983	3.7894	2.9983	6.7693	8.8783
h (MPa)	7.3288×10^4	2.4649×10^4	7.3288×10^4	2.9714×10^4	6.2785×10^4
H* (-)	0.4218	0.4088	0.4218	0.543	0.6281
K _c (h^{-1})	1.0020×10^{-4}	5.2579×10^{-5}	1.0020×10^{-4}	6.3709×10^{-4}	1.4046×10^{-4}
v (-)	2.8	2.7955	2.8	5.3106	2.7977

6 EXPERIMENTAL STUDY OF A WELDED BRANCHED HEADER PIPE

The medium bore main pipe was welded symmetrically, at right angles, to the branch. The main body of the vessel (pipe) was 465mm in outer-diameter with a wall thickness of 20mm; the branch was 111mm in inner-diameter with a wall thickness of 8mm. The mean radius to thickness ratio of the main pipe was $R/t_p = 11.125$, and the mean radius to thickness ratio of the branch r/t_b was 7.438. Both the main pipe and branch were pressurised to a constant level

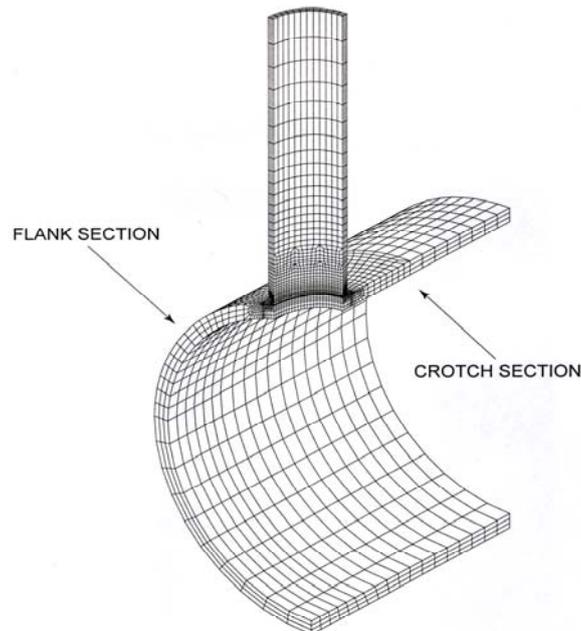


Figure 9: Finite element brick mesh of medium bore vessel [13].

of 4.0 MPa, at a constant uniform temperature of 590°C. The parent and weld materials have the same Young's Modulus of 160 GPa at 590°C.

For the class of low alloy ferritic steels studied here, it is established practice to stress relieve welds using an appropriate heat treatment; and, in the CDM analyses the welds have been assumed to be initially stress free.

The weld geometry was specified at two geometrical locations, namely the crotch and flank sections. The flank plane is normal to the crotch plane, and contains the axis of the branch. The geometry of the weld has been defined by Hayhurst *et al* [13].

7 FINITE ELEMENT FORMULATION

A single quadrant of symmetry of the branch is shown in Figure 9; the planes of symmetry are the flank and crotch planes. The volume has been modelled with six sided bricks. Each brick is then filled with 24 constant strain tetrahedra; the total number of nodes is 54,687, and the number of degrees of freedom is 164,061, and the number of elements is 240,048. The internal pressure conditions of 4MPa were applied using nodal forces normal to the internal surfaces of the pipe and branch; and the pipe and branch endloads were applied using nodal forces parallel to the pipe and branch axes. Zero normal displacements were applied to nodes on the two planes of symmetry.

The CDM numerical analysis involves the solution of a combined boundary-initial value problem. Having obtained the elastic solution at $t = 0$, the current rates of damage and creep strain are determined using the CDM-based constitutive equations set (8). These rate equations have been solved as a system of ordinary differential equations using a fourth-order Runge-Kutta method [14]. This procedure involves the repeated solution of a boundary-value problem to determine the field quantities required for the numerical solution. The integration technique uses a time step control algorithm based on error estimation; and, time steps are reduced until absolute and relative errors on normalised creep strain rate are met. When an element reaches the failure condition, $\omega_2 \geq 0.7$, it is removed, and the boundary-value problem is redefined. The field variables from the previous iterations become the new starting values for the next iteration. The procedure is then repeated until complete failure of the damage vessel occurs.

The first solution route taken has been to use the *FRONTAL* numerical solver written by Scott [15] for use on parallel architecture computers; and, to employ a heuristic time step algorithm due to Vakili-Tahami *et al* [16]. Both of these developments have resulted in a speed-up of the running of the code by a factor of 35 relative to that for a single processor [17]. However, for the 164 061 degrees of freedom the problem took 70 days to run. The second solution route taken has been to use a parallel iterative Preconditioned Conjugate Gradient (PGC) solver, written by Salimi and Hayhurst [18] as a direct replacement for the frontal solver discussed above. The speed-up achieved from parallelism scaled up to 64 processors, and this reduced the computation time to 3 days [19].

Table 2: Comparison of predicted and experimental times to first failure, and lifetimes

	FRONTAL Solver	PGC solver
First Failure Time	8 826 h	8 820 h
Rupture Lifetime	18 242 h	18 233 h
Experimental Lifetime	20 038 h	20 038 h

8 PRESENTATION OF NUMERICAL RESULTS

For the solutions carried out with both the FRONTAL and PGC solvers, the location of the initiation of failure, i.e. the first element to fail, was identical. It occurred at the weld toe in the crotch plane at approx 8 825 h, c.f Table 2. Failure progressed within the Type IV material with greatest concentration in the crotch plane; a contiguous failure path was predicted at approximately 18 235 h at which point the steam contained within the branch would be expected to leak out. The failure path is shown on the damage plots of ω_2 in Fig 10;

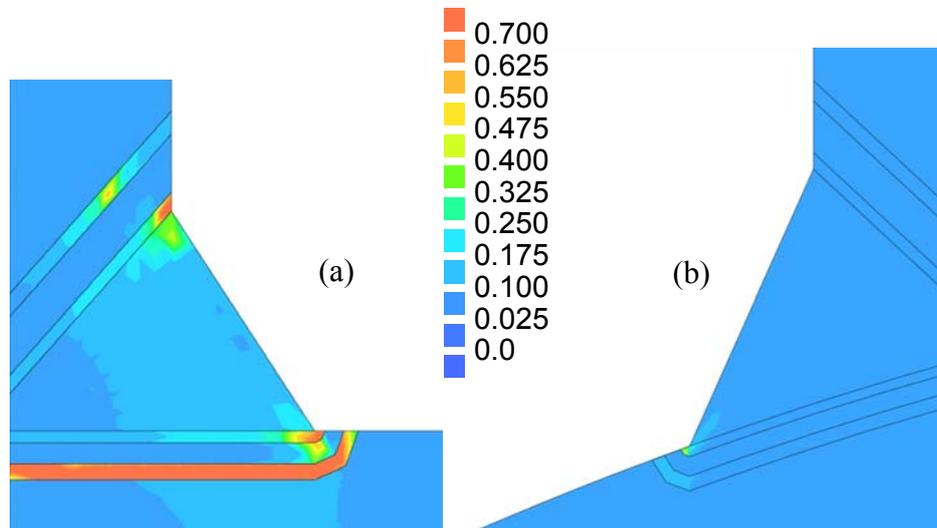


Figure 10: Damage contour plots, ω_2 , predicted for diametral sections of the branch at $t/t_f=0.99$, (a) crotch section, and (b) flank section.

the red zone of Fig 10a shows the failure, or steam leak path. The levels of damage are everywhere else low. Damage levels on the orthogonal Flank plane are extremely small.

Table 2 shows that the time to first failure and the lifetime predicted by the two methods are extremely close. However, the predicted lifetimes are of the order of 9% lower than the experimental value. When interpreted in terms of a design stress this corresponds to a stress which is almost 2% too high; however it is a most acceptable result.

9 DISCUSSION AND CONCLUSIONS

The paper has shown the complexity of modelling interacting creep damage mechanisms and has demonstrated a formalism that is based within the Continuum Damage Mechanics

framework for accurate physical modelling and extrapolation. The latter is justified on the basis that the differential equations that describe damage evolution are physics-based. The damage evolution equations are coupled with the creep strain rate equation, and are also coupled through the interaction with elasticity.

The difficulty of identifying the strongest damage mechanisms and the subsequent determination of the corresponding constitutive materials constants has been addressed using an appropriately defined Kernal function and appropriate optimisation methods.

A low alloy ferritic steel branched welded pipe, for which experimental test data is available at 590°C has been analysed using the above techniques. The weldment that connects the two orthogonal pipes, has four materials zones, each of which has been characterised in constitutive materials terms. The derived equations have been used as input to two different finite elements schemes.

Since the CDM finite element solution is computationally demanding, the speed-ups that computer parallelism can provide have been exploited. It is the equation solver that occupies the majority of the solution time, and therefore has to be fast. Firstly a FRONTAL equation solver has been used that takes 70 days to run on 4 processors; and secondly a PGC equation solver has been employed that takes three days to run on 64 processors. Both methods give near identical results, with the lifetimes being close to those observed experimentally.

REFERENCES

- [1] Perrin, I.J. and Hayhurst, D.R., *Continuum Damage Mechanics Analyses of Type IV Creep Failure in Ferritic Steel Crossweld Specimens*, (1999) *Int. J. Press. Ves. & Piping*, **76**: 599-617.
- [2] Mustata, R., Hayhurst, J.R., Hayhurst, D.R., and Vakili-Tahami, F., *CDM Predictions of Creep Damage Initiation and growth in Ferritic Steel Weldments in a Medium Bore Branched Pipe under Constant Pressure at 590°C using a 4 Material Weld Model*, (2006) *Archive of Applied Mechanics*, **75**, 475-495.
- [3] Hayhurst, D.R., Dimmer, P.R. and Morrison, C.J., *Development of Continuum Damage in the Creep Rupture of Notched Bars*, (1984), *Phil. Trans. Roy. Soc. Lond. (A)*, **311**, 103-129.
- [4] Hayhurst, D.R., *CDM Mechanisms-Based Modelling of Tertiary Creep: Ability to Predict the Life of Engineering Components*, (2005), *Archives of Mechanics*, **57**, 2-3, 71-100.
- [5] Hayhurst, D.R., *Creep Rupture under Multiaxial States of Stress*. *J. Mech. Phys. Solids*, (1970), **20**: 381-390.
- [6] Hayhurst, D.R., Vakili-Tahami, F. and Radford, R., *Continuum Damage Mechanics: From Creep Rupture under Homogeneous Stress to Creep Crack Growth*. In *Continuous Damage and Fracture*, 385-396, Ed. Ahmed Benallal, Elsevier. (2000).
- [7] Hayhurst, D.R., *Materials Databases and Mechanisms-Based Constitutive Equations for use in Design*, Chap., 5, 501-562, *Creep and Damage in Materials and Structures*, Springer-Verlag, Vienna & New York, No. 399, Edited by: J.J.Skrzypek and H. Altenbach, (1999).

-
- [8] Hayhurst, D.R., Lin, J. and Hayhurst, R.J. *Failure in notched tension bars due to high-temperature creep: Interaction between nucleation controlled cavity growth and continuum cavity growth*, Int. J. of Solids and Structure, (2008) **45**, 2233-2250.
- [9] Othman, A.M, Hayhurst, D.R. and Dyson, B.F. *Skeletal point stresses in circumferentially notched tension bars undergoing tertiary creep modelled with physically-based constitutive equations*, (1993), Proc. R. Soc Lond., A **441**, 343-358.
- [10] Cocks, A.C.F. and Ashby, M.F., *On creep fracture by void growth*, (1982), Prog. Mater. Sci. **27**, 189-244.
- [11] Dunne, F.P.E.F., Othman, R. Hall, F.R. and Hayhurst, D.R., *Representation of uni-axial creep curves using continuum damage mechanics*, (1990), I. J. Mech Sci, **32**, 11, 445-957.
- [12] Kowalewski, Z.L., Hayhurst, D.R. and Dyson, B.F., *Mechanisms-based creep constitutive equations for an aluminium alloy*, (1994), J. Strain Analysis, **29**, 309-316.
- [13] Hayhurst, D.R. Hayhurst, R. J. and Vakili-Tahami, F. *Continuum damage mechanics predictions of creep damage initiation and growth in ferritic steel weldments in a medium bore branched pipe under constant pressure at 590°C using a five-material weld model*, (2005), Proc. R. Soc., A **461**, pp. 2303-2326.
- [14] Hayhurst, D.R., and J.T. Henderson, *Creep Stress Redistribution in Notched Bars*, (1977) Int. J. Mech. Sci., **19**, 133-146.
- [15] Scott, J. *HSL_MP62 Symmetric finite-element system: multiple-front method*. AEA Technology Engineering Software, The Gemini Building, Harwell International Business Centre. Oxford, (2002).
- [16] Vakili-Tahami, F., Hayhurst, D.R. and Wong, M.T., *High-temperature creep rupture of low alloy ferritic steel but-welded pipes subjected to combined internal pressure and end loadings*, (2005), Phil. Trans. R. Soc., **A363**, 2629-2661.
- [17] Hayhurst, R. J., Vakili-Tahami, F. and Hayhurst D. R. 2009 *Verification of 3-D parallel CDM software for the analysis of creep failure in the HAZ region of Cr-Mo-V crosswelds*, Int. J. Press. Vessels & Piping, **86**, 475-485.
- [18] Salimi, B. and Hayhurst, D. R., *Reduced Numerical Solution Times for Combined Boundary-Initial Value Problems using Parallel Computing*, (2009), Engineering Computations, **27**, No.6, 746-772.
- [19] Salimi, B. and Hayhurst, D. R. *Continuum Damage mechanics predictions of the Effect of Bi-axiality of loading on the Creep Rupture Lifetime of a Ferritic Steel medium Bore Branched Pipe*, Research Report No. DMM. 10 01, The University of Manchester, School of Mechanical, Aerospace and Civil Engineering, (2009).

A FINITE ELEMENT FOR NONLOCAL ELASTIC ANALYSES

FRANCESCO MAROTTI DE SCIARRA

Department of Structural Engineering
University of Naples Federico II
Via Claudio 21, 80125 Napoli, Italy
e-mail: marotti@unina.it

Key words: Nonlocal elasticity, Regularization, Nonlocal finite element method.

Abstract. A nonlocal elastic behaviour of integral type is modeled assuming that the nonlocality lies in the constitutive relation. The diffusion processes of the nonlocality are governed by an integral relation containing a recently proposed symmetric spatial weight function expressed in terms of an attenuation function. Starting from the variational formulation associated with the structural boundary-value problem in the context of nonlocal elasticity, a nonlocal finite element model is proposed and a 1D example is proposed.

1 INTRODUCTION

It is well-known that one of the main drawbacks of local elasticity consists in the fact that many problems, such as sharp crack-tip in continuum fracture mechanics, lead to stress singularities in classical elastic theories.

A possible solution consists in considering a continuum approach in which there are information regarding the behaviour of the material microstructure by assuming that an elastic material can transmit information to neighbouring points within a certain distance. Such a distance is the internal length scale and is an essential material parameter which accounts for nonlocal effects in the continuum. Nonlocal variables turn then out to be weighted average of the corresponding local variables over the material points of the structure and the internal length controls the weighting process related to a state variable at a given point.

A continuum theory for elastic material with long range cohesive forces can be found in the pioneristic work of Kröner. A nonlocal elastic theory is presented by Eringen but a simplified and more effective nonlocal theory is contributed in [1] by assuming that nonlocality appears only in the constitutive relation. It is shown that several problems related to stress singularities in local elasticity, such as crack-tip problems, disappear by adopting the nonlocal theory. An elastic model in a geometrically linear range endowed with the nonlocal elastic material model is dealt with in [2] in which the extension to nonlocal linear elasticity of the classical principles of the total potential energy, complementary energy and mixed Hu-Washizu principle are also provided.

In the present paper, starting from the nonlocal elastic constitutive model proposed by Eringen and co-workers, the thermodynamic framework and the boundary-value problem for nonlocal elasticity are formulated and the complete set of nonlocal mixed variational principles is then provided. A recently proposed, in the context of damage mechanics,

symmetric spatial weight function which preserves constant fields is considered. A firm variational basis to the nonlocal model is provided. A consistent symmetric nonlocal finite element procedure is then derived starting from the nonlocal counterpart of the displacement-based variational formulation. A piecewise homogeneous bar is solved by the recourse to the proposed nonlocal finite element method for an imposed displacement and different attenuation functions. The solutions obtained are in a good agreement each other and no pathological behaviours at the boundary are present.

2 NONLOCAL ELASTICITY

The nonlocal elastic model is based on the idea that the long range forces arising in a homogeneous isotropic elastic structure are described by the following constitutive relation [3,4]:

$$\bar{\sigma}(\mathbf{x}) = (R\sigma)(\mathbf{x}) = (RE\boldsymbol{\varepsilon})(\mathbf{x}) = \int_{\Omega} W(\mathbf{x}, \mathbf{y})\mathbf{E}(\boldsymbol{\varepsilon}(\mathbf{y}))d\mathbf{y} \quad (1)$$

The linear regularization operator R transforms the local stress field σ into the related nonlocal stress $\bar{\sigma}$ since its value at the point \mathbf{x} of the body Ω depends on the entire field σ . In linear isotropic elasticity, the elastic operator is $\mathbf{E}=K(\mathbf{1} \otimes \mathbf{1})+2G\mathbf{I}^*$ where K and G denote the bulk and the shear moduli respectively, $\mathbf{I}^*=\mathbf{I}-1/3(\mathbf{1} \otimes \mathbf{1})$ is the fourth-order deviatoric projection tensor being \mathbf{I} and $\mathbf{1}$ the fourth-order and the second-order identity tensors respectively.

From a mechanical standpoint, the space weight function W describes the mutual long-range elastic interaction. The function W is positive, have its maximum for $\mathbf{x}=\mathbf{y}$ and decreases monotonically and rapidly to zero approaching the boundary of the interaction zone. The space weight function W vanishes, or it approaches to zero, for $\|\mathbf{x}-\mathbf{y}\| \geq r$ where r is the chosen influence distance,

A nonlocal behavior is present for high space variation of the local stress σ so that it results $R=I$ for uniform fields σ being I the identity operator. Accordingly the weight function W must fulfill the normalizing condition:

$$\int_{\Omega} W(\mathbf{x}, \mathbf{y})d\mathbf{y} = 1 \quad (2)$$

for any \mathbf{x} in Ω . In order to impose such a condition, also for points close to the boundary of the body in which the interaction zone is deprived of a contribution, the following expression is considered in the sequel:

$$W(\mathbf{x}, \mathbf{y}) = \left[1 - \alpha \frac{V(\mathbf{x})}{V_{\infty}} \right] \delta(\mathbf{x}, \mathbf{y}) + \frac{\alpha}{V_{\infty}} g(\mathbf{x}, \mathbf{y}) \quad (3)$$

which is similar to the one proposed in [5] within the context of nonlocal damage. In the equation (3), V is the representative volume:

$$V(\mathbf{x}) = \int_{\Omega} g(\mathbf{x}, \mathbf{y})d\mathbf{y} , \quad (4)$$

V_{∞} is the value assumed by the representative volume V for an unbounded body, the symbol $\delta(\mathbf{x}, \mathbf{y})$ denotes the Dirac delta distribution and α is an adimensional scalar parameter. The scalar function $g(\mathbf{x}, \mathbf{y})$ is a symmetric attenuation function.

Due to the symmetry of the weight function W , the regularization operator is self-adjoint, i.e. $R=R'$ where R' denotes the dual operator.

3 THERMODYNAMIC FRAMEWORK

Let us analyze the thermodynamic framework for the nonlocal elastic model. The first principle of thermodynamics (see e.g. [6]) for a nonlocal behaviour can be formulated as follows:

$$\int_{\Omega} \dot{e}(\mathbf{x})d\mathbf{x} = \langle \bar{\boldsymbol{\sigma}}, \dot{\boldsymbol{\varepsilon}} \rangle + \int_{\Omega} \dot{Q}(\mathbf{x})d\mathbf{x} \quad (5)$$

where e is the internal energy density depending on strain $\boldsymbol{\varepsilon}$ and entropy s . The heat supplied to an element of volume is $dQ/dt = -\text{div}\mathbf{q}$ being \mathbf{q} the heat flux and “div” is the divergence operator.

The relation (5) can be written pointwise as follows:

$$\dot{e}(\mathbf{x}) = \bar{\boldsymbol{\sigma}}(\mathbf{x}) * \dot{\boldsymbol{\varepsilon}}(\mathbf{x}) + P(\mathbf{x}) + \dot{Q}(\mathbf{x}) \quad (6)$$

where the nonlocality residual function P takes into account the energy exchanges between neighbor particles [7]. The residual P fulfils the insulation condition:

$$\int_{\Omega} P(\mathbf{x})d\mathbf{x} = 0 \quad (7)$$

since the body is a thermodynamically isolated system with reference to energy exchanges due to nonlocality.

The second principle of thermodynamics is enforced in its classical pointwise form:

$$\dot{s}(\mathbf{x})T(\mathbf{x}) + \text{div}\mathbf{q}(\mathbf{x}) - \nabla T(\mathbf{x}) * \frac{\mathbf{q}(\mathbf{x})}{T(\mathbf{x})} \geq 0 \quad (8)$$

everywhere in Ω where ds/dt is the internal entropy production rate per unit volume and T is the absolute temperature. The symbol ∇ denotes the gradient operator.

The thermodynamic laws (6) and (8) yield the non-negative dissipation at a given point of the body:

$$D(\mathbf{x}) = \bar{\boldsymbol{\sigma}}(\mathbf{x}) * \dot{\boldsymbol{\varepsilon}}(\mathbf{x}) - \dot{\phi}(\mathbf{x}) + P(\mathbf{x}) - s(\mathbf{x})\dot{T}(\mathbf{x}) - \nabla T(\mathbf{x}) * \frac{\mathbf{q}(\mathbf{x})}{T(\mathbf{x})} \geq 0 \quad (9)$$

where $\phi = e - sT$ is the free energy.

Considering isothermal processes, the inequality (9) becomes:

$$D(\mathbf{x}) = \bar{\boldsymbol{\sigma}}(\mathbf{x}) * \dot{\boldsymbol{\varepsilon}}(\mathbf{x}) - \dot{\phi}(\mathbf{x}) + P(\mathbf{x}) \geq 0. \quad (10)$$

The free energy function at a point \mathbf{x} of the body Ω is defined according to the relation:

$$\phi(\boldsymbol{\varepsilon}(\mathbf{x})) = \frac{1}{2} (R\mathbf{E}\boldsymbol{\varepsilon})(\mathbf{x}) * \boldsymbol{\varepsilon}(\mathbf{x}). \quad (11)$$

A direct evaluation shows that, for a piecewise homogeneous material, the operators R and \mathbf{E} commute with respect to the scalar product in $L^2(\Omega)$ so that the following equality holds $\langle R\mathbf{E}\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2 \rangle = \langle \boldsymbol{\varepsilon}_1, R\mathbf{E}\boldsymbol{\varepsilon}_2 \rangle$ for any strain $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$.

The body energy dissipation L follows from the integration of (10) to get:

$$L = \langle \bar{\boldsymbol{\sigma}}, \dot{\boldsymbol{\varepsilon}} \rangle - \int_{\Omega} \dot{\phi}(\mathbf{x})d\mathbf{x} \geq 0. \quad (12)$$

Note that the global free energy is the functional of the strain $\boldsymbol{\varepsilon}$ obtained by integrating the specific free energy (11) over the whole domain of the body:

$$\Phi(\boldsymbol{\varepsilon}) = \int_{\Omega} \phi(\boldsymbol{\varepsilon}(\mathbf{x})) d\mathbf{x} = \frac{1}{2} \langle R\mathbf{E}\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon} \rangle \quad (13)$$

and the complementary potential turns then out to be the quadratic functional:

$$\Phi^*(\bar{\boldsymbol{\sigma}}) = \frac{1}{2} \langle \bar{\boldsymbol{\sigma}}, (R\mathbf{E})^{-1} \bar{\boldsymbol{\sigma}} \rangle. \quad (14)$$

Recalling the equality (13), the body energy dissipation (12) becomes:

$$\langle \bar{\boldsymbol{\sigma}}, \dot{\boldsymbol{\varepsilon}} \rangle - \langle R\mathbf{E}\boldsymbol{\varepsilon}, \dot{\boldsymbol{\varepsilon}} \rangle \geq 0. \quad (15)$$

The relation (15) must hold for any admissible deformation mechanism so that, following widely used arguments [6], the following state law is obtained:

$$\bar{\boldsymbol{\sigma}}(\mathbf{x}) = (R\mathbf{E}\boldsymbol{\varepsilon})(\mathbf{x}) = d\phi(\boldsymbol{\varepsilon}(\mathbf{x})). \quad (16)$$

It is then apparent that the relation (15) holds as an equality. Moreover the dissipation (10) can be viewed as the integrand of (15) so that the inequality (10) vanishes according to the reversible nature of the model:

$$D(\mathbf{x}) = \bar{\boldsymbol{\sigma}}(\mathbf{x}) * \dot{\boldsymbol{\varepsilon}}(\mathbf{x}) - \dot{\phi}(\mathbf{x}) + P(\mathbf{x}) = 0 \quad (17)$$

and the explicit expression for the nonlocality residual function at a given point of the body is given by:

$$P(\mathbf{x}) = \dot{\phi}(\mathbf{x}) - \bar{\boldsymbol{\sigma}}(\mathbf{x}) * \dot{\boldsymbol{\varepsilon}}(\mathbf{x}) = \frac{1}{2} (R\mathbf{E}\dot{\boldsymbol{\varepsilon}})(\mathbf{x}) * \boldsymbol{\varepsilon}(\mathbf{x}) - \frac{1}{2} (R\mathbf{E}\boldsymbol{\varepsilon})(\mathbf{x}) * \dot{\boldsymbol{\varepsilon}}(\mathbf{x}). \quad (18)$$

Finally the nonlocal elastic relation can be expressed in the following equivalent forms in terms of the free energy functional Φ and its conjugate Φ^* :

$$\bar{\boldsymbol{\sigma}} = d\Phi(\boldsymbol{\varepsilon}) \Leftrightarrow \boldsymbol{\varepsilon} = d\Phi^*(\bar{\boldsymbol{\sigma}}) \Leftrightarrow \Phi(\boldsymbol{\varepsilon}) + \Phi^*(\bar{\boldsymbol{\sigma}}) = \langle \bar{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon} \rangle \quad (19)$$

where the last equality represents the Fenchel's relation.

4. THE NONLOCAL ELASTIC STRUCTURAL PROBLEM

In order to develop the structural model, it is convenient to formulate the constitutive relations in a global form, i.e. in terms of quantities pertaining to the whole structure. In the sequel such quantities will be referred to as fields. For a continuous model such fields are functionals defined in the domain Ω occupied by the body and belong to suitable functional spaces.

Local subdifferential (or differential) relations, enforced almost everywhere in Ω , can be equivalently expressed in global form by integrating the relevant functions over the domain Ω .

It can be proved that if the local function is convex (concave), the corresponding global one turn out to be convex (concave) in the corresponding fields. An analogous definition holds for any other functional to be defined over the whole body .

Let $\mathbf{u} \in U$ be the displacement field which is square integrable in Ω together with its distributional derivatives up to the order m [8]. Conforming displacement fields fulfill linear constraint conditions and belong to a closed linear subspace $L \subset U$.

The kinematic operator \mathbf{B} is a bounded linear operator from U to the Hilbert space of square integrable strain fields $\boldsymbol{\varepsilon} \in D$.

Denoting by F the subspace of external forces, which is dual of U , the continuous operator

\mathbf{B}' is the equilibrium operator and is dual of \mathbf{B} . Let $\ell = \{\mathbf{t}, \mathbf{b}\}$, belonging to F , be the load functional where \mathbf{t} and \mathbf{b} denote the tractions and the body forces.

In a geometrically linear range, the compatibility condition and the equilibrium equation are given by:

$$\boldsymbol{\varepsilon} = \mathbf{B}(\mathbf{u} + \mathbf{w}), \quad \mathbf{f} = \mathbf{B}'\boldsymbol{\sigma}. \quad (20)$$

The external relation between reactions $\mathbf{r} \in F$ and displacements $\mathbf{u} \in U$ can be given in terms of two conjugate concave functionals Y and Y^* by means of the following equivalent relations:

$$\mathbf{r} \in \partial Y(\mathbf{u}) \Leftrightarrow \mathbf{u} \in \partial Y^*(\mathbf{r}) \Leftrightarrow Y(\mathbf{u}) + Y^*(\mathbf{r}) = \langle \mathbf{r}, \mathbf{u} \rangle \quad (21)$$

where the symbol ∂ denotes the superdifferential of concave functionals [9]. The last equality represents the Fenchel's relation.

In the case of external frictionless bilateral constraints with non-homogeneous boundary conditions, the admissible set of displacements is given by the subspace $L = \mathbf{w} + L_o$ where L_o collects conforming displacements which satisfy the homogeneous boundary conditions. Then the functional Y turns out to be the indicator of L_o , i.e. $Y(\mathbf{u}) = 0$ if $\mathbf{u} - \mathbf{w}$ belongs to L_o and $-\infty$ otherwise.

The relations governing the nonlocal elastic structural problem for a given load history $\ell(t)$ are:

$$\begin{cases} \mathbf{B}'\boldsymbol{\sigma} = \ell + \mathbf{r} \\ \mathbf{B}(\mathbf{u} + \mathbf{w}) = \boldsymbol{\varepsilon} \\ \bar{\boldsymbol{\sigma}} = d\Phi(\boldsymbol{\varepsilon}) \\ \mathbf{u} \in \partial Y^*(\mathbf{r}) \end{cases} \quad (22)$$

The variational formulation in the complete set of the state variables is provided by the next statement.

Proposition 1 - *The set of state variables $(\mathbf{u}, \bar{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon}, \mathbf{r})$ is a solution of the saddle problem:*

$$\min_{\boldsymbol{\varepsilon}} \max_{\mathbf{r}} \text{stat}_{\mathbf{u}, \bar{\boldsymbol{\sigma}}} M(\mathbf{u}, \bar{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon}, \mathbf{r}) \quad (23)$$

where:

$$M(\mathbf{u}, \bar{\boldsymbol{\sigma}}, \boldsymbol{\varepsilon}, \mathbf{r}) = \Phi(\boldsymbol{\varepsilon}) + Y^*(\mathbf{r}) + \langle \bar{\boldsymbol{\sigma}}, \mathbf{B}(\mathbf{u} + \mathbf{w}) - \boldsymbol{\varepsilon} \rangle - \langle \ell + \mathbf{r}, \mathbf{u} \rangle \quad (24)$$

if and only if it is a solution of the nonlocal elastic structural problem (22).

The variational formulation in terms of displacements \mathbf{u} is obtained by enforcing the external constraint relation (22)₄ in terms of the Fenchel's relation (21)₃ and the compatibility condition (22)₂ in the expression of the functional M to get:

Proposition 2 - *The displacement \mathbf{u} is a solution of the convex optimization problem:*

$$\min_{\mathbf{u}} P(\mathbf{u}) \quad (25)$$

where:

$$P(\mathbf{u}) = \Phi(\mathbf{B}(\mathbf{u} + \mathbf{w})) - Y(\mathbf{u}) - \langle \ell, \mathbf{u} \rangle \quad (26)$$

if and only if it is a solution of the nonlocal elastic structural problem (22).

The potential P is the nonlocal counterpart of the classical total potential energy in elasticity.

5. A NONLOCAL FINITE ELEMENT

The nonlocal total potential energy functional P can be adopted in order to develop a finite element procedure for the proposed nonlocal elastic model. Using a conforming finite element discretization, let Ω_e ($e=1, \dots, N$) be the domain decomposition induced by the meshing of the domain Ω . The unknown displacement field $\mathbf{v}(\mathbf{x})$ is given, for each element, in the interpolated form $\mathbf{v}_h^e(\mathbf{x}) = \mathbf{N}_e(\mathbf{x})\mathbf{q}_e$ with $\mathbf{x} \in \Omega_e$ where \mathbf{q}_e is the vector collecting the nodal displacement of the e -th finite element and $\mathbf{N}_e(\mathbf{x})$ is the chosen shape-function matrix.

The conforming displacement field $\mathbf{v}_h = \{\mathbf{v}_h^1, \mathbf{v}_h^2, \dots, \mathbf{v}_h^N\}$ satisfies the homogeneous boundary conditions and the interelement continuity conditions. The rigid-body displacements are ruled out by imposing the conformity requirement. The displacement parameters \mathbf{q}_e can be expressed in terms of nodal parameters \mathbf{q} by means of the standard assembly operator A_e according to the parametric expression $\mathbf{q}_e = A_e \mathbf{q}$. The interpolated counterpart of the nonlocal total potential energy P can be obtained by adding up the contributions of each non-assembly element and imposing the conforming requirement to the interpolating displacement to get:

$$P_h(\mathbf{v}_h) = \frac{1}{2} \langle \mathbf{REB}(\mathbf{v}_h + \mathbf{w}_h), \mathbf{B}(\mathbf{v}_h + \mathbf{w}_h) \rangle - \langle \ell, \mathbf{u} \rangle, \quad (27)$$

with $\mathbf{v}_h \in L_o$. The interpolated nonlocal total potential energy P_h can then be explicitly written as follows:

$$\begin{aligned} P_h(\mathbf{v}_h) &= \frac{1}{2} \sum_{e=1}^N \int_{\Omega_e} \left(1 - \alpha \frac{V(\mathbf{x})}{V_\infty} \right) \mathbf{E}_e \mathbf{B}(\mathbf{v}_h^e + \mathbf{w}_h^e)(\mathbf{x}) * \mathbf{B}(\mathbf{v}_h^e + \mathbf{w}_h^e)(\mathbf{x}) dx + \\ &+ \frac{\alpha}{V_\infty} \sum_{e=1}^N \sum_{m=1}^N \int_{\Omega_e} \int_{\Omega_m} g(\mathbf{x}, \mathbf{y}) \mathbf{E}_m \mathbf{B}(\mathbf{v}_h^m + \mathbf{w}_h^m)(\mathbf{y}) * \mathbf{B}(\mathbf{v}_h^e + \mathbf{w}_h^e)(\mathbf{x}) dy dx + \\ &- \sum_{e=1}^N \int_{\Omega_e} \mathbf{b}(\mathbf{x}) * \mathbf{v}_h^e(\mathbf{x}) dx - \sum_{e=1}^N \int_{S_e} \mathbf{t}(\mathbf{x}) * \Gamma \mathbf{v}_h^e(\mathbf{x}) dx \end{aligned} \quad (28)$$

where $S_e = \partial\Omega \cap \partial\Omega_e$.

The matrix form of the discrete problem is obtained by imposing the stationarity of P_h with respect to \mathbf{v}_h which is given by:

$$\begin{aligned} &\sum_{e=1}^N \int_{\Omega_e} \left(1 - \alpha \frac{V(\mathbf{x})}{V_\infty} \right) \mathbf{E}_e \mathbf{B}(\mathbf{v}_h^e + \mathbf{w}_h^e)(\mathbf{x}) * \mathbf{B} \delta \mathbf{v}_h^e(\mathbf{x}) dx + \\ &+ \frac{\alpha}{V_\infty} \sum_{e=1}^N \sum_{m=1}^N \int_{\Omega_e} \int_{\Omega_m} g(\mathbf{x}, \mathbf{y}) \mathbf{E}_m \mathbf{B}(\mathbf{v}_h^m + \mathbf{w}_h^m)(\mathbf{y}) * \mathbf{B} \delta \mathbf{v}_h^e(\mathbf{x}) dy dx = \\ &= \sum_{e=1}^N \int_{\Omega_e} \mathbf{b}(\mathbf{x}) * \delta \mathbf{v}_h^e(\mathbf{x}) dx + \sum_{e=1}^N \int_{S_e} \mathbf{t}(\mathbf{x}) * \Gamma \delta \mathbf{v}_h^e(\mathbf{x}) dx \end{aligned} \quad (29)$$

for any $\delta \mathbf{v}_h^e \in L_o$.

Defining the component submatrices and subvectors in the form:

$$\begin{aligned}
 \mathbf{K}_{ee}^l &= \int_{\Omega_e} (\mathbf{BN}_e)^T(\mathbf{x})\mathbf{E}_e(\mathbf{BN}_e)(\mathbf{x})d\mathbf{x} \\
 \mathbf{K}_{ee}^{nl} &= -\frac{\alpha}{V_\infty} \int_{\Omega_e} (\mathbf{BN}_e)^T(\mathbf{x})\mathcal{V}(\mathbf{x})\mathbf{E}_e(\mathbf{BN}_e)(\mathbf{x})d\mathbf{x} \\
 \mathbf{K}_{ee} &= \mathbf{K}_{ee}^l + \mathbf{K}_{ee}^{nl} \\
 \mathbf{K}_{em}^{nl} &= \frac{\alpha}{V_\infty} \int_{\Omega_e} \int_{\Omega_m} (\mathbf{BN}_e)^T(\mathbf{x})g(\mathbf{x},\mathbf{y})\mathbf{E}_m(\mathbf{BN}_m)(\mathbf{y})d\mathbf{y}d\mathbf{x}
 \end{aligned} \tag{30}$$

and

$$\begin{aligned}
 \mathbf{f}_e^l &= \int_{\Omega_e} \mathbf{N}_e^T(\mathbf{x})\mathbf{b}(\mathbf{x})d\mathbf{x} + \int_{S_e} \mathbf{N}_e^T(\mathbf{x})\mathbf{t}(\mathbf{x})d\mathbf{x} - \int_{\Omega_e} (\mathbf{BN}_e)^T(\mathbf{x})\mathbf{E}_e(\mathbf{BN}_e)(\mathbf{x})d\mathbf{x}\mathbf{w} \\
 \mathbf{f}_e^{nl} &= \frac{\alpha}{V_\infty} \int_{\Omega_e} (\mathbf{BN}_e)^T(\mathbf{x})\mathcal{V}(\mathbf{x})\mathbf{E}_e(\mathbf{BN}_e)(\mathbf{x})d\mathbf{x}\mathbf{w} + \\
 &\quad - \frac{\alpha}{V_\infty} \int_{\Omega_e} \int_{\Omega_m} (\mathbf{BN}_e)^T(\mathbf{x})g(\mathbf{x},\mathbf{y})\mathbf{E}_m(\mathbf{BN}_m)(\mathbf{y})d\mathbf{y}d\mathbf{x}\mathbf{w}
 \end{aligned} \tag{31}$$

$$\mathbf{f}_e = \mathbf{f}_e^l + \mathbf{f}_e^{nl}$$

the matrix form of the discrete problem is:

$$\sum_{e=1}^N A_e^T \mathbf{K}_{ee} A_e \mathbf{q} + \sum_{e=1}^N \sum_{m=1}^N A_e^T \mathbf{K}_{em}^{nl} A_m \mathbf{q} = \sum_{e=1}^N A_e^T \mathbf{f}_e. \tag{32}$$

The integration appearing in (30)₁ is performed elementwise so that \mathbf{K}_{ee}^l turns out to be the standard stiffness matrix while \mathbf{K}_{ee}^{nl} and \mathbf{K}_{em}^{nl} in (30) turn out to be the nonlocal symmetric stiffness matrices reflecting the nonlocality of the model. The elements of the matrix \mathbf{K}_{em}^{nl} vanish if the related elements are too far with respect to the influence distance r . Accordingly the matrix \mathbf{K}_{em}^{nl} is banded with a band width larger than in the standard stiffness matrix.

Hence the solving linear equation system follows from (29) and is given by:

$$\mathbf{K}\mathbf{q} = (\mathbf{K}^l + \mathbf{K}^{nl})\mathbf{q} = \mathbf{f} \tag{33}$$

where the global stiffness matrix \mathbf{K} is symmetric and positive definite.

In the case of a local elastic behaviour, the nonlocal terms disappear and the solving equation system reduces to the standard local finite element method given by $\mathbf{K}^l\mathbf{q}=\mathbf{f}^l$.

6. A COMPUTATIONAL EXAMPLE

The elastic bar reported in Fig. 1 is solved by the proposed nonlocal finite element approach.



Figure 1: A one-dimensional bar in tension.

The bar has a unit cross-section and a length $L=100\text{ cm}$. It is clamped at the end $x=0$ and is subjected to a given displacement w at the other end $x=L$. The bar is piecewise homogeneous

and the Young modulus has the following expression $E(x) = \beta E_0$ for $0 \leq x \leq L/2$ and $E(x) = E_0 = 21 \times 10^4 \text{ MPa}$ for $L/2 \leq x \leq L$ with the parameter β varying from 0 to 1. Three different weight functions W_1 , W_2 and W_3 of the type (3) are employed in which the attenuation function g is, respectively, given by the Gauss-like function:

$$g_1(\mathbf{x}, \mathbf{y}) = \frac{1}{l\sqrt{2\pi}} \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2l^2}\right) \quad (34)$$

the bi-exponential function:

$$g_2(\mathbf{x}, \mathbf{y}) = \frac{1}{2l} \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|}{2l}\right) \quad (35)$$

and the bell-shaped polynomial function:

$$g_3(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{1}{l\sqrt{2\pi}} \exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2l^2}\right) & \text{if } \|\mathbf{x} - \mathbf{y}\| \leq r \\ 0 & \text{if } \|\mathbf{x} - \mathbf{y}\| > r \end{cases} \quad (36)$$

where it is assumed $l=r/6$. The internal length is $l=2 \text{ cm}$, the influence distance is $r=12 \text{ cm}$ and the material parameter is $\alpha=-1$. The imposed displacement at the end $x=L$ is $w=0.2 \text{ cm}$.

A series of computations have been accomplished by using the above data, the three space weight functions W_1 , W_2 , W_3 and different jumps of the elastic modulus provided by the parameter β .

In the case of a homogeneous material, i.e. for $\beta=1$, strains and stresses coincide to the classical solution for homogeneous media independently of the internal length.

The strain plot ε are provided in Figs. 2 for different values of β .

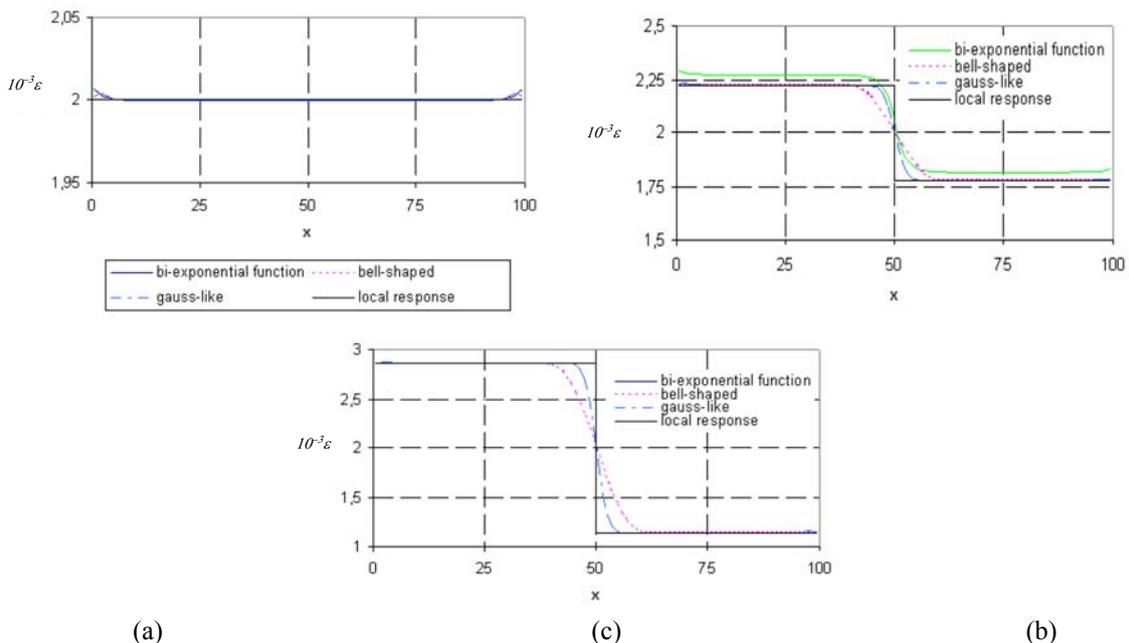


Figure 2: Strain plots of the bar in tension for different attenuation functions g .

The relations (34)-(36) of the attenuation function g are considered in the expressions of

the symmetric function W_1 , W_2 , W_3 and different values of the ratio β are also addressed. In Fig. 2a the homogeneous bar with $E_1 = E_2$ ($\beta = 1$) is considered. In Fig. 2b the piecewise homogeneous bar with $E_1 = 0.8E_o$ and $E_2 = E_o$ ($\beta = 0.8$) is solved and in Fig. 2c the piecewise homogeneous bar with $E_1 = 0.4E_o$ and $E_2 = E_o$ ($\beta = 0.4$) is considered.

As expected, the solution for $\beta=1$ reported in Fig. 2(a) coincides to the local one and the value $\varepsilon = w_b/L = 2 \times 10^{-3}$ is attained independently of the choice of the attenuation function g in the expression of the space weight function W . On comparing the nonlocal behaviour with the local one in the Figs. 2(b) and 2(c) for different values of the ratio β , it is apparent the presence in the nonlocal response of a narrow layer around the middle section of the bar in which the strain ε smoothly varies with more or less slope depending on the considered attenuation function g .

The comparison shows that the use of the bell-shaped and Gauss-like attenuation functions in the expression of the spatial weight function W provides the best fit of the constant strain for different values of β . Moreover, the Gauss-like function presents a narrow layer around the middle section of the bar with a sharper slope than the one corresponding to the bell-shaped attenuation function.

The displacement profiles corresponding to the considered values of β are reported in Fig. 3 in which the discontinuity in the middle section of the bar is apparent in the case of nonhomogeneity.

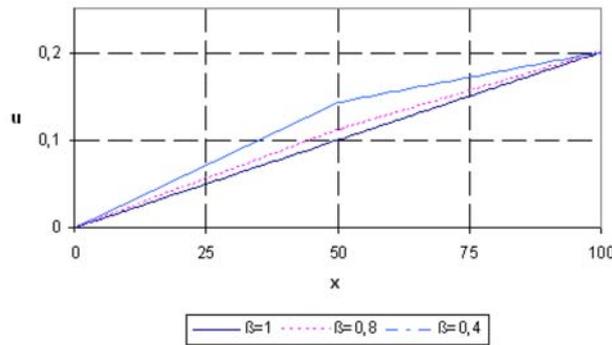


Figure 3: Displacement plots of the bar in tension for different values of the ratio β .

The stress plots are reported in Fig. 4.

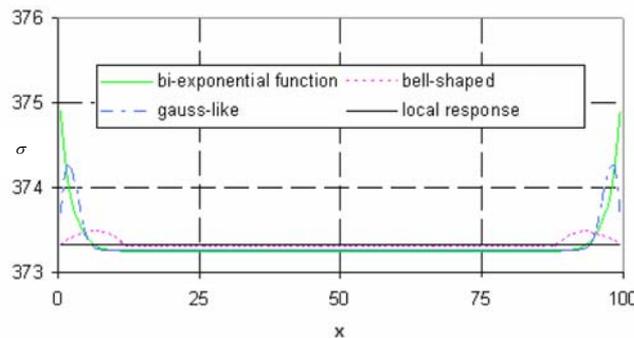


Figure 4: Stress plots of the bar in tension for different attenuation functions g .

The stress is evaluated for a piecewise homogeneous bar with $E_1 = 0.8E_0$ and $E_2 = E_0$ ($\beta = 0.8$) considering the attenuation functions g (34)-(36) in the expression of the symmetric function W .

No boundary effects are present in the stress field in a layer near the end cross-sections.

Acknowledgement: Research support by "Ministero dell'Istruzione, dell'Università e della Ricerca" of Italy is kindly acknowledged.

REFERENCES

- [1] Eringen, A.C. Line crack subjected to shear. *Int. J. Fract.* (1978) **14**:367-379.
- [2] Polizzotto, C. Nonlocal elasticity and related variational principles. *Int. J. Solids Structures* (2001) **38**:7359-7380.
- [3] Eringen, A.C. and Kim, B.S. Stress concentration at the tip of a crack. *Mech. Res. Commun.* (1974) **1**:233-237.
- [4] Eringen, E.C., Speziale C.G. and Kim, B.S. Crack-tip problem in nonlocal elasticity. *J. Mech. Phys. Solids* (1977) **25**:339-355.
- [5] Borino, G., Failla, B. and Parrinello, F. A symmetric nonlocal damage theory. *Int. J. Solids Struct.* (2003) **40**:3621-3645.
- [6] Lemaitre, J. and Chaboche, J.L. *Mechanics of Solids Materials*. Cambridge University Press, Cambridge, UK, (1994).
- [7] Edelen, D.G.B. and Laws, N. On the thermodynamics of systems with nonlocality. *Arch. Rat. Mech. Anal.* (1971) **43**:24-35.
- [8] Showalter, R. E. *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*. American Mathematical Society (1997).
- [9] Rockafellar, R.T. *Convex Analysis*. Princeton, Princeton University Press, (1970).

A Geometric Theory of Plasticity

V. P. Panoskaltsis¹, D. Soldatos¹ and S. P. Triantafyllou²

¹Department of Civil Engineering, Demokritos University of Thrace,
12 Vassilissis Sofias Street, Xanthi, 67100, GREECE
vpanoska@civil.duth.gr

²Institute of Structural Analysis and Aseismic Research, National Technical University of
Athens, Zografou Campus, Athens, 15773, GREECE

Key words: Metric, generalized plasticity, covariance, balance of energy, reversibility, dissipation, Lie derivative, second law of thermodynamics, internal variables

Summary: *A new geometric formulation of rate-independent generalized plasticity is presented. The formulation relies crucially on the consideration of the physical (referential) metric as a primary internal variable and does not invoke any decomposition of the kinematical quantities into elastic and plastic parts. On the basis of a purely geometrical argument the transition to classical plasticity is demonstrated. The covariant balance of energy is systematically employed for the derivation of the mechanical state equations. It is shown that unlike the case of finite elasticity, in finite plasticity, the covariant balance of energy does not yield the Doyle-Ericksen formula, unless a further assumption is made. As an application, a new material model is developed and is tested numerically for the solution of several problems of large scale plastic flow.*

1 Introduction

The kinematics of large deformation plasticity has been for a long time a subject of serious debate. For instance, Nemat-Nasser [1] on the basis of the principle of energy conservation proposed an additive decomposition of the rate of deformation tensor. This decomposition constitutes the basic kinematical assumption adopted in several works of computational interest, such as those of Nagtegaal and De Jong [2], Key and Krieg [3], Atluri [4] and Panoskaltsis et al. [5]. Green and Naghdi [6, 7] considered the additive decomposition of the Lagrangian strain tensor, while Lee and Liu [8] and Lee [9] among others (e.g. Mandel [10], Kratochvil [11], Dashner [12], Lubliner [13, 14], Simo [15], Le and Stumpf [16]) advocated the multiplicative decomposition of the deformation gradient together with the concept of an elastically relaxed (intermediate) configuration. Additional suggestions constitute the combination of both additive and multiplicative theories proposed by Simo and Ortiz [17]. However, as it is pointed out in a review paper by Naghdi [19], almost all authors introduce some measure of plastic strain that can be either a primitive quantity or a solution of an evolution equation, from which the plastic strain

is obtained after suitable initial conditions have been specified. Such an assumption has been subjected to the serious questioning of Gilman [20], among others. The basic objective of this paper is to propose a new formulation of the theory of plasticity, in the general case of finite deformations, which is not based on the introduction of an “artificial” quantity that will stand for plastic strain, but on the *referential metric* which is a natural quantity. The powerful concept of referential metric was introduced by Valanis [22] (who termed it “physical metric”) and it was advanced further by Valanis and Panoskaltsis [23]. The approach has its origins in the realization that since the body is embedded into some ambient space, it is automatically endowed with two distinct metrics, namely the ambient space metric and the body one. These metrics are of different thermomechanical origin and their interrelation in the course of deformation will inevitably specify both the elastic and the plastic (dissipative) properties of the body. As a result, the body (referential) metric is considered as the primary internal variable, while the metric of the ambient space is considered as the control variable, leading naturally to a strain space approach, which offers several advantages in the formulation of an elastic-plastic theory (see, for instance Naghdi [19]). Moreover, in our work we introduce non-Euclidean spaces providing some important advantages, that are not supplied by the classical Euclidean ones. In particular, the present approach can describe several internal material structures, which may differ vastly from the classical Euclidean ones, namely directional densities, curved material structures, pre-formed materials, pre-stressed reference configurations and the presence of dislocation fields, which may change a Euclidean internal structure to a non-Euclidean one (see, Valanis and Panoskaltsis [23]). Second, the introduction of manifold underlying spaces, besides allowing for the extraction of valuable geometrical information, provides the necessary mathematical tools for a covariant formulation of the theory. Generalized plasticity theory (Lubliner [13, 14, 21]) is used in this paper, since it is more general and versatile than classical plasticity. Also, as we have proved in another publication [41], classical plasticity is a particular case of generalized plasticity theory.

The present paper is organized as follows: In section 2 generalized plasticity is presented in a covariant setting. For this purpose, manifold structure is considered not only for the body of interest and the ambient space, but also for the state space, that is the set of all realizable states over a material point. The involvement of the standard pull-back/push-forward operations (e.g. Marsden and Hughes [24, p. 67]), leads to the introduction of the convected Lie derivative (e.g. [24, p. 95], which eventually leads to a covariant formulation of the theory. Loading-unloading criteria, in both the reference and the spatial configurations, are derived as well. The spatial covariant balance of energy, which constitutes the keystone for any covariant theory, relativistic or non-relativistic, is presented in section 3. It is emphasized that even though this concept has been extensively studied within the context of finite elasticity this is not the case within the context of an inelastic theory. Our development shows that, unlike the case of finite elasticity, the covariant balance of energy does not yield the Doyle-Ericksen formula, unless a further assumption is made. As an application, a material model is derived and its predictions for several problems of large scale plastic flow are presented in section 4.

2 Constitutive theory

Since we deal with large scale plastic flow, we follow the geometrical approach proposed within the context of non-linear elasticity by Marsden and Hughes [24]. Accordingly, we consider both the body of interest and the ambient space S as three dimensional Riemannian manifolds with (covariant) metrics \mathbf{G} and \mathbf{g} , respectively. In particular, let Ω be the reference configuration of the body of interest with points labeled by (X^1, X^2, X^3) and define a motion of the body within the ambient space as the time dependent mapping \mathbf{x} :

$$\begin{aligned} \mathbf{x} : \Omega \rightarrow S, \quad x^1 &= x^1(X^1, X^2, X^3, t), \\ x^2 &= x^2(X^1, X^2, X^3, t), \\ x^3 &= x^3(X^1, X^2, X^3, t), \end{aligned} \quad (1)$$

which maps the points of the reference configuration onto the points $\mathbf{x} = \mathbf{x}(\mathbf{X}, t)$ of the current (spatial) configuration.

Then the deformation gradient is defined as the tangent map of (1), i.e.

$$\mathbf{F}(\mathbf{X}, t) = \frac{\partial \mathbf{x}(\mathbf{X}, t)}{\partial \mathbf{X}}, \quad (2)$$

with determinant $J(\mathbf{X}, t) = \det[\mathbf{F}(\mathbf{X}, t)] > 0$. Additionally, we consider the right Cauchy-Green deformation tensor defined as the pull back of the spatial metric \mathbf{g} and the Finger deformation tensor \mathbf{c} defined as the push-forward of the referential metric \mathbf{G} :

$$\mathbf{C} = \mathbf{x}^*(\mathbf{g}) = \mathbf{F}^T \mathbf{g} \mathbf{F}, \quad (3)$$

$$\mathbf{c} = \mathbf{x}_*(\mathbf{G}) = \mathbf{F}^{-T} \mathbf{G} \mathbf{F}^{-T}. \quad (4)$$

A detailed analysis of the geometrical meaning of those quantities can be found in Valanis [22] (see also Yavari et al. [28]).

Generalized plasticity (e.g. Lubliner [13, 14, 21]) is a *local* internal variable theory of rate independent behavior, which is based primarily on the assumption that plastic deformation may take place on loading but not on unloading. The theory may, in principle, be formulated equivalently with respect to the stress or the strain (deformation) space. Since we deal with manifold underlying spaces and their corresponding metrics, a strain space formulation of the theory seems more natural. In turn, in the absence of thermal effects¹ the mechanical state at the referential point \mathbf{X} with coordinates X^1, X^2, X^3 is determined by the control variable, which is identified by the right Cauchy-Green tensor and the internal variable vector. The latter is assumed to be composed by the referential metric \mathbf{G} and an additional internal variable vector \mathbf{Q} . In general, the referential metric is unknown and several considerations must be made for its determination, including experimental procedures (Valanis and Panoskaltsis [23]). To this end, it is emphasized that for the elastic-plastic (dissipative) continuum studied here, the referential metric is a function of the history of deformation (see Valanis [22], Valanis and Panoskaltsis

¹ The extension of the theory to the non-isothermal regime is a non-trivial task and may be done along the lines presented in this work in conjunction with some developments given in Marsden and Hughes [24] and Lubliner [14].

[23]) and its consideration among the internal variables has a concrete physical basis. It is also emphasized that the only one case in which the referential metric is constant in the course of deformation is that of an elastic (non-dissipative) material, like the one discussed in the covariant approach of Marsden and Hughes [24]. The additional internal variable vector may be composed by hardening parameters (e.g. [15, 30]), evolving anisotropy directions (e.g. [29]) and multiple material metrics arising as a physical possibility when changes in the internal structure are due to multiple internal deformation mechanisms [22]. The internal variable vector \mathbf{Q} is assumed to be covariant in the sense that under mapping (1) obeys the general transformation law, that is:

$$q^{j_1 \dots j_r}_{i_1 \dots i_s} = \frac{\partial x^{j_1}}{\partial X^{I_1}} \dots \frac{\partial x^{j_r}}{\partial X^{I_r}} \frac{\partial X^{K_1}}{\partial x^{i_1}} \dots \frac{\partial X^{K_s}}{\partial x^{i_s}} Q^{I_1 \dots I_r}_{K_1 \dots K_s}, \quad (5)$$

where $q^{j_1 \dots j_r}_{i_1 \dots i_s}$ are the components of \mathbf{q} which is the push-forward of \mathbf{Q} , i.e. $\mathbf{q} = \mathbf{x}_* (\mathbf{Q})$.

The state space $\mathcal{D} = (\mathbf{C}, \mathbf{G}, \mathbf{Q})$ is assumed to be attached at the point \mathbf{X} so that the set $\{\mathbf{X}\} \times \mathcal{D}$ is a fiber of \mathbf{X} . A dynamical process \mathbf{P} may be identified by the local vector bundle mapping (e.g. Abraham et al. [27, p. 167]):

$$\mathbf{P} : \Omega \times \mathcal{D} \rightarrow S \times \mathcal{D}',$$

defined as:

$$\mathbf{P}((\mathbf{X}, t), \mathbf{C}, \mathbf{G}, \mathbf{Q}) = ((\mathbf{x}, t), \mathbf{x}_*(\mathbf{C}), \mathbf{x}_*(\mathbf{G}), \mathbf{x}_*(\mathbf{Q})) = ((\mathbf{x}, t), \mathbf{g}, \mathbf{c}, \mathbf{q}). \quad (6)$$

Accordingly, the material state in the current configuration at the point \mathbf{x} with coordinates x^1, x^2, x^3 is determined by the spatial metric \mathbf{g} and the internal variable vector (\mathbf{c}, \mathbf{q}) . A *local process* Ψ in the state space \mathcal{D} is defined as a curve in \mathcal{D} , i.e. as a mapping:

$$\Psi : I \in \mathbb{R} \rightarrow \mathcal{D}$$

defined as:

$$\Psi(t) = (\mathbf{C}(t), \mathbf{G}(t), \mathbf{Q}(t)).$$

The direction and the speed of such a process is determined by the tangent vector,

$$\dot{\Psi} : \mathcal{D} \rightarrow T\mathcal{D}, \text{ with } \dot{\Psi}(t) = (\dot{\mathbf{C}}(t), \dot{\mathbf{G}}(t), \dot{\mathbf{Q}}(t)),$$

or simply $\dot{\Psi} = (\dot{\mathbf{C}}, \dot{\mathbf{G}}, \dot{\mathbf{Q}})$, where $T\mathcal{D}$ is the tangent space of \mathcal{D} . Since $\dot{\mathbf{C}}$ is always known under deformation control, the components $(\dot{\mathbf{G}}, \dot{\mathbf{Q}})$ have to be determined. The latter are assumed to be given as:

$$\begin{aligned} \dot{\mathbf{G}} &= \mathbf{A}(\mathbf{C}, \mathbf{G}, \mathbf{Q}, \dot{\mathbf{C}}), \\ \dot{\mathbf{Q}} &= \mathbf{B}(\mathbf{C}, \mathbf{G}, \mathbf{Q}, \dot{\mathbf{C}}). \end{aligned} \quad (7)$$

or equivalently,

$$\begin{aligned} L_v \mathbf{c} &= \mathbf{a}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}, L_v \mathbf{g}), \\ L_v \mathbf{q} &= \mathbf{b}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}, L_v \mathbf{g}), \end{aligned} \quad (8)$$

where \mathbf{A} and $\mathbf{B} : T\mathcal{D} \times \mathcal{D} \rightarrow T\mathcal{D}$ and \mathbf{a} and $\mathbf{b} : T\mathcal{D}' \times \mathcal{D}' \rightarrow T\mathcal{D}'$ are vector fields in $T\mathcal{D}$ and $T\mathcal{D}'$, respectively and hence they are considered as tensorial functions of the denoted arguments and $L_v(\cdot)$ stands for the Lie derivative, defined as the convected derivative relative to the current configuration (e.g., [25, 27]). Further, the dependence of the functions \mathbf{a} , \mathbf{b} on the deformation

gradient \mathbf{F} , because of the push-forward operation by which Equation (8) is derived from Equation (7) is noteworthy.

Rate independence implies that Equations (7) and (8) are invariant under a replacement of t by $\chi(t)$ where $\chi(\cdot)$ is any monotonically increasing function. Then a necessary and sufficient condition for rate independence is that \mathbf{A} and \mathbf{B} be homogeneous to the first degree (see [14]), that is for any positive number α :

$$\begin{aligned} \mathbf{A}(\mathbf{C}, \mathbf{G}, \mathbf{Q}, \alpha \cdot \dot{\mathbf{C}}) &= \alpha \cdot \mathbf{A}(\mathbf{C}, \mathbf{G}, \mathbf{Q}, \dot{\mathbf{C}}), \\ \mathbf{B}(\mathbf{C}, \mathbf{G}, \mathbf{Q}, \alpha \cdot \dot{\mathbf{C}}) &= \alpha \cdot \mathbf{B}(\mathbf{C}, \mathbf{G}, \mathbf{Q}, \dot{\mathbf{C}}) \end{aligned} \quad (9)$$

or equivalently in the spatial description:

$$\begin{aligned} \mathbf{a}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \alpha \cdot L_v \mathbf{g}) &= \alpha \cdot \mathbf{a}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}, L_v \mathbf{g}), \\ \mathbf{b}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \alpha \cdot L_v \mathbf{g}) &= \alpha \cdot \mathbf{b}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}, L_v \mathbf{g}). \end{aligned} \quad (10)$$

A local process is defined as elastic if it lies entirely in a six dimensional submanifold of \mathcal{D} defined by $(\mathbf{G}, \mathbf{Q}) = \text{constant}$; otherwise is defined as plastic. The basic concept of generalized plasticity is that of the elastic range of a state (e.g. Pipkin and Rivlin [31], Lucchesi and Podio-Guidugli [32], Bertram [33]). The latter is introduced as a submanifold of the state space, defined as the manifold comprising the values of \mathbf{C} that can be reached by an elastic process from the current state. It is further assumed that the boundary of the elastic range is a five dimensional manifold, the points of which have coordinate neighborhoods, and which is attached to the interior in much the same way a face of a cube is attached to the interior. The latter manifold may be defined as a loading surface at (\mathbf{G}, \mathbf{Q}) (e.g., Lubliner [14], Eisenberg and Phillips [34]). A state within the elastic range may be defined as plastic if it lies on a loading surface and elastic otherwise. Accordingly, the rate equations for the internal variables may be derived on the basis of the defining property of a plastic state and the irreversibility of a process from such a state. Thus, if $(\mathbf{C}, \mathbf{G}, \mathbf{Q})$ is a plastic state and \mathbf{N} is the outward normal to the loading surface in the state under consideration, then a simple form for the functions \mathbf{A} and \mathbf{B} which fulfills both requirements together with the homogeneity conditions is:

$$\begin{aligned} \mathbf{A} &= \Gamma \mathbf{\Lambda}(\mathbf{C}, \mathbf{G}, \mathbf{Q}) \langle \mathbf{N} : \dot{\mathbf{C}} \rangle, \\ \mathbf{B} &= \Gamma \mathbf{M}(\mathbf{C}, \mathbf{G}, \mathbf{Q}) \langle \mathbf{N} : \dot{\mathbf{C}} \rangle, \end{aligned} \quad (11)$$

where Γ is a scalar function of the state variables related to the yielding properties of the continuum and which must be positive at any plastic state and zero at any elastic one and $\langle \cdot \rangle$ stands for the Macauley bracket defined as:

$$\langle x \rangle = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0. \end{cases}$$

Finally, $\mathbf{\Lambda}$ and \mathbf{M} are assumed to be non-vanishing functions of the state variables which are associated with the direction of the plastic flow.

In view of Equations (11) the final form for the rate equations for the internal variables may be stated as:

$$\begin{aligned} \dot{\mathbf{G}} &= \Gamma \mathbf{\Lambda}(\mathbf{C}, \mathbf{G}, \mathbf{Q}) \langle \mathbf{N} : \dot{\mathbf{C}} \rangle, \\ \dot{\mathbf{Q}} &= \Gamma \mathbf{M}(\mathbf{C}, \mathbf{G}, \mathbf{Q}) \langle \mathbf{N} : \dot{\mathbf{C}} \rangle. \end{aligned} \quad (12)$$

From the rate Equations (12) one can derive directly the loading-unloading criteria for generalized plasticity theory as:

$$\left\{ \begin{array}{ll} \Gamma(\mathbf{C}, \mathbf{G}, \mathbf{Q}) = 0 & \text{elastic state,} \\ \Gamma(\mathbf{C}, \mathbf{G}, \mathbf{Q}) \neq 0 & \text{and } \mathbf{N} : \dot{\mathbf{C}} < 0 \text{ elastic process,} \\ & \mathbf{N} : \dot{\mathbf{C}} = 0 \text{ neutral process,} \\ & \mathbf{N} : \dot{\mathbf{C}} > 0 \text{ plastic process.} \end{array} \right. \quad (13)$$

To this end it is emphasized that the loading-unloading criteria play a paramount role for the numerical implementation for a generalized plasticity based model (see Panoskaltis et al. [5, 30]).

The equivalent spatial formulation can be derived either in a similar manner (see [5]), or by performing a push-forward operation to Equations (12) as:

$$\left. \begin{array}{l} L_v \mathbf{c} = \gamma \boldsymbol{\lambda}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) \langle \mathbf{n} : L_v \mathbf{g} \rangle, \\ L_v \mathbf{q} = \gamma \boldsymbol{\mu}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) \langle \mathbf{n} : L_v \mathbf{g} \rangle, \end{array} \right\} \quad (14)$$

where $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$ and \mathbf{n} are the push-forwards of the functions Λ , \mathbf{M} , \mathbf{N} in the current configuration and γ is the expression of Γ in terms of the spatial variables. The loading - unloading criteria follow in a similar manner as:

$$\left\{ \begin{array}{ll} \gamma(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) = 0 & \text{elastic state,} \\ \gamma(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) \neq 0 & \text{and } \mathbf{n} : L_v \mathbf{g} < 0 \text{ elastic process,} \\ & \mathbf{n} : L_v \mathbf{g} = 0 \text{ neutral process,} \\ & \mathbf{n} : L_v \mathbf{g} > 0 \text{ plastic process.} \end{array} \right. \quad (15)$$

The rate Equations (12) or (14) along with the mathematical expression of the loading (hyper) surfaces in the state space, which in general are assumed to be given as single parameter families of the form:

$$\Phi(\mathbf{C}, \mathbf{G}, \mathbf{Q}) = K, \quad (16)$$

or equivalently:

$$\varphi(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) = k, \quad (17)$$

constitute the basic ingredients of rate-independent generalized plasticity. It is concluded that the theory does not employ the concept of the yield surface as a basic ingredient. Nevertheless, the concept of the yield surface may be introduced on the basis of purely geometric arguments and an equivalent approach for classical plasticity may be obtained.

3 The covariant balance of energy and the stress tensor

The keystone element of any covariant constitutive theory, relativistic or non-relativistic, is that of the *covariant balance of energy*. A first approach to this concept, within the context of non-linear elasticity and manifold underlying spaces, is given in Marsden and Hughes [24] (see also Simo and Marsden [26]). In particular, these authors by *postulating the invariance of the energy density under arbitrary spatial diffeomorphisms*, that is spatial transformations which may

change the whole ambient space, derived the conservation and balance laws of classical mechanics, together with the Doyle-Ericksen formula (see Doyle and Ericksen [36]). The concept has been studied further by Yavari et al. [28], where particular emphasis is placed on the transformation properties of the balance of energy under *arbitrary referential diffeomorphisms*; this approach leads to the notion of the *configurational forces*, which are forces acting on the reference configuration. Nevertheless, the concept seems to have passed largely unnoticed within the context of materials with internal variables, which do not appear explicitly in the balance laws (Bertram [33]). Its generalization for the elastic-plastic body studied herein, constitutes the primary objective of this section.

The principle of balance of energy may be stated as an axiom as follows:

Let U be any open subset of Ω and let $\mathbf{P} : U \times \mathcal{D} \rightarrow S \times \mathcal{D}'$ be a fixed dynamical process of the body within the space $S \times \mathcal{D}'$, which is modeled by the local vector bundle map (6). Let $\rho(\mathbf{x}, t)$ be the mass density, $e(\mathbf{x}, t)$ the internal energy function per unit mass, $\mathbf{v}(\mathbf{x}, t)$ the spatial velocity, $\mathbf{b}(\mathbf{x}, t)$ the external body force per unit mass and $\mathbf{t}(\mathbf{x}, t, \mathbf{m})$ the Cauchy traction vector, where \mathbf{m} is the unit normal to the boundary $\partial_{\mathbf{x}}(U)$. Then, for a purely mechanical theory the balance of energy axiom reads as follows:

Axiom (Balance of Energy): The dynamical process satisfies balance of energy if:

$$\frac{d}{dt} \int_{x(U)} \rho \left(e + \frac{1}{2} \langle \mathbf{v}, \mathbf{v} \rangle \right) dv = \int_{x(U)} \rho \langle \mathbf{b}, \mathbf{v} \rangle dv + \int_{\partial_{\mathbf{x}}(U)} \langle \mathbf{t}, \mathbf{m} \rangle da \quad (24)$$

where $\langle \cdot, \cdot \rangle$ stands for the inner product in S and dv , da are the volume and area elements respectively, in the current configuration. The covariant balance of energy axiom *within the context of the elastic-plastic body* in question may be stated as follows:

Axiom (Covariant Balance of Energy): For the fixed dynamical process $\mathbf{P} : U \times \mathcal{D} \rightarrow S \times \mathcal{D}'$ which satisfies the balance of energy axiom, consider an arbitrary superposed spatial diffeomorphism $\xi : (S, \mathcal{D}') \rightarrow (S, \mathcal{D}')$. Postulate that the new dynamical process $\bar{\mathbf{x}} = \xi \circ \mathbf{x}$ satisfies the balance of energy axiom, *provided that the metric \mathbf{g} is replaced by $\xi^* \mathbf{g}$, the Finger deformation tensor \mathbf{c} is replaced by $\xi^* \mathbf{c}$, the additional internal variable vector \mathbf{q} is replaced by $\xi^* \mathbf{q}$ and velocities, forces, accelerations, etc. are transformed according to the standard laws of tensorial calculus (e.g., Marsden and Hughes [24, p. 163]), i.e.:*

$$\frac{d}{dt} \int_{\bar{x}(U)} \bar{\rho} \left(\bar{e} + \frac{1}{2} \langle \bar{\mathbf{v}}, \bar{\mathbf{v}} \rangle \right) d\bar{v} = \int_{\bar{x}(U)} \bar{\rho} \langle \bar{\mathbf{b}}, \bar{\mathbf{v}} \rangle d\bar{v} + \int_{\partial_{\bar{\mathbf{x}}}(U)} \langle \bar{\mathbf{t}}, \bar{\mathbf{m}} \rangle d\bar{a}. \quad (25)$$

We assume that unlike the classical elastic case discussed for instance in Marsden and Hughes [24], where the functions e , ρ , etc. depend solely on the spatial metric, in the case of an inelastic material with internal variables, these functions depend on the internal variables as well. This plausible consideration follows from the fact that the internal variables -even though not controllable- they are measurable (e.g. Maugin [37, p. 277]) and play an equally weighted role to the continuum response.

The basic assumption for postulating Equation (25) lies crucially on the adopted transformation law for the internal energy density, namely:

$$\bar{\mathbf{x}} = \xi(\mathbf{x}), \quad \bar{e}(\bar{\mathbf{x}}, t, \mathbf{g}, \mathbf{c}, \mathbf{q}) = e(\mathbf{x}, t, \xi^* \mathbf{g}, \xi^* \mathbf{c}, \xi^* \mathbf{q}). \quad (26)$$

As it is noted in Simo and Marsden [26], this assumption is rather natural and is based on the fact that referential tensors such as the right Cauchy-Green tensor \mathbf{C} , the referential metric \mathbf{G} and the additional internal variable vector \mathbf{Q} , remain unchanged under the superposed spatial diffeomorphism. In particular, if $\xi : (S, \mathbf{g}, \mathbf{c}, \mathbf{q}) \rightarrow (S, \bar{\mathbf{g}}, \bar{\mathbf{c}}, \bar{\mathbf{q}})$ is a spatial diffeomorphism, the new metric $\bar{\mathbf{g}}$ and the new internal variable vector $(\bar{\mathbf{c}}, \bar{\mathbf{q}})$ must be such that $\bar{\mathbf{C}} = \mathbf{C}$ and $(\bar{\mathbf{G}}, \bar{\mathbf{Q}}) = (\mathbf{G}, \mathbf{Q})$, that is:

$$\bar{\mathbf{g}} = \xi \circ \mathbf{x}_* (\bar{\mathbf{C}}) = \xi \circ \mathbf{x}_* (\mathbf{C}) = \xi_* (\mathbf{g}) \quad (27)$$

and similarly,

$$(\bar{\mathbf{c}}, \bar{\mathbf{q}}) = \xi \circ \mathbf{x}_* [(\bar{\mathbf{G}}, \bar{\mathbf{Q}})] = \xi \circ \mathbf{x}_* [(\mathbf{G}, \mathbf{Q})] = \xi_* (\mathbf{c}, \mathbf{q}). \quad (28)$$

Furthermore, by extending the classical elastic case, one can define the set

$$O_{(\mathbf{g}, \mathbf{c}, \mathbf{q})} = \left\{ (\xi_* (\mathbf{g}), \xi_* (\mathbf{c}, \mathbf{q})) \mid \xi : (S, \mathcal{D}') \rightarrow (S, \mathcal{D}') \text{ is a diffeomorphism} \right\}$$

as the orbit of the state variables.

The covariance axiom may be systematically used in order to find restrictions on constitutive equations of materials and in particular, when deformation is the primary variable, to show how the stress can be derived from the internal energy density. For instance, Doyle and Ericksen [36, p. 57] proved that for an elastic continuum, the Cauchy stress tensor $\boldsymbol{\sigma}$ is given as:

$$\boldsymbol{\sigma} = 2\rho \frac{\partial e}{\partial \mathbf{g}}. \quad (29)$$

For the rate independent continuum, with the rate equations taking the form (14), the procedure is the following: Equation (25) is evaluated at time $t = t_0$ for which

$$\xi \Big|_{t=t_0} = \text{identity and } \mathbf{w} = \frac{\partial \xi}{\partial t} \Big|_{t=t_0},$$

is the velocity of ξ (at $t = t_0$).

Then, by applying a standard procedure (see Marsden and Hughes [24, pp. 166, 167]), which involves the transport theorem, the divergence theorem and the Cauchy tetrahedron, we obtain the equations of conservation of mass, balance of momentum, balance of moment of momentum, together with the additional identity:

$$\int_{x(U)} \left[\rho (\dot{\bar{e}} - e) - \frac{1}{2} \boldsymbol{\sigma} : L_{\mathbf{w}} \mathbf{g} \right] dv = 0. \quad (30)$$

The definition of the Lie derivative yields:

$$\begin{aligned} \dot{\bar{e}} &= \dot{e} + \frac{\partial e}{\partial \mathbf{g}} : \frac{d}{dt} \Big|_{t=t_0} \xi^* (\mathbf{g}) + \frac{\partial e}{\partial \mathbf{c}} : \frac{d}{dt} \Big|_{t=t_0} \xi^* (\mathbf{c}) + \frac{\partial e}{\partial \mathbf{q}} : \frac{d}{dt} \Big|_{t=t_0} \xi^* (\mathbf{q}) = \\ &= \dot{e} + \frac{\partial e}{\partial \mathbf{g}} : L_{\mathbf{w}} \mathbf{g} + \frac{\partial e}{\partial \mathbf{c}} : L_{\mathbf{w}} \mathbf{c} + \frac{\partial e}{\partial \mathbf{q}} : L_{\mathbf{w}} \mathbf{q}. \end{aligned} \quad (31)$$

Substitution of (31) into (30) yields:

$$\int_{x(U)} \left[\rho \frac{\partial e}{\partial \mathbf{g}} : L_{\mathbf{w}} \mathbf{g} + \rho \frac{\partial e}{\partial \mathbf{c}} : L_{\mathbf{w}} \mathbf{c} + \rho \frac{\partial e}{\partial \mathbf{q}} : L_{\mathbf{w}} \mathbf{q} - \frac{1}{2} \boldsymbol{\sigma} : L_{\mathbf{w}} \mathbf{g} \right] dv = 0. \quad (32)$$

Substitution of the rate Equations (12) into Equation (31) yields

$$\int_{x(U)} \left\{ \left[\rho \frac{\partial e}{\partial \mathbf{g}} + \rho \frac{\partial e}{\partial \mathbf{c}} : \boldsymbol{\lambda}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) \mathbf{n} + \rho \frac{\partial e}{\partial \mathbf{q}} : \boldsymbol{\mu}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) \mathbf{n} - \frac{1}{2} \boldsymbol{\sigma} \right] : L_{\mathbf{w}} \mathbf{g} \right\} dv = 0, \quad (33)$$

from which and by noting that $L_{\mathbf{w}} \mathbf{g}$ can be arbitrarily specified, we derive

$$\rho \frac{\partial e}{\partial \mathbf{g}} + \rho \frac{\partial e}{\partial \mathbf{c}} : \boldsymbol{\lambda}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) \mathbf{n} + \rho \frac{\partial e}{\partial \mathbf{q}} : \boldsymbol{\mu}(\mathbf{g}, \mathbf{c}, \mathbf{q}, \mathbf{F}) \mathbf{n} - \frac{1}{2} \boldsymbol{\sigma} = 0. \quad (34)$$

Therefore, the covariance axiom does not yield the Doyle-Ericksen formula, unless a further assumption is made, namely that an unloading process (i.e. a process with $\mathbf{n} : L_{\mathbf{v}} \mathbf{g} \leq 0$) from a plastic state is *quasi-reversible*, which means that in such a process the plastic dissipation, defined as

$$d_p = - \frac{\partial e}{\partial \mathbf{c}} : L_{\mathbf{v}} \mathbf{c} - \frac{\partial e}{\partial \mathbf{q}} : L_{\mathbf{v}} \mathbf{q} \left(= - \frac{\partial e}{\partial \mathbf{G}} : \mathbf{G} - \frac{\partial e}{\partial \mathbf{Q}} : \dot{\mathbf{Q}} \right)$$

vanishes. If this is the case, the Doyle-Ericksen formula can be derived as in the classical elastic case directly from Equation (32).

It is interesting to note that Lubliner [14] arrives at a similar result by working entirely in the reference configuration and on the basis of the second law of thermodynamics, expressed in the form of the Clausius-Planck inequality. In our covariant approach we bypass the second law, focusing on all transformations of a given process, unlike the second law of thermodynamics where we focus on all processes (Marsden and Hughes [24, p. 201]).

4 A model problem

In order to illustrate the application of the presented concepts to the constitutive modeling of solid materials a specific model is developed. The formulation of the model is motivated by classical metal plasticity and in particular it comprises von-Mises loading surfaces with both isotropic and kinematical hardening.

Since we deal with large scale inelastic flow, the kinematics of the problem together with the covariance principle, suggest that a formulation of the model in terms of the spatial metrics and their Lie derivatives is more fundamental. Further, in the current configuration the spatial metric has usually a diagonal form, which makes the computations simpler than those in the reference configuration, where the (Lagrangian) metric \mathbf{C} is fully populated (e.g. see [29]). Accordingly, the stress response is assumed to be hyperelastic, governed by an isotropic strain energy function, which is given in terms of the invariants of the tensor $\mathbf{g}\mathbf{b}$, where \mathbf{b} is the left Cauchy-Green tensor, defined as the push-forward of the reciprocal (contravariant) metric \mathbf{G}^{-1} , i.e. $\mathbf{b} = \mathbf{x}_* (\mathbf{G}^{-1}) = \mathbf{F}\mathbf{G}\mathbf{F}^T$, as:

$$\rho_0 \psi = \lambda \frac{(I_3)^2 - 1}{4} - \frac{\lambda}{2} \ln \sqrt{I_3} + \frac{1}{2} \mu I_3^{-\frac{1}{3}} [I_1 - 3],$$

where λ and μ are Lamé' type of parameters, ρ_0 is the referential density and I_1 and I_3 are the first and the third invariants of the tensor \mathbf{g} .

Then, the Kirchhoff stress tensor $\boldsymbol{\tau}$, ($\boldsymbol{\tau} = J\boldsymbol{\sigma}$) is given by the Doyle-Ericksen formula (see [24, p. 204]) as:

$$\boldsymbol{\tau} = 2\rho_0 \frac{\partial \psi}{\partial \mathbf{g}} = \lambda \frac{I_3 - 1}{2} \mathbf{g}^{-1} + \mu I_3^{-\frac{1}{3}} \text{dev} \mathbf{b}, \quad (35)$$

where $\text{dev}(\cdot)$ stands for the deviatoric operator in the current configuration and is given as:

$$\text{dev}(\cdot) = (\cdot) - \frac{1}{3} [\mathbf{g} : (\cdot)] \mathbf{g}^{-1}.$$

The loading surfaces are assumed to be given by a von-Mises type of expression in the form:

$$\varphi(\boldsymbol{\tau}, \mathbf{g}, a, \mathbf{q}) = \sqrt{(\tau^{ij} - q^{ij})(\tau^{kl} - q^{kl})g_{ik}g_{jl} - \frac{1}{3}(\tau^{kl}g_{kl})^2} - \sqrt{\frac{2}{3}}(\sigma_y + \delta \bar{H}a), \quad (36)$$

where a is a scalar internal variable, that stands for the description of the isotropic hardening of the von-Mises loading surface and \mathbf{q} is a purely deviatoric tensorial internal variable (back-stress), which stands for its kinematic hardening. Finally, σ_y denotes the uniaxial yield stress and δ and \bar{H} are two model parameters related to the hardening properties of the material.

The evolution of the contravariant metric is assumed to be given by a normality flow rule, which resembles the one derived on the basis of the maximum plastic dissipation by Simo [15], within the context of classical multiplicative elastoplasticity:

$$\mu I_3^{-\frac{1}{3}} L_v \mathbf{b} = -2 \frac{1}{\beta} \frac{\langle \varphi \rangle}{|\varphi|} \mathbf{n} \langle \mathbf{n} : L_v \mathbf{g} \rangle, \quad (37)$$

where β is an additional parameter. For the rate equations for the evolution of the hardening variables, motivated from the infinitesimal theory (e.g. see Simo and Hughes [38, p. 90]), we consider the following forms:

$$\dot{a} = \sqrt{\frac{2}{3}} \frac{1}{\beta} \frac{\langle \varphi \rangle}{|\varphi|} \langle \mathbf{n} : L_v \mathbf{g} \rangle, \quad (38)$$

$$L_v \mathbf{q} = \frac{2}{3} (1 - \delta) \bar{H} L_v \mathbf{b}. \quad (39)$$

Finally, the normal vector $\mathbf{n} = \frac{\partial \varphi}{\partial \mathbf{g}}$ to the loading surfaces, after lengthy computations (see Simo [15]), can be found to be:

$$\mathbf{n} = \frac{\partial \varphi}{\partial \mathbf{g}} = \bar{\mu} \left(\mathbf{m} + \frac{\|\text{dev} \boldsymbol{\tau}\|}{\bar{\mu}} \text{dev}[\mathbf{m}^2] \right), \quad (40)$$

where $\bar{\mu} = \mu I_3^{-\frac{1}{3}} I_1$ and $\mathbf{m} = \frac{\text{dev} \boldsymbol{\tau}}{\|\text{dev} \boldsymbol{\tau}\|}$ is the normal vector to the yield surface in the stress space.

An equivalent expression of the model in the reference configuration can be derived in a similar

manner like the one discussed in section 2, by a pull-back operation to the basic equations (e.g. [16, 29, 30]).

The proposed model can be implemented numerically by employing a predictor-corrector scheme like the one proposed within the context of a classical elastic-plastic formulation by Simo and Hughes [38, pp. 311-321]. In sharp contrast to the classical elastic-plastic case, *for the model in question the state variables are not constrained to lie within the manifold enclosed by the yield surface, because of the absence of the concept of the yield surface.* Accordingly, *unlike the classical elastic-plastic case where the governing equations define a unilaterally constrained problem of evolution* (e.g. Simo and Ortiz [17], Simo and Hughes [38, pp. 311-321]), in the present case *the governing equations define a differential system which must obey the loading-unloading conditions stated in Section 2* (see Equations (13), (15)). The resulting system is in general highly non-linear and it can be solved by an iterative technique. Algorithmic details related to the crucial role played by the loading-unloading criteria for the numerical implementation of a generalized plasticity model, within the context of the infinitesimal theory, can be found in Panoskaltsis et al. [39]. Additional details, encompassing several algorithmic forms of the time continuous loading-unloading criteria within the context of large deformation plasticity and algorithmic approximations for the Lie derivatives appearing in the formulation, are given in Panoskaltsis et al. [5, 30].

Due to lack of space one problem of large scale plastic flow is considered, that of finite shear. This problem has been used extensively as a testing problem within the context of large deformation constitutive theory (e.g. [4, 5, 18, 40] and is defined by:

$$x^1 = X^1 + \gamma X^2, \quad x^2 = X^2, \quad x^3 = X^3,$$

where γ is the shearing parameter. The material parameters are similar to those considered by Simo and Hughes [38, p. 326] (where a related problem, i.e. the elastic-plastic upsetting of an axisymmetric billet is examined):

$$\lambda = 833.33 \text{ MPa}, \quad \mu = 384.62 \text{ MPa}, \quad \sigma_y = 1 \text{ MPa}, \quad \bar{H} = 3 \text{ MPa}, \quad \delta = 0.2.$$

In our example both *isotropic and kinematic hardening mechanisms* are considered. The referential metric is assumed initially to be equal to the Euclidean metric.

The stress-deformation curves predicted by the model for different values of the plastic parameter β are shown in Figures 1 and 2. The predicted response corresponds to stresses which increase monotonically with strain and the oscillating response, reported among others by Atluri [4] in a finite shear problem, due to the use of the (corrotational) Jaumann rate, does not appear. The salient feature of a generalized plasticity predicted stress-strain curve, according to which reloading, following unloading from a plastic state, results in plastic behavior before the stress level from which the unloading began (see [30, 35]), is verified.

5 Concluding Remarks

A new approach to large deformation plasticity has been proposed. The approach considers the *referential, "physical" metric as the primary internal variable*, the time derivative of which accounts for the description of the plastic (dissipative) mechanisms within the material.

The proposed approach has several advantages over the classical approaches to large deformation elastic-plastic theory since:

- (1) It is based on a physical quantity namely the physical metric, which can be determined by experimental procedures and does not consider neither the introduction of an “artificial” primary measure accounting for plastic deformation, nor any decomposition of the kinematical quantities in elastic and plastic parts.
- (2) By considering the “physical” metric as the primary internal variable and the spatial metric as the control variable, a natural strain-space formulation of the theory is achieved.
- (3) It can describe several internal structures, which may differ vastly from the classical Euclidean one.
- (4) It constitutes a natural extension of the well established theory of elasticity of Marsden and Hughes [24] to the elastic-plastic range.
- (5) It can be extended naturally to a covariant one.

Furthermore, in the course of this development two additional novel features are presented:

- (1) The derivation of the Doyle-Ericksen formula within the context of an elastic-plastic continuum.
- (2) The derivation of a simple elastic-plastic model within a strain-space formulation and its numerical implementation for the solution of large scale plastic flow problems.

References

- [1] Nemat-Naser, S.: On finite deformation elasto-plasticity. *Int. J. Solids Structures* 18, 857-872 (1982)
- [2] Nagtegaal, J.C., De Jong, J.E.: Some computational aspects of elastic-plastic large strain analysis. *Int. J. Num. Methods Engrg.* 17, 15-41 (1981)
- [3] Key, S.W., Krieg, R.D.: On the numerical implementation of inelastic time dependent and time independent, finite strain constitutive equations in structural mechanics. *Computer Methods Appl. Mech. Engrg.* 33, 439-452 (1982)
- [4] Atluri, S.N.: On constitutive relations at finite strain: Hypo-elasticity and elasto-plasticity with isotropic or kinematic hardening. *Computer Methods Appl. Mech. Engrg.* 43, 137-171 (1984)
- [5] Panoskaltsis, V.P., Polymenakos L.C., Soldatos D. Eulerian structure of generalized plasticity: Theoretical and computational aspects. *ASCE, J. Engrg. Mech.*, 134, 354-361 (2008)
- [6] Green, A. E., Naghdi, P.M.: A general theory of an elastic-plastic continuum. *Arch. Rat. Mech. Anal.* 18, 251-281 (1965)
- [7] Green, A. E., Naghdi, P.M.: A thermodynamic development of elastic-plastic continua. In: Parker, H., Sedov, L.I. (Eds.), *Proc. IUTAM Symp. On Irreversible Aspects of Continuum Mechanics and Transfer of Physical Characteristics in Moving Fluids*, Springer-Verlag, 117–131 (1966)
- [8] Lee, E.H., Liu, D.T.: Finite strain elastic-plastic theory with application to plane-wave analysis. *J. Appl. Phys.* 38, 19-27 (1967)

- [9] Lee, E.H.: Elastic-plastic deformation at finite strains. *J. Appl. Mech.* **36**, 1-6, (1969)
- [10] Mandel, J.: *Plasticité classique et viscoplasticité*. CISM, Udine 1971, Springer-Verlag, Vienna, New York (1972).
- [11] Kratochvil, J.: On finite strain theory of elastic–inelastic materials. *Acta Mech.* **16**, 127-142 (1973).
- [12] Dashner P. A.: Invariance considerations in large strain elasto-plasticity *ASME J. Appl. Mech.* **53**, 55-60 (1986)
- [13] Lubliner, J.: Normality rules in large-deformation plasticity. *Mech. Matls* **5**, 29-34 (1986)
- [14] Lubliner, J.: Non-isothermal generalized plasticity. In *Thermomechanical Couplings in solids*, eds. H. D. Bui and Q. S. Nyugen, 121-133 (1987)
- [15] Simo, J. C.: A Framework for finite strain elastoplasticity based on maximum plastic dissipation and multiplicative decomposition Part I: Continuum formulation. *Computer Methods Appl. Mech. Engrg.* **66**, 199-219 (1988)
- [16] Le, K.H., Stumpf, H.: Constitutive equations for elastoplastic bodies at finite strain: thermodynamic implementation. *Acta Mech.* **100**, 155-170, (1993)
- [17] Simo, J.C., Ortiz, M.: A unified approach to finite deformation elastoplastic analysis based on the use of hyperelastic constitutive equations. *Computer Methods Appl. Mech. Engrg.* **49**, 221-245 (1985)
- [18] Agah-Tehrani, A., Lee, E.H., Mallet, R.L., Onat, E.T.: The theory of elastic-plastic deformation at finite strain with induced anisotropy modeled as combined isotropic-kinematic hardening. *J. Mech. Phys. Solids* **35**, 519-539 (1987)
- [19] Naghdi, P.M.: A critical review of the state of finite plasticity. *Z. Angew. Math. Phys.* **41**, 315-387, (1990)
- [20] Gilman, J.J.: Physical nature of plastic flow and fracture, In Lee, E.H., Symonds, P.S. (Eds), *Plasticity, Proc. 2. Symp. On Naval Structural Mechanics*, Pergamon Press, Oxford, 43-99 (1960)
- [21] Lubliner, J.: An axiomatic model of rate-independent plasticity. *Int. J. Solids Structures* **16**, 709-713 (1980)
- [22] Valanis, K.C.: The concept of physical metric in thermodynamics. *Acta Mech.* **113**, 169-184 (1995)
- [23] Valanis K.C., Panoskaltsis V.P.: Material metric, connectivity and dislocations in continua. *Acta Mech.* **175**, 77-103 (2005)
- [24] Marsden, J. E., Hughes, T. J. R.: *Mathematical foundations of elasticity*, Dover Publications, New York (1994)
- [25] Stumpf, H., Hoppe, U.: The application of tensor algebra on manifolds to nonlinear continuum mechanics - Invited survey article. *Z. Angew. Math. Mech.* **77**, 327-339 (1997)
- [26] Simo, J. C., Marsden, J. E.: On the rotated stress tensor and the material version of the Doyle-Ericksen formula. *Arch. Ration. Mech. Anal.* **86**, 213-231 (1984)
- [27] Abraham, R., Marsden, J.E., Ratiu, T: *Manifolds, tensor analysis and applications*, 2nd ed., Springer-Verlag, New Work (1988)

- [28] Yavari, A., Marsden, J.E., Ortiz, M.: On spatial and material covariant balance laws in elasticity. *J. Math. Phys.* **47**, 1-53 (2006)
- [29] Miehe, C.: A constitutive frame of elastoplasticity at large strains based on the notion of a plastic metric. *Int. J. Solids Structures*. **35**, 3859-3897 (1998)
- [30] Panoskaltsis, V.P., Polymenakos, L.C., Soldatos, D.: On large deformation generalized plasticity. *J. Mech. Matls Struct.* **3**, 441-457 (2007)
- [31] Pipkin, A. C., Rivlin, R. S.: Mechanics of rate-independent materials. *Z. Angew. Math. Physik* **16**, 313-326 (1965)
- [32] Lucchesi, M., Podio-Guidugli, P.: Materials with elastic range: a theory with a view toward applications: Part II *Arch. Ration. Mech. Anal.* **110**, 9-42 (1992)
- [33] Bertram, A.: An alternative approach to finite plasticity based on material isomorphisms. *Int. J. Plasticity* **52**, 353-374 (1998)
- [34] Eisenberg, M.A., Phillips, A.; A theory of plasticity with non-coincident yield and loading surfaces. *Acta Mech.* **11**, 247-260 (1971)
- [35] Lubliner, J.: On loading, yield and quasi-yield hypersurfaces in plasticity theory. *Int. J. Solids Structures* **11**, 1011-1016 (1975)
- [36] Doyle, T.C., Ericksen, J.L.: Nonlinear elasticity. *Advances in Applied Mechanics*, Academic Press, New York (1956)
- [37] Maugin, G.A.: *The thermodynamics of plasticity and fracture* Cambridge University Press (1992)
- [38] Simo, J. C. and Hughes, T. J. R.: *Computational inelasticity*, Springer-Verlag, New York (1997)
- [39] Panoskaltsis, V. P., Bahuguna, S. and Soldatos, D.: A general consistent integration scheme for rate-independent generalized plasticity. In eds. D. R. J. Owen, E. Onate and E. Hinton, COMPLAS, International Conference on Computational Plasticity, Fundamentals and Applications, Barcelona, Spain (1997)
- [40] Haupt, P., Tsakmakis, C.: On kinematic hardening and large plastic deformations. *Int. J. Plasticity* **2**, 279-293 (1986)
- [41] Panoskaltsis, V. P., Soldatos, D., Triantafyllou, S., P.: The concept of physical metric in rate – independent generalized plasticity. *Acta Mech.* in press. Published on line: DOI 10.1007/s00707-010-0417-3 (2011)

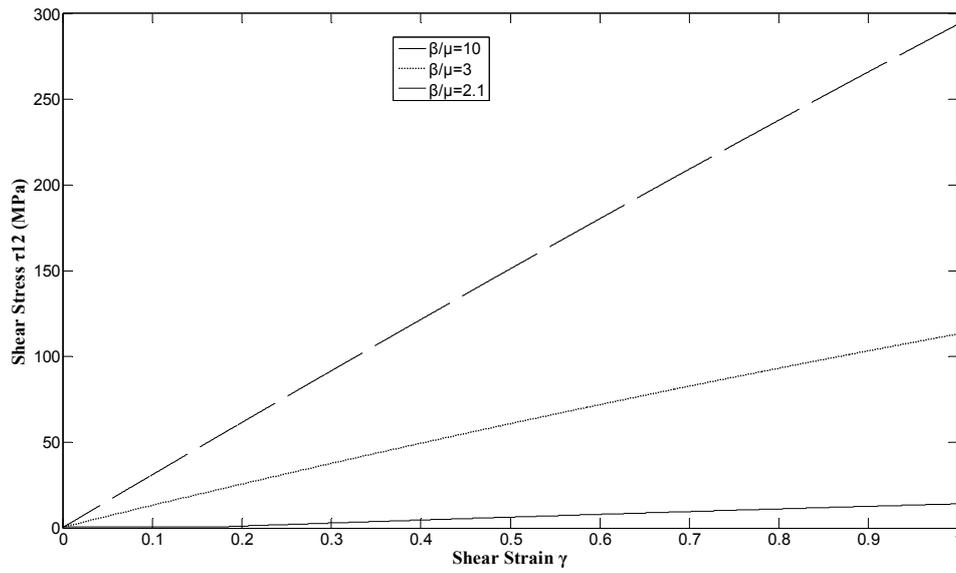


Figure 1: Finite Shear. Shear stress τ_{12} vs. shear strain γ .

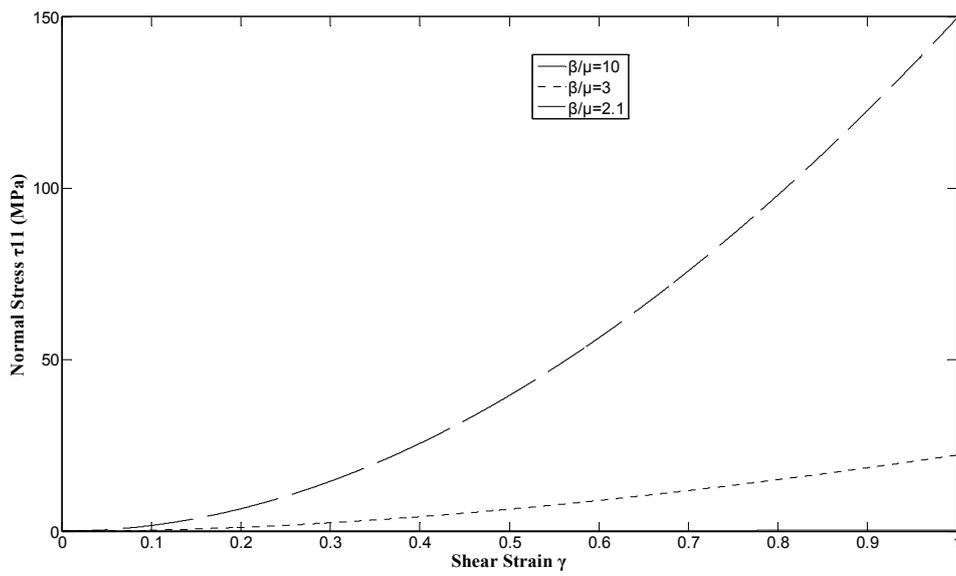


Figure 2: Finite Shear. Normal Stress τ_{11} vs. shear strain γ .

ANALYSIS OF ROCK MASSIF BASED ON THE THEORY OF DAMAGE

JAROSLAV KRUIS* AND TOMÁŠ KOUDELKA*

*Department of Mechanics, Faculty of Civil Engineering
Czech Technical University in Prague
Thákurova 7, 166 29 Prague, Czech Republic
e-mail: jk@cml.fsv.cvut.cz, koudelka@cml.fsv.cvut.cz

Key words: Damage, Heat Transfer, Coupled Problems, Non-linear Systems of Equations, Arc-length Methods

Abstract. This paper deals with analysis of rock massif in connection with a planned deposit of nuclear waste. This problem has emerged with renaissance of nuclear power plants which are now considered as a stable and relatively clean source of electrical energy. The deposits are located deep in rock massif and analysis of stress redistribution caused by cave drifting and temperature changes is required. Difficulties with solution of systems of non-linear algebraic equations were discovered and two variants of the arc-length method were tested.

1 Introduction

Renaissance of nuclear power plants in many countries is accompanied with solution of problem with nuclear waste. One method considered is the excavation of large underground caves where the waste can be stored with the help of very ingenious technology. The artificially made caves are located deep under the surface in rock massif. This topic is in the center of attention of many research group [1, 2, 3].

The final analysis of deposit of nuclear waste has to contain not only mechanical part but also transport processes have to be studied in connection with possible leakage. Therefore, hydro-thermo-mechanical analysis will be required. The main emphasis will be laid on the coupling of damage parameters with material diffusivity and permeability. In this preliminary study, only one way coupling, where the heat transfer influences the mechanical analysis, is taken into account.

Special attention has to be devoted to the solver of non-linear algebraic equations which are obtained by discretisation of the problem solved by the finite element method. The arc-length method is used in this paper because of the softening branch of the stress-strain diagrams. Some difficulties were observed and the classical spherical arc-length method was compared with the linearized method.

Threedimensional model of the whole rock massif with excavated caves is usually very large, i.e. it contains at least hundreds of thousands degrees of freedom. Such large problems are hardly solvable on single processor computers. In fact, there are two possibilities how to solve these problems efficiently. One is based on domain decomposition methods which can be performed on parallel computers [4, 5, 6]. Second possibility is based on adaptive methods where the mesh and the degree of approximation polynomials are changed with respect to error estimates [7].

The paper is organized as follows. Section 2 describes very briefly material models used. Section 3 summarizes variants of the arc-length method. Section 4 deals with numerical results.

2 Material Models

As was mentioned before, this contribution deals with preliminary analysis of rock where the mechanical load cases are accompanied by thermal load. The mechanical analysis is based on the damage theory. The simplest isotropic damage model was not used with respect to different strength in tension and compression. The used orthotropic damage model is described in the next subsection.

2.1 Orthotropic Damage Model

The main drawback of the scalar isotropic damage model is that it uses only one damage parameter for all principle directions regardless of tension or compression. Once the damage parameter caused by exceeding limit strain in one principle direction evolves, it reduces stiffness in all remaining principle directions even though they should not be influenced. This drawback is not significant in the case of the one-dimensional stress state such as pure bending but it becomes more important especially for the three-dimensional stress state.

That led to development of the more advanced damage model which can describe better the 3D problems. In reference [8], the authors proposed general anisotropic model for concrete which contains nine material parameters. The laboratory measurements of the required material parameters has to be performed but it caused difficulties for certain cases. Additionally, the model required a significant number of internal variables that have to be stored. These difficulties led to development of a simplified version of the model which is based on six material parameters - three for tension and another three parameters for compression.

The model is based on the following stress-strain relation

$$\sigma_\alpha = (1 - H(\varepsilon_\alpha)D_\alpha^t - H(-\varepsilon_\alpha)D_\alpha^c)[(3K - 2G)\varepsilon_v + 2G\varepsilon_\alpha], \quad (1)$$

where the index α stands for the index of principle components of the given quantity. The model defines two sets of damage parameters D_α^t and D_α^c for tension and compression, respectively. In the equation (1), the symbol $H()$ denotes the Heaviside function, K is the bulk modulus, G is the shear modulus and ε_v stands for volumetric strain.

There are many evolution laws that can be used for D_t and D_c description. In our problems, the two evolution laws for the damage parameters are used similarly to the laws used in the scalar isotropic damage model. The first law gives better results for compression but the determination of the material parameters is more complicated. It can be written in the form

$$D_\alpha^\beta = \frac{A_\beta(|\varepsilon_\alpha^\beta| - \varepsilon_0^\beta)^{B_\beta}}{1 + A_\beta(|\varepsilon_\alpha^\beta| - \varepsilon_0^\beta)^{B_\beta}}, \quad (2)$$

where β represents indices t or c which are used for tension and compression. A_β , B_β are material parameters controlling the peak value and slope of the softening branch and ε_0^β is the strain threshold. The second law involves correction of the dissipated energy with respect to the size of elements and it describes tension better. It is defined by the non-linear equation (3) which can be solved using the Newton method

$$(1 - D_\alpha^\beta)E|\varepsilon_\alpha^\beta| = f_\beta \exp\left(-\frac{D_\alpha^\beta h |\varepsilon_\alpha^\beta|}{w_0^\beta}\right). \quad (3)$$

In the above equation, f_β represents the tensile or compressive strength and w_0^β controls the initial slope of the softening branches. More details about the implemented models can be found in [9, 10, 11].

2.2 Models for Transport Processes

With respect to limited space, detailed description of material models for transport processes are not included in this paper. They can be found e.g. in references [12, 13, 14]. Efficient computer implementation of transport processes can be found in [15].

3 Arc-length Method

With respect to the softening part of stress–strain diagram, the systems of non-linear algebraic equations are solved by one of arc-length methods. Detailed description of the method can be found in references [16, 17].

The equilibrium condition of a structure after discretization by the finite element method has the form

$$\mathbf{f}_{int}(\mathbf{d}) = \mathbf{f}_c + \lambda \mathbf{f}_p \quad (4)$$

where \mathbf{d} denotes the vector of nodal displacements, \mathbf{f}_{int} denotes the vector of internal forces, \mathbf{f}_c denotes the vector of constant prescribed forces, $\lambda \mathbf{f}_p$ denotes the vector of proportionally changing forces and λ denotes the scalar load-level multiplier. The vector of unbalanced forces has the form

$$\mathbf{r}(\mathbf{d}, \lambda) = \mathbf{f}_c + \lambda \mathbf{f}_p - \mathbf{f}_{int}(\mathbf{d}) \quad (5)$$

and it is the residual. The relationship between the forces \mathbf{f} and the displacements \mathbf{d} has to be obtained by an incremental process where increments of length along the curve \mathbf{f} — \mathbf{d} are used. Within each increment, iteration process is needed in order to attain the equilibrium state. Let the i -th increment be known, i.e. the vector \mathbf{d}_i and the parameter λ_i are known and $\mathbf{r}(\mathbf{d}_i, \lambda_i) = \mathbf{0}$. Expansion of the residual has the form

$$\mathbf{r}(\mathbf{d}_{i+1}, \lambda_{i+1}) = \mathbf{r}(\mathbf{d}_i, \lambda_i) + \frac{\partial \mathbf{r}(\mathbf{d}_i, \lambda_i)}{\partial \mathbf{d}} \delta \mathbf{d}_i + \frac{\partial \mathbf{r}(\mathbf{d}_i, \lambda_i)}{\partial \lambda} \delta \lambda_i = -\mathbf{K}_{i,0} \delta \mathbf{d}_{i,1} + \mathbf{f}_p \delta \lambda_{i,1} = \mathbf{0} \quad (6)$$

Let the vector $\delta \mathbf{d}_{i,1}$ be in the form

$$\delta \mathbf{d}_{i,1} = \delta \lambda_{i,1} \mathbf{v}_{i,1} \quad (7)$$

The length of arc in the first iteration within the increment can be written

$$(\delta \mathbf{d}_{i,1})^T \delta \mathbf{d}_{i,1} + \psi^2 (\delta \lambda_{i,1})^2 \mathbf{f}_p^T \mathbf{f}_p = (\delta \lambda_{i,1})^2 \mathbf{v}_{i,1}^T \mathbf{v}_{i,1} + \psi^2 (\delta \lambda_{i,1})^2 \mathbf{f}_p^T \mathbf{f}_p = (\Delta l)^2 \quad (8)$$

where the scaling parameter ψ was defined. The increment of the scalar load multiplier has the form

$$\delta \lambda_{i,1} = \pm \frac{\Delta l}{\sqrt{\mathbf{v}_{i,1}^T \mathbf{v}_{i,1} + \psi^2 \mathbf{f}_p^T \mathbf{f}_p}} \quad (9)$$

New system of equations has to be solved in second iteration

$$\mathbf{f}_c + (\lambda_i + \delta \lambda_{i,1}) \mathbf{f}_p - \mathbf{f}_{int}(\mathbf{d}_i + \delta \mathbf{d}_{i,1}) - \mathbf{K}_{i,1} \delta \mathbf{d}_{i,2} + \mathbf{f}_p \delta \lambda_{i,2} = \mathbf{0} \quad (10)$$

Cumulative quantities are defined

$$\Delta \mathbf{d}_{i,j} = \Delta \mathbf{d}_{i,j-1} + \delta \mathbf{d}_{i,j} \quad (\Delta \mathbf{d}_{i,1} = \delta \mathbf{d}_{i,1}) \quad (11)$$

$$\Delta \lambda_{i,j} = \Delta \lambda_{i,j-1} + \delta \lambda_{i,j} \quad (\Delta \lambda_{i,1} = \delta \lambda_{i,1}) \quad (12)$$

and they are schematically depicted in Figure 1. Equation (10) can be split into two systems

$$\mathbf{K}_{i,1} \mathbf{u}_{i,2} = \mathbf{f}_c + (\lambda_i + \Delta \lambda_{i,1}) \mathbf{f}_p - \mathbf{f}_{int}(\mathbf{d}_i + \Delta \mathbf{d}_{i,1}) \quad (13)$$

$$\mathbf{K}_{i,1} \mathbf{v}_{i,2} = \mathbf{f}_p \quad (14)$$

where the previously defined notation is used. The length of arc has now the form

$$\|\Delta \mathbf{d}_{i,1} + \mathbf{u}_{i,2} + \delta \lambda_{i,2} \mathbf{v}_{i,2}\|^2 + \psi^2 \|\Delta \lambda_{i,1} \mathbf{f}_p + \delta \lambda_{i,2} \mathbf{f}_p\|^2 = (\Delta l)^2 \quad (15)$$

which is the quadratic equation

$$a_1 (\delta \lambda_{i,2})^2 + a_2 (\delta \lambda_{i,2}) + a_3 = 0 \quad (16)$$

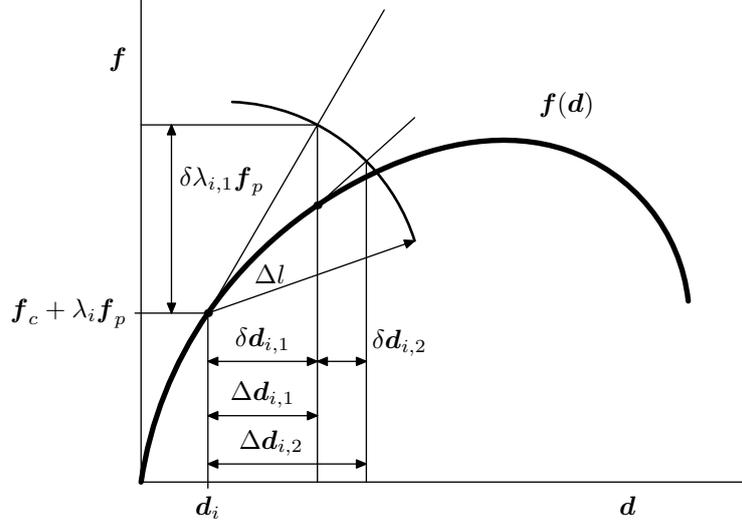


Figure 1: Non-linear force–displacement relationship.

with coefficients

$$a_1 = \mathbf{v}_{i,2}^T \mathbf{v}_{i,2} + \psi^2 \mathbf{f}_p^T \mathbf{f}_p \quad (17)$$

$$a_2 = 2\mathbf{v}_{i,2}^T (\Delta \mathbf{d}_{i,1} + \mathbf{u}_{i,2}) + 2\Delta \lambda_{i,1} \psi^2 \mathbf{f}_p^T \mathbf{f}_p \quad (18)$$

$$a_3 = (\Delta \mathbf{d}_{i,1} + \mathbf{u}_{i,2})^T (\Delta \mathbf{d}_{i,1} + \mathbf{u}_{i,2}) + (\Delta \lambda_{i,1})^2 \psi^2 \mathbf{f}_p^T \mathbf{f}_p - (\Delta l)^2 \quad (19)$$

The increment $\delta \lambda_{i,2}$ is obtained from the quadratic equation (16). New values are again substituted to the residual and equality to the zero vector is checked. The algorithm is summarized in Table 1 and it is called the spherical arc-length method. If the scaling parameter ψ is equal to zero, the method is called the cylindrical arc-length method.

Solution of the quadratic equation (16) is straightforward but only one root has to be used for next computation. One of the criteria used has the form

$$\cos \theta = \frac{\Delta \mathbf{d}_{i,j+1}^T \Delta \mathbf{d}_{i,j}}{(\Delta l)^2} \rightarrow \max \quad (20)$$

Substitution of (11) and (12) leads to the form

$$\cos \theta = \frac{1}{(\Delta l)^2} \Delta \mathbf{d}_{i,j}^T (\Delta \mathbf{d}_{i,j} + \mathbf{u}_{i,j+1} + \delta \lambda_{i,j+1} \mathbf{v}_{i,j+1}) \quad (21)$$

Both roots of the equation (16) are substituted to the expression (21) and the root leading to the larger value is selected.

Linearized form of the arc-length leads to the expression

$$\delta \lambda_{i,j+1} = \frac{-\frac{1}{2} l_{i,j} - \Delta \mathbf{d}_{i,j}^T \mathbf{u}_{i,j+1}}{\Delta \mathbf{d}_{i,j}^T \mathbf{v}_{i,j+1} + \psi^2 \Delta \lambda_{i,j} \mathbf{f}_p^T \mathbf{f}_p} \quad (22)$$

and no root selection procedure is needed.

Table 1: Algorithm of the Arc-length Method

$\lambda_0 = 0, \mathbf{d}_0 = \mathbf{0}$ For $i = 0, 1, 2, \dots$ $\Delta\lambda_{i,0} = 0, \Delta\mathbf{d}_{i,0} = \mathbf{0}, \mathbf{r}_{i,0} = \mathbf{0}$ For $j = 0, 1, 2, \dots$ $\mathbf{u}_{i,j+1} = \mathbf{K}_{i,j}^{-1} \mathbf{r}_{i,j}$ $\mathbf{v}_{i,j+1} = \mathbf{K}_{i,j}^{-1} \mathbf{f}_p$ $a_1 = \mathbf{v}_{i,j+1}^T \mathbf{v}_{i,j+1} + \psi^2 \mathbf{f}_p^T \mathbf{f}_p$ $a_2 = 2\mathbf{v}_{i,j+1}^T (\Delta\mathbf{d}_{i,j} + \mathbf{u}_{i,j+1}) + 2\Delta\lambda_{i,j} \psi^2 \mathbf{f}_p^T \mathbf{f}_p$ $a_3 = \ \Delta\mathbf{d}_{i,j} + \mathbf{u}_{i,j+1}\ ^2 + (\Delta\lambda_{i,j})^2 \psi^2 \mathbf{f}_p^T \mathbf{f}_p - (\Delta l)^2$ $a_1 (\delta\lambda_{i,j+1})^2 + a_2 (\delta\lambda_{i,j+1}) + a_3 = 0 \Rightarrow \delta\lambda_{i,j+1}$ $\delta\mathbf{d}_{i,j+1} = \mathbf{u}_{i,j+1} + \delta\lambda_{i,j+1} \mathbf{v}_{i,j+1}$ $\Delta\mathbf{d}_{i,j+1} = \Delta\mathbf{d}_{i,j} + \delta\mathbf{d}_{i,j+1}$ $\Delta\lambda_{i,j+1} = \Delta\lambda_{i,j} + \delta\lambda_{i,j+1}$ $\mathbf{r}_{i,j+1} = \mathbf{f}_c + (\lambda_i + \Delta\lambda_{i,j}) \mathbf{f}_p - \mathbf{f}_{int}(\mathbf{d}_i + \Delta\mathbf{d}_{i,j})$ if $\ \mathbf{r}_{i,j+1}\ < \varepsilon$, stop $\lambda_{i+1} = \lambda_i + \Delta\lambda_i$ $\mathbf{d}_{i+1} = \mathbf{d}_i + \Delta\mathbf{d}_i$
--

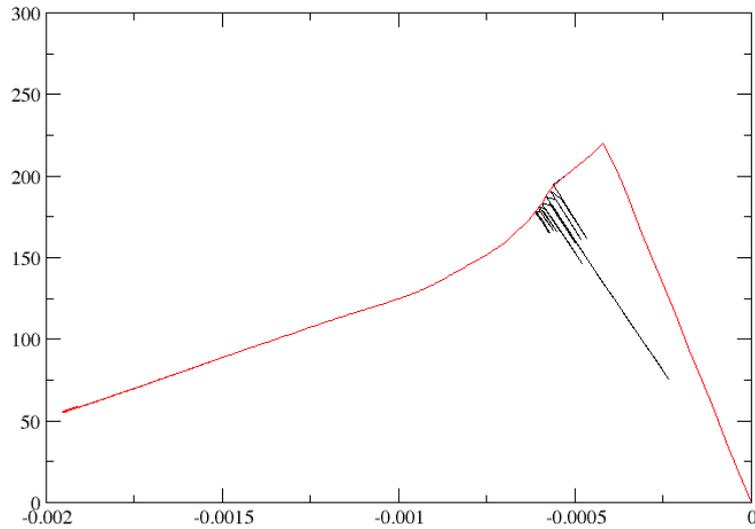


Figure 2: Difficulties with non-linear solvers in softening branch.

4 Numerical results

Two problems were numerically modelled in this preliminary study. First problem was represented by a compression of cylinder which was used for material parameter identification. Cylindrical rock samples were tested in laboratory [1] and the Young modulus, stress-strain relationships and the compressive strength were obtained. Difficulties were observed during the numerical analysis in softening branch of stress-strain relationship. The spherical arc-length method converges very slowly while the linearized arc-length method exhibits incorrect unloading–loading cycles. The situation is depicted in Figure 2 where the red curve is obtained by the spherical method and the black curve by the linearized method.

Second problem was represented by a rock massif with excavated caves. Two different caves were taken into account. Distribution of the damage parameter is depicted for both cases in Figure 3.

5 Conclusions

Analysis of rock massif based on the orthotropic damage material model was performed. Load cases were represented by the dead load and thermal load. Identification of material parameters was done with the help of laboratory tests. Difficulties with solution of non-linear systems of equations were observed.

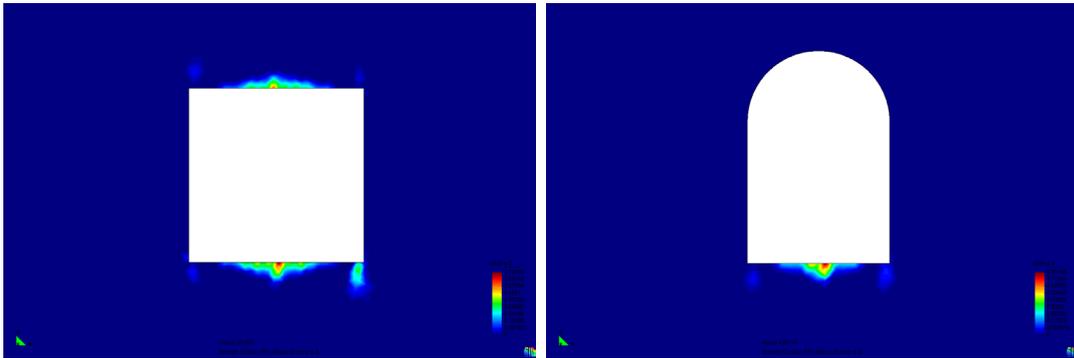


Figure 3: Distribution of the damage parameter in the vicinity of caves.

Acknowledgement:

Financial support for this work was provided by project number P105/10/1682 of Czech Science Foundation. The financial support is gratefully acknowledged.

REFERENCES

- [1] Staub, I. and Andersson, J.C. *Äspö Pillar Stability Experiment*. Swedish Nuclear Fuel and Waste Management Co., Sweden, (2004).
- [2] Blaheta, R. and Byczanski, P. and Kohut, R. and Kolcun, A. and Šňupárek, R. Large-scale modelling of T-M phenomena from underground deposition of the spent nuclear fuel. In *EUROCK 2005: Impact of Human Activity on the Geological Environment*, pages 49-55, (2005).
- [3] Blaheta, R. and Byczanski, P. and Jakl, O. and Kohut, R. and Kolcun, A. and Krečmer, K. and Starý, J. Large scale parallel FEM computations of far/near stress field changes in rocks. *Future Generation Computer Systems-The International Journal of Grid Computing Theory Methods and Applications*, (2006) **22**:449–459.
- [4] Kruis, J. *Domain Decomposition Methods for Distributed Computing*. Saxe-Coburg Publications, Kippen, Stirling, Scotland, (2006).
- [5] Kruis, J. *Domain Decomposition Methods on Parallel Computers*. In Topping, B. H. V. and Mota Soares, C. A., editors, *Progress in Engineering Computational Technology*, pages 299–322, Saxe-Coburg Publications, Stirling, Scotland, UK, (2004).
- [6] Kruis, J. The FETI Method and its Applications: A Review. In Topping, B.H.V. and Iványi, P., editors, *Parallel, Distributed and Grid Computing for Engineering*, pages 199-216, Saxe-Coburg Publications, Stirling, Scotland, UK, (2009).
- [7] Šolín, P. and Segeth, K. and Doležel, I. *Higher-Order Finite Element Methods*. Chapman & Hall/CRC, USA, (2004).

- [8] Papa, E. and Taliercio, A. Anisotropic Damage Model for the Multiaxial Static and Fatigue Behaviour of Plain Concrete. *Eng. Fract. Mech.* (1996) V. 55, N. 2:163–179.
- [9] Koudelka, T. and Krejčí, T. An Anisotropic Damage Model for Concrete in Coupled Problems. In Topping, B.H.V. and Papadrakakis, M., editors, *Proceedings of the Ninth International Conference on Computational Structures Technology*, Civil-Comp Press, Stirlingshire, UK, (2008), paper 157.
- [10] Krejčí, T. and Koudelka, T. and Šejnoha, J. and Zeman, J. Computer Simulation of Concrete Structures subject to Cyclic Temperature Loading. In Topping, B.H.V. and Costa Neves, L.F. and Barros, R.C., editors, *Proceedings of the Twelfth International Conference on Civil, Structural and Environmental Engineering Computing*, Civil-Comp Press, Stirlingshire, UK, (2009), paper 131.
- [11] Koudelka, T. and Krejčí, T. and Šejnoha, J. Analysis of a Nuclear Power Plant Containment. In Topping, B.H.V. and Costa Neves, L.F. and Barros, R.C., editors, *Proceedings of the Twelfth International Conference on Civil, Structural and Environmental Engineering Computing*, Civil-Comp Press, Stirlingshire, UK, (2009), paper 132.
- [12] Schrefler, B.A. and Lewis, R.W. *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media*. 2nd Edition, John Wiley & Sons, (1998).
- [13] Šejnoha, J. and Bittnar, Z. and Krejčí, T. and Kruis, J. Computational Aspects in Thermo-Hydro-Mechanical Analysis of Porous Media, Part I: Transport Processes. In Topping, B. H. V. and Bittnar, Z., editors, *Computational Structures Technology*, pages 153–182, Civil-Comp Press Ltd, Stirling, Scotland, UK, (2002).
- [14] Černý, R. and Maděra, J. and Kočí, J. and Vejmelková, E. Heat and moisture transport in porous materials involving cyclic wetting and drying. In *Fourteenth International Conference on Computational Methods and Experimental Measurements*, pages 3–12, WIT Press, Wessex, UK, (2009).
- [15] Kruis, J. and Koudelka, T. and Krejčí, T. Efficient computer implementation of coupled hydro-thermo-mechanical analysis. *Mathematics and Computers in Simulation* (2010) **80**:1578–1588.
- [16] Crisfield, M.A. *Non-linear Finite Element Analysis of Solids and Structures*. John Wiley & Sons Ltd, Chichester, UK, (1991).
- [17] Bittnar, Z. and Šejnoha, J. *Numerical Methods in Structural Mechanics*. ASCE Press, New York, USA, (1996).

CONTACT-DAMAGE COUPLED MODELLING OF FRP REINFORCEMENTS UNDER VARIABLE LOADING TIMES

G. MAZZUCCO^{*}, V.A. SALOMONI^{*} AND C. MAJORANA^{*}

^{*} Department of Structural and Transportation Engineering
Faculty of Engineering, University of Padua
Via F.Marzolo, 9 – 35131 Padua.

e-mail: mazzucco@dic.unipd.it, salomoni@dic.unipd.it, majorana@dic.unipd.it

Key words: Contact, Damage, FRP Debonding, Creep.

Abstract. In the last years FRP (Fiber Reinforced Polymer) technology has been developed to repair damaged concrete structures. In this work it is proposed to investigate the complex mechanism of stress-strain evolution at the FRP interface, during different loading programs (short or long-time loadings), until complete debonding. This study has been performed by means of a fully three-dimensional approach within the context of damage mechanics, to appropriately catch transversal effects as well as normal stresses, developing a realistic and comprehensive study of the delamination process. The adhesion properties have been reconstructed through a contact model incorporating an elastic-damage constitutive law, relating inter-laminar stresses acting in the sliding direction. A F.E. research code (FRPCON) has been developed, including a numerical procedure accounting for Mazars's damage law inside the contact algorithm. The code is able to describe the delamination process considering the different surface preparation of the concrete part as well. The long-time behaviour of these composite structures has been studied by means of two visco-elastic formulations: i) Bazant's B3 law has been considered for the concrete component, where creep effect is composed by three different terms, i.e. the elastic part, basic creep and drying creep; ii) for FRP's fibres and matrix a micromechanical approach has been implemented. The experimental results of long-time bending tests have been used to calibrate and validate the numerical models.

1 INTRODUCTION

The externally bonded fibre reinforced polymer (FRP) technology has been successfully used in the last years, particularly for recovering structures. In civil engineering practice, this method (generally used for masonry and concrete materials) can be subjected to delamination phenomena: the rupture is localized in the first concrete's layer close to the FRP, named interface zone. In literature several studies of stress field evolution at the interface, obtained in close forms [1, 2] and *via* numerical models [3-6], can be found. Generally, the delamination process in flexural beams is influenced by shear stress concentration at the joint, normally evaluated experimentally by single or double shear tests [7, 8] or bending tests [8].

The contribution of surface preparation in concrete beams (before FRP bonding) to increase strength, as proposed in [9-11], is a new aspect which can be considered for a deeper understanding of the delamination mechanism. Different techniques of surface preparation aim to increase concrete's roughness at the interface zone, allowing for a better bonding with FRP and an increasing strength under ultimate limit loads.

A three-dimensional finite element code has been specifically developed to simulate debonding processes, influenced by surface preparation, by means of a contact-damage algorithm [3, 11] able to represent the entire delamination progression under short or long time applied loads. Transient analyses to simulate long term effects have been carried out considering visco-elastic materials characterized by different compliance functions, for concrete materials (by using B3's theory [12]) and for FRP materials (by using a micromechanical theory developed in [13]) as well. For additional information on long term analyses see [11].

2 THE DELAMINATION PROCESS

Delamination is a progressive detachment of FRP sheets from concrete surface after exceeding the joint strength. Experimental evidences [9, 10] establish that delamination normally affects a thin layer of concrete close to FRP sheets (i.e. at the interface), due to the fact that the maximum shear stress of concrete is smaller than the adhesive's one. From the numerical point of view, the interface can be represented with a physical constant thickness (t_a) known as *adhesive layer* or, as previously stated, *interface zone* [5-8]. In standard practice, concrete surface is subjected to mechanical treatments for enhancing the asperities before FRP bonding, hence increasing adhesion strength (Figure 1 a). After bonding, the real adhesive thickness depending on concrete surface's roughness does not result constant along the joint, hence it appears as more appropriate to assume t_a as the average asperity height (Figure 1 b), so including concrete and adhesive as well.

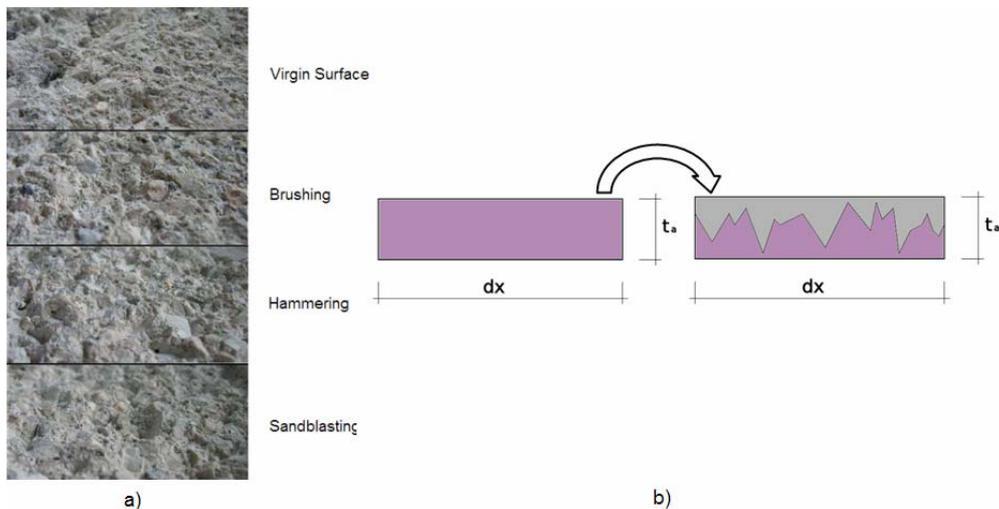


Figure 1: a) Different types of surface treatment; b) interface thickness with asperity inclusions.

If the interface zone is considered as composed by adhesive and concrete (i.e. asperities), necessarily homogenised, the constitutive characteristics can be assumed as

$$\mathbf{D}^C = \mathbf{D}^C(V_c, V_a)$$

$$D_{ij}^C = \frac{D_{c,ij} V_c^{\%} + D_{a,ij} V_a^{\%}}{V_c^{\%} + V_a^{\%}} \quad (1)$$

where V_c is the volume of asperities, V_a the volume of adhesive, $V_k\%$ the volumetric percentage of k-material ($k = a, c$) and $D_{k,ij}$ are the constitutive tensor components of k-material.

3 CONTACT MODEL

FRP-concrete bonding at the interface has been numerically modelled by means of the contact mechanics theory. If considering two bodies, Ω^1 and Ω^2 (Figure 2), e.g. representative of concrete and FRP, two surfaces can be identified, Γ^1 (with $\Gamma^1 \in \Omega^1$, named *slave*) and Γ^2 (with $\Gamma^2 \in \Omega^2$ named *master*), where contact is possible. The *closed contact* condition is achieved and the two bodies are in contact if the contact surface $\Gamma^C = \Gamma^1 \cap \Gamma^2 \neq 0$.

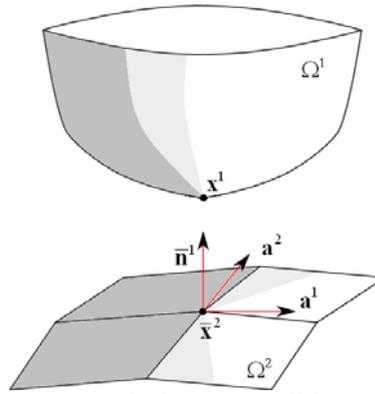


Figure 2: Contact condition.

Contact is defined considering the fundamental conditions [14]:

- Non-penetration conditions

$$\begin{aligned} (\bar{\mathbf{u}}^2 - \mathbf{u}^1) \cdot \mathbf{n}^1 + g &\geq 0 \quad \text{on } \Gamma_c \\ g &= (\bar{\mathbf{X}}^2 - \mathbf{X}^1) \cdot \mathbf{n}^1 \quad \text{on } \Gamma_c \end{aligned} \quad (2)$$

where \mathbf{u}^i (with $i = 1, 2$) are the displacement vectors, \mathbf{X}^i the vectors' position in the reference configuration, g the gap function (the distance between two points in contact) and \mathbf{n} the normal vector (Figure 2).

- Action-reaction conditions

$$\mathbf{t}^2 + \bar{\mathbf{t}}^1 = 0 \quad \text{on } \Gamma_c \quad (3)$$

where \mathbf{t}^i are the stress vectors.

- Kuhn-Tucker conditions

$$\left[\mathbf{t} - (\mathbf{D}^C : \boldsymbol{\varepsilon}) \mathbf{n}^1 \right] \left[(\bar{\mathbf{x}}^2 - \mathbf{x}^1) \cdot \mathbf{n}^1 \right] = 0 \quad \text{on } \Gamma_c \quad (4)$$

where \mathbf{D}^C is the constitutive tensor and $\boldsymbol{\varepsilon}$ the strain tensor of the interface (which has a physical volume, as explained in the previous Section).

In the developed three-dimensional numerical code, characterized by quadratic brick elements (20 nodes), master and slave surfaces have been defined by the faces of brick elements (Figure 3). The closed contact condition have been considered in the *contact pair*, defined through a slave node and a master point where the gap function g is evaluated.

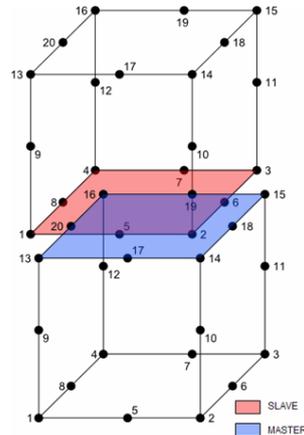


Figure 3: Master and slave surfaces in brick elements.

The master point $\bar{\mathbf{x}}^2$ in a contact pair is chosen as the point with a minimum distance from the slave node \mathbf{x}^1 ; generally it does not coincide with a master *node* but with a generic point belonging to Γ^2 . If the gap function value is less than a minimum distance, the contact is defined as closed. To describe concrete-FRP adhesion by means of a contact algorithm, the minimum distance for considering closed contact has been assumed equal to the asperity height t_a .

To associate stress and strain tensors to the interface zone at each slave node (and for every contact pair) an *element* with volume $\Delta V = \Delta x \cdot \Delta y \cdot t_a$ has been considered, where the base $S = \Delta x \cdot \Delta y$ is geometrically defined by mesh discretization (Figure 4 a) and the volume's height has been assumed equal to t_a (Figure 4 b).

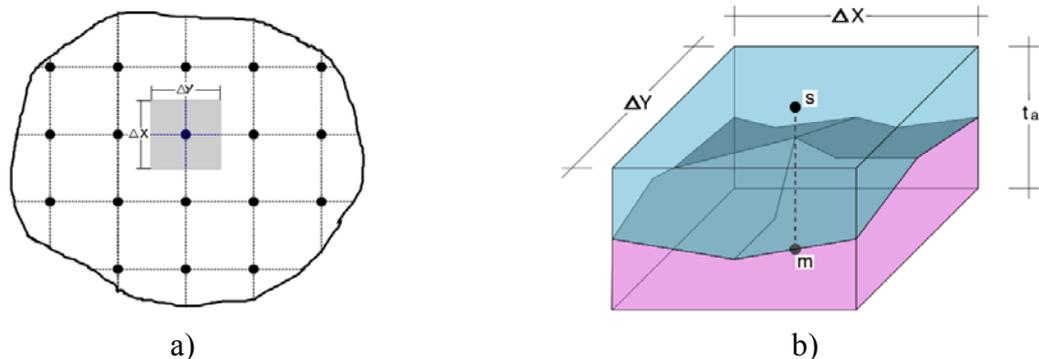


Figure 4: a) Reference base at the interface in the discretized slave surface; b) typical *interface volume* for every contact pair (concrete plus adhesive).

The strain tensor has been evaluated by considering linear displacement variations inside

ΔV [2]; the stress field at the interface between FRP and concrete has been defined once slave and master displacement components for every contact pair ($u^s, v^s, w^s, u^m, v^m, w^m$), as well as the constitutive tensor \mathbf{D}^C , had been known

$$\boldsymbol{\sigma} = \mathbf{D}^C : \boldsymbol{\varepsilon} \quad (5)$$

4 DAMAGE MODEL

A contact procedure alone is not sufficient to simulate the stress-strain evolution at the contact zone during delamination processes. Considering that the detachment between FRP and concrete occurs in the first concrete layer closed to the joint, to represent the loss of adhesion the Mazars' damage law [15] (not reported here for sake of brevity) has been associated to the contact algorithm. In this way bonding/debonding phenomena are driven by the evolution of damage at the interface.

Being the mechanical characteristics of the interface zone during delamination dependant on damage variable d , the tensor \mathbf{D}^C is correspondingly modified by damage, $\mathbf{D}^C = \mathbf{D}^C(d)$. Hence delamination occurs if, during loading, the damage variable assumes a unit value at a *contact pair*: the contact is consequently open.

5 MODELS AND RESULTS

Numerical models have been carried out to first calibrate and then validate the procedure with available experimental results [9]. The setup of the single shear test (Figure 5) consisted in one concrete prism $100 \times 100 \times 300 \text{ mm}^3$ connected to one 50 mm wide sheet of carbon fiber-reinforced polymer (CFRP).

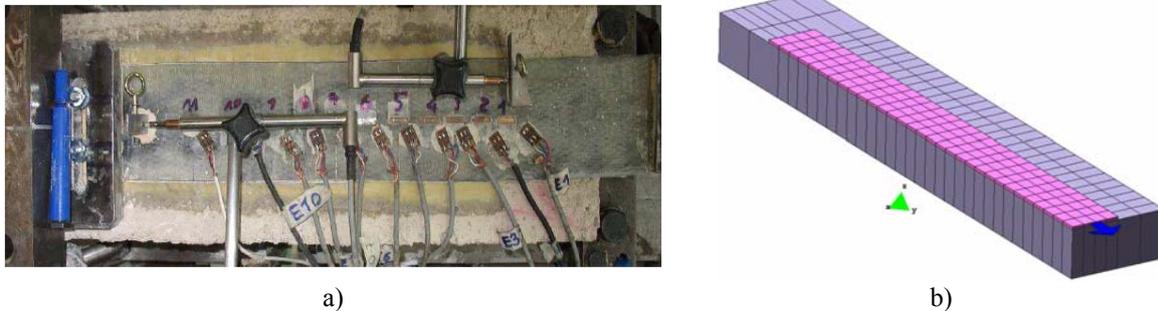


Figure 5: a) Experimental single shear test; b) adopted discretization.

Several types of surface preparation have been considered with different asperity heights (see Table 1).

Under an increasing applied load, the elastic state is overcome at the interface and subsequently delamination starts; the damage evolution of Figure 6a is representative of the delamination process, characterized by a typical shear stress distribution (Figure 6 b).

A comparison between experimental and numerical results in terms of ultimate loads has hence confirmed the correctness of the adopted procedure (Table 1).

Experimental bending tests on concrete beams reinforced by FRP sheets under long time loads [11] have been additionally considered for validating the model. The bending test setup is reported in Figure 7, where $100 \times 100 \times 600 \text{ mm}^3$ concrete prisms have been strengthened by

50×400 mm² carbon fibres sheets (their thickness has been evaluated in 0.165 mm).

The numerical simulations have accounted for a viscous behaviour for the considered materials; specifically, model B3 [12] has been chosen for representing concrete creep and a micromechanical model [13] for FRP sheets creep as well.

Table 1: Experimental and numerical ultimate limit loads with different surface preparation techniques.

Surface preparation technique	Asperity height t_a [mm]	Experimental Ultimate load [kN]	Numerical Ultimate load [kN]
Sandblasting	2.7	25.31	25.89
Hammering	2.4	23.48	22.98
Brushing	1.6	17.62	15.85

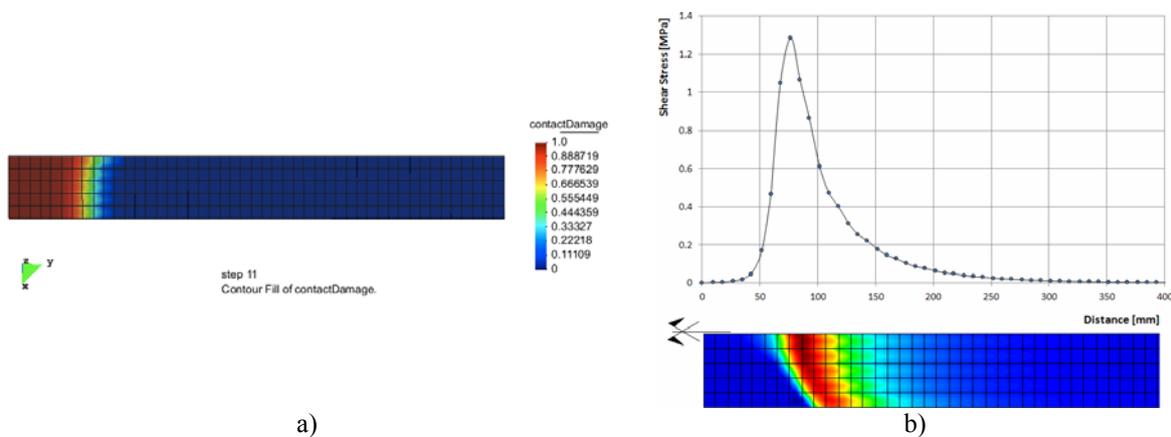


Figure 6: a) Damage evolution during debonding; b) interface shear stresses.

Numerical strains at the joint have been compared with experimental strains evaluated at strain gauges, externally applied to the CFRP reinforcement; Figure 8 refers to strain gauge 1 (applied at 5 mm from the middle of the beam).

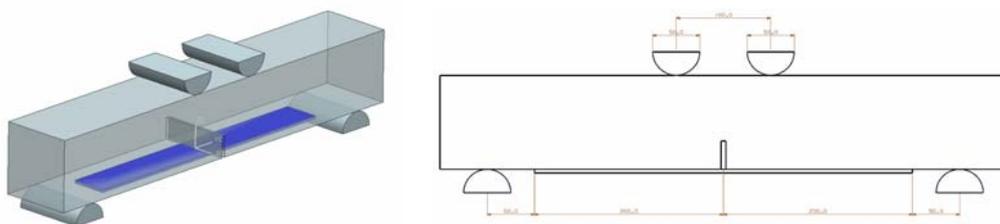


Figure 7: Bending test setup.

6 CONCLUSIONS

Composite beams, made of concrete strengthened with FRP sheets, have been here investigated considering short and long time applied loads, starting from already available experimental evidences and referring to a specifically developed 3D numerical model. The whole system (concrete plus FRP and adhesive) has been represented via three different physical objects: the concrete base, the interface zone composed by adhesive and concrete

asperities and the strengthening bonded FRP strip. The adhesion between layers has been modelled by means of a contact model whose elastic-damage constitutive law relates interlaminar stresses acting in the sliding direction. Long term effects have been studied considering appropriate compliance functions (B3 model for concrete and a micromechanical model for FRP). The research F.E. code has demonstrated to be able to simulate delamination processes and long time stress-strain evolutions. By comparing the numerical results with those of a wide experimental investigation, in terms both of ultimate load and strain vs. time, it has been shown that such an approach is able to catch delamination from a three-dimensional point of view and its evolution during the entire loading process.

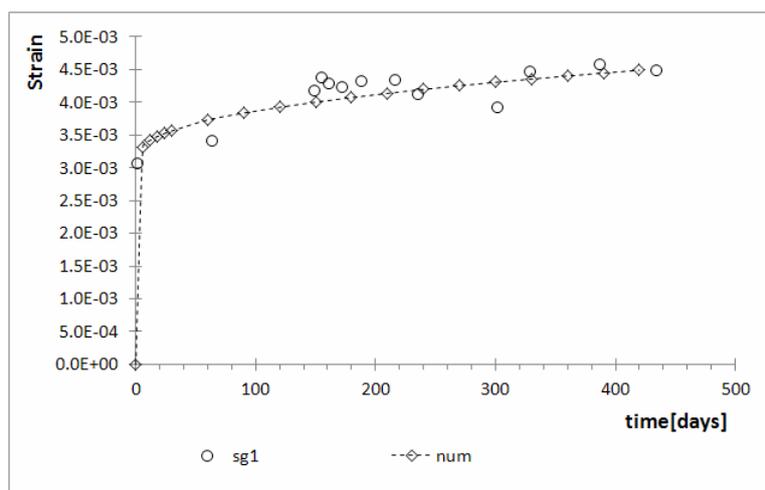


Figure 8: Comparison between experimental and numerical interfacial strains.

REFERENCES

- [1] Liu, Z. and Zhu, B. Analytical solutions for R/C beams strengthened by externally bonded steel plate. *J. of Tongji University* (1994) 167-176.
- [2] Smith, ST. and Teng, JG. Interfacial stresses in plated beams. *Eng. Struct.* (2001) **23**(7):857-71.
- [3] Mazzucco, G., Salomoni, V.A., Pellegrino, C. and Majorana, C.E. Three-dimensional modelling of bond behaviour between concrete and FRP reinforcement. *Eng. Comp.* (2011) **28**(1):5-29.
- [4] Bruno, D. and Greco, F. Mixed mode delamination in plates: a refined approach. *Int. J. Solids Struc.*, (2001) **38**:9149-9177.
- [5] Ascione, F. Ultimate behaviour of adhesively bonded FRP lap joints. *Comp. Part B* (2009) **40**:107-115.
- [6] Coronado, C.A. and Lopez, M.M. Damage Approach for the Prediction of Debonding Failure on Concrete Elements Strengthened with FRP. *J. Comp. Constr.* (2007) **11**(4):391-400.
- [7] Mazzotti, C., Savoia, M. and Ferracuti, B. A new single-shear set-up for stable debonding of FRP-concrete joints. *Constr. Buil. Mat.* (2009) **23**:1529-1537.
- [8] Pellegrino, C. Tinazzi, D. and Modena, C. Experimental Study on Bond Behavior between Concrete and FRP Reinforcement. *J. Comp. Constr.* (2008) **12**(2):180-189.
- [9] Iovinella, I. *Influenza dei trattamenti superficiali sull'efficacia del rinforzo esterno in FRP*. PhD Thesis, Napoli, Italy (2009).
- [10] Toutanji, H. and Ortiz, G. The effect of surface preparation on the bond interface between FRP sheets and concrete members. *Comp. Struc.* (2001) **53**:457-462.
- [11] Mazzucco, G. *Experimental and Numerical analysis of bond behaviour between concrete and*

- FRP*. PhD Thesis, Trento, Italy (2011).
- [12] Bazant, Z. and Baweja, S. Creep and Shrinkage prediction model for analysis and design of concrete structures: Model B3. *Adam Neville Symposium: Creep and Shrinkage – Structural Design Effects, ACI SP-194, A*, Al-Manaseer Ed., Am. Concr. Inst., Farmington Hills, Michigan (2000) 1-83.
- [13] Ascione, L., Berardi, V.P. and D’Aponte, A. Il Comportamento a lungo termine dei materiali compositi fibrorinforzati: un modello micromeccanico. *XXXII Congresso AIAS*, Salerno, Italy (2003).
- [14] Wriggers, P. *Computational Contact Mechanics*. Springer-Verlag, Berlin (2006).
- [15] Mazars, J. A. Description of Micro and Macroscale Damage of Concrete structures. *Eng. Frac. Mech.* (1986) **25**:729-737.

EFFECT OF AGGREGATES AND ITZ ON VISCO-DAMAGED RESPONSE OF CONCRETE AT THE MESO SCALE LEVEL

G. XOTTA, V.A. SALOMONI AND C.E. MAJORANA

Department of Structural and Transportation Engineering
Faculty of Engineering, University of Padua
via F. Marzolo 9 – 35131 Padua, Italy
e-mail: xotta@dic.unipd.it, salomoni@dic.unipd.it, majorana@dic.unipd.it

Key words: meso-scale modeling, model B3, damage, spalling.

Abstract. A deep knowledge on the behavior of concrete materials at the mesoscale level requires, as a fundamental aspect, to characterize aggregates and specifically, their thermal properties if fire hazards (e.g. spalling) are accounted for. The assessment of aggregates performance (and, correspondingly, concrete materials made of aggregates, cement paste and ITZ –interfacial transition zone-) is crucial for defining a realistic structural response as well as damage scenarios. Particularly, it is assumed that concrete creep is associated to cement paste only and that creep obeys to the B3 model proposed by Bazant and Baweja since it shows good compatibility with experimental results and it is properly justified theoretically. The fully coupled 3D F.E. code NEWCON3D has been adopted to perform meso-scale analyses of concrete characterized by aggregates of different types and different thermal conductivities. Damage maps allows for defining an appropriate concrete mixture for responding to spalling and for characterizing the coupled behaviour of ITZ as well.

1 INTRODUCTION

The effect of aggregates on the visco-damaged response of concrete at the meso-scale level is here considered; particularly, model B3 [1, 2] and Mazars' law [3, 4] have been chosen and implemented in the FE code NEWCON3D [5] when considering creep and damage, respectively.

As regards the viscous response, many experiments available in literature have shown that in concrete the source of creep is the cement paste, instead aggregates do not creep in the range of stresses encountered in service. It is therefore reasonable to consider concrete as a composite formed by one aging viscoelastic phase (cement paste) and one elastic phase (aggregate) [6]. Hence model B3, generally adopted to characterize the creep features of concrete, can be successfully used even to model cement paste creep alone [7].

Model B3 has been first validated within NEWCON3D to fit experimental tests at the macro-level [8] and subsequently adopted to perform predictive creep and shrinkage analyses at the meso-level, where concrete is considered as a three-phase composite made of cement paste, aggregate and ITZ.

Additionally, the damaged behavior of concrete at the meso-level has been considered both

to understand the influence of ITZ, the weakest region of the composite material, on the overall mechanical behavior and to define an appropriate concrete mixture for responding to spalling.

2 MODEL B3

Model B3, first developed by Bazant and Baweja in 1995 [1, 2], characterizing concrete creep and shrinkage in the design of concrete structures, shows good compatibility with experimental results and it is better theoretically justified if compared to previous models.

The model distinguishes between Basic Creep (time-dependent deformations where no moisture exchanges with the environment occur) and Drying Creep (additional creep strain accounting for drying).

The compliance function of this model at time t , if a unit uniaxial constant stress is applied at time t' , takes the following form:

$$J(t, t') = q_1 + C_0(t, t') + C_d(t, t', t_0) \quad (1)$$

where q_1 is the instantaneous strain due to a unit stress, $C_0(t, t')$ is the compliance function for basic creep and $C_d(t, t', t_0)$ the additional compliance function due to drying, t the current age, t' the age at loading and t_0 the age at the start of drying.

Specifically, the total basic creep compliance is:

$$C_0(t, t') = q_2 \cdot Q(t, t') + q_3 \cdot \ln[1 + (t - t')^n] + q_4 \cdot \ln\left(\frac{t}{t'}\right) \quad (2)$$

instead the additional creep due to drying is:

$$C_d(t, t', t_0) = q_5 \cdot [\exp\{-\delta H(t)\} - \exp\{-\delta H(t')\}]^{1/2} \quad (3)$$

where $H(t)$ and $H(t')$ are spatial averages of pore relative humidity.

As regards q_1, q_2, q_3, q_4, q_5 parameters, Z.P. Bazant has provided a series of relations based on a statistical survey data of the Data Base by RILEM.

3 MAZARS' DAMAGE MODEL

The damage model considered in NEWCON3D is the non-local Mazars' damage one. The stress-strain law is expressed as:

$$\sigma = A_0 (I - D) \otimes \varepsilon_e \quad (4)$$

where σ and ε_e are stress and strain tensors, D is the damage parameter and A_0 the initial stiffness matrix of the material.

The response of the material takes the following form:

$$f(\varepsilon, A, K_0) = \tilde{\varepsilon} - K(D) \quad (5)$$

where $\tilde{\varepsilon}$ is the equivalent strain and $K(D)$ the hardening/softening parameter, initially equal to K_0 .

Particularly, the response in traction or compression is described by the damage parameters D_t and D_c :

$$D = \alpha_t D_t + \alpha_c D_c, \quad D_t = F_t(\tilde{\varepsilon}) \quad \text{and} \quad D_c = F_c(\tilde{\varepsilon}) \quad (6)$$

$$F_i(\varepsilon_i) = 1 - \frac{(1 - A_i)K_0}{\tilde{\varepsilon}} - \frac{A_i}{\exp[B_i(\tilde{\varepsilon} - K_0)]} \quad (i = t, c) \quad (7)$$

where α_t and α_c are weight coefficients and K_0 , A_i and B_i are parameters that can be determined from experimental tests [9].

Considering a non-local approach, the model computes a variable $\bar{\varepsilon}$:

$$\bar{\varepsilon}(\underline{x}) = \frac{l}{V_r(\underline{x})} \int_V \tilde{\varepsilon}(\underline{s}) \alpha(\underline{s} - \underline{x}) dv \quad (8)$$

where $\tilde{\varepsilon}$ is the equivalent strain, \underline{x} the coordinate of the current Gauss point, \underline{s} the coordinate of the generic Gauss point, α the weight function depending on the characteristic length l and V_r is the characteristic volume. For further explanations the reader is referred to [3, 4, 10, 11].

3 CONCRETE AT THE MESO-SCALE LEVEL

Concrete has a highly heterogeneous microstructure and its composite behavior is exceedingly complex. For obtaining a deeper understanding, theoretical studies based on micromechanics analysis of the interaction between various components of concrete have been developed for deducing its macroscopic constitutive behavior. However, microstructure and properties of the individual components of concrete and their effects on the macroscopic overall response have not been deepened enough; such aspects are now analyzed for simulating concrete behavior from the computational viewpoint [12].

As a composite material, concrete is a mixture of cement paste with aggregates inclusions of different sizes. The components of the heterogeneous material have different properties. The way they react on loading varies too. Variation in stiffness and strength of the components has influence on the global stiffness and fracture behaviour of the material. Different thermal expansion coefficients of the components result in internal stresses (*eigenstresses*) when the global temperature changes. Heat of hydration during hardening and temperature changes during fire are important features as well. Differences in porosity of the components influence the transport and with that also hygral dilation. Different chemical compositions have influence on internal reactions taking place inside the material, which can also be a function of ingress of species [13].

However, concrete is not just a two-phase composite; it has been found that the presence of grains in the paste causes a thin layer of matrix material surrounding each inclusion to be more porous than the bulk of the surrounding cement paste matrix. This layer is called the interfacial transition zone (ITZ), which is known to play an important role in the properties of a concrete composite [14, 15]. The ITZ has a layered structure, a lower density than the bulk matrix and it is more penetrable by fluids and gases [15, 16]; therefore, the ITZ greatly influences the overall permeability of concrete [17]. Additionally, due to its complex structure, the ITZ appears to be the weakest region of the composite material when exposed to external loads. Experiments have demonstrated that the elastic modulus of concrete is strictly related to the elastic modulus and volume fraction of ITZ regions [18, 19]. However, in presence of low w/c ratios and/or fine mineral admixtures (e.g. silica fume), the ITZ may be

absent or difficult to detect. Therefore, the ITZ is not necessarily an intrinsic feature of concrete but depends on factors such as the presence of admixtures, type of mixing, w/c ratio, etc. [20].

As regarding creep features, it is experimentally established that in concrete the only phase subjected to creep is the cement paste; aggregates have essentially an elastic response and cement paste consequently acts as an aging viscoelastic phase. Hence the cement paste is characterized by the aging viscoelastic compliance function $J(t, t')$; particularly, as previously shown, the compliance function adopted here is the one of model B3, being the physical basis of the assumptions used in its derivation also valid for mortar, [6, 7].

4 NUMERICAL ANALYSES.

4.1 Validation of B3 model

To validate B3 model within NEWCON3D, a series of tests by L'Hermite et al. [8] have been taken as reference; specifically, a $7 \times 7 \times 28$ cm³ prism was considered, with the same characteristics as reported in literature (see Table 1) and subjected first to basic creep only (specimen kept in water) and subsequently to drying creep (specimen cured in water; at $t_0 = 2$ days exposed to drying).

Table 1: Parameters used for validation of the model (see [21])

Size of the sample [cm ³]	3.5×3.5×14
Elastic Modulus [MPa]	28522.1
f_{cm28} [MPa]	36.3
Cement Content [kg/m ³]	350
Water Content [kg/m ³]	171.5
Aggregate-cement ratio	4.82
Age at the start of drying t_0 [d]	2
Age of loading t' [d]	7 - 28 - 90
Axial compressive stress [MPa]	9.07
Environmental Relative Humidity	50%
Environmental Temperature [°C]	20

The sample has been additionally loaded by an axial compressive stress of 9.07 MPa; Figure 1 reports the curves of the compliance function (including drying creep), numerically obtained via the code NEWCON3D once model B3 has been implemented. Three different loading times are considered (7, 28 and 90 days) and a comparison with the curves given by Bazant [21] is depicted, showing a good agreement between numerical and experimentally-based results.

Additionally, the spatial averages of pore relative humidity $H(t)$ and $H(t')$ within drying creep have been replaced with the current relative humidity obtained at each time step from the coupled **u**-H-T system of equations; hence, it has been possible to effectively estimate the humidity variation contribution on the creep term (Figure 2).

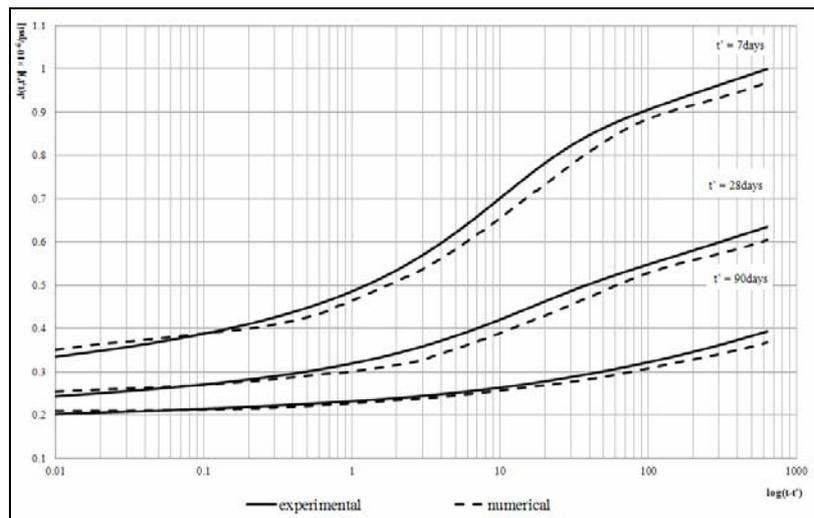


Figure 1: Comparison between compliance curves reported in [8] and obtained via NEWCON3D (dashed lines).

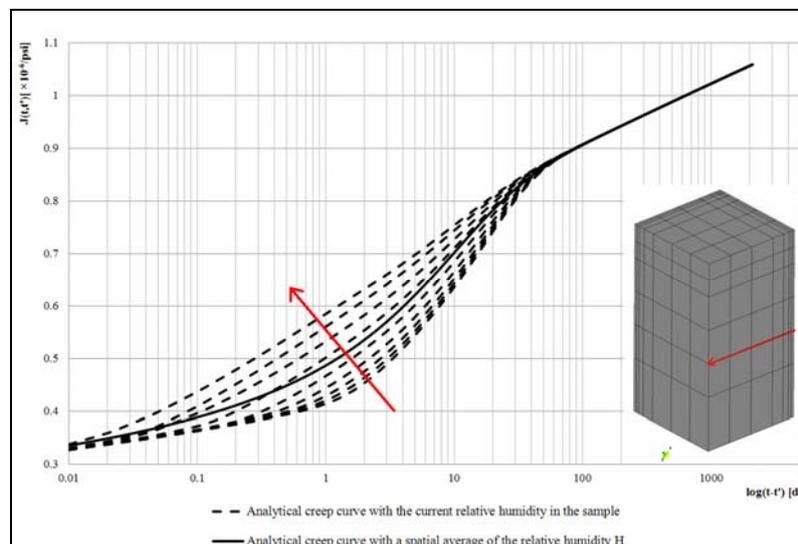


Figure 2: Comparison between analytical creep curves obtained with the spatial average relative humidity (solid line) and with the current relative humidity (dashed lines).

4.2 Numerical Analysis of concrete at the meso-level

If concrete is considered as a composite material made of cement paste and aggregates, the only phases subjected to creep are cement paste and ITZ, whereas aggregates behave elastically, as reported in [6]-[7].

The adopted parameters are listed in **Table 2** and **Table 3**.

The tests by L’Hermite et al. are again taken as reference; the sample is first subjected to basic creep only (specimen kept in water) and subsequently to drying creep (specimen cured in water; at $t_0 = 2$ days exposed to drying at 50% relative humidity and 20°C). The sample is additionally loaded by an axial compressive stress of 9.07 MPa at time $t' = 7$ days.

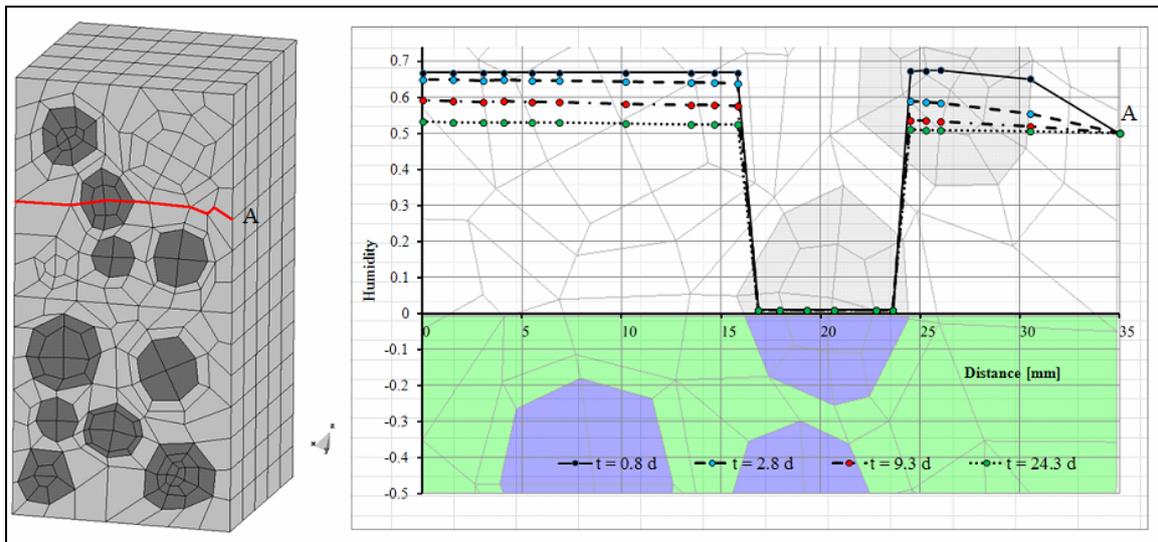
Table 2: Meso-scale analysis: parameters for cement paste.

Elastic Modulus [MPa]	28522.1
f_{cm28} [MPa]	36.3
Cement Content [kg/m^3]	350
Water Content [kg/m^3]	171.5
Aggregate-cement ratio	1.5

Table 3: Meso-scale analysis: parameters for aggregates.

Elastic Modulus [MPa]	67000
-----------------------	-------

Figure 3 depicts the evolution of relative humidity within the sample at different times along a section (“A”) passing through an aggregate; the physical barrier exerted by the aggregate towards the flux of humidity is clearly evidenced. Consequently, some delay in drying can be noticed when comparing humidity fluxes from the macro-scale analysis (the results are not reported here for sake of brevity).

**Figure 3:** Evolution of relative humidity in the sample at different times for the red line.

A comparison between creep curves for two different points in the cement paste is shown in **Figure 4**: as expected, in proximity of the aggregate, creep effects (hence deformations) are reduced due to a “stiffening effect” coming from the aggregates themselves.

The same concrete sample ($3.5 \times 3.5 \times 7 \text{ cm}^3$, in symmetry conditions) has been further considered by including the ITZ and allowing a damage triggering effect; in this way a first estimate of the role of ITZ on the overall mechanical behavior of the concrete sample has been obtained, in view of defining an appropriate concrete mixture for e.g. responding to spalling under high temperature conditions.

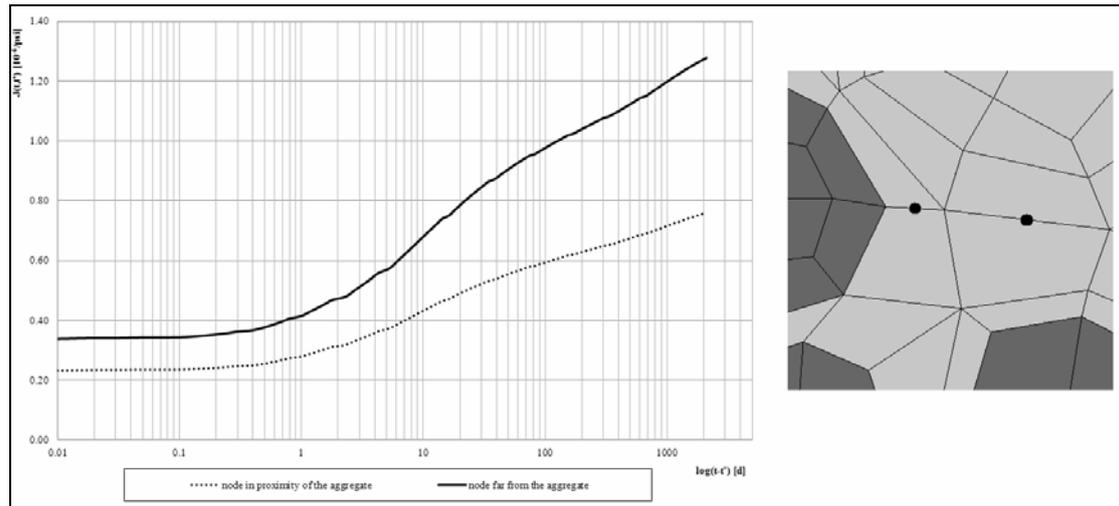


Figure 4: Comparison of the creep function for two nodes of the cement paste.

In **Table 4** and **Table 5** the parameters used for the three different phases as well as for Mazars' damage law are listed.

Table 4: Parameters for cement paste, aggregates and ITZ.

	CEMENT PASTE	ITZ	AGGREGATE
Elastic Modulus [MPa]	28522.1	10000	67000
Poisson's Ratio	0.15	0.2	0.2
Reference Diffusivity along x,y,z [mm^2/d]	150	300	0
Thermal Conductivity along x,y,x [$\text{N}/\text{d K}$]	110000	220000	170000

Table 5: Parameters used for the isotropic Mazar's damage model (cement paste and ITZ).

k_0	1×10^{-4}
A_t	1.2
B_t	5000
A_c	1
B_c	1000
B	1

The sample is subjected to a compressive load, in displacements control so to simulate the non linear material response in the softening regime.

The evolution of damage within the sample is shown in **Figure 5** (reference line on the right) at three different times; particularly, it can be noticed that the peaks (maximum damage) occur in the ITZ, the weakest zone of concrete.

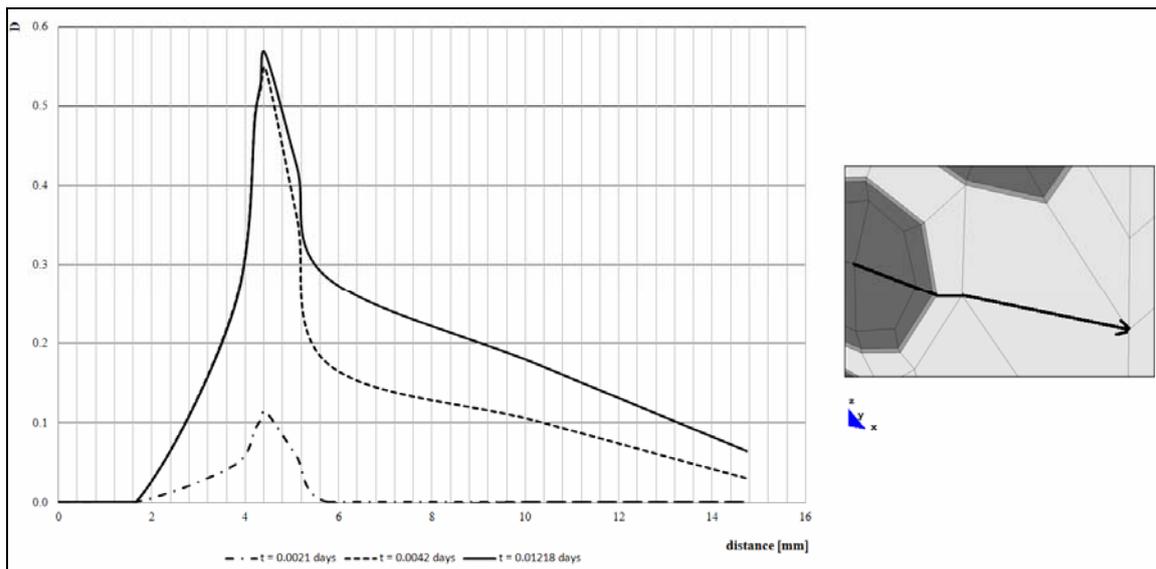


Figure 5: Evolution of damage within the sample at different times for the reference line on the right.

5 CONCLUSIONS

From the previous results some key conclusions can be drawn.

First, the three-dimensionality of the geometric description of concrete at the meso-level can be now appreciated, since 2D or axi-symmetric sections often used in the past were not able to describe with enough precision the complex behaviour of concrete as a composite material. Now this effect can be systematically incorporated due to the availability of powerful computers.

Then, creep of cement paste and ITZ, described by consolidated and complete models as the B3 one presented by Z.P. Bazant (carefully calibrated on the basis of well known experimental results like the ones given by R. L'Hermite), allows to incorporate in the model the complex reality of creep, which is not only a matter of fluid flow and pressure dissipation but also the result of chemical-physical reactions.

Again, the description of concrete as a composite material, in connection with porous media analysis, allows for understanding the hygro-thermal and mechanical response of concrete, first of all in terms of hygral and thermal changes in a material where aggregate inclusions (incapsulated by ITZ concave volumes) appear with some statistical distribution inside concrete (as originally described by F.H.Wittmann, even if without the ITZ effect, [22]).

Hygral barriers due to the presence of aggregates can be seen only at this modelling level. On the other side, thermal conductivity properties dominate the thermal conduction in the sample.

Finally, from the mechanical viewpoint, the remarkable damage peak effect arising from the inclusion of ITZ, if compared with the less pronounced peak when ITZ is disregarded from the analysis, is here reported.

REFERENCES

- [1] Bazant, Z.P. and Baweja, S. Creep and shrinkage prediction model for analysis and design of concrete structures: Model B3. *Adam Neville Symposium: Creep and Shrinkage – Structural Design Effects*, ACI SP – 194 (2000) 1–83.
- [2] ACI Committee 209, Guide for Modeling and Calculating Shrinkage and Creep in Hardened Concrete *ACI Report 209, 2R-08*, Farmington Hills (2008).
- [3] Majorana, C.E. Influenza del danno sul comportamento termoigrometrico e meccanico del continuo. *Estratto dal giornale del genio civile*, Luglio, Agosto e Settembre 1989.
- [4] Mazars, J., Pijaudier-Cabot, G. and Pulikowski, J. Steel-concrete bond analysis with non local continuous damage. *Rapport Interne No. 96*, March 1989.
- [5] Salomoni, V., Majorana, C.E., Mazzucco, G., Xotta, G. and Khoury, G.A. Multiscale modelling of Concrete as a Fully Coupled Porous Medium. *Concrete Materials: Properties, Performance and Applications*, Ch. 3, NOVA Science Publishers (2009) 171-231.
- [6] Granger, L. and Bazant, Z. Effect of Composition on Basic Creep of Concrete and Cement Paste. *Journal of engineering mechanics* (1995) **121**.
- [7] Baweja, S., Dvorak, G.J. and Bazant, Z.P. Triaxial Composite Model for Basic Creep of Concrete. *Journal of Engineering Mechanics* (1998) 959-965.
- [8] L’Hermite, R., Mamillan, M. and Lefèvre, C. Nouveaux résultats de recherches sur la déformation et la rupture du béton. *Ann. Inst. Batiment Trav. Publics* (1965) **18**: 323-360.
- [9] Mazars, J. Application de la mécanique de l’endommagement au comportement non linéaire et la rupture du béton de structure. *Thèse de Doctorat d’Etat, L.M.T.*, Université de Paris, France (1984).
- [10] Pijaudier–Cabot, J. Continuum Damage Theory – Application to Concrete. *J Engrg Mech*, ASCE (1989) **115**: 345-365.
- [11] Marotti de Sciarra, F. A general theory for nonlocal softening plasticity of integral-type. *International Journal of Plasticity* (2008) **24**: 1411-1439.
- [12] Wriggers, P. and Moftah, S.O. Mesoscale Models for concrete: Homogenisation and damage behavior. *Finite El An Des* (2006) **42**: 623-636.
- [13] Schlangen, E., Koenders, E.A.B. and Van Breugel, K. Influence of internal dilation on the fracture behavior of multi-phase materials. *Engrg Frac Mech* (2007) **74**: 18-33.
- [14] Ollivier, J.P., Maso, J.C. and Bourdette, B. Interfacial Transition Zone in Concrete. *Adv Cem Bas Mat* (1995) **2**:30-38.
- [15] Scrivener, K.L., Crumby, A.K. and Laugesen, P. The Interfacial Transition Zone (ITZ) Between Cement Paste and Aggregate in Concrete. *Interface Science* (2004) **12**: 411-421.
- [16] Liao, K.Y., Chang, P.K., Peng, Y.N. and Yang, C.C. A study on characteristics of interfacial transition zone in concrete. *Cem Con Res* (2004) **34**(6): 977-89.
- [17] Garboczi, E.J., Bentz, D.P. and Schwartz, L.M. Modeling the Influence of the Interfacial Zone on the DC Electrical Conductivity of Mortar. *Adv. Cement-Based Mater.* (1995) **2**(5): 169–81.
- [18] Zheng, J.J. and Zhou, X.Z. A numerical method for predicting the elastic modulus of concrete made with two different aggregates. *J Zhejiang Univ SCIENCE A* (2006) **7**(II): 293-296.

- [19] Simeonov, P. and Ahmad, S. Effect of transition zone on the elastic behavior of cement-based composites. *Cem Conc Res* (1995) **25**(1): 165–76.
- [20] Garboczi, E.J., Bentz, D.P. and Shane, J.D. Effect of the interfacial zone on the conductivity of Portland cement mortars. *J Am Cer Soc* (2000) **83**(5): 1137-1144.
- [21] Bazant, Z. and Kim, J. Improved prediction model for time-dependent deformations of concrete: Part1-7, in *Materials and Structures* (1992) **25**: 21-28.
- [22] Wittman, F.H. Surface tension, shrinkage and strength of hardened cement paste. *Mat. Struct* (1968) **1**(6): 547-552.

NON-LINEAR MODELLING, DESIGN AND PRODUCTION OF STEEL BLAST-RESISTANT DOORS AND WINDOWS

VALENTINA A. SALOMONI*, GIANLUCA MAZZUCCO*, GIOVANNA XOTTA*,
RICCARDO L. FINCATO*, MARCO SCHIAVON**, CARMELO E. MAJORANA*

*Department of Structural and Transportation Engineering
Faculty of Engineering, University of Padua
Via F.Marzolo, 9 – 35131 Padua, Italy

e-mail: salomoni@dic.unipd.it, mazzucco@dic.unipd.it, xotta@dic.unipd.it, fincato@dic.unipd.it,
majorana@dic.unipd.it, www.dic.unipd.it

** Wellco S.p.A. Engineering and Construction,
v.le Mazzini 65/5, 31049 Valdobbiadene (TV), Italy
e-mail: m.schiavon@wellco.it, www.wellco.it

Key words: Blast, Impact, Plasticity, Steel doors.

Abstract. Numerical-experimental results are here described, derived from an innovative experience at both national and international level, related to modelling, designing and producing steel blast-resistant doors and windows. Their capability to sustain thermal loads due to fire hazards is additionally accounted for. The activity has been developed within a collaboration between Wellco S.p.A. and some researchers of the Department of Structural and Transportation Engineering of the University of Padua, Italy. The study has been conducted to define and characterize the non-linear response of a large number of doors and steel framed windows, with the objective of sustaining dynamic loads from explosive hazards of fixed magnitude, variable design and clearing times. The local overcome in the strength limit (with correspondent plastic response) and possible formation of plastic hinges has been critically discussed. Numerical models have allowed for refining first design sketches and subsequently understanding the real thermo-mechanical behaviour for the investigated structures. Experimental tests on typical steel doors at 1:1 scale have been performed at the Laboratory of Construction Materials of the same Department above. Such tests had the objective of “a-posteriori” verifying the correctness of the already available numerical results, validating the adopted procedures and correspondingly guaranteeing the doors’ structural efficiency even under dynamic loads higher than design ones.

1 INTRODUCTION

The work comes from a joint collaboration in the field of Blast Resistant buildings, doors and windows. Particularly, steel doors and windows have been investigated following a request of an international client constructing gas plants in Eastern Europe.

Doors and windows effectively represent the most peculiar elements when designing blast resistant buildings, e.g. if their re-opening after explosion is requested for safety reasons

(escape of personnel after the blast) [1]. Such an aspect is largely binding, essentially in the numerical modelling phases, being in fact its fulfilment to be guaranteed by controlling specific parameters. It is hence possible to admit that doors and windows enter the plastic regime (also considering that Ultimate Limit States are accounted for [1], [2]), but this requires additional verifications of well-defined ductility and rotation ratios. Consequently, the correct element behaviour is not affected and safety/rescue operations are ensured.

Doors and windows designed and constructed by Wellco S.p.A. are of various types and dimensions: from the one-shutter 1000×2000 mm² door up to larger 3500×4500 mm² ones.

In relation to doors and windows' dimensions [3]-[5], to value and duration of the explosive load (fixed), first design lines have been developed and subsequently the whole problem has been investigated by verifying procedures and characterizing the dynamic response of the structural elements. Non-linear (for material and geometry) analyses have been conducted also considering frames, joints, plates, hinges, glasses and opening devices. Procedures and methodology of analysis had already been known from a previous joint experience [6]-[9].

In the following the main results related to one door type only have been reported for sake of brevity, as well as the lines followed in agreement with International Recommendations; additional analyses, not described here, have also allowed for designing doors and windows under impulsive loads as well as thermal ones, satisfying the Italian requirements for REI60 or REI120.

2 F.E. MODELLING

2.1 Geometry

Finite Element models have been set up to simulate the door's behaviour in its closed configuration; beam-type elements have been used for defining the main structure (frame), **Figure 1**, whereas shell-type ones have characterized the internal and external steel plates.

Beam elements present a transversal section in agreement with the design one, to allow for defining a correct stiffness to internal and border elements (**Figure 2**); the number of horizontal stiffeners has been defined proceeding via a series of repeated analyses to obtain a structural response to guarantee the appropriate functional door's behaviour. Again, steel plates have been modeled to reproduce the design drawings (**Figure 3**).

2.2 Constraint conditions

Each shutter is connected to the edge wall through hinges (which number has been determined again via repeated analyses), modeled with rigid constraints to allow for free rotations. The counter-frame has been represented by the introduction of springs with equivalent stiffness, active in compression only; such a stiffness has been evaluated by considering a three-dimensional local model to which an imposed unit displacement has been applied (**Figure 4**). The contact between shutters has been additionally considered by interposing *link* elements, inactive if the response leads them to move away from each other (**Figure 5**), and closure points (representative of the real closure system, **Figure 6**).

During the *rebound* phase (qualifying the dynamic response and corresponding to the door bending in opposite direction with respect to the applied load) it has been assumed that the

only active constraints are exclusively represented by closure points and hinges.

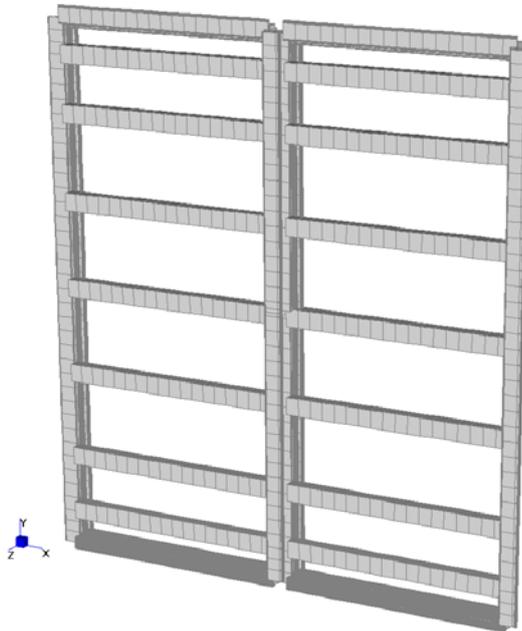


Figure 1: Main frame model, double shutter-type door.

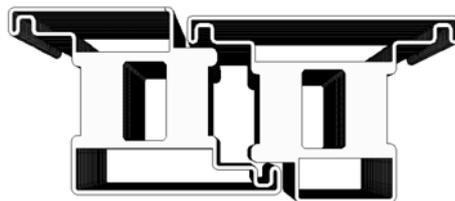


Figure 2: Internal and border elements.



Figure 3: Typical horizontal section.

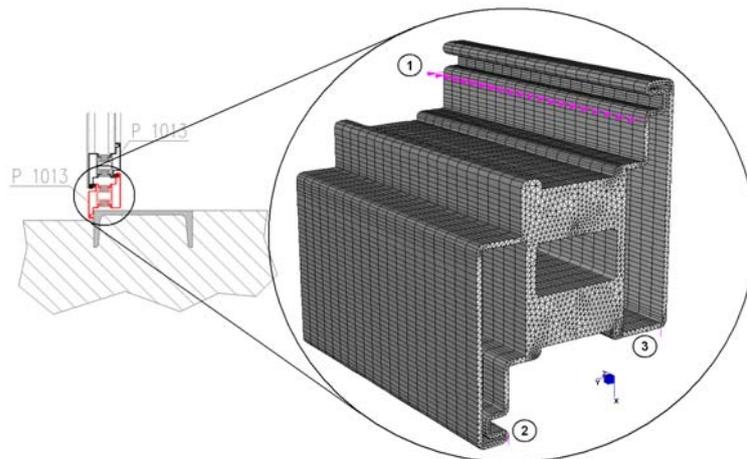


Figure 4: 3D model for counter-frame stiffness definition.

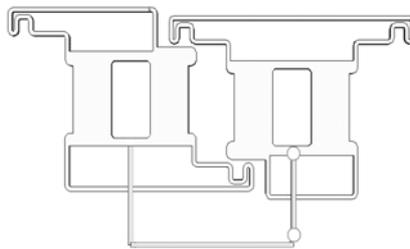


Figure 5: Shutter-to-shutter contact elements.

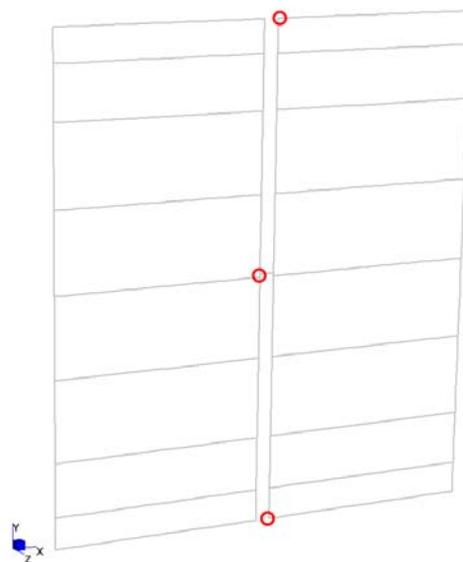


Figure 6: Schematic representation of closure points (in red).

2.3 Dynamic analyses

Once the (design) peak value of the blast load has been defined, as well as the time required for dissipating overpressure (t_d), it has been assumed to consider the impact of a plane frontal wave considering, in agreement with the Regulations, specific values for peak reflected pressure (P_r), stagnation pressure (P_s) and clearing time (t_c), realizing diagrams of the type of **Figure 7**.

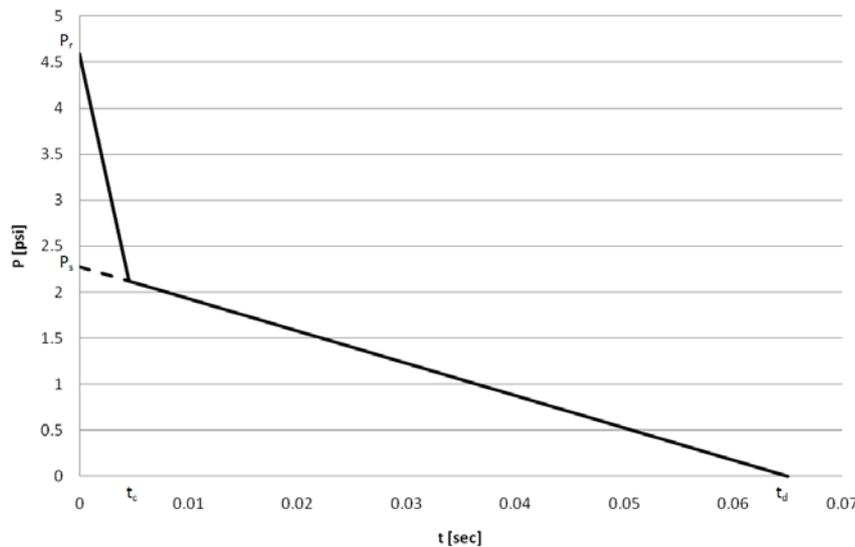


Figure 7: Typical pressure-time diagram for plane frontal blast wave.

The analyses have additionally included effects of dynamic damping to take into account a possible reduction in stress deriving from internal frictions and yielding of some elements; particularly, damping effects coming from the formation of plastic hinges have been represented by assigning an elasto-plastic behaviour to the material, whereas a fixed damping ratio has allowed for evaluating damping from internal frictions.

In **Figure 8** a typical displacement evolution in damped and undamped configurations is reported: the maximization of effects (peaks of maximum and minimum) is reached in both situations; this comes from the fact that the blast is rapidly exhausted (red curve) and the damping contribution is highlighted for longer times only (larger than t_d). Consequently, concerning the design phase, maximum actions only have to be considered and not the entire loading history; such an aspect has allowed for developing essentially undamped analyses, reducing computational times without losing in approach generality and/or underestimating the real response.

2.4 Analysis of results

For the considered door, the analyses have highlighted an elastic response for the internal frame in the peak phase (**Figure 9**), with occurrence of out-of-plane displacements compatible with the correct structural behaviour of the whole door, both in the peak and rebound phase

(Figure 10).

Further, constraints reactions (variable with the structure's oscillation consequent to the explosion) have been analysed and maximum values have been taken as reference; in general, internal hinges appeared to be overloaded, due to the door's bending, and their verification has been developed in agreement with Eurocode 3 (Figure 11); anchoring bolts of the perimetric counter-frame have been checked as well.

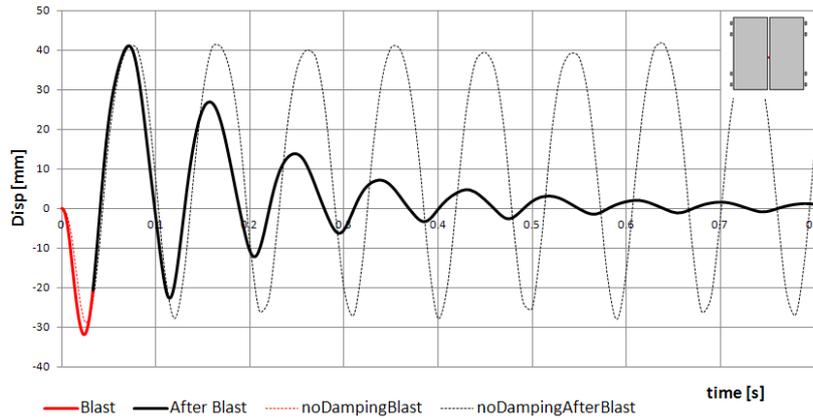


Figure 8: Displacement vs. time for damped and undamped analyses.

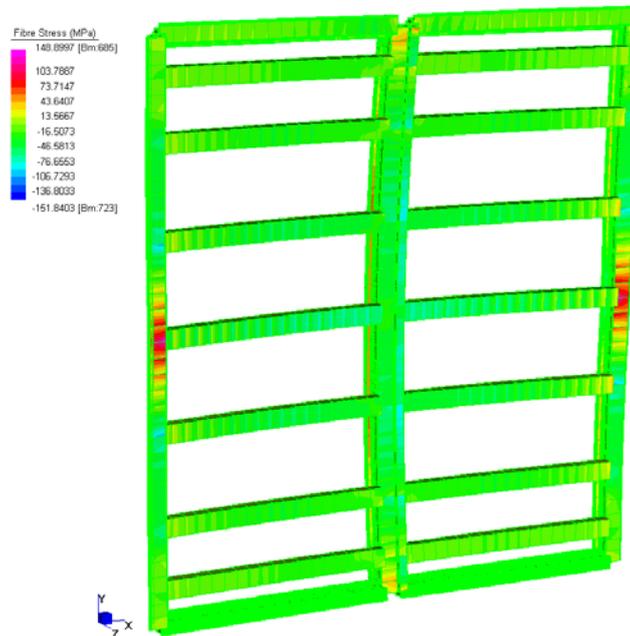


Figure 9: Maximum stresses in the peak phase, internal frame.

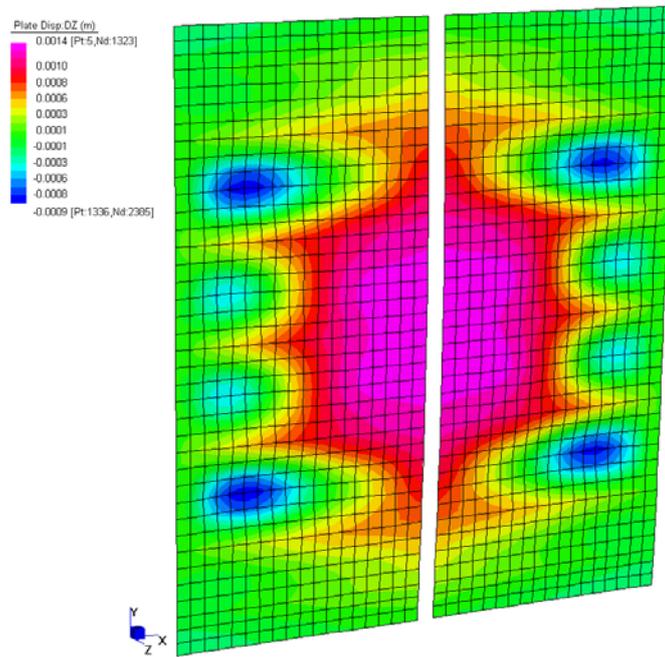


Figure 10: Contour map of out-of-plane displacements during rebound (steel plates).

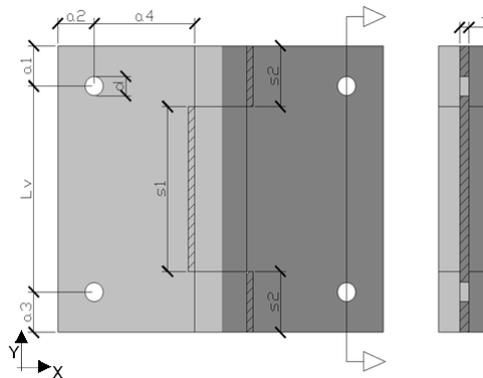


Figure 11: Geometric scheme for hinges' verification.

The design of the door has been then completed via tensile/compressive, bending and shear (and coupled actions) verifications for each structural element following the procedures of the Ultimate Limit States; as required by EC3, design and resistant actions ratios have been controlled to check their being lower than one, both in the peak and rebound phase, as well as close to these states. When such a limit had not been satisfied and correspondingly for some elements a plastic regime had been evidenced, the respect of the additional limits provided by ASCE standards in term of ductility and rotation ratios have been controlled.

3 EXPERIMENTAL TESTS

The results obtained from the numerical analyses in terms of strength and deformability have been subsequently compared with those coming from laboratory tests, conducted on a real scale door, confirming the correctness of approaches and methodologies as well as the door capability to sustain dynamic loads even larger (nearly double) than the design ones. The tests have been developed at the Laboratory of Construction Materials of the Department of Structural and Transportation Engineering in Padua, Italy. Even in an “ultimate” configuration, the requirement of a door reopening has been guaranteed, proving its efficiency in response and technical realization for dynamic regimes.

The test scheme has been planned to (dynamically) reconstruct the explosive event even without using blast-reproducing devices, hence minimizing costs connected to the entire test set-up but anyway ensuring a correspondence between tests and real behaviour. It has consequently been chosen to perform an impacting mass with fixed weight to be thrown against the door; the nature of the impact is so local, being the impact area not distributed on the whole door’s surface (**Figure 12**), but such a condition has been verified to be more severe in the evaluation of the door behaviour, so once again ensuring a more precautionary situation.



Figure 12: Test set-up: impacting element (left) and position of displacement transducers (right).

The design condition has been reconstructed making reference to energetic equivalences,

by matching the kinetic energy associated to the impacting mass with the work done by the blast load; in this way a “design” height has been determined, such as to certify that the impact could lead to pressure values equivalent to the explosion ones.

Via such an approach it has not been possible to take into account the (real) transient nature of the pressure wave; a possible, consequent, underestimate in the blast effects is however associable to the rebound phase only, but these have been evaluated as negligible: in fact, the maximum displacements used in the energetic equivalence above are numerically derived and consequently they come from having included real quantities such as reflected pressure (higher than the design one) and wave duration. It is additionally verifiable that a structure is more sensitive towards a variation in the peak pressure rather than in a different time distribution of the pressure wave itself.

In the methodological definition of the tests and in the subsequent discussion of results even effects coming from deformable constraints have been included as well (in fact, in the test the door is not restrained to any edge wall).

It has been observed from the displacement values measured by transducers (**Figure 13**) that: a) the whole system response (door and supporting frames) results damped, favorable condition for a structure designed to respond to dynamic loads; b) the peaks in the curve subsequent to the first one are effectively “fictitious” (i.e. non reproducible in reality), because consequent to the repeated impact of the mass and strictly related to the planned tests (it is however possible to say that such a phenomenon represents a condition in favor to safety, being the door more stressed); c) a rebound is evidenced: this could be amplified in reality, due to a depressurization consequent to the explosion, but such an aspect is believed to have no consequences on the real capacity of the door for sustaining loads, nor on its reopening (**Figure 14**).

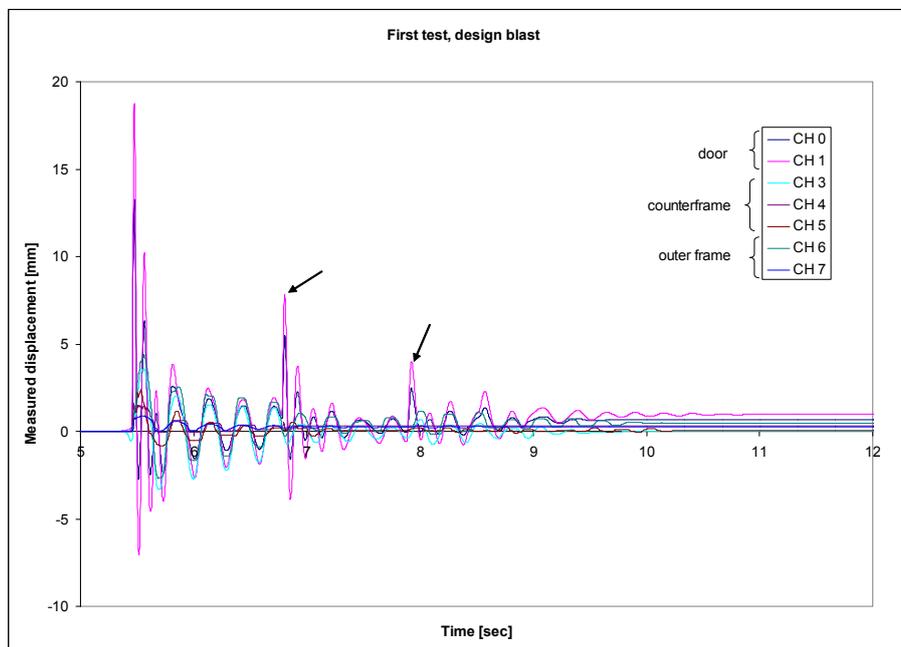


Figure 13: Measured displacements, test with design pressure.

The tests (conducted in 3 subsequent phases), as anticipated, have allowed for demonstrating the validity of the hypotheses and of the adopted procedures, as well as to prove the agreement between numerical and experimental results. Even including (necessary) simplifying assumptions, the tests have shown a correct structural behaviour for the door, both under the design load and for nearly double ones, not only without evidencing collapse phenomena (even locally), but also qualifying the post-explosion functional character of the door itself.



Figure 14: Door reopening after design blast.

4 CONCLUSIONS

Numerical-experimental results have been briefly described, referring to an innovative experience, at both national and international level, in modelling, designing and realizing steel blast-resistant doors and windows.

The study has been conducted to define and characterize the non-linear response of a large number of doors and windows with steel frame, with the objective of sustaining dynamic loads from explosive hazards of fixed magnitude and variable design and clearing times.

The local overcome in the strength limit (i.e. generating a plastic response) and possible formation of plastic hinges has been critically discussed and examined in relation to prescribed Regulations and Recommendations.

The numerical models have allowed for refining first design sketches and subsequently understanding the real structural behaviour of the investigated structures.

Experimental tests on typical steel doors at 1:1 scale have been conducted with the objective of “a-posteriori” verifying the correctness of the already available numerical results, validating the adopted procedures and correspondingly guaranteeing the doors’ structural efficiency even under dynamic loads even higher than design ones. Their capability to sustain

thermal loads due to fire hazards has been additionally accounted for.

REFERENCES

- [1] ASCE Report. *Design of Blast Resistant Buildings in Petrochemical Facilities*, NY (1997).
- [2] UNI EN 1993-1-1. *Eurocode 3: Design of steel structures. Part 1-1: General rules and rules for buildings* (1993 and 2005).
- [3] UNI EN 13541:2000. *Glass in building – Security Glazing – Testing and classification of resistance against explosion pressure* (2000).
- [4] UNI EN 572-1:2004. *Glass in building - Basic soda lime silicate glass products - Part 1: Definitions and general physical and mechanical properties* (2004).
- [5] prEN 13474-2 (CENT/TC129/WG8). *Glass in building - Design of glass panes - Part 2: Design for uniformly distributed loads* (2000).
- [6] Majorana C.E., Salomoni V.A. Preliminary design and numerical analyses of steel blast doors and windows, *V Asia-Pacific Conference on Shock and Impact Loads on Structures*. 301-307 (2003) (ISBN: 981-04-8563-8).
- [7] Salomoni V.A. Predicted Responses of Structures Subjected to Blast and Blast-Induced Phenomena, *2nd Int. Conf. On Protection of Structures Against Hazards*. Singapore, Dec. 01-03 (2004) (ISBN: 981-05-0613-9).
- [8] Naval Facilities Engineering Command. *Blast resistant structures - Design Manual 2.08*. Philadelphia, PA (1986).
- [9] BP International Limited, Research & Engineering Centre. *Procedure for the Design of Buildings subject to Blast Loadings. RP 4-6*. BP Group Recommended Practices and Specifications for Engineering: Engineering Practices Group, Chertsey Road, Sunbury-on-Thames, Middlesex, TW16 7LN, UK (1993).

SUBLOADING SURFACE PLASTICITY MODEL ALGORITHM FOR 3D SUBSIDENCE ANALYSES ABOVE GAS RESERVOIRS

VALENTINA A. SALOMONI^{*}, RICCARDO L. FINCATO^{*}

^{*} Department of Structural and Transportation Engineering
Faculty of Engineering, University of Padua
Via F.Marzolo, 9 – 35131 Padua, Italy
e-mail: salomoni@dic.unipd.it, fincato@dic.unipd.it, www.dic.unipd.it

Key words: Subsidence, Plasticity, Subloading surface, Environmental impact, Gas recovery.

Abstract. The coupled hydro-mechanical state in soils coming from consolidation/subsidence processes and undergoing plasticity phenomena is here evaluated by means of the subloading surface model. The most important feature of this theory is the abolition of the distinction between the elastic and plastic domain, as it happens in conventional elastoplastic models. This means that plastic deformations are generated whenever there is a change in stress and a smoother elasto-plastic transition is produced. The plasticity algorithm has been implemented in the PLASCON3D FE code (on the basis of a previous 2D version), coupling hydro-(thermo)-mechanical fields within a saturated porous medium (locally partially saturated at reservoir level due to the possible presence of a gas phase) subjected to external loads and water/gas withdrawals from deep layers (aquifers/reservoirs). The 3D model has been first calibrated and validated against examples taken from literature, and then subsidence analyses at regional scales due to gas extractions have been developed to predict the evolution of settlements and pore pressure in soils for long-term scenarios.

1 INTRODUCTION

Surface subsidence due to withdrawal of underground fluids occurs in many parts of the world, see for instance the case book of Poland [1]. Underground fluids involved are either water from superficial aquifers or gas and oil from usually deeper reservoirs. Such surface settlement is a particular threat if it is experienced in low lying areas, close to the sea, e.g. Groningen in the Netherlands (gas), Venice (water) and Ravenna (water and gas) in Italy, Wilmington (oil) in the USA. Surface subsidence of this kind is almost exclusively understood in terms of drop of pressure in the aquifers or in the reservoir: i.e. withdrawal of these underground fluids results in a reduction of their pressure downhole; this in turn increases the part of the overburden carried by the skeleton of the reservoir rocks causing compaction. The compaction manifests itself, through deformation of the overlying strata, as surface settlement.

In case of a single fluid (water) involved, compaction can easily be explained by the principle of Terzaghi [2] which states that the compression of a porous medium is controlled by changes of effective stresses, i.e. variations of the difference between total stresses and the

pressure of the fluid in the pores. However, when more fluids are involved or more phases of the same fluid, the Terzaghi traditional expression of effective stress alone is not sufficient to completely justify measured compaction and the concepts of unsaturated soil mechanics with appropriate stress measures and elastoplasticity concepts are needed. Drop of reservoir pressure is not the only mechanism leading to reservoir compaction and suction effects must also be accounted for at least for some types of extracted fluids and some reservoir rocks.

It is then proposed in [3-5] that capillary effects and structural collapse can not be ruled out as significant factors in the development of subsidence occurring above gas fields. These phenomena seem to provide sound explanations for continuing surface settlements when reservoir pore pressures stabilise and for additional settlements occurring even after the end of gas production. However, it is to be said that for the investigated area here considered, undergoing subsidence, there is no direct experimental evidence on samples from the field to show the key effect of capillarity on subsidence itself and hence any additional consideration could be largely speculative with many assumptions that are not justified enough.

Again, the discussion about the contribution of capillary effects when performing reservoir compaction and subsidence analyses at regional scale is out of scope for the present paper. The idea is to make use of unconventional plasticity [6] by means of the subloading surface model [7-11] for predicting softening behaviour of soil as well as reducing computational efforts when performing fully coupled hydro-mechanical subsidence analyses in three-dimensional domains [12], as demonstrated below. The reader is referred to [13] for discussions about modelling strain-softening from the computational point of view.

2 THE PROBLEM OF SUBSIDENCE ABOVE GAS RESERVOIRS

The particular subsidence problem solved here is first briefly summed up as follows [14-18]. It is supposed that there are several gas reservoirs at different levels, and some of the reservoirs have an edge aquifer. It is further assumed that in each respective domain there is only one fluid: water in the aquifers and gas in the reservoirs. The gas, upon exploitation of the reservoirs, may be substituted by encroaching water, which comes from the edge aquifers and from possible leaky aquitards. Capillary effects due to the simultaneous presence of gas and water in the reservoirs are not accounted for, as previously explained.

Under these assumptions the following balance equations can be written, where the chosen macroscopic field variables are displacements \mathbf{u} and water pressure p_w .

Linear momentum balance equation for the mixture solid + water or solid + gas

$$\nabla \cdot \boldsymbol{\sigma} = \mathbf{0} \quad (1)$$

where $\boldsymbol{\sigma}$ is the total stress tensor; no variable body forces are accounted for.

An averaged density of the mixture of the form

$$\rho = (1 - \phi)\rho_s + \phi\rho_\pi \quad \pi = w \text{ (water) or } g \text{ (gas), } s = \text{solid} \quad (2)$$

ϕ being the porosity, is assumed in the following.

Flow conservation equation for the aquifers and aquitards

$$-\nabla \cdot \left\{ \frac{\mathbf{k}}{\mu} \nabla (p_w + \rho_w g h) \right\} + \left(\mathbf{m}^T - \frac{\mathbf{m}^T \mathbf{D}_T}{3K_s} \right) \frac{\partial \boldsymbol{\varepsilon}}{\partial t} + \left[\frac{(1 - \phi)}{K_s} + \frac{\phi}{K_w} - \frac{1}{(3K_s)^2} \mathbf{m}^T \mathbf{D}_T \mathbf{m} \right] \frac{\partial p_w}{\partial t} = 0 \quad (3)$$

where \mathbf{m} is a vector with components equal to unity for the normal stress components and zero for the shear stress components, K_w the bulk modulus, K_s the averaged bulk modulus of the solid grains, k the absolute permeability matrix of the medium, \mathbf{D}_T the tangent matrix, $\boldsymbol{\varepsilon}$ the total strain of the skeleton, μ the dynamic viscosity of water.

Instead of writing a similar mass balance equation as Eq. (3) for the gaseous phase, we consider its integral form for the whole reservoir volume. This is valid if the reservoir volume is small compared to the analysed cross section and its thickness is small compared to the depth of burial [15, 16, 18, 19]. The conservation equation assumes hence the form of a

Material balance equation for the reservoir

$$GB_{gi} = (G - G_p)B_g + W_e - W_p B_w \quad (4)$$

where G is the initial free gas in place, B_g the gas formation volume factor, B_w the water formation volume factor, G_p the cumulative gas production, W_p the cumulative water production and W_e the influx from the adjacent aquifer and from leaky aquitards; the index i denotes initial conditions. Its incremental form can be found in [13].

The model is completed by the constitutive relation for solid mechanical behaviour relating the effective stress $\boldsymbol{\sigma}'$ and the adopted strain measure. In general, for small displacement gradients, it can be written as

$$\dot{\boldsymbol{\sigma}}' = \mathbf{E}(\dot{\boldsymbol{\varepsilon}} - \dot{\boldsymbol{\varepsilon}}^p) \quad (5)$$

$\dot{\boldsymbol{\varepsilon}}^p$ being the plastic strain rate, and \mathbf{E} the fourth order elasticity tensor.

3 MODELLING PLASTICITY – THE SUBLOADING SURFACE MODEL

The subloading surface model is a particular elasto-plastic model falling within the framework of unconventional elastoplasticity [6], an extended elastoplasticity theory such that the interior of the yield surface is not a purely elastic domain, but rather a plastic deformation is induced by the rate of stress inside the yield surface [7-11]. Its main features are briefly recalled here.

In the subloading surface model the conventional yield surface is renamed the normal yield surface, since its interior is not regarded as a purely elastic domain. The plastic deformation develops gradually as the stress approaches the normal yield surface, exhibiting a smooth elastic-plastic transition. Thus the subloading surface model fulfils the smoothness condition [11, 20-22], which is defined as the stress rate-strain rate relation (or the stiffness tensor) changing continuously for a continuous change of stress rate. Strain accumulation is predicted for a cyclic loading with an arbitrary stress amplitude, where the magnitude of accumulated strain depends continuously on the stress amplitude because of the fulfillment of the smoothness condition. Inelastic deformation occurs immediately when the stress point once again moves outward the current yield surface. Zero diameter yield surface bounding surface models, nested surface models, and subloading models have this attribute, but do not display any purely elastic response [6].

A subloading surface is also introduced (together with the normal yield one), which always passes through the current stress point $\boldsymbol{\sigma}$ and keeps a shape similar to that of the normal yield surface and a similar orientation with respect to the origin of stress space, i.e. $\boldsymbol{\sigma} = \mathbf{0}$.

The ratio of similarity is named normal yield ratio and governs the approach of the

subloading surface to the normal one, i.e. if $R = 0$ the subloading surface is a point coinciding with the origin of the stress space, whereas $0 < R < 1$ represents the subyield state and with $R = 1$ the stress lies directly on the normal surface.

The subloading surface can be described by the scalar-valued tensor function

$$f(\boldsymbol{\sigma}) = RF(H) \quad (6)$$

where the scalar H is the isotropic hardening/softening variable; in agreement with [10] the normal yield surface takes e.g. the form

$$F = F_0 \exp\left(\frac{H}{\rho' - \gamma}\right) \quad (7)$$

in which F_0 is the initial value of F , ρ' and γ the slopes of the normal consolidation and swelling curves respectively in $\ln v - \ln p$ space (v being the specific volume and $p = -\text{tr}(\boldsymbol{\sigma})/3$).

The extended consistency condition for the subloading surface is obtained by differentiating Eq. (6), which leads to

$$\text{tr}\left(\frac{\partial f(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} \dot{\boldsymbol{\sigma}}\right) = \dot{R}F + RF'\dot{H} \quad (8)$$

together with considering the evolution rule of the normal yield ratio, given by

$$\dot{R} = U \|\dot{\boldsymbol{\varepsilon}}^p\| \quad \text{for } \dot{\boldsymbol{\varepsilon}}^p \neq \mathbf{0} \quad (9)$$

where $\dot{\boldsymbol{\sigma}}$ is the proper objective co-rotational stress rate, $\dot{\boldsymbol{\varepsilon}}^e = \mathbf{E}^{-1}\dot{\boldsymbol{\sigma}}$, and U is a monotonically decreasing function of R satisfying the condition

$$\begin{cases} U = +\infty & \text{for } R = 0 \\ U = 0 & \text{for } R = 1 \\ (U < 0 & \text{for } R > 1) \end{cases} \quad (10)$$

The associated flow rule is assumed as

$$\dot{\boldsymbol{\varepsilon}}^p = \bar{\lambda} \bar{\mathbf{N}} \quad (11)$$

where $\bar{\lambda}$ is the positive proportional factor representing the increment of plastic deformation along the direction given by the normalized outward normal of the subloading surface $\bar{\mathbf{N}}$

$$\bar{\lambda} = \frac{\text{tr}(\bar{\mathbf{N}}\dot{\boldsymbol{\sigma}})}{\bar{M}^p} \quad (12)$$

$$\bar{\mathbf{N}} \equiv \frac{\partial f(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} \left/ \left\| \frac{\partial f(\boldsymbol{\sigma})}{\partial \boldsymbol{\sigma}} \right\| \right. \quad (13)$$

being \bar{M}^p the plastic modulus.

The loading criterion is finally given [22, 23]

$$\begin{cases} \dot{\boldsymbol{\varepsilon}}^p \neq \mathbf{0}: & \text{tr}(\bar{\mathbf{N}}\mathbf{E}\dot{\boldsymbol{\varepsilon}}) > 0 \\ \dot{\boldsymbol{\varepsilon}}^p = \mathbf{0}: & \text{tr}(\bar{\mathbf{N}}\mathbf{E}\dot{\boldsymbol{\varepsilon}}) \leq 0 \end{cases} \quad (14)$$

For additional details, see [13].

4 APPLICATION CASE: 3D SUBSIDENCE ANALYSIS ON REGIONAL SCALE

Surface subsidence due to extraction of underground fluids (water, hydrocarbons) plays an important role in reservoir engineering. For decades a great deal of attention has been directed towards modelling this phenomenon, also because it affects historical cities, like Venice and Ravenna in Italy [3-5, 15, 16, 18, 24-25]. Subsidence analyses are computationally intensive by involving problems of regional scale and very long time spans: e.g. in the case of the Groningen gas field, subsidence predictions for the year 2050 have been made from the year 1973 on. The subsidence surfaces have been obtained with different models and codes, e.g. in [26] with ECLIPSE [27]; in [25] with a quasi three-dimensional hydrologic model and a three-dimensional uncoupled structural model and in [9, 18] with a fully coupled consolidation model.

Apart from exceptions (as e.g. in [9]), three-dimensional subsidence models have assumed a linear elastic response for the solid skeleton and not much has been done for modelling possible interactions in case several reservoirs at different levels are distributed over a large area, as it is the case of the Northern Adriatic region, Italy [28] (**Figure 1**). Here the pools' depth of burial ranges between 900 and 4000 m and the horizontal area involved is about 19000 Km². In addition, the different pools are not scheduled to be put in production at the same time [30], which complicates the situation further.



Figure 1: Location of gas pools in the Northern Adriatic Sea [29].

At present, creep, plasticity [28] and capillary effects [3-5] are envisaged among the possible processes yielding a retarded sinking over the reservoir; as already explained before, the choice here is to refer to an unconventional elastoplastic model, also being its potentialities independent on the specific application case treated here.

The effects of the exploitation of four of the gas reservoirs shown in **Figure 1**, located at three different depths and undergoing different production histories [30], are here analysed; the region covers an area of 40×40 km² and has a depth of 1300 m; it is discretized by about 500 20-node isoparametric elements (additional results, not reported here for sake of brevity, refer to 980 and 2940 elements as well). Free flux on the horizontal and vertical boundaries of the investigated area is assumed. The main material parameters are shown in **Table 1** [16, 30]; the grains are assumed to be incompressible and the clayey layers to behave in agreement with the subloading surface model when accounting for plasticity effects. The geomechanical data have been obtained through analysis of master-logs at our disposal, which are

representative of the investigated area, whereas the plastic variables have been taken from the previous examples, appropriately scaled to take into account the effect of depth.

Table 1: Material data for subsidence analysis

Soil stratum #	E [MPa]	ν	k_i [m/day]	Depths [m]
1	$1.13 \cdot 10^4$	0.17	0.2208	1300÷1254
2	$1.00 \cdot 10^4$	0.17	$0.865 \cdot 10^{-4}$	1300÷1254
3 & Reservoir # 1	$1.13 \cdot 10^4$	0.17	0.2208	1300÷1254
4	$1.00 \cdot 10^4$	0.17	$0.865 \cdot 10^{-4}$	1300÷1254 & 1300÷1070
5, 7, 9	$1.14 \cdot 10^4$	0.30	0.7985	1254÷1070
6, 8, 10	$0.322 \cdot 10^4$	0.38	$0.865 \cdot 10^{-4}$	1254÷1070
11 & Reservoir # 4	$1.14 \cdot 10^4$	0.30	0.7985	1070÷1027
12	$0.322 \cdot 10^4$	0.38	$0.865 \cdot 10^{-4}$	1027÷860
13 & Reservoirs # 2, 3	$0.898 \cdot 10^4$	0.15	0.9752	860÷848
14	$0.555 \cdot 10^4$	0.37	$0.865 \cdot 10^{-4}$	848÷600
15	$0.224 \cdot 10^4$	0.39	$0.865 \cdot 10^{-4}$	600÷0

As evidenced by **Table 1** and **Figure 2**, some planimetric variability for the soil strata has been additionally introduced just to be closer to the real configuration of the subsoil, e.g. considering the available seismic section of [30]; so 7 *macro-levels* are present, including 15 different soil strata. The horizontal projection of the investigated pools can be seen in **Figure 3**, together with the mean radius of their productive levels. The exploitation points (wells) are assumed to be equally distributed above each reservoir such as to allow for the assumption of a constant drop of pressure inside it. The pressure histories (**Figure 4**), obtained from previous reservoir simulators (starting from available gas production records developing in 10 years) are applied as boundary conditions to the nodes of each pool. A computationally more expensive alternative would be to apply the outflow given from the production schedule (if available).

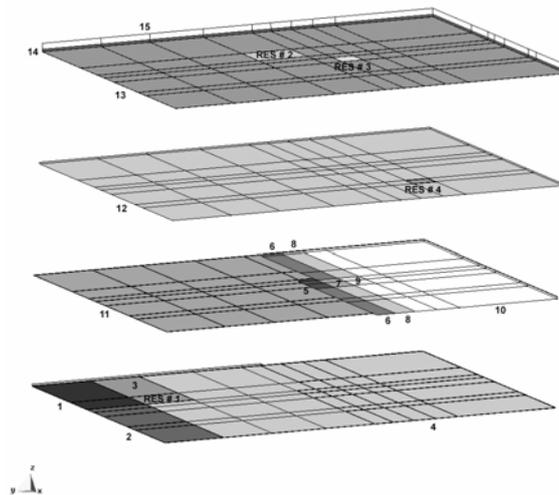


Figure 2: Schematic representation of the soil strata distribution: macro-levels are superimposed from surface (top) to bottom (see **Table 1**).

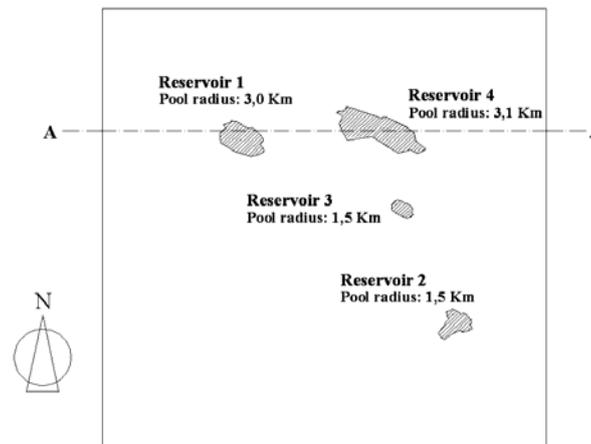


Figure 3: Horizontal projection of the investigated reservoirs for subsidence analysis.

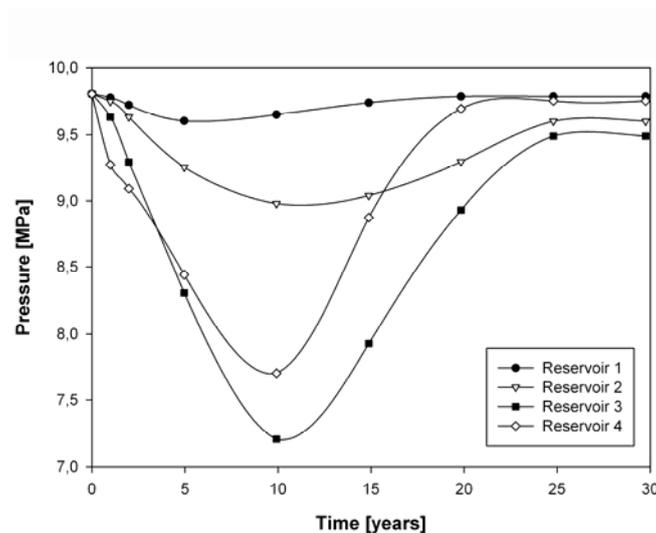


Figure 4: Reservoirs' pressure histories.

The analysis has been pushed up to 30 years from the beginning of exploitations, when a general pressure recovery has already been attained (**Figure 4**); the results in terms of surface subsidence above each reservoir are shown in **Figure 5**, accounting for linear elasticity and unconventional elasto-plasticity as well. The effect of interaction among the different reservoirs can be seen from the shifting in time of the maximum value of subsidence as compared with the minimum of reservoir pressure: this phenomenon is also to be partly ascribed to the presence of clay layers confining the pools, but it is particularly evident when plasticity is introduced: as an extreme situation, maximum subsidence can not be reached even after 30 years; a “residual” delayed land subsidence has clearly appeared, so confirming the usefulness of the proposed unconventional plasticity model for modelling continuing surface settlements when reservoir pore pressures stabilize and for additional settlements

occurring even after the end of gas production.

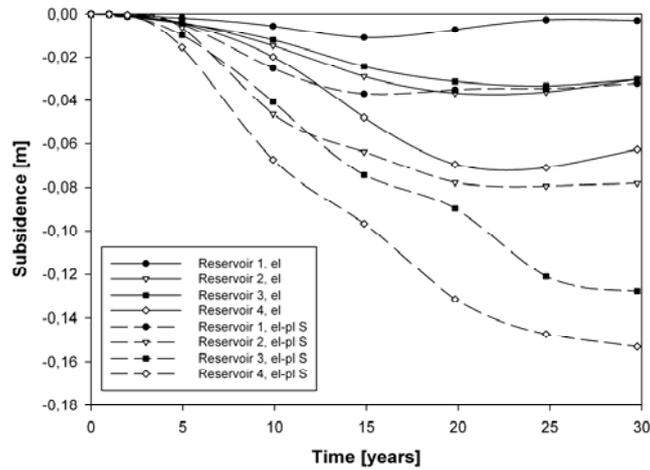


Figure 5: History of surface subsidence above the reservoirs.

The subsidence bowl is depicted in **Figure 6** and **Figure 7**, referring to the evolution of surface subsidence of Section A-A, **Figure 3**; apart from repeating the general trend shown in **Figure 5**, they evidence a subsidence bowl which appears, independently on time, to be slightly wider when elasto-plasticity is accounted for. It is to be underlined that the time scales involved, as well as the orders of magnitude for the resulting subsidence, agree well with what evidenced by [28] and [30] (the latter referring to linear elasticity only), with similar (or equal, as in the latter case) cumulative gas production histories and geological/geomechanical subsoil configurations.

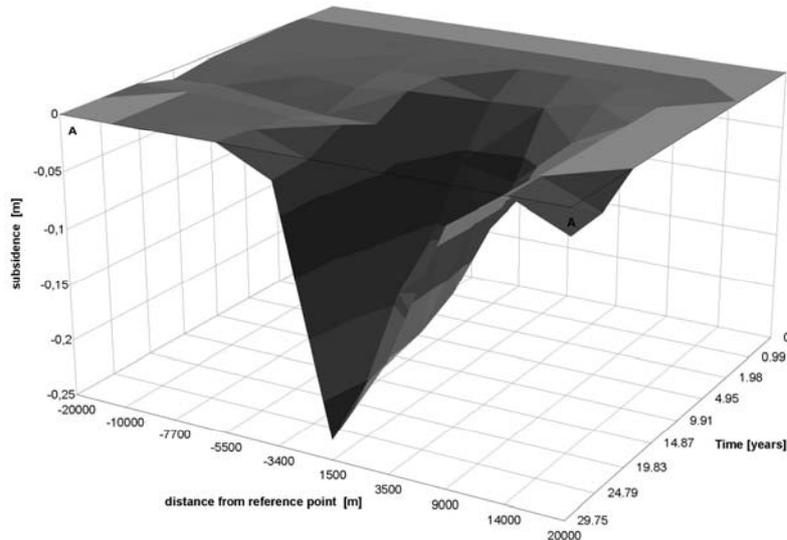


Figure 6: Subsidence bowl, linear elastic case.

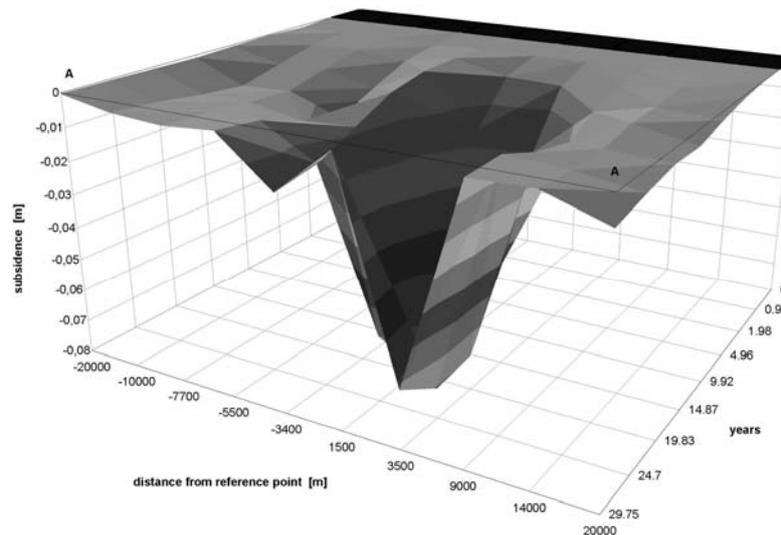


Figure 7: History of surface subsidence above the reservoirs.

5 CONCLUSIONS

The coupled hydro-mechanical state in soils coming from consolidation/subsidence processes and undergoing plasticity phenomena has been evaluated by means of the subloading surface model, allowing for predicting a smooth response for smooth monotonic loading, considering the sign of $\text{tr}(\overline{\mathbf{NED}})$ only in the loading criterion, automatically drawing back of a stress to the normal yield surface even if it goes out from the surface itself. Hence a rough numerical calculation with a large loading step is allowed and return-mapping iterative techniques can subsequently be skipped, so enhancing speedup and efficiency of large scale coupled analyses, as required when modelling subsidence in 3D domains and for long-term scenarios. The plasticity algorithm has been implemented in the PLASCON3D FE code, coupling hydro-thermo-mechanical fields within a saturated (locally partially saturated) porous medium subjected to external loads and water/gas withdrawals from deep layers (aquifers/reservoirs).

The plastic deformation due to the change of stress inside the yield surface exhibiting a smooth elastic-plastic transition has been described, as well as a first ability of describing softening behaviours has been shown.

Regional subsidence analyses due to gas extractions have been possible with reduced computational efforts when introducing unconventional elasto-plasticity in the code. It has been demonstrated that the time scales involved, as well as the orders of magnitude for the resulting subsidence, agree well with what evidenced by [28] and [30] (the latter referring to linear elasticity only), with similar (or equal, as in the latter case) cumulative gas production histories and geological/geomechanical subsoil configurations. Particularly, the effects of interaction among exploitations have been estimated, as well as the phenomenon of residual land subsidence near abandoned gas fields has been successfully modelled: the estimation of this delayed environmental cost of gas pumping is generally neglected, whereas it clearly

appears of being fundamental for an increased awareness of the consequence that gas production may have on future coastline stability relatively far from the gas field [28].

REFERENCES

- [1] Poland J. (Ed.). *Guidebook to Studies of Land Subsidence due to Groundwater withdrawal*. UNESCO, Paris (1984).
- [2] Terzaghi K. Die Berechnung der Durchlässigkeitsziffer des Tones aus dem Verlauf der hydrodynamischen Spannungserscheinungen. *Akademie der Wissenschaften in Wien, Sitzungsberichte, Mathematisch-naturwissenschaftliche Klasse, Part Iia* (1923) **132**(3/4): 125-138.
- [3] Schrefler BA, Bolzon G, Salomoni V, Simoni L. On compaction in gas reservoirs. *Atti dell'Accademia Nazionale dei Lincei - Rendiconti Lincei: Scienze Fisiche e Naturali*, s. IX (1997) **VIII**(4): 235-248.
- [4] Simoni L, Salomoni V, Schrefler BA. Elastoplastic subsidence models with and without capillary effects. *Comp. Meth. Appl. Mech. Engrg.* (1999) **171**(3-4): 491-502.
- [5] Menin A, Salomoni VA, Santagiuliana R, Simoni L, Gens A, Schrefler BA. A mechanism contributing to subsidence above gas reservoirs. *Int. J. Comp. Meth. Engrg. Sci. Mech.* (2008) **9**(5): 270-287.
- [6] Drucker DC. Conventional and unconventional plastic response and representation. *J. Appl. Mech. Rev.* ASME (1988) **41**(4):151-167.
- [7] Hashiguchi K, Ueno M. Elastoplastic constitutive laws of granular materials. *Proceedings 9th Int. Conf. Soil Mech. Found. Engrg.*, Special Session 9, Tokyo, Japan (1977): 73-82.
- [8] Hashiguchi K. Constitutive equations of elastoplastic materials with elastic-plastic transitions. *J. Appl. Mech.* ASME (1980) **47**(2): 266-272.
- [9] Hashiguchi K. Subloading surface model in unconventional plasticity. *Int. J. Solids Struct.* (1989) **25**(8): 917-945.
- [10] Hashiguchi K, Saitoh K, Okayasu T, Tsutsumi S. Evaluation of typical conventional and unconventional plasticity models for prediction of softening behaviour of soils. *Geotech.* (2002) **52**(8): 561-578.
- [11] Hashiguchi K. *Elastoplasticity theory*. In: F Pfeiffer, P Wriggers (Eds.), *Lecture notes in applied and computational mechanics*. Springer: Berlin (2009) **42**: 393 p.
- [12] Yale DP. Coupled geomechanics-fluid flow modelling: effects of plasticity and permeability alteration. *SPE/ISRM Rock Mechanics Conference*. Irving, TX, USA, October 20-23 (2002) **SPE/ISRM 78202**: 10 p.
- [13] Salomoni V.A., Fincato R. 3D subsidence analyses above gas reservoirs accounting for an unconventional plasticity model. *Int. J. Num. Anal. Meth. Geomech.* (2011) (in press).
- [14] Lewis RW, Schrefler BA. *A finite element simulation of the subsidence of gas reservoirs undergoing a waterdrive*. In: RH Gallagher, DH Norrie, JT Oden, OC Zienkiewicz (Eds.), *Finite Element in Fluids*. Wiley: Chichester (1982) **4**: 179-199.
- [15] Schrefler BA, Lewis RW, Majorana CE. Subsidence above volumetric and waterdrive gas reservoirs. *Int. J. Num. Meth. Fluids* (1981) **1**(2): 101-15.
- [16] Lewis RW, Schrefler BA. *The finite element method in the deformation and consolidation of porous media*. Wiley: Chichester (1987).
- [17] Lewis RW, Schrefler BA. *The finite element method in the static and dynamic*

- deformation and consolidation of porous media*. Wiley: Chichester (1998).
- [18] Schrefler BA, Wang X, Salomoni V, Zuccolo G. An efficient parallel algorithm for three-dimensional analysis of subsidence above gas reservoirs. *Int. J. Num. Meth. Fluids* (1999) **31**(1): 247-60.
- [19] Ferronato M, Gambolati G, Teatini P. On the role of reservoir geometry in waterdrive hydrodynamics. *J. Pet. Sci. Engrg.* (2004) **44**: 205-221.
- [20] Hashiguchi K. Mechanical requirements and structures of cyclic plasticity models. *Int. J. Plasticity* (1993) **9**(6): 721-748.
- [21] Hashiguchi K. The extended flow rule in plasticity. *Int. J. Plasticity* (1997) **13**(1): 37-58.
- [22] Hashiguchi K. Fundamentals in constitutive equation: continuity and smoothness conditions and loading criterion. *Soils Found.* (2000) **40**(3): 155-161.
- [23] Hashiguchi K. On the loading criterion. *Int. J. Plasticity* (1994) **10**(8): 871-878.
- [24] Carbognin L, Gatto P, Mozzi G, Gambolati G. *Land subsidence of Ravenna and its similarities with the Venice case*. In: SK Saxena (Ed.), *Evaluation and Prediction of Subsidence*. ASCE: NY, USA (1978): 254-266.
- [25] Gambolati G, Ricceri G, Bertoni W, Brighenti G, Vuillermin E. Mathematical simulation of the subsidence of Ravenna. *Water Res. Research* (1991) **27**(11): 2899-2918.
- [26] Comerlati A, Ferronato M, Gambolati G, Putti M, Teatini P. Fluid-dynamic and geomechanical effects of CO₂ sequestration below the Venice Lagoon. *Env. Engrg. Geosc.* (2003) **XII**(3): 211-226.
- [27] Schlumberger-GeoQuest. *Eclipse Simulator*. Version 2003A. Technical Description, Houston, TX, USA (2003).
- [28] Baù D, Gambolati G, Teatini P. Residual land subsidence near abandoned gas fields raises concern over Northern Adriatic coastland. *Eos* (2000) **81**(22): 245-249.
- [29] *Il Sole 24 Ore*. July 16 (2008) **195** (in Italian).
- [30] AGIP. *Progetto Alto Adriatico – Studio di impatto ambientale*. AGIP, San Donato, Italy (1996) (in Italian).

APPLICATION OF THE PROPER GENERALIZED DECOMPOSITION METHOD TO A VISCOELASTIC MECHANICAL PROBLEM WITH A LARGE NUMBER OF INTERNAL VARIABLES AND A LARGE SPECTRUM OF RELAXATION TIMES

M. HAMMOUD*, M. BERINGHIER* AND J.C. GRANDIDIER*

* Institut P' -Département Physique et Mécanique des Matériaux
UPR CNRS 3346 - ENSMA

Téléport 2, 1 avenue Clément Ader, BP 40109, F-86961 Futuroscope Cedex, France.

E-mail: {hammoudm, marianne.beringhier, J.C.grandidier}@ensma.fr,

Web page: <http://www.ensma.fr>

Key words: PGD, Internal variables, Relaxation time.

Abstract. We here extend the use of the PGD to the case of a viscoelastic mechanical problem with a large number of internal variables and with a large spectrum of relaxation times. Such a number of internal variables leads to solving a system of non linear differential equations which correspond to the return to the equilibrium state. The feasibility and the robustness of the method are discussed in a simple case; a future application is the simulation of a polymer reaction under cyclic loading.

1 INTRODUCTION

To solve a problem with a large number of degrees of freedom (dofs), numerical techniques, as parallel computing and domain decomposition, can be used. In the case of a multiphysical problem or of a problem with a large number of internal variables, it leads to solving a large number of differential equations.

The PGD method, based on the radial approximation [4], has proved to be efficient for solving problems with a large number of dofs [2, 3], and particularly in the case of a coupled thermo mechanical problem [1].

We here extend the use of the PGD [1] to the case of a viscoelastic mechanical problem with a large number of internal variables and with a large spectrum of relaxation times (50 to 100 times [5]). Such a number of internal variables leads to solving a system of non linear differential equations which correspond to the return to the equilibrium state. The feasibility and the robustness of the method are discussed in a simple case; a future application is the simulation of a polymer reaction under cyclic loading.

Section 2 introduces the equations of the viscoelastic mechanical problem. In section 3, the PGD is used to solve the problem with internal variables. While in section 4, we present the numerical results of a problem with one internal variable but with different relaxation times.

2 RESOLUTION OF VISCOELASTIC PROBLEM WITH INTERNAL VARIABLES

2.1 Equations of problem

Let us consider a one-dimensional problem in space (noted x). The generic form of the mechanical problem with internal variables is written as:

$$\frac{\partial \sigma}{\partial x} + f = 0 \quad (1)$$

$$\frac{dz_j}{dt} = -\frac{1}{\tau_j} (z_j - z_j^\infty) \quad (2)$$

Where

$$\sigma = E_v \frac{\partial u}{\partial x} - \sum_{j=1}^m z_j \quad (3)$$

$$z_j^\infty = E_{ij}^\infty \frac{\partial u}{\partial x} \quad (4)$$

Where z_j^∞ represents the relaxed status of the internal variable z_j ; τ_j is the relaxation time; E_{ij}^∞ is the relaxed Young's modulus and E_v is the vitreous Young's modulus of the material.

The vitreous Young's modulus depends on the E_∞ modulus and the relaxed one E_r and its form is: $E_v = E_\infty + \sum_j^m E_{ij}^\infty$ where $E_r = \sum_j^m E_{ij}^\infty$.

Problems (1)-(2) are assumed to be defined on the domain: $\Omega = \Omega_x \times \Omega_t$, where $\Omega_x = [0, L_x]$ and $\Omega_t = [0, L_t]$. The initial conditions are equal to zero and the boundary conditions are written as: $\underline{\underline{\sigma}} \cdot \underline{\underline{n}} = \underline{\underline{F}}$ on $\partial\Omega_f$ and $\underline{\underline{u}} = \underline{\underline{0}}$ on $\partial\Omega_u$.

Remark: The subscript j concerns the internal variables, and it varies from 1 to m , where m is the number of internal variables. If we consider a problem with m internal variables, it means that the equation (2) is reported m times. The specificity of each equation is related to the relaxation time of this internal variable. The displacement field and the internal variable are then coupled.

2.2 Use of the PGD to solve the viscoelastic problem

The aim of the separated representation method is to compute N couples of functions $\{(A_i, B_i), (C_i, D_i), i = 1, \dots, N\}$ such that $\{A_i, C_i, i = 1, \dots, N\}$ and $\{B_i, D_i, i = 1, \dots, N\}$ are defined respectively in Ω_x and Ω_t , and the solutions u and z of the coupled problem can be written in the following separated form:

$$\begin{cases} u(x,t) \approx \sum_{i=1}^N A_i(x)B_i(t) \\ z_j(x,t) \approx \sum_{k=1}^N C_{kj}(x)D_{kj}(t) \end{cases} \quad (5)$$

Once N separated variables functions are computed, the next separated variables functions to be computed are called $R(x)$, $S(t)$, $V_j(x)$ and $W_j(t)$ in this step. They are solutions of the Galerkin variational formulation related to equations (1) and (2):

$$\int_{\Omega} (div \sigma + f) u^* d\Omega + \int_{\partial\Omega_f} (\sigma(l) - F) u^* d\Omega = 0 \quad \forall u^* \quad (6)$$

$$\int_{\Omega} \left(\frac{dz_j}{dt} + \frac{1}{\tau_j} (z_j - z_j^\infty) \right) z_j^* d\Omega = 0 \quad \forall z_j^* \quad (7)$$

With the trial and test fields written as follows:

$$u(x,t) = \sum_{i=1}^n A_i(x)B_i(t) + R(x)S(t), \quad z_j(x,t) = \sum_{i=1}^n C_{ij}(x)D_{ij}(t) + V_j(x)W_j(t) \quad (8)$$

$$u^*(x,t) = R^*(x)S(t) + R(x)S^*(t), \quad z_j^*(x,t) = V_j^*(x)W_j(t) + V_j(x)W_j^*(t) \quad (9)$$

2.3 Specificities related to the internal variables

The displacement field and the internal variables are completely integral. We choose to solve equation (6) and then equation (7) at each step of enrichment in order to decouple the problem. The process is initialized with the solution of an elastic problem without internal variable. Once the displacement field is computed (6), the solution is introduced in equation (4) in order to compute the internal variable in equation (2). This internal variable is placed in equation (3) to compute the solution of the viscoelastic problem. This process is iterated until convergence.

In a step of enrichment, the method of fixed point is used to compute the functions $R(x)$, $S(t)$ and then $V_j(x)$ and $W_j(t)$. In what follows, the equations to compute these functions are:

For $R(x)$:

$$E_v \beta_t \frac{\partial^2 R}{\partial x^2} = -E_v \sum_{i=1}^n \alpha_i^t \frac{\partial^2 A_i}{\partial x^2} + \sum_{i=1}^n \sum_{j=1}^m \gamma_i^j \frac{\partial C_{ij}}{\partial x} + \sum_{j=1}^m \delta_t^j \frac{\partial V_j}{\partial x} - \theta_t \quad (10)$$

For $S(t)$:

$$E_v \beta_x S(t) = -E_v \sum_{i=1}^n \alpha_x^i B_i(t) + \sum_{i=1}^n \sum_{j=1}^m \gamma_x^j D_{ij}(t) + \sum_{j=1}^m \delta_x^j W_j(t) - \theta_x \quad (11)$$

For $V_j(x)$:

$$\left(\gamma_{ij}^j + \frac{1}{\tau_j} \theta_t^j\right) V_j = - \sum_{i=1}^n \beta_i^i C_{ij} - \frac{1}{\tau_j} \sum_{i=1}^n \delta_{ij}^i C_{ij} + \frac{B}{\sqrt{\tau_j}} \sum_{i=1}^n \beta_{ij}^i \frac{\partial A_i}{\partial x} + \frac{B}{\sqrt{\tau_j}} \frac{\partial R}{\partial x} \Gamma_t^i \quad (12)$$

For $W_j(t)$:

$$\frac{\partial W_j}{\partial t} \gamma_{xj}^j + \frac{1}{\tau_j} W_j \gamma_{xj}^j = - \sum_{i=1}^n \beta_{xj}^i \frac{\partial D_{ij}}{\partial t} - \frac{1}{\tau_j} \sum_{i=1}^n \beta_{xj}^i D_{ij} + \frac{B}{\sqrt{\tau_j}} \sum_{i=1}^n \alpha_{xj}^i B_i + \frac{B}{\sqrt{\tau_j}} S \delta_{xj}^i \quad (13)$$

Where :

$$\begin{aligned} \beta_t &= \int_{\Omega_t} S^2(t) dt, \alpha_t^i = \int_{\Omega_t} B_i(t) S(t) dt, \gamma_t^j = \int_{\Omega_t} D_{ij}(t) S(t) dt, \delta_t^j = \int_{\Omega_t} W_j(t) S(t) dt, \theta_t = \int_{\Omega_t} (\sigma(t) - F) S(t) dt. \\ \beta_x &= \int_{\Omega_x} R(x) \frac{\partial^2 R(x)}{\partial^2 x} dx, \alpha_x^i = \int_{\Omega_x} R(x) \frac{\partial^2 A_i(x)}{\partial^2 x} dx, \gamma_x^j = \int_{\Omega_x} R(x) \frac{\partial C_{ij}(x)}{\partial x} dx, \delta_x^j = \int_{\Omega_x} R(x) \frac{\partial V_j(x)}{\partial x} dx \\ \theta_x &= \int_{\Omega_x} (\sigma(t) - F) R(x) dx, \theta_t^j = \int_{\Omega_t} W_j^2(t) dt, \beta_t^j = \int_{\Omega_t} \frac{\partial D_{ij}(t)}{\partial x} W_j(t) dt, \delta_{ij}^j = \int_{\Omega_t} D_{ij}(t) W_j(t) dt, \alpha_{ij}^j = \int_{\Omega_t} B_i(t) W_j(t) dt, \\ \Gamma_t^j &= \int_{\Omega_t} S(t) W_j(t) dt, \gamma_{ij}^j = \int_{\Omega_t} \frac{\partial W_j(t)}{\partial x} W_j(t) dt, \gamma_{xj}^j = \int_{\Omega_x} V_j^2(x) dx, \beta_{xj}^j = \int_{\Omega_x} C_{ij}(x) V_j(x) dx, \\ \alpha_{xj}^i &= \int_{\Omega_x} \frac{\partial A_i(x)}{\partial x} V_j(x) dx, \delta_{xj}^i = \int_{\Omega_x} \frac{\partial R(x)}{\partial x} V_j(x) dx. \end{aligned}$$

After each iteration l , two residuals are computed; R_u related to equation (1) and R_{z_j} related to equation (2).

$$R_u = \frac{\|u_{l+1} - u_l\|}{\|u_l\|} \leq \varepsilon \quad (14)$$

$$R_{z_j} = \max \left(\frac{\|z_{j,l+1} - z_{j,l}\|}{\|z_{j,l}\|} \right) \leq \varepsilon \quad (15)$$

Where $\| \cdot \|$ stands for the L^2 norms. The iterative procedure is stopped when $\max(R_u, R_{z_j})$ is small enough. The solution of the problem is then given by equation (5).

3 APPLICATION TO A PROBLEM WITH ONE INTERNAL VARIABLE

The simulation test is a 10 mm long one-dimensional bar subjected to a load $F(x,t) = G(x=10) \times H(t)$ at $x = 10$ mm as shown in figure 1 and with null boundary conditions at $x = 0$ mm. The time L_t equals 10 s. The parameters of material are given in the table 1. The time step equals 0,1 sec, and the space one equals 0,25 mm. We here consider a problem with one internal variable with different relaxation times and study its influence on the displacement field.

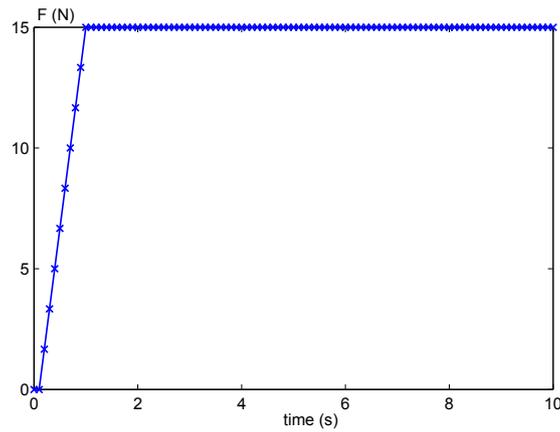


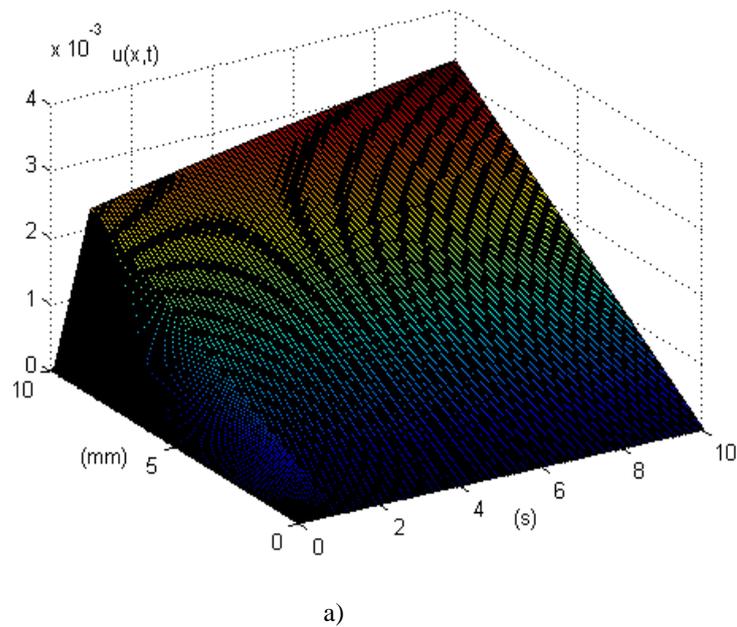
Figure 1: Creep load at the extremity of the bar

Table 1: Mechanical parameters of the material

E_v (MPa)	E_∞ (MPa)
1140	140

3.1 Results and Discussion

For τ equals 1 second, the solution is reached with 10 iterations with $(\|A_i\| \times \|B_i\|) = (0.01 - 0.07 - 0.009 - 0.0025 - 0.0001 - 0.002 - 0.001 - 0.0004 - 0.0001 - 0.00003)$ for the displacement field; and $(\|C_i\| \times \|D_i\|) = (10.41 - 7.53 - 0.85 - 0.19 - 0.03 - 0.21 - 0.09 - 0.03 - 0.009 - 0.002)$ for the internal variable.



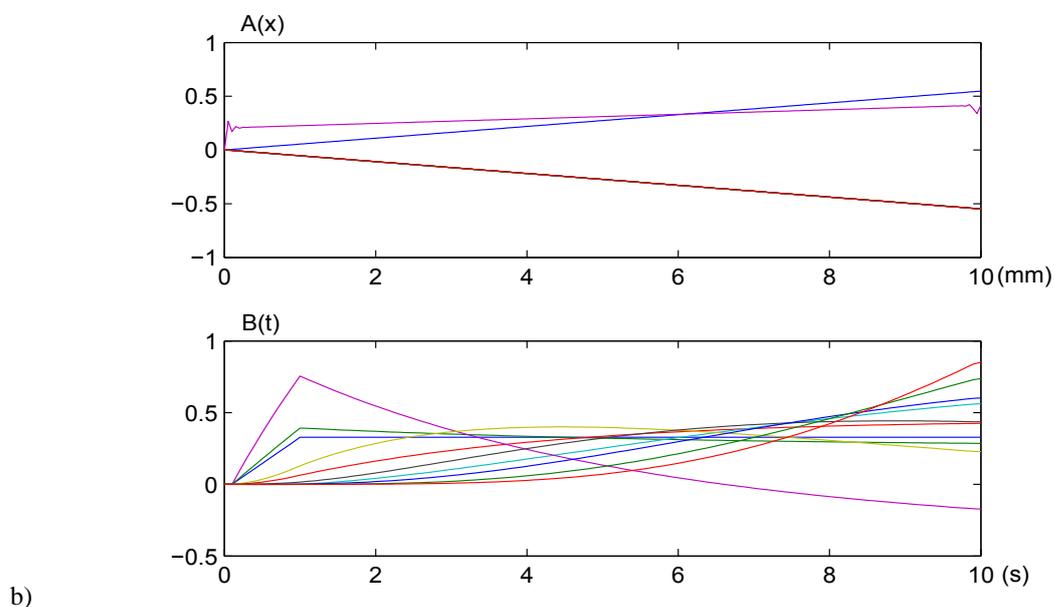


Figure 2: (a) Displacement field and (b) Spatial and Temporal modes.

Figure 2a shows the displacement field computed. Temporal and spatial modes for the displacement field are represented in Figure 2b. Figure 3 shows the convergence of the solution via the residuals (equations (14) and (15)) compared to the iteration number.

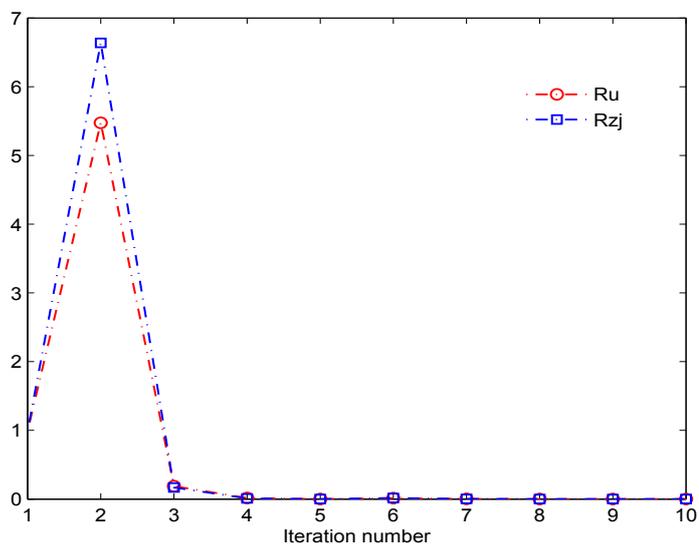
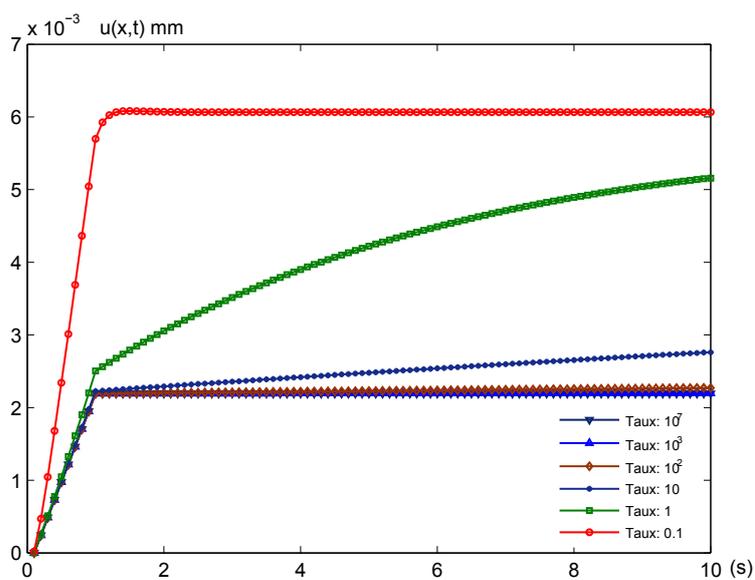
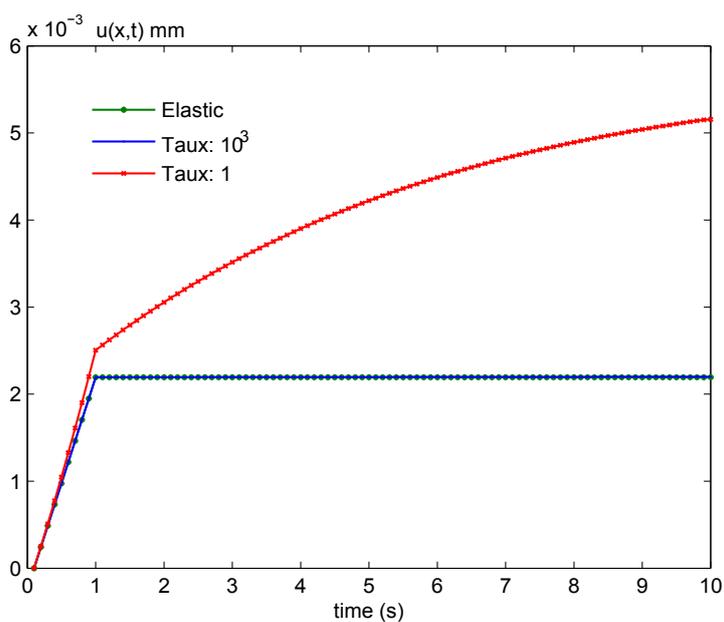


Figure 3: Convergence of the PGD



a)



b)

Figure 4: (a) Influence of relaxation time on the displacement field, (b) and comparison with elastic solution

The same simulation is done with different value of relaxation time. Based on the value of τ , the solution was not reached with the same number of modes which increases when τ decreases. Figure 4a shows the influence of the relaxation time on the displacement field.

We can observe that the influence of the internal variable is important when the value of the relaxation time is weak. Figure 4b shows that when the relaxation time is high, the reached solution is that of the elastic problem. In that case, fewer modes are needed to represent the solution. In table 2, we show the number of iterations with respect to the variation of the relaxation time τ .

Table 2: Influence τ of on the number of modes

τ (s)	0.1	1	10	100	1000
iteration	15	10	9	5	4

4 CONCLUSION

In this work, the PGD was validated in the case of a viscoelastic problem with one internal variable. The solution was computed with different relaxation time, and we observed that the number of modes needed to describe the displacement field depends on the relaxation time.

These results are encouraging and will be extended to the case with many internal variables and relaxation time in order to simulate the behavior of the *PE-HD* polymer under cyclic loading. It will be our future application.

REFERENCES

- [1] M. Béringhier, M. Gueguen, J.C. Grandidier. Solution of strongly coupled multiphysics problems using space-time separated representations – Application to thermoviscoelasticity, *Archives of Computational Methods in Engineering*, (2010), doi: 10.1007/s11831-010-9050-5.
- [2] F. Chinesta, A. Ammar, E. Cueto. Proper generalized decomposition of multiscale models, *International Journal for Numerical Methods in Engineering*, (2010), Volume 83, 1114-1132,
- [3] A. Ammar, B. Mokdad, F. Chinesta, R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids - Part II: Transient simulation using space-time separated representations, *Journal of Non-Newtonian Fluid Mechanics*, (2007), 144, 98-121.
- [4] P. Ladevèze. Nonlinear computational structural mechanics - New approaches and non-incremental methods of calculation, (1999), *Springer Verlag*.
- [5] Ch. Cnat. A thermodynamic theory of relaxation based on a distribution of non-linear processes, *Journal of Non-Crystalline Solids*, Volumes 131-133, Part 1, (1991), Pages 196-199.

THERMO-MECHANICAL MODELLING OF DYNAMIC TENSILE EXTRUSION TEST

L. PERONI^{*}, M. SCAPIN^{*} AND M. AVALLE^{*}

^{*} Dipartimento di Meccanica (DIMEC)
Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
e-mail: lorenzo.peroni@polito.it, web page: <http://www.polito.it>

Key words: Dynamic Tensile Extrusion, Johnson-Cook, Zerilli-Armstrong, Steinberg-Guinan, Mechanical Threshold Stress, thermo-mechanical, LS-DYNA.

Abstract. Recently, the researchers of the Los Alamos National Laboratory (LANL) have been developed a technique to investigate high strain-rates in which a sphere is launched at high speed (300-800 m/s) through a conical die (Dynamic Tensile Extrusion test).

The system can use a set-up similar to a Taylor test (like a light gusgun) but induces in the material high levels of tensile stresses and allows to investigate the fracture and damage of the material in these extreme conditions. Due to the high strain-rates and hydrostatic component of stress that occur during the extrusion in the die, the material is also subject to a rapid heating.

In this work, the DTE test has been simulated with the explicit Lagrangian code LS-DYNA. In order to correctly reproduce the various aspects of the test it is necessary to use an elasto-plastic material model that takes into account strain-rate, thermal softening, pressure influence and damage and that adequately describes the thermal coupling effects and the equation of state (EOS) of the material. Besides, due to the high distortion of the Lagrangian mesh, the use of a mesh adapting technique is absolutely necessary.

1 INTRODUCTION

A lot of different experimental techniques can be applied to identify material mechanical properties in function of loading conditions, stress/strain distribution, pressure influence, thermal contribution, wave or shockwave propagation and, finally, temperature and strain-rate ranges. In general, it is not possible to completely characterize a material with a single experimental test because each test is limited in strain-rate. At low strain-rate (from 10^{-5} s^{-1} to 10^{-1} s^{-1}) the material transformation is considered to be isotherm without any propagation and dynamic phenomena. The experimental test can be conducted on a standard electro-mechanical or servo-hydraulic machine. In the medium strain-rate test (from 10^{-1} s^{-1} to 10^2 s^{-1}) the deformation process is still isotherm but the dynamic component of the response starts to be relevant. In this case the tests are performed via pneumatic machine and drop fall. Finally, in the experimental characterization under high value of strain-rate (from 10^2 s^{-1}) the wave propagation (elastic, plastic or shock waves) is very important and the phenomenon evolution is adiabatic. This implies the overheating of the specimen due to the heat conversion of a

percentage of the mechanical work since the phenomenon duration does not allow the conduction or convection heat dispersion. The experimental test to perform in these conditions are, for example, compression and tensile test on Split Hopkinson Bar, the Taylor test, the flyer impact test and the Dynamic Tensile Extrusion (DTE).

In this work the attention is focused on the material characterization under high value of strain-rate and plastic strain. These input conditions are get via a Dynamic Tensile Extrusion: the experimental concept was developed by Gray III et al. at the Los Alamos National Laboratory [1]. The main objective is the development of a suitable numerical method able to reproduce all the features of the experimental test: quite relevant pressure dependence, strain and strain-rate hardening and coupled thermal softening. Obviously, the choice of the material models used for the mechanical behaviour description is correlated to the goodness of the model to be able to describe these fundamental aspects. In this work the completely empirical Johnson-Cook model, the semi-empirical Steinberg-Guinan and Zerilli-Armstrong models and the phenomenological Mechanical Threshold Stress model are taken into account for what concerns the deviatoric behaviour in addition to a Mie-Grüneisen equation of state for the description of the hydrodynamic behaviour. The numerical simulations are performed with the FEM code LS-DYNA [2], which includes implicit and explicit solver with coupled thermo-mechanical and high non-linear capabilities. The numerical analysis is performed on copper spherical projectile and the numerical results are compared with the experimental results obtained in [1,3,4] and the numerical results obtained via the FEM code MARC in [5].

2 EXPERIMENTAL TEST DESCRIPTION

The Dynamic Tensile Extrusion (DTE) concept was proposed by Gray III et al. in [1]. The test is suitable to evaluate the material mechanical response in case of high plastic strain and strain-rate. These conditions are obtained via a gas-gun facilities without the use of a high-explosive launch system. The experimental results used for the numerical model validation are obtained in [1,3,4].

In case of copper, a spherical projectile (diameter 7.60 mm) is accelerated in a high-pressure helium gas gun at ~400 m/s and extruded through a high-strength steel die (entrance diameter of 7.62 mm, exit diameter of 2.28 mm and angle of 10°). In case of tantalum a spherical projectile (diameter 7.58 mm) is accelerated in a high-pressure helium gas gun at 384 and 553 m/s and extruded through a high-strength steel die (entrance diameter of 7.62 mm, exit diameter of 2.8 mm and angle of 9°).

For both materials the sequence of the segments exiting the die are captured via a high-speed photography [1,3,4] and the results for tantalum are shown in Fig.1 within the reassembly of the recovered segments in sequence of exit. Comparing the results for tantalum specimens at different launch speed it is evident higher velocity implies high elongation and propensity to plastic instabilities of the extruded material. Both in case of copper and tantalum, the material appears to be not fully extruded since the rear portion of the specimen halts in the die.

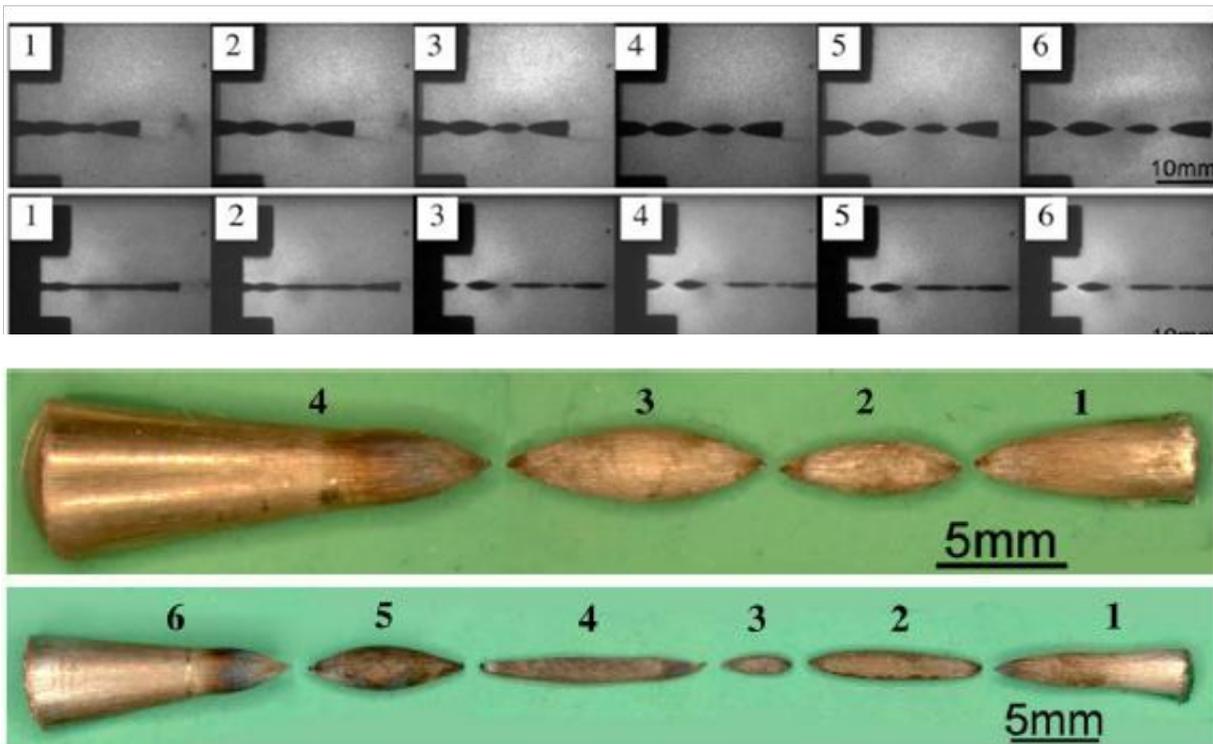


Figure 1: high speed images of the DTE on tantalum launched at 384 m/s and 553 m/s; topology of the recovered extruded Ta sphere launched at 384 m/s and 553 m/s [1,4]

3 NUMERICAL MODEL

The numerical simulations are carried out via the FEM code LS-DYNA [2]. The model is a 2D axis-symmetric model and the solution is obtained by an explicit solver. In order to avoid too high distortion of the Lagrangian mesh, the use of a mesh adapting technique is absolutely necessary (Figure 2).

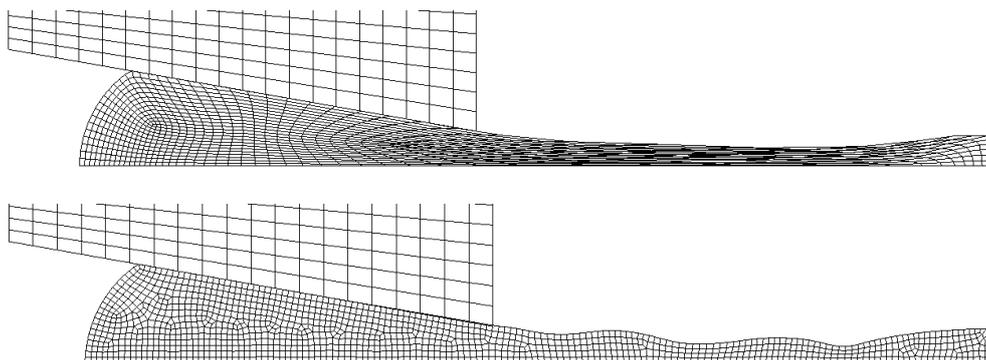


Figure 2: effects of mesh adaptivity on a simulation of DTE: without an adaptive algorithm the mesh is heavy distorted due to the high level of plastic strain reached; with an adaptive algorithm, which remeshes the model at constant time interval, the elements maintain a regular shape and a fixed characteristic dimension

In such kind of problems the choice of the material model (strength model and equation of state) is of fundamental importance: the numerical model has to be able to take into account all the mechanical variables that play a key role in the material deformation process and in the resultant stress and velocity fields. The variable of interest are strain, strain-rate, temperature and pressure.

The phenomenon is strongly thermo-structural coupled (Figure 3): the structural-mechanics material conditions produce the rise in temperature that consequently modifies the mechanical material response. Usually, due to the rise in temperature, there is a decrease of the mechanical strength but also a modification (reduction) in the effect of the strain-rate on the flow stress.

As a matter of fact, the material properties may depend strongly upon the local temperature and for these problems it becomes necessary to compute the temperature throughout the material as time increases. In particular, thermal softening phenomena are essentially due to heat conversion of mechanical work occurring at high strain-rate where strain is localized.

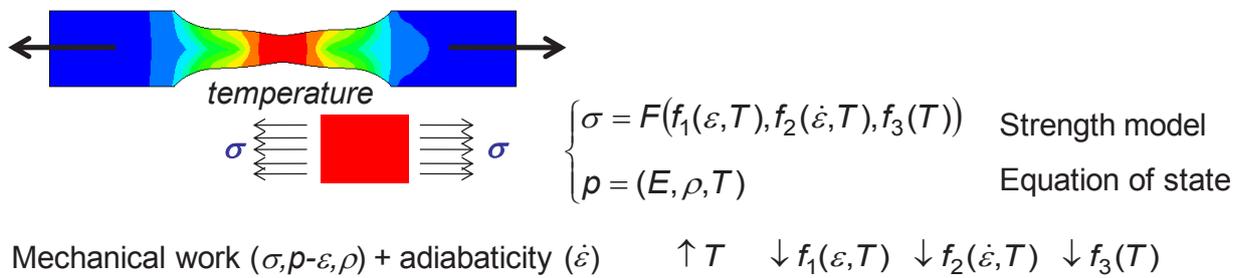


Figure 3: scheme of the thermo-structural coupling

In general, starting from $\dot{\varepsilon} \geq 10^2 \text{ s}^{-1}$ thermal diffusion (conduction and convection) can be neglected and thermal softening can be evaluated under adiabatic assumption. Given this last hypothesis and the further assumption of uniform stress, strain and temperature fields, the material temperature could be analytically computed as a function of plastic work

$$T = T_r + \frac{\beta}{\rho \cdot C_p} \cdot \int_0^{\varepsilon_p} \sigma(\varepsilon, \dot{\varepsilon}, T) \cdot d\varepsilon \quad (1)$$

where ρ is the material density, C_p is the specific heat at constant pressure and β is the Taylor-Quinney coefficient [6,7] that represents the portion of plastic work that is converted into heat (for metals about 0.9). The temperature increase is not uniform but grows up in the material with a distribution corresponding to plastic work.

In a more general manner, assuming that the stress tensor and internal energy are state functions of the strain and the temperature, it is possible to evaluate the temperature increment in the material once the strength model and the equation of state of the material are known.

3.1 Strength material model

In past decades a lot of material models for the description of the deviatoric behaviour are proposed. The classification model makes a distinction between empirical, semi-empirical and

phenomenological models. The empirical models have not physical basis, but are obtained by interpolation of the experimental data. On the other hand the phenomenological models are obtained starting from the transformation in the material occurring during a deformation process.

Table 1: Visco-plastic material models

Johnson-Cook [8]	$\sigma_y = \left(A + B \varepsilon_{pl}^n \right) \left(1 + C \ln \frac{\dot{\varepsilon}_{pl}}{\dot{\varepsilon}_0} \right) \left(1 - \left(\frac{T - T_r}{T_m - T_r} \right)^m \right)$	empirical multiplicative model: temperature T does not influence strain and strain-rate effects (athermal model); no pressure and density influences; melting temperature and shear modulus constant
Zerilli-Armstrong [9]	$\sigma_y = \Delta\sigma_G + B_1 \sqrt{\varepsilon_{pl}} e^{(-\beta_0 + \beta_1 \ln(\dot{\varepsilon}))T} + \frac{k_e}{\sqrt{l}} \quad (\text{FCC})$ $\sigma_y = \Delta\sigma_G + B_0 e^{(-\beta_0 + \beta_1 \ln(\dot{\varepsilon}))T} + K_0 \varepsilon_{pl}^n + \frac{k_e}{\sqrt{l}} \quad (\text{BCC})$	dislocation-based semi-empirical model: athermal and thermal stress components; no pressure and density influences; melting temperature and shear modulus constant
Steinberg-Guinan [10]	$\sigma_y = \sigma_A \frac{G}{G_0}$ $\frac{G}{G_0} = \left[1 + h_1 P v^{1/3} - h_2 (T - 300) \right]$ $\sigma_A = \sigma_0 \left[1 + \beta (\varepsilon_{pl,i} + \varepsilon_{pl}) \right]^n < \sigma_{MAX}$ $T_m = T_{m0} \exp[2a(1-v)] v^{-2(\gamma_0 - a - 1/3)}$	semi-empirical rate independent model (athermal model); shear modulus function of pressure, density and temperature; melting temperature function of density
Mechanical Threshold Stress [12,14]	$\sigma_y = \hat{\sigma}_A + Gs \hat{\sigma}_T$ $s = \left\{ 1 - \left[\frac{kT \ln(\dot{\varepsilon}_0 / \dot{\varepsilon})}{g_0 G b^3} \right]^{1/q} \right\}^{1/p}$ $\theta = (a_0 + a_1 \ln(\dot{\varepsilon}) + a_2 \dot{\varepsilon}) \left[1 - \frac{\tanh(\alpha \hat{\sigma}_T / \hat{\sigma}_S)}{\tanh(\alpha)} \right]$ $\hat{\sigma}_S = \hat{\sigma}_{S0} \left(\frac{\dot{\varepsilon}}{\dot{\varepsilon}_{S0}} \right)^{\frac{Gb^3 A}{kT}}$ $G = G_0 - \frac{G_T}{e^{t/T} - 1}$	dislocation mechanism based model (mechanical threshold stress as internal state variable): athermal and thermal (thermal activation and controlled drag) stress components; no pressure and density influences; shear modulus function of temperature

Models such that proposed by Johnson-Cook (J-C) [8] is purely empirical model and it is one of the most widely used. An example of semi-empirical model is the Zerilli-Armstrong

(Z-A) model [9], that is obtained on the basis of the dislocation mechanics theory and presents a different formulation for BCC and FCC materials. Another example of semi-empirical model is the Steinberg-Cochran-Guinan-Lund (S-C-G-L) model [10,11], which was first developed for the description of high strain-rates behaviour [10], and after extended to low strain-rates [11]. A phenomenological more complex dislocation-based model is the Mechanical Threshold Stress (MTS) model [12].

In general, a visco-plasticity model express the flow stress as a combination of athermal and thermal components. The athermal part of the stress can be in general correlated with strain, strain-rate, temperature and pressure, while the thermal component identifies the temperature influence on the strain-rate effects on the flow stress (temperature-strain-rate coupling).

The J-C model [8] is the most simple model able to predict the mechanical behaviour of the materials under different loading conditions. Besides, as mentioned before, it is one of the most used material model, so it is implemented in many FEM codes and it is quite easy to find in literature the values of J-C parameters for different materials. In the J-C model only the athermal stress component is taken into account and it expresses the flow stress as reported in Table 1. In more details, A is the elastic limit strength, B and n are the work hardening parameters. C and $\dot{\epsilon}_0$ express the strain-rate sensitivity coefficients and m describe the thermal softening. In more details, $\dot{\epsilon}_0$ represents the quasi-static strain-rate threshold, that, ideally, represents the highest strain-rate for which the strain-rate effects on the flow stress are negligible. Finally, the thermal effects are described, in function of the temperature T , the reference temperature at which there are not any thermal effects T_r and the melting temperature at which the material mechanical strength goes to zero T_m . In this condition the material loses its shear strength and starts to behave like a fluid. The J-C model is a multiplicative model, in which the effects of plastic strain, strain-rate and temperature are assumed to act independently. The J-C model is quite inaccurate to model the material behaviour in case of high pressure conditions. As a matter of fact, it neglects the influence of the pressure and changes in volume on the flow stress. Besides, it considers the melting temperature as a constant, while the solid-liquid transition is influenced by the density.

The Z-A [9] model is based on simplified dislocation mechanics, and as the J-C model, it neglects the pressure and densities influences and considers the shear modulus and the melting temperature to be constants. At difference from J-C model it takes into account also the thermal component of the stress. Since it is partially physical-based, it presents a different flow stress formulation for FCC and BCC materials. In reference to the Table 1, in the Z-A models, l represents the polycrystal grain diameter and $\Delta\sigma_g$, B_0 , B_1 , β_0 , β_1 , K_0 , n and k_e are experimental constants based on a dislocation mechanics analysis of the plastic deformation mechanism operative in the two different crystal lattice structures.

A more accurate model is the S-G material model [10] in which both the flow stress and the shear modulus are function of temperature, strain and also pressure. In reference to Table 1, v is the relative volume, β and n are the work-hardening parameters and $\epsilon_{pl,i}$ is the initial equivalent plastic strain, normally equal to zero. The subscript 0 refers to the condition in which $T=300$ K, $P=0$ and $\epsilon=0$, the reference state. h_1 and h_2 are proportional to the derivative of the yield stress respect to the pressure and the derivative of the shear modulus respect to the temperature, respectively. Also for this model, if the temperature overcomes the melting

temperature then the material strength goes to zero, but in this case the melting temperature is not constant but is a density function. The S-G model is strain-rate independent (athermal stress component), according to the fact that for strain-rates greater than $\sim 10^5 \text{ s}^{-1}$ the rapid decrease of rate dependent effects with increasing dynamic stress may be explained by the increase in temperature with increasing stress [10]. This assumption is reasonable in case of FCC material, like copper, but it is quite inaccurate in case of BCC materials [11].

The last material model takes into account in this work is the MTS model [12] which is quite different from the previous ones because it uses the mechanical threshold stress as the internal state variable (instead of the strain) and treats evolution of this variable using physically based and phenomenological based expression. In reference to Table 1, the flow stress is separated into two components: the athermal component $\hat{\sigma}_A$ characterizes the rate independent interactions of dislocations with long-range barriers such as grain boundaries, while the thermal component $\hat{\sigma}_T$ characterizes the rate dependent interaction with short range obstacles. The factor s is related only to the thermally activated controlled glide in according to the results of [13] that indicates at strain-rate lower than $\sim 10^4 \text{ s}^{-1}$ the rate controlling deformation mechanism is thermal activation while the contribution of dislocation drag is negligible. s is evaluated via an Arrhenius rate equation, where $\dot{\epsilon}_0$ is a constant, k is the Boltzmann constant, g_0 is the normalized activation energy, G is the shear modulus, b is the magnitude of the Burger's vector and p and q are constants that characterize the shape of the obstacle profile. In the definition of the strain hardening rate θ , the first terms is the hardening due to dislocation accumulation (a_0 , a_1 and a_2 are material constants) and the second one is the dynamic recovery rate, that is related to the threshold stress at zero-hardening $\hat{\sigma}_S$, correlated to the saturation threshold stress for deformation at 0 K $\hat{\sigma}_{S0}$ ($\dot{\epsilon}_{S0}$, and A are constants). Finally, the shear modulus G temperature dependence is expressed according to the results found in [14].

In addition in LS-DYNA [2] it is possible to combine each material strength model with an element failure model. In particular, it is possible to define one (or more) condition for the deletion (erosion) of the elements: when the condition is reached the element is cancelled from the model. In case of the DTE simulations the condition is imposed on the value of the maximum effective strain.

Similarly, the standard LS-DYNA J-C model provides also a failure model for which the strain at fracture is given by

$$\varepsilon_f = \left(D_1 + D_2 e^{D_3 \sigma^*} \right) \left(1 + D_4 \ln \dot{\epsilon}^* \right) \left(1 + D_5 T^* \right) \quad (2)$$

where D_1 , D_2 , D_3 , D_4 and D_5 are the failure parameters and σ^* is the ratio between the pressure and the effective stress. The material fracture occurs when the damage parameter

$$D = \sum \frac{\Delta \varepsilon_p}{\varepsilon_f} \quad (3)$$

reaches the value of 1.

3.2 Equation of state

The other fundamental aspect to consider is the implementation of an equation of state to simulate the hydrodynamic material behaviour. An EOS is a constitutive relation between state variables and describes the state of the matter. Usually it expresses a thermodynamic variable (such as pressure) in function of two other independent state variables (such as density and internal energy).

In general the equations of state can be divided into two categories: analytical EOS and tabular EOS.

In LS-DYNA [2] the equations of state for metals are in general partitioned into two terms: a cold pressure and a thermal pressure (see equation 4). The first terms $P_C(\mu)$ is a function of the density and is hypothetically evaluated along a 0 K isotherm, while the second one $P_T(\mu, E)$ depends on both the density and the internal energy per unit volume of the system

$$P(\mu, E) = P_C(\mu) + P_T(\mu, E) \quad (5)$$

where $\mu = \rho/\rho_0 - 1$. It is clear that for a compressed material $\mu > 0$, for an expanded material $\mu < 0$ and if no load is applied to the material $\mu = 0$. More precisely the EOS implemented have the general form

$$P(\mu, E) = A(\mu) + B(\mu)E \quad (6)$$

so the pressure is a general function of the density but linearly depending by the energy per unit volume.

The analytical equation of state implemented in LS-DYNA and used in this work is the Mie-Grüneisen EOS, that is able to describe the behaviour of a solid material. It defines the pressure for compressed materials as

$$P = \frac{\rho_0 C_0^2 \mu \left[1 + \left(1 - \frac{\gamma_0}{2} \right) \mu - \frac{a}{2} \mu^2 \right]}{\left[1 - (S_1 - 1)\mu - S_2 \frac{\mu^2}{1 + \mu} - S_3 \frac{\mu^3}{(1 + \mu)^2} \right]^2} + (\gamma_0 + a\mu)E \quad (7)$$

and for expanded materials as

$$P = \rho_0 C_0^2 \mu + (\gamma_0 + a\mu)E \quad (8)$$

where ρ_0 is the initial density, C_0 the sound speed, γ_0 the Grüneisen parameter that defines the effect on the atoms vibration consequent to the change in energy (or temperature), S_i are the coefficients of the polynomial that defines the shock wave velocity as function of particle velocity and a is the first volume correction to γ_0 .

4 NUMERICAL RESULTS AND DISCUSSION

A great number of simulation are performed in order to investigate the effects of the various numerical model aspects on the results of a DTE on copper specimen.

This type of test is used as a benchmark of numerical model capabilities to reproduce a phenomenon in extreme conditions of strain, strain-rate, temperature and pressure. In this preliminary phase, the interest is focused on the understanding of the more important model

parameters and modelling choices to correctly reproduce the real behaviour. In the future, when experimental data of this kind of test will be available, it will be possible to use a combined experimental and numerical approach to obtain the strength model parameters of material under investigation. For this reason, the value of the material model parameters used in these simulations are those reported in the original works of the model authors [8-12].

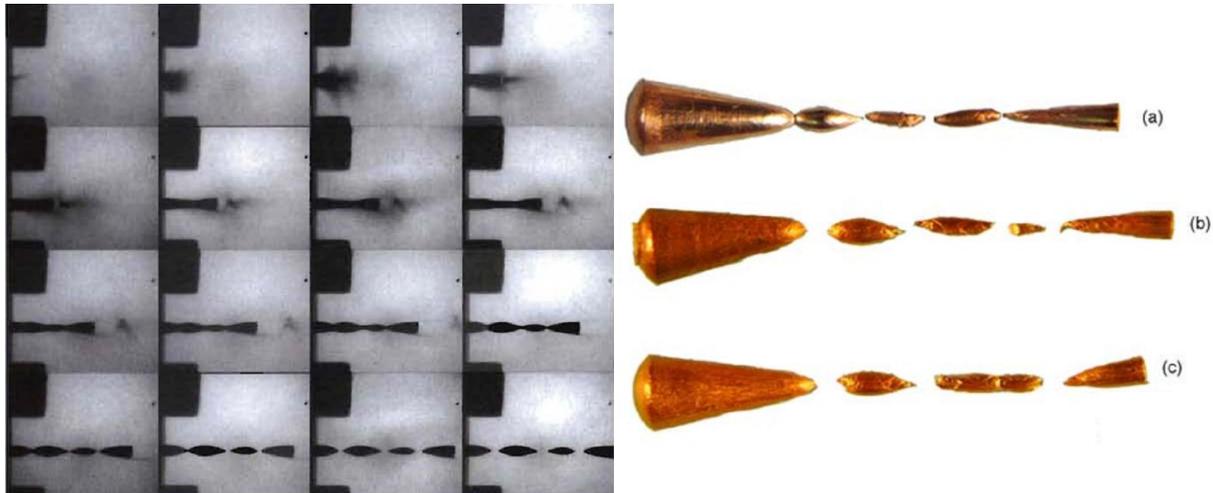


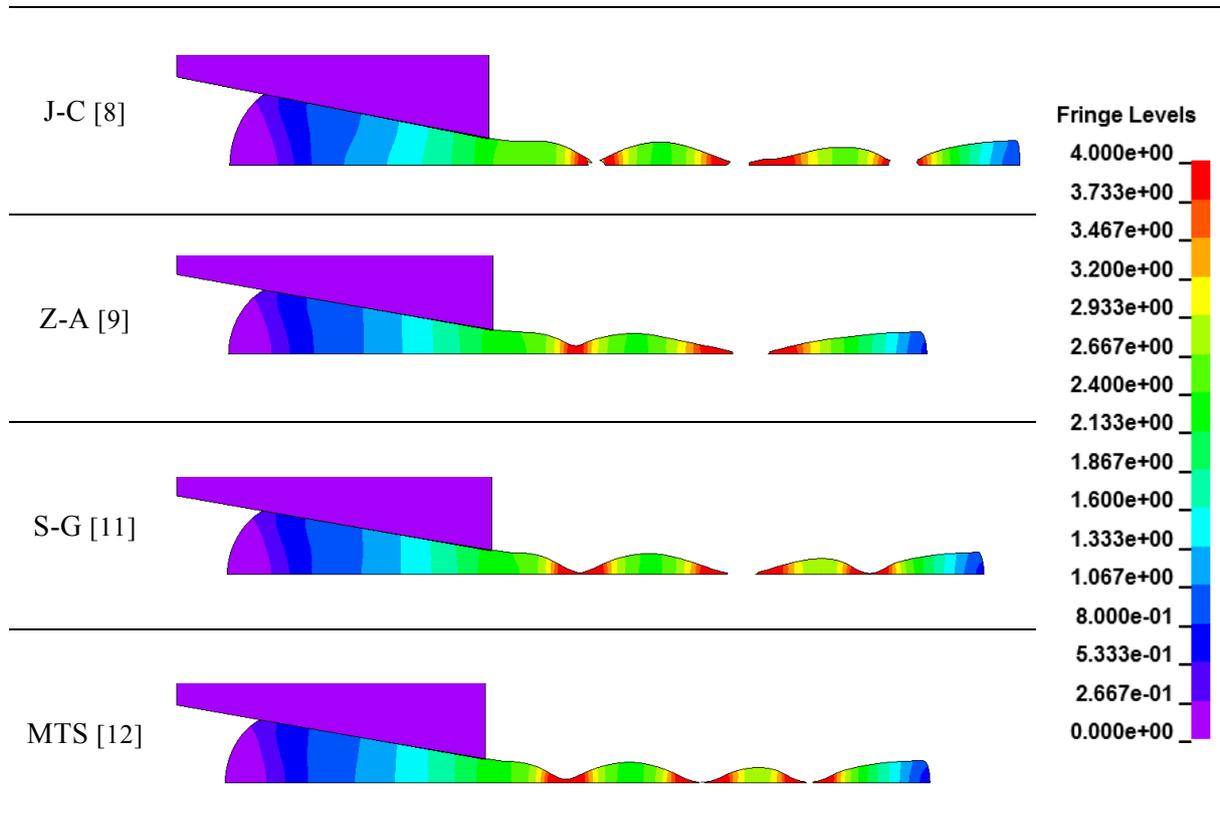
Figure 4: high-speed imaging showing the deformation of an extruded Copper sphere which entered the die at 400 m/s (time intervals 4 μ s) [3]; dynamic tensile extrusion of Copper for (a) 65 μ m, (b) 118 μ m and (c) 185 μ m grain size [1]

Figure 4 shows some results presented in [1,3] in terms of shape evolution during the test and final geometry of the recovered samples (no other experimental measurement is available).

Table 2 summarizes the set of numerical results in terms of plastic strain for the copper sphere (diameter 7.60 mm) launched at 400 m/s. The die (entrance diameter of 7.62 mm, exit diameter of 2.28 mm and angle of 10°) is modelled as a rigid body ($E=200$ GPa) and the contact with the specimen is frictionless. The results are shown at the same time step (48 μ s after the impact).

For all the material models, the deformed shape has a good likelihood with the experimental test results even if the model parameters are not optimized. The level of plastic strain is comparable for all the models, as showed in Table 2, but the speed of the projected pieces is different due to a different flow stress evaluation that implies a different energy dissipation (plastic work).

As mentioned before, these results are obtained with the data available in [8-12]; with a change of some material parameters the results of the numerical simulations becomes much more similar to the experimental ones. In an empirical model (such J-C) all the parameters affect the results and could be changed to obtain a best fit, but this may be very complicate without results from other type of tests or a suitable optimization strategy. On the other hand, with a phenomenological model the number of parameters that is sensible to change is usually limited.

Table 2: numerical results in terms of plastic strain for the simulation of the DTE test on copper sphere launched at 400 m/s: comparison between different material models

To understand the typical distribution of mechanical and thermo-dynamical quantities in such problems, in Table 3 the results for S-G material model at two different time are reported. At 21 μ s after the impact, the sample is still in the die, subjected to high compression stress, strain and strain-rate with an increase in the particle velocity from the initial 400 m/s up to about 1 km/s. The pressure level is in the order of 2 GPa, and the temperature increment is about 200 K.

When the material is extruded outward, the stress, strain and strain-rate become positive (tension) and due to the high velocity and plastic instability, the strain level grows up and the deformation is localized. In this phase the pressure falls down but the temperature grows up in the heavy deformed portions, with an increment of about 500 K.

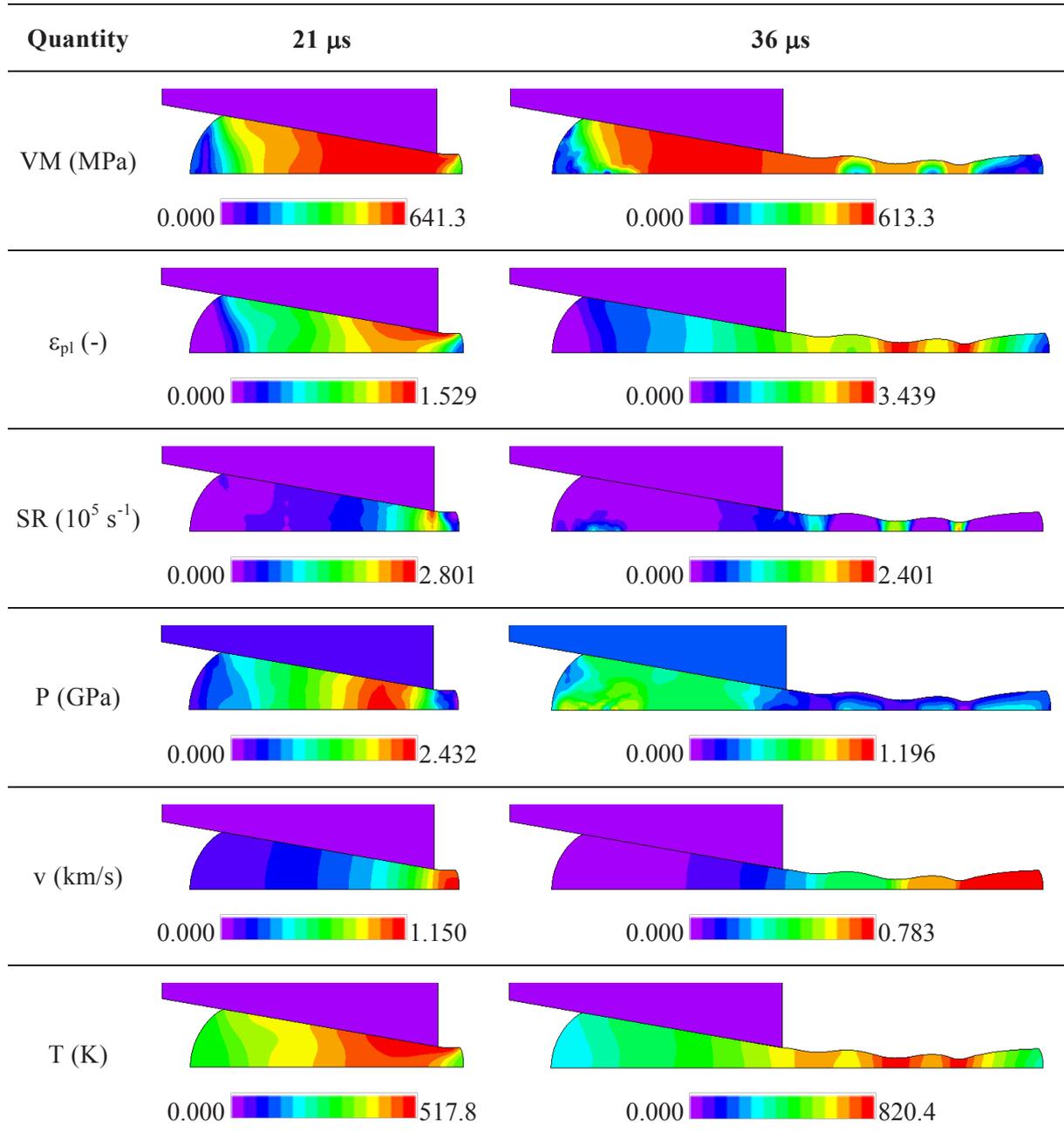
CONCLUSION

Recently, the researchers of the Los Alamos National Laboratory (LANL) developed a technique to investigate high strain-rates in which a sphere is launched at high speed (300-800 m/s) through a conical die (Dynamic Tensile Extrusion test).

The system can use a set-up similar to a Taylor test (like a light gusgun) but induces in the material high levels of tensile stresses and allows to investigate the fracture and damage of the material in these extreme conditions. Due to the high strain-rates and hydrostatic component

of stress that occur during the extrusion in the die, the material is also subject to a rapid heating.

Table 3: numerical results for the simulation of the DTE test on copper sphere launched at 400 m/s: S-G strength material model



In this work, the DTE test was simulated with the explicit Lagrangian code LS-DYNA. In order to correctly reproduce the various aspects of the test it is necessary to use an elasto-plastic material model that takes into account strain-rate, thermal softening, pressure influence

and damage and that adequately describes the thermal coupling effects and the equation of state (EOS) of the material.

Several strength material models both empirical and phenomenological were compared.

Besides, due to the high distortion of the Lagrangian mesh, the use of a mesh adapting technique was absolutely necessary.

The simulations were compared with available experimental results from tests performed on pure copper, at Los Alamos National Laboratory.

REFERENCES

- [1] Gray III, G. T., et al., Influence of shock prestraining and grain size on the dynamic tensile extrusion response of copper: experiments and modeling. *14th American Physical Society Topical Conference on Shock Compression of Condensed Matter*, Baltimore, MD (2005)
- [2] Gladman, B., et al., *LS-DYNA® Keyword User's Manual – Volume I – Version 971*. LSTC (2007)
- [3] Burkett, M.W. and Clancy, S.P., Eulerian Hydrocode Modelling of a Dynamic Tensile Extrusion Experiment. *Proceedings of the 11th Hypervelocity Impact Symposium* (2010)
- [4] Cao, F., Cerreta, E.K., Trujillo, C.P. and Gray III, G.T., Dynamic tensile extrusion response of tantalum. *Acta Materialia* (2008) **56**:5804-5817
- [5] Bonora, N., Flater, P.J., Ruggero, A., Ianniti, G. and Nixon, M.E., Dynamic tensile extrusion (DTE): a new experimental technique for the validation of constitutive modeling – Preliminary results. *Workshop IGF* (2010)
- [6] Macdougall, D., Determination of the plastic work converted to heat using radiometry. *Exp. Mech.* (2000) **40**:298-306
- [7] Hodowany, J., Ravichandram, G., Rosakis, A.J. et al., *Exp. Mech.* (2000) **40**:113-123
- [8] Johnson, G. R. and Cook, W. A., A constitutive model and data for metals subjected to large strains, high strain rates and high temperatures, *7th International Symposium on Ballistic*, (1983) 541-547
- [9] Steinberg, D.J., Cochran, S.G. and Guinan, M.W., A constitutive model for metals applicable at high-strain rate. *Journal of Applied Physics*, (1980) **51**:1498
- [10] Steinberg, D.J. and Lund, C.M., A constitutive model for strain rates from 10^{-4} to 10^6 s⁻¹. *Journal de physique, Symposium C3*, (1988) **49**:433-440
- [11] Zerilli, F.J. and Armstrong, R.W., Dislocation-mechanics-based constitutive relations for material dynamics calculations. *Journal of Applied Physics*, (1987) Vol. 61
- [12] Follansbee, P.S. and Kocks, U.F., A constitutive description of the deformation of copper based on the use of the mechanical threshold. *Acta Metallurgica*, (1988) **36** (1):81-93
- [13] Regazzoni, G., Kocks, U.F. and Follansbee, P.S., Dislocation kinetics at high strain rates. *Acta Metall.* (1987) **12**:2865-2875
- [14] Varshni, Y.P., Temperature Dependence of the Elastic Constants. *Physical Review B* (1970) Vol. 2, No. 10

COMPARATIVE PATIENT-SPECIFIC FSI MODELING OF CEREBRAL ANEURYSMS

KENJI TAKIZAWA*, TYLER BRUMMER†, TAYFUN E. TEZDUYAR†
AND PENG R. CHEN‡

*Department of Modern Mechanical Engineering and
Waseda Institute for Advanced Study, Waseda University
1-6-1 Nishi-waseda, Shinjuku-ku, Tokyo 169-8050, JAPAN

†Mechanical Engineering, Rice University – MS 321
6100 Main Street, Houston, TX 77005, USA

‡Cerebrovascular and Neuro-Endovascular Program
Skull Base Program, Department of Neurosurgery
Mischer Neuroscience Institute
University of Texas Medical School at Houston
6400 Fannin, Houston, TX 77030, USA

Key words: Cardiovascular fluid mechanics, Cerebral aneurysms, Patient-specific modeling, Fluid–structure interaction, Space–time techniques

Abstract. We consider a total of ten cases, at three different locations, half of which ruptured sometime after the images were taken. We use the stabilized space–time FSI technique developed by the Team for Advanced Flow Simulation and Modeling, together with a number of special techniques targeting arterial FSI modeling. We compare the ten cases based on the wall shear stress, oscillatory shear index, and the arterial-wall stress. We also investigate how simpler approaches to computer modeling of cerebral aneurysms perform compared to FSI modeling.

1 INTRODUCTION

Arterial fluid mechanics modeling is now a significant part of computational biomechanics research. Much of this has been in patient-specific modeling of cerebral arteries with aneurysm, taking into account the fluid–structure interaction (FSI) between the blood flow and arterial walls. This class of research has been benefiting much from computational mechanics techniques targeting FSI modeling in general. The Deforming-Spatial-Domain/Stabilized Space–Time (DSD/SST) formulation [1, 2] was developed by the Team for Advanced Flow Simulation and Modeling (T★AFSM) for flow computations with moving boundaries and interfaces, including FSI. The formulation is based

on the Streamline-Upwind/Petrov-Galerkin (SUPG) [3] and Pressure-Stabilizing/Petrov-Galerkin (PSPG) [1] methods. The DSD/SST formulation is used with the mesh update methods [4] developed by the T★AFSM. New-generation DSD/SST formulations were introduced by the T★AFSM in [5]. The stabilized space–time FSI (SSTFSI) technique, which is based on the new-generation DSD/SST formulations, was also introduced in [5]. The SSTFSI technique, with special techniques developed by the T★AFSM for arterial FSI, has been extensively used for arterial modeling, with emphasis on cerebral aneurysms [6, 7]. The special techniques include methods for calculating an estimated zero-pressure (EZP) arterial geometry [8, 9], a special mapping technique for specifying the velocity profile at an inflow boundary with non-circular shape [10], techniques for using variable arterial wall thickness [10, 9], mesh generation techniques for building layers of refined fluid mechanics mesh near the arterial walls [11, 9], a recipe for pre-FSI computations that improve the convergence of the FSI computations [6, 8], and techniques [12] for the projection of fluid–structure interface stresses, calculation of the wall shear stress (WSS) and calculation of the oscillatory shear index (OSI).

In this paper, which is a short version of a recently-submitted journal article [7], we focus on comparative patient-specific FSI modeling of cerebral aneurysms. We have ten cases, at three different locations, half of which ruptured sometime after the images were taken. We compare these cases based on the WSS, OSI and the arterial-wall stress. We also investigate how simpler methods perform compared to FSI modeling.

2 COMPUTATIONAL TECHNIQUES AND GENERAL CONDITIONS

We use the SSTFSI technique [5], with special techniques for arterial FSI. The special techniques include a mapping technique for specifying the inflow velocity profile [10], methods [12] for the projection of fluid–structure interface stresses and calculation of the WSS and OSI, and the Separated Stress Projection (SSP) technique [13, 9]. Special boundary condition techniques [9] are used for inclined inflow and outflow planes. The fully-discretized, coupled fluid, structure and mesh-moving equations are solved with the quasi-direct coupling technique (see Section 5.2 in [5]). In iteratively solving the linear systems involved at every nonlinear iteration, we use “Selective Scaling” technique (see Remark 14 in [5]) to shift the emphasis between the fluid and structure parts. In some cases, we also use selective scaling to shift the emphasis between the parts of the fluid equations corresponding to the momentum conservation and incompressibility constraint [14].

The fluid and structure properties can be found in [9]. At the inflow we specify the velocity profile as a function of time, by using a special technique [9]. The Womersley parameter, which appears in that technique, is defined as $\Upsilon = r_B \sqrt{2\pi/\nu T}$. Here r_B is the average radius of the inflow cross-sectional area, which comes from the image-based data, ν is the kinematic viscosity, and T is the period of the cardiac cycle, which is taken as 1 s. The volumetric flow rate (which is calculated based on a velocity waveform that represents the cross-sectional maximum velocity) is scaled by a factor. The scaling factor is determined in such a way that the scaled flow rate, when averaged over the cardiac

cycle, yields a target WSS for Poiseuille flow over an equivalent cross-sectional area. The target WSS is 10 dyn/cm^2 in the current T★AFSM computations. The time-step size is $3.333 \times 10^{-3} \text{ s}$. The number of nonlinear iterations per time step is 6. The number of GMRES [15] iterations per nonlinear iteration for the fluid+structure block was chosen such that mass balance is satisfied to within at most 5% for each case. For all six nonlinear iterations the fluid scale is 1.0 and the structure scale is 100. In three of the cases, the fluid scales for the momentum conservation and incompressibility constraint are 1.0 and 10. For the mesh moving block the number of GMRES iterations is 30. All computations were completed without any remeshing. For additional description of the computational techniques used and general conditions, see [7].

3 CASE STUDIES

Ten cases are studied from three locations: 4 Middle Cerebral Artery (MCA), 4 Anterior Communicating Artery (Acom), and 2 Basilar Artery. Half of each location ruptured sometime after the images were taken. Figures 1–3 show the lumen geometries. The

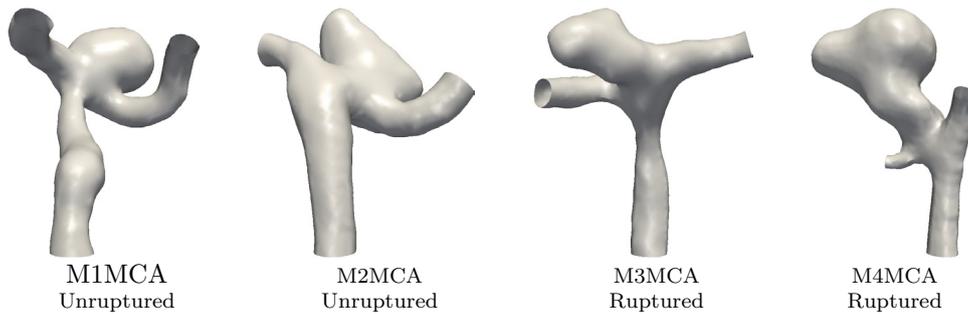


Figure 1: Arterial lumen geometry obtained from voxel data for the MCA models.

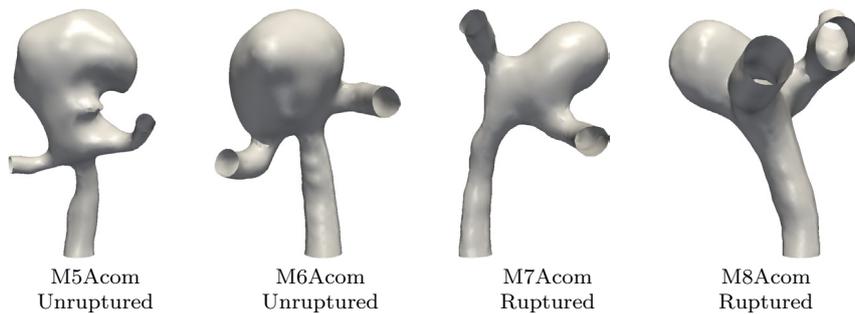


Figure 2: Arterial lumen geometry obtained from voxel data for the Acom models.

physical parameters are shown in Table 1. The number of nodes vary between 8,000 and 18,000 for the structure (hexahedral) meshes and between 33,000 and 60,000 for the fluid (tetrahedral) meshes. For all models the maximum WSS occur at the maximum inflow

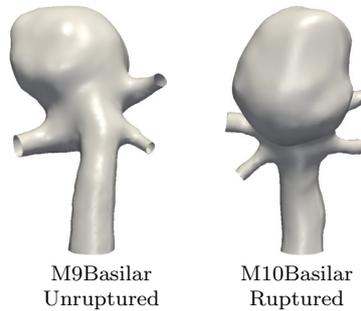


Figure 3: Arterial lumen geometry obtained from voxel data for the Basilar models.

Model	D_1	D_{O1}	D_{O2}	D_{O3}	Υ	Q_{\max}
M1MCA	2.43	2.42	1.87		1.52	0.51
M2MCA	1.56	1.41	1.38		0.97	0.12
M3MCA	2.50	1.49	1.43		1.57	0.56
M4MCA	1.70	1.21	0.81		1.06	0.08
M5Acom	3.05	1.78	1.75		1.91	1.08
M6Acom	3.13	2.12	2.12		1.96	1.20
M7Acom	1.02	0.90	0.80		0.64	0.04
M8Acom	1.94	2.31	2.17	1.38	1.21	0.25
M9Basilar	2.60	1.31	1.01	0.88	1.63	0.64
M10Basilar	3.03	1.34	1.04	1.01	1.90	1.06

Table 1: Physical parameters. Diameters are in mm and peak volumetric flow rate is in ml/s. M10Basilar has a fourth outflow with diameter 0.93 mm.

flow rate of the cardiac cycle. Maximum and average WSS values are shown in Figure 4. Figures 5–9 show the OSI for all the models. The maximum structural stress in space and time occurs at the peak pressure. Figure 10 shows the maximum stress and maximum variation in stress for all the models. As a point of reference, we note from [16] that the breaking strength of saccular aneurysms is in the range of 730–1,900 kPa.

4 NUMERICAL-PERFORMANCE STUDIES

We investigate how simpler approaches to modeling of our ten cases compare to FSI modeling. The three simpler modeling techniques are computing the blood flow with the artery shape held fixed at the average pressure (92 mm Hg), computing the arterial wall deformation with a prescribed, time-dependent pressure, and computing the blood flow with the prescribed arterial shape coming from that arterial-wall computation. We refer to these modeling techniques as “Rigid Artery (RA)”, “Structure (S)”, and “Prescribed Shape (PS)” in this paper. For the RA and PS computations we compare the WSS and OSI. For the Structure computations, we compare the arterial-wall stress.

Figure 11 shows the maximum and average WSS for the FSI, RA, and PS techniques. We see the maximum WSS being almost the same for the FSI and PS computations. The shape for the PS comes from a structural mechanics only computation where the viscous forces from the fluid are not accounted for. This gives the PS a slightly smaller shape than the FSI shape resulting in WSS that is on average 2.5% higher than FSI. For the

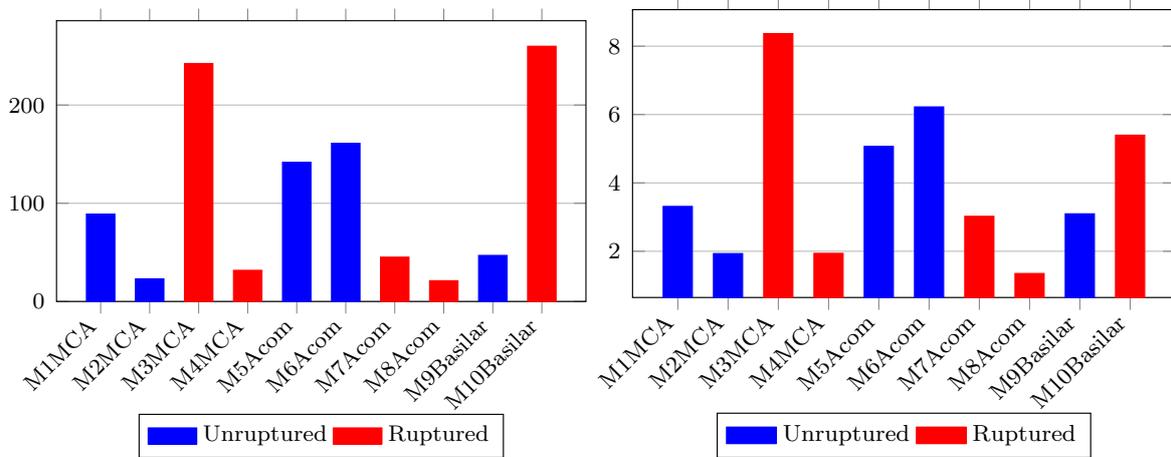


Figure 4: Maximum and average WSS (dyn/cm²) in space and time.

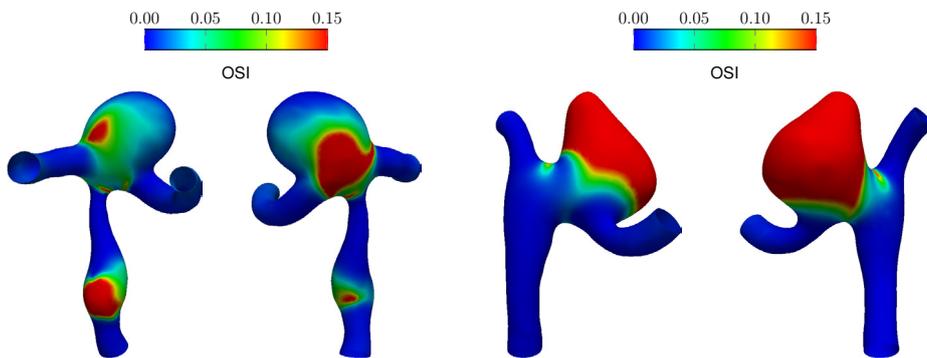


Figure 5: OSI for M1MCA and M2MCA.

RA computations, we see a clear pattern of higher WSS than the FSI. The RA shape is inflated to the average pressure, which, over the interval of peak inflow flow rate, is smaller than the FSI shape over that same interval. For this reason, we see the average WSS to be 4.2% higher than FSI. Using M5Acom as a sample model, we compare the spatial distribution of the OSI obtained with the PS and RA techniques (Figures 12 and 13) to those obtained with the FSI computation. The differences between the RA and FSI computations show the need for computing with a deformable structure. Figure 14 shows the maximum arterial-wall stress for the FSI and Structure techniques. The differences are less than 1%.

5 CONCLUDING REMARKS

We have presented an extensive comparative study based on patient-specific FSI modeling of cerebral aneurysms. We considered a total of ten artery models, coming from three different locations, half of which ruptured sometime after the images were taken.

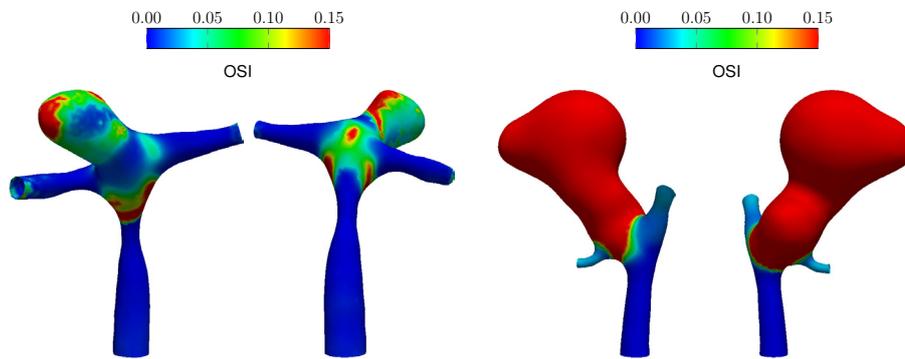


Figure 6: OSI for M3MCA and M4MCA.

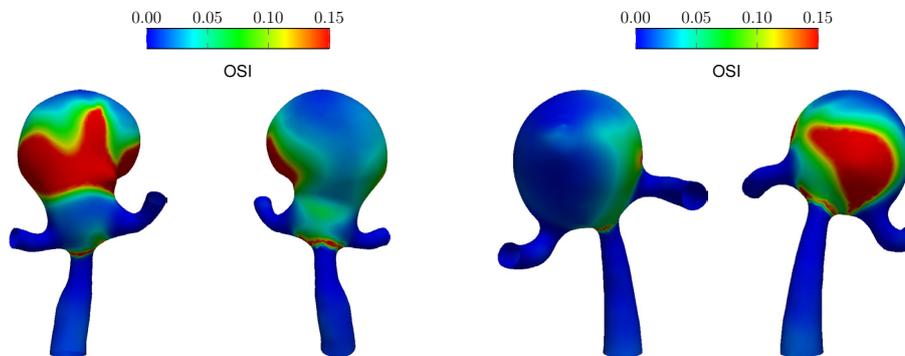


Figure 7: OSI for M5Acom and M6Acom.

We used the SSTFSI technique developed by the T★AFSM, together with a number of special techniques targeting arterial FSI modeling, which were also developed by the T★AFSM. We compared the WSS, OSI and the arterial-wall stress. We also showed how simpler approaches perform compared to FSI modeling. The simpler approaches were computing the blood flow with the artery shape held fixed, computing the arterial wall deformation with a prescribed, time-dependent pressure, and computing the blood flow with the prescribed arterial shape coming from that arterial wall computation.

ACKNOWLEDGMENT

This work was supported partly by John & Ann Doerr Fund for Computational Biomedicine. It was also supported partly by NSF Grant CNS-0821727. The 3DRA research at the Memorial Hermann was supported by the Weatherhead Foundation. We thank Dr. Ryo Torii for the velocity wave form.

REFERENCES

- [1] T.E. Tezduyar, “Stabilized finite element formulations for incompressible flow computations”, *Advances in Applied Mechanics*, **28** (1992) 1–44.

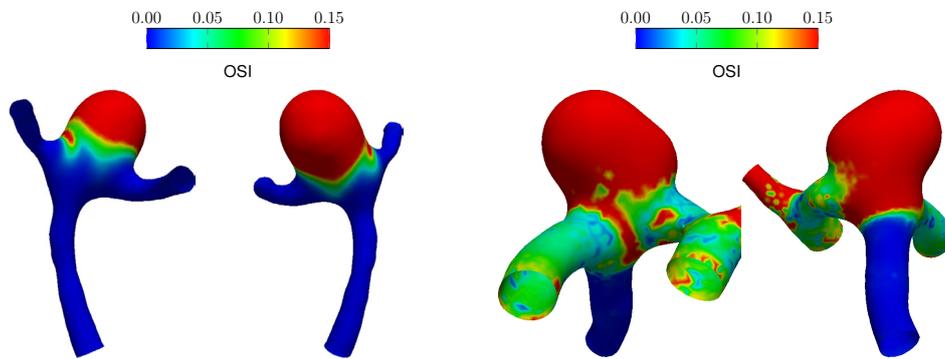


Figure 8: OSI for M7Acom and M8Acom.

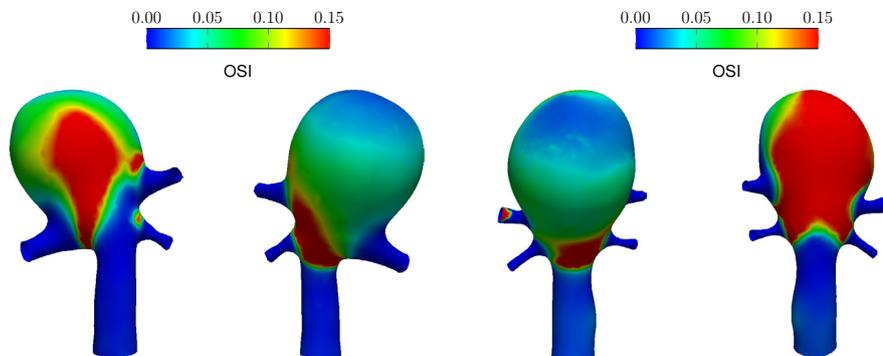


Figure 9: OSI for M9Basilar and M10Basilar.

- [2] T.E. Tezduyar, “Computation of moving boundaries and interfaces and stabilization parameters”, *International Journal for Numerical Methods in Fluids*, **43** (2003) 555–575.
- [3] A.N. Brooks and T.J.R. Hughes, “Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations”, *Computer Methods in Applied Mechanics and Engineering*, **32** (1982) 199–259.
- [4] T. Tezduyar, S. Aliabadi, M. Behr, A. Johnson, and S. Mittal, “Parallel finite-element computation of 3D flows”, *Computer*, **26** (1993) 27–36.
- [5] T.E. Tezduyar and S. Sathe, “Modeling of fluid–structure interactions with the space–time finite elements: Solution techniques”, *International Journal for Numerical Methods in Fluids*, **54** (2007) 855–900.
- [6] T.E. Tezduyar, S. Sathe, T. Cragin, B. Nanna, B.S. Conklin, J. Pausewang, and M. Schwaab, “Modeling of fluid–structure interactions with the space–time finite

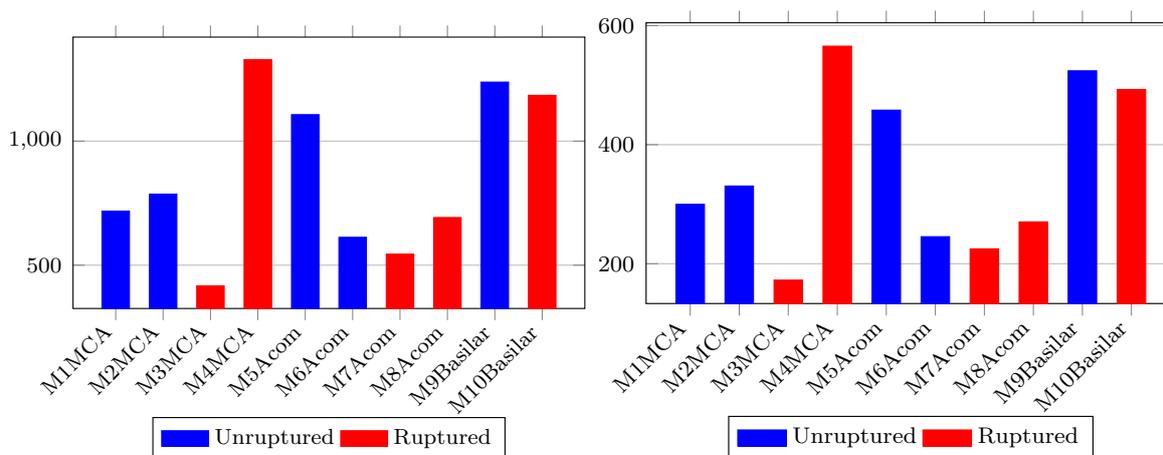


Figure 10: Maximum stress and maximum variation in stress (kPa) in space and time.

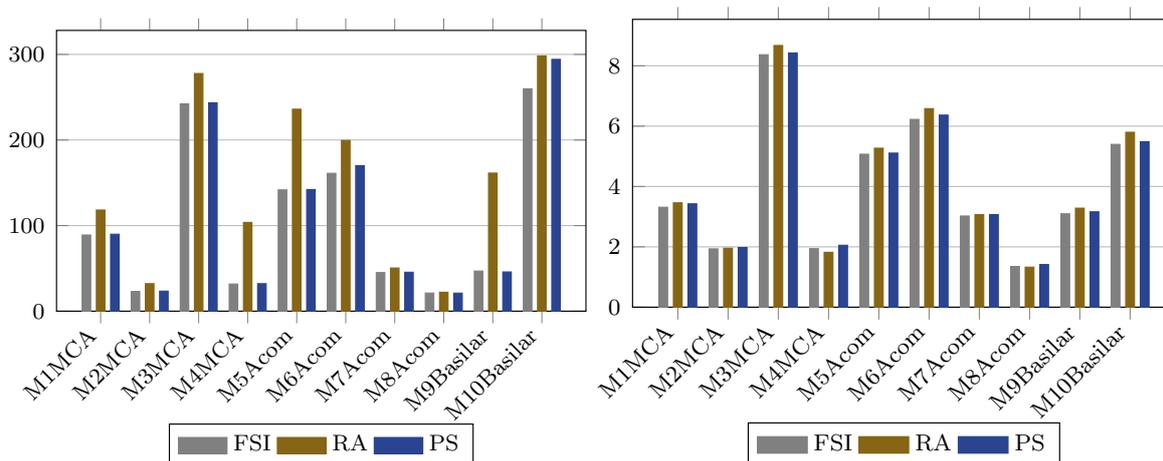


Figure 11: Maximum and average WSS (dyn/cm²) in space and time.

elements: Arterial fluid mechanics”, *International Journal for Numerical Methods in Fluids*, **54** (2007) 901–922.

- [7] K. Takizawa, T. Brummer, T.E. Tezduyar, and P.R. Chen, “A comparative study based on patient-specific fluid–structure interaction modeling of cerebral aneurysms”, *Journal of Applied Mechanics*, submitted, 2011.
- [8] T.E. Tezduyar, S. Sathe, M. Schwaab, and B.S. Conklin, “Arterial fluid mechanics modeling with the stabilized space–time fluid–structure interaction technique”, *International Journal for Numerical Methods in Fluids*, **57** (2008) 601–629.
- [9] T.E. Tezduyar, K. Takizawa, T. Brummer, and P.R. Chen, “Space–time fluid–structure interaction modeling of patient-specific cerebral aneurysms”, *International*

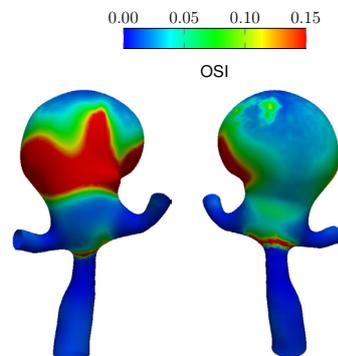


Figure 12: PS OSI for M5Acom.

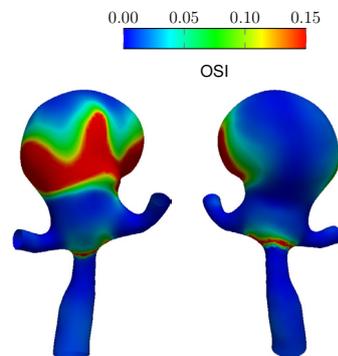


Figure 13: RA OSI for M5Acom.

Journal for Numerical Methods in Biomedical Engineering, published online, DOI: 10.1002/cnm.1433, February 2011.

- [10] K. Takizawa, J. Christopher, T.E. Tezduyar, and S. Sathe, “Space–time finite element computation of arterial fluid–structure interactions with patient-specific data”, *International Journal for Numerical Methods in Biomedical Engineering*, **26** (2010) 101–116.
- [11] T.E. Tezduyar, M. Schwaab, and S. Sathe, “Sequentially-Coupled Arterial Fluid–Structure Interaction (SCAFSI) technique”, *Computer Methods in Applied Mechanics and Engineering*, **198** (2009) 3524–3533.
- [12] K. Takizawa, C. Moorman, S. Wright, J. Christopher, and T.E. Tezduyar, “Wall shear stress calculations in space–time finite element computation of arterial fluid–structure interactions”, *Computational Mechanics*, **46** (2010) 31–41.
- [13] T.E. Tezduyar, S. Sathe, J. Pausewang, M. Schwaab, J. Christopher, and J. Crabtree, “Interface projection techniques for fluid–structure interaction modeling with moving-mesh methods”, *Computational Mechanics*, **43** (2008) 39–49.

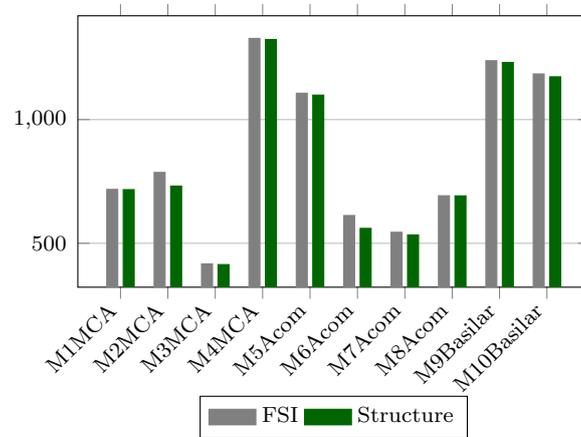


Figure 14: Maximum stress (kPa) in space and time.

- [14] K. Takizawa, C. Moorman, S. Wright, J. Purdue, T. McPhail, P.R. Chen, J. Warren, and T.E. Tezduyar, “Patient-specific arterial fluid–structure interaction modeling of cerebral aneurysms”, *International Journal for Numerical Methods in Fluids*, **65** (2011) 308–323.
- [15] Y. Saad and M. Schultz, “GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems”, *SIAM Journal of Scientific and Statistical Computing*, **7** (1986) 856–869.
- [16] D.J. Macdonald, H.M. Finlay, and P.B. Canham, “Directional wall strength in saccular brain aneurysms from polarized light microscopy”, *Annals of Biomedical Engineering*, **28** (2000) 533–542.

MULTI-SCALE ANALYSIS ON CAVITATION DAMAGE AND ITS MITIGATION FOR THE SPALLATION NEUTRON SOURCE

KOHEI OKITA^{*}, KENJI ONO^{*}, SHU TAKAGI[†] AND YOICHIRO MATSUMOTO[†]

^{*} Functionality Simulation and Information Team, VCAD System Research Program, RIKEN
Hirosawa 2-1, Wako-shi, Saitama, 351-0198, Japan
e-mail: {okita, keno}@riken.jp, www.riken.go.jp

[†] Department of Mechanical Engineering, the University of Tokyo
Hongo 7-3-1, Bunkyo-ku, Tokyo, 113-8656, Japan
e-mail: {takagi, ymats}@mech.t.u-tokyo.ac.jp, www.u-tokyo.ac.jp

Key words: Pressure Wave Propagation, Bubbly Flows, Spallation Neutron Source, Cavitation Erosion, Fluid Structure Interaction

Abstract. *Impact of injecting microbubbles on the thermal expansion due to the nuclear spallation reaction were examined numerically. Since the mercury density is higher than the density of solid wall, the interaction between mercury and solid wall must be taken into account. Our approach is to solve the momentum and energy conservation equations and the time development of elastic stress for both bubbly fluid and elastic solid. The Keller equation is employed to reproduce the nonlinear oscillation of bubble with considering the thermal dumping effect by the reduced order model. The continuum phase of liquid mercury is coupled with the discrete phase of microbubbles using the Euler-Lagrange method. As the results, the bubble cloud develops around the center of inertia of motion induced by the thermal expansion. The elasticity of the wall affects on the migration of the center of inertia away from the wall. The injection of microbubbles is effective to decrease the pressure rise due to thermal expansion for both rigid and elastic wall conditions when the void fraction of microbubbles is higher than the volume rate of thermal expansion of liquid mercury.*

1 INTRODUCTION

Japan proton accelerator research complex (J-PARC) was constructed as a spallation neutron source in Japan. The heat generation of the nuclear spallation reaction causes the thermal expansion of liquid mercury, which produces high pressure waves. When the pressure waves hit a casing wall, cavitation occurs and erodes the wall [1]. To mitigate the cavitation erosion, a method of introducing gas bubbles into liquid mercury has been proposed [2]. The method has expected that the microbubbles absorb the thermal expansion of liquid mercury and attenuate the pressure waves.

The propagation of pressure waves caused by a thermal shock in liquid mercury containing microbubbles has been numerically investigated in the previous work [3]. The influences of the injecting bubble size and void fraction on the absorption of the thermal expansion of liquid mercury and the attenuation of pressure waves have clarified as follows. Firstly, if the void fraction is higher than the volume rate of thermal expansion of liquid mercury, the

pressure rise due to thermal expansion decreases with decreasing bubble radius, because of the increase of the natural frequency of bubble in bubbly mixture. Secondly, as the bubble radius increases, the peak of pressure waves which propagate at the sound speed of mixture decreases due to the dispersion effect of bubbly mixture. In the case of liquid mercury containing large bubbles whose natural frequency is lower than the frequency of thermal shock, the pressure waves propagate at the sound speed of the liquid mercury and increase the peak pressure at the wall. Thirdly, the comparison with and without heat transfer through the gas liquid interface shows that the pressure waves are attenuated by the thermal damping effect even if the decrease of the void fraction makes the behaviour of bubbles nonlinear.

Since the density of liquid mercury ($13,579\text{kg/m}^3$) is higher than the density of solid wall (316SS, $7,946\text{kg/m}^3$), the impact of the interaction between liquid mercury and wall on cavitation erosion should be considered when the pressure waves hit the wall. Therefore, we extend the previous work to consider the elasticity of wall in a two-dimensional computation.

Our approach is to solve the mass, momentum and energy conservation equations for bubbly fluid with the equation of state of liquid mercury to reproduce the pressure rise due to thermal shock. The Euler-Lagrange method is used for coupling the liquid mercury and microbubbles. The nonlinear oscillation of bubble dynamics is described by the Keller equation with an assumption of spherical bubble. The thermal damping effect due to the heat transfer through the bubble interface is taken into account using the reduced order model. The elastic wall is represented by solving the equations for elastic body.

In the present paper, firstly, basic equations are introduced to reproduce the propagation of pressure waves caused by a thermal shock in liquid mercury containing microbubbles with taking into account the interaction between bubbly fluids and elastic wall. Secondly, the numerical method and model are mentioned. Finally, the impacts of the interaction of the thermal expansion of liquid mercury and the wall on the pressure rise due to thermal expansion and the distribution of bubble clouds are discussed.

2 BASIC EQUATIONS

2.1 Discrete phase of bubbles

To describe the dynamics of spherical bubble of radius R , the Keller equation [4][5] is employed

$$\begin{aligned} & \left[1 - \frac{1}{c_s} \frac{dR}{dt} \right] R \frac{d^2 R}{dt^2} + \frac{3}{2} \left[1 - \frac{1}{3c_s} \frac{dR}{dt} \right] \left(\frac{dR}{dt} \right)^2 \\ & = \frac{1}{\rho} \left[1 + \frac{1}{c_s} \frac{dR}{dt} + \frac{1}{c_s} R \frac{d}{dt} \right] \left(P_G - \frac{2\sigma}{R} - \frac{4\mu}{R} \frac{dR}{dt} - P_L \right) \end{aligned} \quad (1)$$

where c_s is the sound speed of surrounding liquid, ρ is liquid density, σ is surface tension and μ is liquid viscosity. If the bubble boundary moves with a velocity much lower than the speed of sound in the gas, the pressure of gas inside bubble P_G can be taken as uniform and is expressed as

$$\frac{dP_G}{dt} = -3\gamma P_G \frac{1}{R} \frac{dR}{dt} + \frac{3(\gamma-1)}{R} K \left. \frac{\partial T}{\partial r} \right|_{r=R}, \quad (2)$$

where γ is the ratio of specific heats [6]. The second term on the right hand side represents the heat transfer between the gas inside bubble and the surrounding liquid. The temperature gradient at the bubble boundary is estimated by a reduced order model [7] instead of directly solving the energy equation for the temperature distribution inside bubble. The reduced order model is briefly described as follows.

The temperature gradient at the bubble boundary is modeled as

$$\left. \frac{\partial T}{\partial r} \right|_{r=R} = \frac{\operatorname{Re}(\tilde{L}_P)}{|\tilde{L}_P|^2} (T_0 - T_b) + \frac{\operatorname{Im}(\tilde{L}_P)}{|\tilde{L}_P|^2} \frac{1}{\omega_N} \frac{dT_b}{dt}, \quad (3)$$

where $T_b = P_G R^3 T_b / P_{G0} R_0^3$ is the representative temperature of the gas inside bubble. The natural frequency of bubble ω_N is described by

$$\omega_N^2 = \frac{\operatorname{Re}(\Upsilon_N) P_G T_0}{\rho R^2 T_b} - \frac{2\sigma}{\rho R^3} - \frac{4}{\rho^2 R^4} \left[\mu + \frac{\operatorname{Im}(\Upsilon_N) P_G T_0}{4\omega_N T_b} \right]^2 \quad (4)$$

where Υ_N is a complex function defined by

$$\Upsilon_N = \frac{3\alpha_N^2 \gamma}{\alpha_N^2 + 3(\gamma-1)(\alpha_N \coth \alpha_N - 1)} \quad (5)$$

with

$$\alpha_N^2 = i \frac{\gamma \omega_N P_G R^2}{(\gamma-1) T_b K_0}. \quad (6)$$

The thermal penetration length \tilde{L}_P is also a complex function and is described as

$$\tilde{L}_P = \frac{\alpha_N^2 - 3(\alpha_N \coth \alpha_N - 1)}{\alpha_N^2 (\alpha_N \coth \alpha_N - 1)} R \quad (7)$$

The temperature gradient at a boundary can be obtained by solving equations (3) to (7). Then the pressure inside is updated according to Eq.(2).

In the present study, bubbles are treated in a Lagrangian way. Due to the assumption of the no-slip condition, the bubble velocity is the same as the bubbly mixture velocity. The position of the each bubble is then traced by

$$\mathbf{x}_B = \int \mathbf{u}(\mathbf{x}_B) dt. \quad (8)$$

2.2 Continuum phase of bubbly fluid and elastic solid

Since the acoustic Mach number is low, the advection term assumes to be negligible. Then, the momentum equation can be represented

$$\rho \frac{\partial \mathbf{v}}{\partial t} = \nabla \cdot \boldsymbol{\sigma}. \quad (9)$$

The total strain rate is

$$\mathbf{e} = \frac{1}{2} \left\{ \nabla \mathbf{v} + (\nabla \mathbf{v})^T \right\}, \quad (10)$$

and includes the elastic strain rate, thermal strain rate and the volumetric strain rate due to bubble oscillation as $\mathbf{e} = \mathbf{e}_E + \mathbf{e}_T + \mathbf{e}_B$. Assuming the small strain, Hook's law for isotropic linear elasticity is employed for the elastic strain. The time development of elastic stress can be described

$$\frac{\partial \boldsymbol{\sigma}}{\partial t} = -\frac{\partial p}{\partial t} \mathbf{I} + 2\mu \left(\mathbf{e} - \frac{1}{3} \text{tr}(\mathbf{e}) \mathbf{I} \right), \quad (11)$$

where μ is shear modulus. The time development of pressure is defined by

$$\frac{\partial p}{\partial t} = -\zeta \left[\text{tr}(\mathbf{e}) - \text{tr}(\mathbf{e}_B) - \text{tr}(\mathbf{e}_T) \right] \quad (12)$$

where ζ is bulk modulus. The volume change due to bubble oscillation is described using the void fraction of bubbles f_G [8]

$$\text{tr}(\mathbf{e}_B) = \frac{1}{1-f_G} \frac{\partial f_G}{\partial t}. \quad (13)$$

And the thermal volumetric strain rate can be described using the thermal expansion rate at constant pressure $\beta = (1/V)(\partial V/\partial T)_p$ [9] as

$$\text{tr}(\mathbf{e}_T) = \frac{\beta}{C_p} T \frac{\partial S}{\partial t}. \quad (14)$$

Both a temperature diffusion and the heat generation due to the nuclear spallation reaction are taken into account in the energy equation

$$\rho T \frac{\partial S}{\partial t} = \nabla \cdot (k \nabla T) + Q. \quad (15)$$

Substituting Eq.(15) to Eq.(14), the thermal volumetric strain rate becomes

$$\text{tr}(\mathbf{e}_T) = \frac{\beta}{\rho C_p} \left[\nabla \cdot (k \nabla T) + Q \right]. \quad (16)$$

On the other hand, using the thermodynamic relation as $TdS = C_p dT - (\beta T/\rho) dp$ [9], the time development of temperature can be described

$$\rho C_p \frac{\partial T}{\partial t} = \beta T \frac{\partial p}{\partial t} + \nabla \cdot (k \nabla T) + Q. \quad (17)$$

In the present study, two volume fractions are introduced. Firstly, the volume fraction of solid phase is employed to distinguish solid phase from fluid phase. Secondly, the volume fraction of gas is defined as the volume ratio of bubble in the fluid phase. The mixture density of fluids and solid is described as $\rho = (1 - \phi_s)\rho_F + \phi_s\rho_S$, where the density of fluid phase can be assumed as $\rho_F = (1 - f_G)\rho_L$ because of low density and low void fraction of gas. Then the mixture density becomes

$$\rho = (1 - \phi_s)(1 - f_G)\rho_L + \phi_s\rho_S. \quad (18)$$

Specific heat and thermal expansion rate at constant pressure are represented respectively,

$$\rho C_p = (1 - \phi_s)(1 - f_G)\rho_L C_{pL} + \phi_s\rho_S C_{pS} \quad (19)$$

and

$$\beta = (1 - \phi_s)(1 - f_G)\beta_L + \phi_s\beta_S. \quad (20)$$

The coefficients, bulk modulus, shear modulus and thermal conductivity are calculated by the harmonic average using the volume fraction of solid phase as

$$\frac{1}{\zeta} = \frac{1 - \phi_s}{\zeta_L} + \frac{\phi_s}{\zeta_S}. \quad (21)$$

The density and bulk modulus of liquid phase is obtained from pressure and temperature through the equation of state for liquid mercury [3].

3 NUMERICAL METHODS

The basic equations are solved using a second-order finite difference scheme based on the FDTD method [10]. To resolve the rebound of bubble collapse, the Keller equation is integrated in adaptive time increments, which are always smaller than the time increment for the integration of the basic equations for the mixture. Bubbles are described by the representative bubble at the Lagrange point and coupled with mixture phase by the Euler-Lagrange method, which requires the interpolation of physical values between Euler and Lagrange points. The void fraction at the Euler point is obtained by

$$f_G(\mathbf{x}) = \int f_G(\mathbf{x}') \delta_\varepsilon(\mathbf{x}' - \mathbf{x}) d\mathbf{x}' = \sum_k \frac{\frac{4}{3}\pi n_{Bk} R_k^3}{V_k} D_\varepsilon(\mathbf{x}_{Bk} - \mathbf{x}) \quad (22)$$

where n_B is the number weight of the representative bubble and D_ε is defined by

$$D_\varepsilon(\mathbf{x}) = \begin{cases} \prod_{i=1}^2 \left[\frac{\Delta x_i}{2\varepsilon} \left\{ 1 + \cos\left(\frac{\pi x_i}{\varepsilon}\right) \right\} \right], & \text{for } |x_i| \leq \varepsilon \\ 0, & \text{othe wise} \end{cases}. \quad (23)$$

The width of the smooth delta function ε is taken as Δx . On the other hand, the pressure and velocity of the mixture at the bubble position, which are required to solve Eq. (1) and Eq. (8), are interpolated, respectively, as follows:

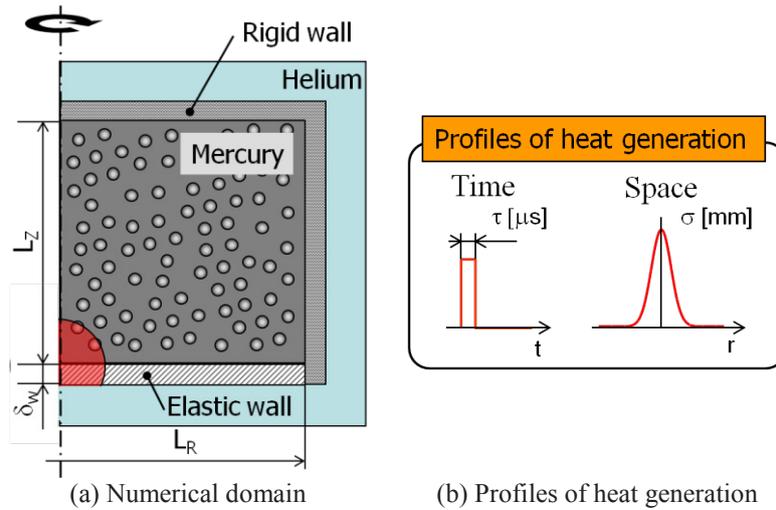


Figure 1: Schematic diagram of the numerical model. The bubbly liquid mercury is surrounded by the rigid and elastic wall of the cylinder, in which helium microbubbles uniformly distribute. The heat generation is introduced at the center bottom.

$$P_L(\mathbf{x}_B) = \sum_{\mathbf{x}} p(\mathbf{x}) D_\varepsilon(\mathbf{x} - \mathbf{x}_B) \quad \text{and} \quad \mathbf{u}(\mathbf{x}_B) = \sum_{\mathbf{x}} \mathbf{u}(\mathbf{x}) D_\varepsilon(\mathbf{x} - \mathbf{x}_B). \quad (24)$$

4 NUMERICAL MODEL

Bubbly liquid mercury in a cylinder is considered as shown in Fig.1(a). Helium microbubbles uniformly distribute in the liquid mercury. The sizes of numerical domain are $L_R=30\text{mm}$ and $L_Z=30\text{mm}$, which is resolved by 200×200 grid points. The width of bottom wall is $\delta_w=1.5\text{mm}$ resolved by 10 grid points. The properties of the solid material of 316SS are density $\rho_S=7,964\text{kg/m}^3$, Young's modulus $E_S=195\text{GPa}$ and Poisson ratio $\nu_S=0.27$. A heat generation is introduced around the center bottom of the cylinder to represent a spallation neutron reaction. The profiles of the heat generation are $1\mu\text{s}$ rectangular pulse in time and Gaussian distribution of the standard deviation $\sigma=5\text{mm}$ in space as shown in Fig.1(b). That induces the thermal expansion of liquid mercury and pressure rises. Pressure waves then propagate through the bubbly liquid mercury. We chose the maximum of the heat generation as $Q_{\max}=26.7 \times 10^{12}\text{W/m}^3$ to reproduce the pressure rise of the order of tens MPa in the case without microbubbles.

5 RESULTS

Influences of the elasticity of the bottom wall are examined without microbubbles. The time evolution of the pressure distribution for the elastic wall is compared with that for the rigid wall in Fig.2. In the case of the elastic wall, the pressure around the center bottom changes as plus, minus and plus in time. The pressure propagates upward unlike with the spherical pressure propagation of the rigid wall. The profiles of the pressure at points of $z=0, 4.5\text{mm}$ on the axis of the cylinder are shown in Fig.3 with comparing for the rigid and elastic wall conditions. The elasticity of the solid wall decreases the pressure fluctuation induced by the thermal expansion at $z=0\text{mm}$ as shown in Fig.3(a). However, as shown in Fig.3(b), the

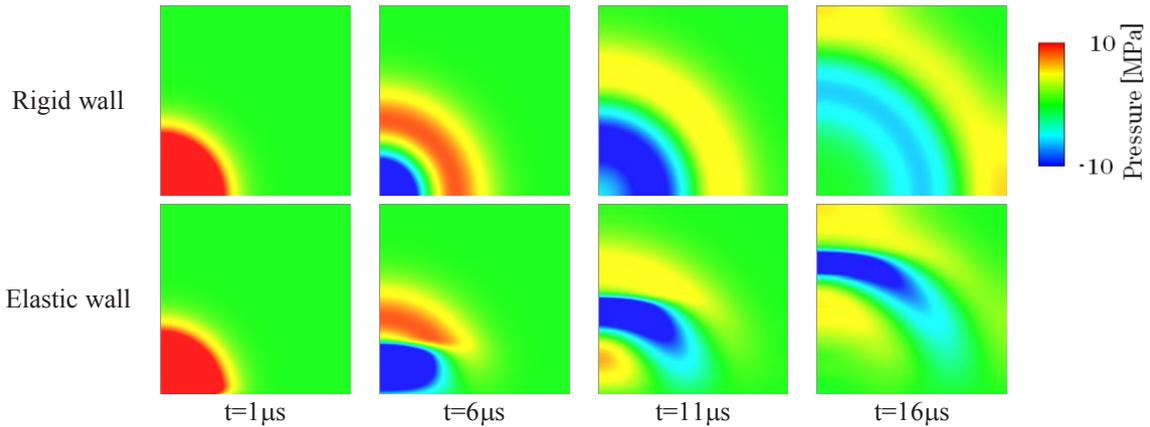


Figure 2: Influence of elasticity of bottom wall on the development of pressure distribution without microbubbles.

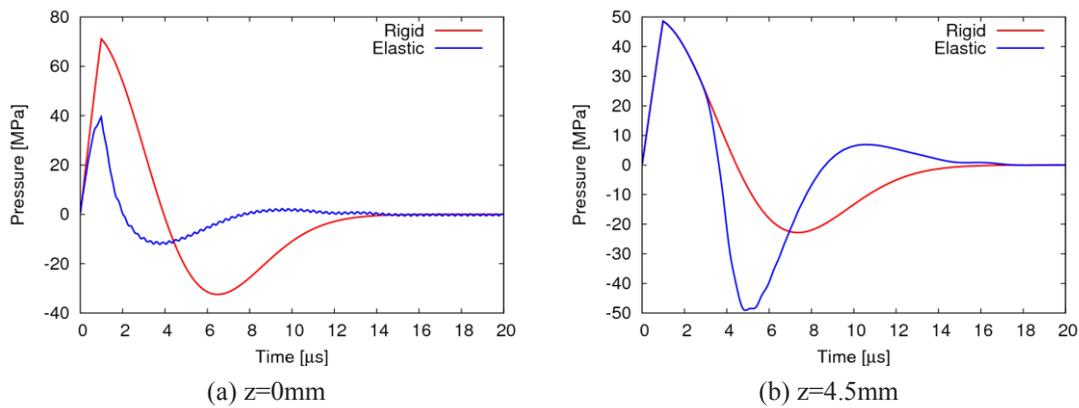


Figure 3: Influence of elasticity of bottom wall on the profiles of the pressure at $z=0, 4.5\text{mm}$ without microbubbles.

negative pressure becomes lower than that for the rigid wall. This is because the center of inertia of the motion induced by the thermal expansion is moved upward away from the center of heat generation owing to the elasticity of the wall, whose density of the solid wall is lower than the mercury density. Such negative pressure is closely related with the growth of bubble clouds. We have focused on the production of the bubble cloud after the thermal expansion of liquid mercury.

Influences of the elasticity of the bottom wall on the production of the bubble cloud are examined by injecting microbubbles to liquid mercury with bubble diameter $d=10\mu\text{m}$ and initial void fraction $f_{G0}=0.05\%$. Figure 4 shows the time evolution of the pressure and void fraction distributions for the rigid and elastic wall conditions. As shown in Fig.4(a), pressure waves spherically propagate under the rigid wall condition at the slower speed than the sound speed of mercury. The bubble clouds also spherically grow up at the center bottom. On the other hand, the bubble clouds grow up away from the wall in the case of the elastic wall. The shape of the bubble clouds is like a squashed sphere as shown in Fig.4(b). The pressure fluctuations are observed between the bubble clouds and the elastic wall. The profiles of the pressure and void fraction at points of $z=0, 4.5\text{mm}$ on the axis of the cylinder are shown in

Fig.5. The pressure increase to around 60MPa for the rigid wall and to around 35MPa for the elastic wall at the point of $z=0\text{mm}$. The peaks of pressure for both conditions are lower than the peaks without microbubbles. This is due to the contraction of the injected microbubbles. As shown in Fig.5(a), the pressure at $z=0\text{mm}$ fluctuates in the case of the elastic wall. Negative pressure as shown in Fig.3 is disappeared owing to the microbubbles. The void fraction at $z=0\text{mm}$ for the rigid wall is higher than that for the elastic wall. Contrary, the void fraction at $z=4.5\text{mm}$ for the rigid wall is lower than that for the elastic wall. This is because the center of inertia of motion induced by the thermal expansion is away from the wall as shown in Fig.4. Thus the bubble cloud develops around the center of inertia of motion induced by the thermal expansion of liquid mercury. The elasticity of the wall impacts on the migration of the center of inertia away from the wall.

Here, the volume rate of thermal expansion of liquid mercury is estimated as

$$\Delta f_L = \frac{\beta_L Q_{\max} \Delta \tau}{\rho_L C_{pL}}. \quad (25)$$

It becomes $\Delta f_L=0.26\%$ for $\beta_L=182 \times 10^{-6} \text{K}^{-1}$, $\rho_L=13,579 \text{kg/m}^3$, $C_{pL}=139 \text{J/kg}\cdot\text{K}$, $Q_{\max}=26.7 \times 10^{12} \text{W/m}^3$ and $\Delta \tau=1 \mu\text{s}$. So it is required for the mitigation of the pressure rise that the initial void fraction of microbubbles is higher than the volume rate of thermal expansion of liquid mercury $\Delta f_L=0.26\%$.

Next, calculations under the condition of the initial void fraction of $f_{G0}=0.5\%$ are performed. The diameter of microbubbles is $10 \mu\text{m}$. Figure 6 shows the time evolution of the pressure and void fraction distribution for the rigid and elastic wall condition. The time evolution of the pressure distribution for both is similar to that without microbubbles as shown in Fig.2. But the time scale with microbubbles is much longer due to the slower sound speed of bubbly liquid. On the other hand, the time evolution of the void fraction distribution well follows the pressure distribution. The void fraction decreases in the high pressure region and decreases in the low pressure region.

Figure 7 shows the profiles of the pressure and void fraction for the rigid and elastic wall conditions. Obviously, the pressure rise decreases in both conditions as shown in Fig.7(a) and (b). The intensity of pressure at $z=0\text{mm}$ is around 2.5MPa, which is much lower than that without microbubbles. The fluctuations of pressure at $z=0\text{mm}$ and 4.5mm are observed in $t=0\sim 4 \mu\text{s}$. This is due to the bubble oscillation with the natural frequency of bubble in bubbly liquid [3]. For the elastic wall, the pressure at $z=0\text{mm}$ continuously fluctuates owing to the oscillation of the elastic wall. As shown in Fig.7(c) and (d), the void fraction at $z=0\text{mm}$ decreases from 0.5% to 0.25% in $t=0\sim 1 \mu\text{s}$. The difference of the void fraction is around 0.25%, which approximately corresponds to the volume rate of thermal expansion of liquid mercury $\Delta f_L=0.26\%$. Thus, the injection of microbubbles is effective to decrease the pressure rise due to thermal expansion of liquid mercury for both rigid and elastic wall conditions when the void fraction of microbubbles is higher than the volume rate of thermal expansion of liquid mercury.

6 CONCLUSIONS

To investigate the cavitation erosion and its mitigation for the nuclear spallation source, the impacts of injecting microbubbles on the thermal expansion due to a nuclear spallation

reaction were examined numerically. The interaction between liquid mercury and solid wall was taken into account by solving the momentum and energy conservation equations and the time development of elastic stress for both fluid and solid phases. Additionally, the nonlinear oscillation of bubbles was reproduced using Keller equation with considering the thermal damping effect by the reduced order model.

As the result of the calculation, the bubble cloud develops around the center of inertia of motion induced by the thermal expansion of liquid mercury. The elasticity of the wall affects on the migration of the center of inertia away from the wall.

The injection of microbubbles is effective to decrease the pressure rise due to thermal expansion of liquid mercury for both rigid and elastic wall conditions when the void fraction of microbubbles is higher than the volume rate of thermal expansion of liquid mercury.

REFERENCES

- [1] Futakawa M., Naoe T., Tsai, C.C., Ishikura S., Ikeda Y., Soyama H., Date H., Cavitation erosion in mercury target of spallation neutron source, *Proceedings of Fifth International Symposium on Cavitation, Osaka, November 1-4 (2003)*.
- [2] Hansen G., Butzek M., Glückler H., Hanslik R., Sotner H., Soukhnov V., Stelzer H., Wolters J., Engineering work for the ESS target station. *Proceedings of the Sixth International Topical Meeting on Nuclear Applications of Accelerator Technology, San Diego, June 1-5(2003)*.
- [3] Okita K., Takagi S., Matsumoto Y., Propagation of Pressure Waves, Caused by a Thermal Shock, in Liquid Metals Containing Gas Bubbles, *Journal of Fluid Science and Technology*, (2008) **3**(1):116-128.
- [4] Keller J.B., Kolodner I.I., Damping of underwater explosion bubble oscillations, *J. Appl. Phys.* (1956) **27**(10):1152-1161.
- [5] Prosperetti A., Lezzi A., Bubble dynamics in a compressible liquid. part 1. first-order theory, *J. Fluid Mech* (1986) **168**:457-478.
- [6] Prosperetti A., The thermal behavior of oscillating gas bubbles, *J. Fluid Mech.* (1991) **222**:587-616.
- [7] Sugiyama K., Takagi S. and Matsumoto Y., A New reduced-Order Model for the Thermal Damping Effect on Radial Motion of a Bubble (2nd Report, Validation of the Model by Numerical Simulation), *Trans. JSME Ser. B* (2004) **71**:1011-1019.
- [8] Commander K.W. and Prosperetti A., Linear pressure waves in bubbly liquids: Comparison between theory and experiments, *J. Acoust. Soc. Am.* (1989) **85**(2):732-746.
- [9] Batchelor G.K., 1.5 Classical thermodynamics in *An Introduction to Fluid Dynamics*, Cambridge University Press (2000).
- [10] Yee K.S., Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media, *IEEE Trans. Antennas Propag.* (1996) **14**(8):301-307.

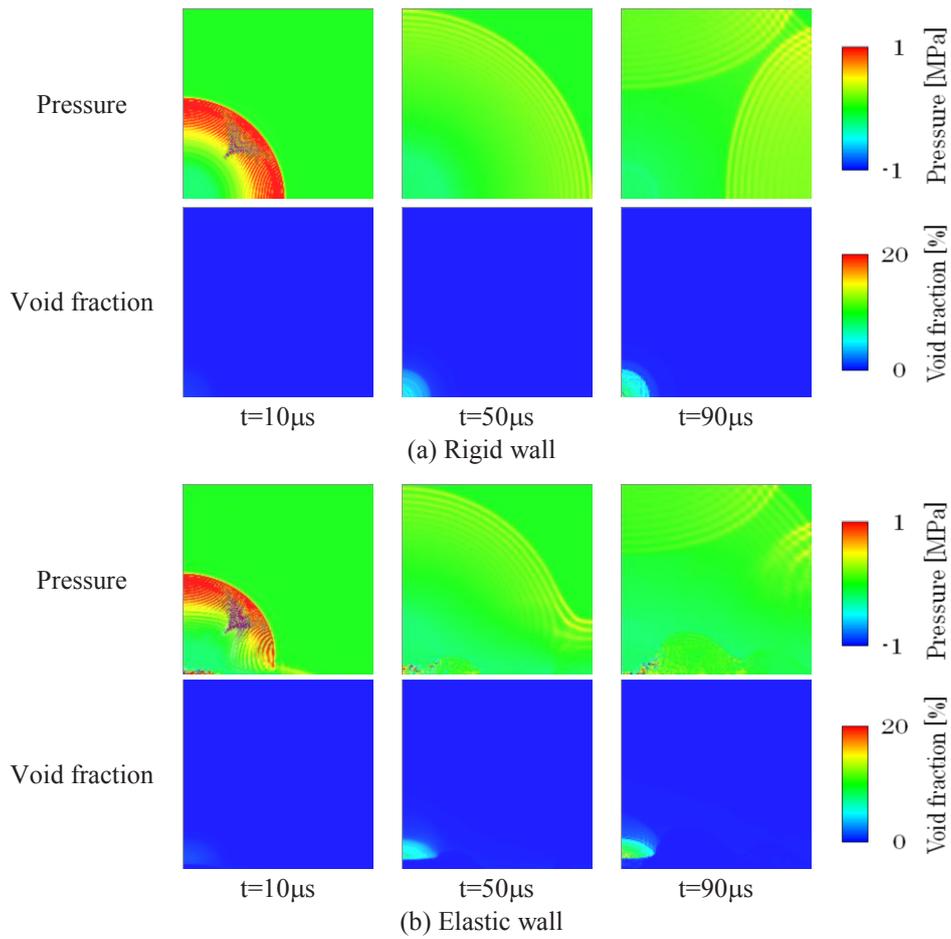


Figure 4: Time evolution of the pressure and void fraction distribution for $f_{G0}=0.05\%$.

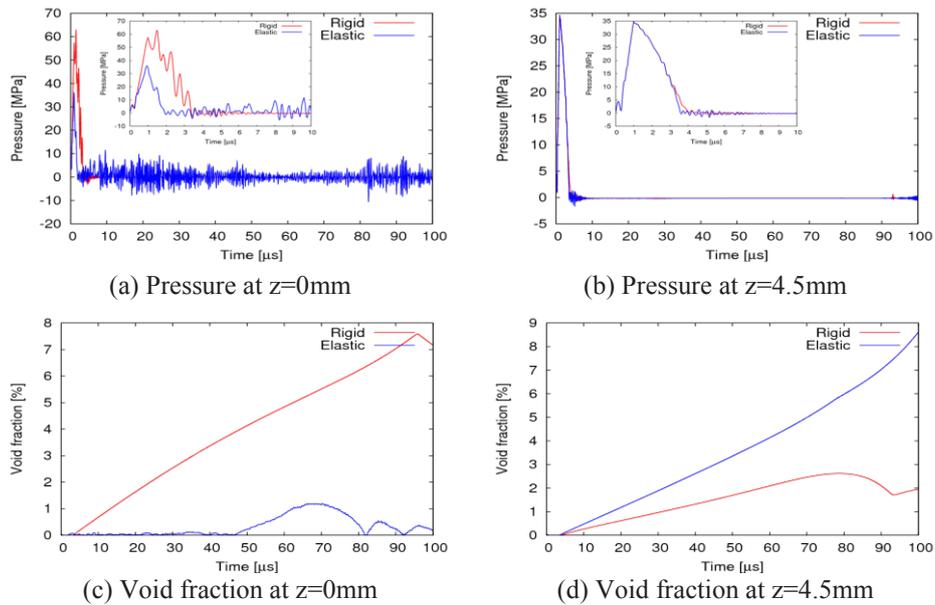


Figure 5: Profile of the pressure and void fraction at points of $z=0, 4.5\text{mm}$ for $f_{G0}=0.05\%$.

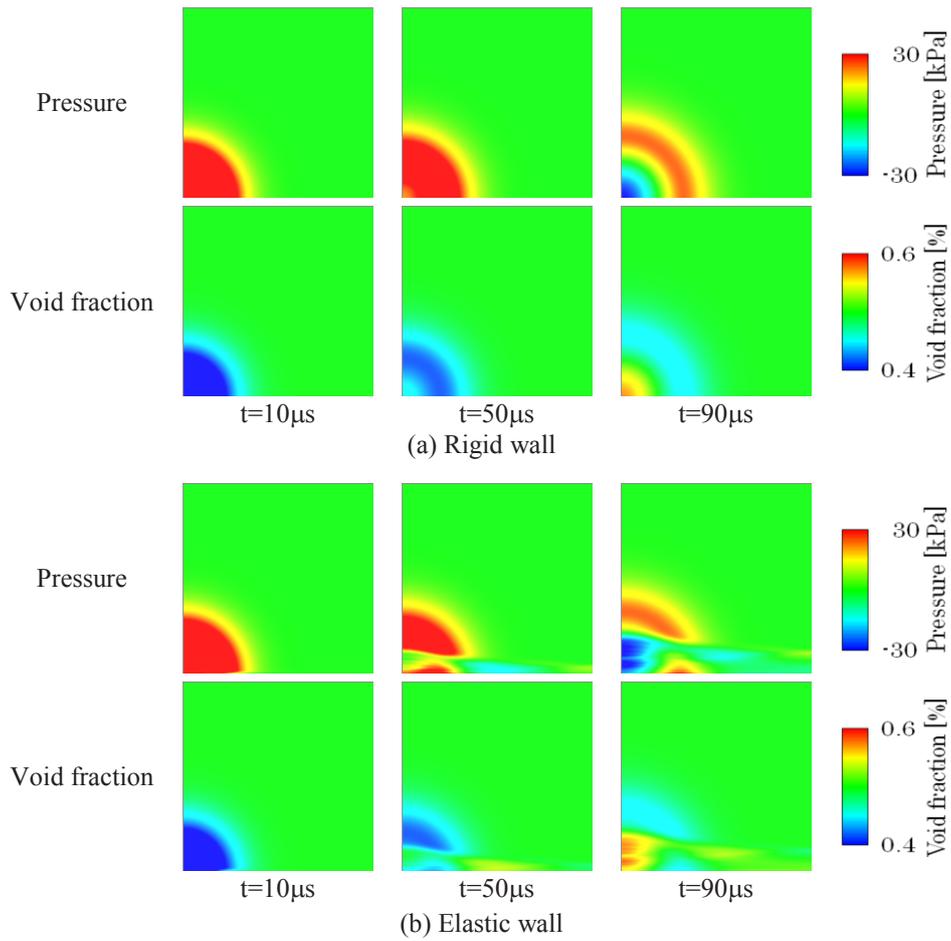


Figure 6: Time evolution of the pressure and void fraction distribution for $f_{G0}=0.5\%$.

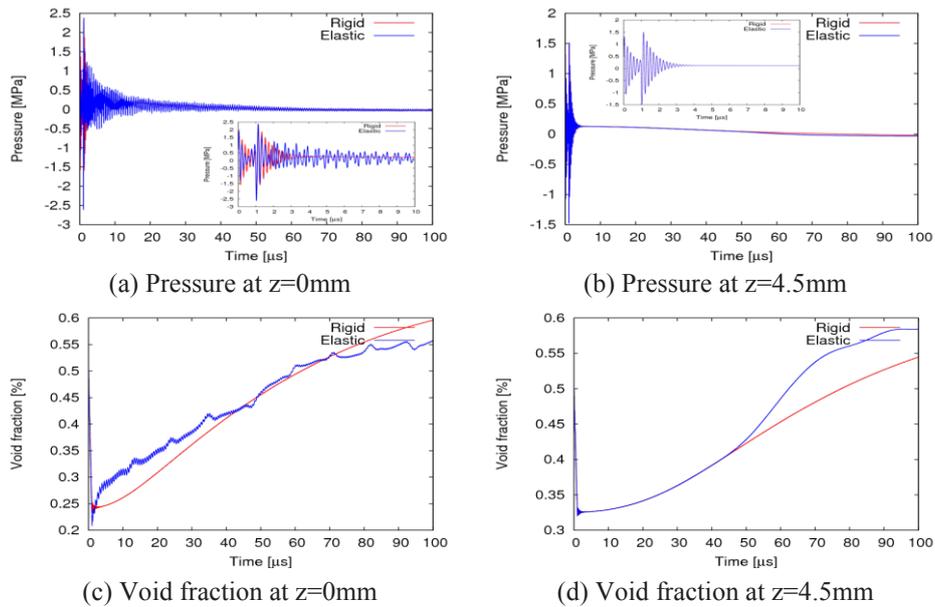


Figure 7: Profiles of the pressure and void fraction at points of $z=0, 4.5\text{mm}$ for $f_{G0}=0.5\%$.

MULTISCALE SPACE–TIME COMPUTATION TECHNIQUES

KENJI TAKIZAWA* AND TAYFUN E. TEZDUYAR†

*Department of Modern Mechanical Engineering and
Waseda Institute for Advanced Study, Waseda University
1-6-1 Nishi-waseda, Shinjuku-ku, Tokyo 169-8050, JAPAN

†Mechanical Engineering, Rice University – MS 321
6100 Main Street, Houston, TX 77005, USA

Key words: Fluid-structure interaction, Space–time formulations, Multiscale techniques, NURBS, Space–time variational multiscale method

Abstract. A number of multiscale space–time techniques have been developed recently by the Team for Advanced Flow Simulation and Modeling (T★AFSM) for fluid–structure interaction computations. As part of that, we have introduced a space–time version of the residual-based variational multiscale method. It has been designed in the context of the Deforming-Spatial-Domain/Stabilized Space–Time formulation, which was developed earlier by the T★AFSM for computation of flow problems with moving boundaries and interfaces. We describe this multiscale space–time technique, and present results from test computations.

1 INTRODUCTION

A number of multiscale space–time techniques [1, 2, 3, 4, 5] have been developed recently by the Team for Advanced Flow Simulation and Modeling (T★AFSM) for fluid–structure interaction (FSI) computations. These have been mostly multiscale techniques based on effective ways of dealing with the different spatial or temporal scales that may be involved in the fluid and structure parts of the problem. They have been tested in conjunction with the Deforming-Spatial-Domain/Stabilized Space–Time (DSD/SST) formulation [6, 7] and stabilized space–time FSI (SSTFSI) technique [8], both developed by the T★AFSM. In addition, recently we have introduced a multiscale space–time technique [9] that is based on representing the different flow scales involved in the fluid mechanics part, so that we could have a good turbulence model for high Reynolds number flows. This multiscale technique, which we call “DSD/SST-VMST”, is the space–time version of the residual-based variational multiscale method [10, 11, 12, 13, 14, 15]. The technique has also been successfully tested in 3D computations [16]. This paper is a short version of the

journal paper [9]. We describe the DSD/SST-VMST technique, and present results from test computations.

2 GOVERNING EQUATIONS AND SPACE–TIME FORMULATION OF INCOMPRESSIBLE FLOWS

2.1 Governing equations

Let $\Omega_t \subset \mathbb{R}^{n_{sd}}$ be the spatial domain with boundary Γ_t at time $t \in (0, T)$. The subscript t indicates the time-dependence of the domain. The Navier–Stokes equations of incompressible flows are written on Ω_t and $\forall t \in (0, T)$ as

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u}\mathbf{u}) - \mathbf{f} \right) - \nabla \cdot \boldsymbol{\sigma} = \mathbf{0}, \quad (1)$$

$$\nabla \cdot \mathbf{u} = 0, \quad (2)$$

where ρ , \mathbf{u} and \mathbf{f} are the density, velocity and the external force, respectively. The stress tensor $\boldsymbol{\sigma}$ is defined as $\boldsymbol{\sigma}(p, \mathbf{u}) = -p\mathbf{I} + 2\mu\boldsymbol{\varepsilon}(\mathbf{u})$, with $\boldsymbol{\varepsilon}(\mathbf{u}) = ((\nabla\mathbf{u}) + (\nabla\mathbf{u})^T)/2$. Here p is the pressure, \mathbf{I} is the identity tensor, $\mu = \rho\nu$ is the viscosity, ν is the kinematic viscosity, and $\boldsymbol{\varepsilon}(\mathbf{u})$ is the strain-rate tensor. The essential and natural boundary conditions Eq. (1) are represented as $\mathbf{u} = \mathbf{g}$ on $(\Gamma_t)_g$ and $\mathbf{n} \cdot \boldsymbol{\sigma} = \mathbf{h}$ on $(\Gamma_t)_h$, where $(\Gamma_t)_g$ and $(\Gamma_t)_h$ are complementary subsets of the boundary Γ_t , \mathbf{n} is the unit normal vector, and \mathbf{g} and \mathbf{h} are given functions. A divergence-free velocity field $\mathbf{u}_0(\mathbf{x})$ is specified as the initial condition.

2.2 Space–time variational formulation

A space–time variational formulation of incompressible flows (see for example [6, 17, 18, 7]) is written over a sequence of N space–time slabs Q_n , where Q_n is the slice of the space–time domain between the time levels t_n and t_{n+1} , and P_n is the lateral boundary of Q_n . We denote the trial and test functions spaces for the velocity and pressure as $\mathbf{u} \in \mathcal{S}_{\mathbf{u}}$, $p \in \mathcal{S}_p$, $\mathbf{w} \in \mathcal{V}_{\mathbf{u}}$ and $q \in \mathcal{V}_p$. In deriving the variational formulation, we start with multiplying Eqs. (1) and (2) with the corresponding test functions, integrating them over Q_n , and setting it equal to zero:

$$\int_{Q_n} \mathbf{w} \cdot \rho \left(\frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u}\mathbf{u}) - \mathbf{f} \right) dQ - \int_{Q_n} \mathbf{w} \cdot \nabla \cdot \boldsymbol{\sigma} dQ + \int_{Q_n} q \nabla \cdot \mathbf{u} dQ = 0. \quad (3)$$

We integrate by parts all the terms except for the external force and enforce the essential (i.e. strong Dirichlet) and natural boundary conditions over $(P_n)_g$ and $(P_n)_h$, the complementary subsets of P_n . That gives us the following variational formulation: find $\mathbf{u} \in \mathcal{S}_{\mathbf{u}}$

and $p \in \mathcal{S}_p$ such that $\forall \mathbf{w} \in \mathcal{V}_u$ and $\forall q \in \mathcal{V}_p$

$$\begin{aligned}
 & \int_{\Omega_{n+1}} \mathbf{w}_{n+1}^- \cdot \rho \mathbf{u}_{n+1}^- d\Omega - \int_{\Omega_n} \mathbf{w}_n^+ \cdot \rho \mathbf{u}_n^- d\Omega - \int_{Q_n} \frac{\partial \mathbf{w}}{\partial t} \cdot \rho \mathbf{u} dQ - \int_{(P_n)_h} (\mathbf{w} \cdot \rho \mathbf{u}) (\mathbf{n} \cdot \mathbf{v}) dP \\
 & + \int_{(P_n)_h} (\mathbf{w} \cdot \rho \mathbf{u}) (\mathbf{n} \cdot \mathbf{u}) dP - \int_{Q_n} \nabla \mathbf{w} : \rho \mathbf{u} \mathbf{u} dQ - \int_{Q_n} \mathbf{w} \cdot \rho \mathbf{f} dQ - \int_{(P_n)_h} \mathbf{w} \cdot \mathbf{h} dP \\
 & + \int_{Q_n} \boldsymbol{\varepsilon}(\mathbf{w}) : \boldsymbol{\sigma} dQ + \int_{P_n} \mathbf{q} \mathbf{n} \cdot \mathbf{u} dP - \int_{Q_n} \nabla q \cdot \mathbf{u} dQ = 0,
 \end{aligned} \tag{4}$$

where the notation $(\cdot)_n^-$ and $(\cdot)_n^+$ denotes the values at t_n as approached from below and above, and $\mathbf{v} = \frac{d\mathbf{x}}{dt}$ is the velocity of the spatial-domain boundary.

2.3 Scale separation

In the variational multiscale techniques [10, 11, 12, 13] the ‘‘coarse-scale’’ and ‘‘fine-scale’’ are separated as follows:

$$\mathcal{S}_u = \overline{\mathcal{S}}_u \oplus \mathcal{S}'_u, \quad \mathcal{S}_p = \overline{\mathcal{S}}_p \oplus \mathcal{S}'_p, \quad \mathcal{V}_u = \overline{\mathcal{V}}_u \oplus \mathcal{V}'_u, \quad \mathcal{V}_p = \overline{\mathcal{V}}_p \oplus \mathcal{V}'_p. \tag{5}$$

The coarse-scale part of Eq. (4) is written as follows:

$$\begin{aligned}
 & \int_{\Omega_{n+1}} \overline{\mathbf{w}}_{n+1}^- \cdot \rho \mathbf{u}_{n+1}^- d\Omega - \int_{\Omega_n} \overline{\mathbf{w}}_n^+ \cdot \rho \mathbf{u}_n^- d\Omega - \int_{Q_n} \frac{\partial \overline{\mathbf{w}}}{\partial t} \cdot \rho \mathbf{u} dQ - \int_{(P_n)_h} (\overline{\mathbf{w}} \cdot \rho \mathbf{u}) (\mathbf{n} \cdot \mathbf{v}) dP \\
 & + \int_{(P_n)_h} (\overline{\mathbf{w}} \cdot \rho \mathbf{u}) (\mathbf{n} \cdot \mathbf{u}) dP - \int_{Q_n} \nabla \overline{\mathbf{w}} : \rho \mathbf{u} \mathbf{u} dQ - \int_{Q_n} \overline{\mathbf{w}} \cdot \rho \mathbf{f} dQ - \int_{(P_n)_h} \overline{\mathbf{w}} \cdot \mathbf{h} dP \\
 & + \int_{Q_n} \boldsymbol{\varepsilon}(\overline{\mathbf{w}}) : \boldsymbol{\sigma} dQ + \int_{P_n} \overline{\mathbf{q}} \mathbf{n} \cdot \mathbf{u} dP - \int_{Q_n} \nabla \overline{q} \cdot \mathbf{u} dQ = 0.
 \end{aligned} \tag{6}$$

From [10, 11, 12, 13], the fine-scale solutions are represented by the strong-form residuals of the coarse-scale:

$$\mathbf{u}' = -\frac{\tau_M}{\rho} \mathbf{r}_M(\overline{\mathbf{u}}, \overline{p}), \tag{7}$$

$$p' = -\rho \nu_C r_C(\overline{\mathbf{u}}), \tag{8}$$

where

$$\mathbf{r}_M(\mathbf{u}, p) = \rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} - \mathbf{f} \right) + \nabla p - 2\nabla \cdot \mu \boldsymbol{\varepsilon}(\mathbf{u}), \tag{9}$$

$$r_C(\mathbf{u}) = \nabla \cdot \mathbf{u}, \tag{10}$$

and τ_M and ν_C are stabilization parameters measured in units of time and kinematic viscosity, respectively.

Remark 1 *More on the fine-scale approximation in conjunction with the Green’s operator can be found in [10, 11, 12, 13].*

2.4 DSD/SST formulation

In the DSD/SST method [6, 17, 18, 7, 8], the space–time finite element interpolation functions are continuous within a space–time slab, but discontinuous from one space–time slab to another. The finite-dimensional trial and test functions spaces for the velocity and pressure are denoted as $\mathbf{u}^h \in (\mathcal{S}_{\mathbf{u}}^h)_n$, $p^h \in (\mathcal{S}_p^h)_n$, $\mathbf{w}^h \in (\mathcal{V}_{\mathbf{u}}^h)_n$ and $q^h \in (\mathcal{V}_p^h)_n$.

2.4.1 Fine-scale discretization

The fine-scale solutions are evaluated over each element from Eqs. (7) and (8) with $\mathbf{u}^h \in (\mathcal{S}_{\mathbf{u}}^h)_n$ and $p^h \in (\mathcal{S}_p^h)_n$:

$$\mathbf{u}' = -\frac{\tau_M}{\rho} \mathbf{r}_M(\mathbf{u}^h, p^h), \quad (11)$$

$$p' = -\rho \nu_C r_C(\mathbf{u}^h). \quad (12)$$

Remark 2 *When the polynomial order of the shape functions is less than two, the last term in Eq. (9) vanishes.*

There are various ways of defining τ_M and ν_C . For τ_M we use the definition

$$\tau_M = \tau_{\text{SUPG}}, \quad (13)$$

where τ_{SUPG} comes from [7], specifically the definition as given by Eqs. (107)–(109) in [7], which can also be found as the definition given by Eqs. (7)–(9) in [8]. For ν_C , we consider ν_{LSIC} definition given in [8]:

$$\nu_C = \nu_{\text{LSIC}} = \tau_{\text{SUPG}} \|\mathbf{u}^h - \mathbf{v}^h\|^2, \quad (14)$$

where \mathbf{v}^h is the mesh velocity, and the definition from [14]:

$$\nu_C = \left(\tau_M \sum_{i=1}^{n_{sd}} G_{ii} \right)^{-1}, \quad (15)$$

where

$$G_{ij} = \sum_{k=1}^{n_{sd}} \frac{\partial \xi_k}{\partial x_i} \frac{\partial \xi_k}{\partial x_j}, \quad (16)$$

and $\boldsymbol{\xi}$ is the vector of element coordinates. In our computations we evaluate the stabilization parameters at $\boldsymbol{\xi} = \mathbf{0}$.

Remark 3 *The τ_{SUGN12} component of the τ_{SUPG} definition given by Eqs. (107)–(109) in [7] is the space–time version of the original definition in [19]. These definitions sense, in addition to the element geometry, the order of the interpolation functions. Some τ definitions do that and some do not. The definitions in Sections 3.3.1 and 3.3.2 of [20], for example, are among those that do not.*

Remark 4 Remark 3 is applicable also when the interpolation functions are NURBS functions. This includes classical p -refinement and also k -refinement, except when used in conjunction with periodic B-splines.

Remark 5 In meshes made of NURBS, for quadrilateral (or hexahedral) elements that degenerate to triangles (or tetrahedra), we calculate τ_{SUGN12} , τ_{SUGN1} when applicable, and “ h_{RGN} ” embedded in the τ_{SUGN3} definition in a special way. Instead of letting the sum of the magnitudes involved in the expression degenerate, we first add together the basis functions associated with the coalescing control points, and then apply the expression using the modified basis functions. In other words, we do not degenerate the expression, but instead apply the expression to the degenerated basis functions. This special way is applicable also in the context of finite element meshes.

2.4.2 Coarse-scale discretization

Spatially discretized version of Eq. (6) is written as follows: find $\mathbf{u}^h \in (\mathcal{S}_{\mathbf{u}}^h)_n$ and $p^h \in (\mathcal{S}_p^h)_n$ such that $\forall \mathbf{w}^h \in (\mathcal{V}_{\mathbf{u}}^h)_n$ and $\forall q^h \in (\mathcal{V}_p^h)_n$:

$$\begin{aligned}
 & \int_{\Omega_{n+1}} (\mathbf{w}^h)_{n+1}^- \cdot \rho ((\mathbf{u}^h)_{n+1}^- + (\mathbf{u}')_{n+1}^-) d\Omega - \int_{\Omega_n} (\mathbf{w}^h)_n^+ \cdot \rho ((\mathbf{u}^h)_n^- + (\mathbf{u}')_n^-) d\Omega \\
 & - \int_{Q_n} \frac{\partial \mathbf{w}^h}{\partial t} \cdot \rho (\mathbf{u}^h + \mathbf{u}') dQ + \int_{(P_n)_h} (\mathbf{w}^h \cdot \rho (\mathbf{u}^h + \mathbf{u}')) (\mathbf{n}^h \cdot (\mathbf{u}^h + \mathbf{u}' - \mathbf{v}^h)) dP \\
 & - \int_{Q_n} \nabla \mathbf{w}^h : \rho (\mathbf{u}^h + \mathbf{u}') (\mathbf{u}^h + \mathbf{u}') dQ - \int_{Q_n} \mathbf{w}^h \cdot \rho \mathbf{f}^h dQ - \int_{(P_n)_h} \mathbf{w}^h \cdot \mathbf{h}^h dP \\
 & + \int_{Q_n} \boldsymbol{\varepsilon}(\mathbf{w}^h) : (\boldsymbol{\sigma}(p^h, \mathbf{u}^h) + \boldsymbol{\sigma}') dQ + \int_{P_n} q^h \mathbf{n}^h \cdot (\mathbf{u}^h + \mathbf{u}') dP \\
 & - \int_{Q_n} \nabla q^h \cdot (\mathbf{u}^h + \mathbf{u}') dQ = 0.
 \end{aligned} \tag{17}$$

Here $\boldsymbol{\sigma}' \equiv \boldsymbol{\sigma} - \boldsymbol{\sigma}^h$ is introduced temporarily. We set the fine-scale solution to zero at the spatial and temporal boundaries, use the assumption $\boldsymbol{\varepsilon}(\mathbf{w}^h) : 2\mu \nabla \mathbf{u}' = 0$ (see [12, 21]), and obtain the following form:

$$\begin{aligned}
 & \int_{\Omega_{n+1}} (\mathbf{w}^h)_{n+1}^- \cdot \rho (\mathbf{u}^h)_{n+1}^- d\Omega - \int_{\Omega_n} (\mathbf{w}^h)_n^+ \cdot \rho (\mathbf{u}^h)_n^- d\Omega - \int_{Q_n} \frac{\partial \mathbf{w}^h}{\partial t} \cdot \rho (\mathbf{u}^h + \mathbf{u}') dQ \\
 & + \int_{(P_n)_h} (\mathbf{w}^h \cdot \rho \mathbf{u}^h) (\mathbf{n}^h \cdot (\mathbf{u}^h - \mathbf{v}^h)) dP - \int_{Q_n} \nabla \mathbf{w}^h : \rho (\mathbf{u}^h + \mathbf{u}') (\mathbf{u}^h + \mathbf{u}') dQ \\
 & - \int_{Q_n} \mathbf{w}^h \cdot \rho \mathbf{f}^h dQ - \int_{(P_n)_h} \mathbf{w}^h \cdot \mathbf{h}^h dP + \int_{Q_n} \boldsymbol{\varepsilon}(\mathbf{w}^h) : \boldsymbol{\sigma}(p^h + p', \mathbf{u}^h) dQ \\
 & + \int_{P_n} q^h \mathbf{n}^h \cdot \mathbf{u}^h dP - \int_{Q_n} \nabla q^h \cdot (\mathbf{u}^h + \mathbf{u}') dQ = 0.
 \end{aligned} \tag{18}$$

2.4.3 Comparison with the original DSD/SST formulation

We can further rearrange the terms in the formulation given by Eq. (18) to compare it with the original DSD/SST formulation (with the advection term retained in the conservation-law form) and obtain the following:

$$\begin{aligned}
 & \int_{Q_n} \mathbf{w}^h \cdot \rho \left(\frac{\partial \mathbf{u}^h}{\partial t} + \nabla \cdot (\mathbf{u}^h \mathbf{u}^h) - \mathbf{f}^h \right) dQ + \int_{Q_n} \boldsymbol{\varepsilon}(\mathbf{w}^h) : \boldsymbol{\sigma}(p^h, \mathbf{u}^h) dQ - \int_{(P_n)_h} \mathbf{w}^h \cdot \mathbf{h}^h dP \\
 & + \int_{Q_n} q^h \cdot \nabla \mathbf{u}^h dQ + \int_{\Omega_n} (\mathbf{w}^h)_n^+ \cdot \rho ((\mathbf{u}^h)_n^+ - (\mathbf{u}^h)_n^-) d\Omega \\
 & - \sum_{e=1}^{(n_{el})_n} \int_{Q_n^e} \left[\rho \left(\frac{\partial \mathbf{w}^h}{\partial t} + \mathbf{u}^h \cdot \nabla \mathbf{w}^h \right) + \nabla q^h \right] \cdot \mathbf{u}' dQ - \sum_{e=1}^{(n_{el})_n} \int_{Q_n^e} \nabla \cdot \mathbf{w}^h p' dQ \\
 & - \sum_{e=1}^{(n_{el})_n} \int_{Q_n^e} \rho \mathbf{u}' \cdot (\nabla \mathbf{w}^h) \mathbf{u}^h dQ - \sum_{e=1}^{(n_{el})_n} \int_{Q_n^e} \rho \mathbf{u}' \cdot (\nabla \mathbf{w}^h) \mathbf{u}' dQ = 0. \tag{19}
 \end{aligned}$$

Here each Q_n is decomposed into elements Q_n^e , where $e = 1, 2, \dots, (n_{el})_n$. The subscript n used with n_{el} is for the general case where the number of space–time elements may change from one space–time slab to another.

Remark 6 *The last two terms correspond to the Reynolds stress and cross-stress, respectively. We call this formulation DSD/SST-VMST (i.e. the version with the variational multiscale turbulence model).*

Remark 7 *If we exclude the last two terms, the formulation is the same as the original DSD/SST formulation (with the advection term retained in the conservation-law form) under the conditions that $\tau_{PSPG} = \tau_{SUPG}$ and $\nu_C = \nu_{LSIC}$. The 6th and 7th terms are the SUPG/PSPG and LSIC (least-squares on incompressibility constraint) stabilization terms, respectively. We name this DSD/SST-SUPS (i.e. the version with the SUPG/PSPG stabilization).*

Remark 8 *One of the main differences between the ALE and DSD/SST forms of the variational multiscale method is that the DSD/SST formulation retains the fine-scale time derivative term $\frac{\partial \mathbf{u}'}{\partial t} \Big|_{\boldsymbol{\xi}}$. Dropping this term is called the “quasi-static” assumption (see [15] for the terminology). This is the same as the WTSE option in the DSD/SST formulation (see Remark 2 of [8]). We believe that this makes a significant difference, especially when the polynomial orders in space or time are higher (see Section 6 in [9]).*

3 TEST COMPUTATIONS WITH FLOW PAST AN AIRFOIL

The airfoil is NACA 64-618 and the geometry is approximated with quadratic B-splines. The computational domain is $(-5, 10) \times (-5, 5)$. The leading edge is located at $(0, 0)$.

The angle of attack is 0° . The length and velocity scales are the chord length and inflow velocity, respectively. The Reynolds number is 6.0×10^6 . We compute the problem with the DSD/SST-DP-SUPS and DSD/SST-DP-VMST techniques, using the ν_{LSIC} definition given by Eq. (14) and neglecting the $2\nabla \cdot \mu \boldsymbol{\varepsilon}(\mathbf{u})$ term in Eq. (9). With both techniques, we use two different meshes, one made of quadratic B-splines and one made of linear finite elements. First we manually generate a “frame” control mesh made of quadratic B-splines, which has 8 patches and is shown in Figure 1. Then, by a knot-insertion process that

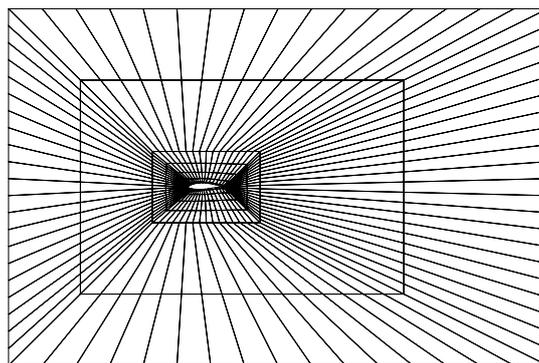


Figure 1: Frame control mesh made of quadratic B-splines. The mesh has 8 patches.

involves little manual intervention, we generate a refined control mesh made of quadratic B-splines, which has 1,681 control points and 1,400 elements. To generate the mesh made of linear finite elements, we start with a quadrilateral mesh generated by interpolating the NURBS geometry at each knot intersection. We subdivide each quadrilateral element into two triangles. The resulting mesh has 1,450 nodes and 2,780 elements. Both meshes are shown in Figure 2. The boundary conditions consist of a uniform velocity at the inflow boundary, zero stress at the outflow boundary, no-slip conditions on the airfoil, and slip conditions at the top and bottom boundaries. The time-step size is 0.01. The number of nonlinear iterations per time step is 3, with 30, 60 and 270 GMRES iterations for the first, second and third nonlinear iterations, respectively. Figures 3–6 show the pressure coefficient and velocity magnitude for the four test computations. Table 1 shows the drag and lift coefficients for the four test computations, together with the measured values from Figure 2a in [22].

4 CONCLUDING REMARKS

A number of multiscale space–time techniques have been developed recently by the T★AFSM for FSI computations, mostly multiscale techniques based on effective ways of dealing with the different spatial or temporal scales that may be involved in the fluid and structure parts of the problem. In addition, recently we have introduced a multiscale space–time technique that is based on representing the different flow scales involved in the fluid mechanics part, thus giving us a good turbulence model for high Reynolds number

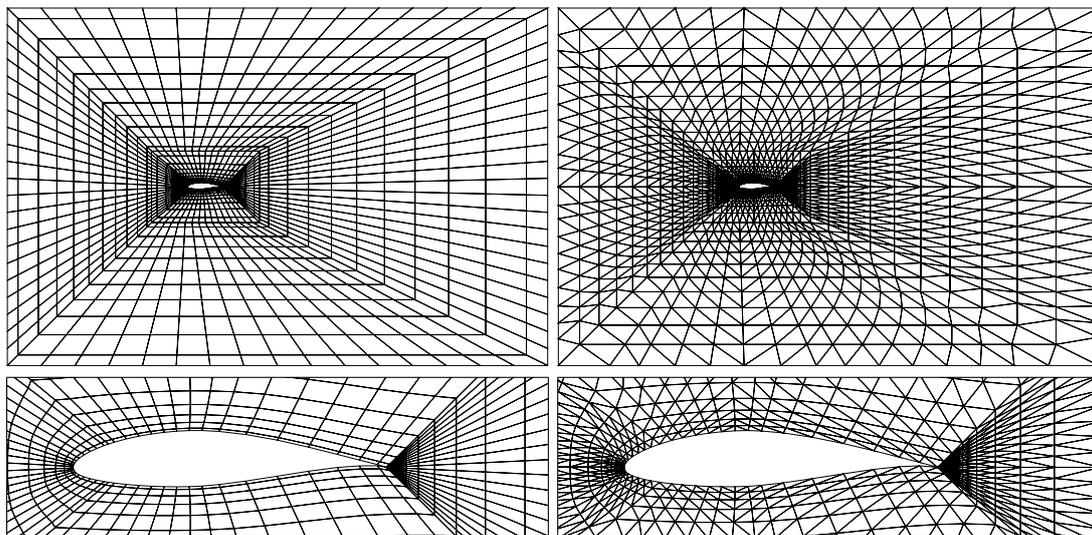


Figure 2: Left: Refined control mesh made of quadratic B-splines (1,681 control points and 1,400 elements). Right: Mesh made of linear finite elements (1,450 nodes and 2,780 elements).

		C_D	C_L
SUPS	Linear Finite Elements	0.0012	0.34
	Quadratic B-Splines	0.0044	0.41
VMST	Linear Finite Elements	0.0010	0.45
	Quadratic B-Splines	0.0032	0.52
Experimental Data From [22]		0.0050	0.45

Table 1: Drag and lift coefficients, C_D and C_L , for the computations and the measured values from [22].

flows. This multiscale technique is the space–time version of the variational multiscale method. We described the technique and presented results from test computations. These computations, and also the 3D computations [16] carried out using the technique, show that the technique is working well even with meshes that would normally be suitable for Reynolds-averaged Navier–Stokes (RANS) type computations. This justifies our expectation that the technique can also be used with meshes that would normally be suitable for a detached-eddy simulation (DES) type computation [23]. The same observation was made in [24] for the residual-based variational multiscale method using ALE and NURBS [15].

ACKNOWLEDGMENT

This work was supported by ARO Grant W911NF-09-1-0346. It was also supported in part by NSF Grant CRCNS-0903949, by NASA Grant NNX09AM89G, and by the Rice Computational Research Cluster funded by NSF Grant CNS-0821727.

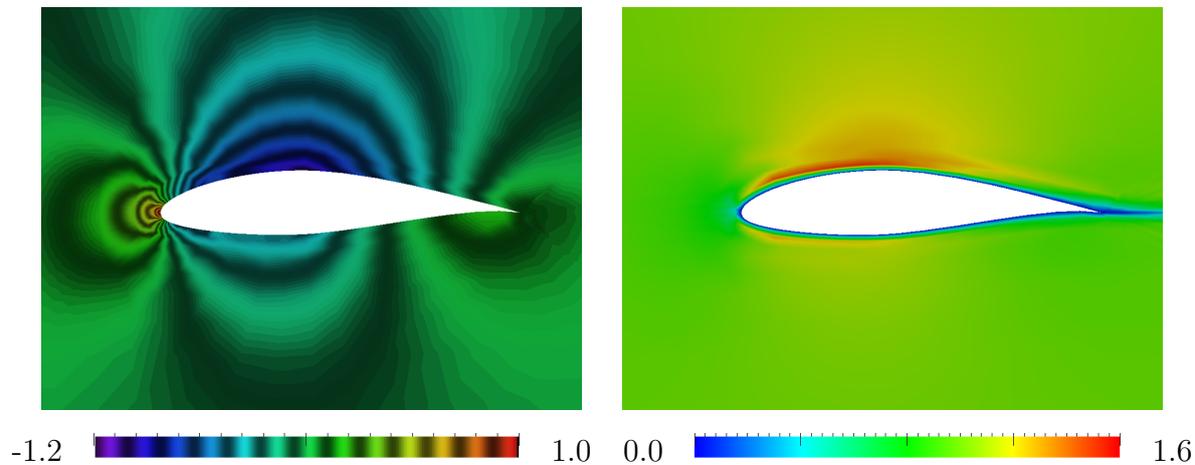


Figure 3: Computed with the DSD/SST-DP-SUPS technique and the mesh made of linear finite elements. Pressure coefficient (left) and velocity magnitude (right).

REFERENCES

- [1] T.E. Tezduyar and S. Sathe, “Enhanced-discretization space-time technique (ED-STT)”, *Computer Methods in Applied Mechanics and Engineering*, **193** (2004) 1385–1401.
- [2] T.E. Tezduyar, M. Schwaab, and S. Sathe, “Sequentially-Coupled Arterial Fluid–Structure Interaction (SCAFSI) technique”, *Computer Methods in Applied Mechanics and Engineering*, **198** (2009) 3524–3533.
- [3] T.E. Tezduyar, K. Takizawa, C. Moorman, S. Wright, and J. Christopher, “Multiscale sequentially-coupled arterial FSI technique”, *Computational Mechanics*, **46** (2010) 17–29.
- [4] T.E. Tezduyar, K. Takizawa, C. Moorman, S. Wright, and J. Christopher, “Space-time finite element computation of complex fluid–structure interactions”, *International Journal for Numerical Methods in Fluids*, **64** (2010) 1201–1218.
- [5] K. Takizawa, S. Wright, C. Moorman, and T.E. Tezduyar, “Fluid–structure interaction modeling of parachute clusters”, *International Journal for Numerical Methods in Fluids*, **65** (2011) 286–307.
- [6] T.E. Tezduyar, “Stabilized finite element formulations for incompressible flow computations”, *Advances in Applied Mechanics*, **28** (1992) 1–44.
- [7] T.E. Tezduyar, “Computation of moving boundaries and interfaces and stabilization parameters”, *International Journal for Numerical Methods in Fluids*, **43** (2003) 555–575.

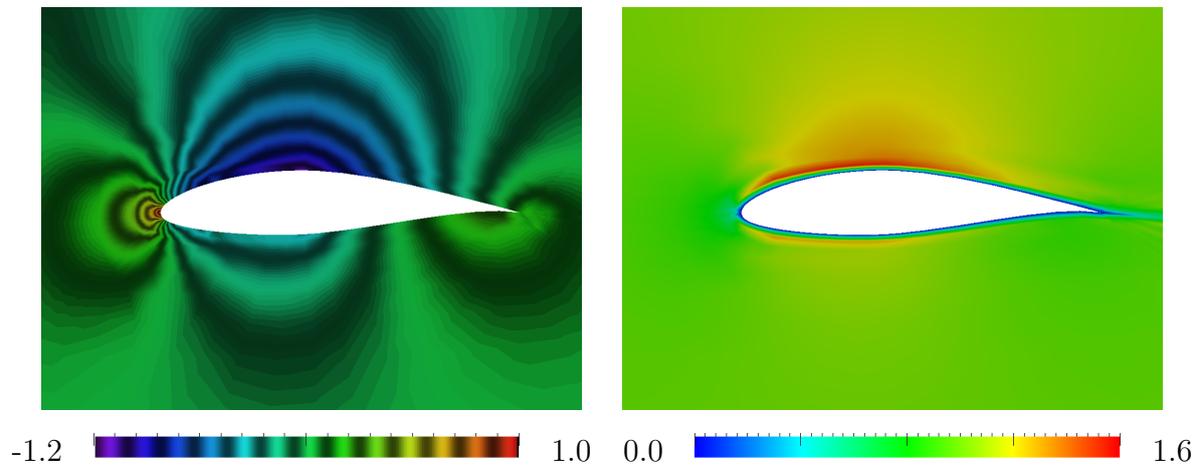


Figure 4: Computed with the DSD/SST-DP-VMST technique and the mesh made of linear finite elements. Pressure coefficient (left) and velocity magnitude (right).

- [8] T.E. Tezduyar and S. Sathe, “Modeling of fluid–structure interactions with the space–time finite elements: Solution techniques”, *International Journal for Numerical Methods in Fluids*, **54** (2007) 855–900.
- [9] K. Takizawa and T.E. Tezduyar, “Multiscale space–time fluid–structure interaction techniques”, *Computational Mechanics*, published online, DOI: 10.1007/s00466-011-0571-z, February 2011.
- [10] T.J.R. Hughes, “Multiscale phenomena: Green’s functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles, and the origins of stabilized methods”, *Computer Methods in Applied Mechanics and Engineering*, **127** (1995) 387–401.
- [11] T.J.R. Hughes, A.A. Oberai, and L. Mazzei, “Large eddy simulation of turbulent channel flows by the variational multiscale method”, *Physics of Fluids*, **13** (2001) 1784–1799.
- [12] T.J.R. Hughes and G. Sangalli, “Variational multiscale analysis: the fine-scale Green’s function, projection, optimization, localization, and stabilized methods”, *SIAM Journal of Numerical Analysis*, **45** (2007) 539–557.
- [13] Y. Bazilevs, V.M. Calo, J.A. Cottrell, T.J.R. Hughes, A. Reali, and G. Scovazzi, “Variational multiscale residual-based turbulence modeling for large eddy simulation of incompressible flows”, *Computer Methods in Applied Mechanics and Engineering*, **197** (2007) 173–201.

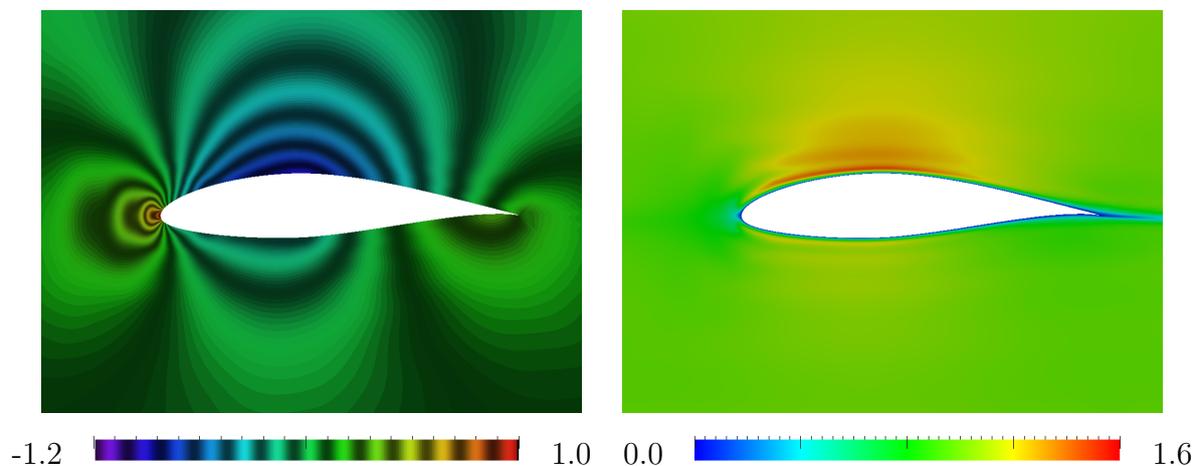


Figure 5: Computed with the DSD/SST-DP-SUPS technique and the mesh made of quadratic B-splines. Pressure coefficient (left) and velocity magnitude (right).

- [14] Y. Bazilevs and I. Akkerman, “Large eddy simulation of turbulent Taylor–Couette flow using isogeometric analysis and the residual–based variational multiscale method”, *Journal of Computational Physics*, **229** (2010) 3402–3414.
- [15] Y. Bazilevs, M.-C. Hsu, I. Akkerman, S. Wright, K. Takizawa, B. Henicke, T. Spelman, and T.E. Tezduyar, “3D simulation of wind turbine rotors at full scale. Part I: Geometry modeling and aerodynamics”, *International Journal for Numerical Methods in Fluids*, **65** (2011) 207–235.
- [16] K. Takizawa, B. Henicke, T.E. Tezduyar, M.-C. Hsu, and Y. Bazilevs, “Stabilized space–time computation of wind-turbine rotor aerodynamics”, *Computational Mechanics*, to appear, 2011.
- [17] T.E. Tezduyar, M. Behr, and J. Liou, “A new strategy for finite element computations involving moving boundaries and interfaces – the deforming-spatial-domain/space–time procedure: I. The concept and the preliminary numerical tests”, *Computer Methods in Applied Mechanics and Engineering*, **94** (1992) 339–351.
- [18] T.E. Tezduyar, M. Behr, S. Mittal, and J. Liou, “A new strategy for finite element computations involving moving boundaries and interfaces – the deforming-spatial-domain/space–time procedure: II. Computation of free-surface flows, two-liquid flows, and flows with drifting cylinders”, *Computer Methods in Applied Mechanics and Engineering*, **94** (1992) 353–371.
- [19] T.E. Tezduyar and Y.J. Park, “Discontinuity capturing finite element formulations for nonlinear convection-diffusion-reaction equations”, *Computer Methods in Applied Mechanics and Engineering*, **59** (1986) 307–325.

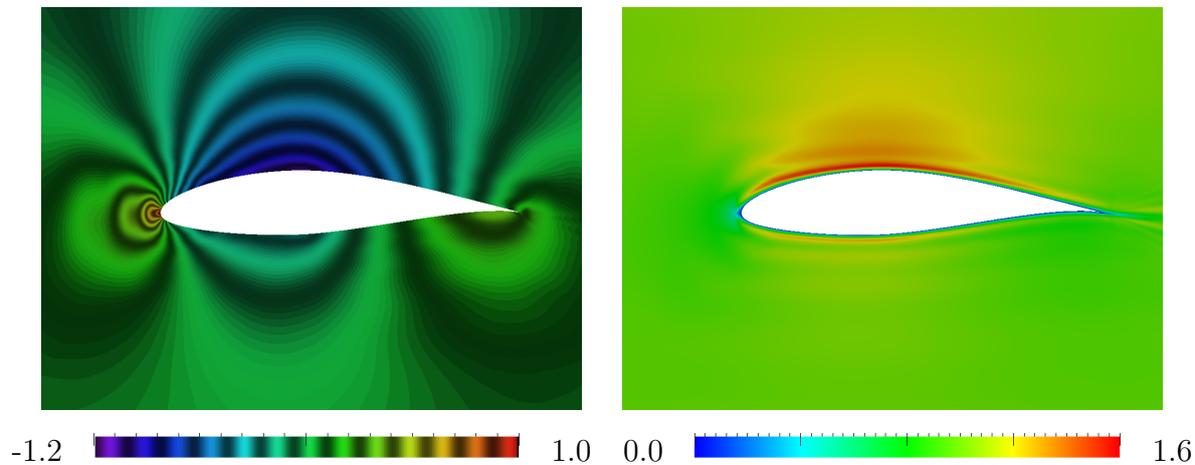


Figure 6: Computed with the DSD/SST-DP-VMST technique and the mesh made of quadratic B-splines. Pressure coefficient (left) and velocity magnitude (right).

- [20] F. Shakib, T.J.R. Hughes, and Z. Johan, “A new finite element formulation for computational fluid dynamics: X. The compressible euler and navier-stokes equations”, *Comput. Methods Appl. Mech. and Engrg.*, **89** (1991) 141–219.
- [21] T.J.R. Hughes and A.A. Oberai, “Calculation of shear stress in Fourier–Galerkin formulations of turbulent channel flows: projection, the Dirichlet filter and conservation”, *Journal of Computational Physics*, **188** (2003) 281–295.
- [22] W.A. Timmer, “An overview of NACA 6-digit airfoil series characteristics with reference to airfoils for large wind turbine blades”, in *Proceedings of AIAA 47th Aerospace Sciences Meeting*, AIAA Paper 2009-268, Orlando, Florida, (2009).
- [23] P.R. Spalart, “Strategies for turbulence modelling and simulations”, *International Journal of Heat and Fluid Flow*, (2000) 252–263.
- [24] Y. Bazilevs, C. Michler, V.M. Calo, and T.J.R. Hughes, “Isogeometric variational multiscale modeling of wall-bounded turbulent flows with weakly enforced boundary conditions on unstretched meshes”, *Computer Methods in Applied Mechanics and Engineering*, **199** (2010) 780–790.

A REFINED 1D FE MODEL FOR THE APPLICATION TO AEROELASTICITY OF COMPOSITE WINGS

Alberto Varello^{*†}, Marco Petrolo^{*†} and Erasmo Carrera^{*}

^{*}Department of Aeronautic and Space Engineering, Politecnico di Torino
Corso Duca degli Abruzzi 24, 10129 Torino, Italy
e-mail: erasmo.carrera@polito.it, www.mul2.com

[†]Institut Jean Le Rond d'Alembert, UMR 7190 CNRS, Univ Paris 6
Case 162, Tour 55-65, 4, Place Jussieu, 75252 Paris, France
e-mail: alberto.varello@polito.it, www.mul2.com

Key words: Aeroelasticity, FE-VLM Coupling, Refined Beam Models, Composite Wings, Unified Formulation

Abstract. The extension of a hierarchical one-dimensional structural model to aeroelasticity is the subject of the present paper. The aerodynamic model is based on the Vortex Lattice Method, VLM, whereas the refined 1D model is based on the Carrera Unified Formulation, CUF. Airfoil in-plane deformation and warping are introduced by enriching the displacement field over the cross-section of the wing. Linear to fourth-order expansions are adopted and classical beam theories (Euler-Bernoulli and Timoshenko) are obtained as particular cases. The VLM aerodynamic theory is coupled with the structural finite element model via an appropriate adaptation of the Infinite Plate Spline method. The aeroelastic tailoring is investigated for several wing configurations (by varying aspect ratio, airfoil geometry and sweep angle) and an excellent agreement with MD NASTRAN solution is provided for structural and aeroelastic cases. The effectiveness of higher-order models for an accurate evaluation of aeroelastic response of isotropic and composite wings is shown.

1 INTRODUCTION

Composite materials are widely used nowadays in a large variety of applications and engineering fields. The advantages related to their spread are becoming so significant that composites are by now a must for state-of-the-art manufacturing technology.

The requirements of weight saving and structural efficiency for aerospace systems such as rotor blades, aircraft wings, and helicopter rotor blades are leading to a wide use of structures in the form of composite thin-walled beams. This makes the accurate evaluation of the response of deformable lifting bodies (LBs) when subjected to steady and unsteady

aerodynamic loadings an even more challenging issue for the aeroelastic design of aerospace vehicles [8]. Over the last decades, many significant contributions have been given in structural, aerodynamic, and aeroelastic coupling modeling [12]. A relevant review article about recent advances in describing fluid-structure interaction is found in Dowell and Hall [9].

Composite beam-like structures can be analyzed by means of beam models provided that a number of non-classical effects such as large torsional warping are incorporated. Detailed structural and aeroelastic models are essential to fully exploit non-classical effects in design due to the properties characterizing advanced composite material structures, such as anisotropy, heterogeneity and transverse shear flexibility [14]. A detailed review of the recent development of beam models can be found in [20]. A considerable amount of work was done in trying to improve the global response of classical beam theories [11, 19] using appropriate shear correction factors, as described by Timoshenko [19]. El Fatmi [10] improved on the displacement field over the beam cross-section by introducing a warping function to refine the description of normal and shear stress of the beam. An asymptotic type expansion in conjunction with variational methods (VABS) was proposed by Yu and co-workers [22]. Generalized beam theories (GBT) have improved classical theories by using a piece-wise beam description of thin-walled sections [18].

Among the various extensions of refined beam models to aeroelasticity, the work done by Librescu and Song [16] on divergence instability of swept-forward wings made of composite materials is mentioned. A thin-walled beam model was implemented which incorporated the anisotropy of the material, transverse shear deformation and warping effects.

Carrera and co-authors have recently proposed refined 1D theories with only generalized displacement variables for the analysis of compact and thin-walled sections/airfoils. Higher-order finite elements are obtained in the framework of the Carrera Unified Formulation, CUF, which considers the order of the model as a free-parameter of the analysis. This technique has been developed over the last decade for plate/shell models [2] and it has recently been extended to beam static and dynamic modeling [4, 5, 6].

The present work couples a refined 1D model based on CUF with the Vortex Lattice Method (VLM) for the analysis of static aeroelastic response of aircraft wings. The aerodynamic load transferring is based on the work presented by Demasi and Livne [8] via the Infinite Plate Spline method [13]. In the proposed formulation (see Varello et al. [21]), the computation of linear steady aerodynamic loads refers to the VLM presented by Katz and Plotkin [15].

2 PRELIMINARIES

A beam-like structure with axial length L is considered. A local cartesian coordinate system is defined in Fig. 1. The cross-section of the structure is Ω and the beam longitudinal axis is along the y direction. The displacement vector of a generic point is:

$$\mathbf{u}(x, y, z) = \left\{ \begin{matrix} u_x & u_y & u_z \end{matrix} \right\}^T \quad (1)$$

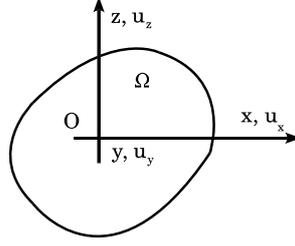


Figure 1: Beam's cross-section geometry and coordinate system.

The stress, $\boldsymbol{\sigma}$, and the strain, $\boldsymbol{\varepsilon}$, are grouped in vectors as follows:

$$\begin{aligned} \boldsymbol{\sigma}_p &= \left\{ \begin{matrix} \sigma_{zz} & \sigma_{xx} & \sigma_{zx} \end{matrix} \right\}^T & \boldsymbol{\varepsilon}_p &= \left\{ \begin{matrix} \varepsilon_{zz} & \varepsilon_{xx} & \varepsilon_{zx} \end{matrix} \right\}^T \\ \boldsymbol{\sigma}_n &= \left\{ \begin{matrix} \sigma_{zy} & \sigma_{xy} & \sigma_{yy} \end{matrix} \right\}^T & \boldsymbol{\varepsilon}_n &= \left\{ \begin{matrix} \varepsilon_{zy} & \varepsilon_{xy} & \varepsilon_{yy} \end{matrix} \right\}^T \end{aligned} \quad (2)$$

The subscripts n and p refer to quantities related to the cross-section Ω and the out-of-plane direction, respectively. In case of small displacements with respect to the length L , the linear strain-displacement relationships hold and are written in a compact notation:

$$\begin{aligned} \boldsymbol{\varepsilon}_p &= \mathbf{D}_p \mathbf{u} \\ \boldsymbol{\varepsilon}_n &= \mathbf{D}_n \mathbf{u} = \mathbf{D}_{np} \mathbf{u} + \mathbf{D}_{ny} \mathbf{u} \end{aligned} \quad (3)$$

where \mathbf{D}_p , \mathbf{D}_{np} , and \mathbf{D}_{ny} are differential matrix operators. Constitutive laws are introduced for beams made of linear elastic orthotropic materials:

$$\begin{aligned} \boldsymbol{\sigma}_p &= \tilde{\mathbf{C}}_{pp} \boldsymbol{\varepsilon}_p + \tilde{\mathbf{C}}_{pn} \boldsymbol{\varepsilon}_n \\ \boldsymbol{\sigma}_n &= \tilde{\mathbf{C}}_{pn}^T \boldsymbol{\varepsilon}_p + \tilde{\mathbf{C}}_{nn} \boldsymbol{\varepsilon}_n \end{aligned} \quad (4)$$

where matrices $\tilde{\mathbf{C}}_{pp}$, $\tilde{\mathbf{C}}_{pn}$, and $\tilde{\mathbf{C}}_{nn}$ are:

$$\tilde{\mathbf{C}}_{pp} = \begin{bmatrix} \tilde{C}_{11} & \tilde{C}_{12} & 0 \\ \tilde{C}_{12} & \tilde{C}_{22} & 0 \\ 0 & 0 & \tilde{C}_{44} \end{bmatrix}, \quad \tilde{\mathbf{C}}_{pn} = \begin{bmatrix} 0 & \tilde{C}_{16} & \tilde{C}_{13} \\ 0 & \tilde{C}_{26} & \tilde{C}_{23} \\ \tilde{C}_{45} & 0 & 0 \end{bmatrix}, \quad \tilde{\mathbf{C}}_{nn} = \begin{bmatrix} \tilde{C}_{55} & 0 & 0 \\ 0 & \tilde{C}_{66} & \tilde{C}_{36} \\ 0 & \tilde{C}_{36} & \tilde{C}_{33} \end{bmatrix} \quad (5)$$

For the sake of brevity, the dependence of the coefficients \tilde{C}_{ij} on Young's moduli, Poisson's ratios, shear moduli, and the fiber orientation angle θ is not reported here. It can be found in the book by Reddy [17]. θ is defined as the angle from the x axis to the 1-material axis on the $x-y$ plane.

3 REFINED ONE-DIMENSIONAL THEORY AND FE FORMULATION

According to the framework of the CUF [2, 7], the displacement field is assumed to be an expansion of generic functions F_τ , which depend on the cross-section coordinates x and z :

$$\mathbf{u}(x, y, z) = F_\tau(x, z) \mathbf{u}_\tau(y) \quad \tau = 1, 2, \dots, N_u = N_u(N) \quad (6)$$

The number of terms N_u depends on the expansion order N , which is a free parameter of the formulation. Mac Laurin's polynomials $x^i z^j$ are chosen as cross-section functions F_τ and hence Eq. 6 is a Taylor-like expansion. Most displacement-based theories can be formulated on the basis of the above generic kinematic field. Classical beam models such as Timoshenko's (TBM) [19] and Euler-Bernoulli's (EBBM) [11] are easily derived from the linear expansion, $N = 1$ [5]. Models having constant and linear distributions of the in-plane displacements components, u_x and u_z , require opportunely reduced material coefficients to overcome Poisson's locking effect [3]. For the sake of brevity, the explicit expression for these coefficients is not reported here, but can be found in [4].

The FE approach is herein adopted to discretize the structure along the y axis. By introducing the shape functions N_i and the nodal displacement vector \mathbf{q} , the displacement field becomes:

$$\mathbf{u}(x, y, z) = F_\tau(x, z) N_i(y) \mathbf{q}_{\tau i} \quad i = 1, 2, \dots, N_N \quad (7)$$

where $\mathbf{q}_{\tau i} = \left\{ q_{u_{x\tau i}} \quad q_{u_{y\tau i}} \quad q_{u_{z\tau i}} \right\}^T$ contains the degrees of freedom of the τ^{th} expansion term corresponding to the i^{th} element node. Elements with a number of nodes N_N equal to 4 are formulated and hereinafter referred as B4. A cubic approximation along the y axis is adopted [1].

The stiffness matrix of the elements is built via the Principle of Virtual Displacements:

$$\delta L_{int} = \int_V (\delta \boldsymbol{\varepsilon}_n^T \boldsymbol{\sigma}_n + \delta \boldsymbol{\varepsilon}_p^T \boldsymbol{\sigma}_p) dV = \delta L_{ext} \quad (8)$$

where L_{int} stands for the strain energy and L_{ext} is the work of external loadings. δ stands for virtual variation. By using Eqs. 3, 4, and 7 the internal virtual work becomes:

$$\delta L_{int} = \delta \mathbf{q}_{\tau i}^T \mathbf{K}^{ij\tau s} \mathbf{q}_{s j} \quad (9)$$

$\mathbf{K}^{ij\tau s}$ is the 3×3 *fundamental nucleus* of the structural stiffness matrix. For the sake of brevity, its explicit expression is not reported here, but it can be found in [20]. It should be noted that no assumptions on the expansion order have been made. Therefore, it is possible to obtain refined beam models without changing the formal expression of the nucleus components. Shear locking is corrected through selective integration [1].

4 STRUCTURAL AND AERODYNAMIC NOTATIONS

The proposed beam model can easily analyze non-planar wing configurations with arbitrary orientation in the 3D space, such as tapered wings with dihedral and sweep angles.

The aerodynamic method here chosen is the Vortex Lattice Method (VLM) [15]. The aerodynamic mesh, which consists in a lattice of N_{AP} quadrilateral panels, lies on a reference trapezoidal surface with 2 edges parallel to the wind direction. As shown in

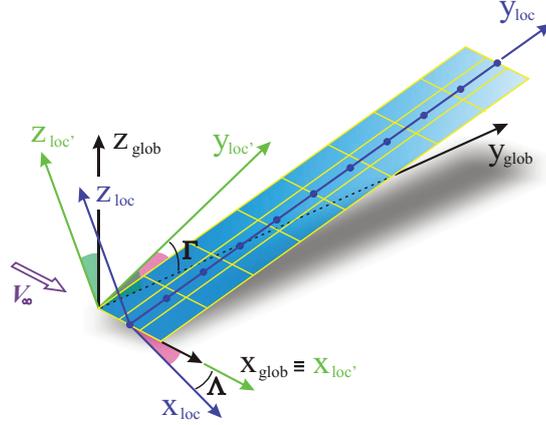


Figure 2: One-dimensional structural mesh and two-dimensional aerodynamic mesh of the wing structure.

Fig. 2, a second coordinate system $x_{loc'} - y_{loc'} - z_{loc'}$ is introduced so that the reference surface lies on the $x_{loc'} - y_{loc'}$ plane. A global coordinate system $x - y - z$ is placed on the airfoil's leading edge point at the root wing section so that x and $x_{loc'}$ axes are both parallel to the free stream velocity \mathbf{V}_∞ (see Fig. 2).

The wing is modeled with a straight beam. The structural FE mesh is contained along the y_{loc} axis, which is on the trapezoidal reference surface. The fundamental nucleus (see Eq. 9) was derived in the local coordinate system. The notation is slightly modified by introducing the subscript “loc” to reflect this fact:

$$\delta L_{int} = \delta \mathbf{q}_{\tau i \text{ loc}}^T \mathbf{K}_{loc}^{ij\tau s} \mathbf{q}_{s j \text{ loc}} = \delta \mathbf{q}_{\tau i}^T [\mathbf{e}^T \cdot \mathbf{K}_{loc}^{ij\tau s} \cdot \mathbf{e}] \mathbf{q}_{s j} = \delta \mathbf{q}_{\tau i}^T \mathbf{K}^{ij\tau s} \mathbf{q}_{s j} \quad (10)$$

where \mathbf{e} is the 3×3 rotation matrix relating the global and the local coordinate systems.

5 SPLINING AND AEROELASTIC FORMULATION

The present advanced beam model allows a very accurate calculation of the displacement field at any point of the three-dimensional wing. Based on this property, the Infinite Plate Spline method [8, 13] was shown (see [21]) to be the ideal choice for the aerodynamic load transferring with the advanced multi-fidelity beam model presented in this work. For the sake of brevity, the final expressions relating the displacements at aerodynamic load points and slopes at aerodynamic control points to the nodal DOFs vector of the whole structure are reported:

$$\tilde{\mathbf{z}}_{loc'} = \tilde{\mathbf{A}}_3^* \cdot \mathbf{q} \quad \frac{d\mathbf{z}_{loc'}}{dx_{loc'}} = \frac{d\mathbf{z}_{loc'}}{dx} = \mathbf{A}_3 \cdot \mathbf{q} \quad (11)$$

The coordinate system for the splining is the local' one. More details about the adaptation of the IPS to the CUF-beam via a set of pseudo-structural points can be found in [21, 20]. The derivation of aerodynamic loads is now faced. According to the VLM [15], the

pressures acting on the wing are transferred as lift forces located on load points of the aerodynamic panels and perpendicular to the wind direction:

$$\mathbf{L} = \frac{1}{2} \rho_{\infty} V_{\infty}^2 \mathbf{I}^D \cdot \Delta \mathbf{p} = \frac{1}{2} \rho_{\infty} V_{\infty}^2 \mathbf{I}^D \cdot [\mathbf{A}^D]^{-1} \cdot \mathbf{w} \quad (12)$$

where $\Delta \mathbf{p}$ contains the dimensionless pressure acting on all the load points, normalized with respect to the dynamic pressure. \mathbf{I}_D contains the panels' geometrical data. The VLM allows the dimensionless normalwash \mathbf{w} , normalized with respect to V_{∞} , to be described as a function of vector $\Delta \mathbf{p}$ by means of the Aerodynamic Influence Coefficient Matrix \mathbf{A}^D .

In the steady case, considering that the structure changes configuration when it deforms, the *boundary condition* used for the Vortex Lattice formulation imposes the dimensionless normalwash to equal the slope at the control points:

$$\mathbf{w} = \frac{d\mathbf{z}_{loc'}}{dx_{loc'}} \quad (13)$$

The transfer from loads at the aerodynamic points to the *energetically* equivalent loads at structural nodes is performed via the Principle of Virtual Displacements. All the lift forces are parallel to the $\mathbf{z}_{loc'}$ axis, hence:

$$\begin{aligned} \delta W &= \delta \tilde{\mathbf{z}}_{loc'}^T \cdot \mathbf{L} = \delta \mathbf{q}^T \cdot \tilde{\mathbf{A}}_3^{*T} \cdot \frac{1}{2} \rho_{\infty} V_{\infty}^2 \mathbf{I}^D \cdot [\mathbf{A}^D]^{-1} \cdot \mathbf{A}_3 \cdot \mathbf{q} = \delta \mathbf{q}^T \cdot \mathbf{L}_{str} \\ &\Rightarrow \mathbf{L}_{str} = -\mathbf{K}_{aero} \cdot \mathbf{q} \end{aligned} \quad (14)$$

where the negative sign is adopted for the sake of convenience. Such a term can go to the left hand side of the aeroelastic equation system and summed up to the structural stiffness matrix:

$$\mathbf{K}_{str} \cdot \mathbf{q} = \mathbf{L}_{str} = -\mathbf{K}_{aero} \cdot \mathbf{q} \quad (15)$$

$$\left[\mathbf{K}_{str} + \mathbf{K}_{aero} \right] \cdot \mathbf{q} = \mathbf{0} \quad \mathbf{K}_{aeroelastic} \cdot \mathbf{q} = \mathbf{0} \quad (16)$$

where $\mathbf{K}_{aeroelastic}$ is the aeroelastic stiffness matrix. The stiffness of the system now takes into account the aerodynamic loads due to the deformed configuration. From Eq. 16 it appears that there is no motion. It occurs because the angle of attack so far considered is zero. To solve this problem, a different from zero angle of attack is assigned to the wing and the corresponding aerodynamic lift forces acting on the panels are computed. The wall tangency condition is now imposed at the control point of each panel by setting the dimensionless normalwash to be equal to the local slope:

$$w = \tan(\pi - \alpha) \quad (17)$$

where the angle of attack α is a small quantity (linear aerodynamic model). Following the same procedure used to build \mathbf{L}_{str} , the energetically equivalent nodal loads \mathbf{L}_{RHS} (the subscript RHS means Right Hand Side) are:

$$\mathbf{L}_{RHS} = \frac{1}{2} \rho_{\infty} V_{\infty}^2 \tan(\pi - \alpha) \tilde{\mathbf{A}}_3^{*T} \cdot \mathbf{I}^D [\mathbf{A}^D]^{-1} \mathbf{d} \quad (18)$$

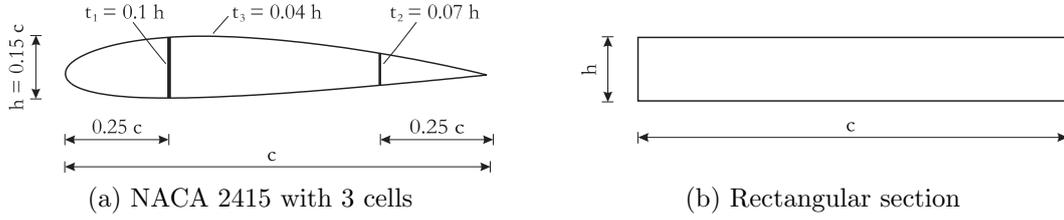


Figure 3: Cross-sections used for the wing configurations.

where \mathbf{d} stands for a $N_{AP} \times 1$ vector of ones. More details about this procedure can be found in [20]. In conclusion, the final aeroelastic equation to be solved is:

$$\mathbf{K}_{\text{aeroelastic}} \cdot \mathbf{q} = \mathbf{L}_{\text{RHS}} \quad (19)$$

Equation 19 allows nodal displacement vector \mathbf{q} to be computed. Now that the right hand side is different from zero, we have a solution.

6 RESULTS AND DISCUSSION

A number of wings with different geometries, layout and loadings are considered. Two different solutions are investigated and compared in this work. The first one coincides with the static structural analysis, hereinafter referred as SSA, and involves only the structural stiffness matrix by disabling the aerodynamic matrix \mathbf{K}_{aero} . The second solution is the static aeroelastic analysis (SAA) which solves the aeroelastic system (Eq. 19) by adding the aerodynamic stiffness matrix to the elastic one. Unless otherwise specified, the wings are subjected to a pure aerodynamic loading (vector \mathbf{L}_{RHS}). Cantilever boundary condition on half-wings is accounted for and the symmetry condition is exploited in the aerodynamic computation.

A swept tapered wing is first taken into account. A half-wing is modeled via a cantilever beam by using the following data: length $L = 5$ m, root chord $c_{\text{root}} = 1.6$ m, taper ratio $\lambda = 0.25$, and sweep angle $\Lambda = +13.5^\circ$. The cross-section is a thin-walled NACA 2415 airfoil, which is subdivided into three cells by two longerons along the spanwise direction at 25% and 75% of the chord (see Fig. 3a). An isotropic aluminium (Young's modulus $E = 69$ GPa and Poisson's ratio $\nu = 0.33$) is considered.

A convergence study is carried out to evaluate the combined effect of the number of finite elements N_{EL} and the expansion order N on the solution. The mechanics of the beam is described in terms of the maximum vertical displacement $u_{z\text{max}}$, which is located at the trailing edge of the tip cross-section. The results for SSA and SAA are shown in Tables 1 and 2, respectively. $u_{z\text{max}}$ increases with N for any mesh, to such an extent that no remarkable differences are detected for high-order expansion. The numerical convergence on N_{EL} is achieved both for SSA and SAA. Linear theories EBBM, TBM, and $N = 1$ are unable to handle any torsional behavior. Poisson's locking correction is not sufficient to make them effective in computing the maximum displacement.

Table 1: Convergence study: effect of the number of $B4$ elements on $u_{z\max}$ [mm] for different beam models. Swept tapered wing. $V_\infty = 50$ m/s, $\alpha = 3^\circ$, 4×40 panels. SSA.

N_{EL}	EBBM	TBM	$N = 1$	$N = 2$	$N = 3$	$N = 4$
2	4.2749	4.2829	4.2909	4.4309	4.9598	5.1036
5	3.1768	3.1842	3.1965	3.6408	3.8045	3.8858
10	3.0401	3.0473	3.0605	3.4701	3.5785	3.6316
20	3.0071	3.0144	3.0277	3.4097	3.4854	3.5377
40	2.9990	3.0062	3.0196	3.3920	3.4440	3.4802
	-13.69 %	-13.49 %	-13.10 %	-2.383 %	-0.886 %	+0.155 %

MD NASTRAN-solid (sol 101): 3.4748

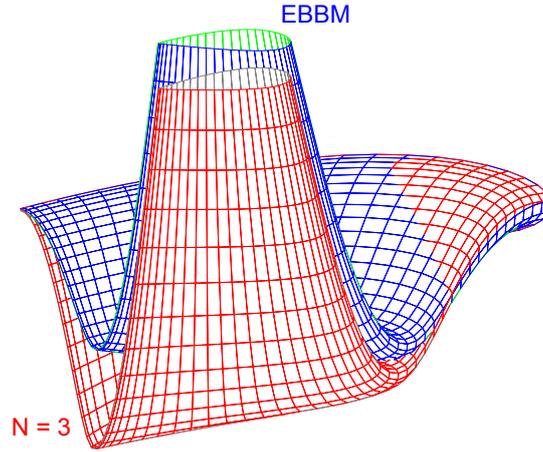

 Figure 4: Tridimensional deformation of the swept tapered wing. $P_z = -7.2$ kN. SAA.

 Table 2: Convergence study: effect of the number of $B4$ elements on $u_{z\max}$ [mm] for different beam models. Swept tapered wing. $V_\infty = 50$ m/s, $\alpha = 3^\circ$, 4×40 panels. SAA.

N_{EL}	EBBM	TBM	$N = 1$	$N = 2$	$N = 3$	$N = 4$
2	4.2747	4.2827	4.2904	4.4236	4.9456	5.0915
5	3.1767	3.1841	3.1959	3.6307	3.7930	3.8743
10	3.0400	3.0472	3.0598	3.4597	3.5678	3.6206
20	3.0071	3.0143	3.0270	3.3993	3.4749	3.5269
40	2.9989	3.0061	3.0189	3.3816	3.4337	3.4695
	-13.40 %	-13.20 %	-12.83 %	-2.353 %	-0.849 %	+0.185 %

MD NASTRAN-shell (sol 144): 3.4631

On the contrary, as far as the SSA is concerned, the refined 1D models (5445 DOFs) are very close to MD NASTRAN-solid results (sol 101 - 10^6 DOFs). Besides, excellent agreement with aeroelastic MD NASTRAN-shell results (sol 144 - 2135 DOFs) is obtained for the aeroelastic analysis (SAA). Figure 3 shows the tridimensional deflection of the wing, drawn by means of a large scale factor, if a bending force ($P_z = -7.2$ kN) is combined to the aerodynamic load. The third-order model clearly illustrates the torsional effect, whereas the limits of EBBM are once again made evident.

Table 3: Effect of the free stream velocity V_∞ [m/s] on $u_{z\max}$ [mm]. Unswept wing with rectangular section. 20 B4 elements. $\alpha = 1^\circ$, 10×50 panels. SSA vs SAA.

	$V_\infty = 30$		$V_\infty = 50$		$V_\infty = 70$	
	SSA	SAA	SSA	SAA	SSA	SAA
<i>EBBM</i>	68.650	68.611	190.693	190.398	373.756	372.628
	-0.05573 %		-0.15459 %		-0.30189 %	
$N = 2$	62.640	68.236	173.999	224.454	341.038	599.986
	+8.93451 %		+28.99720 %		+75.92906 %	
$N = 4$	66.947	73.397	185.963	244.464	364.487	668.562
	+9.63479 %		+31.4584 %		+83.42555 %	

Table 4: Effect of tailoring on the twist $\Delta u_{z\ TIP}$ [mm] of the tip cross-section. Unswept wing with rectangular section. 20 B4 elements. $V_\infty = 50$ m/s, $\alpha = 1^\circ$, 10×50 panels. Orthotropic material. SSA vs SAA.

θ	SSA	SAA	SAA	SAA
	$N = 4$	$N = 4$	NASTRAN	% Diff
-60°	2.9935	3.5297	3.5407	-0.3107
-30°	4.3186	5.5433	5.5408	+0.0451
0°	1.0209	1.0708	1.0741	-0.3072
30°	-2.4578	-2.1753	-2.1638	+0.5315
60°	-1.1199	-1.0576	-1.0439	+1.3124
90°	1.0396	1.0914	1.1065	-1.3647

An unswept wing of length $L = 5$ m with a rectangular cross-section is introduced. Referring to Fig. 3b, the chord c is equal to 1 m and the height is $h = 20$ mm. Table 3 reports a parametric study on $u_{z\max}$, placed at the leading edge of the tip cross-section,

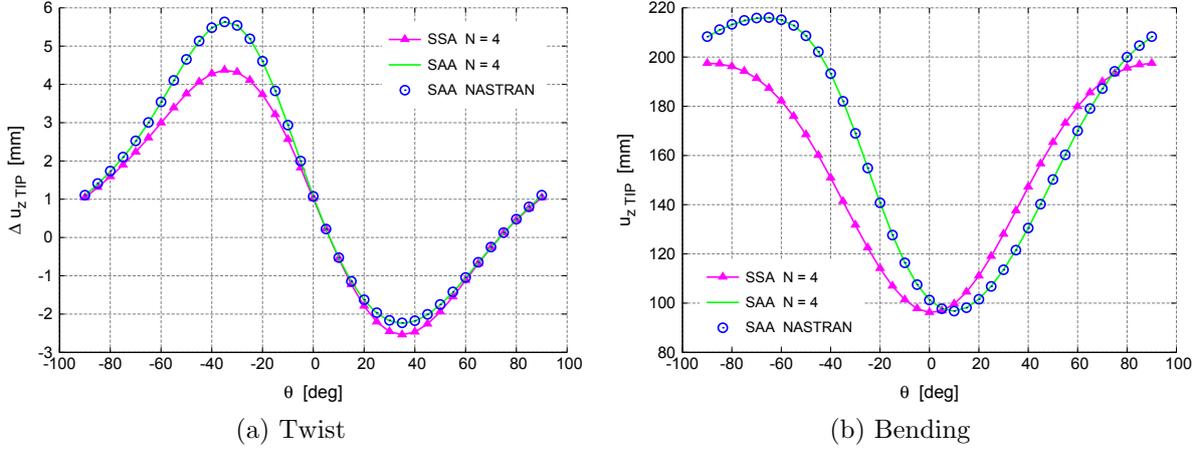


Figure 5: Effect of tailoring on the twist and bending of the tip cross-section. Unswept wing with rectangular section. Orthotropic material. SSA vs SAA.

as the free stream velocity changes. While $u_{z \max}$ increases linearly with the square of V_∞ for SSA, the same does not occur for SAA. The contribution of \mathbf{K}_{aero} to system stiffness becomes more evident as V_∞ increases and the difference from SSA can become very significant for $V_\infty = 70$ m/s. This difference increases with the expansion order N , which enhances the flexibility of the structure. EBBM is ineffective in describing the difference between structural and aeroelastic behavior.

The length L of the unswept wing is further extended to 10 m. The chord c is equal to 1 m and the height of the rectangular section is $h = 100$ mm. A composite material is introduced to analyze the well-known aeroelastic tailoring. Young's modulus along the longitudinal axis E_L is equal to 20.5 GPa, whereas those along the transverse directions are equal to 10 GPa. Poisson's ratio $\nu = 0.25$ and the shear modulus $G = 5$ GPa are the same in all directions. Table 4 shows the effect of the lamination θ on the torsion of the tip cross-section due to the only aerodynamic pressure. The quantity Δu_z is defined as the difference of u_z between leading and trailing edges. The aeroelastic tailoring is more evident for the wing twist evaluation as it is presented in Fig. 5a.

The comparison of SSA and SAA underlines the importance of the contribution of \mathbf{K}_{aero} in the case $N = 4$ to evaluate the aeroelastic behavior of composite wings. While the curve related to SSA is essentially symmetrical with respect to the $\theta = 0^\circ$ lamination, the aeroelastic analysis shows a trend which is far from symmetrical. In general, the aeroelastic analysis leads the twist of the unswept wing to be higher compared to the structural solution as the lamination changes, especially for negative values of θ . The same result occurs for bending behavior as shown in Fig. 5b, where only the SSA case obtains an almost symmetrical curve. The excellent agreement between the fourth-order beam model and MD NASTRAN-shell (sol 144) to describe the aeroelastic response of orthotropic wings with generic orientation is again striking.

7 CONCLUSIONS

This paper has extended the Vortex Lattice Method and a refined one-dimensional structural model for the analysis of anisotropic wings. Advanced 1D finite elements have been obtained via the Carrera Unified Formulation which allows any order theory to be obtained in a hierarchical manner. The static aeroelastic and structural response of wings with different geometries and cross-sections has been analyzed. Isotropic and composite materials has been considered.

The CUF-VLM structural and coupling models have been assessed and compared with MD NASTRAN results. The effectiveness of higher-order models for an accurate analysis of aircraft wings exposed to a free stream has been shown in respect to classical theories. Comparison of structural and aeroelastic solutions has underlined the importance of the contribution of aerodynamic stiffness. The effect of aeroelastic tailoring has been investigated and excellent agreement with MD NASTRAN in evaluating the aeroelastic response of composite wings has been documented. Future works will focus on aeroelastic static and dynamic stability analyses (divergence and flutter).

REFERENCES

- [1] K.J. Bathe. *Finite element procedure*. Prentice hall, Upper Saddle River, New Jersey, 1996.
- [2] E. Carrera. Theories and finite elements for multilayered plates and shells: a unified compact formulation with numerical assessment and benchmarking. *Archives of Computational Methods in Engineering*, 10(3):216–296, 2003.
- [3] E. Carrera and S. Brischetto. Analysis of thickness locking in classical, refined and mixed multilayered plate theories. *Composite Structures*, 82(4):549–562, 2008.
- [4] E. Carrera and G. Giunta. Refined beam theories based on a unified formulation. *International Journal of Applied Mechanics*, 2(1):117–143, 2010.
- [5] E. Carrera, G. Giunta, P. Nali, and M. Petrolo. Refined beam elements with arbitrary cross-section geometries. *Computers and Structures*, 88(5–6):283–293, 2010. DOI: 10.1016/j.compstruc.2009.11.002.
- [6] E. Carrera, M. Petrolo, and A. Varello. Advanced beam formulations for free vibration analysis of conventional and joined wings. *Journal of Aerospace Engineering*, 2010. In press.
- [7] L. Demasi. ∞^3 hierarchy plate theories for thick and thin composite plates: the generalized unified formulation. *Composite Structures*, 84(3):256–270, 2008.

- [8] L. Demasi and E. Livne. Dynamic aeroelasticity of structural nonlinear configurations using linear modally reduced aerodynamic generalized forces. *AIAA Journal*, 47(1):71–90, 2009.
- [9] E.H. Dowell and K.C. Hall. Modeling of fluid-structure interaction. *Annual Review of Fluid Mechanics*, 33:445–490, 2001.
- [10] R. El Fatmi. Non-uniform warping including the effects of torsion and shear forces. Part I: A general beam theory. *International Journal of Solids and Structures*, 44(18-19):5912–5929, 2007.
- [11] L. Euler. *De curvis elasticis*. Bousquet, Lausanne and Geneva, 1744.
- [12] Y.C. Fung. *An Introduction to the Theory of Aeroelasticity*. Dover Publications, 2008.
- [13] R. Harder and R.N. Desmarais. Interpolation using surface splines. *Journal of Aircraft*, 9(2):189–192, 1972.
- [14] K. Kapania and S. Raciti. Recent advances in analysis of laminated beams and plates, part I: Shear effects and buckling. *AIAA Journal*, 27(7):923–935, 1989.
- [15] J. Katz and A. Plotkin. *Low-Speed Aerodynamics*. Cambridge University Press, 2001.
- [16] L. Librescu and O. Song. On the static aeroelastic tailoring of composite aircraft swept wings modelled as thin-walled beam structures. *Composites Engineering*, 2:497–512, 1992.
- [17] J.N. Reddy. *Mechanics of laminated composite plates and shells. Theory and Analysis*. CRC Press, 2nd edition, 2004.
- [18] N. Silvestre and D. Camotim. First-order generalised beam theory for arbitrary orthotropic materials. *Thin-Walled Structures*, 40(9):791–820, 2002.
- [19] S.P. Timoshenko and J.N. Goodier. *Theory of elasticity*. McGraw-Hill, New York, 1970.
- [20] A. Varello, E. Carrera, and L. Demasi. Vortex lattice method coupled with advanced one-dimensional structural models. *Journal of Aeroelasticity and Structural Dynamics*, 2011. Submitted.
- [21] A. Varello, L. Demasi, E. Carrera, and G. Giunta. An improved beam formulation for aeroelastic applications. In *51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, AIAA Paper 2010-3032, Orlando, Florida, 12-15 April 2010.
- [22] W. Yu, V.V. Volovoi, D.H. Hodges, and X. Hong. Validation of the variational asymptotic beam sectional analysis (VABS). *AIAA Journal*, 40(10):2105–2113, 2002.

EXPERIMENTAL BENCHMARK AND NUMERICAL VALIDATION OF A FREE HEAVING AIRFOIL

J.J.H.M. STERENBORG*, A.H. VAN ZUIJLEN* AND H.BIJL*

*Delft University of Technology, faculty of Aerospace Engineering,
Kluyverweg 1 2629 HS Delft,
The Netherlands
e-mail: j.j.h.m.sterenborg@tudelft.nl

Key words: Heave, fluid-structure interaction, experimental benchmark, URANS, panel code, theodorsen

Abstract. In order to validate fluid-structure interaction solvers, a one degree of freedom (1 DOF) aeroelastic experiment is performed. A rigid wing with an harmonically actuated flap, is suspended by springs to allow a free heaving motion. Displacements and time dependent aerodynamic forces are measured for reduced flap frequencies ranging from $k = 0.1$ to $k = 0.3$. Simulations with three codes of different complexity level are performed for validation purposes: theodorsens model, a 2D panel code and 2D URANS, all coupled to a 1 DOF structural model. Results presented by bode diagrams, show differences in both the displacement and the lift between numerical work and experiment. Although there is an offset, consistency is found between displacements, forces and phase angles in the system for all simulations and the experiment.

1 INTRODUCTION

In many applications the interaction between fluid and structures is important to consider. Examples are deformations of aircraft, buildings, bridges and wind turbines due to airloads. Fluid-structure interactions (FSI) can be investigated in several ways. It is very common to use a numerical approach for solving a particular case [1, 2, 6]. However, the problem with many numerical FSI solvers is that the validation with measurement data is relatively unexplored. Some validations are performed, but as far as noticed they all have there restrictions, e.g. very low Reynolds numbers [7] or prescribed motions [3]. In general one can say that without a decent validation it means that it is not proven that the equations that are solved, tackle the problem in the correct way.

Another option is to perform field measurements: existing structures equipped with sensors gather the necessary data. Field measurements are e.g. performed on wind turbines. However, the problem with such methods is that it is almost impossible to

know the exact boundary conditions. This implies the data can often only be used for a qualitative assessment. Validation of numerical solvers is not likely to be possible with these datasets.

A third option is to perform experiments in a wind tunnel with a (scale) model. This provides direct insight in fluid-structure interactions, although also in wind tunnel measurements one has to deal with uncertainties and measurement errors. Benefit is that when uncertainties and measurement errors are considered and quantified, the measurement data can be used for validation purposes.

This research is performed to be able to validate FSI solvers for Reynolds numbers in the order of one million. The focus is on a controlled experiment in the wind tunnel involving FSI. The complexity level is reduced as much as possible and has led to a one degree of freedom aeroelastic experiment: a free heaving airfoil with an actuated 20% chord trailing edge flap. Forces and displacements are measured by sensors.

Next to the experiment, for validation an URANS solver, a 2D panel code and Theodorsens model, all coupled to a structural model, are used to simulate the same test case. The solutions are compared with the experimental results.

2 EXPERIMENT AND METHODS

For both the experiment and the numerical modeling it is beneficial to keep the complexity level as moderate as possible. Therefore, a rigid body is considered with only one degree of freedom: the vertical or heaving motion. Since the structure is rigid the structural model can be modeled by a mass-damper-spring system, as shown in Figure 1.

The wing (mass m) is excited by a time dependent lift force F . The lift force is fluctuating due to a prescribed, harmonic actuated flap motion and the wing motion. By applying different flap frequencies the response of the system and thus the fluid-structure interactions can be assessed. Since the wing is mounted in between side panels, another simplification is made: the flow can be approximated to be two dimensional for angles of attack in the linear part of the $c_l - \alpha$ curve. This assumption is checked by observing the spanwise flow component with tufts covering about 75% of the span width, although some losses of lift near the edges are unavoidable.

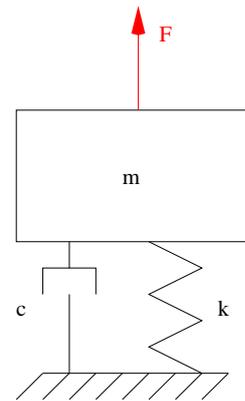


Figure 1: 1 DOF structural model

2.1 EXPERIMENTAL SETUP

The experiment is conducted in the Open Jet Facility (OJF) of Delft University of Technology.

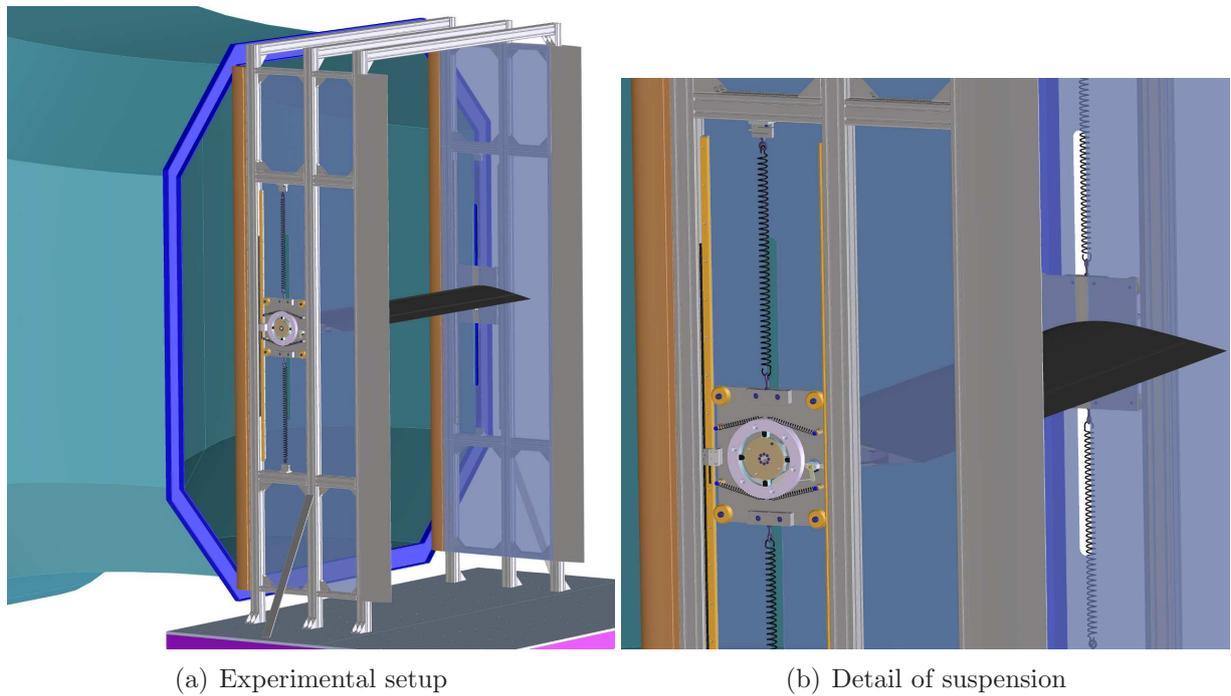


Figure 2: Experimental setup: (a) overview and (b) detail of suspension

The wind tunnel is characterized by an open test section of 2.85m x 2.85m. The typical turbulence level in this tunnel at the measurement wind speed of 21m/s is below 0.5%. At the test section an adjustable flat table is present to mount a test model. The plateau is positioned about 20 cm below the exit, such that the jet can expand also downwards.

On the plateau a supporting frame with a height of 3.2m, a width of 2m and a length of 1.03m is mounted. This frame accommodates sensors, a rails system, spring systems and two side plates that can translate along the rails. In Figure 2(a) an schematic overview is given of the experimental setup.

The wing is attached with bearings to two side plates that can slide along rails in vertical direction, as can be seen in 2(b). The side plates are suspended to the frame by 4 sets of springs, that are pre-tensioned. The amount of pre-tension is not important as long as it is such that the springs for each possible position of the wing are loaded. The springs have a joint theoretical stiffness of $k_s = 8550.3\text{N/m}$. The mechanism for the pitching motion, as can be seen in the same figure is locked. The carbon-fibre wing is an extrusion of a DU96-W180 airfoil with a full span 20% chord trailing edge flap. The flap is hinged at the lower surface and actuated by two servo engines installed in the wing. Turbulator strips are applied at the suction surface at 2% of the chord from the leading edge and on the pressure side at 5% of the chord. The wing has a chord of 0.5m, yielding a Reynolds number of $\text{Re} = 700000$ and the span is 1.8m. The natural frequency of the entire setup is estimated to be 2.58Hz.

Various measurement devices are used in the setup, which are read out with a rate of 2000Hz. An inductive potentiometer in the wing measures the flap angle with an accuracy of $\pm 0.1^\circ$. The angle of attack is kept constant and is measured with an inclinometer with the same accuracy: $\pm 0.1^\circ$. The vertical displacement is measured by a linear encoder with an accuracy of 0.1mm. Forces are measured in two ways: two full wheatstone bridges (strain gauges) on each wing axis measure forces perpendicular and parallel to the wing chord, loadcells on each spring set (4x) measure the vertical force in the springs, with an accuracy of $\pm 0.2\text{N}$. All signals are filtered using a low-pass Butterworth filter with a corner frequency 8Hz above the highest dominating frequency.

2.2 FREQUENCY RESPONSE AND FLAP ACTUATION

For dynamic systems the frequency response can be assessed by applying input signals with different frequencies. For the system under consideration, this means the harmonic flap actuation frequency must be changed. The harmonic motion is characterized by the reduced frequency, according to equation (1):

$$k = \frac{\pi f c}{V}, \quad (1)$$

where f is the flap frequency, c is the chord and V the freestream velocity. The reduced frequency is ranging from $k = 0.1$ up to $k = 0.3$, with steps of $\Delta k = 0.025$. This range is chosen to be around the natural frequency of the system, which is about $k = 0.2$ and with large enough tails to capture the frequency response to a large extent. Hereby, the limitations are given by the servo engines. The flap amplitude is set to $\delta_{flap} = \pm 2^\circ$. The angle of attack is chosen to be $\alpha = 0$, in order to assure attached flow. This implies that when the flap is actuated in a harmonic manner, the responses are harmonic as well.

The frequency response is found by looking at the gain M and the phase angle Ψ between combinations of input and output signals. In the frequency domain, for a standard mass-damper-spring system, the transfer function $G(s)$, gain M and phase angle Ψ can be written as:

$$G(s) = \frac{1}{(s + \zeta\omega_n + j\omega_d)(s + \zeta\omega_n - j\omega_d)}, \quad (2)$$

$$M = |G(s)|, \quad (3)$$

$$\Psi = \angle G(s) = \arctan\left(\frac{\text{Im}[G(s)]}{\text{Re}[G(s)]}\right), \quad (4)$$

with the natural frequency ω_n , the damping ratio ζ and the damped natural frequency

ω_d given by:

$$\begin{aligned}\omega_n &= \sqrt{\frac{k_s}{m}}, \\ \zeta &= \frac{c}{2\sqrt{k_s m}}, \\ \omega_d &= \omega_n \sqrt{1 - \zeta^2}.\end{aligned}\tag{5}$$

In the equations m is the total mass, k_s is the combined spring stiffness and $j = \sqrt{-1}$. The mass-damper-spring system relation (2) is valid in case the lift force is used as input and the displacement is the output, meaning the experimental results can be checked accordingly. However, the lift is determined as a function of the displacement and thus the transfer function (2) is used in the reversed manner. Result is that the found phase lag between lift and displacement is in fact a lead.

2.3 FORCE DERIVATION

The forces measured by the strain gauges (SG) and the load cells (LC) are not the lift and drag forces. Consider the wing on which the lift force L , an added mass force, the drag force D , a vertical inertial force and the strain gauges forces in normal F_n and tangential direction F_t are acting on the left and the right axis. For the horizontal (x-direction) and vertical (y-direction) force equilibrium the following set of equations can be derived:

$$D = (F_{t,left} + F_{t,right}) \cos \alpha + (F_{n,left} + F_{n,right}) \sin \alpha,\tag{6}$$

$$L = (m_{wing} + m_a) \ddot{y} + (F_{n,left} + F_{n,right}) \cos \alpha - (F_{t,left} + F_{t,right}) \sin \alpha.\tag{7}$$

In formula (7) the mass of the wing is given by m_{wing} and the added mass is given by m_a . The added mass force is determined by considering the difference in force responses to a initial excitation of the wing and the force response to a similar initial excitation, but than with a equivalent mass (slender bar) instead of the wing. The added mass is found to be 2.31kg. Using the forces measured by the load cells F_{lc} , equations can be derived for the lift force, making use of the structural model as shown in Figure 1.

$$L = (m_s + m_a) \ddot{y} + (c_a + c_s) \dot{y} + F_{lc},\tag{8}$$

In this equation the damping coefficients c_a and c_s appear. The aerodynamic damping coefficient c_a is found by applying an initial displacement to the wing in tunnel on conditions and measure the decay of the displacement. For the structural damping c_s the wing must be replaced by a bar with equivalent weight and the same procedure can be followed as for obtaining the aerodynamic damping. The total damping coefficient c per test case can also be found by equating relations (7) and (8). The spring stiffness k_s can be checked by measuring the vertical forces and the offset of the wing in vertical direction

from the mid of the structure in tunnel off conditions. For the combined damping coefficient c_{tot} a typical value of about 100Ns/m is observed and for the structural stiffness $k_s = 8225.3\text{N/m}$ is measured. The stiffness k_s is 325N/m less compared to the theoretical value.

3 NUMERICAL WORK

Three simulations are performed with solvers of different complexity level: Theodorsens model, a panel code and an unsteady Reynolds averaged Navier-Stokes (URANS) flow solver, all coupled to a structural model. The simplest model is Theodorsens model, marching in time with a explicit 4th-order Runge-Kutta scheme. It is loosely coupled to the 1 DOF structural model, without the possibility for subiterations.

A more complex aerodynamic model is a free wake 2D panel code. To an existing code a structural model is coupled with an explicit Euler time stepping scheme and the same loose coupling as for Theodorsens model. A routine is implemented to move the flap panels according to the prescribed actuation. The flap motion is performed each time step, before the FSI displacement is computed. The total displacement for each panel is simply the sum of both contributions, since the body is rigid.

The model with the highest complexity level is an URANS flow solver coupled to a generic structure solver. The flow solver handles unstructured, hexahedral meshes and turbulence is taken care of by the Spalart-Allmaras (SA) turbulence model. The SA turbulence model is chosen because it is able to adequately solve attached flow cases. The FSI solver is loosely coupled and allows for subiterations. The flap motion is implemented in the source code in the same manner as for the 2D panel code.

As mentioned for dynamic systems it is important to predict/know the frequency response of the system. Therefore, for all three different codes a similar like interval of reduced frequencies as for the experiment is covered. Time and spatial convergence studies are performed as well, but not treated in this paper.

4 RESULTS

The setup of the experiment allows to concentrate on the force equilibrium and the displacements in the vertical direction. Since the responses are found to behave harmonically according to the input of the flap, all variables are represented as a function of the flap phase angle. The flap phase angle ϕ is given by:

$$\phi = 360(ft - k), \quad k = 0, 1, 2, \dots \quad \text{and} \quad 0^\circ \leq \phi \leq 360^\circ. \quad (9)$$

In this relation t is the time and f the flap frequency. The harmonic behaviour of all responses has another advantage: all signals can be phase-averaged over a large number of samples to average out small fluctuations. This is done for both the experimental results as well as for the numerical work.

For a reduced frequency range of $k = 0.1$ to $k = 0.3$ the focus is on the lift force and the vertical displacement, both as a function of ϕ , and the phase lag between the signals.

Also the phase lag of the signal with respect to the flap deflection is considered. Use is made of the bode plot to summarize the frequency responses of the system variables for the ease of comparison.

4.1 EXPERIMENTAL RESULTS

For each reduced frequency a measurement is performed of about 5 minutes. From this data set, phase averaged signals can be constructed for one full period ($\phi = 0^\circ$ to $\phi = 360^\circ$). In Figure 3 an example is given of a phase averaged result for the lift coefficient for $k = 0.2$ with the phase angle ϕ . Lift coefficients are computed making use of equation (7) and (8). In the same figure also the spread of the data is given by making use of boxplots indicating the the mean and the spread of the data.

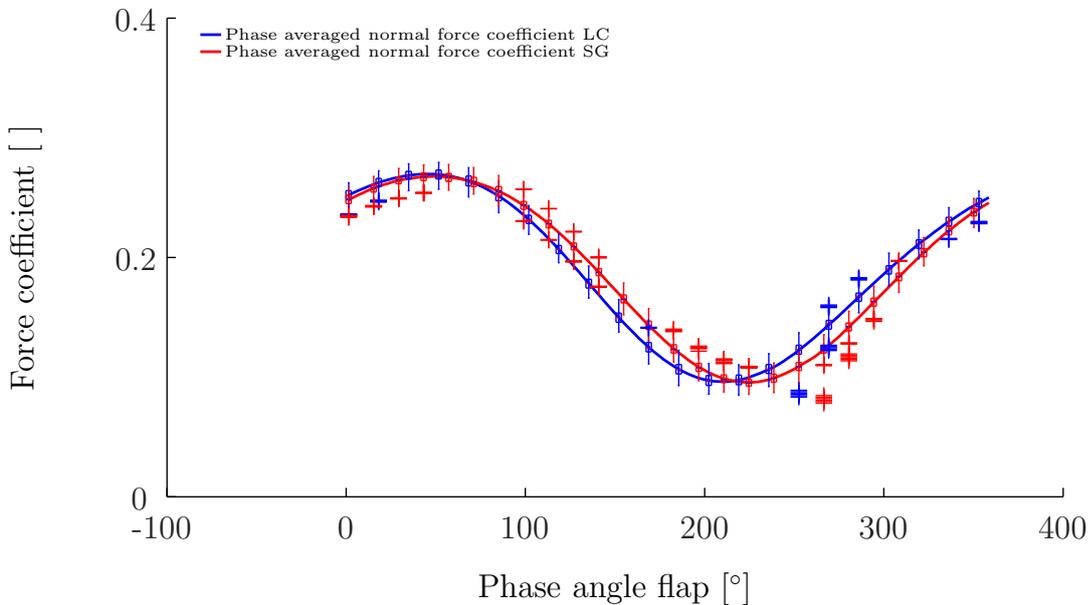


Figure 3: Phase-averaged lift coefficient, $k = 0.2$, $v = 21m/s$, $\delta_{flap} = \pm 2^\circ$

From the figure some important conclusions can be drawn. The deviations of the lift coefficients are acceptable, since 50% of the data is within 4% of the mean lift coefficient and 75% of the data is within about 10% of the mean value. Furthermore, it is noticed there is a small difference in the phase between the lift coefficient based on the strain gauges and the load cells of about maximum 8° , with the load cells leading. Most concerning however, is the mean value of the lift which is about $c_{l,mean} = 0.1878$. This is far from the static lift coefficient $c_{l,\alpha=0} = 0.285$ found in previous measurements. Possible cause might be an offset in the angle of attack or an wind tunnel related change in effective angle of attack.

To get a better understanding of the interaction between the flow and the structure, and than in particular the phase lags, the phase-averaged lift and displacement with the flap phase angle ϕ are plotted together for $k = 0.2$ in Figure 4.

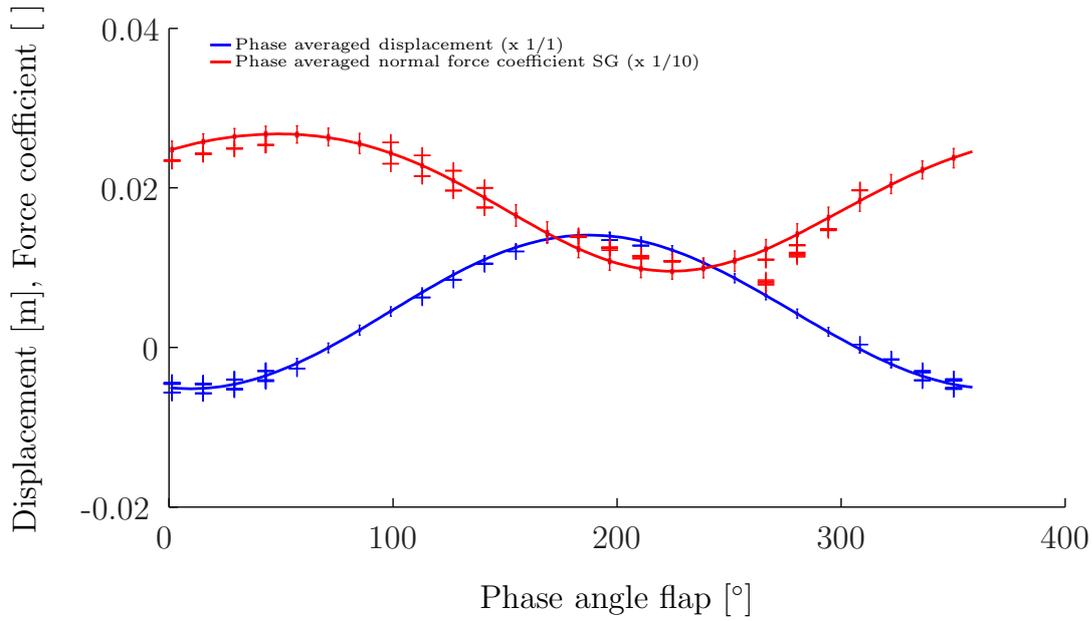


Figure 4: Phase-averaged lift coefficient, $k = 0.2$, $v = 21m/s$, $\delta_{flap} = \pm 2^\circ$

The displacement y is plotted with respect to the initial position (wind off). It can be seen that the displacement is oscillating around a mean value of about $y = 0.0047m$ with an amplitude of $0.0096m$. It can be seen that the lift generation and the vertical displacement of the wing are out of phase with a lag in the lift of about 148° . Observing the same figure, the delays of the aerodynamic force and kinematic response of the wing with respect to the flap deflection is found. Apparently, the maximum lift for a reduced frequency of $k = 0.2$ is reached before the flap is fully deflected downwards. In previous research [5], it was found that in the case of an oscillating flap (no moving wing) the maximum lift occurs at maximum downward flap deflection. This implies that the motion of the wing and the resulting change in the effective angle of attack, causes the shift of the maximum lift with the flap deflection with respect to non-moving wing case.

As mentioned all previous demonstrated results can be combined in a bode diagram for all measurements. In Figure 5 the gain and the phase delay are presented for $k = 0.1$ to $k = 0.3$. Also the exact solution is plotted for the transfer function (2).

The bode diagram shows that for frequency ratios $\ll 1$ the phase delays are small and increase with increasing frequency ratio up to about $\phi = 180^\circ$. Furthermore it can be seen that the trends are according to the exact solution for a mass-damper-spring system.

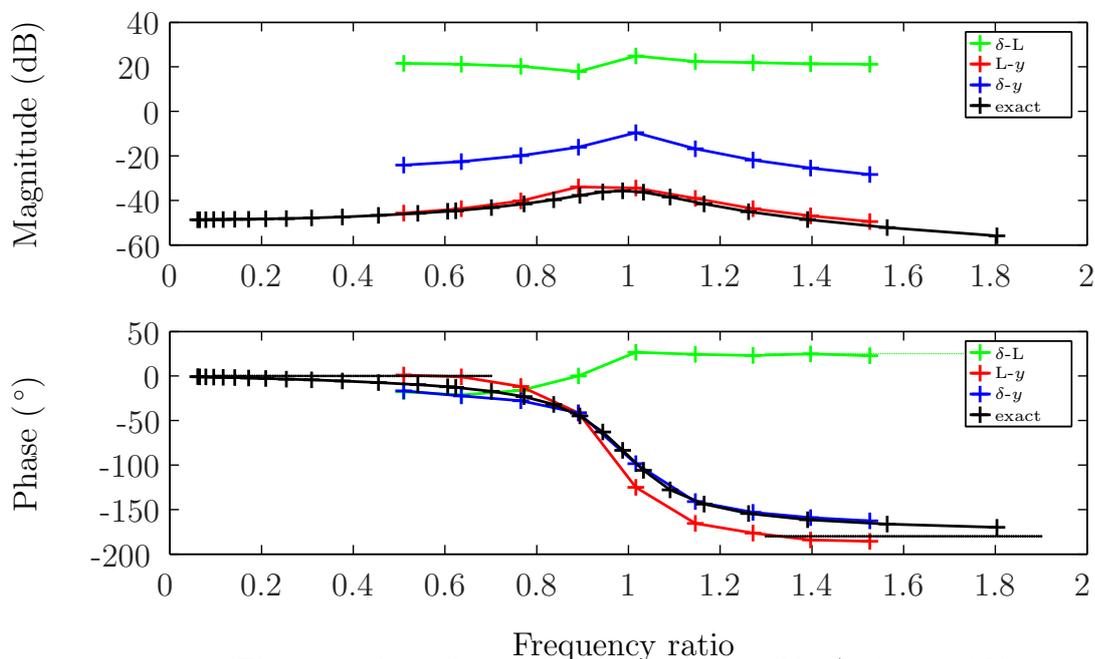


Figure 5: Bode diagram for experiment, $v = 21\text{m/s}$, $\delta_{flap} = \pm 2^\circ$

4.2 NUMERICAL RESULTS

The numerical results are presented for Theodorsens model, the 2D panel code and the URANS model separately. For all simulations the free stream velocity is $v = 21\text{m/s}$ and the flap is deflected for each reduced frequency in a similar fashion as in the experiment: $\delta_{flap} = \pm 2^\circ$. The computations are performed for the same reduced frequencies: $k = 0.1$ to $k = 0.3$. For Theodorsens model and the panel code the structural parameters are scaled for a 2D approach: $m = 16.75\text{kg}$, $c_{tot} = 56.64\text{Ns/m}$ and $k_s = 4922.45\text{N/m}$, whereas the 2D URANS simulation do not need scaling.

Theodorsens model is written according to previous work [4]. It is a 2D approximation without any modeling of wind tunnel or supporting structure interferences. The used time step is $\Delta t = 0.005\text{s}$. In Figure 6 the bode diagram is given for the simulations performed with the coupled Theodorsens model, along with the exact solution.

It follows from the bode diagram that both the magnitude and the phase show a similar tendency as the experimental results and the exact solution. However, only the magnitude of the flap-displacement curve is at nearly the same level as for the experiment. The other curves show quite an offset with respect to the experimental data, which means the predicted lift, as expected, is different compared to the experiment: for $k = 0.2$ $c_{l,mean} = 0.27$ whereas the experimental value is about $c_{l,mean} = 0.185$. As a consequence the displacements is shifted with $\Delta y = 0.0027\text{m}$ such that $y_{mean} = 0.0074\text{m}$. The phase angles are of the same order as the exact solution.

Just like Theodorsens model, the 2D panel code does also not correct for aerodynamic

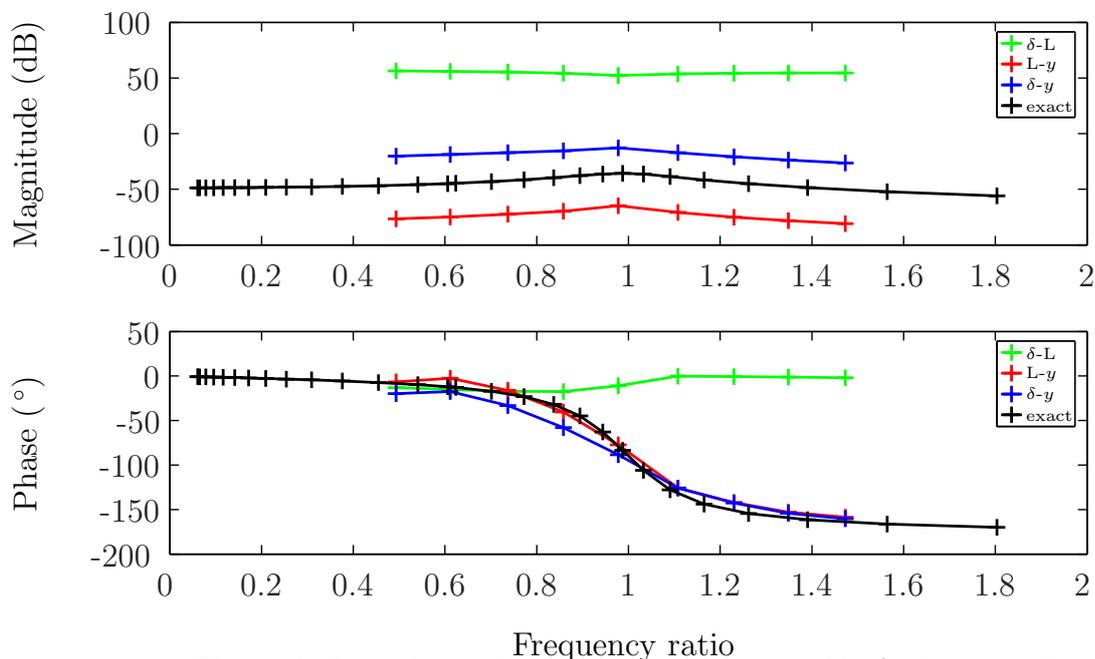


Figure 6: Bode diagram for theodorsens model, $v = 21\text{m/s}$, $\delta_{flap} = \pm 2^\circ$

interference. The airfoil is reconstructed with 200 panels and the free wake contains maximum 1500 vortex elements. The time step is similar as used in Theodorsens model: $\Delta t = 0.005$. Figure 7 shows the bode diagram for the 2D panel code coupled to the 1 DOF structural model.

The gain of the various inputs is very similar to the gains found with Theodorsens model. For the phase angles it is seen that the trends are similar as for the experiment, except for the $L-y$ phase lag for frequency ratios > 1 . For $k = 0.2$ the mean lift coefficient is $c_{l,mean} = 0.255$ and the mean displacement is $y_{mean} = 0.0114\text{m}$. The predicted lift is of the same order as found with Theodorsen, the displacement is not corresponding the change in lift.

The URANS flow solver, with the turbulence modeled by the Spalart-Allmaras equation, is marching in time with $\Delta t = 0.005$. The free stream turbulence level is set to be of the same order as for the experiment. Computations are performed with an unstructured hexahedral mesh containing 70k cells. The bode plot is found in Figure 8.

It can be clearly seen that the results found with the URANS solver and Theodorsens model are very similar. The mean lift coefficient, which is about $c_{l,mean} = 0.299$, is the most off from the experimental lift coefficient, but comparable to the static value of $c_{l,\alpha=0} = 0.285$. The mean displacement is scaled according to the lift coefficient to a value of $y_{mean} = 0.0081\text{m}$. For the phase angles a very close approximation is found with respect to the exact solution, just like the solution found with Theodorsens model.

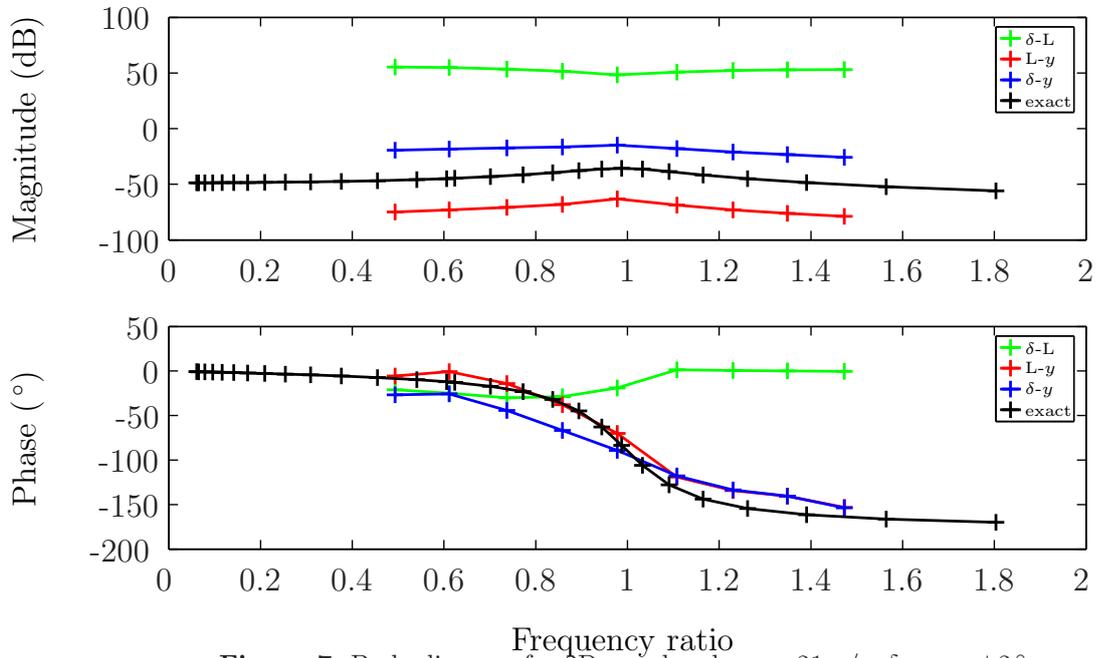


Figure 7: Bode diagram for 2D panel code, $v = 21\text{m/s}$, $\delta_{flap} = \pm 2^\circ$

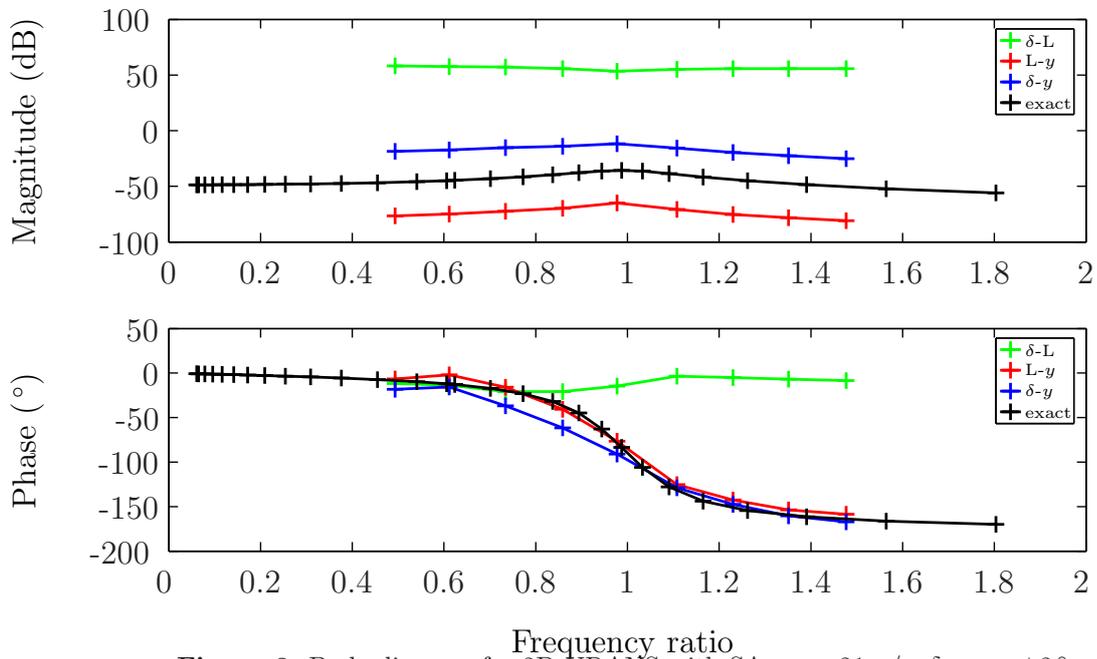


Figure 8: Bode diagram for 2D URANS with SA, $v = 21\text{m/s}$, $\delta_{flap} = \pm 2^\circ$

5 CONCLUSIONS

A 1 DOF aeroelastic experiment is performed: a freely heaving rigid wing with a 20% trailing edge flap. The flap is actuated harmonically with an amplitude of $\delta_{flap} = \pm 2^\circ$

over an reduced frequency range of $k = 0.1$ to $k = 0.3$. Amongst others the aerodynamic lift, the wing displacement and the flap angle are measured. For validation simulations are performed with Theodorsens model, a 2D panel code and a 2D URANS solver with SA turbulence modeling, all coupled to a 1 DOF structural model.

The experiment shows good agreement with the exact solution for a standard mass-damper-spring system. Concerning is the offset of the found mean lift coefficient with respect to the static lift coefficient for the same constant angle of attack: $c_{l,mean} = 0.1878$ and $c_{l,\alpha=0} = 0.285$ respectively. The good coherence of the experimental results with the exact solution indicate there might be an offset in the angle of attack, but further research is needed. The phase angles are according the expectancy: almost no lag for frequencies much smaller than the eigenfrequency and for frequency above the eigenfrequency asymptotic behaviour is found towards $\phi = 180^\circ$.

The numerical solutions confirm that the experiment is off, since the mean lift values are all close to the mean static lift value. The found displacements are consist with the offset in the lift, except for a slightly higher prediction of the displacement for the panel code. The phase angles are close to expectancy.

References

- [1] A.D. Gardner et al. Simulation of Oscillating Airfoils and Moving Flaps Employing the DLR-TAU Unsteady Grid Adaptation. In *New Results in Numerical Fluid and Experimental Fluid Mechanics VI*, volume Vol. 96/2008 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, pages 170–177. Springer Berlin/Heidelberg, 2008.
- [2] Joseph A. Garcia. Numerical investigation of nonlinear aeroelastic effects on flexible high aspect ratio wings. *Journal of aircraft*, 2005.
- [3] P. Gerontakos and T. Lee. Piv study of flow around unsteady airfoil with dynamic trailing-edge flap deflection. *Experiments in Fluids*, 45:955–972, 2008.
- [4] J. Gordon Leishman. Unsteady lift of a flapped airfoil by indicial concepts. *Journal of aircraft*, 31.
- [5] R.C.J. Lindeboom, J.J.H.M. Sterenborg, and C.Simao. Determination of unsteady loads on a du96w180 airfoil with actuated flap using particle image velocimetry. In *The Science of making Torque from Wind*, 2010.
- [6] V. Riziotis, A. Hizanidi, and S.G. Voutsinas. Aeroelastic stability analysis of wind turbine blades using CFD techniques. *NTUA publication*.
- [7] J.M.A. Stijnen, J. de Hart, P.H.M. Bovendeerd, and F.N. van de Vosse. Evaluation of a fictitious domain method for predicting dynamic response of mechanical heart valves. *Journal of fluids and structures*, 19, 2004.

FREQUENCY PARAMETRIZATION TO NUMERICALLY PREDICT FLUTTER IN TURBOMACHINERY

M. Philit*, L. Blanc[†], S. Aubert[‡], W. Lolo[†], P. Ferrand* and F. Thouverez[†]

*Laboratoire de Mécanique des Fluides et d'Acoustique (LMFA)
Université de Lyon, Ecole Centrale de Lyon; Université Lyon 1; INSA Lyon; CNRS UMR 5509
36, avenue Guy de Collongue, 69134 Ecully cedex, France
e-mail: mickael.philit@ec-lyon.fr, webpage: <http://lmfa.ec-lyon.fr>

[†]Laboratoire de Tribologie et de Dynamique des Systèmes (LTDS)
Ecole Centrale de Lyon; CNRS UMR 5513
36 avenue Guy de Collongue, 69134 Ecully cedex, France
e-mail: wilfried.lolo@ec-lyon.fr, webpage: <http://ltds.ec-lyon.fr>

[‡] FLUOREM
Centre scientifique Auguste Moiroux
64 Chemin des Mouilles, 69134 Ecully Cedex, France
e-mail: stephane.aubert@fluorem.com, webpage: <http://www.fluorem.com>

Keywords: Fluid-Structure Coupling, Aeroelasticity, Turbine Blade Design, Flutter Margin Prediction, Meta-modeling

Abstract. In the quest for performance, modern turbomachinery designs are increasingly prone to flutter hazards. Unfortunately, their prediction is currently too expensive and inaccurate for industrial purpose. A significant step towards faster methods would consist in substituting a sequential algorithm to the classical iterative ones encountered in loose coupling strategies. The approach proposed here makes it possible through the use of a meta-model taking into account the sensitivity to design variables. This parametrized method is evaluated on a standard well referenced turbine configuration.

1 INTRODUCTION

Flutter is a critical issue presently facing the designers of turbomachinery blades. Indeed, modern turbomachinery configurations involve lighter and slender blades for efficiency. Meanwhile, the span tends to increase because of higher mass flow. Under these conditions, blade designs are more likely to react to the dynamic loading effects due to unsteady aerodynamics and conversely the flow itself is more likely to be affected by blade motion. In the worst case, a self-excited vibration can develop into resonance until the blade fails. These aeroelastic phenomena should of course be avoided and hence it is extremely important to be able to predict such unsteady aerodynamic interactions. Although experiments are of great importance in the understanding of these phenomena, they are sometimes hazardous. So, numerical simulation is the preferred mean to study flutter.

According to [1], two trends are significant in such simulation code structure. On the one hand, non-linear, fully integrated solvers have been developed for airfoil and wing and can be extended to turbomachinery blades. Fluid and structural domains are discretized together and the associated problems are solved simultaneously. Such tools are really complex and still require huge computer resources. On the other hand, loosely coupled, linearized solvers allow for a faster simplified modelling. They are based on partially integrated solvers with alternate fluid and structure calculations synchronized more or less sophisticatedly depending on the way the data are transferred from one computation to the next via the fluid - structure interface [2, 3, 4]. The process is all the more CPU time-consuming as it is iterative.

The present paper is in line with the traditional approach of loose coupling. Indeed, the high stiffness of targeted blades allows to carry out fluid and structure computations in turn. But, instead of several iterations, the proposed approach is based on a single standard fluid computation followed by a structure computation *supplemented by high order derivatives* which provide information on the sensitivity of the fluid loading to variations of design parameters and operating conditions. The meta-model based on this parametrization is subjected to an optimization process which gives directly the characteristics of the coupled fluid-structure behavior without loop.

The paper outline is as follows. Firstly, theoretical issues are detailed: the proposed strategy is positioned with respect to classical approaches and the parametrization formalism is expressed in the specific case of a single parameter dependency. Secondly, the proposed methodology is supported by an application case to validate the flow parametrization and the ensuing meta-model.

2 A NEW LOOSE COUPLING STRATEGY

2.1 Fluid-structure behavior modelization

Structure and fluid calculations are carried out in turn to characterize the coupled system stability. They share geometric hypothesis. Bladed disks composed of identical sectors are considered. Thus cyclic symmetry properties are used to simulate one reference sector, including a single blade. The contribution of every sector is taken into account through periodic boundary conditions with phase lag, the so-called ‘‘Inter Blade Phase Angle’’ (IBPA).

First, the outline of the applied structure calculations is presented. No structural damping and no other excitation forces than the fluid loading induced by the considered blade displacement are assumed. A finite element discretization of the motion equation, written for the meshed reference sector, leads for a given IBPA value to:

$$\mathbf{M} \frac{d^2 \mathbf{u}}{dt^2} + \mathbf{K} \mathbf{u} = \mathbf{f}, \quad (1)$$

where \mathbf{u} is the vector of local structural displacements relative to equilibrium state under loading, \mathbf{f} is the vector of aerodynamic forces, \mathbf{M} and \mathbf{K} are respectively mass and stiffness complex matrices (taking into account centrifugal effects if necessary). In the framework of linear stability analysis, the unknown \mathbf{u} and the source term \mathbf{f} are modeled as:

$$\mathbf{u} = A \mathbf{\Phi} e^{i\tilde{\omega}t}, \quad (2)$$

$$\mathbf{f} = A \mathbf{\Psi} e^{i\tilde{\omega}t} \quad (3)$$

where A is the displacement amplitude, $\mathbf{\Phi}$ and $\tilde{\omega}$ are the displacement mode shape and complex angular frequency, $\mathbf{\Psi}$ is the fluid modal loading. For a prescribed flow (i.e. $\mathbf{\Psi}$), Eq. 1 is then recasted as the following eigenvalue problem, the unknown being the mode $(\tilde{\omega}, \mathbf{\Phi})$:

$$(\mathbf{K} - \tilde{\omega}^2 \mathbf{M}) \mathbf{\Phi} = \mathbf{\Psi}(\tilde{\omega}, \mathbf{\Phi}). \quad (4)$$

It is to notice that, in general, this equation leads to complex eigenvalues :

$$\tilde{\omega} = \omega + i\alpha \text{ with } \omega, \alpha \in \mathbb{R} \quad (5)$$

For stable configurations, ω is the damped natural angular frequency while α , positive, is the decay rate.

Next, the outline of the applied fluid calculations is presented. A viscous perfect gas subjected to turbulent flows is considered. A finite volume discretization of the Reynolds

Averaged Navier Stokes (RANS) equations supplemented with a turbulence model, written for the meshed fluid domain, leads to:

$$\frac{d}{dt} (\mathbf{J}(\mathbf{p}) \mathbf{q}) + \mathbf{F}(\mathbf{p}, \mathbf{q}) = \mathbf{0} \quad (6)$$

where \mathbf{q} is the vector of nodal conservative and turbulent variables, \mathbf{J} is the vector of mesh cells volume, \mathbf{p} is the vector determining solid walls shape and boundary conditions, \mathbf{F} is the non-linear function expressing the balance of convective and viscous fluxes. In the framework of linear stability analysis, the imposed instantaneous operating conditions \mathbf{p} and the unknown \mathbf{q} are modeled as :

$$\mathbf{p} = \bar{\mathbf{p}} + \delta\mathbf{p}e^{i\omega t}, \quad (7)$$

$$\mathbf{q} = \bar{\mathbf{q}} + \delta\mathbf{q}e^{i\omega t} \quad (8)$$

where $(\bar{\mathbf{p}}, \bar{\mathbf{q}})$ are the mean values and $(\delta\mathbf{p}, \delta\mathbf{q})$ are the small perturbation harmonic amplitudes at the prescribed angular frequency ω defined in Eq. 5. Then, the classical time-linearized Navier-Stokes approximation [5], where higher order terms are neglected, is applied. For a prescribed blade motion (i.e. $\bar{\mathbf{p}}, \delta\mathbf{p}, \omega$), Eq. 6 is then recasted on the one hand as the Steady RANS equations (SRANS, Eq. 9) and on the other hand as the Linearized RANS equations (LRANS, Eq. 10), the unknown being $(\bar{\mathbf{q}}, \delta\mathbf{q})$:

$$\mathbf{F}(\bar{\mathbf{q}}, \bar{\mathbf{p}}) = \mathbf{0} \quad (9)$$

$$\left\{ i\omega\mathbf{J}(\bar{\mathbf{p}}) + \frac{\partial\mathbf{F}}{\partial\mathbf{q}}(\bar{\mathbf{q}}, \bar{\mathbf{p}}) \right\} \delta\mathbf{q} = - \left\{ i\omega\bar{\mathbf{q}}\frac{\partial\mathbf{J}}{\partial\mathbf{p}}(\bar{\mathbf{p}}) + \frac{\partial\mathbf{F}}{\partial\mathbf{p}}(\bar{\mathbf{q}}, \bar{\mathbf{p}}) \right\} \delta\mathbf{p} \quad (10)$$

Thus, fluid-structure coupling arises from the terms Ψ and $\delta\mathbf{p}$. Indeed, in Eq. 4, Ψ is a function of the unsteady fluid behavior and in Eq. 10, $\delta\mathbf{p}$ is a function of the unsteady wall displacement, namely:

$$\Psi = \Psi(\delta\mathbf{q}) \quad (11)$$

$$\delta\mathbf{p} = \delta\mathbf{p}(\tilde{\omega}, \Phi) \quad (12)$$

Furthermore, $\tilde{\omega}$ is shared by Eqs. 4 and 10, but it is an output of Eq. 4 whereas its real part is an input of Eq. 10.

As a conclusion, to characterize the fluid-structure coupled system stability, the unknown to be calculated are $(\tilde{\omega}, \Phi, \delta\mathbf{q})$ from Eqs. 4, 10, 11, 12. From these, the balance of mechanical energy exchanged between the fluid and the blade over one cycle will determine if the unsteady fluid loading is prone to amplify ($\alpha < 0$) or to damp ($\alpha > 0$) the blade oscillation.

2.2 Classical approaches

Modes calculation is often initialized in vacuum, i.e. $\Psi = 0$. Let $(\tilde{\omega}_0, \Phi_0)$ be a solution of interest of Eq. 4 for such conditions. Usually, it is assumed that the mode shape Φ_0 remains unaffected by the flow : only the eigenfrequency $\tilde{\omega}$ is modified. As a result, the angular frequency shift, $\Delta\tilde{\omega} = \tilde{\omega} - \tilde{\omega}_0$, is sought instead of $\tilde{\omega}$ itself. $\Delta\tilde{\omega}$ is evaluated from the residual of Eq. 4 projected on the mode shape of interest Φ_0 :

$$R(\Delta\tilde{\omega}) = \Phi_0^H (\mathbf{K} - (\tilde{\omega}_0 + \Delta\tilde{\omega})^2 \mathbf{M}) \Phi_0 - \Phi_0^H \Psi(\tilde{\omega}_0 + \Delta\tilde{\omega}, \Phi_0). \quad (13)$$

where superscript H indicates the Hermitian conjugate. It is to notice that obviously, $R(0) \neq 0$ if Φ_0 and $\Psi(\tilde{\omega}_0, \Phi_0)$ are not orthogonal. $\Delta\tilde{\omega}$ is sought such that it minimizes $R(\Delta\tilde{\omega})$, i.e. as the solution of:

$$\frac{\partial R^H R}{\partial \Delta\tilde{\omega}} = 0. \quad (14)$$

Ψ being an implicate intricate function of $\Delta\tilde{\omega}$, an iterative procedure is classically applied. $R(\Delta\tilde{\omega})$ is replaced by:

$$R^{(n)} = \Phi_0^H \left(\mathbf{K} - (\tilde{\omega}_0 + \Delta\tilde{\omega}^{(n)})^2 \mathbf{M} \right) \Phi_0 - \Phi_0^H \Psi(\tilde{\omega}_0 + \Delta\tilde{\omega}^{(n-1)}, \Phi_0). \quad (15)$$

$\Psi(\tilde{\omega}_0 + \Delta\tilde{\omega}^{(n-1)}, \Phi_0)$ is computed from Eqs. 11 and 10, using the previous iteration angular frequency $\tilde{\omega}^{(n-1)} = \tilde{\omega}_0 + \Delta\tilde{\omega}^{(n-1)}$. Then, $\Delta\tilde{\omega}^{(n)}$ is updated from Eq. 14 recasted as:

$$(\tilde{\omega}_0 + \Delta\tilde{\omega}^{(n)}) \Phi_0^H \mathbf{M} \Phi_0 R^{(n)} = 0 \quad (16)$$

Such a step from $\Delta\tilde{\omega}^{(n-1)}$ to $\Delta\tilde{\omega}^{(n)}$ is repeated as long as the projected fluid modal loading $\Phi_0^H \Psi$ changes significantly. The convergence of this procedure might be slow, or even hazardous, according to the dependence of Ψ on $\Delta\tilde{\omega}$.

2.3 Proposed parametrized sequential approach

The starting point of the proposed approach consists in modelizing the fluid modal loading by a truncated Taylor series expansion [6], as:

$$\Psi(\tilde{\omega}_0 + \Delta\tilde{\omega}) \simeq \Psi_0(\tilde{\omega}_0) + \frac{\partial \Psi}{\partial \omega}(\tilde{\omega}_0) \Delta\tilde{\omega} + \frac{1}{2} \frac{\partial^2 \Psi}{\partial \omega^2}(\tilde{\omega}_0) \Delta\tilde{\omega}^2 \quad (17)$$

Minimizing the residual $R(\Delta\tilde{\omega})$ reduces then in finding once the roots of the third order polynomial $\frac{\partial R^H R}{\partial \Delta\tilde{\omega}} = 0$ (Eq. 14), without further iterations. The novelty is in computing the derivatives of Ψ from those of $\delta \mathbf{q}$. Namely, deriving Eqs. 11 and 10 yields for first order terms:

$$\frac{\partial \Psi}{\partial \omega}(\tilde{\omega}_0) = \frac{\partial \Psi}{\partial \delta \mathbf{q}}(\delta \mathbf{q}_0) \frac{\partial \delta \mathbf{q}}{\partial \omega}(\tilde{\omega}_0) \quad (18)$$

with

$$\left\{ i\omega_0 \mathbf{J}(\bar{\mathbf{p}}) + \frac{\partial \mathbf{F}}{\partial \mathbf{q}}(\bar{\mathbf{q}}, \bar{\mathbf{p}}) \right\} \frac{\partial(\delta \mathbf{q})}{\partial \omega} = - \left\{ i\omega_0 \bar{\mathbf{q}} \frac{\partial \mathbf{J}}{\partial \mathbf{p}}(\bar{\mathbf{p}}) + \frac{\partial \mathbf{F}}{\partial \mathbf{p}}(\bar{\mathbf{q}}, \bar{\mathbf{p}}) \right\} \frac{\partial(\delta \mathbf{p})}{\partial \omega} - i \left\{ \mathbf{J}(\bar{\mathbf{p}}) \delta \mathbf{q}_0 + \bar{\mathbf{q}} \frac{\partial \mathbf{J}}{\partial \mathbf{p}}(\bar{\mathbf{p}}) \delta \mathbf{p}_0 \right\} \quad (19)$$

Recursive derivations yield then to higher order terms, such as $\frac{\partial^2 \Psi}{\partial \omega^2}$. From the initial solution in vacuum ($\tilde{\omega}_0, \Phi_0$), the proposed approach builds thus a parametrized meta-model of the flow unsteadiness comprising $\left(\delta \mathbf{q}_0, \frac{\partial \delta \mathbf{q}}{\partial \omega}, \frac{\partial^2 \delta \mathbf{q}}{\partial \omega^2}, \dots \right)$, which leads to a meta-model of the fluid loading $\left(\Psi_0, \frac{\partial \Psi}{\partial \omega}, \frac{\partial^2 \Psi}{\partial \omega^2}, \dots \right)$. Using a Taylor series expansion, the residual $R(\Delta \tilde{\omega})$ (Eq. 13) is recasted as a polynomial in $\Delta \tilde{\omega}$, easily minimized. Iterations between structure and fluid calculations are therefore avoided. It is to notice that the left hand side matrices in Eqs. 10 and 19 are identical. So, $\left(\delta \mathbf{q}_0, \frac{\partial \delta \mathbf{q}}{\partial \omega}, \frac{\partial^2 \delta \mathbf{q}}{\partial \omega^2}, \dots \right)$ are solutions of the same linear system with multiple RHS. Therefore, computational time can be significantly reduced by taking advantage of this feature.

2.4 Numerical methods

In the present paper, mechanical Eqs. 4 and 14 are solved with AnsysTM tools and MatlabTM. The steady flow solution (Eq. 6) is computed with TurbflowTM solver [7, 8]. Eq. 10 is solved by the Fluorem's LRANS solver called Turb'LinTM, while Fluorem's Parametrized RANS(PRANS) solver called Turb'SensTM is used for Eq. 19.

The same numerical setup is shared by the three flow solvers. Fluid domain is discretized with multi-block structured meshes. Convective fluxes are evaluated with JST scheme [9] in which blade motion is taken into account through grid deformation (ALE formulation). Turbulence effects are modeled by Kok $k-\omega$ model [10], but turbulent variables are assumed constant during LRANS and PRANS computations following the classical "frozen turbulence" hypothesis. To solve the linear systems, a GMRES Krylov algorithm is used instead of a more usual time marching algorithm as it is proven to be more stable and robust [11].

The partial derivatives in Eqs. 10 and 19 are exactly computed by a set of functions generated by Fluorem "in house" Automatic Derivation (AD) Tool, applied to the core \mathbf{F} function (Eq. 6) of Turb'Flow. This tool was originally intended to parametrize steady flow solutions [12, 13].

3 VALIDATION ON A TURBINE CONFIGURATION

3.1 Test case overview

A cascade aeroelastic configuration is chosen so that it provides realistic flow boundary conditions and can exhibit flutter properties. It is a well referenced subsonic turbine configuration (Standard Configuration number 4 in [14]).

Results of the measurements are reported for four different IBPAs (270°, 180°, 90° and 0°) by Bölcs and Fransson [14]. The experimental apparatus involves a highly loaded annular cascade of blades that are bending in a controlled oscillatory manner. The frequency of the oscillation remains nearly constant as well as the amplitude. Investigations have been performed for the first bending mode at a constant exit Mach number of 0.9. The experimental reduced frequency is then $k = 0.107$. The full finite element model and the single sector simulated are presented on Fig. 1. Fig. 2 presents the steady Mach number distribution obtained when solving Eq. 6.

The local response of the flow is quantified through the complex pressure coefficient defined for a surface element ($\mathbf{n}dS$) of the blade, as:

$$\tilde{c}_p = \frac{\Psi(\Delta\tilde{\omega}) \cdot \mathbf{n}}{dS(P_t - P_s)} \quad (20)$$

where $P_t - P_s$ is the dynamic pressure taken upstream.

The global energy exchange between the blade and the flow is resolved according to Fransson [14]. The unsteady aerodynamic work coefficient is integrated over the entire blade for a cycle of oscillation yielding to the aerodynamic damping coefficient :

$$\Xi = \frac{-1}{(P_t - P_s)} \Im(\Phi_0^H \Psi(\Delta\tilde{\omega})) \quad (21)$$

By definition, if the coefficient is negative, the forcing of the blade adds energy making the system unstable. Thus, this coefficient will be evaluated for illustrative purpose.

3.2 Coupling validation

Herein, the new methodology is demonstrated for an IBPA of 0 degree with regards to the frequency parameter. Mode shape, IBPA and operating conditions are held fixed. According to experimental data, IBPA 0 case is near the stability frontier.

Following Eq. 4, a modal analysis in vacuum provides the first bending mode shown in Fig. 1. Modal frequency is computed at 149.91 Hz. The result is consistent with experimental data.

Fig. 3 shows the unsteady pressure coefficient defined in Eq. 20, represented through its module and phase in degrees over the blade as well as its derivatives with respect to angular frequency (computed with Eqs. 10 and 19).

In order to investigate the validity of Eq. 17 approximation, three characteristic locations on the blade are chosen as illustration. Numerical probes are thus picked out at the leading edge, and on the pressure and suction sides (green, red and blue points displayed in Fig 2 respectively). These three points have been chosen because of their different sensitivity to the steady aerodynamic flow (Fig. 3). On the Fig. 4, the \tilde{c}_p module behaves almost as a parabolic curve and the phase is non-linear. At leading edge the module and phase depend quasi-linearly on the frequency as shown in Fig. 4. On the pressure side, the phase has a parabolic profile.

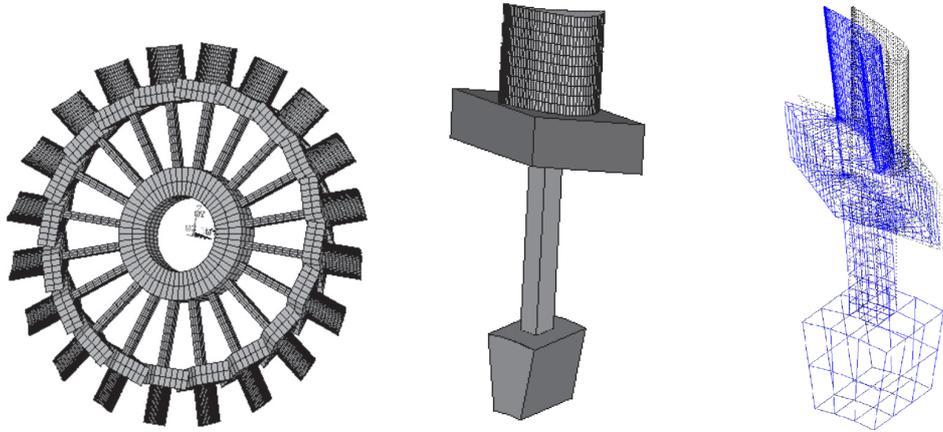


Figure 1: Model of the bladed disk: complete model (left) , reference sector (middle), targeted mode shape (right)

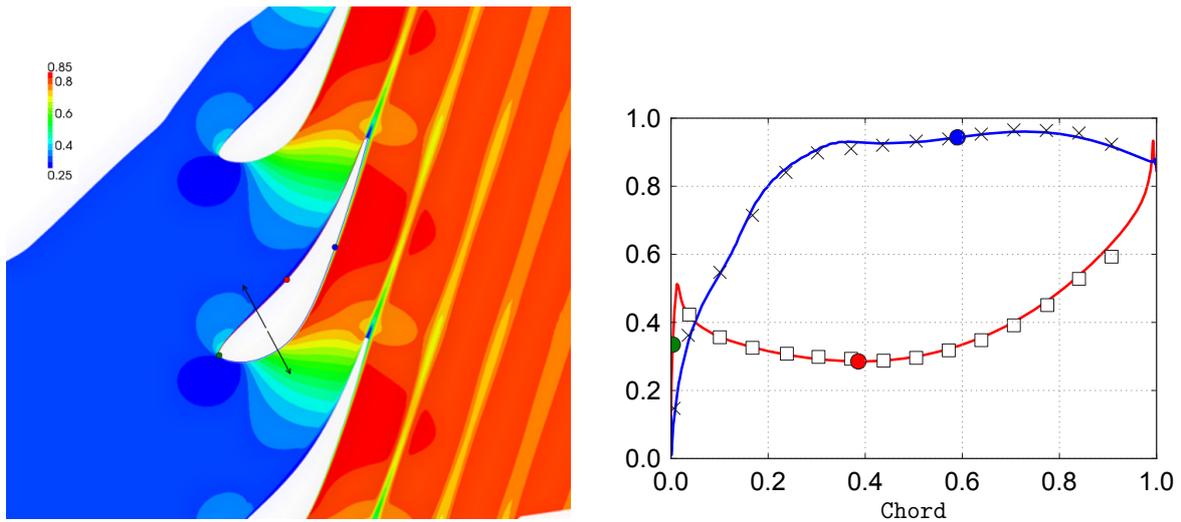


Figure 2: Mach number in blade-to-blade view

Overall, the approximant is well suited. At each position on the blade, the different approximations (lines) are obtained by extrapolating \tilde{c}_p from Eqs. 20 and 17. They agree very well with direct computation of \tilde{c}_p (symbols) using only Eq. 10, and particularly on the phase component. This is promising because the stability analysis is highly related to the phase status. From a local aerodynamic point of view, the model behaves well, it is thus expected to do as well from a global point of view.

As the result of this parametric study, the damping coefficient (Eq. 21) versus the oscillation frequency is presented in Fig. 5. The first order model takes only into account the first derivative of loadings giving a rough idea on the linear local (in terms of angular

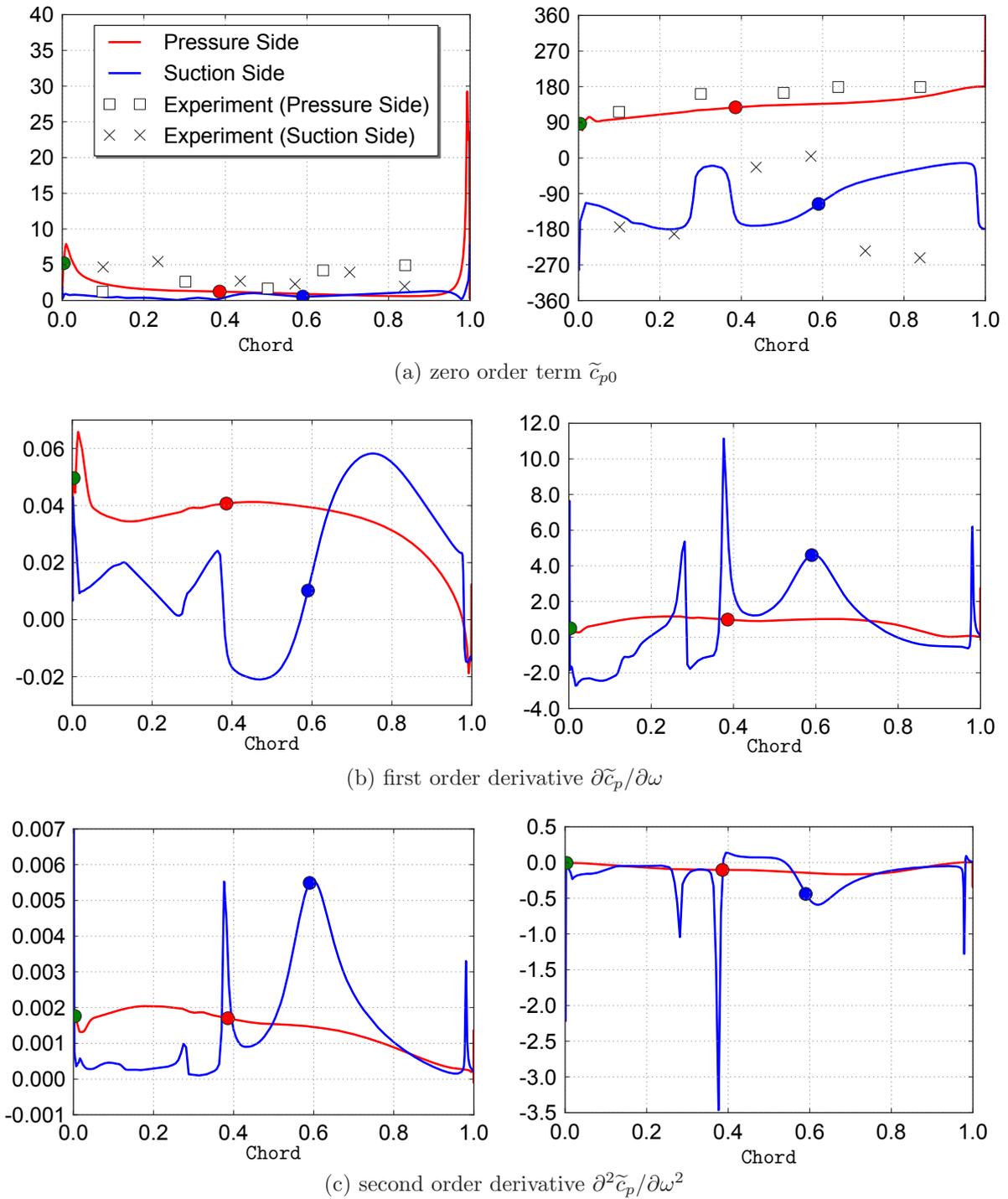


Figure 3: Amplitude (left) and phase (right) of the pressure coefficient and its derivatives

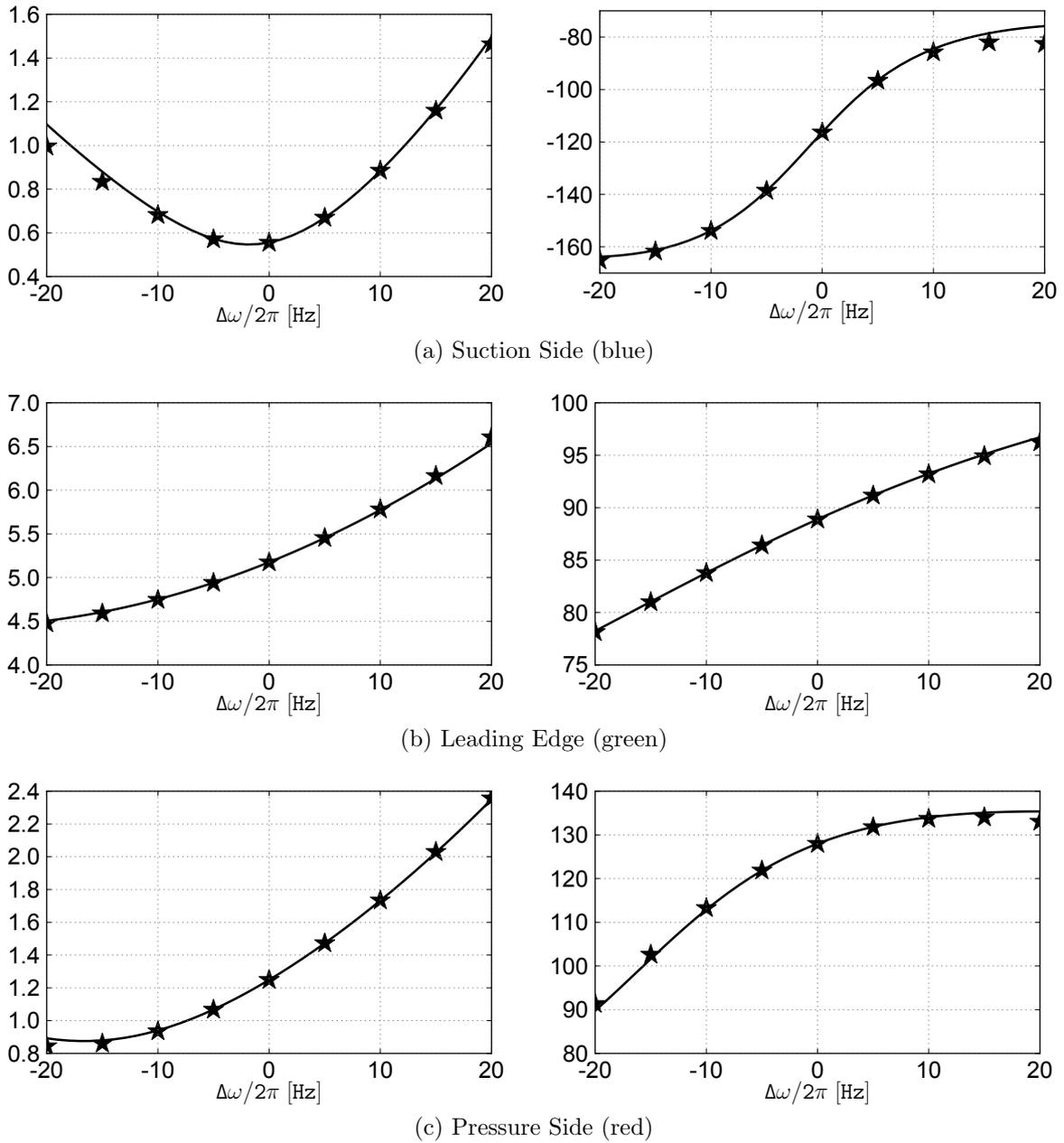


Figure 4: Comparison at numerical probes of pressure coefficient amplitude (left) and phase (right)

frequency) behavior of the flow. Finally the second order model brings a good agreement with direct computations by adding the effect of second order derivatives. In this case, one could notice that with the second order global damping coefficient approximation the system is predicted marginally stable when angular frequency is reduced. On the contrary,

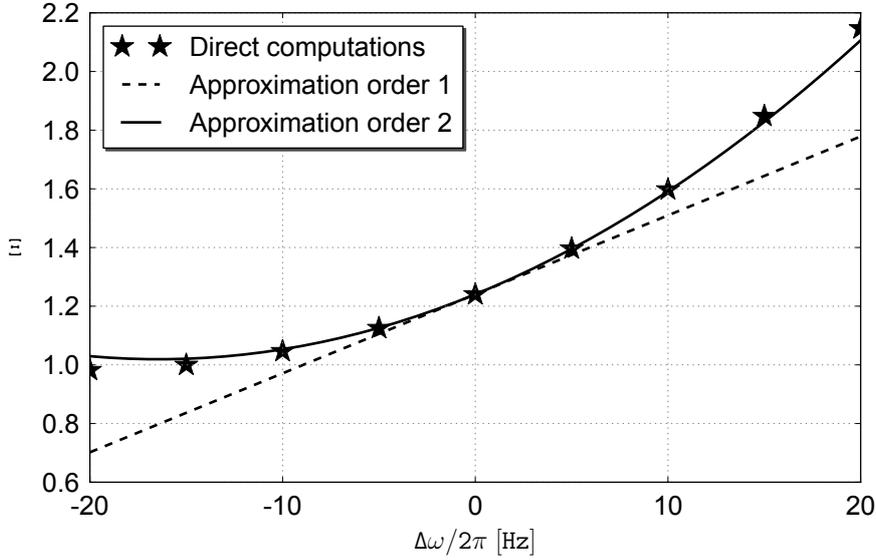


Figure 5: Aerodynamic damping sensitivity to frequency

IBPA	freq. in vacuum (Hz)	freq. in flow (Hz)	relative variation (%)	Ξ
0°	149.91	149.87	-0.03	1.24
90°	149.88	150.20	0.21	8.40
180°	149.85	150.58	0.49	6.03
270°	149.88	150.31	0.28	-0.77

Table 1: Frequency shifts and aerodynamic damping coefficient for different values of IBPA

with first order only, it turns to be unstable.

Because residue R 's minimization (see Eq. 13) gives the frequency shifts, a special attention is paid to the validation of its estimation. Thus, the evolution of $R^H R$ according to frequency is drawn on Fig. 6 for different IBPA values: a comparison is performed between functions obtained through Eq. 17 and computed points using Eq. 15. The agreement is very satisfactory: the minima are found at the same frequencies for both methods whatever the considered IBPA. Moreover the approximate residues fit the reference values in a wide frequency range, around $\pm 10\%$ of the frequency estimated in vacuum.

The frequency shifts associated with the minima are gathered in tab. 1. These shifts are relatively small, revealing quite a weak influence of the fluid on the frequency. Larger variations would be expected with a lower stiffness of the structure, but this has to be tested numerically.

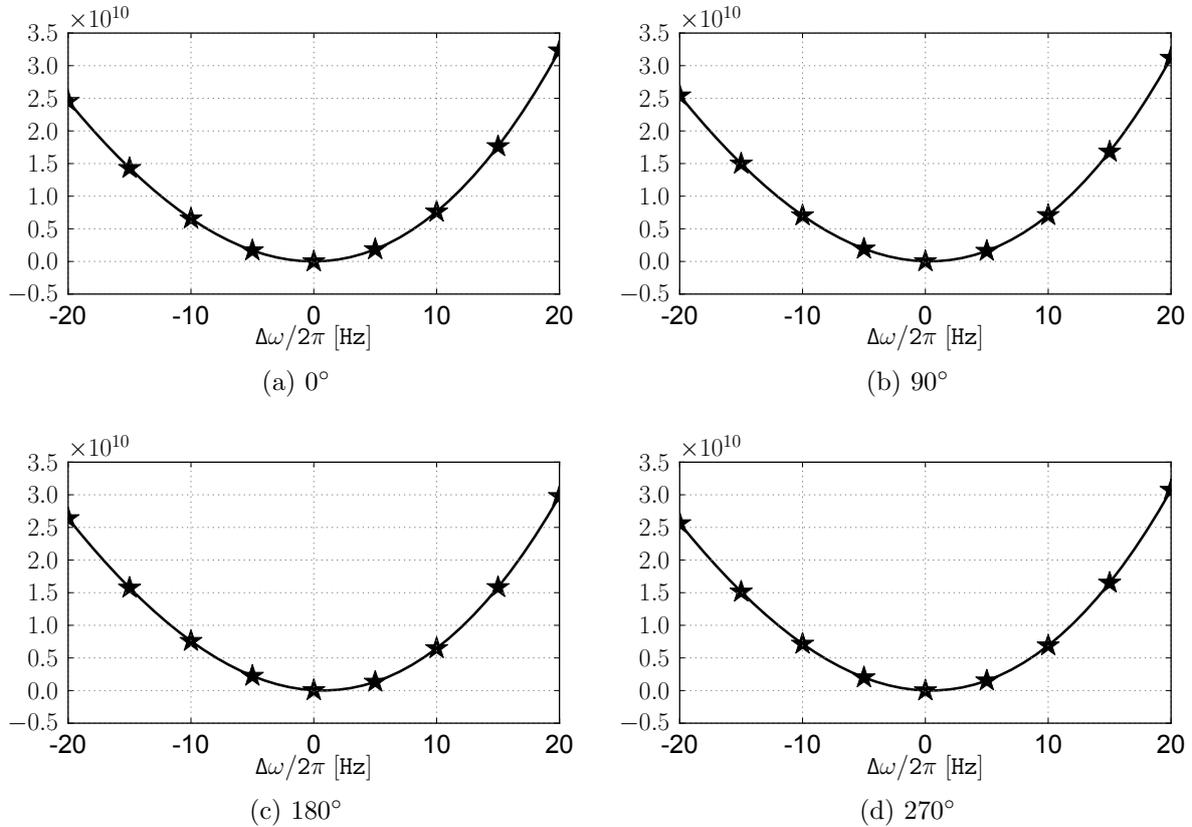


Figure 6: $R^H R$ evolution for different IBPA values

4 CONCLUSIONS

The feasibility of a non-iterative fluid-structure loose coupling approach dedicated to turbomachinery flutter problematics has been demonstrated. The new strategy is based on a parametrization that is made possible thanks to an "in-house" Automatic Derivation Tool that provides the aerodynamic force gradients with respect to angular frequency. The flow influence on mode's eigenfrequency has been investigated to provide, in a time-saving way, the characteristics of the mode. However, any design variable affecting stability margin could have been chosen, such as variations stemming from either structural response (implying mode shape or IBPA) or from operating conditions (rotating velocity, Mach number, flow incidence).

The gain of CPU time obtained by avoiding the burden of iterating inside the fluid-structure coupling procedure lets foresee applications of this approach in other domains where loose coupling is used, such as forced response computation, with or without mistuning calculation. The formalism could also be useful in presence of structural nonlinearities.

ACKNOWLEDGEMENTS

- The authors, who are members of the french ANR project ANR-08-2009 CapCAO (standing for “parametrization-aided optimized aeroelastic design”), wish to thank ANR (Agence National de la Recherche) for its generous sponsoring.
- CINES (Centre Informatique National de l’Enseignement Supérieur), which provides the computational support of the present work, is also gratefully acknowledged.
- The authors would like to especially thank Vsevolod Kharyton and Jean-Pierre Lainé for their helpful contributions.

REFERENCES

- [1] Marshall, J. G. & Imregun, M. Review of aeroelasticity methods with emphasis on turbomachinery applications. *Journal of Fluids and Structures* **10**, 237–267 (1996).
- [2] He, Z., Epureanu, B. I. & Pierre, C. Parametric study of the aeroelastic response of mistuned bladed disks. *Computers and Structures* **85**, 852–865 (2007).
- [3] Tran, D. M. Multi-parameter aerodynamic modeling for aeroleastic coupling in turbomachinery. *Journal of Fluids and Structures* **25**, 519–534 (2009).
- [4] Tran, D. M., Liauzun, C. & Labaste, C. Methods of fluid-structure coupling in frequency and time domains using linearized aerodynamics for turbomachinery. *Journal of Fluids and Structures* **17**, 519–534 (2003).
- [5] Holmes, D. G., Mitchell, B. E. & Lorence, C. B. Three dimensional linearized Navier-Stokes calculations for flutter and forced response. In *Proceedings of the 8th International Symposium of Unstead Aerodynamics and Aeroelasticity of Turbomachines*, 211–224 (1997).
- [6] Karpel, M. Design for active and passive flutter suppression and gust alleviation. *NASA CR3482* (1981).
- [7] Smati, L., Aubert, S., Ferrand, P. & Massao, F. Comparison of numerical schemes to investigate blade flutter. In *Proceedings of the 8th International Symposium of Unstead Aerodynamics and Aeroelasticity of Turbomachines* (1997).
- [8] Smati, L., Aubert, S. & Ferrand, P. Numerical study of unsteady shock motion to understand transonic flutter. In *Proceedings of Euromech 349 on Simulation of fluid-structure interaction in aeronautics* (1996).
- [9] Jameson, A., Schmidt, W. & Turkel, E. Numerical solutions of the Euler equations by finite volume methods with Runge-Kutta time stepping schemes. *AIAA paper 81-1259* (1981).

- [10] Kok, J. Resolving the dependence on free-stream value for $k\text{-}\omega$ turbulence model. *AIAA Journal* **38**, 1292–1295 (2000).
- [11] Chassaing, J.-C., Gerolymos, G. A. & Jérémiasz, J.-G. GMRES solution of compressible linearized Navier-Stokes equations without pseudo-time-marching. In *44th AIAA Aerospace Sciences Meeting* (2006).
- [12] Aubert, S. *et al.* Optimisation of a gas mixer using a new parametric flow solver. In *ECCOMAS Computational Fluid Dynamics Conference*, 17 (2001).
- [13] Moreau, S. *et al.* Parametric study of a fan blade cascade using a new parametric flow solver Turb’Opty. In *ASME Joint US-European Fluids Engineering Conference*, 10 (Montreal, Canada, 2002).
- [14] Bölcs, A. & Fransson, T. Aeroelasticity in turbomachines. Comparison of theoretical and experimental cascade results. *Communication du laboratoire de thermique appliquée et de turbomachines de l’EPFL* **13** (1986).

NEURAL NETWORKS AS SURROGATE MODELS FOR NONLINEAR, TRANSIENT AERODYNAMICS WITHIN AN AEROELASTIC COUPLING-SCHEME IN THE TIME DOMAIN

K. Lindhorst*, M. C. Haupt* and P. Horst*

*Institute of Aircraft Design and Lightweight Structures (IFL)
Technische Universität Braunschweig
Hermann-Blenk-Straße 35, D-38108 Braunschweig, Germany
e-mail: k.lindhorst@tu-bs.de, www.ifl.tu-braunschweig.de

Key words: Aeroelasticity, Reduced Order Modelling, Neural Networks, NLR7301

Abstract. In this paper the creation of a nonlinear, transient surrogate model is described that can be used within an aeroelastic coupling-scheme in the transonic range. The method is based on the theory of artificial neural networks as well as the autoregressive moving average method (ARMA). It can be shown that the method is able to approximate the nonlinear aeroelastic behaviour of the NLR7301 airfoil. Also limit cycle oscillations can be approximated with acceptable accuracy.

1 INTRODUCTION

Aeroelasticity is an important aspect in modern aircraft design and cannot be neglected in the optimization process of flight performance. The challenge herein is the complexity of an aeroelastic simulation, which is normally separated in two coupled subsystems, the aerodynamic and the structural subsystem. This coupled problem can be solved with high accuracy by combining a computational fluid dynamics solver (CFD) with a computational structure mechanics solver (CSM) within a CFD-CSM-coupling scheme. On the one hand with such a coupling scheme influences due to separation, transition or shocks can be considered, but on the other hand the computational effort is very high.

In general the solution of the aerodynamic subsystem needs much more computational effort than the structural subsystem, especially in simplified systems like a 2D airfoil or 3D wing sections. Furthermore in aeroelasticity only a fraction of the calculated aerodynamic values are needed, for example the global lift and global pitching moment of the airfoil or the pressure distribution on the coupling surface. However the whole aerodynamic system has to be solved completely each timestep.

Therefore the aim within this paper is the creation of a surrogate model for the nonlinear,

transient aerodynamic subsystem that provides only the necessary data and allows much faster aeroelastic analysis with high accuracy. Such a surrogate model would also allow more complex aeroelastic calculations like manouvers as well as atmospheric disturbances like gusts within an acceptable time.

Reduced order modelling in aeroelasticity became an important field of research in the last years. For flutter boundary prediction methods based on Hopf bifurcation were successfully applied. According to Henshaw et al. [1] harmonic balance (HB), high order harmonic balance (HOHB), center manifold, normal form and numerical continuation methods can be used for bifurcation analysis. Another approach is the eigenvalue realization algorithm (ERA) [1, 6], which identifies a linear state-space-formulation of a given system. Furthermore proper orthogonal decomposition (POD) is widely used for model reduction [12]. Lucia et al. [6] also uses POD in combination with the second order Volterra series to approximate nonlinear behaviour.

Finally different types of neural networks have been used to predict the behaviour of aeroelastic systems. Voitcu for example uses classic multilayer perceptron networks (MLP-ANN) to predict a closed aeroelastic system with structural nonlinearities [9, 10]. In the contrary Won proposes a radial basis function network (RBF-ANN) to approximate the aerodynamic behaviour of the AGARD 445.6 wing [13].

The approach chosen in this paper is based on Won's RBF networks but is modified to allow an efficient multiple input multiple output mapping.

2 ARTIFICIAL NEURAL NETWORKS

Artificial neural networks (ANN) are mathematical models of biological neural networks, which are naturally also known as brains. An ANN is a quite variable mapping method that can be used in many different fields like physics, biologics, economics and engineering, especially in control engineering. There exist many modifications of artificial neural networks but most of them are based on the multilayer perceptron network (MLP-ANN), which is explained shortly in the following. For more details the reader is referred to Hagan et al. [4].

2.1 Multilayer Perceptron Neural Network

A MLP-ANN consists of one or more layers, which also consist of one or more neurons. In general the architecture¹ is optional but it has strong influence on the networks prediction precision. So the number of layers and neurons that are necessary for a proper system identification depends on the complexity and nonlinearity of the observed system. In figure 1 a feed forward MLP-ANN with two layers is shown.

Like in biological neural networks the neurons are the backbones of the ANN. A neuron

¹The architecture is the outer topology meaning the number of layers as well as the inner topology meaning the number of neurons per layer

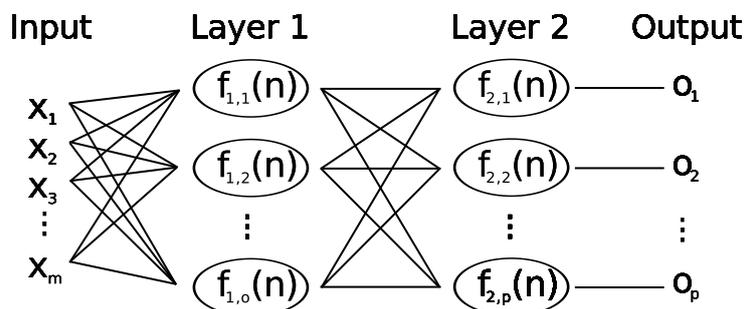


Figure 1: Two layer feed forward MLP-ANN

in a MLP-ANN consists of a bias b , its weighted connections to other neurons w_i and a transfer function $f(n)$ also called activation function. In general the neurons transfer function is user-defined, but often set to linear $f(n) = n$ or hyperbolic functions $f(n) = \tanh(n)$. The answer a_n of a single neuron to j given inputs x is calculated with equation 1.

$$a_n = f_n(b_n + \sum_{k=0}^j w_{n,k}x_k) \quad (1)$$

An important property of neural networks is that the ability to learn. The learning process of a MLP-ANN is the optimization of the bias values and weight factors in order to match the target output. The network learning process is the difficult part in ANN usage, because the training set as well as the architecture must be well chosen.

An important method for the network training is the backpropagation algorithm, which allows the systematic training of multilayer networks. The backpropagation can be used with several optimization methods like steepest descent, Levenberg-Marquardt or Newtons method. The backpropagation as well as the optimization algorithms are described in Hagan et al. [4].

Well designed networks should be able to predict the behaviour of nonlinear multiple input multiple output (MIMO) systems, but finding the best architecture can be quite difficult.

2.2 Radial Basis Function Artificial Neural Network

Another type of neural network is mentioned by Won [13] called radial basis function networks (RBF-ANN). This network uses radial basis functions, for example gaussian functions as shown in equation 4, in order to approximate nonlinear systems. The network Won proposes is quite a simple summation of the gaussian functions (see eq. 3), which have the ability as universal function approximators.

$$f_n(x) = e^{-\frac{(x-c_n)^2}{\sigma_n^2}} \quad (2)$$

$$o(x) = \sum_{j=1}^k e^{-\frac{(x-c_j)^2}{\sigma_j^2}} \quad (3)$$

This kind of networks has a more simple architecture than the MLP-ANN but the mapping of different types of inputs to target outputs is inefficient because the RBF-neurons must cover the whole input range. Furthermore Won proposes a simple training algorithm using the pseudo inverse of the neuron outputs. This training method led to unsatisfying results during the investigations for this paper, which will be explained in section 3.1.

Therefore in this paper a hybrid of both network types is used. Neurons with gaussian RBF are used in a special layer called prelayer to capture nonlinear effects but the network itself is a small MLP-ANN. Furthermore the prelayer is divided into neuron clusters in order to increase the efficiency. Due to this clustering the network is called clustered artificial neural network (CANN).

3 CLUSTERED ARTIFICIAL NEURAL NETWORKS

The clustered artificial neural network is in fact a MLP-ANN with an extra layer placed between the input values and the multilayer perceptron network (cf. fig. 2). The prelayer consists of neuron clusters that are designed for one type of input and which are not involved in the training process. The neurons of each cluster have the same gaussian transfer functions as shown in equation 4. The neuron centers c_n are distributed equidistantly over the range of input values given by the training set. The width σ of the gaussian functions of each cluster is set to $\sigma_n = 2(c_n - c_{n-1})$, so it is defined by the distance between the neuron centers within the current cluster. The number of neurons N per cluster is optional but should be chosen depending on the range of input values as well as the training set.

A given input value is therefore transformed into N values which can be described as an input vector $\underline{C}(x)$. The entries of $\underline{C}(x)$ are values between 0 and 1 whereupon the value depends on the distance between neuron center to the input value: If the input x equals the center c the neurons output is 1, but the larger the distance between x and c is the smaller is the neurons output. This means that the position of the maximum value of the entries depends on the scalar value of x . This vector is then fed into the ANN and mapped to an output.

$$\underline{C}(x) = [f_1(x), f_2(x), \dots, f_N(x)] = [e^{-\frac{(x-c_1)^2}{\sigma^2}}, e^{-\frac{(x-c_2)^2}{\sigma^2}}, \dots, e^{-\frac{(x-c_N)^2}{\sigma^2}}] \quad (4)$$

3.1 Properties and limitations of CANN

The main advantage of the CANN is its simple architecture. The prelayer replaces the hidden layers of the classic MLP-ANN so that only the output layer remains. Thus it is

possible to use only one neuron with a linear transfer function for each output to cover nonlinearities within the cluster range. In fact for more complex problems the prelayer may also be combined with a real multilayer network but for the current problem one neuron per output was used. According to that the network training can be performed quite easily, because the network only consists of one layer with linear neurons which have to be optimized. In fact this leads to a linear optimization problem which can be solved with the pseudo-inverse as Won proposed [13]. Unfortunately the input matrix, which has to be inverted is conditioned badly which leads to poor results of the pseudo-inverse training process. For this reason the steepest descent algorithm with conjugate gradient method was used in this paper which is described by Hagan [4]. The investigations showed that the simply structured and easily implemented CANN-method is able to predict nonlinear behaviour with acceptable precision. It is important to say, that a MLP-ANN with a well-designed architecture might be more precise than a CANN but the topology of this well-fitted architecture must be defined first.

A further limitation of the method is the ability of extrapolation. The cluster boundaries, which are set during the training process, define the covered input range. Input values laying out off these cluster boundaries lead to zero valued prelayer output vectors $\underline{v}(i) = \underline{0}$. Thus the input values outside the cluster boundaries are neglected by the prelayer. This aspect is not a real disadvantage, because a reliable extrapolation of nonlinear systems is in fact not possible.

3.2 Usage as Transient Predictor

Until now a nonlinear mapping method for MIMO systems was described. In order to apply the method to transient, timediscretized systems an approach similar to the autoregressive moving average (ARMA) is used, like it is proposed by Won [13]. The ARMA method predicts the response of a linear, timediscretized single input single output system at the time t with the sum of m past inputs x and n past outputs o like it is shown in equation 5. The coefficients a_j and b_k have to be determined so that the model error is minimized.

$$o(t) = \sum_{j=0}^m a_j x(t-j) + \sum_{k=1}^n b_k o(t-k) \quad (5)$$

This approach is a simple neural network with a single linear neuron without a bias. Hence this method is also applicable to more complex networks like MLP-ANNs as well as CANNs. In contrast to ARMA in this paper the gradients of the past n inputs are observed instead of the outputs, because the usage of past outputs has a destabilizing effect.

Furthermore m is set to 1 so only the input of the current timestep is used and the time history of motion is only respected by the gradients. The advantage herein is that the

CANN is able to learn easily the approximation of the static as well as the transient behaviour of the system, because in the static case the gradients are set to zero and only the remaining input is mapped to the target output. Another aspect is, that the gradients contain the information of the time step size implicitly.

With these modifications and the assumption that there exists one layer behind the prelayer with one neuron per target output the approximation of $\underline{o}(t)$ can be described with equation 6.

$$\underline{o}(t) = \underline{a}_0 \underline{C}_x(x(t)) + \sum_{k=0}^n \underline{b}_k \underline{C}_x(\dot{x}(t-k)) \quad (6)$$

The degrees of freedom of the observed two dimensional, rigid aeroelastic model are the pitch angle α and the plunge excitation h . Regarding that only the velocity of plunge motion \dot{h} and not the plunge excitation h itself influences the aerodynamics, h is neglected as input parameter. Furthermore the gradients are scaled with the freestream flow velocity u_∞ to cover the aerodynamic dependency of the freestream flow conditions. For the same reason the dimensionless lift and pitching moment coefficient are used as target outputs. So the input parameters are α , $\frac{\dot{\alpha}}{u_\infty}$, $\frac{\dot{h}}{u_\infty}$ and the target output parameters are the lift coefficient C_L and the pitching momentum coefficient C_M . A scheme of the used CANN is shown in figure 2.

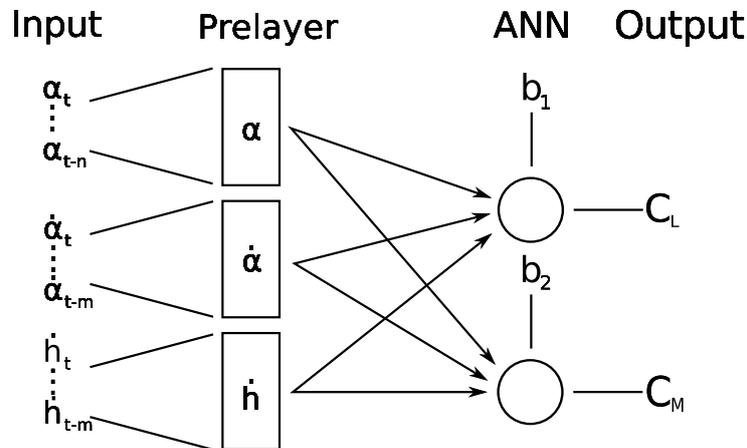


Figure 2: CANN-scheme used for transient aeroelastic systems

4 RESULTS

The method is demonstrated on the NLR7301 airfoil which is shown in figure 3. The CFD data are calculated with the *TAU*-code of the *German Aerospace Center (DLR)* [2, 3]. Weber [11] as well as Tang [7] showed that viscous effects have to be respected

for proper limit cycle calculation. Both of them suggest the Spalart-Allmaras turbulence model which is therefore also used in this paper. To neglect influences due to the transition location a fully turbulent flow is assumed.

Furthermore the number of neurons per cluster in the prelayer is set to $N = 30$ and the number of observed former gradients is set to $m = 75$, which is chosen due to parameter studies.

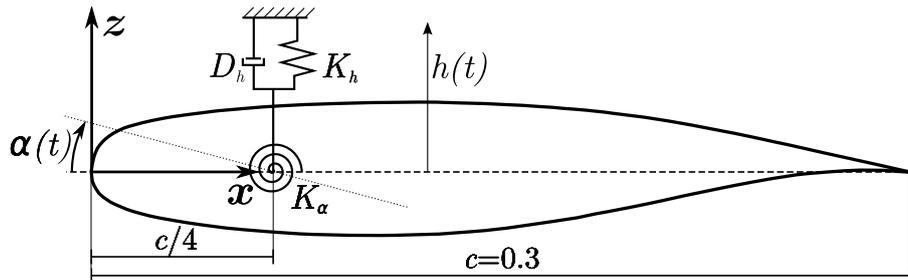


Figure 3: Geometry of the investigated NLR7301 airfoil

4.1 Network Training

In the following pitch and plunge motion of the CANN-CSD-coupling are compared with the results of the CFD-CSD-coupling at various Mach numbers (see Tab. 2). It is important to say that a network is identified at a fixed Mach number, so for each of the investigated Mach numbers at $Ma = 0.753$, $Ma = 0.7$ and $Ma = 0.65$ an according network was trained. On the other hand the freestream flow parameters pressure p_∞ , velocity u_∞ and density ρ_∞ change with the temperature T_∞ at a fixed Mach as well as a fixed Reynolds number. The free stream parameters are connected with the Sutherland model, the ideal gas law, the Mach number definition and the Reynolds number definition. According to that the flow conditions can be manipulated even if the Mach and Reynolds number is fixed, which is why different free stream temperatures T_∞ are investigated with the same network at each Mach number.

In application the network must correctly react on different frequencies and amplitudes. This flexibility must be learned during the training which is why it is necessary to provide as different training sets as possible to the network. The three networks in this paper were trained with a standardized bundle of forced motion training sets to show the practicability of the method. It is important that the training sets include different amplitudes and frequencies to ensure the required flexibility.

An example for a bundle of forced motion training sets is shown in figure 4. In addition to the shown transient sets a small set of 33 static calculations with $\alpha_{Static} = -8^\circ \dots 8^\circ$ is also involved in the network training to ensure correct approximation of static cases.

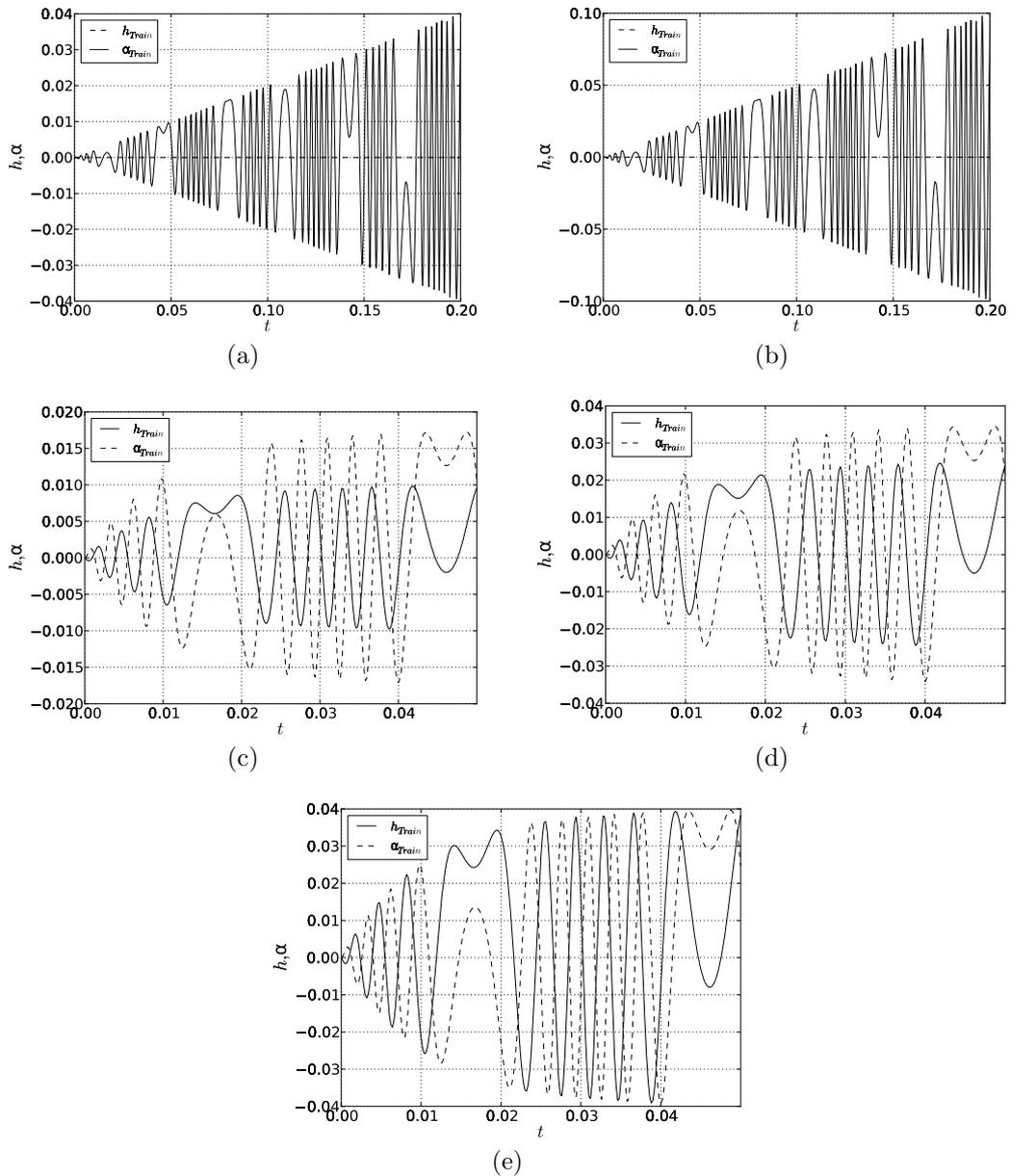


Figure 4: Example of a bundle of training set for one Mach number

4.2 Usage as Surrogate Model

After the network was trained it can be used in a two degree of freedom coupling scheme with a structure model. In this paper the structure model of Tang [7] is used for the coupled investigations of which the parameters are given in table 1. The governing equation of motion is given in 7.

Table 1: Parameters of the structure model by Tang [7]

m	26.64kg	$I_{c/4}$	0.086m ² kg
s_α	0.378mkg	K_h	1.21 · 10 ⁶ N/m
K_α	6.68 · 10 ³ Nm/rad	D_h	82.9kg/s
D_α	0.197m ² kg/rad/s		

The deformation and load transfer between the structure and the surrogate model is quite trivial: The output values of the structure model h and α are used as network input and the network output values C_L and C_M are given back to the structure model. The lift $L(t)$ and the pitching moment $M(t)$ can then be recalculated from C_L and C_M with the current flow conditions.

Furthermore an iterative staggered coupling with Aitken relaxation and first order predictor is used as it is described by Unger [8]. For the time integration of the structural system the Newmark method is used (see Hughes [5]).

$$\begin{bmatrix} m & s_\alpha \\ s_\alpha & I_{c/4} \end{bmatrix} \begin{pmatrix} \ddot{h} \\ \ddot{\alpha} \end{pmatrix} + \begin{bmatrix} D_h & 0 \\ 0 & D_\alpha \end{bmatrix} \begin{pmatrix} \dot{h} \\ \dot{\alpha} \end{pmatrix} + \begin{bmatrix} K_h & 0 \\ 0 & K_\alpha \end{bmatrix} \begin{pmatrix} h \\ \alpha - \alpha_0 \end{pmatrix} = \begin{pmatrix} L(t) \\ M(t) \end{pmatrix} \quad (7)$$

LCO case

First the LCO case at $Ma = 0.753$ is observed. Weber as well as Tang performed their LCO investigations without a structural damper and with an initial angle of attack of $\alpha_0 = 0.6^\circ$. In order to match their results no damper is used in this case, so $D_h = D_\alpha = 0$ and the same initial angle of attack is applied. Here the first time steps are critical, because the surrogate model needs the gradients of former time steps (see equation 6), which are not available during this phase. Therefore the unknown gradients are set to zero.

In figure 5 the predicted motion of the CANN-CSM-model is compared with the calculated motion of the CFD-CSM-scheme. The amplitudes of the predicted motion increases faster than the reference solution between $t = 0.2$ and $t = 0.8$ (see figure 5 on the left hand side) but the final LCO amplitude is predicted quite well as it can be seen in figure 5 on the right hand side. The temporal offset between the predicted and the calculated motion can be explained with small differences in the frequencies. In figure 6 (a) and (b) the corresponding C_L and C_M coefficients are shown. In both diagrams the nonlinear effects can be recognized due to the nonelliptic shape of the hysteresis loop.

The computational effort of the CFD-CSM-scheme for the presented LCO-case is about 92 hours on two Intel i7 cores with 2.8 GHz. In contrast to that the surrogate model needs

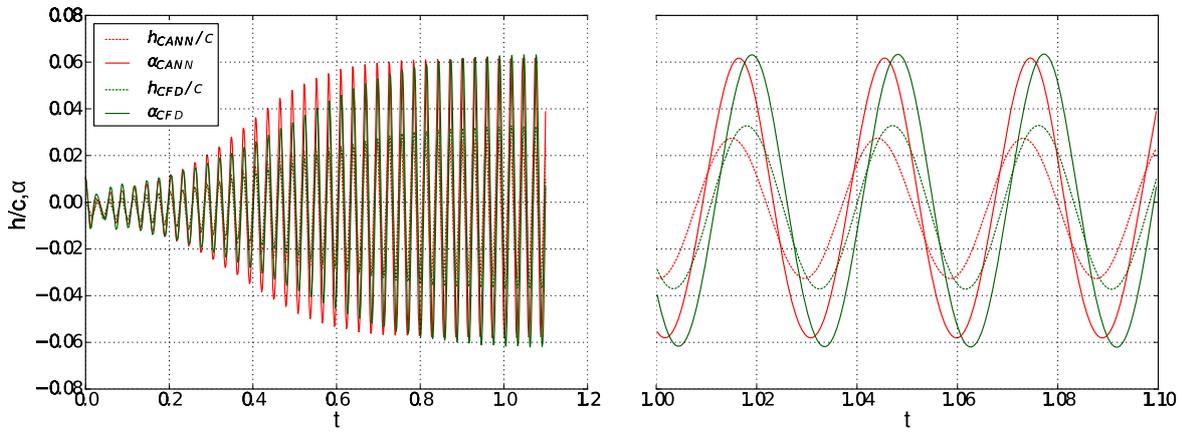


Figure 5: Comparison of CFD- and CANN-results of the LCO-case

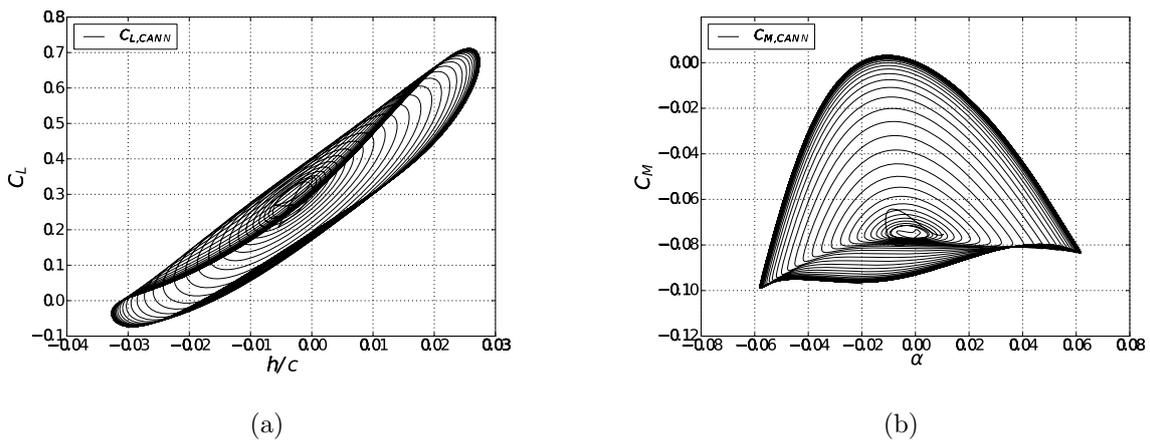


Figure 6: a) Approximated lift coefficient vs. h ; b) Approximated pitching moment coefficient

for the same calculation 6 minutes and 13 seconds on a single Intel i7 core with 2.8 GHz. So the surrogate model needs 0.1% of the time with half of the processor power to predict the LCO behaviour.

It is important to say, that the creation of the surrogate model takes also about 23 hours per Mach number, because training sets have to be calculated (ca. 21h) and the network has to be trained as well (ca. 2h).

Investigation of other Mach numbers

Furthermore two other Mach numbers with different free stream conditions are investigated, which can be found in table 2. In all observed cases the Reynolds number and

Reynolds length are constant at $Re = 1.723 \cdot 10^6$ and $L_{Re} = 0.3m$. In these investigations the dampers D_h and D_α are set to the values of table 1 and the initial angle to $\alpha_0 = 0^\circ$. Furthermore during the first 100 steps ($t_{100} = 0.02$) a forced motion is applied to the coupled model with $\alpha(t) = \frac{\pi}{180} \sin(200t)$ and $h(t) = 0$. The training sets for both networks were calculated at a free stream temperature of $T = 273.15K$.

Table 2: Aerodynamic parameters of the observed cases

Parameter	Case 1	Case 2	Case 3	Case 4	Case 5	Case 6
Ma	0.7	0.7	0.7	0.65	0.65	0.65
$T_\infty[K]$	173.15	273.15	373.15	173.15	273.15	373.15
$u_\infty[m/s]$	184.635	231.902	271.047	171.447	215.337	251.687

In figures 7, 8 and 9 the predicted and the calculated airfoil motion is compared. In the beginning of each time series during the transient non-periodic phase the prediction matches the calculation with quite good accuracy. Even the dependency of the free stream conditions are covered in an acceptable manner. But it is also noticeable that after the transient starting phase the difference between the predicted and calculated amplitude increases with the time. This means that the stability limit of the surrogate model does not match the stability limit of the CFD-system. This lack of accuracy may be fixed with a better chosen training set.

Nevertheless it can be summarized, that the presented method is able to represent the aerodynamic system behaviour and can therefore substitute the CFD code within a CFD-CSM-coupling scheme for efficient aeroelastic analysis.

5 CONCLUSIONS

A method is presented, which is able to substitute the CFD-code within a CFD-CSM-coupling scheme with two degrees of freedom. Also nonlinear, transient effects can be covered by the method, which is demonstrated at the LCO-case as well as several examples at different transonic Mach numbers. Even the dependency of different free stream flow conditions at a fixed Mach number can be regarded, but the surrogate model also shows a lack of accuracy respective to the stability behaviour of the system.

Another important aspect is the computational effort needed for the creation of the surrogate model. It can be shown that the system behaviour can be identified with several not specially chosen trainings sets. For more complex aeroelastic investigations like manouvers or influences caused by gusts the choice of training sets should be optimized for a proper approximation of the stability behaviour.

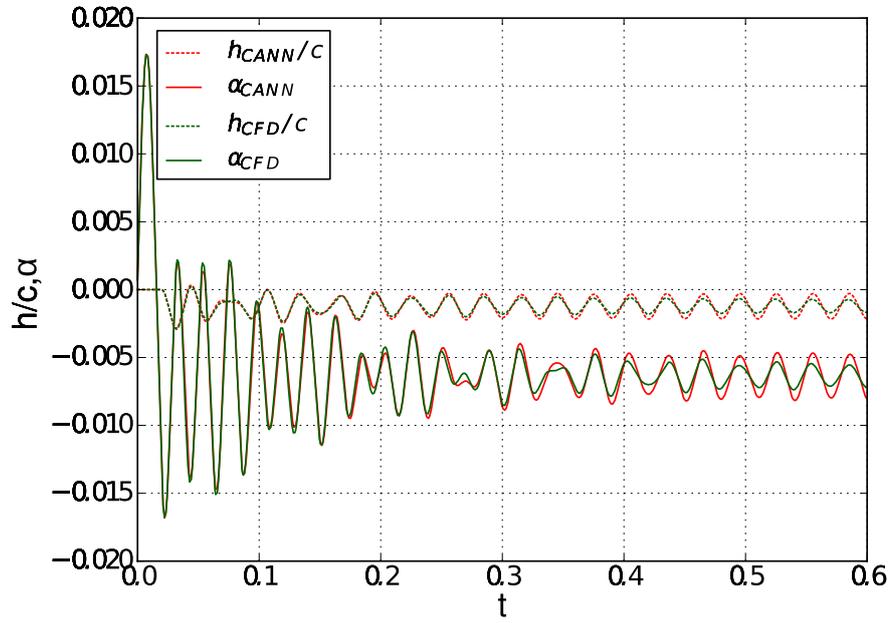
6 ACKNOWLEDGMENTS

As part of the "ComFliTe" project the partial funding of this research under contract 20A0801D of the Federal Ministry of Economics and Technology through the German Aerospace Center is gratefully acknowledged. The authors thank J. Brink-Spalink of Airbus, F. Selimefendigil and S. Goertz of DLR for valueable discussions and their support.

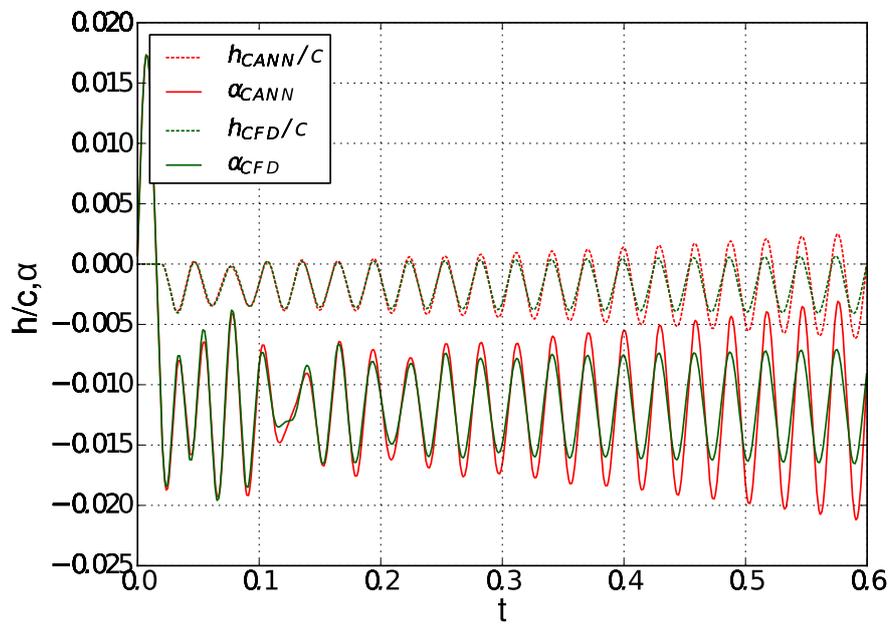
REFERENCES

- [1] M. de C. Henshaw, K. Badcock, G. Vio, C. Allen, J. Chamberlain, I. Kaynes, G. Dimitriadis, J. Cooper, M. Woodgate, A. Rampurawala, D. Jones, C. Fenwick, A. Gaitonde, N. Taylor, D. Amor, T. Eccles, and C. Denley. Non-linear aeroelastic prediction for aircraft applications. *Progress in Aerospace Sciences*, 43(4-6):65 – 137, 2007.
- [2] M. Galle, T. Gerhold, and J. Evans. Parallel computation of turbulent flows around complex geometries on hybrid grids with the DLR-TAU code. In *Ecer, A., Emerson, D. R. (eds.), Proc. 11th Parallel CFD Conf., May 23-26 1999, Williamsburg, VA, North Holland*, 1999.
- [3] H. V. Gerhold, T. and D. Schwamborn. On the validation of the DLR-TAU code. In *New Results in Numerical and Experimental Fluid Mechanics, Notes on Numerical Fluid Mechanics*. Vieweg, 1999.
- [4] M. T. Hagan, H. B. Demuth, and M. Beale. *Neural Network Design*. PWS Publishing, 1996.
- [5] T. J. Hughes and J. McCulley. *The Finite Element Method*. Prentice-Hall International Editions, 1987.
- [6] D. J. Lucia, P. S. Beran, and W. Silva. Aeroelastic system development using proper orthogonal decomposition and volterra theory. In *44th AIAA Structures, Structural Dynamics and Materials Conference*, 7-10 April 2003.
- [7] L. Tang, R. E. Bartels, P.-C. Chen, and D. D. Liu. Numerical investigation of transonic limit cycle oscillations of a two-dimensional supercritical wing. *Journal of Fluids and Structures*, 17:29–41, 2003.
- [8] R. Unger, M. Haupt, and P. Horst. Coupling techniques for computational non-linear transient aeroelasticity. *Journal of Aerospace Engineering*, 222:435–447, 2008.
- [9] O. Voitcu and Y. S. Wong. A neural network approach for nonlinear aeroelastic analysis. In *43th AIAA Structures, Structural Dynamics and Materials Conference, 22-25 April, 2002, Denver, Colorado*, 2002.

- [10] O. Voitcu and Y. S. Wong. An improved neural network model for nonlinear aeroelastic analysis. In *44th AIAA Structures, Structural Dynamics and Materials Conference, 7-10 April, 2003, Norfolk, Virginia, 2003*.
- [11] S. Weber, K. D. Jones, J. A. Ekaterinaris, and M. F. Platzler. Transonic flutter computations for the nlr 7301 supercritical airfoil. *Aerospace Science and Technology*, 5:293–304, 2001.
- [12] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA Journal*, pages 2323–2330, 2002.
- [13] K. Won, H. Tsai, M. Sadeghi, and F. Liu. Non-linear impulse methods for aeroelastic simulations. In *AIAA-2005-4845, presented at the 23rd AIAA Applied Aerodynamics Conference, Toronto, Ontario, June 6-9, 2005, 2005*.

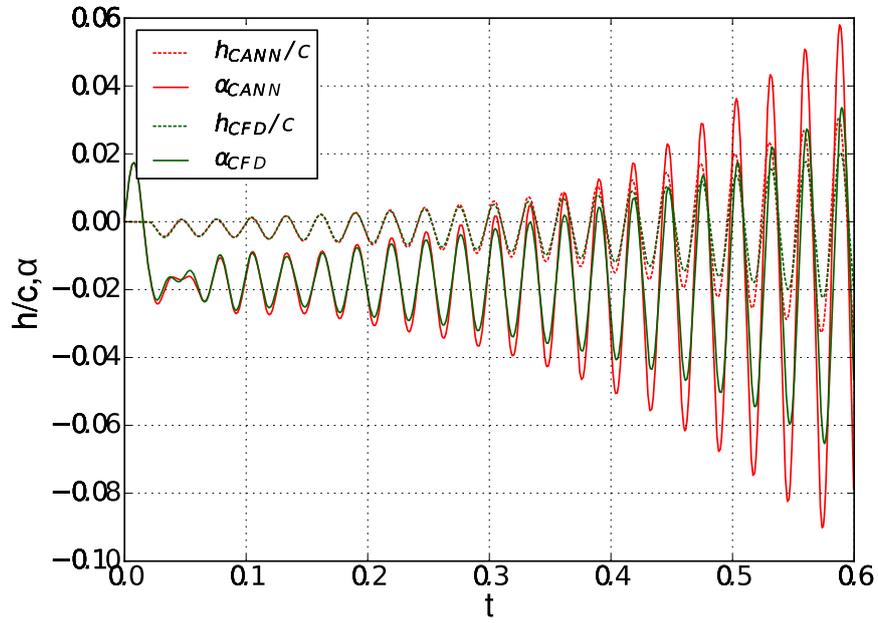


(a)

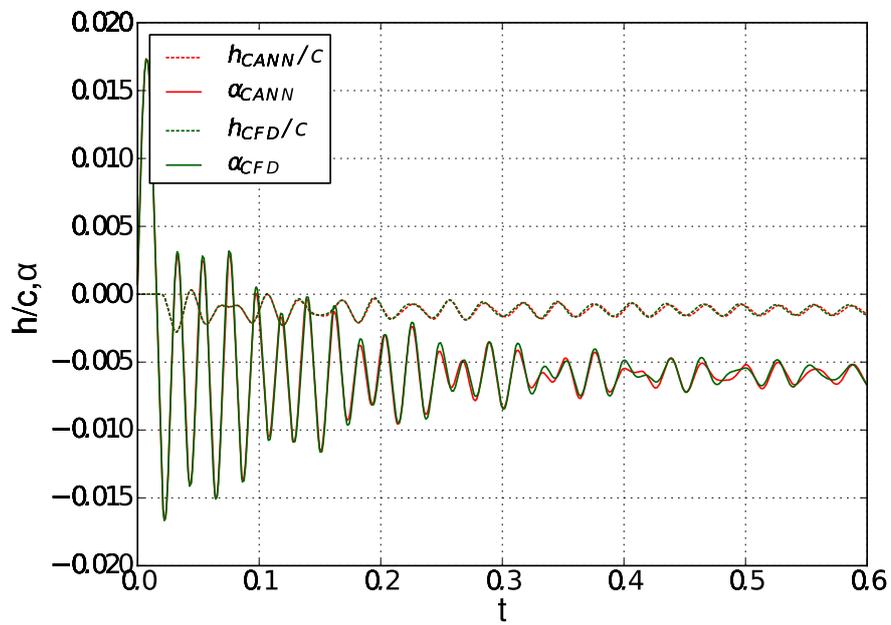


(b)

Figure 7: CANN vs. CFD Cases 1 & 2

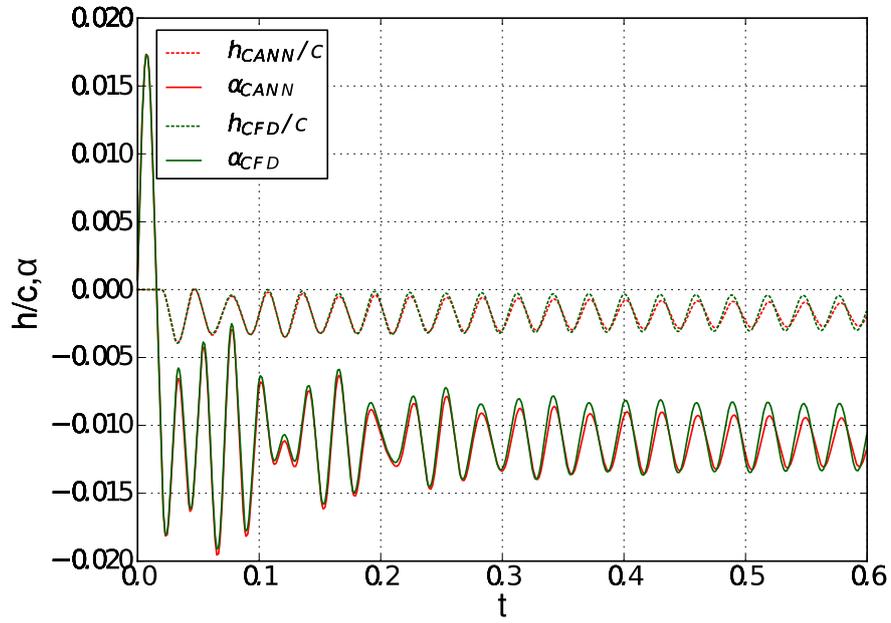


(a)

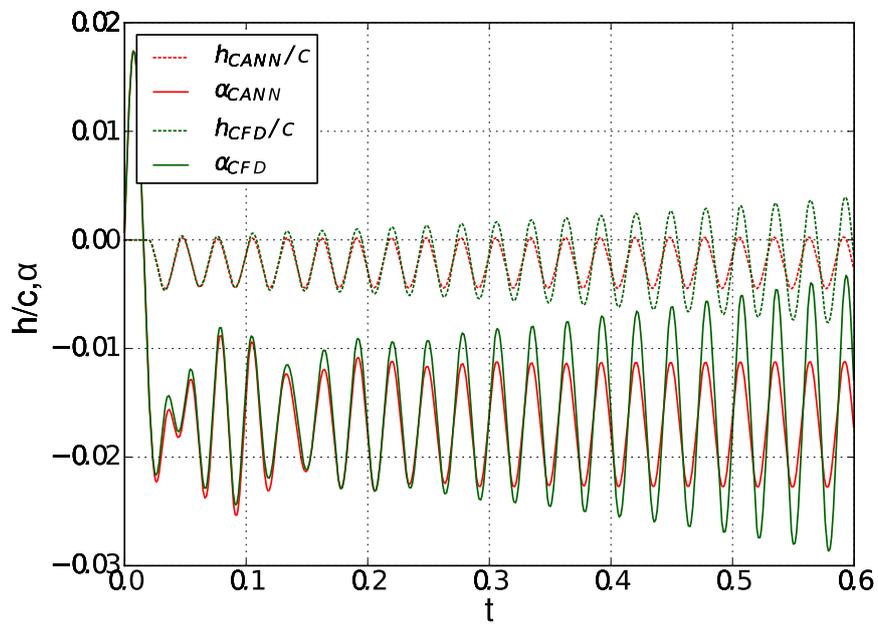


(b)

Figure 8: CANN vs. CFD Cases 3 & 4



(a)



(b)

Figure 9: CANN vs. CFD Cases 5 & 6

WIND–STRUCTURE INTERACTION ON CONSTRUCTION STAGES FOR UNBALANCED SEGMENTAL BRIDGES

ALEJANDRO HERNÁNDEZ* AND JESÚS G. VALDÉS*

* Departamento de Ingeniería Civil, División de Ingenierías, Campus Guanajuato, Sede Belén
Universidad de Guanajuato
Av. Juárez 77, Zona Centro, Guanajuato, Gto., México. CP 36000
e-mail: alexhm_mx@yahoo.com, valdes@ugto.mx

Key words: wind–structure interaction, segmental bridges, construction stages.

Abstract. Behavior due to wind–structure interaction of segmental bridges on construction stages is presented. The paper deals with the analysis of the bridge “Viaducto km 61+000” located near the city of Xalapa, in the state of Veracruz, Mexico. This 470 m long bridge with 17.9 m wide cast in place box girder will carry four lanes of traffic over a deep valley, with piers height from 26 m to 113 m. The superstructure consists of two central spans of 145 m long and two approach spans of 90 m long. The roadway geometry has vertical and horizontal curvatures along the entire length of the bridge. The substructure consists of cast–in–place rectangular hollow piers on bored pile foundations and spread footings. The superstructure is erected by the unbalanced cantilever method using form travelers. Computational fluid dynamics based on the finite element method is used to simulate wind forces acting on the structure, which are coupled to computational structure dynamics on construction stages. Thus, a stabilized fluid flow formulation is presented together with an ALE scheme while geometrically non–linear solid dynamics finite elements are used to simulate the bridge behavior. Both solutions are coupled together using a strong coupling technique to perform an aeroelastic analysis of the bridge. Differences obtained among numerical approach and code requirements are presented.

1 INTRODUCTION

On construction stages, most of civil constructions are vulnerable to nature acting forces like wind or earthquakes, in the specific case of bridges, forces are resisted in one or two places at most on construction stages, making a very vulnerable structure and needing a careful construction process. A good behavior estimate on construction stages permit to achieve a successful constructions without undesirable incidents.

“Viaducto km 61+000” is a 470 m long bridge now in construction with two central spans of 145 m and two approach spans of 90 m. Figure 1 shows a view of bridge when it will be finished. The superstructure is supported by three cast–in–place rectangular hollow piers from 26 m to 113 m height on bored pile foundations and spread footings.

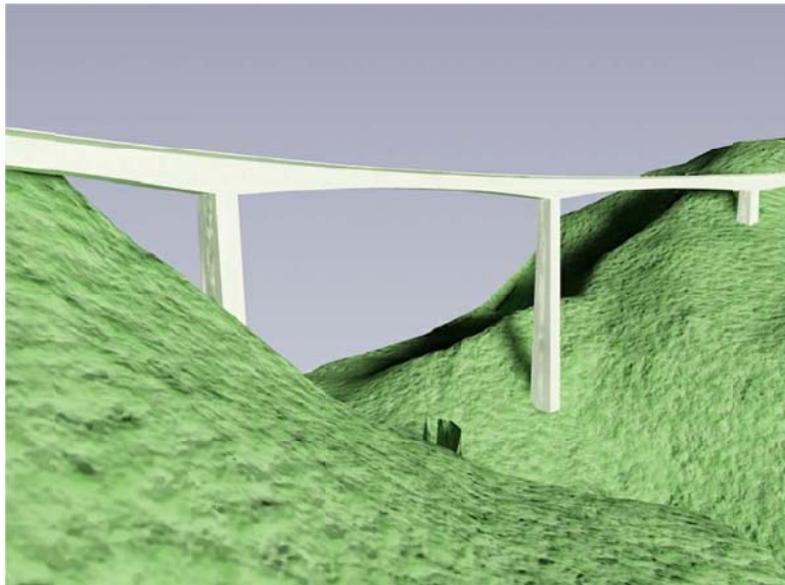


Figure 1 Simulated bridge view when it will be finished

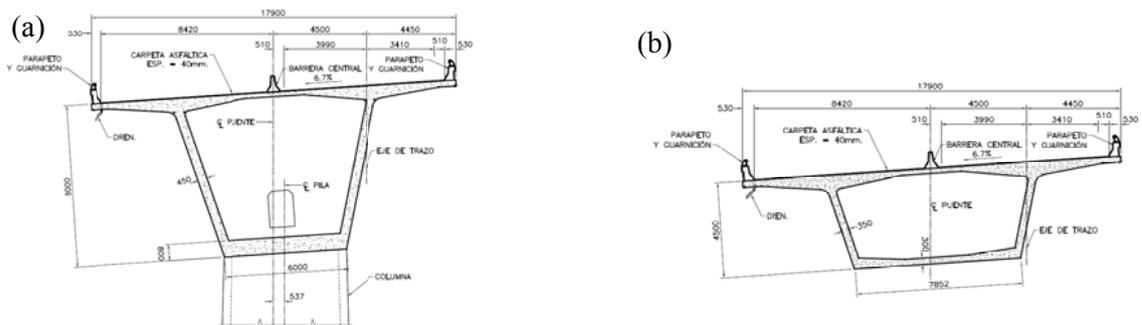


Figure 2 Superstructure section geometry (a) Pile segment (b) Mid central span

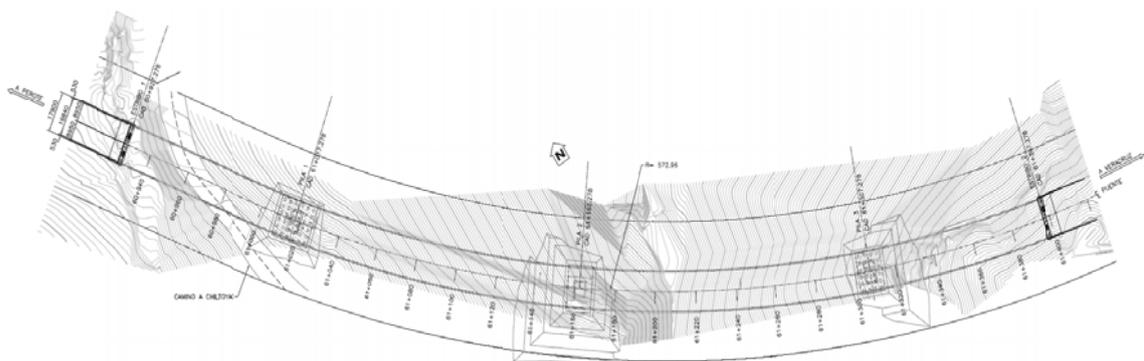


Figure 3 Bridge plan view

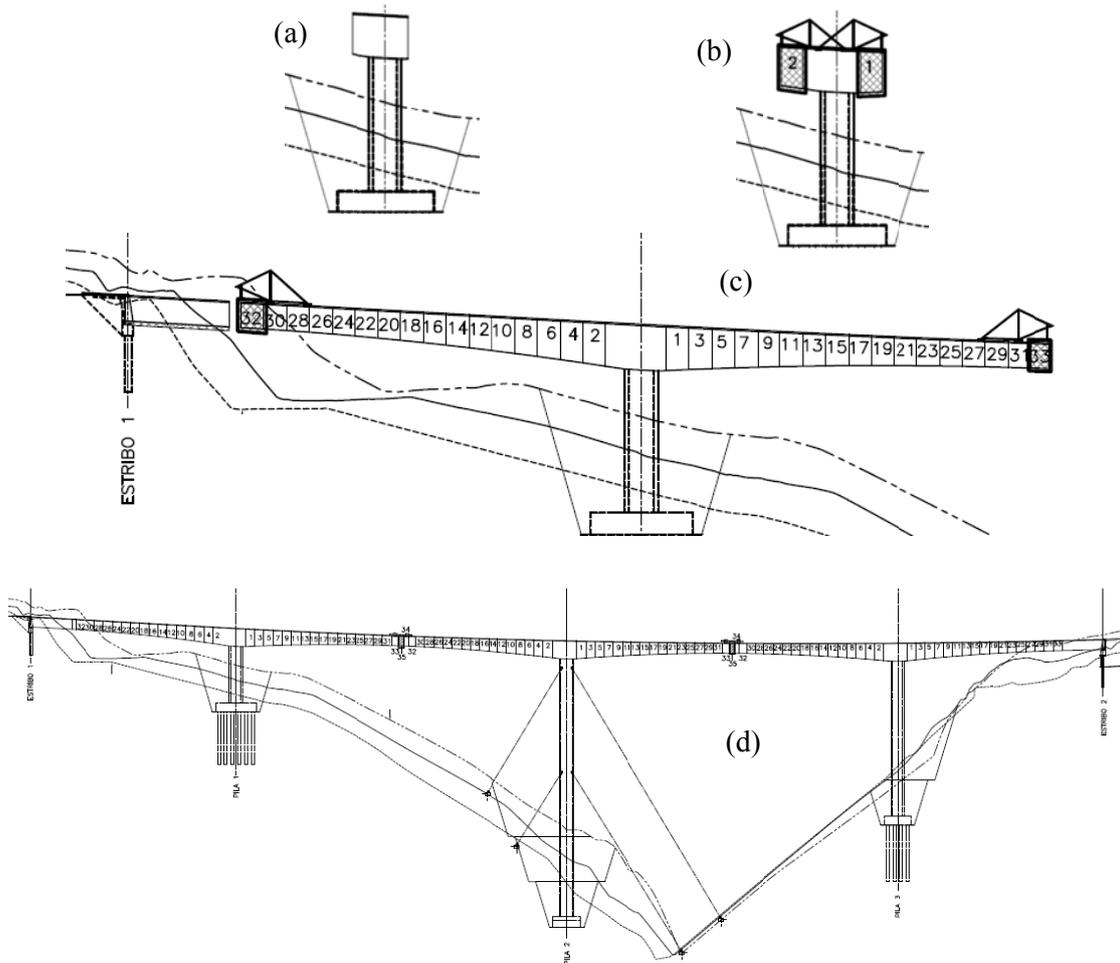


Figure 4 Description of some bridge construction stages (a) Construction of pile and pile segment (b) Form travelers colocation on pile segment (c) Segments construction sequence in cantilever (d) Construction of closure segments

The superstructure is a 17.9 m wide cast-in-place box girder with variable section (see Figure 2) and will carry four lanes of traffic. The roadway has vertical and horizontal curvature along the entire length of the bridge as can be showed in Figure 3 and Figure 4(d). Superstructure construction procedure consists of use two form travelers in unbalanced cantilever above each pile, which is described in Figure 4. Before construct closure segments are placed, bridge substructures are isostatic inverted pendulums, which are very vulnerable to nature forces like wind or earthquakes.

In this paper finite element analysis considering wind-structure interaction at some construction stages is presented. Acting forces obtained from code requirements [1] and finite element analysis considering wind-structure interaction are compared.

3 FLUID–STRUCTURE MODELATION

3.1 Structure

For structural purposes, a geometrically nonlinear solid analysis is employed. Mathematical expressions to solve the structural part are obtained from lineal momentum conservation equation, once discretized with finite elements consists of solve the next equation

$$\mathbf{f}^{\text{int}}(\mathbf{u}_{n+1}) + \mathbf{M}\ddot{\mathbf{u}}_{n+1} = \mathbf{f}^{\text{ext}}(\mathbf{u}_{n+1}) \quad (1)$$

where

\mathbf{f}^{int} = Internal forces

\mathbf{f}^{ext} = External forces

\mathbf{M} = Mass

\mathbf{u} = Displacements vector

$\ddot{\mathbf{u}}$ = Acceleration field

Generalized- α method is used to compute time integration for equation (1) because other traditional methods like β -Newmark technique produce spurious results with geometrically nonlinear finite element analysis.

Figure 5 shows a 3D-view of the structural model to be used to compute the dynamic of the structure due to wind action.

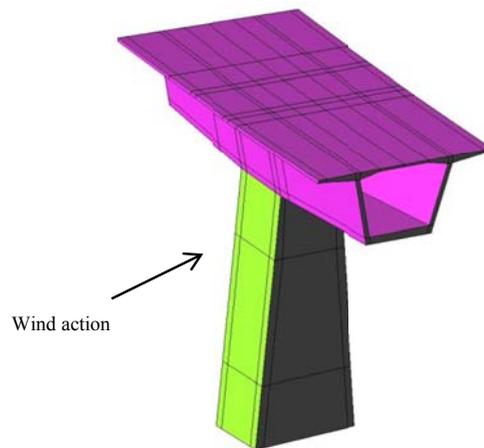


Figure 5 Structure model to compute structural dynamic system

3.2 Fluid

An incompressible fluid formulation has been used to simulate wind action due to wind velocity is lower than 0.3 Mach. Navier–Stokes equations are used to model fluid flow expressed as

$$\begin{aligned} \mathbf{M}\dot{\mathbf{v}}_{n+1} + \mathbf{K}\mathbf{v}_{n+1} - \mathbf{G}\mathbf{p}_{n+1} &= \mathbf{f}_{n+1}^{\text{ext}} \\ \mathbf{G}^T \mathbf{v}_{n+1} &= 0 \end{aligned} \quad (2)$$

where

- \mathbf{v} = Velocity
- \mathbf{p} = Pressure field
- $\dot{\mathbf{v}}$ = Acceleration
- \mathbf{M} = Mass matrix
- \mathbf{K} = Matrix with convective and viscous terms
- \mathbf{G} = Matrix to include pressure terms or to consider the incompressibility fluid

For dynamic fluid flow analysis, equations (2) can be expressed as

$$\begin{aligned} (\mathbf{v}_h^{n+1}, \mathbf{w}_h) + c(\mathbf{v}_h^{n+1}, \mathbf{v}_h^{n+1}, \mathbf{w}_h) - b(p_h^{n+1}, \mathbf{w}_h) + a(\mathbf{v}_h^{n+1}, \mathbf{w}_h) &= (\mathbf{b}_h^{n+1}, \mathbf{w}_h) \\ b(q_h, \mathbf{v}_h^{n+1}) &= 0 \end{aligned} \quad (3)$$

In the equation (3) notation used in [2] and [3] has been employed and the meaning can be founded in [4]. To analyze the structure in a faster way, equations (3) has been uncoupled using fractional step method that can be founded in [3], and considering this expressions are fully eulerian, an ALE formulation has been used to take in count the structure movement and move the domain fluid as well. The uncoupled equations are expressed as

$$\begin{aligned} (\dot{\mathbf{u}}_h^{n+\alpha_m^f}, \mathbf{w}_h) + c(\tilde{\mathbf{c}}_h^{n+\alpha_f^f}, \tilde{\mathbf{v}}_h^{n+\alpha_f^f}, \mathbf{w}_h) - b(p_h^n, \mathbf{w}_h) + a(\tilde{\mathbf{v}}_h^{n+\alpha_f^f}, \mathbf{w}_h) + \\ \tau(\tilde{\mathbf{c}}_h^{n+\alpha_f^f} \cdot \nabla \tilde{\mathbf{v}}_h^{n+\alpha_f^f} + \nabla p_h^n - \pi_h^n, \tilde{\mathbf{c}}_h^{n+\alpha_f^f} \cdot \nabla \mathbf{w}_h) &= (\mathbf{b}_h^{n+1}, \mathbf{w}_h) \\ -\frac{\Delta t \gamma^f}{\alpha_m^f} (\nabla [p_h^{n+1} - p_h^n], \nabla q_h) - \tau(\tilde{\mathbf{c}}_h^{n+\alpha_f^f} \cdot \nabla \tilde{\mathbf{v}}_h^{n+\alpha_f^f} + \nabla p_h^{n+1} - \pi_h^n, \nabla q_h) &= b(q_h, \tilde{\mathbf{v}}_h^{n+1}) \\ \frac{\alpha_m^f}{\Delta t \gamma^f} (\mathbf{v}_h^{n+1} - \tilde{\mathbf{v}}_h^{n+1}, \mathbf{w}_h) - b(p_h^{n+1} - p_h^n, \mathbf{w}_h) &= 0 \\ (\pi_h^{n+1}, \eta_h) - (\tilde{\mathbf{c}}_h^{n+\alpha_f^f} \cdot \nabla \tilde{\mathbf{v}}_h^{n+\alpha_f^f} + \nabla p_h^{n+1}, \eta_h) &= 0 \end{aligned} \quad (4)$$

Equations (4) are formulated in a four implicit steps for each time step. The first step consists of solve system at an intermediate velocity, which is a nonlinear formulation. Once found the intermediate velocity, in the second step the final pressure is computed. Final velocity is calculated in third step, and finally, the complete system is stabilized in the fourth step. The generalized- α method is used for time integration of the fractionated step. A complete analytical deduction of generalized- α method can be founded for incompressible fluids in [4].

Figure 6 shows a view of used mesh to compute fluid flow around the structure, where structure model (Figure 5) fits exactly.

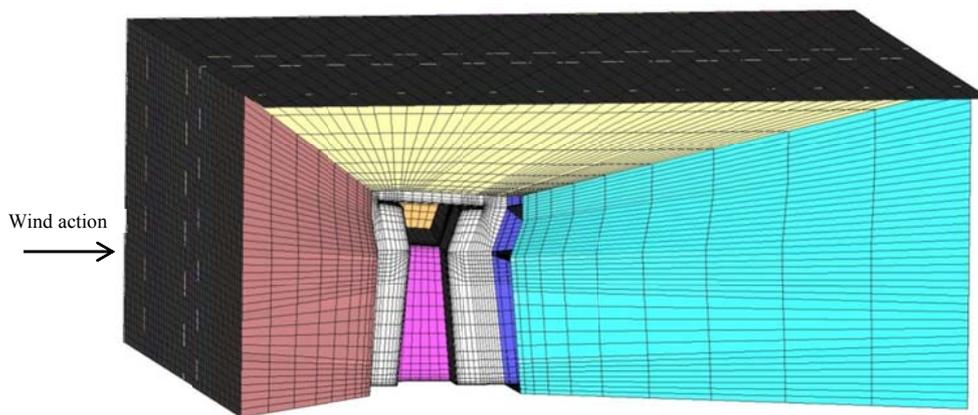


Figure 6 Internal mesh used to compute fluid flow around the structure

3.3 Fluid–structure interaction

Structure and the fluid are solved in a domain that contains the structure; both calculations are employed to predict the simultaneous effects, solving fluid and structure at the same time like in real world. The couple fluid–structure interaction problem can be performed using a monolithic solution [5] or partitioned scheme [6], [4].

Principal disadvantages of monolithic solutions are related to have all programmed in one code, adding the problem of increase the degrees of freedom, and for hence, computational time. Partitioned methods permit to have specialized codes to each part of the problem, in this case, structure and fluid codes. Having separated codes permit to reduce computational time to achieve solution. Disadvantage of partitioned approach is related to convergence of coupling solutions, which can be solved employing Aitken schemes that solve the coupling for problems where the added mass effect is not significant. This solution is used in [4] giving great result for aeroelastic problems like treated in [7] and is used for problem herein presented.

Calculus process consists of predict displacements of structure using the structural solver (CSD) considering the structural dynamic response of the previous time step. Then the displacement predictions of the structure are passed to the mesh solver (CMD) to match the fluid mesh with deformed structure. When fluid mesh is adjusted, a fluid solver (CFD) is used to compute the dynamic fluid flow. Finally acting fluid forces on structure are passed to structural solver to update structure displacements. This depicted procedure is computed until convergence criteria is achieved, ending the time step. Algorithms for coupling calculations can be founded in [4].

4 STUDIED PROBLEM

In this paper, construction above Pile 1 is presented (see Figure 4) focusing in results obtained for FSI after construct the fifth superstructure segment. Figure 5 shows the finite element model used to compute structural dynamics of the bridge. Mesh used to simulate wind action on structure is showed in Figure 6 when the place occupied by the structure can be seen inside.

Wind action is established using requirements specified by the Mexican Federal Electric Commission Code, (*Comisión Federal de Electricidad, CFE*) [8], which is the code applicable in Mexico for wind design at the construction site. Wind velocity profile for analysis is showed in Figure 7.

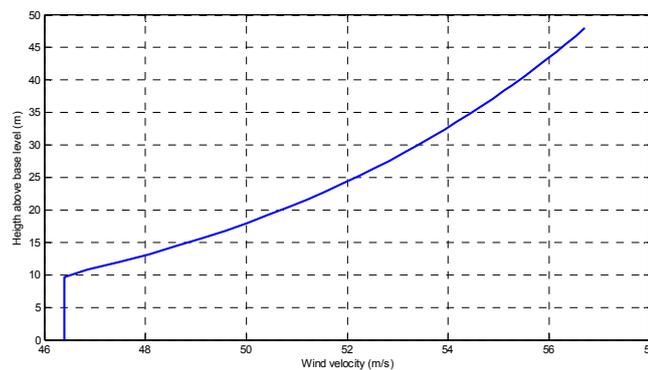


Figure 7 Wind velocity profile for analysis

Figure 8 shows wind pressure and structure deformation at several time steps for one complete vibration cycle, representing the dynamic behavior of the structure. Moreover Figure 9 shows the fluid state at one time step around the structure. At the scale showed in Figure 9 structure deformation cannot be appreciated, but it is considered.

Table 1 shows some results obtained from code regulations established in [1] and [9] corresponding to a static analysis and compared to FSI analysis described in this paper. As can be seen, obtained displacements at the top of the structure are greater with FSI than code requirements. Shear forces and overturning moments computed at pile base have a lower value with FSI compared to code analysis. The results shows that even displacement predictions with FSI are greater, forces are not, this is because the static analysis is not enough to predict inertial forces acting in the structure, but conservative forces considered in the static analysis results in greater reactions at pile base.

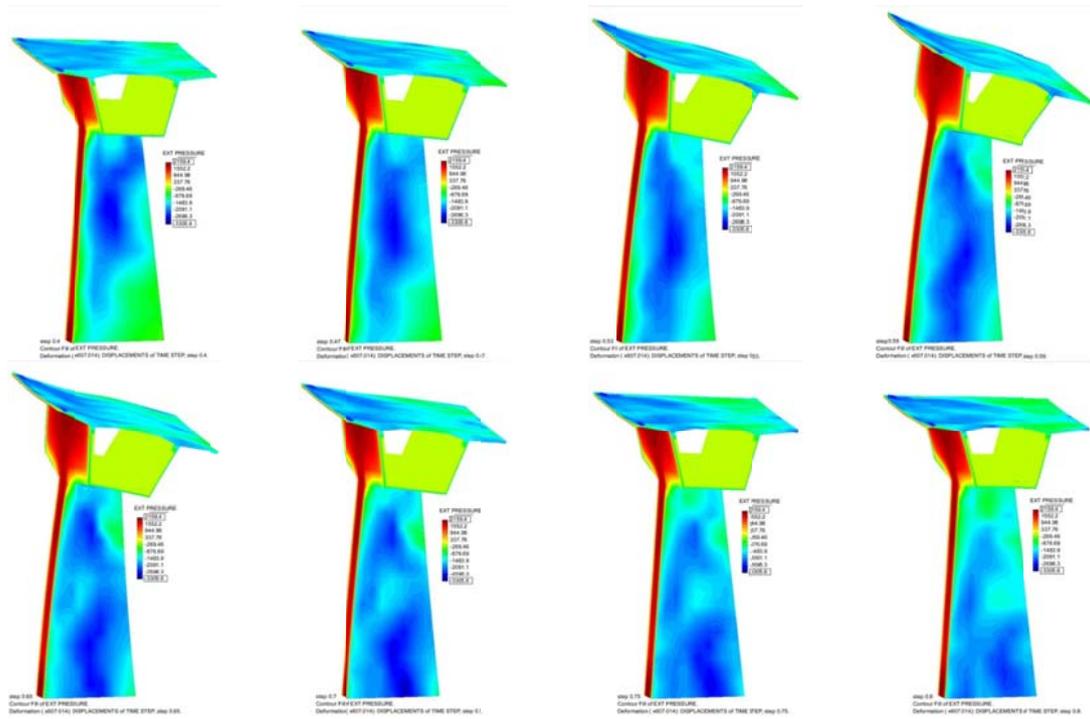


Figure 8 Wind pressures [Pa] and deformation variation in time for one cycle of wind induced vibration

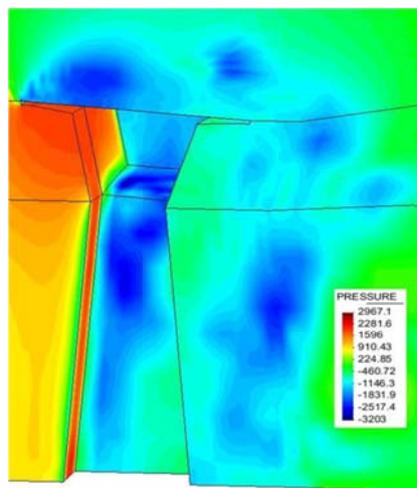


Figure 9 Fluid pressure [Pa] near structure

Table 1 Some maximum values obtained from Code and FSI analyses

Construction Stage	Top displacement (m)		Base shear (kN)		Overturning Moment (MN-m)	
	Code	FSI	Code	FSI	Code	FSI
5	0.0022	0.0083	1318.608	980.490	34.349	23.212

5 CONCLUSIONS

Result shows that static analysis derived of code requirements is a very conservative approach to estimate acting forces in initial constructions stages of analyzed bridge. This kind of approximate analysis is relatively easy to perform, but is maybe too expensive for this particular project. Analysis and comparative in final construction stages are necessary to determine security against wind action of the bridge, as well on complete constructed bridge.

6 ACKNOWLEDGEMENTS

The authors would like to thank the collaboration with CIMNE's classrooms that let us use their Software GiD in this work. Also a special mention to International Bridge Technologies de México SA de CV to provide all information about "Viaducto km 61+000", we hope to apply the obtained knowledge in future projects to achieve a more reliable structures.

REFERENCES

- [1] AASHTO. *Standard Specification for Highway Bridges, 17th Ed.* American Standard Association of State Highway and Transportation Officials, (2002).
- [2] Gunzburger M.D. *Finite Element Methods for Viscous Incompressible Flow.* San Diego Academic Press (1989).
- [3] Codina R. Pressure stability in fractional step finite element methods for incompressible flows. *Journal of Computational Fluids* (2001) **70**:112–140
- [4] Valdés J.G. *Fluid–Structure Interaction Analysis.* Verlag Ed. D.M. Germany, (2008)
- [5] Hubner B. A monolithic approach to fluid–structure interaction using space–time finite elements. *Computer Methods in Applied Mechanics and Engineering.* (2004) **193**:2087–2107.
- [6] Wuchner R. *Mechanik und numeric der formfinding und fluid–struktur–interaction von membrantragwerken.* Doctoral Thesis. Munich Technical University, (2006)
- [7] Valdés J.G., Hernández H. and Hernández A. Análisis aeroelástico de chimeneas de acero utilizando elementos no–lineales de lámina sin rotación. *Revista Internacional de Métodos Numéricos para el Cálculo y Diseño en Ingeniería.* (2010) **26**:233–240.
- [8] CFE. *Manual de Diseño de Obras Civiles – Diseño por viento.* Comisión Federal de Electricidad, México, (2008).
- [9] SCT. *Normas de Proyectos de Puentes y Estructuras SCT.* Secretaría de Comunicaciones y Transportes (2004).

AN OPTIMIZED RETURN MAPPING ALGORITHM FOR THE BARCELONA BASIC MODEL

M. PERTL^{*}, M. HOFMANN[†] AND G. HOFSTETTER^{*}

^{*} Unit for Strength of Materials and Structural Analysis
Institute for Basic Sciences in Civil Engineering
University of Innsbruck, Technikerstraße 13, 6020 Innsbruck, Austria
e-mail: michael.pertl@uibk.ac.at, web page: <http://www.uibk.ac.at/bft>

[†] ALPINE BeMo Tunneling GmbH
Bernhard-Höfel-Straße 11, 6020 Innsbruck, Austria
e-mail: matthias.hofmann@alpine-bemo.com, web page: <http://www.bemo.net/>

Key words: partially saturated soil, stress update algorithm, embankment dam

Abstract. For the most well-known constitutive model for partially saturated soils, the Barcelona Basic Model, an optimized return mapping algorithm is proposed, which is characterized by analytical integration of the hardening law and by solving only a nonlinear scalar equation at the integration point level. To investigate the performance of the proposed algorithm several implicit and explicit stress update algorithms are compared at the integration point level. Finally, the proposed stress update algorithm is applied to a 2D solid-fluid coupled numerical simulation of water flow through a homogeneous embankment dam.

1 INTRODUCTION

The development of constitutive models for partially saturated soil and the implementation into FE-programs are ongoing research topics. The latter requires selection of a suitable stress update algorithm. In addition to accuracy, robustness and efficiency of the employed stress update algorithm play a decisive role especially for large-scale FE-analyses. This is the motivation for developing an optimized return mapping algorithm for the most well-known constitutive model for partially saturated soils, the Barcelona Basic Model (BBM). In the pioneering work [1] basic concepts of modeling the behavior of partially saturated soils were introduced, e. g. the application of two independent stress parameters, consisting in the particular case of net stress and matric suction. The original version of the BBM was developed further by e. g. [7, 14, 2]. However, it is employed here in its original version, because the latter was agreed as the basis for extensive benchmark activities within the framework of the MUSE network [9].

Following the ideas of [5] the proposed algorithm is derived from the general formulation of the return mapping algorithm [13]. Whereas the latter requires solving a system of several nonlinear equations at the integration point level, the former is characterized by analytical integration of the hardening law and by solving only a nonlinear scalar equation at the integration point level.

To investigate the performance of the proposed algorithm several stress update algorithms are compared at the integration point level. This is done on the basis of two sets of material parameters for the BBM by prescribing different volumetric and deviatoric strain increments at constant matric suction. The investigated stress update algorithms include both implicit and explicit integration schemes. The latter include a forward Euler integration scheme and a semi-explicit integration algorithm [10]. The Richardson extrapolation method, described in [3], is used as the basis for sub-stepping with error control, which is an essential ingredient especially of explicit stress update algorithms. Moreover, enhancements of implicit integration methods with sub-stepping and error control techniques are investigated. In addition, a fifth-order Runge–Kutta stress update algorithm with error control is included in this investigation [4].

Finally, the proposed stress update algorithm is applied to a 2D solid-fluid coupled numerical simulation of water flow through a homogeneous embankment dam. The governing equations of the finite element formulation are based on a three-phase model for partially saturated soils, see e. g. [8]. Since in such structural analyses various hydro-mechanical loading conditions are encountered at different integrations points, the robustness and efficiency of the proposed stress update algorithm can be demonstrated.

2 THE BARCELONA BASIC MODEL

The BBM is formulated in terms of the net stress tensor $\boldsymbol{\sigma}''$ and the matric suction p^c . The net stress

$$\boldsymbol{\sigma}'' = \boldsymbol{\sigma} - p^a \mathbf{I} \quad (1)$$

is the total stress $\boldsymbol{\sigma}$ in excess of the pore air pressure p^a , and the matric suction

$$p^c = p^a - p^w \quad (2)$$

is the difference between the air pressure p^a and the water pressure p^w . For stress states located within the elastic domain, enclosed by the yield surface, the elastic volumetric and deviatoric strain rates are given as

$$\dot{\epsilon}_v^e = \frac{\kappa}{1+e} \frac{\dot{p}''}{p''} + \frac{\kappa_s}{1+e} \frac{\dot{p}^c}{p^c + p_{atm}}, \quad \dot{\epsilon}_{ij}^e = \frac{\dot{s}_{ij}}{2G}, \quad (3)$$

with the material parameters κ and κ_s , representing the elastic stiffness for changes of the mean net pressure $p'' = (\boldsymbol{\sigma}'' : \mathbf{I})/3$ and for changes of the matric suction p^c , respectively. e , p_{atm} , \dot{s}_{ij} and G denote the void ratio, the atmospheric air pressure, the deviatoric stress rate and the shear modulus, respectively. It follows from (3₁) that the elastic volumetric

strain rate $\dot{\varepsilon}_v^e$ depends on both the mean net pressure p'' and the matric suction p^c . Within the elastic domain the stress point (p'', e) lies on the unloading-reloading line (URL) with slope κ . For isotropic plastic conditions it lies on the isotropic compression line (ICL) with the suction-dependent slope

$$\lambda(p^c) = \lambda(0) [(1 - r) e^{-\beta p^c} + r] . \quad (4)$$

$\lambda(p^c)$ describes the soil stiffness during plastic loading in a hydrostatic test for a given matric suction p^c in terms of the respective stiffness $\lambda(0)$ at saturated conditions and the material parameters r and β .

The intersection point of the URL and the ICL is denoted as the preconsolidation pressure p_0'' . The ICL is defined by the slope $\lambda(p^c)$ and the void ratio $e = N(p^c) - 1$ at $p'' = 1$ with $N(p^c)$ denoting the respective specific volume. From the volumetric behavior of the BBM follows

$$e = N(p^c) - 1 - \lambda(p^c) \ln \left(\frac{p_0''}{1} \right) + \kappa \ln \left(\frac{p_0''}{p''} \right) , \quad (5)$$

from which

$$p_0'' = \exp \left(\frac{-N(p^c) + 1 + e}{\kappa - \lambda(p^c)} \right) p''^{\frac{\kappa}{\kappa - \lambda(p^c)}} \quad (6)$$

is obtained. The yield surface is defined as

$$f = J_2 - \frac{M^2}{3} (p'' + p_s'') (p_0'' - p'') \quad (7)$$

with the second invariant of the deviatoric stress tensor $J_2 = s_{ij}s_{ij}/2$ and

$$p_s''(p^c) = k_s p^c , \quad p_0'' = p_{ref}'' \left(\frac{(p_0'')^*}{p_{ref}''} \right)^{\frac{\lambda(0) - \kappa}{\lambda(p^c) - \kappa}} . \quad (8)$$

In (7) M defines the slope of the critical state line. p_s'' and p_0'' both depend on the matric suction according to (8). For negative values of p'' the intersection of the yield surface (7) with the plane $J_2 = 0$ is given by p_s'' according to (8₁) with the material parameter k_s describing the increase in cohesion due to the matric suction. The preconsolidation pressure p_0'' and the one for saturated conditions $(p_0'')^*$ are located on the so called loading collapse yield curve (LC curve) according to (8₂). This curve is the intersection of the yield surface with the plane $J_2 = 0$ for positive values of p'' . Here, p_{ref}'' serves as a reference pressure such that for $(p_0'')^* = p_{ref}''$ (8₂) degenerates to $p_0'' = p_{ref}'' = \text{const.}$

The plastic strain rate is determined from the non-associated flow rule

$$\dot{\varepsilon}_{ij}^p = \dot{\gamma} \frac{\partial g}{\partial \sigma_{ij}''} \quad (9)$$

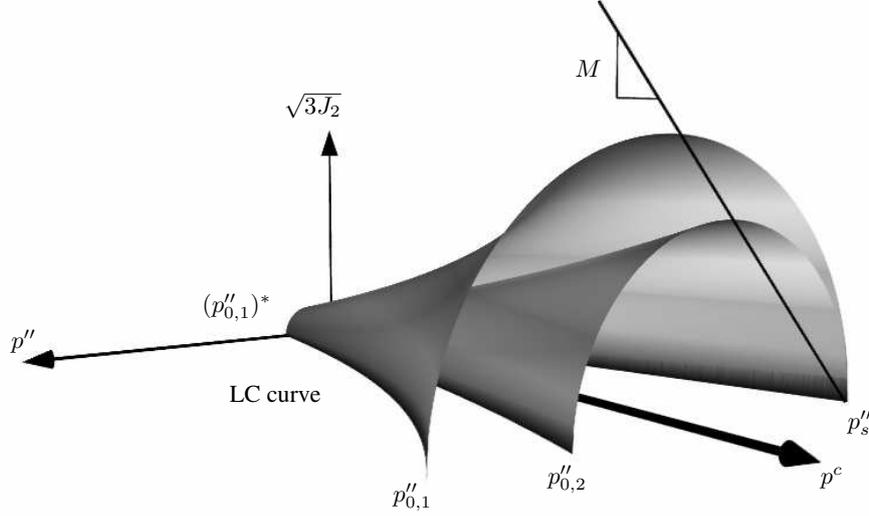


Figure 1: Yield surface of the BBM for different values of the preconsolidation pressure

with the flow potential

$$g = \alpha J_2 - \frac{M^2}{3} (p'' + p_s'') (p_0'' - p'') , \quad (10)$$

where α is a constant. The hardening law relates the rate of the preconsolidation pressure at saturated conditions $(\dot{p}_0'')^*$, which serves as the hardening parameter, to the volumetric plastic strain rate $\dot{\varepsilon}_v^p$ by

$$(\dot{p}_0'')^* = (p_0'')^* \frac{1 + e}{\lambda(0) - \kappa} \dot{\varepsilon}_v^p . \quad (11)$$

(11) describes the evolution of the yield surface. The latter is shown for two different values of $(p_0'')^*$ in Fig. 1.

3 AN OPTIMIZED RETURN MAPPING ALGORITHM

For deriving a computationally efficient version of the return mapping algorithm the flow rule (9) is split into a volumetric and deviatoric part

$$\dot{\varepsilon}_v^p = \dot{\gamma} 3 \frac{\partial g}{\partial I_1''} , \quad \dot{e}_{ij}^p = \dot{\gamma} \frac{\partial g}{\partial J_2} s_{ij} . \quad (12)$$

Backward Euler integration of (12) yields

$$\begin{aligned}\Delta\varepsilon_v^p &= \gamma 3 \frac{\partial g}{\partial I_1''} = \Delta\varepsilon_v - \Delta\varepsilon_v^e, \\ \Delta e_{ij}^p &= \gamma \frac{\partial g}{\partial J_2} s_{ij} = \Delta e_{ij} - \Delta e_{ij}^e,\end{aligned}\tag{13}$$

where $\gamma = \dot{\gamma}\Delta t$. Note that quantities with the subscript n refer to the converged values at the previous time instant t_n , whereas all other quantities refer to the values at the current time instant t_{n+1} . In case of a constant value of G

$$\begin{aligned}s_{ij} = 2Ge_{ij}^e &= 2G [(e_{ij} - e_{ij,n}^p) - (e_{ij}^p - e_{ij,n}^p)] \\ &= s_{ij}^{Trial} - 2G\Delta e_{ij}^p\end{aligned}\tag{14}$$

follows from (3₂). Inserting (13) into (14) gives

$$\left(1 + \gamma \frac{\partial g}{\partial J_2} 2G\right) s_{ij} = s_{ij}^{Trial}.\tag{15}$$

The term enclosed by the brackets is a scalar quantity, hence, s_{ij} and s_{ij}^{Trial} differ only by a scalar factor. Thus, from (15) it follows

$$\left(1 + \gamma \frac{\partial g}{\partial J_2} 2G\right)^2 J_2 = J_2^{Trial}.\tag{16}$$

Making use of $\gamma = (\Delta\varepsilon_v - \Delta\varepsilon_v^e)/(3\partial g/\partial I_1'')$, resulting from (13₁), yields

$$\left(3 \frac{\partial g}{\partial I_1''} - \frac{\partial g}{\partial J_2} 2G(\Delta\varepsilon_v^e - \Delta\varepsilon_v)\right)^2 J_2 - \left(3 \frac{\partial g}{\partial I_1''}\right)^2 J_2^{Trial} = 0.\tag{17}$$

In (17) the incremental volumetric strain $\Delta\varepsilon_v$ is known from the current estimate of the displacement increment at t_{n+1} . J_2 and $\Delta\varepsilon_v^e$ in (17) can be replaced by

$$J_2 = \frac{M^2}{3} (p'' + p_s'')(p'' - p'')\tag{18}$$

and

$$\Delta\varepsilon_v^e = \frac{\kappa}{1+e} \ln\left(\frac{p''}{p_n''}\right) + \frac{\kappa_s}{1+e} \ln\left(\frac{p^c + p_{atm}}{p_n^c + p_{atm}}\right)\tag{19}$$

following from (7) and from integration of the rate constitutive equation (3₁). The rate of the void ratio is given by

$$\dot{e} = -(1+e)\dot{\varepsilon}_v.\tag{20}$$

Integration of (20) yields the value of the void ratio at t_{n+1}

$$e = (1+e_n) \exp(-\Delta\varepsilon_v) - 1.\tag{21}$$

(17) together with (18), (19), (21) and (6) represents a nonlinear scalar equation for the unknown I_1'' (or $p'' = I_1''/3$), which can be solved, e.g., by the Newton method. Once p'' has been determined from this equation, it is inserted into (6), yielding p_0'' , and the latter into the recast form of (8₂) yielding the hardening parameter $(p_0'')^*$.

4 COMPARISON OF DIFFERENT STRESS UPDATE ALGORITHMS

To investigate the performance of the proposed stress update algorithm, the following stress update algorithms are considered in the subsequent comparative study:

- (a) an explicit stress update algorithm, characterized by forward integration of the constitutive rate equation

$$\dot{\boldsymbol{\sigma}} = \mathbb{C}^{ep} \dot{\boldsymbol{\varepsilon}} + \mathbf{C}^{p^c, ep} \dot{p}^c \quad (22)$$

with $\mathbb{C}^{ep} = \partial \boldsymbol{\sigma} / \partial \boldsymbol{\varepsilon}$ and $\mathbf{C}^{p^c, ep} = \partial \boldsymbol{\sigma} / \partial p^c$ denoting the constitutive tangent operators, which is combined with adaptive sub-stepping and error control based on the Richardson extrapolation method;

- (b) a general return mapping algorithm [13], characterized by backward Euler integration of the rate equations for the plastic strains (9) and the hardening variable (11) and by enforcing the condition $f = 0$ for the yield function (7) at t_{n+1} , which requires solving a system of nonlinear equations, consisting of the consistency parameter, the net stress and the hardening variable;
- (c) the optimized return mapping algorithm described in section 3 and [6] respectively;
- (d) a semi-explicit stress update algorithm [10], characterized by explicit integration of the rate equations for the plastic strains (9) and the hardening variable (11) and by enforcing the condition $f = 0$ for the yield function (7) at t_{n+1} for determining the consistency parameter, which is combined with adaptive sub-stepping and error control based on the Richardson extrapolation method;
- (e) the implicit fifth-order Runge–Kutta integration algorithm with error control RADAU5, proposed in [4].

A comparison of the investigated stress update algorithms with respect to the accuracy is performed on the basis of two different sets of material parameters for the BBM, provided in [1]. To this end, the error of the stresses computed for prescribed combinations of volumetric and deviatoric strain increments are considered. The error is defined as a relative error, related to the "exact" value for the respective stress component computed by the RADAU5 algorithm [4] prescribing an extremely small error tolerance of 10^{-10} .

Fig. 2 shows a comparison of the integration errors obtained by means of the general return mapping algorithm and the optimized return mapping algorithm for the relatively large range of strain increments ranging from 0 up to 3%. A particular point of the diagrams shown in Fig. 2 indicates the error of a single step stress update for a particular combination of volumetric and deviatoric strain increment $(\varepsilon_v, \varepsilon_s)$. E.g., the point $(\varepsilon_v = 2\%, \varepsilon_s = 3\%)$ represents the integration error for the strain increment $\Delta \varepsilon_v = 0.02$ and $\Delta \varepsilon_s = 0.03$, obtained by a single step backward Euler integration. According to Fig. 2 the integration errors for the investigated strain increments reach up to 40%. Contrary

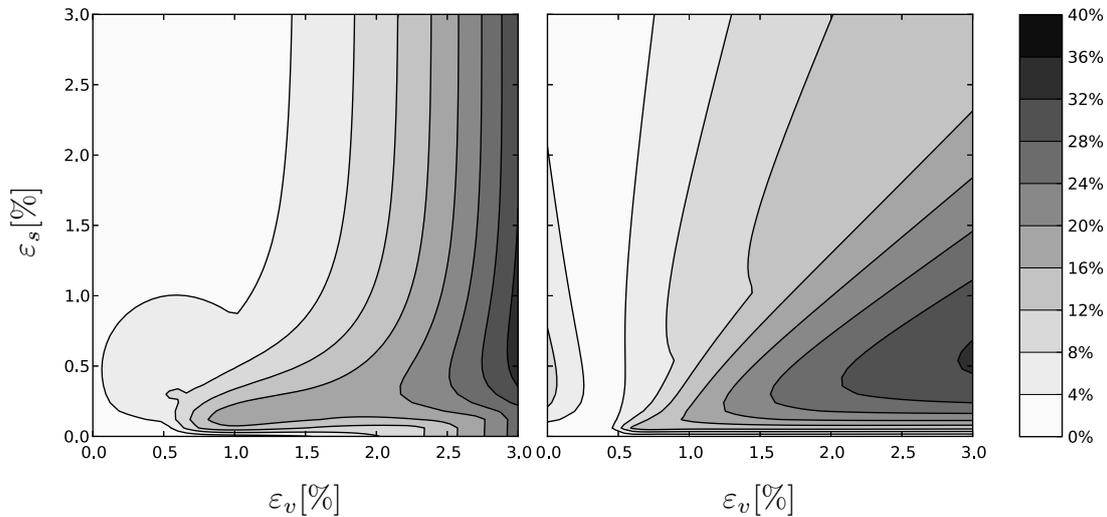


Figure 2: Integration errors for the general return mapping algorithm (left) and the optimized return mapping algorithm (right)

to the general return mapping algorithm the optimized return mapping algorithm gives the exact solution for hydrostatic strain paths due to the analytical integration of the hardening law. Because of the large integration errors, similar to the explicit and semi-explicit stress update algorithm, the return mapping algorithms are also enhanced by adaptive sub-stepping and error control.

For the investigation of the efficiency of the considered stress update algorithms the stresses were computed for 25 combinations of volumetric and deviatoric strain increments of 0.5%, 0.75%, 1.0%, 1.25% and 1.5% for prescribed maximum values of the integration error, ranging from 10^{-1} to 10^{-10} . The mean values of the number of required instructions are shown in the diagrams of Fig. 3 for the resulting mean values of the computed errors.

It follows from Fig. 3 that for a prescribed error tolerance the optimized return mapping algorithm is by far more efficient than the general return mapping algorithm and it is even more efficient than the explicit integration method. The RADAU5 algorithm is very efficient for very small prescribed values of error tolerances.

5 APPLICATION TO A COUPLED FINITE ELEMENT ANALYSIS

The application of the developed optimized return mapping algorithm is demonstrated by a coupled solid-fluid Finite Element analysis of the water flow through a homogeneous earth dam. The cross section of the homogeneous earth dam is shown in Fig. 4, the employed hydraulic parameters and material parameters are reported in [11, 12]. A rigid foundation of the dam is assumed by constraining the displacements on the bottom of the dam. For the undrained part of the base an impermeable boundary is assumed, whereas for the drained part a permeable boundary is considered by applying a mass flux with a

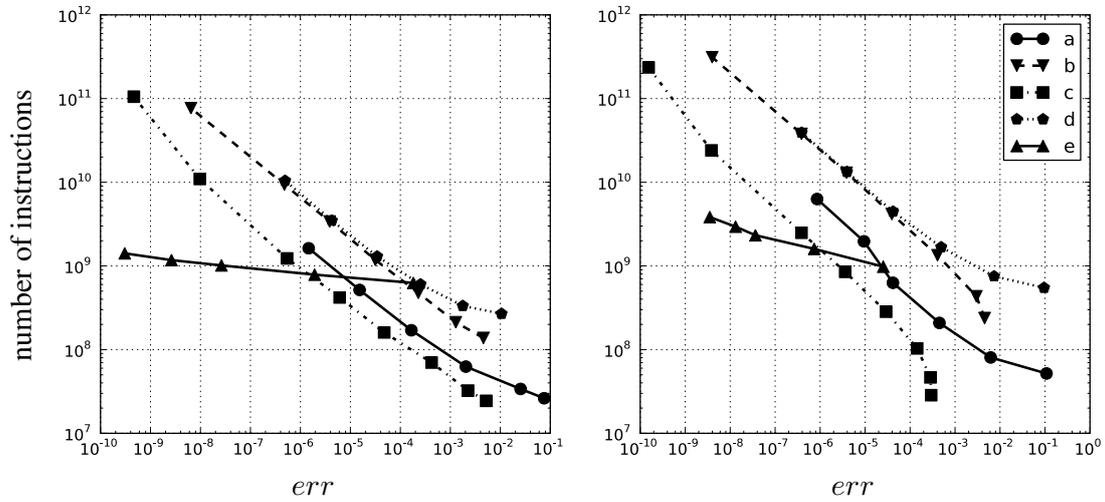


Figure 3: Work precision diagrams for two sets of material parameters: (a) explicit stress update, (b) general return mapping algorithm, (c) optimized return mapping algorithm, (d) semi-explicit stress update algorithm, (e) implicit fifth-order Runge–Kutta algorithm

pressure dependent velocity

$$\begin{aligned} v_n &= k_{sc} p^w \quad \text{for } p^w > 0, \\ v_n &= 0 \quad \text{for } p^w \leq 0, \end{aligned} \tag{23}$$

with p^w denoting the water pressure at the boundary and k_{sc} is a sufficiently large seepage coefficient to approximately enforce the requirement of a zero water pressure for a freely draining surface. Similar boundary conditions for the fluid phase are applied at the free surfaces, i.e. at the upstream slope above the water level, the crest and the downstream slope.

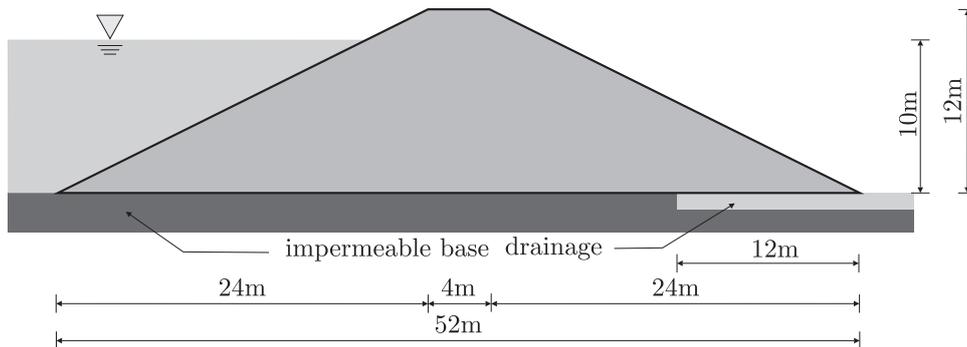


Figure 4: Cross section of the homogeneous earth dam

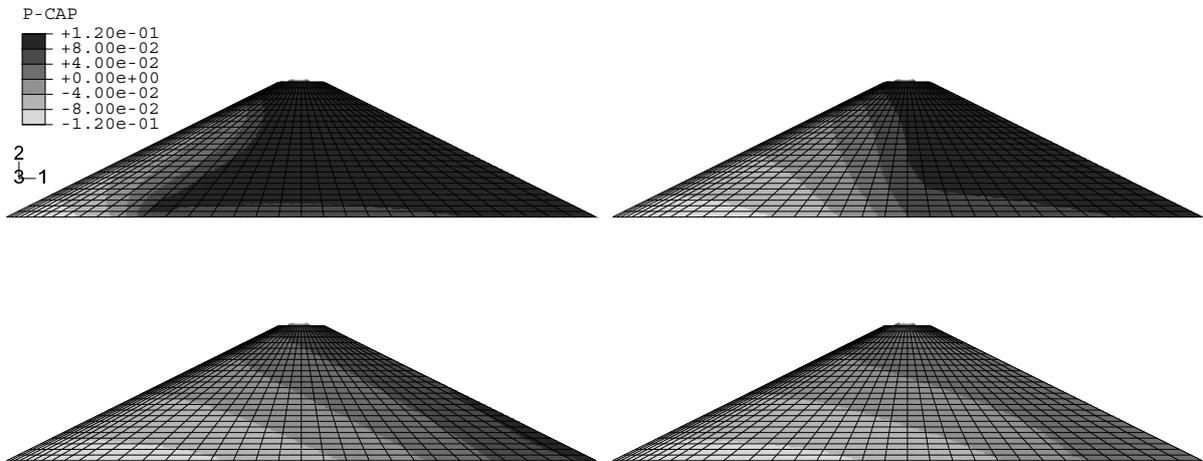


Figure 5: Distribution of matric suction for selected time instants from top to bottom: (a) $t = 20$ d, (b) $t = 60$ d, (c) $t = 310$ d, (d) $t = 800$ d

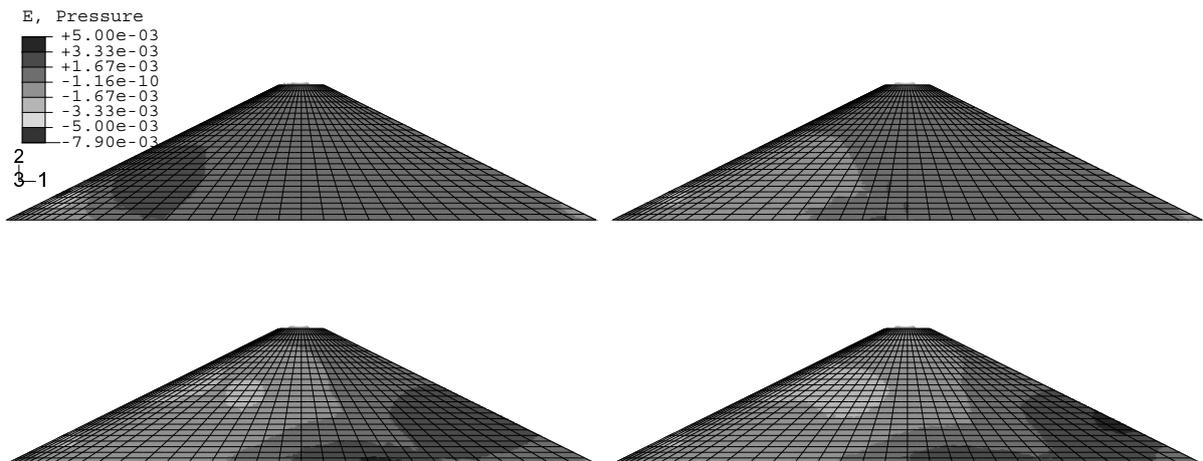


Figure 6: Distribution of the volumetric strain for selected time instants: $t = 20$ d, $t = 60$ d, $t = 310$ d, $t = 800$ d

In the first step of the numerical analysis a constant matric suction of 100 kPa, corresponding to an initial degree of water saturation of $S^w = 0.729$, is assumed for the dam body and the primary stresses due to dead load are computed presuming elastic response. In the second step the net stresses, the void ratio and matric suction, computed in the first step, are taken as initial values, whereas the displacements are set equal to zero and matric suction at the upstream boundary is reduced to zero by specifying the respective boundary conditions. In the subsequent steps of the analysis the transient seepage flow

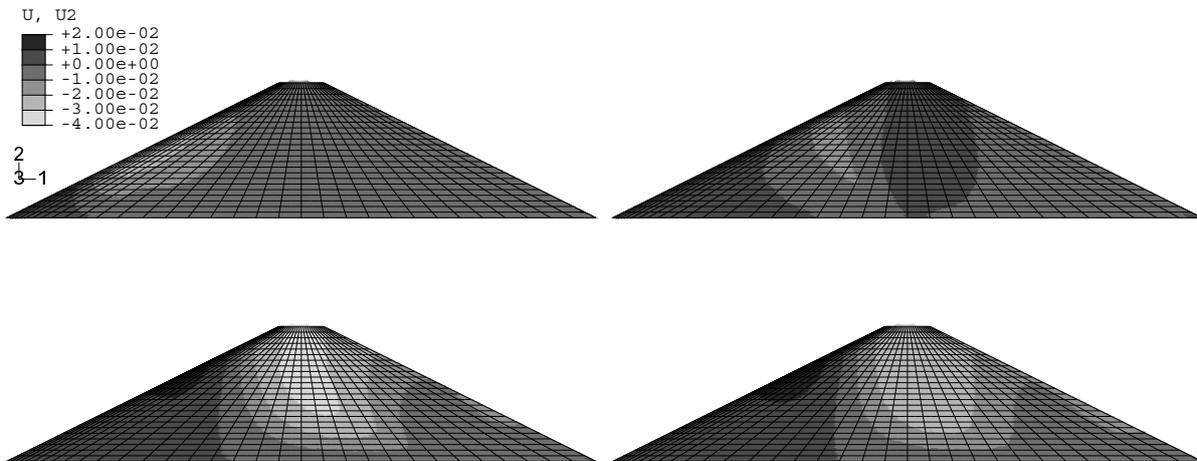


Figure 7: Vertical displacements for selected points of time from top to bottom: (a) $t = 20$ d, (b) $t = 60$ d, (c) $t = 310$ d, (d) $t = 800$ d

due to a water table of $H = 10$ m is computed until steady state conditions are attained.

Figure 5 shows the distribution of matric suction for selected time instants. Negative values of matric suction represent values of the water pressure in fully water saturated regions. 20 days after impoundment only a small region close to the upstream slope is affected by the water flow. With advancing time the water saturated domain is propagating until steady state conditions are attained.

The distribution of the volumetric strain with advancing saturation front is shown in Fig. 6. Volumetric compaction, indicated by positive values of the volumetric strain, occurs in the lower central region and in the downstream regions which are characterized by partial saturation. The computed distribution of the vertical displacements of the soil skeleton (negative values denote settlements) is shown in Figure 7. Shortly after the impoundment settlements in the vicinity of the upstream face of the dam occur, whereas at later stages uplifting is predicted in this region. By contrast, in the upper central region of the dam, the settlements increase with advancing saturation front.

6 CONCLUSIONS

In this paper an optimized return mapping algorithm was proposed, which is characterized by analytical integration of the hardening law and by solving only a nonlinear scalar equation at the integration point level. To investigate the performance of the developed algorithm, several stress update algorithms were compared with respect to accuracy and efficiency: (a) an explicit stress update algorithm, (b) a general return mapping algorithm, (c) the proposed optimized return mapping algorithm, (d) a semi-explicit stress update algorithm, and (e) an implicit fifth-order Runge-Kutta stress update algorithm. Large integration errors were encountered for the return mapping algorithms when larger

strain increments were integrated in one step. Hence, similar to the explicit and semi-explicit stress update algorithm, they were enhanced by adaptive sub-stepping and error control. It was shown that for a prescribed error threshold value the optimized return mapping algorithm is by far more efficient than the general return mapping algorithm and it is even more efficient than the explicit integration method. Finally, to demonstrate the robustness and efficiency of the proposed algorithm within a Finite Element context, a simplified 2D coupled transient numerical simulation of the behavior of an embankment dam due to impoundment was performed.

References

- [1] Alonso, E. E., A. Gens, and A. Josa (1990). A constitutive model for partially saturated soils. *Géotechnique* 40, 405–430.
- [2] Cui, Y. J., P. Delage, and N. Sultan (1995). An elasto-plastic model for compacted soils. In *Unsaturated soils*, Volume 2, pp. 703–709. Balkema, Rotterdam.
- [3] Fellin, W., M. Mittendorfer, and A. Ostermann (2009). Adaptive integration of constitutive rate equations. *Computers and Geotechnics* 36(5), 698–708.
- [4] Hairer, E. and G. Wanner (1996). *Solving Ordinary Differential Equations II, Stiff and Differential-Algebraic Problems* (2nd rev. ed.). Springer Berlin Heidelberg.
- [5] Hickman, R. J. and M. Gutierrez (2005). An internally consistent integration method for critical state models. *International Journal for Numerical and Analytical Methods in Geomechanics* 29(3), 227–248.
- [6] Hofmann, M. (2010). *Integrationsalgorithmen und Parameteridentifikation elasto-plastischer Stoffgesetze für teilgesättigte Materialien*. Ph. D. thesis, Universität Innsbruck.
- [7] Josa, A., A. Balmaceda, A. Gens, and E. E. Alonso (1992). An elastoplastic model for partially saturated soil exhibiting a maximum of collapse. In *3rd international conference on computational plasticity*, Volume 1, Barcelona, pp. 815–826.
- [8] Lewis, R. W. and B. A. Schrefler (1998). *The Finite Element Method in the Static and Dynamic Deformation and Consolidation of Porous Media* (2 ed.). John Wiley & Sons.
- [9] Marie Curie Research Training Network (2009). <http://muse.dur.ac.uk>. (April 17, 2009).
- [10] Mittendorfer, M. (2006). *Interne Differentiation nichtlinearer anelastischer Materialmodelle*. Master’s thesis, Leopold-Franzens-Universität Innsbruck.

- [11] Pertl, M. (2010). *Grundlagen, Implementierung und Anwendung eines Drei-Phasen Modells für Böden*. Ph. D. thesis, Universität Innsbruck.
- [12] Pertl, M., M. Hofmann, and G. Hofstetter (2011). Coupled solid-fluid fe-analysis of an embankment dam. *Frontiers of architecture and civil engineering in China* 5(1), 53–62.
- [13] Simo, J. and T. Hughes (1998). *Computational Inelasticity*. Springer New York.
- [14] Wheeler, S. J. and V. Sivakumar (1995). An elasto-plastic critical state framework for unsaturated soil. *Géotechnique* 45, 35–53.

DYNAMIC BEHAVIOUR OF SATURATED POROELASTIC LAYERS WITH EMBEDDED WALL SUBMITTED TO SEISMIC ACTIONS

GRAZINA, J. C.* AND PINTO, P. L.†

Dep. Engenharia Civil – Fac. Ciências e Tecnologia da Universidade de Coimbra (FCTUC)
Universidade de Coimbra
Polo II, 3030-788 Coimbra, Portugal
*e-mail: graza@dec.uc.pt
†e-mail: ppinto@dec.uc.pt

Key words: Soil-structure interaction, Finite elements model, Poroelastic saturated materials, Time domain analysis, Viscous damping.

Summary. Behaviour of poroelastic saturated materials submitted to dynamic actions is strongly dependent of the solid skeleton permeability and the frequency of the movement. Depending on these quantities, undrained behaviour or fully drained behaviour occurs for total coupled interaction or for null interaction, respectively. Between these limit cases, some relative movement occurs among solid skeleton and fluid, generating viscous damping, which, in turn, modifies the elastic response of the system. This paper presents results of coupled behaviour in poroelastic saturated layers with an embedded impermeable wall, submitted to seismic actions. A Finite Element code developed at the University of Coimbra (FEMEPDYN), with coupled formulation u^s-u^w-p , was used for this purpose. Dynamic responses of poroelastic layers with an embedded wall are compared with non-porous materials for similar conditions. Damping of non-porous materials was previously calculated by calibration of the Rayleigh coefficients in order to match the free field responses for both materials types. Unlike for the free field analysis, results reveal some differences between both materials responses with the presence of the wall, due to the volumetric deformations imposed by the embedded wall. Also, those volumetric deformations have a more uniform distribution in poroelastic layers with permeabilities that represent near total and near fully drained behaviour.

1 INTRODUCTION

Behaviour of poroelastic saturated materials is actually described using two theories: the Biot theory [1, 2], based on the Lagrange classical mechanics, and the Porous Mixtures theory, firstly presented by Fillunger (1913), which involves the concept of volume fractions [3]. Considering some simplifications and the incompressibility of the constituents, both theories result in similar coupled equilibrium equations, where the interaction force between porous solid skeleton and the interstitial fluid is expressed by:

$$R = \frac{n^2 \gamma_w}{k} (v^w - v^s) \quad (1)$$

where k is the permeability, n is the porosity, γ_w is the bulk unit weight of fluid, v^s is the velocity of solid and v^w is the velocity of fluid.

Interaction forces between the solid skeleton and the fluid in a poroelastic saturated material are greatly dependent on the permeability and frequency of the movement. Zienkiewicz et al. [4] presented parametric analyses for wide ranges of permeabilities and frequencies of harmonic vertical excitations in a poroelastic column. These authors found bound parameters related to the limit levels of coupled interaction, from near drained behaviour (quasi-null coupled interaction) to near undrained behaviour (quasi-total coupled interaction), for a range of loading velocities from the quasi-static (consolidation problem in despite to inertial forces) to the rapid loading. For near undrained behaviour, relative movements between solid and fluid are very small and viscous interaction forces developed are insignificant. In the opposite situation, for near full drained behaviour, permeabilities are very high and, consequently, viscous interaction forces are very small, in spite the existence of large relative movements between both phases. In between these limiting situations, relevant viscous interaction forces may be developed for common dynamic loadings, as those originated by foundations of industry equipment or earthquakes.

Viscous damping values due to coupled interaction forces were previously calculated on poroelastic saturated layers subjected to shear free field movements [5]. In these analyses, only shear deformations are induced and, consequently, excess pore pressures are inexistent. Permeabilities for near limit cases of coupled interaction with quasi-null damping were found, as well the damping ratios for middle range permeabilities. The absence of excess pore pressure generation allows the use of non-porous linear elastic materials for the calculations of the dynamic responses: with dry properties for quasi-null coupled interaction and with saturated undrained properties for quasi-total coupled interaction. In between these limit cases, appropriate Rayleigh damping can be used in non-porous materials for the same purpose [6]. Main advantages of the use of non-porous materials in finite elements models are the much lower effort and time computing.

Shear movements induced on saturated porous layers with embedded impermeable walls produce non-homogeneous fields of volumetric deformations and excess pore pressures generation. This factor influences the whole response of the layer and tends to disregard the previously related in free field conditions. Analyses are presented for an artificial seismic action applied at the base of the layers.

2 NUMERICAL MODELLING

2.1 Finite element code

The numerical analyses were carried using the finite element code FEMEPDYN [6], developed at the University of Coimbra. FEMEPDYN code uses the Generalized- α time integration algorithms to perform dynamic calculations in time domain, for both non-porous and porous materials. The porous materials are modelled by the u^s - u^w - p coupled formulation [7, 8]. At each time step this formulation enables the computation of nodal displacements, velocities and accelerations (d.v.a.) of both solid and fluid phases, as well pore pressures

(p.w.p.) at the corner nodes of the mesh elements. For this purpose, finite element meshes are composed by quadrangular hybrid elements with 8 nodes of Q8/C4 type, which ensures fluid pressure continuity between elements (Figure 1a). Meshes for non-porous materials are composed with isoparametric quadrangular elements of 8 nodes.

2.2 Materials properties

Properties of the poroelastic and non-porous materials are presented in Table 1 and Table 2, respectively. Shear modulus values of 20 MPa (G20) and 80 MPa (G80) were considered for both materials types. The pore fluid in the poroelastic analyses is almost incompressible, with a bulk modulus of $K_w=1\times 10^8$ MPa. The permeability coefficients adopted varies in a wide range with the purpose to achieve the two near limit cases of viscous interaction – for the quasi-null coupled behaviour unrealistic high permeability coefficients were used.

The reproductions of poroelastic behaviours were carried with two types of non-porous materials: for the drained behaviours and allowable volumetric deformations with the Poisson coefficient of the solid skeleton ($\nu=0.3$); for the undrained behaviours with a Poisson coefficient of $\nu=0.49$. Differences in results between these non-porous materials should not be detectable in movements without volumetric deformations, as occurs in pure shear column analyses. The bulk unit weights of non-porous materials were calculated using the solid skeleton and pore fluid densities of the poroelastic materials, respectively ρ_s and ρ_w . For movements caused by shear wave propagation, bulk unit weights should be calculated considering saturated and dry materials, respectively for undrained and drained conditions [9]. Densities of saturated and dry materials, ρ_{sat} and ρ_d , are calculated using Equations 2 and shear wave velocities, v_s , are calculated using Equation 3 for respective ρ value.

$$\text{Saturated: } \rho = \rho_{sat} = (1-n)\rho_s + n\rho_w \quad (2)$$

$$\text{Dry: } \rho = \rho_d = (1-n)\rho_s$$

$$v_s = \sqrt{\frac{G}{\rho}} \quad (3)$$

Table 1: Properties of poroelastic materials

G (MPa)	E_s (MPa)	ν	ρ_s (kg/m ³)	ρ_w (kg/m ³)	K_w (MPa)	n	k (m/s)
20	52	0.3	2.6×10^3	1.0×10^3	1×10^5	0.365	1×10^2 to 1×10^{-5}
80	208						

Table 2: Properties of non-porous materials

G (MPa)	E_s (MPa)	ν	Coupled interaction / Material	γ (kN/m ³)	v_s (m/s)
20	52.0	0.3	Null / Dry	16.20	110.1
	59.6	0.49	Total / Saturated	19.78	99.6
80	208.0	0.3	Null / Dry	16.20	220.1
	238.4	0.49	Total / Saturated	19.78	199.2

The embedded wall has a linear elastic behaviour and was considered as a non-porous material in all the analyses. The elastic properties are respective to a common concrete material, with an elasticity modulus of $E=20$ GPa and a Poisson coefficient of $\nu=0.2$.

2.3 Models meshes

The modelled media consists in homogeneous layers with 20 m thickness, settled over a rigid and impervious bedrock material. The width of the FE domain depends of the analyses type: for the free field response of the layer, the existence of only pure shear horizontal movements allows the use of a single column; for the analyses with the embedded wall, a larger mesh of 50.4 m width is used in order to mitigate the influence of the lateral boundaries on wall movements. In both cases, the lateral boundaries have restrictions for vertical displacements and the bottom boundary is impervious and has vertical displacement restrictions. In free vibration analyses, the bottom boundary has also horizontal displacements restrictions. In other analyses, horizontal accelerations (harmonic or seismic) are imposed at the bottom boundary. The stiffness of layer is related with velocity of wave propagation (Eq. 3) and modal frequencies (Eq. 4). Higher values require more refined meshes in wave propagation direction to obtain accurate results. Therefore, for layers G20 and G80, 20 and 40 elements uniformly distributed in vertical direction were adopted in respective meshes.

In the FE models with embedded wall, interface joint elements with no thickness were inserted between wall and surrounding layer. Parameters of constitutive model of these joint elements were chosen with extremely low tangential stiffness and high normal stiffness. In these conditions, an almost free shear displacement is allowed and avoided the opening of gaps in the mesh. Formulation of these joint elements may consider both impervious or permeable interfaces. In the analyses here presented, permeable case is considered. The wall is 12 m deep and 0.4 m thick.

Figure 1 presents a scheme of the hybrid elements used in poroelastic analyses, the FE mesh column used in the free field analyses and the right half of the mesh used in the analyses with the embedded wall.

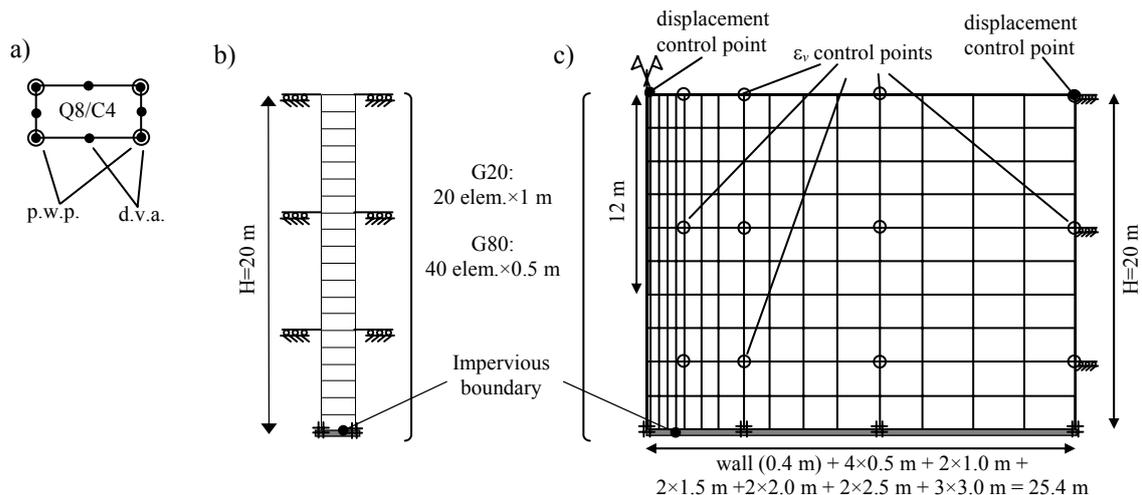


Figure 1: Finite element meshes: a) Hybrid elements; b) Free field; c) Layer with embedded wall.

2.4 Seismic action

The seismic action is an artificial accelerogram generated by the SeismoSignal software [10], so as to match an elastic response spectrum defined by the Eurocode 8 [11]. Parameters of this spectrum were chosen for the seismic zone of Coimbra (Portugal), a seismic action of Type 2 (lower magnitude and epicentre in continental territory) and for Ground Type A (bedrock layer, considering the seismic action applied at the rigid base boundary of the finite element mesh). Figure 2 presents the accelerogram adopted in the analyses with a time duration of 22 seconds, that includes a stationary part of 10 seconds, and a maximum ground acceleration of $a_g=1.3 \text{ m/s}^2$. The time steps used are of 0.01 seconds for G20 layers and 0.005 seconds for G80 layers (regarding the higher frequencies existent in this layers movement).

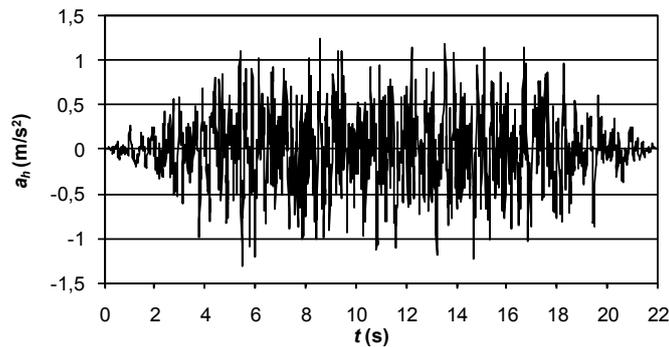


Figure 2: Artificial accelerogram used in the analyses

3 EVALUATION OF VISCOUS DAMPING IN FREE FIELD SHEAR MOVEMENT

3.1 Methodology

Here is presented a description of the method to evaluate the modal frequencies and respective damping ratio of poroelastic saturated layers in free field conditions. Analyses were made for the range of permeabilities presented in Table 1.

Firstly, a free vibration horizontal movement was induced at the layer by releasing a pre-imposed displacement on the top of the column. From the free vibration responses, modal frequencies were calculated using a Discrete Fourier Transformer (DFT) algorithm [12]. Secondly, harmonic shear accelerations, with the previously calculated modal frequencies, were imposed at the base of the layer exciting only one mode at once. The one-mode response tends to be infinitely amplified in the absence of viscous damping, as approximately occurs for permeability values near the limit cases of total and null coupled interaction. For middle range permeabilities, viscous damping restricts amplification of the movements and a steady state response is achieved. As an example, the amplified response for the 1st mode of the layer G20 with $k=1 \times 10^{-1} \text{ m/s}$ is presented in Figure 3. The amplification value obtained, D , allows the calculation of the damping coefficient, ξ , for each vibration mode, using the simplified equation for homogeneous layers [13]:

$$D = \sinh \left((2n-1) \xi \frac{\pi}{2} \right)^{-1} \quad (4)$$

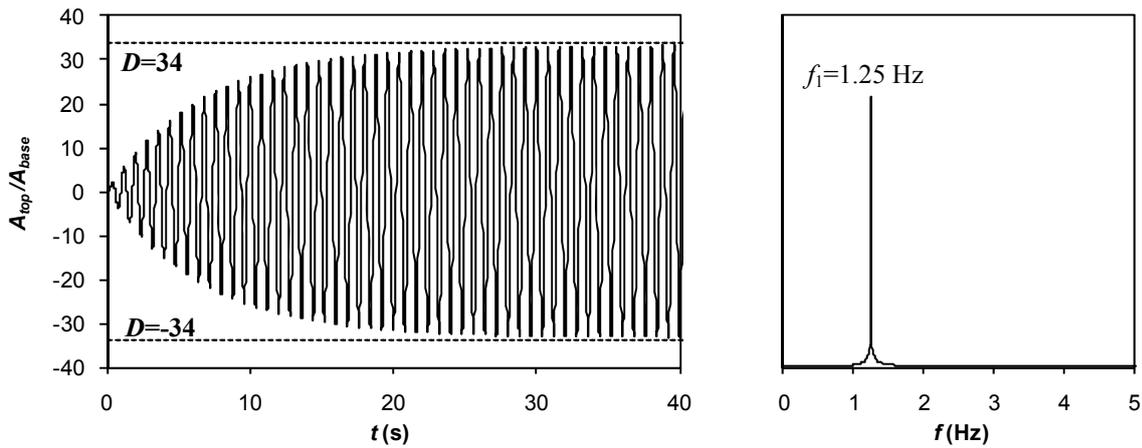


Figure 3: Amplified response and frequency spectra for harmonic action with 1st mode frequency.

3.2 Free vibration responses

The elastic deformations due to a pre-imposed displacement, d_0 , on the top of the column, released at $t=0$ s, induce a pure shear movement in free vibration mode. Depending on the coupled interaction between the fluid and solid skeleton, ruled by permeability, different levels of viscous damping are developed. In Figure 3 are presented the envelopes of the normalized displacements during the first 40 seconds of free vibration movement.

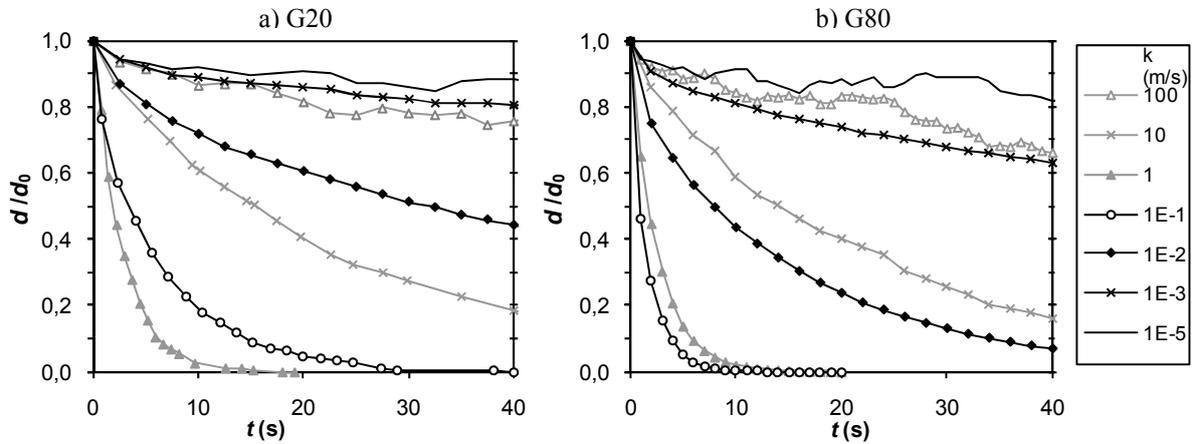


Figure 4: Time variation of displacement amplitudes in free vibration mode for several permeabilities.

The figures show a sharp decrease of displacement amplitudes with decreasing permeabilities from $k=100$ m/s to 1 m/s for G20 and to 0.1 m/s for G80. For these last permeabilities a noticeable higher viscous damping is developed during the movement. For lower permeabilities, viscous damping tends to decrease as can be noticed by the less damped movements. Lower damping occurs for $k=1 \times 10^{-5}$ m/s and $k=100$ m/s, denoting their similitude of behaviour with the undamped coupled limit cases. Anyway, time decreasing amplitudes are noticeable for these higher and lower values of k , particularly for $k=100$ m/s, which means that limit cases are not completely achieved and some residual damping is present.

3.3 Modal frequencies

Analytical value of the n modal frequency for homogeneous layers can be calculated with the well known expression:

$$f_n = (2n-1)v_s/4H \quad (5)$$

where v_s is the shear wave velocity (Eq. 3), and H is the thickness of the layer.

The results of the 1st and 2nd modal frequencies detected from spectral distributions for the aforementioned free vibration responses, as well analytical values obtained from Equation 5, are compiled in Table 3. Cases where dispersive frequency distribution hinders the achievement of accurate frequency higher modes are also presented in brackets.

Analytical and calculated values for non-porous materials are in good agreement, which is a trustful indication that appropriate meshes and time steps had been used. This table shows that modal frequencies of poroelastic materials can be separated in two sets, each one with values similar to those of coupled limit cases calculated with non-porous materials. In each set, the pure shear behaviour of the poroelastic material can be reproduced using the respective non-porous material with an appropriate viscous damping. It is visible that the transition from dry to saturated behaviours occurs suddenly for permeabilities between $k=1\text{m/s}$ to $k=1\times 10^{-1}\text{m/s}$, for both G20 and G80, when viscous damping effect is more notorious (as shown in Figure 4).

Table 3: Modal frequencies for different coupled interaction levels

Coupled interaction	G20		G80		
	Permeability	Modal frequencies		Modal frequencies	
	k (m/s)	f_1 (Hz)	f_2 (Hz)	f_1 (Hz)	f_2 (Hz)
Null / Analytical ($\rho=\rho_d$)	∞	1.376	4.127	2.752	8.255
Null / Non-porous ($\rho=\rho_d$)	∞	1.375	4.120	2.750	8.240
Poroelastic	100	1.375	4.120	2.750	8.240
	10	1.375	4.120	2.750	8.245
	1	1.350	4.125	2.740	8.240
	10^{-1}	1.250	(3.908)	2.540	(8.375)
	10^{-2}	1.245	(3.735)	2.490	(7.490)
	10^{-3}	1.245	3.730	2.490	(7.460)
	10^{-4}	1.245	3.730	2.490	7.460
10^{-5}	1.245	3.730	2.490	7.460	
Total / Analytical ($\rho=\rho_{sat}$)	0	1.245	3.735	2.490	7.470
Total / Non-porous ($\rho=\rho_{sat}$)	0	1.245	3.730	2.490	7.460

() Inaccurate values.

3.4 Viscous damping

The steady-state responses for harmonic actions with modal frequencies allow the calculation of damping ratios, ξ , of the first two modes for the middle range permeabilities. For the lowest and highest permeability values, the residual damping present is not enough to hinder an increasing amplification of the responses, therefore, to establish a ξ value. Also, determination of the ξ values for higher modes is defected of the indelible presence of the lower modes. Despite these limitations, values of ξ presented in Table 4 are in accordance with damped responses presented in Figure 3.

Table 4: Modal damping ratios for different coupled interaction levels

k (m/s)	G20 layer		G80 layer	
	ξ_1 (%)	ξ_2 (%)	ξ_1 (%)	ξ_2 (%)
100	<0.05	--	<0.05	--
10	0.45	0.18	0.23	(0.12)
1	3.74	1.51	2.16	0.76
10^{-1}	1.87	4.68	3.53	6.95
10^{-2}	0.21	--	0.40	1.18
10^{-3}	<0.05	--	<0.05	(0.14)

-- Not detectable values; () Inaccurate values.

4 ANALYSES WITH THE SEISMIC ACTION

4.1 General description

The behaviours of poroelastic and non-porous layers are compared for equivalent viscous damping conditions. For that purpose, damping ratios previously obtained for poroelastic layers (Table 4) are considered for the calculation of Rayleigh damping parameters, a_0 and a_1 used in non-porous materials. These Rayleigh parameters are calculated for the modal frequencies f_1 and f_2 presented in limit coupled cases (Table 3). Limit case frequencies, and therefore dry or saturated non-porous properties, are chosen according to proximity of frequency values observed in poroelastic layer which is intended to be replicated. Table 5 presents the designations and main characteristics of the carried analyses with poroelastic and non-porous materials.

Table 4: Characteristics of the comparative analyses with seismic action.

G (MPa)	Poroelastic		Non-porous			
	Analysis	k (m/s)	Analysis	Coupled inter. / Material	Rayleigh coefficients	
					a_0	a_1
20	G20_k-5	10^{-5}	G20_sat	Total / Sat.	0	0
	G20_k-1	10^{-1}	G20_sat_R	Total / Sat.	0.05421	3.90×10^{-3}
	G20_k1	10	G20_dry_R	Null / Dry	0.07582	2.59×10^{-5}
80	G80_k-5	10^{-5}	G80_sat	Total / Sat.	0	0
	G80_k-1	10^{-1}	G80_sat_R	Total / Sat.	0.42616	2.77×10^{-3}
	G80_k1	10	G80_dry_R	Null / Dry	0.07387	1.88×10^{-5}

4.1 Responses of layers in free field conditions

The time variation of relative displacements on the top of the layers ($u' = u_{top} - u_{base}$) is presented in Figure 5, for the several analyses. In this figure a good match between displacements in poroelastic and non-porous layers is visible, meaning that poroelastic behaviour can be fairly reproduced with non-porous damped materials (or undamped materials for quasi-limit coupled cases), in pure shear conditions. Also frequency spectra (not shown here) are very similar for both materials in these studied cases. A large amplification of displacements is found for $k = 1 \times 10^{-5}$ m/s analyses, particularly visible for the G80 layer, due to the almost absence of viscous damping in these cases.

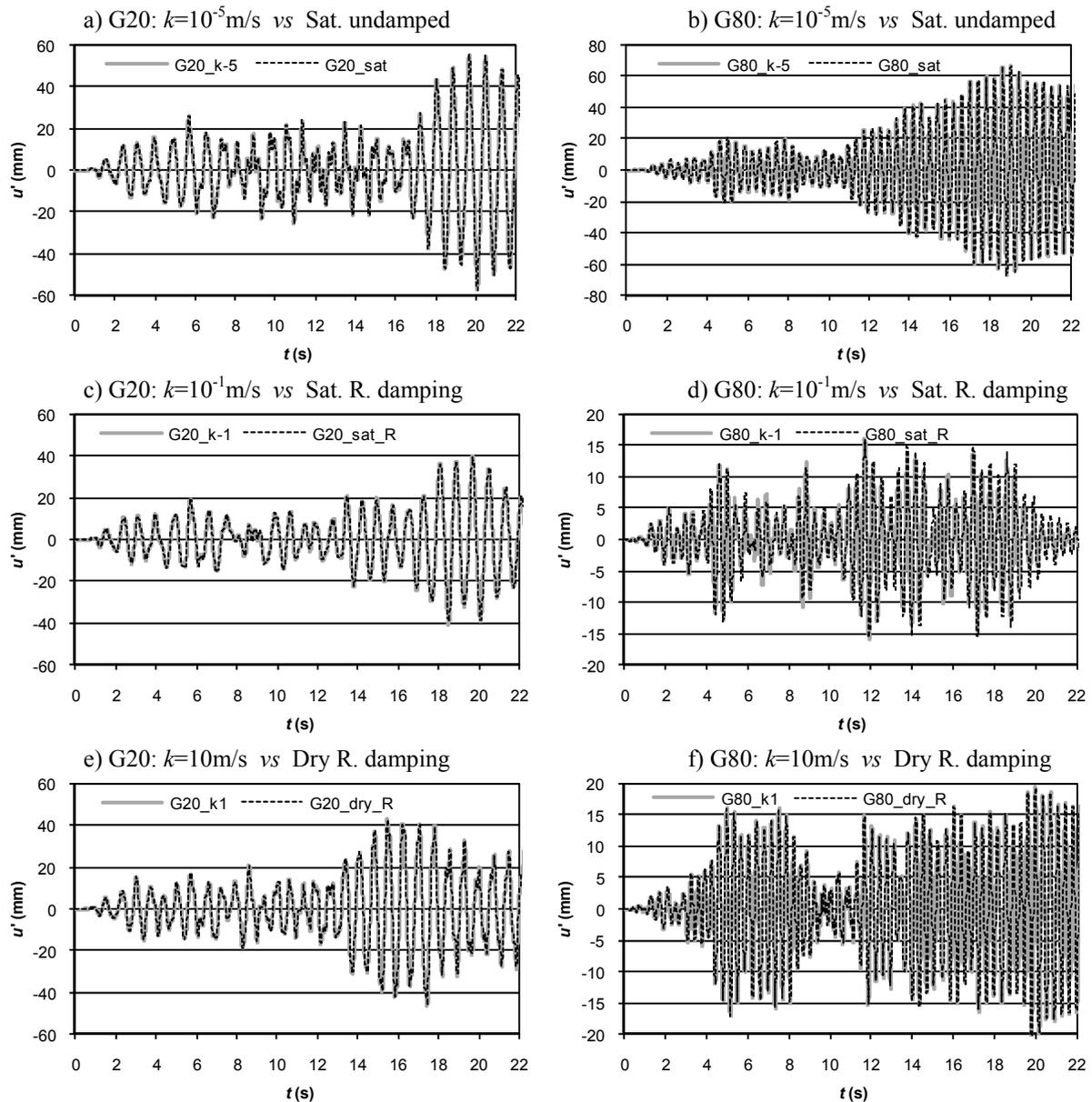


Figure 5: Time variation of displacement amplitudes in free vibration mode for several permeabilities.

4.2 Responses of layers with embedded wall

Some results of the above-mentioned analyses are presented in Figure 6. In this figure, comparisons of relative displacements between poroelastic and non-porous layers are displayed for the top of the layers, above the wall and on the boundary (displacement control points marked in Figure 1c). These results are in agreement with that observed in free field analyses, denoting a good, but not so perfect, match of displacements between both materials. In these figures a higher amplification of responses is observed on the top of the wall, enhanced particularly for the less damped materials as G80_k-5 and G20_k1.

Responses are less amplified at the boundaries, denoting a decreasing influence with the

distance of the wall. Anyway, the boundary displacements have some differences from those of the free field analyses, mainly for the less damped materials, which mean that these boundaries are not sufficiently far from the wall to reproduce the free field.

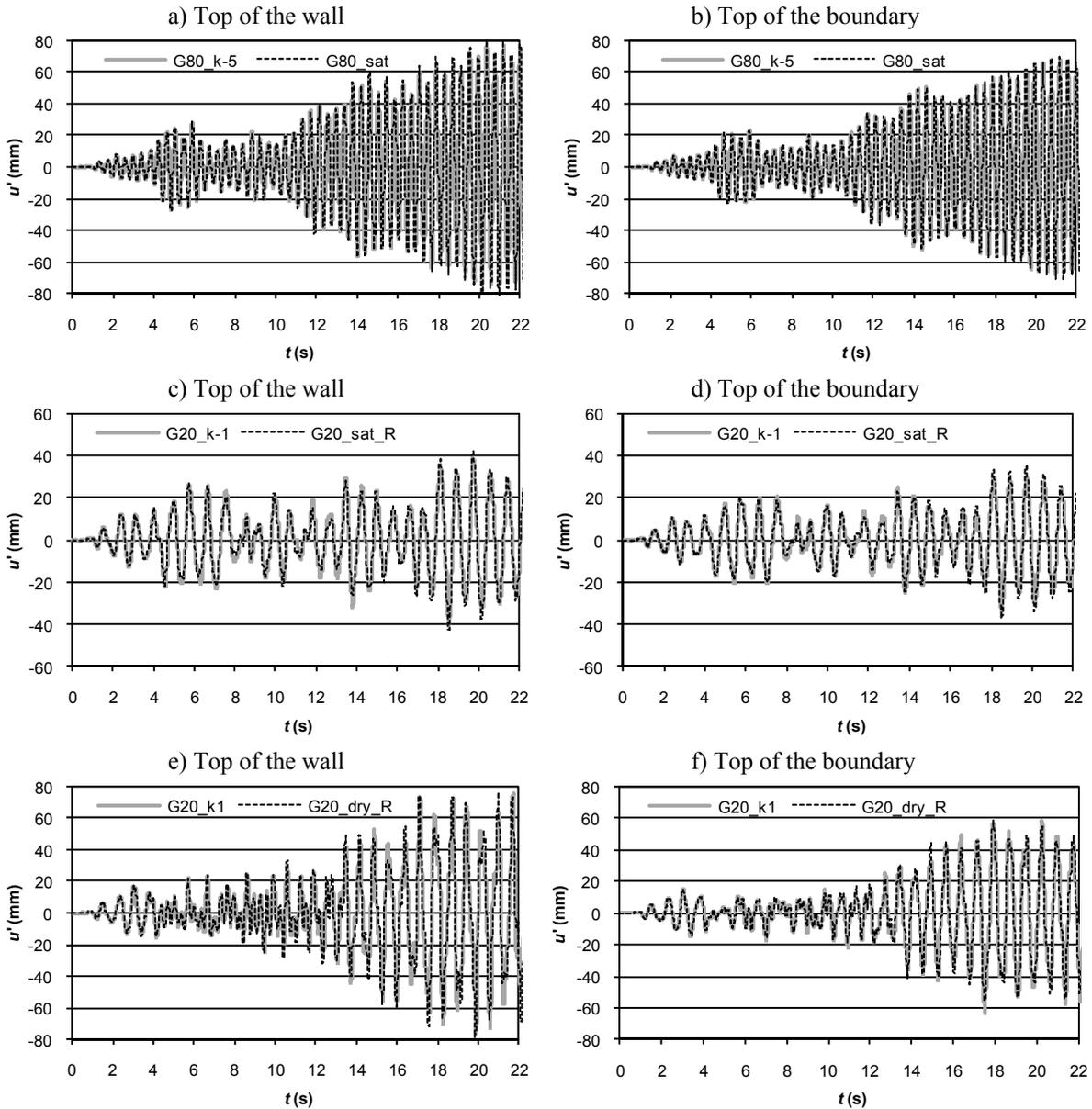


Figure 6: Time variation of displacement amplitudes for the layers with embedded wall.

The volumetric deformations, ε_v , were calculated in some control points (Fig. 1) for several depths and for distances from the wall not less than 2m. The results of the time variation of ε_v are displayed in Figure 7 for the poroelastic layers G20_k-5 and G80_k-1. For both cases, it is evident that results at the boundary points are quite different from those at the inner zone of the layer. This should be due to the slide restrictions imposed at the lateral boundaries layers.

Also, the time variation of ε_v is less uniform from point to point in the layer with higher permeability, due to the fact that for $k=1 \times 10^{-5} \text{ m/s}$ the behaviour is more similar to the undrained case in the whole of the inner zone of the layer.

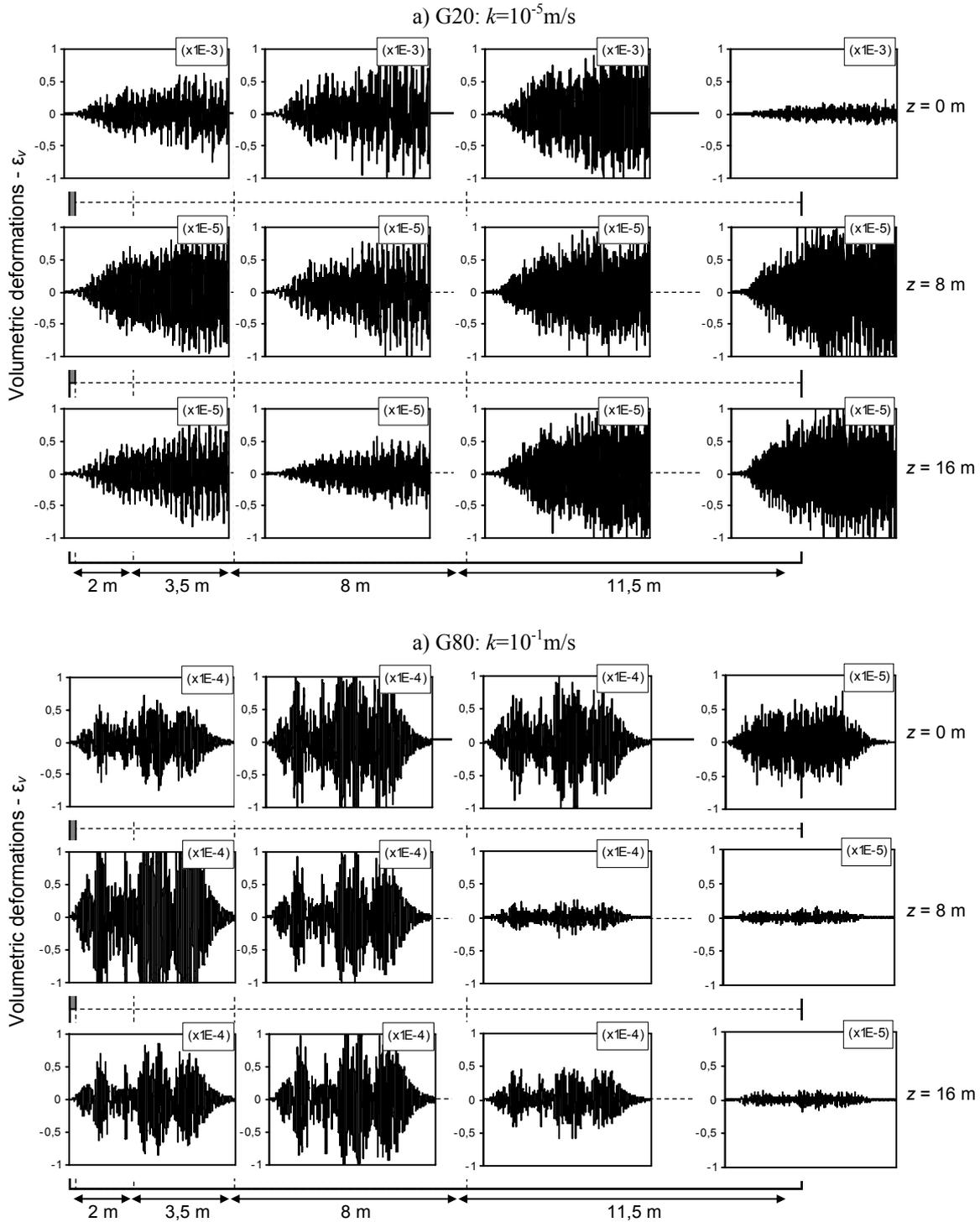


Figure 6: Volumetric deformations in layers G20_k-5 and G80_k-1 with embedded wall.

5 CONCLUSIONS

The results presented in this paper show that non-porous materials with appropriate properties and Rayleigh damping may reproduce almost perfectly pure shear behaviour of poroelastic layers. The good results obtained in the comparative analyses between both materials validate the method used for the determination of Rayleigh coefficients.

Subsequent similar analyses were made for layers with an embedded wall. In these cases, the volumetric deformations induce slight differences between responses of both materials, more noticeable for the less damped cases. Nevertheless, calculations of the responses of layers with the embedded wall reveal fairly rigorous results with non-porous materials, despite the pore fluid pressures presented in poroelastic cases. The use of non-porous materials has the aforementioned benefits in effort and time computation.

For the analyses with the embedded wall in poroelastic materials, a more uniform field of volumetric deformations is observed for permeabilities near limit coupled cases, meaning that over whole inner media has a similar behaviour.

REFERENCES

- [1] Biot, M.A. Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low frequency range. *J. Acoustical Soc. of America* (1956) **28(2)**: 168-178.
- [2] Biot, M.A. Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range. *J. Acoustical Soc. of America* (1956) **28(2)**: 179-191.
- [3] Schanz, M.; Diebels, S. A comparative study of Biot's theory and the linear theory of porous media for wave propagation problems, *Acta Mechanica* (2003) **161**: 213-235.
- [4] Zienkiewicz, O.C.; Chang, C.T.; Bettess, P. Drained, undrained, consolidating and dynamic behaviour assumptions in soils, *Geot echnique* (1980) **30(4)**: 385-395.
- [5] Grazina, J.C.; Pinto, P.L. and Taborda, D. Evaluation of viscous damping due to solid-fluid interaction in a poroelastic layer subjected to shear dynamic actions, NUMGEO (2010), Trondheim, pp. 447-452.
- [6] Grazina, J.C. Modela o din mica com acoplamento viscoso de maci os elastopl sticos. Aplica o a estruturas de suporte flex veis submetidas a ac oes s smicas. *PhD thesis, University of Coimbra* (in Portuguese), (2010).
- [7] Zienkiewicz, O.C.; Chan, A.H.C.; Pastor, M.; Schrefler, B.A. and Shiomi, T. *Computational Geomechanics with Special Reference of Earthquake Engineering*. John Wiley & Sons, Chichester, (1999).
- [8] Arduino, P. and Macari, E.J. Implementation of a porous media formulation for geomaterials. *J. Eng. Mechanics, ASCE*, (2001) **127(2)**, pp. 157-166.
- [9] Gajo A. Influence of viscous coupling in propagation of elastic waves in saturated soil. *J. Geot. Engineering, ASCE*, (1995) **121(GT9)**, pp. 636-644.
- [10] Seisimosignal – “A computer program for signal processing of strong-motion data” [online]. Available from URL: <http://www.seisimosoft.com>.
- [11] Eurocode 8 – NP, EN 1998-1: Design of structures for earthquake resistance, (2010).
- [12] Taborda, D. DFTi – Improved Discrete Fourier Transform Algorithm software: version 1.1.1, (2008).
- [13] Kramer, S.L. *Geotechnical Earthquake Engineering*. Prentice Hall, New Jersey, (1996).

actively secreted by the microbial cells or accumulate in the extracellular space after cell lysis. The EPS matrix serves several functions including cohesiveness between neighboring cells, structural integrity of the entire biofilm, protection from predation, antimicrobials, and others [2]. The ability of certain biofilm-forming microbial species to degrade detrimental (or synthesize useful) chemical species is driving the interest of scientists and engineers in studying and developing techniques based on microbial biofilms. For example, this ability of microbes is observed during the natural attenuation of contaminants dissolved in the groundwater, and is exploited during the treatment of contaminated and waste water in porous medium based bioreactors (e.g., packed and fluidized beds, hollow-fiber membranes).

Porous media are exquisite *hosts* for biofilm-forming bacteria because of their high specific surface [3]. The analysis of biofilm growth in porous media is not trivial since the structure of the system exhibits a *hierarchy of characteristic length scales* that spans several orders of magnitude (from several nanometers in the EPS up to a few hundreds of meters at the aquifer scale – Fig. 1) and, further, there exists an intricate interplay of hydrodynamic, physico-chemical and biological processes occurring at different *characteristic time scales*. In addition, each structural level might be heterogeneous with respect to geometrical and topological characteristics (e.g. pore and grain size, shape and connectivity), physical properties (e.g. fluid density and viscosity), chemical composition (e.g. mineralogy of the solid matrix), as well as biological composition and activity (e.g. number and physiological state of bacterial cells). A generic description of the main processes involved in the formation of biofilms in porous media is given in [4].

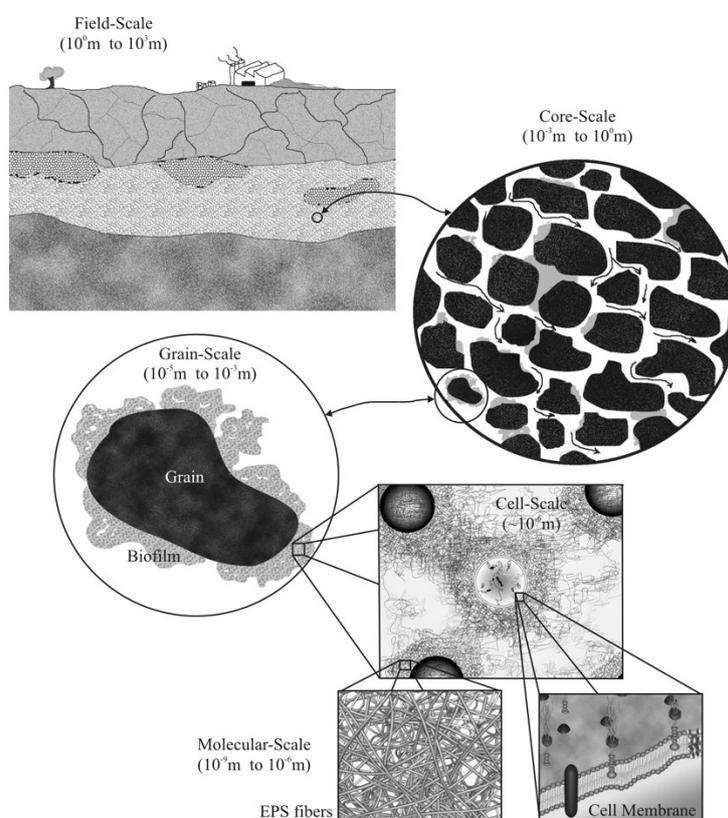


Figure 1: Biofilms in porous media: hierarchical structure of the system (reprinted from [4]).

Mathematical models and simulators are indispensable tools for the study of such complex systems because these can be used to postulate and test hypotheses concerning the underlying mechanisms, to design and interpret experiments, and to optimize the performance of bioreactors. Up to date, significant work has been devoted in the development of models for this process on a single length scale (e.g., biofilm or field scale), neglecting what happens at smaller and/or larger length scales. However, the process under consideration is inherently multiscale, meaning that the different length scales interact strongly to produce the observed behavior. For example, the pressure drop and the residence time distribution of dissolved or particulate matter, which are typically measured at the reactor-scale depend strongly on the detailed spatial distribution of the biofilms at the core-scale.

We have recently developed the first hierarchical, hybrid computer simulator of biofilm growth dynamics in porous media (HiBioSim-PM) [4]. The simulator predicts: i) the structural and biological heterogeneity at the biofilm scale, and ii) the pattern of evolution and the rate of growth of heterogeneous biofilms within the pore space of porous media (core scale). A continuum-based approach for fluid flow and mass transport is combined with individual-based approaches for biofilm growth, detachment, and migration in the pore space. The rationale for the development of this *hybrid* approach is the disparity in characteristic lengths between chemical species and biofilms. Mesoscopic cooperative effects are taken into account explicitly, and the impact of biofilm formation on the spatiotemporal distribution of preferential flowpaths and concentration profiles of dissolved substances is determined. In this work, we present an improved version of our simulator and use it to investigate the impact of biofilm formation on the fate and transport of colloidal particles in a three-dimensional model pore structure.

2 DESCRIPTION OF THE SIMULATOR (HIBIOSIM-PM)

A brief description of the main modules of the simulator is given in the following paragraphs, and more details can be found in [4,5].

2.1 Generation of the pore structure

In the present work, a planar network of three-dimensional pores with rhombohedral topology is used as a model pore structure. This type of porous medium was selected because it models the microfluidic pore networks used in our laboratory for experimental studies of biofilm formation. In general, any type of computer-generated virtual structure, or digitized representations of real porous media can be used in the simulator. In previous works, we have used unconsolidated and consolidated cores of granular porous media [4-6].

2.2 Inoculation of the pore structure

The inoculation sites are determined as follows. First, the flow field within the pore space is determined as described in subsection §2.5. Then, a single microbial cell at a time is inserted randomly at the inflow boundary of the virtual porous medium and moves along the streamlines (§2.6) until it is captured at a grain surface or exits the system. This procedure is repeated until the number of initially attached microbial cells (*first colonists*) equals a prescribed value.

2.3 Biofilm growth

With regard to the biological processes, the biofilm is treated as a population of interacting individual microbial cells, which consume nutrients, grow, proliferate, and synthesize EPS. The growth and proliferation of microbial cells within the biofilm is modeled using a 3D cubic lattice of unit biomass compartments (UBCs). A UBC might contain a single microbial cell (single occupancy UBC) or more than one microbial cells of the same species (multiple occupancy UBC). A state vector is assigned to each UBC. The state vector contains information about the species (for mixed biofilms), the size, and the physiological status (active, dormant or apoptotic) of the microbial cell that occupies the UCB, as well as the volume fraction and the porosity of the EPS hydrogel. The active microbial cells assimilate nutrients and synthesize new cellular mass with rate proportional to their mass. Part of the cellular mass is used for maintenance purposes (endogenous metabolism). Simultaneously, they synthesize and secrete EPS within their UBC with rate proportional to their growth rate. Part of the EPS matrix lyses (say due to enzymatic or hydrolytic action) with rate proportional to its mass. Based on these assumptions the mass balances for bacterial cells and EPS within a UBC are

$$\frac{dX_{\kappa}}{dt} = \mu_{\kappa,g} X_{\kappa} - \mu_{\kappa,m} X_{\kappa} \quad (1)$$

$$\frac{dX_{\pi}}{dt} = Y_{\pi/\kappa} \mu_{\kappa,g} X_{\kappa} - k_{lys} X_{\pi} \quad (2)$$

where X_{κ} , X_{π} denote the mass of cells, EPS over the volume of a UBC, and $\mu_{\kappa,g}$, $\mu_{\kappa,m}$, $Y_{\pi/\kappa}$, k_{lys} are kinetic parameters, which might be defined as functions of the local environmental conditions (nutrient concentrations, mechanical stresses, temperature, pH, etc.). For the simulation results presented in this work we used the following expression for the local specific growth rate of cells

$$\mu_{\kappa,g} = F_n(C_n) F_{\sigma}(I_{\beta}) \quad (3)$$

$$F_n(C_n) = \mu_{max} \frac{C_n}{K_C + C_n} \quad (4)$$

$$F_{\sigma}(I_{\beta}) = \begin{cases} 1 - I_{\beta}/I_{\beta,crit} & \text{if } I_{\beta} \leq I_{\beta,crit} \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Here, C_n is the local concentration of a growth-limiting nutrient, I_{β} is the first invariant of the local stress tensor for the solid components of the biofilm, and μ_{max} , K_C , $I_{\beta,crit}$ are kinetic parameters.

If the specific growth rate is greater than a critical value, $\mu_{\kappa,g,crit}$, the active bacterial cells continuously increase their mass until it exceeds a prescribed upper threshold value, $m_{\kappa,crit}^+$. Then they divide into two equal daughter cells (the number of cells within the UBC doubles and the mass of each cell halves). If the specific growth rate gets lower than the critical value, the cells enter the dormancy state during which the metabolic activity is halted and only consumption of cellular mass for maintenance purposes takes place. Dormant cells may be activated once again if the specific growth rate is restored to a value greater than the critical.

During dormancy, cellular mass decreases continuously until a prescribed lower threshold value, $m_{\kappa, crit}^-$, is reached. Then the cells enter the apoptosis state (programmed cell death), which is irreversible and lysis of the cells occurs with probability

$$p_{ap}(\tau_{ap}) = 1 - \exp(-k_{ap} \tau_{ap}) \quad (6)$$

where τ_{ap} is the time since the cells became apoptotic and k_{ap} is the apoptosis rate constant.

2.4 Cell proliferation and EPS spreading

The microbial cells move within the biofilm because of the internal stresses that develop during the division of cells. In this work, the cellular motion that is caused by cell divisions is modeled on the 3D cubic lattice of UBCs, by implementing a least-action principle. In particular, if the mass of a cell exceeds a corresponding, prescribed maximum value then *biomass percolates* from the UBC, which is occupied by the cell under division, towards the nearest empty-UBC, which contains only fluid and/or EPS hydrogel, along the *path that minimizes the energy of displacement*. First, a random walk procedure is used to generate a large number of paths that connect the UBC, which is occupied by the cell under division, and the nearest empty-UBC, without passing over solid obstacles. Then, the shortest of them is chosen. If there are more than one paths with the same minimum length, then the path that corresponds to the minimum EPS content is chosen. Afterwards, the overgrown UBC displaces its adjacent UBC (defined by the path) and thus generates a temporarily empty-UBC in which it puts half of its biomass, while it retains the other half. The displaced UBC in turn displaces its adjacent UBC and takes its position. This sequence of interactions continues until the end position of the path.

2.5 Fluid flow and fluid-biofilm interaction

With regard to momentum and mass transport processes, the biofilm is treated as a biphasic *continuum* with poroelastic material behavior. Single-phase flow of an incompressible Newtonian fluid (say a dilute aqueous solution) is considered within the pore space of the porous medium, which is occupied partly by fluid and partly by biofilms. Within the fluid regions, the Navier-Stokes equations and the continuity equation result from the formulation of the momentum and mass balance, respectively. Within the porous biofilm, Brinkman's extension of Darcy's law is considered as an appropriate equation to describe the flow along with the conservation of total fluid mass. Further, linear elastic behavior is considered for the solid components of the biofilm [7]. The final equations are

$$\nabla \cdot \mathbf{v}_f = 0 \quad (7)$$

$$\rho_f \frac{\partial \mathbf{v}_f}{\partial t} + \alpha_c \rho_f \nabla \cdot (\mathbf{v}_f \mathbf{v}_f / \varepsilon_\beta) = -\varepsilon_\beta \nabla P_f + \mu_f \nabla^2 \mathbf{v}_f - (1 - \alpha_c) \varepsilon_\beta \frac{\mu_f}{k_\beta} \mathbf{v}_f \quad (8)$$

$$\mathbf{0} = \nabla \cdot \boldsymbol{\sigma}_s + \mathbf{F}_{f \rightarrow s} \quad (9)$$

$$\boldsymbol{\sigma}_s = -(1 - \varepsilon_\beta) P_f + \lambda_s (\nabla \cdot \mathbf{u}_s) + \mu_s [\nabla \mathbf{u}_s + \mathbf{u}_s \nabla] \quad (10)$$

$$\mathbf{F}_{f \rightarrow s} = \varepsilon_\beta \frac{\mu_f}{k_\beta} \mathbf{v}_f \quad (11)$$

Here, \mathbf{v}_f is the local superficial velocity of the fluid, P_f is the intrinsic pressure of the fluid, \mathbf{u}_s is the local displacement of the solids in the biofilm, μ_s , λ_s are the Lamé parameters for the solid, μ_f is the fluid viscosity, ρ_f is the fluid density, ε_β is the local volume fraction of fluid, k_β is the local hydraulic permeability (defined only within the regions of porous biofilms) and α_c is a computational parameter that equals unity within regions of fluid and zero within biofilms. The local hydraulic permeability of the biofilm is calculated as a function of the volume fractions of cells, EPS and water, the average diameter of cells, the average diameter of EPS fibers, and the internal porosity of the EPS (see Appendix in [4]).

Equations (7) and (8) are valid everywhere in the pore space (fluid and biofilm regions), while equation (9) applies only in the biofilm regions. These equations are solved numerically by combining finite difference and finite element methods as follows. First, eqs.-(7),(8) are solved numerically using a staggered grid for the spatial discretization, central finite differences for the viscous and pressure terms and a higher-order upwinding scheme for the inertial terms. Then, eq.-(9) is solved using the Galerkin finite element method on a structured mesh of hexahedral elements with C0-quadratic interpolation functions for the displacement.

2.6 Detachment, migration, and reattachment

If a UBC is adjacent to clear fluid, the average shear stress acting on the surfaces exposed to fluid is calculated. Then, if the exerted shear stress exceeds a designated critical value, the UBC is considered to loose the cohesiveness with adjacent UBCs or solid surfaces. Afterward, the UBC begins to move along the fluid streamlines as if a fluid element (in a first approximation the effects of gravity and drag forces are neglected based on the facts that biofilm is highly porous and its density is very close to that of the aqueous solution). The trajectory of the UBC within the pore space is calculated from the numerical integration of

$$\frac{d\mathbf{r}_p}{dt} = \mathbf{v}_f \quad (12)$$

where \mathbf{r}_p is the position of the mass center of the UBC at time t . The UBC stops moving if it passes over the outflow boundary of the porous medium or if it becomes reattached to grain or biofilm surface, which is exposed to shear stress lower than the critical value. If at least one UBC has been detached, the flow field is updated.

2.7 Solute transport

The fate of a dissolved substance 'A' (nutrient, chemical signal etc.) is determined from the convection-diffusion-reaction equation

$$\frac{\partial}{\partial t}(K_{A,\beta/f} C_A) + \nabla \cdot (\mathbf{v}_f C_A) = \nabla \cdot [\mathbf{D}_{A,eff} \cdot \nabla C_A] + R_A \quad (13)$$

where C_A is the concentration, $K_{A,\beta/f}$ is the partition coefficient between the biofilm and the aqueous solution, $\mathbf{D}_{A,eff}$ is the local effective diffusivity and R_A is the local reaction rate of the dissolved substance A. The diffusion coefficient in the biofilm is calculated using the recently developed model in [8]. Equation (7) is solved using a fractional step method, in which the solution procedure is split up into independent steps corresponding to the

convection, diffusion and reaction processes and each step is solved independently. An explicit in time, higher-order upwinding scheme is used for the convective terms, implicit central differencing is used for the diffusive terms and the explicit fourth order Runge-Kutta method is used for the reaction terms.

3 RESULTS AND DISCUSSION

Here, we present sample simulation results of biofilm growth in a pore network with rhombohedral topology. The values of the physicochemical, biological and operational parameters used in this simulation are the same with those used in [4]. Figure 2 shows snapshots of the spatiotemporal evolution of biofilms within the pore space, at various points in time. Initially, small biofilms are formed at the sites of inoculation. Thereafter, the initial biofilms grow further and, also, new colonies appear downstream. These new colonies are formed by the re-attachment of cells that detached from biofilms near the inflow boundary. By the time, the biofilms grow, spread, merge, and cover completely the surface of pores. Under conditions of constant flow rate through the pore network, the reduction of the clear pore sections results in increased fluid velocities which, in turn, cause increased shear stresses at the fluid-biofilm interface. High shear stresses cause continuous erosion of the biofilm surface and, thus, maintain the pores unplugged (in this simulation).

The impact of biofilm formation on the fate and transport of colloidal particles is determined via the following computer experiment. For a given configuration of biofilms, we perform particle tracking for a large number, typically one million, of particles. As a first approximation, we considered that the drag force, which is exerted by the fluid on the particle, is dominant in the force balance and, thus, the trajectory of each particle can be determined from the numerical solution of Eq.-(12). Every particle enters at a random position of the inflow boundary and moves until either it passes over the outflow boundary, or it is captured by interception at the fluid-solid or fluid-biofilm interface. Figure 3 shows the residence time distributions of particles convected out of the pore network for six different spatial configurations of biofilms, and for two different scenarios regarding the flow through the biofilms: a) impermeable biofilms ($k_{\beta}=0$), and b) permeable biofilms ($k_{\beta}=10\mu\text{m}^2$). Further, Fig.-4A shows the effect of the biofilm volume fraction on the average residence time of particles, which are convected out of the pore network. Two important observations can be made. First, as the amount of biofilms increases, the residence time distribution changes form, narrows, and moves to lower values of particle residence time. This trend was intuitively expected because the fluid velocity in the pores increases as the clear pore section decreases. Second, for a given configuration of biofilms, even imperceptible flow through the biofilms results in decreased fluid velocities in the clear pore regions and, thus, in longer residence times for the particles.

A very interesting behavior is observed with regard to the effect of the biofilm volume fraction on the removal efficiency of particles, which is shown in Fig.-4B. As the amount of biofilm increases up to a certain value, the percentage of captured particles increases monotonically. Beyond that value, further increase in the biofilm volume fraction doesn't affect the particle removal efficiency substantially. This behavior is explained as follows. The initial formation of irregular biofilm colonies increases the available area for particle deposition and, thus, results in increased particle removal efficiency. However, at later stages

of the process, the available area for particle deposition decreases again because the biofilms cover completely the surface of the pores and, further, the fluid-biofilm interface becomes smooth by the action of increased shear stress. Nonetheless, the particle removal efficiency doesn't undergo substantial change because the effect of decreased surface area is counterbalanced by the decrease in the cross section of the clear pore, which results in increased probability for a particle to obtain a limiting trajectory that leads to capture.

In Fig.-4B, we observe also that more permeable biofilms produce significantly higher removal efficiencies of particles. This is explained as follows. Higher biofilm permeability results in increased slip velocity at the fluid-biofilm interface, which allows more particles to approach closer to the interface and, thus, be captured.

4 CONCLUSIONS

- The re-attachment of detached cells and biofilm fragments enhances significantly the downstream migration of biofilms in the porous medium. This is the only computer simulator that accounts explicitly for the fate of detached biofilm fragments. In *all* previous models and simulators of biofilm formation, the detached material is merely treated as “lost”.
- The formation of biofilms alters the geometrical and topological characteristics of the pore structure and the flow field, which in turn strongly affect the fate of moving particles within the porous medium. In particular, biofilm formation results in increased removal efficiency and decreased residence time of particles in the porous medium.

REFERENCES

- [1] Costerton, J.W., Lewandowski, Z., de Beer, D., Caldwell, D., Korber, D. and James, G. Biofilms, the customized microniche. *J. Bacteriol.* (1994)176:2137-2142.
- [2] Rittmann, B.E. The significance of biofilms in porous media. *Water Resour. Res.* (1993) 29(7):2195-202.
- [3] Wolfaardt, G.M., Lawrence, J.R. and Korber D.R. “Function of EPS.” In: Wingender, J., Neu, T.R. and Flemming H.-C. (Eds.) *Microbial Extracellular Polymeric Substances: characterization, structure, and function.* Springer-Verlag (1999) pp.171-200.
- [4] Kapellos, G.E., Alexiou, T.S. and Payatakes A.C. Hierarchical simulator of biofilm growth and dynamics in granular porous materials. *Adv. Water Resour.* (2007)30:1648-67.
- [5] Kapellos G.E. Transport phenomena and dynamics of microbial biofilm growth during the biodegradation of organic pollutants in porous materials: hierarchical theoretical modeling and experimental investigation. Doctoral thesis in Greek. (2008) <http://nemertes.lis.upatras.gr/dspace/handle/123456789/1693>
- [6] Kapellos G.E., Alexiou T.S., Pavlou S. and Payatakes A.C. Hierarchical simulation of biofilm growth dynamics in porous media. In: Starrett, S.K. et al. (Eds.) *Environmental Science and Technology*, American Science Press, Houston, USA, Vol. 1, (2007) 491-502.
- [7] Kapellos, G.E., Alexiou, T.S. and Payatakes A.C. Theoretical modeling of fluid flow through cellular biological media: An overview. *Math. Biosci.* 225(2):83-93 (2010).
- [8] Kapellos, G.E., Alexiou, T.S. and Payatakes A.C. A multiscale theoretical model for diffusive mass transfer in cellular biological media. *Math. Biosci.* (2007) 210(1):177-237.

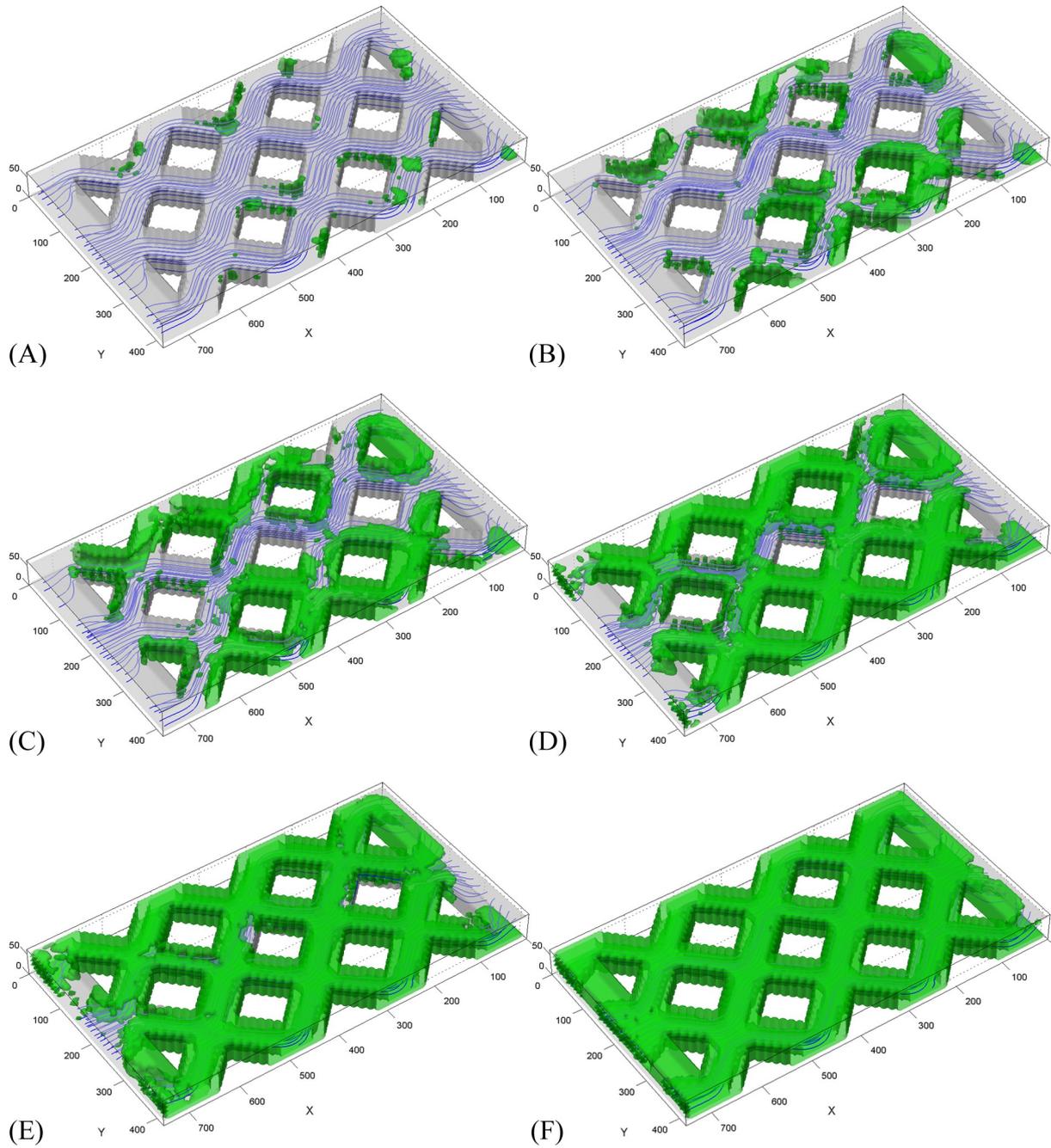


Figure 2: Spatial distribution of biofilms within the pore space at various time instants: (A) $t^*=7$, (B) $t^*=12$, (C) $t^*=15.5$, (D) $t^*=20.5$, (E) $t^*=24.25$, (F) $t^*=30$, where $t^*=\mu_{max}t$ and μ_{max} is the specific growth rate of cells. The blue lines are equidistant streamlines. The biofilms (green color) and the pore space (gray color) are semi-transparent for better visualization.

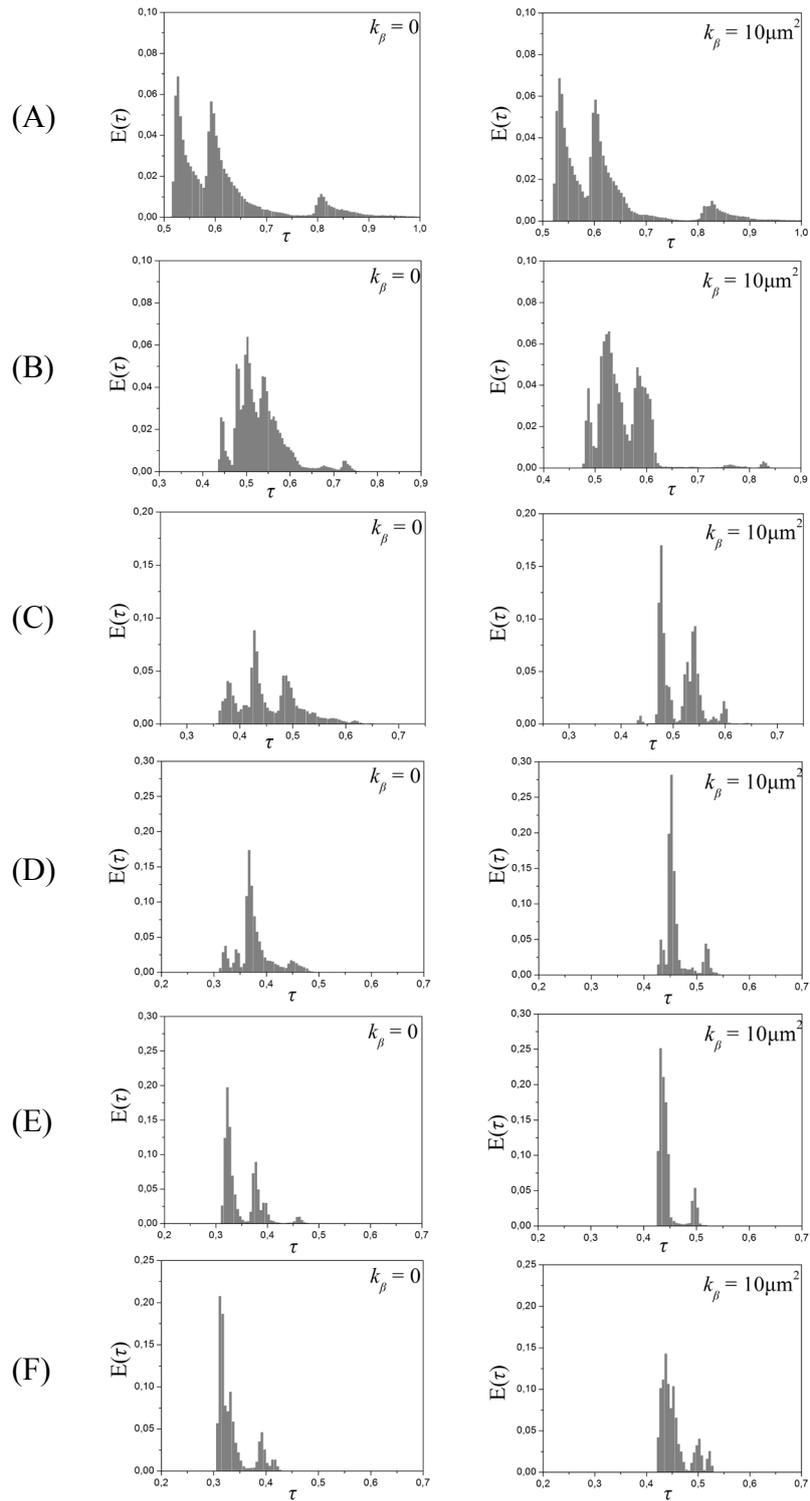


Figure 3: Residence time distributions of particles convected out of the pore network for two scenarios: impermeable biofilms (left column, $k_\beta=0$), and permeable biofilms (right column, $k_\beta=10\mu\text{m}^2$). The capital letters on the outer left denote the corresponding biofilm distributions shown in Figure 2.

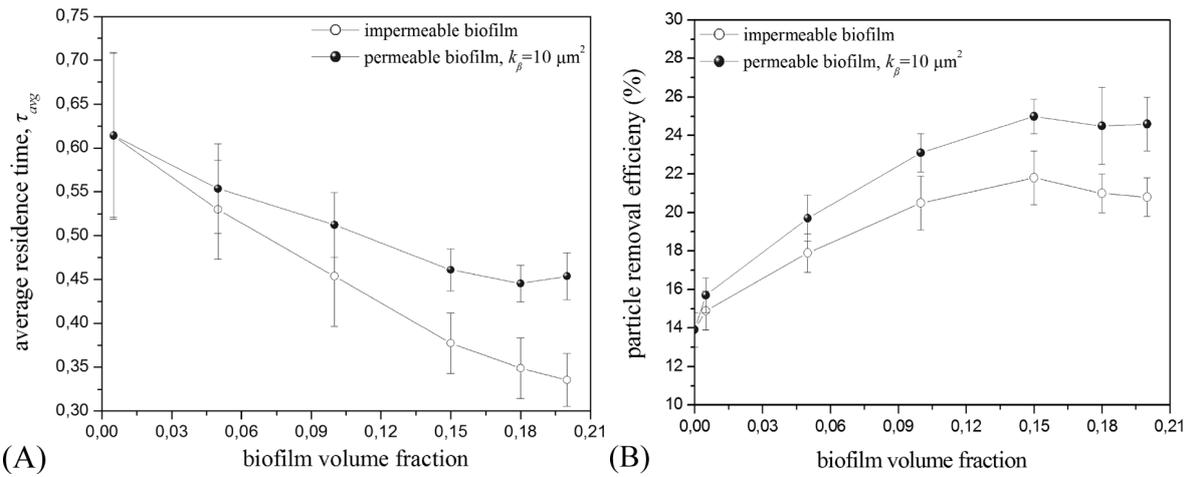


Figure 4: Impact of the biofilm volume fraction on: (A) the average residence time of particles convected out of the pore network, and (B) the efficiency of particle removal (i.e., percentage of captured particles).

NON-LINEAR DYNAMIC SOIL RESPONSE UNDERNEATH A VERTICAL BREAKWATER SUBJECTED TO IMPULSIVE SEA WAVE ACTIONS

M. M. STICKLE^{*}, P. DE LA FUENTE^{*} AND C. OTEO^{**}

^{*} Continuum Mechanics and Structures Department
Universidad Politécnica de Madrid
Avd. Profesor Aranguren s/n, 28040 Madrid, Spain
e-mail: miguelstickle@caminos.upm.es, pdelaf@caminos.upm.es

^{**} Professor on Ground Eng.
C / Torpedero Tucumán 26, 28016 Madrid, Spain
e-mail: carlosoteo@telefonica.net

Key words: Breakwater, Dynamic, Coupled Problems, Interaction, ADÍNDICA

Summary *A theoretical model of the soil-water-structure interaction involved in a breakwater structure subjected to sea wave actions is presented. The model includes i) soil skeleton-pore fluid interaction governed by the $u - p_w$ Generalized Biot formulation [1] including dynamic effects, ii) non-linear soil elastoplastic behaviour described by a novel Generalized Plasticity model [2] coupled with a conservative hyperelastic formulation for the dependence of the elastic stiffness on the stress [3], iii) coupling between the caisson and foundation through a non-linear contact with geometrical compatible formulation incorporating frictional behaviour.*

The numerical solution of the settled governing equations has been fully developed through the Finite Element Method. Furthermore, a program called ADÍNDICA has been created in M Matlab language. ADÍNDICA is a Spanish acronym for “Caisson Breakwater Dynamic Analysis”. Related numerical analyses are developed with reference to precise boundary value problems of specific physical nature. ADÍNDICA code has been able to reproduce adequately the principal characteristics of the caisson oscillations and instantaneous pore pressure generation relation deduced experimentally. Moreover, ADÍNDICA has been able to reproduce satisfactorily the accumulative settlement behaviour of a vertical breakwater structure subjected to series of sea wave impacts including the correlation between accumulated settlements and residual pore pressure.

1 INTRODUCTION

The design of the foundation of marine structures presents a series of difficulties due to the complexity of the forces exerted over the structure, derived from the dynamic swell action and transmitted to the seabed through a complex foundation-structure interaction, as well as the nonlinear soil behaviour, where there is a coupling between solid skeleton and pore water.

These difficulties make the dynamics associated with a seabed underlying a vertical breakwater a uniquely complex task. It seems that the phenomena involved in these dynamics cannot be reproduced with a single model. Therefore it is necessary to couple a series of

models to adequately reproduce each of the determinant aspects concerned, for which the use of numerical techniques becomes indispensable.

Although the phenomenon of wave-induced seabed instability has received great attention among coastal geotechnical engineers since the 80s of last centuries, most of the developed researches [4, 5] have modelled the seabed soil skeleton-pore fluid interaction through the pseudostatic Biot Formulation. This theory, even if it is the base of most of subsequent developments, does not include dynamic terms. However several researches [6, 7] have shown the significant relation of these terms with the wave induced effective stress development.

Most soil models used in the investigations of sea floor dynamics have been limited to the poroelastic model. Only a few contributions [8] have incorporated advanced constitutive relations that are able to represent properly the features of soil response under cyclic loading. This is a prominent aspect within any model proposed to analyze the geomechanical behaviour associated with a breakwater foundation as is needed to investigate the possible degradation process, i.e. the change of the strength and stiffness of the soil with time, mainly due to repetitive loading. The theoretical model for the soil-water-structure interaction presented in this paper includes an advanced sand constitutive model sensitive to cyclic loads.

The caisson-rubble mound interaction phenomenon, responsible of the principal loads transmitted to the foundation, has been investigated mostly through elastic Mass-Spring-Dashpot models [9, 10] where the caisson structure is modelled as a point mass. Therefore, these models are not able to analyze the different interface strain-stress states involved in this contact surface. Few researches have employed frictional contact mechanics of deformable bodies to represent this interaction phenomenon not analyzing geomechanical implication.

In the next chapter the proposed theoretical model for the soil-water-structure interaction involved in a breakwater structure subjected to sea wave actions is presented. Afterwards the Finite Element numerical solution is outlined leading to some related numerical analyses with reference to precise boundary value problems of specific physical nature in order to justify the theoretical model and its numerical approach.

The sign convention of compressive stresses and strains negative while compressive pore pressure positive is assumed.

2 SOIL-WATER-STRUCTURE INTERACTION THEORETICAL MODEL.

The soil-water-breakwater interaction has been modelled coupling different physical systems, therefore independent solution of each system being impossible without simultaneous solution of the others.

The physical systems involved in the soil-water-breakwater interaction analysis are the caisson, the rubble mound and the sea bed (Figure 1). The coupling among these three occurs on domain interface via the boundary conditions imposed there. The rubble mound and the sea bed are already coupled media, where skeleton-pore fluid interaction exists, and the coupling occurs through the governing partial differential equations describing each physical phase.

Sea waves are not modelled as a proper physical system in the proposed theoretical model, representing the sea wave actions exerted over the structure as boundary conditions.

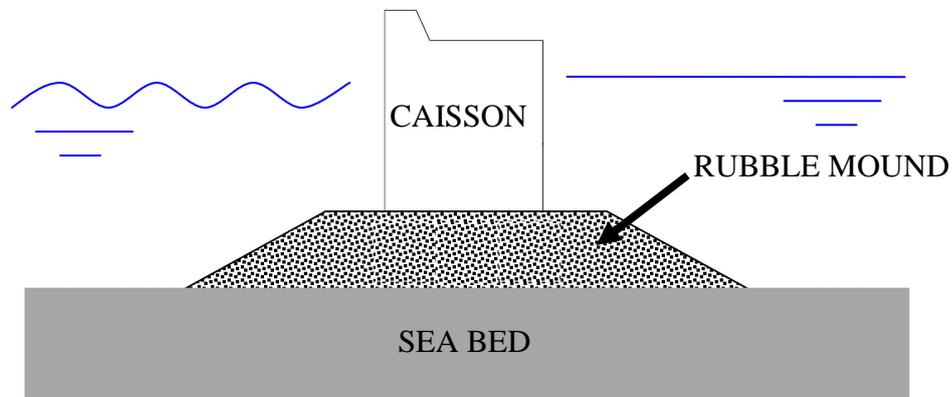


Figure 1: Physical systems involved in the soil-water-breakwater interaction model

The theoretical model for the soil-water-breakwater interaction proposed is developed in two dimensions under plain strain idealization.

Once the sea bed, rubble mound and caisson governing equations are derived, including the couplings involved as well as the initial and boundary conditions, the theoretical model for the soil-water-breakwater interaction proposed will be set.

In (Figure 2), the main parts of the theoretical model proposed in this paper in order to analyze the complex clay like-rubble mound-caisson-swell interaction are schematically shown. The novel theoretical contributions appear in this figure over a dark colour box.

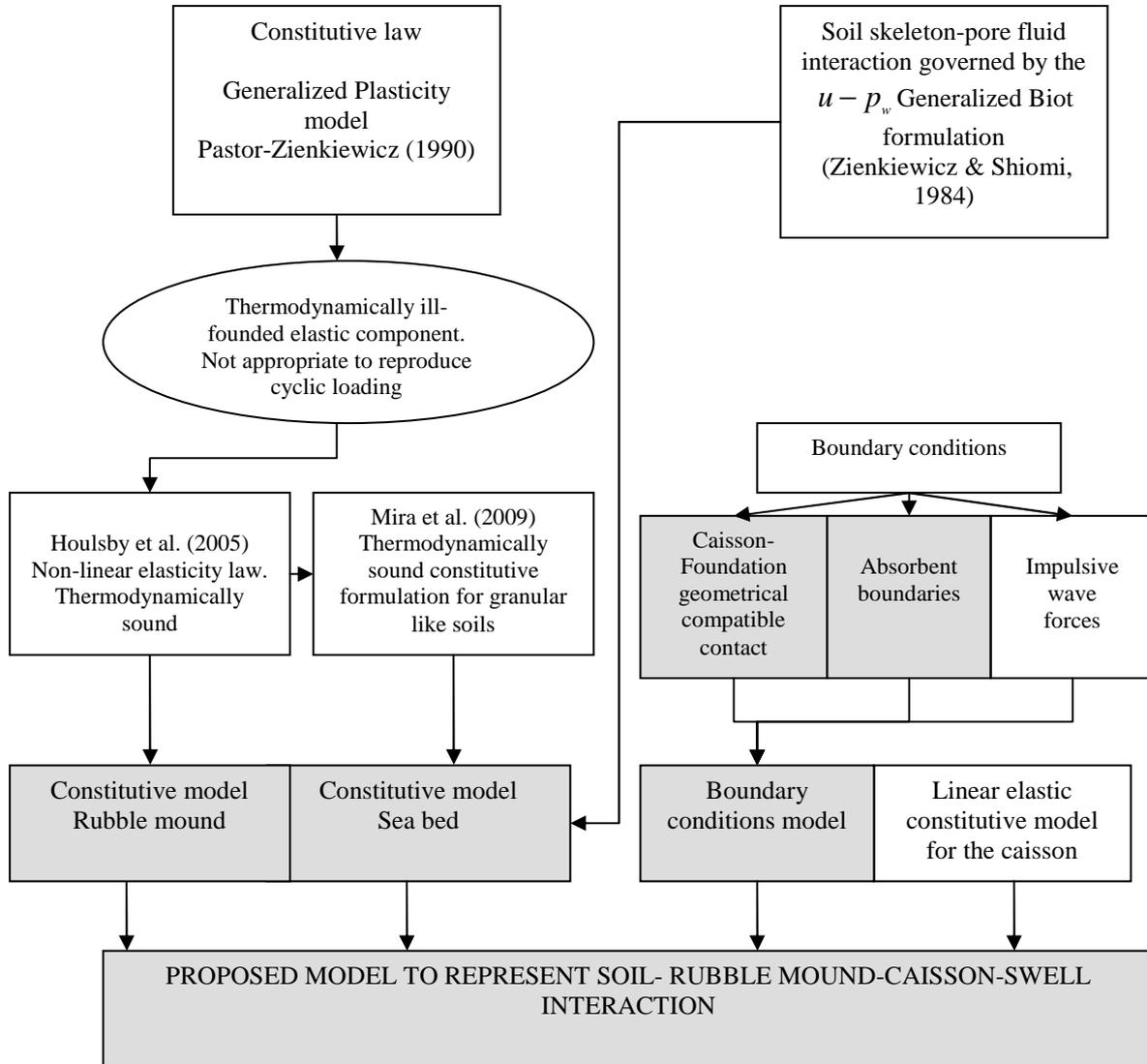


Figure 2: Outline of the theoretical model proposed.

3 FINITE ELEMENT SOLUTION OF THE SETTLED GOVERNING EQUATIONS

Once the kinematic relations as well as the constitutive laws are integrated in the balance equations, a system of five partial differential equations with five field variables is established. The field variables involved are: sea bed skeleton displacement \mathbf{u}^{sb} , pore water pressure p_w^{sb} , rubble mound skeleton displacement \mathbf{u}^{rm} , pore water pressure p_w^{rm} and caisson displacement \mathbf{u}^{ca} .

The system of partial differential equations can be discretized using standard Galerkin techniques, as described in [11]. After spatial discretization, $\mathbf{u}^{sb} \cong \mathbf{N}^u \bar{\mathbf{u}}^{sb}$, $p_w^{sb} \cong \mathbf{N}^p \bar{p}_w^{sb}$, $\mathbf{u}^{rm} \cong \mathbf{N}^u \bar{\mathbf{u}}^{rm}$, $p_w^{rm} \cong \mathbf{N}^p \bar{p}_w^{rm}$, $\mathbf{u}^{ca} \cong \mathbf{N}^u \bar{\mathbf{u}}^{ca}$ the second order ordinary

differential equation system (1)-(3) is obtained

$$\mathbf{M}^{sb} \ddot{\mathbf{u}}^{sb} + \mathbf{C}^{sb} \dot{\mathbf{u}}^{sb} + \int_{\Omega^{sb}} \mathbf{B}^T \boldsymbol{\sigma}'^{sb} d\Omega^{sb} - \mathbf{Q}^{sb} \bar{\mathbf{p}}_w^{sb} - \mathbf{f}^{sb1} = \mathbf{0} \quad (1)$$

$$(\mathbf{Q}^{sb})^T \dot{\mathbf{u}}^{sb} + \mathbf{H}^{sb} \bar{\mathbf{p}}_w^{sb} + \mathbf{S}^{sb} \dot{\bar{\mathbf{p}}}_w^{sb} - \mathbf{f}^{sb2} = \mathbf{0}$$

$$\mathbf{M}^{rm} \ddot{\mathbf{u}}^{rm} + \mathbf{C}^{rm} \dot{\mathbf{u}}^{rm} + \int_{\Omega^{rm}} \mathbf{B}^T \boldsymbol{\sigma}'^{rm} d\Omega^{rm} - \mathbf{Q}^{rm} \bar{\mathbf{p}}_w^{rm} - \mathbf{f}^{rm1} = \mathbf{0} \quad (2)$$

$$(\mathbf{Q}^{rm})^T \dot{\mathbf{u}}^{rm} + \mathbf{H}^{rm} \bar{\mathbf{p}}_w^{rm} + \mathbf{S}^{rm} \dot{\bar{\mathbf{p}}}_w^{rm} - \mathbf{f}^{rm2} = \mathbf{0}$$

$$\mathbf{M}^{ca} \ddot{\mathbf{u}}^{ca} + \mathbf{C}^{ca} \dot{\mathbf{u}}^{ca} + \mathbf{K}^{ca} \bar{\mathbf{u}}^{ca} - \mathbf{f}^{ca} = \mathbf{0} \quad (3)$$

Where $\mathbf{B} = \mathbf{S}\mathbf{N}^u$ and

$$\mathbf{f}^{sb1} = \int_{\Omega^{sb}} (\mathbf{N}^u)^T \boldsymbol{\rho}^{sb} \mathbf{b} d\Omega^{sb} + \int_{\Gamma_t^{sb}} (\mathbf{N}^u)^T \mathbf{t}_{imp}^{sb} d\Gamma_t^{sb} + \mathbf{C}_r^{sb}$$

$$\mathbf{C}_r^{sb} = (\mathbf{R}_1 + \mathbf{R}_2 + \mathbf{R}_3) \dot{\mathbf{u}}^{sb} \quad (4)$$

$$\mathbf{f}^{sb2} = \int_{\Omega^{sb}} (\nabla \mathbf{N}^p)^T \mathbf{k}^{sb} \boldsymbol{\rho}_w^{sb} \mathbf{b} d\Omega^{sb} + \int_{\Gamma_{pw}^{sb}} (\mathbf{N}^p)^T \mathbf{q}_{imp}^{sb} d\Gamma_{pw}^{sb}$$

$$\mathbf{f}^{rm1} = \int_{\Omega^{rm}} (\mathbf{N}^u)^T \boldsymbol{\rho}^{rm} \mathbf{b} d\Omega^{rm} + \int_{\Gamma_t^{rm}} (\mathbf{N}^u)^T \mathbf{t}_{imp}^{rm} d\Gamma_t^{rm} + \mathbf{C}_c^{rm}$$

$$\mathbf{f}^{rm2} = \int_{\Omega^{rm}} (\nabla \mathbf{N}^p)^T \mathbf{k}^{rm} \boldsymbol{\rho}_w^{rm} \mathbf{b} d\Omega^{rm} + \int_{\Gamma_{pw}^{rm}} (\mathbf{N}^p)^T \mathbf{q}_{imp}^{rm} d\Gamma_{pw}^{rm} \quad (5)$$

$$\mathbf{f}^{ca} = \int_{\Omega^{ca}} (\mathbf{N}^u)^T \boldsymbol{\rho}^{ca} \mathbf{b} d\Omega^{ca} + \int_{\Gamma_t^{ca}} (\mathbf{N}^u)^T \mathbf{t}_{imp}^{ca} d\Gamma_t^{ca} + \mathbf{C}_c^{ca} \quad (6)$$

The matrices given in the system (1)-(3) are defined by

$$\mathbf{M}^{sb} = \int_{\Omega^{sb}} (\mathbf{N}^u)^T \boldsymbol{\rho}^{sb} \mathbf{N}^u d\Omega^{sb}, \mathbf{M}^{rm} = \int_{\Omega^{rm}} (\mathbf{N}^u)^T \boldsymbol{\rho}^{rm} \mathbf{N}^u d\Omega^{rm}, \mathbf{M}^{ca} = \int_{\Omega^{ca}} (\mathbf{N}^u)^T \boldsymbol{\rho}^{ca} \mathbf{N}^u d\Omega^{ca} \quad (7)$$

$$\mathbf{Q}^{sb} = \int_{\Omega^{sb}} \mathbf{B}^T \mathbf{m} \mathbf{N}^p d\Omega^{sb}, \mathbf{Q}^{rm} = \int_{\Omega^{rm}} \mathbf{B}^T \mathbf{m} \mathbf{N}^p d\Omega^{rm} \quad (8)$$

$$\mathbf{S}^{sb} = \int_{\Omega^{sb}} (\mathbf{N}^p)^T \frac{1}{Q^{sb}} (\mathbf{N}^p) d\Omega^{sb}, \mathbf{S}^{rm} = \int_{\Omega^{rm}} (\mathbf{N}^p)^T \frac{1}{Q^{rm}} (\mathbf{N}^p) d\Omega^{rm} \quad (9)$$

$$\mathbf{H}^{sb} = \int_{\Omega^{sb}} (\nabla \mathbf{N}^p)^T \frac{\mathbf{k}^{sb}}{\rho_w \cdot g} (\nabla \mathbf{N}^p) d\Omega^{sb}, \mathbf{H}^{rm} = \int_{\Omega^{rm}} (\nabla \mathbf{N}^p)^T \frac{\mathbf{k}^{rm}}{\rho_w \cdot g} (\nabla \mathbf{N}^p) d\Omega^{rm} \quad (10)$$

$$\mathbf{C}^{sb} = \alpha^{sb} \mathbf{M}^{sb} + \beta^{sb} \mathbf{K}^{sb}, \mathbf{C}^{rm} = \alpha^{rm} \mathbf{M}^{rm} + \beta^{rm} \mathbf{K}^{rm}, \mathbf{C}^{ca} = \alpha^{ca} \mathbf{M}^{ca} + \beta^{ca} \mathbf{K}^{ca} \quad (11)$$

$$\mathbf{K}^{sb} = \int_{\Omega^{sb}} \mathbf{B}^T \mathbf{D}^{sb} (\boldsymbol{\sigma}'^{sb}) \mathbf{B} d\Omega^{sb}, \mathbf{K}^{rm} = \int_{\Omega^{rm}} \mathbf{B}^T \mathbf{D}^{rm} (\boldsymbol{\sigma}'^{rm}) \mathbf{B} d\Omega^{rm}, \mathbf{K}^{ca} = \int_{\Omega^{ca}} \mathbf{B}^T \mathbf{D}^{ca} \mathbf{B} d\Omega^{ca} \quad (12)$$

The \mathbf{C}_r^{sb} term in (4) represents the contribution of the radiation boundaries to the discretized governing equations, while \mathbf{C}_c^{rm} and \mathbf{C}_c^{ca} terms, appearing in (5) and (6), respectively, represents the contribution of the rubble mound - caisson contact to the discretized governing equations.

The proper choice of the element type in order to discretize the computational domain is of paramount importance as some elements introduce errors leading to unrealistic limit loads and spurious failure elements [12]. Under Babuska-Brezzi robustness condition, keeping in mind the need of a C^0 interpolation for each field variable, in the present paper a mixed isoparametric lagrangian triangular element has been used, with 6 nodes quadratic interpolation for any skeleton displacement, \mathbf{u}^{sb} , \mathbf{u}^{rm} y \mathbf{u}^{ca} and 3 node linear interpolation for pore water pressure interpolation, p_w^{sb} , p_w^{rm} .

Temporal discretization of the displacements $\bar{\mathbf{u}}^{global} = [\bar{\mathbf{u}}^{sb}, \bar{\mathbf{u}}^{rm}, \bar{\mathbf{u}}^{ca}]^T$ is performed by the Generalized Newmark *GN22* scheme while the excess pore pressure of the sea bed and rubble mound $\bar{\mathbf{p}}_w^{global} = [\bar{\mathbf{p}}_w^{sb}, \bar{\mathbf{p}}_w^{rm}]^T$ are discretized by the *GN11* scheme, leading to the following difference equation systems

$$\begin{aligned} \ddot{\bar{\mathbf{u}}}_{n+1}^{global} &= \ddot{\bar{\mathbf{u}}}_n^{global} + \Delta \ddot{\bar{\mathbf{u}}}_n^{global} \\ \dot{\bar{\mathbf{u}}}_{n+1}^{global} &= \dot{\bar{\mathbf{u}}}_n^{global} + \Delta t \cdot \ddot{\bar{\mathbf{u}}}_n^{global} + \beta_1 \cdot \Delta t \cdot \Delta \ddot{\bar{\mathbf{u}}}_n^{global} \\ \bar{\mathbf{u}}_{n+1}^{global} &= \bar{\mathbf{u}}_n^{global} + \Delta t \cdot \dot{\bar{\mathbf{u}}}_n^{global} + \frac{1}{2} \Delta t^2 \cdot \ddot{\bar{\mathbf{u}}}_n^{global} + \frac{1}{2} \Delta t^2 \cdot \beta_2 \cdot \Delta \ddot{\bar{\mathbf{u}}}_n^{global} \end{aligned} \quad (13)$$

$$\begin{aligned} \dot{\bar{\mathbf{p}}}_{wn+1}^{global} &= \dot{\bar{\mathbf{p}}}_{wn}^{global} + \Delta \dot{\bar{\mathbf{p}}}_{wn}^{global} \\ \bar{\mathbf{p}}_{wn+1}^{global} &= \bar{\mathbf{p}}_{wn}^{global} + \Delta t \cdot \dot{\bar{\mathbf{p}}}_{wn}^{global} + \Delta t \cdot \bar{\beta}_1 \cdot \Delta \dot{\bar{\mathbf{p}}}_{wn}^{global} \end{aligned} \quad (14)$$

After the incorporation of difference equation (13) and (14) in (1)-(3) a non linear algebraic system is obtained where the unknown values are $[\Delta \ddot{\bar{\mathbf{u}}}_n^{sb}, \Delta \dot{\bar{\mathbf{p}}}_{wn}^{sb}, \Delta \ddot{\bar{\mathbf{u}}}_n^{rm}, \Delta \dot{\bar{\mathbf{p}}}_{wn}^{rm}, \Delta \ddot{\bar{\mathbf{u}}}_n^{ca}]$. The Newton-Raphson scheme is used to solved the non linear algebraic in each time step, obtaining the values of the displacements $\bar{\mathbf{u}}_{n+1}^{global}$ and pore water pressure $\bar{\mathbf{p}}_{wn+1}^{global}$ at time t_{n+1} by the difference equations (13) and (14).

4 ADÍNDICA PROGRAM.

The soil-water-structure interaction involved in a breakwater structure subjected to sea wave actions is not restricted to a unique engineering discipline. Taking into account the limitations of commercial codes to deal with this kind of multidisciplinary phenomena, the authors of the present paper have decided to fully develop the numerical solution of the settled governing equations. Furthermore, a program called ADÍNDICA has been created in M Matlab language. ADÍNDICA is a Spanish acronym for “Caisson Breakwater Dynamic Analysis”.

This program can be used to design gravity maritime structures foundations. The users of ADÍNDICA program may be able to analyze the fundamental aspects involved in the geomechanic behaviour associated with the foundation of this kind of structures. These aspects are: *i*) the complex caisson-rubble mound interaction derived from the swell dynamic and cyclic action, allowing an estimation of the stresses transmitted to the seabed, *ii*) the soil skeleton-pore fluid coupling, essential to estimate the pore pressure variation influenced by the elastic compression of the pore fluid (air), by elastic compression and dilatation of the soil skeleton in combination with limited drainage, and *iii*) the possible change of soil strength and stiffness with time mainly due to repetitive loading and /or consolidation, allowing a degradation estimation of the seabed and long term effects. The principal features of ADÍNDICA code [13] are the following:

- Two dimensional finite element program under plane strain and axisymmetric idealization.
- Pre and Postprocessor without leaving Matlab environment. It is also coupled with GID for specific situations.
- Static, consolidation and dynamic problems can be solved.
- Linear and nonlinear problems (material and contact-impact) can be solved.
- Robust and accurate local integration algorithm for elastoplastic constitutive law as well as global load-displacement integration algorithm.
- Point and distributed loads can be defined on the boundary. Both in static and dynamic fashion.
- Stage constructions can be designed

5 RESULTS AND CONCLUSIONS

The large scale model conducted in 2004 by Kudella and Oumeraci [14] in the Large Wave Flume (GWK) of Hannover is numerically reproduced under the scope of the soil-water-structure interaction model proposed in the present paper.

ADÍNDICA code has been able to reproduce adequately the principal characteristics of the caisson oscillations and instantaneous pore pressure generation experimentally deduced (Figure 3, Figure 4, Figure 5). These characteristics are: *i*) the magnitude of the caisson motions induced by impact load at the seaward edge is higher than at the shoreward edge, *ii*) The shape of the p_w records closely follow the shape of the vertical movement records of the caisson edges with the reverse sign. Indeed, the relatively small downward caisson motion at shoreward edge induces a much higher positive pore pressure amplitude than the negative amplitude at seaward edge, where the upward motion is higher and *iii*) The influence of

caisson motions on pore pressure generation decreases with increasing depth.

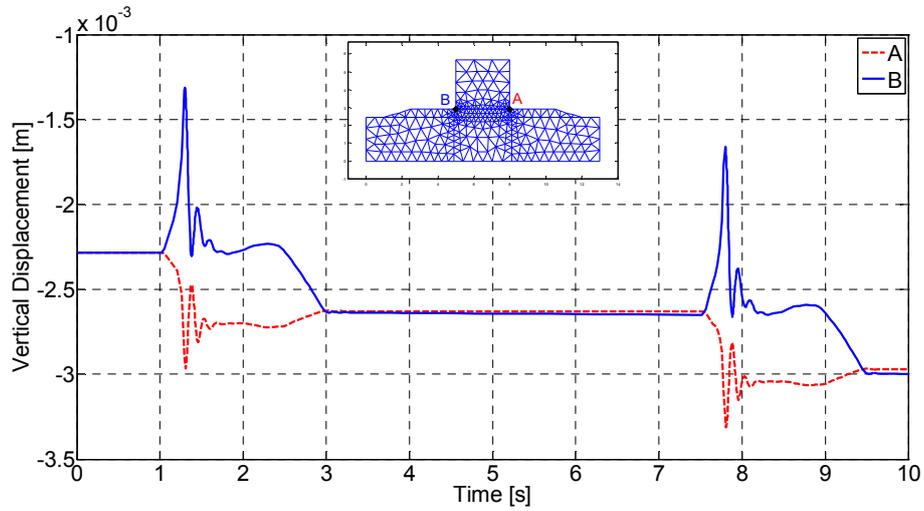


Figure 3: Vertical displacement obtain at caisson edges induced by two impulsive wave actions ($H = 0.6m$, $T = 6.5s$, $h_s = 1.6m$, $h_1 = 0.6m$). ADÍNDICA numerical result.

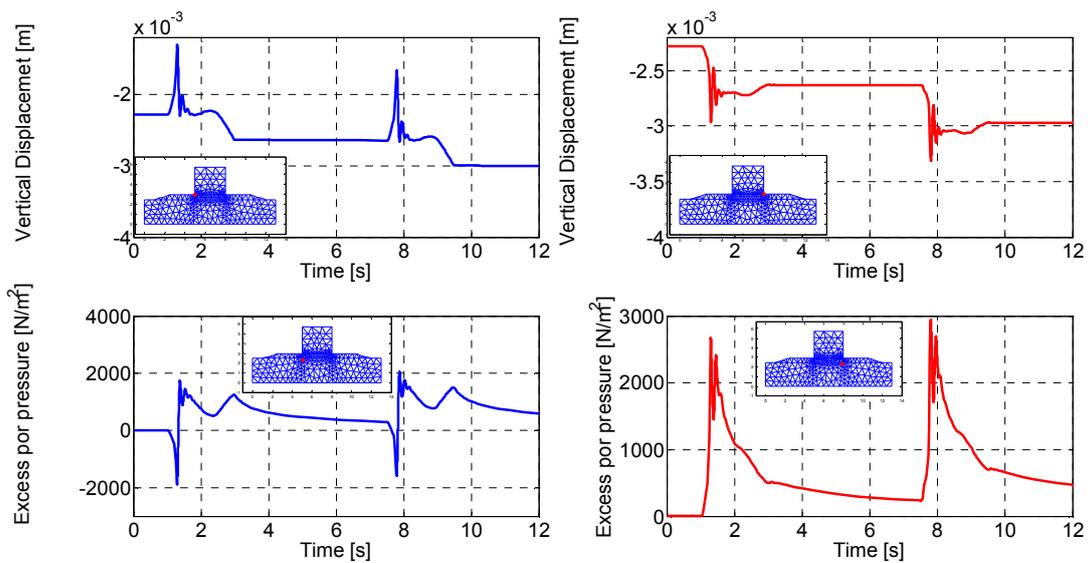


Figure 4: Caisson edges vertical displacement excess pore pressure relation ($H = 0.6m$, $T = 6.5s$, $h_s = 1.6m$, $h_1 = 0.6m$). ADÍNDICA numerical result.

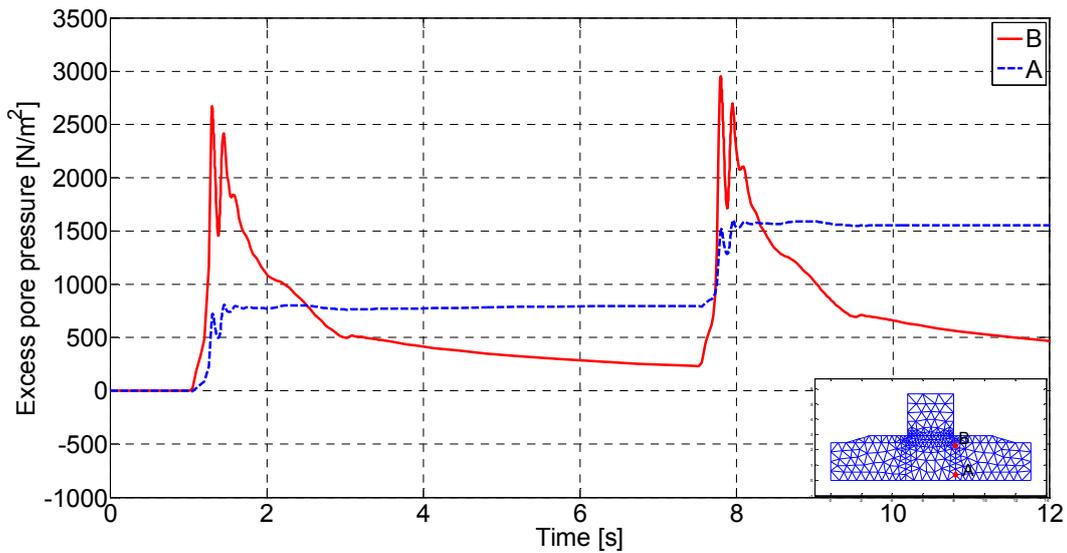


Figure 5: Excess pore pressures obtained at different heights in the sand layer ($H = 0.6m$, $T = 6.5s$, $h_s = 1.6m$, $h_l = 0.6m$). ADÍNDICA numerical result.

ADÍNDICA code is able to reproduce satisfactorily the accumulative settlement behaviour of a vertical breakwater structure subjected to series of sea wave impacts. It is also able to simulate adequately the correlation between accumulated settlements and residual pore pressure (Figure 6).

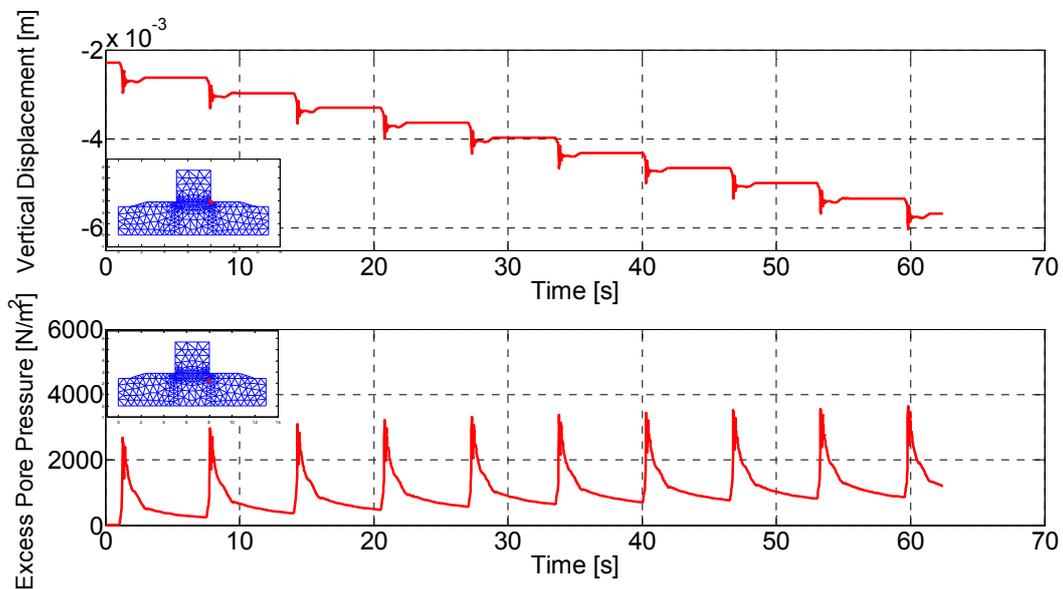


Figure 6: Relation between accumulated settlement and residual pore pressure ($H = 0.6m$, $T = 6.5s$, $h_s = 1.6m$, $h_l = 0.6m$). ADÍNDICA numerical result.

REFERENCES

- [1] Zienkiewicz, O. C., and Shiomi, T., 1984, "Dynamic Behavior of Saturated Porous Media: The Generalized Biot Formulation and its Numerical Solution," *Int. J. Numer. Analyt. Meth. Geomech.*, 8pp. 71-96.
- [2] Pastor, M., Zienkiewicz, O. C., and Chan, A. H. C., 1990, "Generalized Plasticity and the Modeling of Soil Behavior," *Int. J. Numer. Analyt. Meth. Geomech.*, 14pp. 151-190.
- [3] Mira, P., Tonni, L., Pastor, M., 2009, "A Generalized Midpoint Algorithm for the Integration of a Generalized Plasticity Model for Sands," *Int. J. Numer. Methods Eng.*, 77(9) pp. 1201-1223.
- [4] Yamamoto, T., Koning, H. L., Sellmeijer, H., 1978, "On the Response of a Poro-Elastic Bed to Water Waves," *J. Fluid Mech.*, 87pp. 193-206.
- [5] Jeng, D. S., Cha, D. H., Lin, Y. S., 2001, "Wave-Induced Pore Pressure Around a Composite Breakwater," *Ocean Eng.*, 28(10) pp. 1413-1435.
- [6] Jeng, D. S., and Cha, D. H., 2003, "Effects of Dynamic Soil Behavior and Wave Non-Linearity on the Wave-Induced Pore Pressure and Effective Stresses in Porous Seabed," *Ocean Eng.*, 30(16) pp. 2065-2089.
- [7] Ulker, M. B. C., Rahman, M. S., and Guddati, M. N., 2010, "Wave-Induced Dynamic Response and Instability of Seabed Around Breakwater," *Ocean Engineering*, 37pp. 1522-1545.
- [8] Pastor, M., Drenpetic, V., Merodo, J. A. F., 2006, "Una Metodología Racional Para El Estudio De Cimentaciones De Estructuras Marinas," *Ingeniería Civil. CEDEX.*, 141pp. 37-48.
- [9] Goda, Y., 1994, "Dynamic Response of Upright Breakwaters to Impulsive Breaking Wave Forces," *Coastal Engineering*, 22(1-2) pp. 135-158.
- [10] Oumeraci, H., and Kortenhaus, A., 1994b, "Analysis of the Dynamic Response of Caisson Breakwaters," *Coastal Engineering*, 22pp. 159-183.
- [11] Zienkiewicz, O.C., Chan, A.H.C., Pastor, M., 1999, "Computational Geomechanics. With Special Reference to Earthquake Engineering," *JOHN WILEY & SONS*, pp. 383.
- [12] Sloan, W., and Randolph, M. F., 1982, "Numerical Prediction of Collapse Loads using Finite Element Methods," *Int. J. Numer. Analyt. Meth. Geomech.*, 6(1) pp. 47-76.
- [13] Stickle, M. M., 2010, "Sobre La Respuesta Dinámica Del Terreno Bajo La Acción Del Olaje En Cajones Fondeados En Suelos Arcillosos", PhD Tesis, Universidad Politécnica de Madrid.
- [14] Kudella, M., Oumeraci, H., De Groot, M. B., 2006, "Large-Scale Experiments on Pore Pressure Generation Underneath a Caisson Breakwater," *J. Waterw. Port. Coastal Ocean Div. Am. Soc. Civ. Eng.*, 132(4) pp. 310-324.

NUMERICAL MODELING OF SHALLOW LANDSLIDE IMPACTS ON FLEXIBLE PROTECTION SYSTEMS AND ITS VALIDATION WITH FULL SCALE TESTING

Albrecht v. BOETTICHER*, Axel Volkwein*, Roland Wüchner[†], Kai-Uwe
Bletzinger[†], Corinna Wendeler[#]

*Swiss Federal Institute for Forest, Snow & Landscape Research (WSL)
Zürcherstrasse 111
8903 Birmensdorf, Switzerland
e-mail: {albrecht.vonboetticher, volkwein}@wsl.ch, www.wsl.ch/fe/lms

[†]Technical University of Munich (TUM)
80290 Munich, Germany
e-mail: {wuechner, kub}@bv.tum.de, www.st.bv.tum.de

[#]Geobrugg AG - Protection Systems
Aachstrasse 11
8590 Romanshorn, Switzerland
e-mail: corinna.wendeler@geobrugg.com, www.geobrugg.com

Key words: Shallow Landslides, Coupled Problems, Freeface Flow, Natural Hazard Protection, Chain-Link Finite Element

Abstract. Innovative flexible protection barriers against shallow landslides, made from high tensile steel meshes, demanded the development of numerical models for the different mesh types accompanied with 1:1 scale prototype field testing. To generate the barrier model, a special Discrete Element had to be developed that represents the chain link type of steel mesh. It was crucial that it includes the complex out-of-plane height of the mesh that accomodates the local high deformations of the net. The filling process of the barrier is simulated using OpenFOAM, taking the fluid-structure interaction with high deformations into account. In addition to the shear thinning behaviour of the mud, the turbulence in form of a large eddy simulation can be introduced, having regard to the interaction with the complex geometry and free surface by a Dynamic Mixed Model approach for the subgrid scales. The model is validated with 1:1 scale tests where shots of 50 m³ of mud travel down a 40 m long slope and hit the barrier.

1 INTRODUCTION

Shallow landslides are natural hazards of increasing importance, making the development of protection barriers that can retain a shallow landslide impact a very important issue. Flexible protection systems against rockfall or shallow landslides have certain advantages compared to rigid protection walls. Their main advantage is the transfer of the impact energy to anchors over time, with brake elements designed to consume the impact energy through plastic deformation. This way, a flexible protection system can slowly decelerate the impact and control peak forces throughout the structure. With protection nets made out of high-tension steel wire, carried by support ropes that are lead over steel posts to ground anchors, low weight barriers were developed that can be installed without much intervention to the slope. They have the capability to catch high energy rapid mass movements like rockfall or shallow landslides. A successful design of this kind of protection systems needs to trace the dynamic behaviour of all involved parts and needs to adjust their dimensioning to the energy transfer through the system. In general, this procedure demands 1:1 scale testing so there is a high demand for numerical models that reduce the number of tests necessary to find a successful design. This work focuses on two challenging issues within this context, the numerical model for *chain-link* steel meshes, and the simulation of the dynamic interaction of a shallow landslide impact with the barrier. The aim of this work is to improve the current state of the art of shallow landslide barrier design, where the whole filling process still is replaced with a static analysis and an equally distributed force.

2 CHAIN-LINK TYPED STEEL NETS

There are a wide range of different mesh structures. The *chain-link* meshes normally found in the design of standard fence structures gain increasing use in setups using high-strength steel due to its enormous capacity to retain loads while having a low need of steel material. Figure 1 shows the general geometry of a chain link mesh and its representation as *elements*, *segments* and the mesh junctions as *nodes* for numerical discretization. The performance of the mesh is dependent on non-trivial processes in the mesh nodes that control the maximal load the mesh can take and the development of the mesh stiffness under high stress.

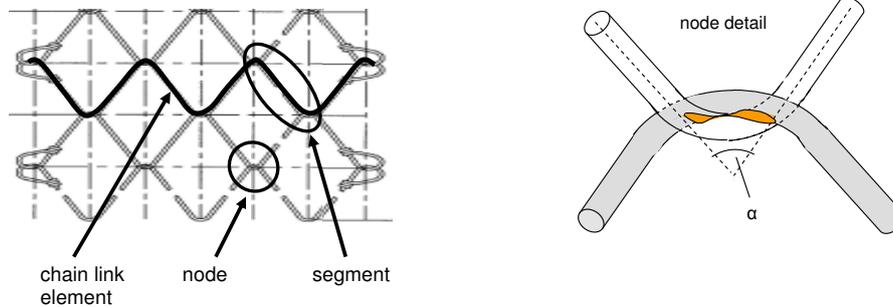


Figure 1: An *element* represents a wire (bundle), a *segment* specifies the element section from net node to net node. At the *node* (right), two elements cross each other in accordance with the mesh angle α and form a contact area between them.

2.1 Angle dependent performance

A series of quasi-static in-plane tension tests (Fig. 2) have demonstrated that the limit load of a segment depends on the mesh angle α . Net samples fixed on all borders and loaded by dragging one border can take about 30% more load when deformed in the direction of the elements than if pulled in an orthogonal direction. This is caused by the increase of the element length that leads to increasing mesh angles, while an orthogonal deformation of the sample causes increasing element widths with decreasing angles.



Figure 2: A TECCO type net is pulled in mesh direction decreasing the mesh angle while deforming.

2.2 Numerical approach of the discrete chain-link finite element

The test samples failed at the net nodes. With detailed modelling of the node geometry in dependency of the mesh angle, the combined normal force, shear and momentum of the wire along the node could be calculated. As failure criteria, the section with the highest normal force and shear interaction was defined. At this section, the steel wire resistance combining shear, normal force and bending was re-checked. With only varying

the steel material properties and the net geometry, the maximal load of plane tensile tests of different mesh types with single wire elements could be predicted with $\pm 5\%$ accuracy.

The net stiffness is not directly related to the stiffness of the steel wire but is dominated by the bending resistance of the segments at the nodes. The out-of-plane-height of the mesh construction provides reserves for locally high deformations that can develop with plastic bending deformations. Two opposite processes occur under loading, one leading to an increase of net stiffness by reducing the out-of-plane-height of the mesh and one that decreases the net stiffness by normal stress - bending interaction. By including these processes to the model, it was possible to simulate the longitudinal and transversal plane tensile tests with good accuracy, matching the maximal deformations with $\pm 6\%$.

The high computational costs needed for such a detailed approach (about three hours to simulate one second of a mesh specimen containing 200 segments on an Intel dual core machine, 3GHz) led to a simplified model that takes advantage of the nearly linear-plastic behaviour but accommodates the dependency to the mesh angle. The discrete chain-link finite element model uses instead the maximal load per mesh width, the maximal load per mesh length and the corresponding maximal deformations as input parameters. This information can be found on the mesh type data sheets and is usually generated from plane tension tests with 1 x 1 m specimen.

2.3 Calibration

The stiffness of the 1 x 1 m specimen under plane loads is not very representative for a large scale net under out-of-plane impact. Therefore, the two maximal deformation input parameters for the discrete chain-link finite element were calibrated for each net type by vertical impact tests, where a 820 kg concrete ball with accelerometers embedded in its core was dropped into a 3.5 x 3.5 m net (Fig. 3 left). The corresponding simulation (Fig. 3 right) shows the discrete chain link finite element integrated in the software FARO^[1]. A comparison of simulation to measurement data obtained with load cells installed in the steel frame, the accelerations measured in the ball and its position captured with high speed cameras (Fig. 3 middle) shows the impact of the ball, its acceleration back in the air and a second impact when the ball drops back in the net.

2.4 Validation

Full scale rockfall tests were taken as validation for the chain-link finite element. Two vertical high impact tests were used, one with a 1600 kg concrete block generating a 500 kJ rockfall impact (Fig. 4) and one with a 3200 kg concrete block and 1000 kJ of impact energy. At both tests, the block reached the net with 25 m/s velocity. The quality of the simulation was observed comparing the simulated and measured time delay of the impact, the maximal rock displacement and the rope forces. The process is reproduced with good accuracy concerning the time delay, and the maximal displacement of the rock from the moment of contact to the moment of standstill is predicted within $\pm 10\%$ difference.

The forces of ropes that are connected to brake elements show good results as in Fig. 3 (middle), overestimating the forces by maximal 13% dependent on the brake element material law. The retaining ropes holding the posts tend to develop dynamic oscillation which is a problem of the rope element. It occurs independent of the chain-link mesh and is subject to further investigation.

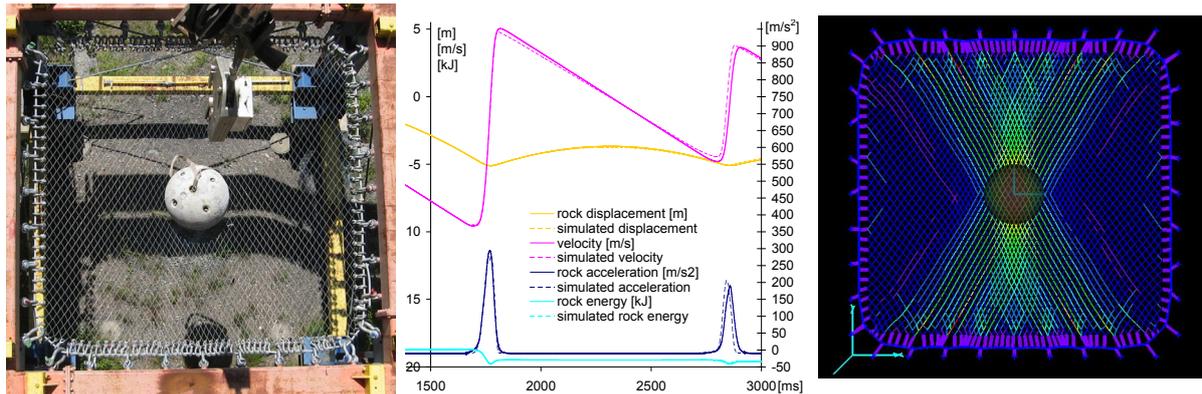


Figure 3: Test site (left), corresponding model (right) and comparison of the calibrated model with curves from experimental data over time in [ms] (middle).

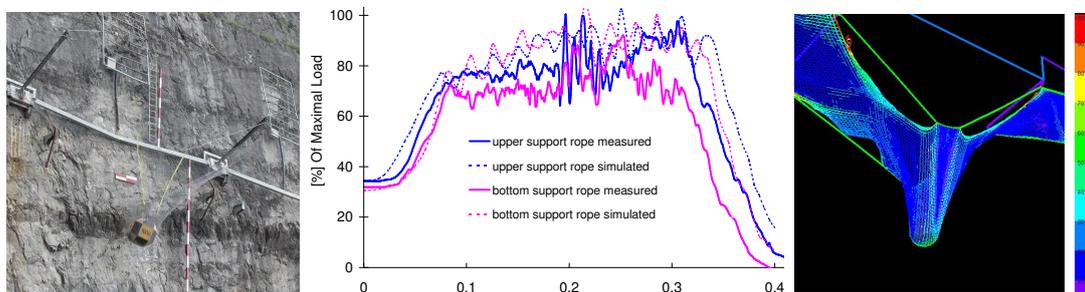


Figure 4: A 500 kg concrete block hits the barrier with 25 m/s (left), calculated and measured rope forces over time in [s] (middle) and the corresponding simulation with colours representing the degree of utilisation (right).

3 SHALLOW LANDSLIDE IMPACT WITH CFD

A crucial factor in the process of development of shallow landslide protection systems is the determination of the landslide impact pressure. Extensive research is in progress to clarify the shallow landslide dynamics and to allow the two dimensional simulation of flow head, density and mean velocity over a slope surface. Derived from depth averaged simulations with two phases, properties like density, water content, flow head and the mean front velocity can be estimated for the moment of impact at the barrier^[2]. These properties were treated as input parameters to a three-dimensional model that represents

the impact area by a finite volume grid that can handle dynamic mesh movements. Due to the local scale of the barrier impact and its short process, the shallow landslide material can be treated as homogenous. It is represented with a two phase freeface flow model processed by using the computational fluid dynamics of OpenFOAM^[3]. The *Volume of Fluid (VOF)* approach was chosen to trace the interface between air and mud, and the mud was represented as a *Herschel-Bulkley* fluid. The solver *InterDyMFoam* was chosen which can handle dynamic mesh deformations, two face flow and turbulence. The shallow landslide impacts studied in this work were dominated by laminar flow, but turbulence in high suspension flows is a focus of ongoing research^{[5][6]}. Therefore, as an option, the *Large Eddy Simulation (LES)* turbulence model of the solver accounting for complex geometry, was combined with the *Dynamic Mixed Model* for the subgrid scales that treats the surface influence and the backscatter effect in an appropriate way^[7]. The corresponding code was developed at LTT Rostock^[4].

3.1 Solver validation

As a check on the performance of the solver, three channel flow experiments with three sets of water inflow to a 2 m long and 0.3 m wide channel were carried out, and the surface velocity was measured with high-speed video tracking suberic particles^[8]. The Reynolds number reached from transient to fully turbulent flow. The simulation using the LES turbulence model, applied the Dynamic Mixed Model for the subgrid scales together with a successful clipping procedure^[9], both implemented by the LTT Rostock. The average surface flow velocity in the simulation fitted the measured results with $\pm 4\%$. Figure 5 presents the simulation grid with the freeface flow velocity.

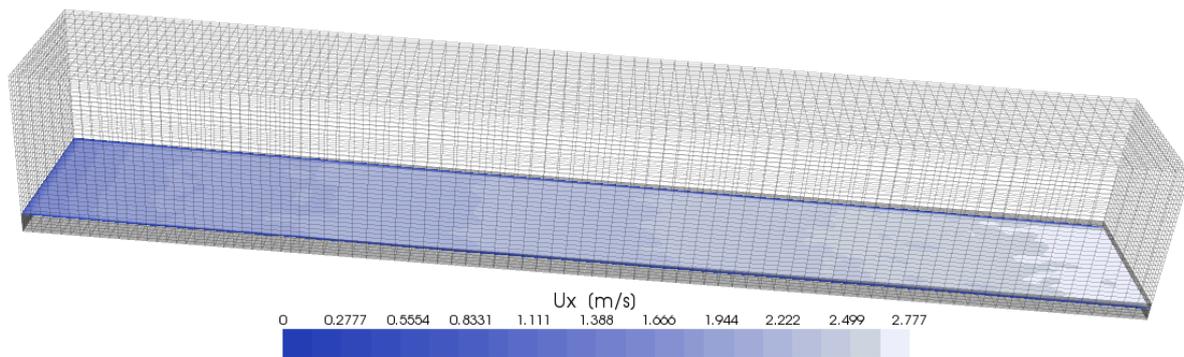


Figure 5: Simulation of a freeface channel flow, 2m long and 0.3m wide, with 20.4 l/s inflow on the left side and 5° inclination.

4 FSI COUPLING

The fluid structure interaction between the FARO structure code and the OpenFOAM CFD simulation was carried out by explicit weak coupling. After a defined interaction interval has passed, FARO sends the barrier deformation to the InterDyMFoam solver

which performs the corresponding mesh deformation to fit the front patch of the simulation grid to the barrier shape defined in FARO, and the InterDyMFoam solver sends the forces at each front cell face to FARO where they are applied to barrier nodes (Fig. 6).

4.1 Boundary condition between non-matching grids

For the development of different barrier types it is necessary to allow maximal precision in the barrier model while guaranteeing a certain independency of the CFD simulation from the geometry details of the construction. The deformation of the barrier model simulates the dynamic behaviour of the structure including the wave propagation through the non-loaded net when the front of the landslide hits the net at its bottom. Thereby, the deformation of the net has strong variations on small scales. For the OpenFOAM solver, it is necessary to avoid deformations where a cell face travels through the cells opposite face causing negative cell volumes and false cell orientations. For that reason, instead of linking the nodes of the finite volume grid boundary to the barrier nodes in a fixed manner, the interaction was implemented as a sliding interface: At every interaction step, each cell node belonging to the finite volume grid boundary searches for barrier nodes that lie momentarily closest to a slope-parallel projection of the cell nodes. Then, the boundary nodes of OpenFOAM perform a displacement along their slope parallel projection so that they get placed along the three closest barrier nodes of the FARO simulation.

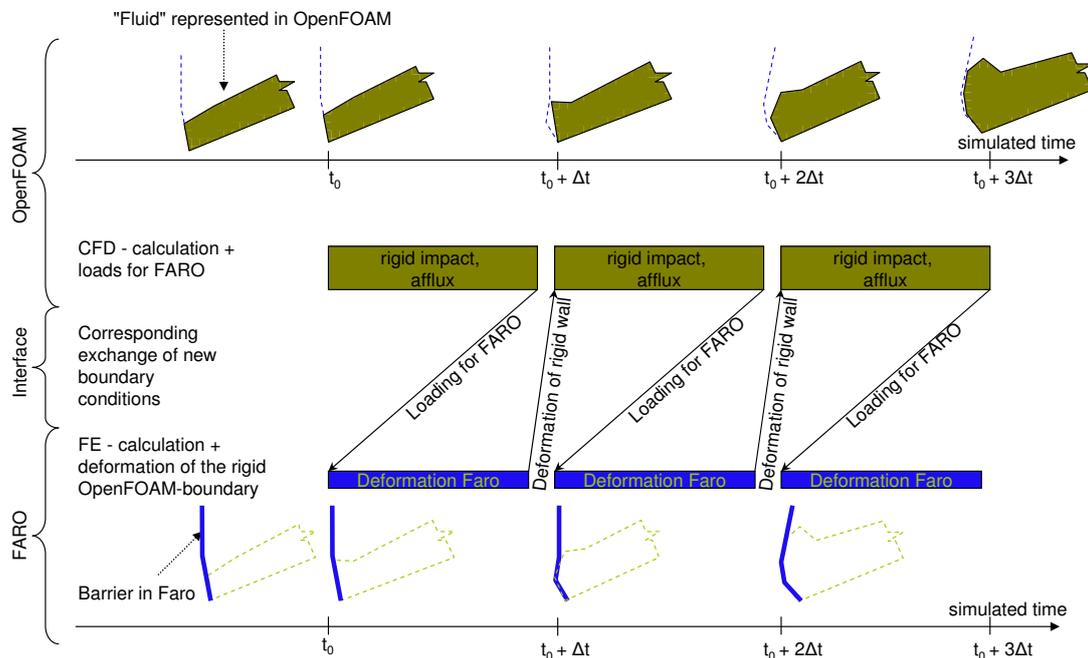


Figure 6: Fluid-structure-interaction between FARO and OpenFOAM: Stepwise update of fluid and barrier dynamics.

4.2 Modelling the area above the barrier

The barrier decelerates the impact and causes the flow to form a wave with a counter-current, which can develop higher than the top end of the barrier. It was necessary to extend the finite volume grid in such way that it reaches about one meter higher than the barrier. This fact caused a concave discontinuity at the front boundary shape. With increasing barrier deformations, cells became false orientated and caused solver instability (Fig. 7). It became necessary to smoothen the edge of the finite volume grid at the upper barrier border (upper support rope) by moving all nodes of the finite volume boundary, which are positioned higher than a cell row below the upper support rope. A successful implementation that can handle different barrier types takes the deformation at the upper support rope and multiplies it with a factor dependent to the point with the maximal barrier deformation. Then this increased upper support rope deformation is applied along all finite volume nodes positioned higher than the upper support rope region, but with an applied linear decrease with height.

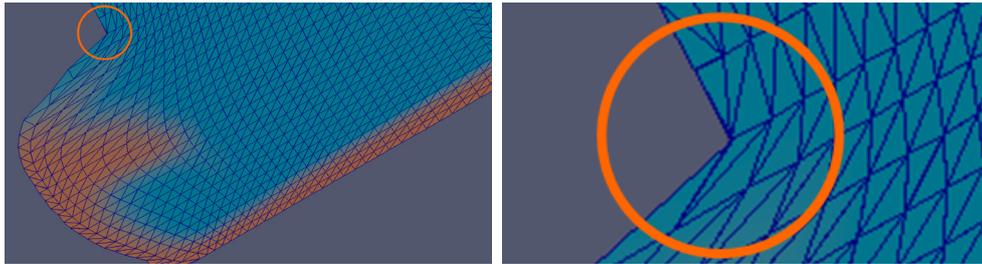


Figure 7: Cut through the simulation grid showing the two phases by colour. The circle marks the concave discontinuity of the front where the back face of the cell interpenetrates the front.

4.3 Data exchange

The exchange of forces and deformations between the OpenFOAM and FARO models is processed via input and output file streams. Initially, the interface writes a file with the position of the nodes at the boundary of the finite volume grid. The FARO code reads and stores these initial positions, and at each interaction step it calculates the discrepancy between these initial positions and their slope-parallel projection on the FARO-barrier. Then the discrepancy is written into a file as displacement vectors, one line for each OpenFoam boundary node. The InterDyMFoam solver reads in these displacement vectors and applies them to the grid over the next interaction interval with its internal dynamic mesh solver. But before starting the next interval, it calculates forces for each OpenFoam boundary node that represent the pressure at the boundary, and writes these force vectors to another file. FARO then reads these forces and distributes them: Each OpenFoam boundary node force is distributed along three FARO nodes, that lie closest to the slope parallel projection of the corresponding OpenFoam boundary node.

The timing of reading and writing, and the control of one simulation waiting for the

other to finish the current interaction interval, is done by two flags written to another two files that store the stage of each simulation. Fig. 8 shows the interface part of the flowchart with the OpenFOAM solver waiting for FARO until FARO has read the forces and written the displacements of the current interaction step. OpenFOAM then reads the displacements, calculates and writes the forces and continues the simulation. Problematic was the file sharing without lock that allows reading the number of CFD interactions while they are written. As a workaround there are currently two file stream readers following each other and comparing the tokens they read. If the tokens match, it is assumed that this token was not overwritten during reading.

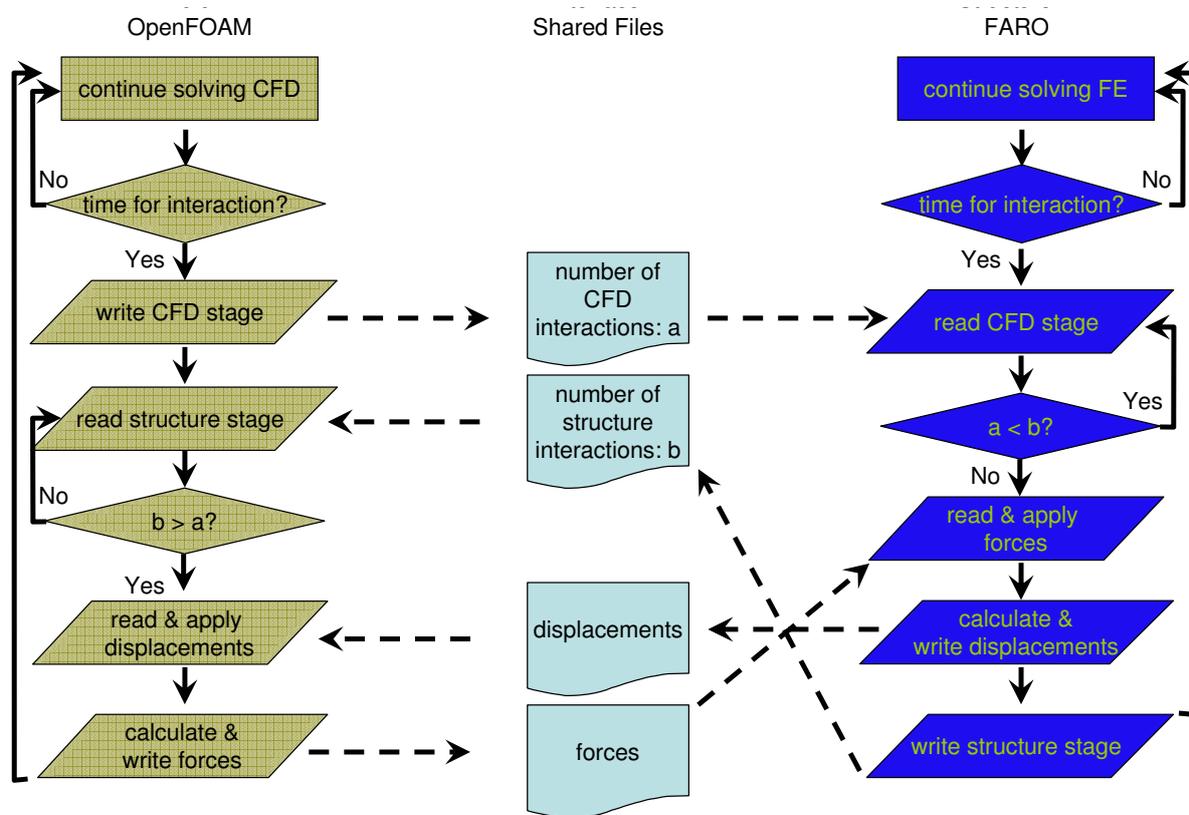


Figure 8: Part of a flowchart of the interface processes in OpenFOAM. From “write CFD stage” to “write structure stage” OpenFOAM stays in the “read structure stage” loop.

4.4 Execution time

The speed of the calculation is limited by the need for exchange between the two codes. It is not an issue of time consumption for data transfer, the files written and read at an interaction stage typically need less than 100 Kb all together. But the solver for the two phase flow could perform much higher timesteps by the use of its automatic time step control that allows speeding up the solution at stages where the flow proceeds with low

courant numbers. The need for a small interaction interval to avoid pressure oscillations limits the time step, so the simulation currently takes four hours for the simulation of one second of impact on a standard dual core PC.

4.5 Testing

A wide number of 1:1 scale test have been realized, were up to 50 cubic meters of mud travel down a 30 m long slope, passing three flow head measurement points carried out with laser, a ground plate taking the weight and tangential force of the flow and two vertical force plates that register the impact pressure. Finally the artificial shallow landslides hit the flexible barrier and activate four tension load cells installed at the support ropes and retaining ropes of the barrier^[10].

4.6 Results

First outputs of one simulated test give promising results: The impact pressures of the model lie below the ones measured by the rigid impact force plates that reached 90000 Pa, see Fig. 10. This fits the assumption that the flexible barrier results in lower pressures than a rigid wall. Upper support rope forces and retaining rope forces fit well to the measured maxima and the development over time (Fig. 9). After the impact, material slowly leaves the barrier towards the sides. That effect is overestimated leading to a faster decrease of static loads in the simulation.

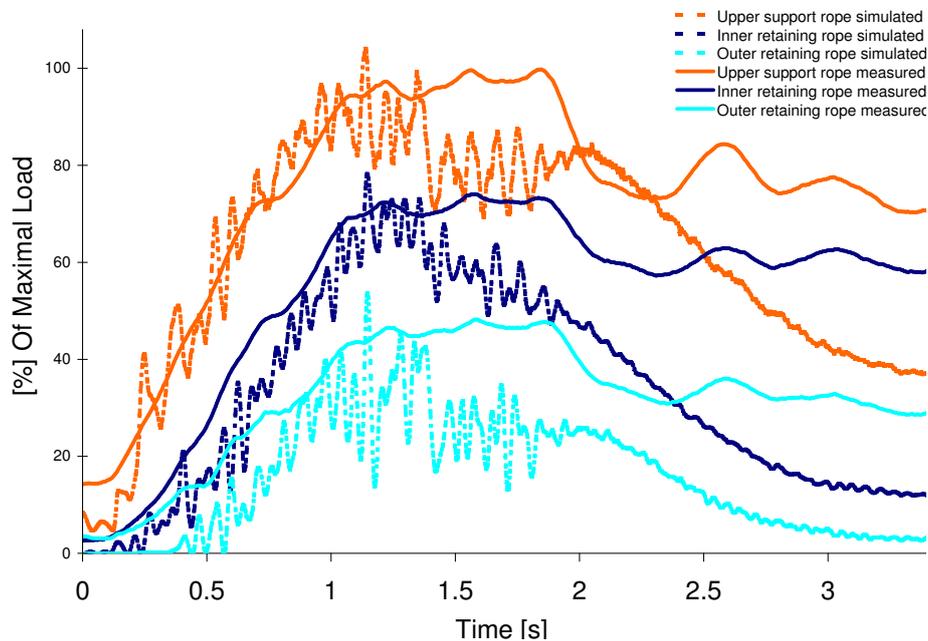


Figure 9: Measured and simulated rope forces taken from the 7th test performed at the WSL-test side in Veltheim, Switzerland.

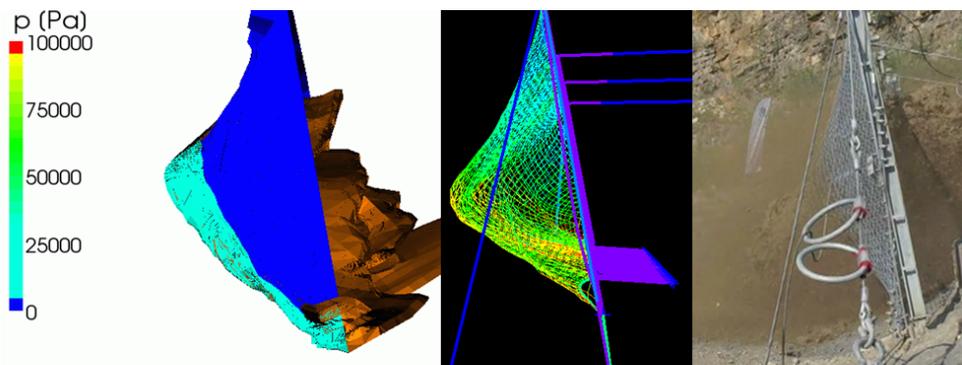


Figure 10: Full impact: The maximal impact wave reaches a height in OpenFOAM that is about the height witnessed at the experiment.

5 CONCLUSIONS

The development of flexible shallow landslide barriers is in need for numerical models. The state of the art to apply Finite-Element simulations for the design of the barrier structure was raised by developing a discrete element for the investigated net types that predicts not only the maximal load but the deformations with good accuracy. The loading of the barrier is conventionally simulated with static and equally distributed forces that do not account for the dynamic interaction. With the coupling of the structure code to a CFD simulation realized in OpenFOAM, the impact and the interaction between barrier and shallow landslide were successfully modelled, and should reduce the efforts needed in full scale prototype testing.

REFERENCES

- [1] Volkwein, A. *Numerische Simulation von flexiblen Steinschlagschutzsystemen*. Ph.D. Diss. Swiss Federal Inst. of Technology Zurich, Switzerland (2004).
- [2] Christen, M., Bartelt, P. and Gruber, U. *RAMMS - a Modelling System for Snow Avalanches, Debris Flows and Rockfalls based on IDL*. Photogramm. Fernerkund. Geoinf., (2007) **4**:289-292.
- [3] OpenFOAM: open source CFD, www.openfoam.com, visited 23.02.2011.
- [4] OpenFOAM LTT Rostock Extensions, www.ltt-rostock.de/mediawiki/index.php/LTTRostockExtensions, visited 23.02.2011.
- [5] Huhtanen, J. P. T. and Karvinen, R. J. *Interaction of Non-Newtonian Fluid Dynamics and Turbulence on the Behavior of Pulp Suspension Flows*. Annual transactions of the Nordic Rheology Society, (2005) **13**:177-186.

- [6] Jakob, M. and Hungr, O. *Debrisflow hazards and related phenomena*. Springer-Praxis Series in Geophysics, (2005) ISBN: 978-3-540-20726-9, p. 166-167.
- [7] Lei Wang and Xi-Yun Lu. *Large eddy simulation of stably stratified turbulent open channel flows with low- to high-Prandtl number*. International Journal of Heat and Mass Transfer (2005) **48**:1883-1897.
- [8] Hegglin, R., *Beiträge zum Prozessverständnis des fluvialen Geschiebetransports*. Master Thesis Institute of Geography, University of Bern, Switzerland (2010).
- [9] Kornev, N. V., Tkatchenko, I. V. and Hassel, E. H. *A simple clipping procedure for the dynamic mixed model based on Taylor series approximation*. Commun. Numer. Meth. Engng (2006) **22**:55-61.
- [10] Bugnion, L.; Volkwein, A., Wendeler, C. and Roth, A. *Large-scale field testing on flexible shallow landslide barriers*. Geophys. Res. Abstr. 12: EGU2010-11755 (2010).

SHEAR DEFORMABLE BEAMS ON NONLINEAR VISCOELASTIC FOUNDATION UNDER MOVING LOADING

EVANGELOS J. SAPOUNTZAKIS* AND ANDREAS E. KAMPITSIS†

* School of Civil Engineering, National Technical University of Athens (NTUA)
Zografou Campus, GR-157 80, Athens, Greece
e-mail: cvsapoun@central.ntua.gr

† School of Civil Engineering, National Technical University of Athens (NTUA)
Zografou Campus, GR-157 80, Athens, Greece
e-mail: evakamb@gmail.com

Key words: Nonlinear Dynamic Analysis, Large Deflections, Moving Loads, Timoshenko Beam, Boundary Element Method, Nonlinear Viscoelastic Foundation.

Abstract. In this paper, a boundary element method is developed for the nonlinear response of shear deformable beams of simply or multiply connected constant cross section, traversed by moving loads, resting on tensionless nonlinear viscoelastic foundation, undergoing moderate large deflections under general boundary conditions. The beam is subjected to the combined action of arbitrarily distributed or concentrated transverse moving loading as well as to axial loading. To account for shear deformations, the concept of shear deformation coefficients is used. Three boundary value problems are formulated with respect to the transverse displacement, to the axial displacement and to a stress functions and solved using the Analog Equation Method, a BEM based method. Application of the boundary element technique yields a system of nonlinear differential – algebraic equations (DAE), which is solved using an efficient time discretization scheme, from which the transverse and axial displacements are computed. The evaluation of the shear deformation coefficient is accomplished from the aforementioned stress function using only boundary integration. Analyses are performed to investigate the effects of various parameters, such as the load velocity, load frequency, shear rigidity, foundation nonlinearity, damping, on the beam displacements and stress resultants and to examine how the consideration of shear and axial compression affect the response of the system.

1 INTRODUCTION

Vibration analysis of beams traversed by moving load is of great interest in the area of high-speed transportation or rocket-sledge technology. This problem can be modelled as a beam on elastic foundation subjected to loading moving at a constant speed.

When the beam deforms the conventional elastic foundation models can sustain both compression as well as tension. In order to address this issue tensionless foundation models were proposed, in which regions of no contact develop beneath the beam. These regions are unknown and the change of the transverse displacement sign provides the condition for the determination of the contact region. Besides, having in mind the magnitude of the arising

compressive forces due to environmental loads such as changes in temperature or moisture or due to the train wheel and the importance of weight saving in engineering structures, the study of nonlinear effects on the analysis of supporting structural elements becomes essential. This non-linearity results from retaining the square of the slope in the strain–displacement relations (intermediate non-linear theory), avoiding in this way the inaccuracies arising from a linearized second – order analysis. Moreover, due to the intensive use of materials having relatively high transverse shear modulus, the error incurred from the ignorance of the effect of shear deformation may be substantial, particularly in the case of heavy lateral loading. All of the aforementioned concepts constitute the motive for a rigorous nonlinear dynamic analysis of shear deformable beams subjected to moving loads and resting on a tensionless nonlinear viscoelastic foundation.

When the beam deflections are small, a wide range of linear analysis tools, such as modal analysis, can be used, and some analytical results are possible. Analytical solutions of problems involving beam vibrations of simple geometry and boundary conditions under moving loads have received a good amount of attention in the literature, with pioneer the work of Krylov [1] and later the one of Timoshenko [2] who determined dynamic stresses in the beam structure. Linear transverse vibrations of a simply supported beam traversed by a constant force moving at a constant velocity were presented by Inglis [3], Lowan [4] and later on by Koloušek [5] and Fryba [6].

Since then, important development has been achieved regarding also *linear* more rigorous dynamic analyses of beams under moving loads employing either analytical or numerical methods. Kargarnovin and Younesian [7-9] studied the response of infinite beams supported by nonlinear or Pasternak-type viscoelastic foundations subjected to harmonic moving loads employing a perturbation methods. Zehsaz et. al. [10] studied the dynamics of railway, as a Timoshenko beam of limited length, lying on a Pasternak viscoelastic foundation, subjected to moving load employing the modal superposition method.

As the beam deflections become larger, the induced *geometric nonlinearities* result in effects that are not observed in linear systems. Chen et. al. [11] performing a geometrically nonlinear analysis with constant axial force presented the dynamic stiffness matrix of an infinite Timoshenko beam on viscoelastic foundation subjected to a harmonic moving load and determined the critical velocities and the resonant frequencies. Kim and Cho [12] presented the vibration and buckling of an infinite beam-column under constant axial force, resting on an elastic foundation and subjected to moving loads of either constant or harmonically varying amplitude with a constant advance velocity, taking into account shear deformation effect.

In this paper, a boundary element method is developed for the nonlinear response of shear deformable beams of simply or multiply connected constant cross section, traversed by moving loads, resting on tensionless nonlinear viscoelastic foundation, undergoing moderate large deflections under general boundary conditions. The beam is subjected to the combined action of arbitrarily distributed or concentrated transverse moving loading as well as to axial loading. To account for shear deformations, the concept of shear deformation coefficients is used. Three boundary value problems are formulated with respect to the transverse displacement, to the axial displacement and to a stress functions and solved using the Analog Equation Method [13], a BEM based method. Application of the boundary element technique yields a system of nonlinear differential–algebraic equations (DAE), which is solved using an

efficient time discretization scheme. The evaluation of the shear deformation coefficient is accomplished from the aforementioned stress function using only boundary integration. The essential features and novel aspects of the present formulation compared with previous ones are summarized as follows.

- i. Shear deformation effect and rotary inertia are taken into account in the nonlinear dynamic analysis of beams subjected to arbitrary (distributed or concentrated) transverse moving, as well as to axial loading.
- ii. The homogeneous linear half-space is approximated by a tensionless three-parameter viscoelastic foundation.
- iii. The beam is supported by the most general nonlinear boundary conditions.
- iv. The proposed model takes into account the coupling effects of bending and shear deformations along the member as well as shear forces along the span induced by the applied axial loading.
- v. The shear deformation coefficients are evaluated using an energy approach, instead of Timoshenko's [14] and Cowper's [15] definitions.
- vi. The effect of the material's Poisson ratio ν is taken into account.
- vii. The proposed method employs a BEM approach (requiring boundary discretization) resulting in line or parabolic elements instead of area elements of the FEM solutions (requiring the whole cross section to be discretized into triangular or quadrilateral area elements), while a small number of line elements are required to achieve high accuracy.

Analyses are performed to investigate the effects of various parameters, such as the load velocity, load frequency, shear rigidity, foundation nonlinearity, damping, on the beam displacements and stress resultants and to examine how the consideration of shear and axial compression affect the response of the system.

2 STATEMENT OF THE PROBLEM

Let us consider a prismatic beam of length l (Fig.1a), of arbitrary constant cross-section of area A (Fig.1b), having at least one axis of symmetry (z -axis). The homogeneous isotropic and linearly elastic material of the beam cross-section, with modulus of elasticity E , shear modulus G and Poisson's ratio ν occupies the two dimensional multiply connected region Ω of the y, z plane and is bounded by the Γ_j ($j=1, 2, \dots, K$) boundary curves, which are piecewise smooth, i.e. they may have a finite number of corners. The beam is supported on a homogeneous tensionless nonlinear three-parameter viscoelastic soil. The foundation model is characterized by the linear Winkler modulus k_L , the nonlinear Winkler modulus k_{NL} , the Pasternak (shear) foundation modulus k_p and the damping coefficient c_z . Taking into account the unbonded contact between beam and subgrade, the interaction pressure at the interface can be only compressive and is represented for the transverse direction by the following relations

$$p_{sz}(x, t) = U(x, t) p_{react}(x, t) \quad (1a)$$

$$p_{react}(x, t) = k_L w(x, t) + k_{NL} w^3(x, t) - k_p \frac{\partial^2 w(x, t)}{\partial x^2} + c \frac{\partial w(x, t)}{\partial t} \quad (1b)$$

where $U(x, t)$ is a unit step function defined as

$$U(x,t) = \begin{cases} 1 & \text{if } p_{react}(x,t) > 0 \\ 0 & \text{if } p_{react}(x,t) \leq 0 \end{cases} \quad (2)$$

The foundation reaction p_{react} of eqn.(1b) takes into account the nonlinear behavior of the subsoil (e.g. ballast and rail-bed) as proposed by Dahlberg [16], demonstrating that the differences between the non-linear and linear models are considerable and a non-linear track model simulates the rail deflection quite well whereas the equivalent linear one cannot. Later, Wu and Thompson [17] presented a similar non-linear model and studied the problem of wheel/track impact employing the finite element method. Moreover, for real sample of the hardening behavior of the foundation one can refer to [18] where detailed field measurement results are presented.

The beam is subjected to the combined action of the arbitrarily distributed or concentrated transverse along the axis of symmetry moving loading $p_z = p_z(x,t)$ with constant velocity V as well as to axial loading $p_x = p_x(x,t)$, as shown in Fig.1a. Under the action of this loading, the displacement field of the beam taking into account shear deformation effect is given as

$$\bar{u}(x,y,z,t) = u(x,t) + z\theta_y(x,t) \quad \bar{w}(x,t) = w(x,t) \quad (3a,b)$$

where \bar{u} , \bar{w} are the axial and transverse beam displacement components with respect to the Cyz system of axes (Fig.1b); $u(x,t)$, $w(x,t)$ are the corresponding components of the centroid C and $\theta_y(x,t)$ is the angle of rotation due to bending of the cross-section with respect to the same point.

Employing the strain-displacement relations of the three - dimensional elasticity for moderate displacements the following strain components can be easily obtained

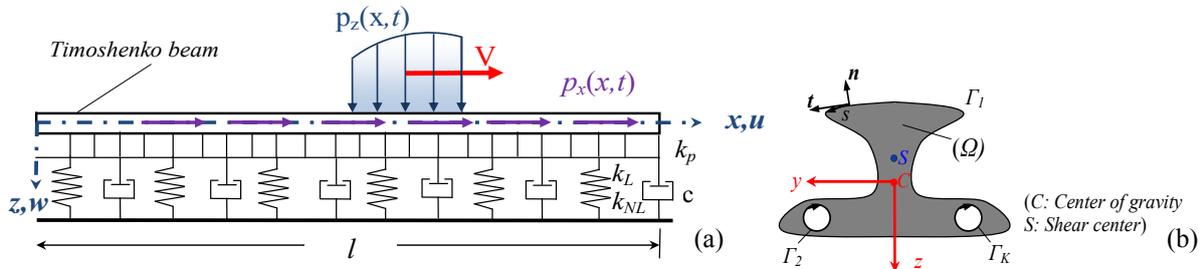


Figure 1: Prismatic beam resting on a nonlinear viscoelastic foundation (a) of an arbitrary mono symmetric cross-section occupying the two dimensional region Ω (b)

$$\varepsilon_{xx} = \frac{\partial \bar{u}}{\partial x} + \frac{1}{2} \left(\frac{\partial \bar{w}}{\partial x} \right)^2 \quad \gamma_{xz} = \frac{\partial \bar{w}}{\partial x} + \frac{\partial \bar{u}}{\partial z} + \frac{\partial \bar{w}}{\partial x} \frac{\partial \bar{w}}{\partial z} \quad \varepsilon_{zz} = \gamma_{yz} = 0 \quad (4a,b,c)$$

where it has been assumed that for moderate displacements $\left(\frac{\partial \bar{u}}{\partial x} \right)^2 \ll \frac{\partial \bar{u}}{\partial x}$, $\left(\frac{\partial \bar{u}}{\partial x} \right) \left(\frac{\partial \bar{u}}{\partial z} \right) \ll \left(\frac{\partial \bar{u}}{\partial x} \right) + \left(\frac{\partial \bar{u}}{\partial z} \right)$. Substituting the displacement components (3) to the strain-displacement relations (4), the strain components can be written as

$$\varepsilon_{xx}(x, z, t) = u' + z\theta_y' + \frac{1}{2}w'^2 \quad \gamma_{xz} = w' + \theta_y \quad (5a,b)$$

where γ_{xz} is the additional angle of rotation of the cross-section due to shear deformation.

Considering strains to be small, employing the second Piola – Kirchhoff stress tensor and assuming an isotropic and homogeneous material, the stress components are defined in terms of the strain ones as

$$\begin{Bmatrix} S_{xx} \\ S_{xz} \end{Bmatrix} = \begin{bmatrix} E & 0 \\ 0 & G \end{bmatrix} \begin{Bmatrix} \varepsilon_{xx} \\ \gamma_{xz} \end{Bmatrix} \quad (6)$$

or employing eqns. (5) as

$$S_{xx} = E \left(u' + z\theta_y' + \frac{1}{2}w'^2 \right) \quad S_{xz} = G \cdot (w' + \theta_y) \quad (7a,b)$$

On the basis of Hamilton's principle, the variations of the Lagrangian equation defined as

$$\delta \int_{t_1}^{t_2} (U - K - W_{ext}) dt = 0 \quad (8)$$

and expressed as a function of the stress resultants acting on the cross section of the beam in the deformed state provide the governing equations and the boundary conditions of the beam subjected to nonlinear vibrations. In eqn.(8), $\delta(\cdot)$ denotes variation of quantities, while U , K , W_{ext} are the strain energy, the kinetic energy and the external load work, respectively given as

$$\delta U = \int_V (S_{xx} \delta \varepsilon_{xx} + S_{xz} \delta \gamma_{xz}) dV \quad \delta K = \frac{1}{2} \int_V \rho (\delta \dot{u}^2 + \delta \dot{w}^2) dV \quad (9a,b)$$

$$\delta W_{ext} = \int_L (p_x \delta u + p_z \delta w - \delta(p_z w)) dx \quad (9c)$$

Moreover, the stress resultants of the beam are given as

$$N = \int_{\Omega} S_{xx} d\Omega \quad M_y = \int_{\Omega} S_{xx} z d\Omega \quad Q_z = \int_{A_z} S_{xz} d\Omega \quad (10a,b,c)$$

Substituting the expressions of the stress components (7) into equations (10), the stress resultants are obtained as

$$N = EA \left(u' + \frac{1}{2}w'^2 \right) \quad M_y = EI_y \theta_y' \quad Q_z = GA_z \gamma_{xz} \quad (11a,b,c)$$

where A is the cross section area, I_y the moment of inertia with respect to z-axis given as

$$A = \int_{\Omega} d\Omega \quad (12)$$

$$I_y = \int_{\Omega} z^2 d\Omega \quad (13)$$

and GA_z is its shear rigidity of the Timoshenko's beam theory, where

$$A_z = \kappa_z A = \frac{1}{a_z} A \quad (14)$$

is the shear area, respectively with κ_z the shear correction factor and a_z the shear deformation coefficient. Substituting the stress components given in eqns. (7) and the strain resultants given in eqns. (5) to the strain energy variation δE_{int} (eqn.9a) and employing eqn. (8), the equilibrium equations of the beam are derived as

$$\rho A \ddot{u} - EA(u'' + w'w'') = p_x \quad (15a)$$

$$\rho A \ddot{w} - (Nw')' - GA_z(w'' + \theta_y') + p_{sz} = p_z \quad EI_y \theta_y'' - GA_z(w' + \theta_y) = \rho I_y \ddot{\theta}_y \quad (15b,c)$$

where (\cdot) , (\prime) denote differentiation with respect to t , x , respectively. Combining equations (15b,c) the following differential equations with respect to u , w are derived as

$$\rho A \ddot{u} - EA(u'' + w'w'') = p_x \quad (16a)$$

$$EI_y w'''' + \rho A \ddot{w} + p_{sz} + \frac{EI_y}{GA_z} \left[(Nw')''' - \rho A \frac{\partial^2 \ddot{w}}{\partial x^2} - p_{sz}'' + p_z'' \right] - (Nw')' - \rho I_y \frac{\partial^2 \ddot{w}}{\partial x^2} - \frac{\rho I_y}{GA_z} \left[\frac{\partial^2 (Nw')'}{\partial t^2} - \rho A \ddot{w} - \ddot{p}_{sz} + \ddot{p}_z \right] = p_z \quad (16b)$$

Eqns. (16) constitute the governing differential equations of a Timoshenko beam, supported on a tensionless nonlinear three-parameter viscoelastic foundation, subjected to nonlinear vibrations due to the combined action of arbitrarily distributed or concentrated transverse moving loading as well as to axial loading. These equations are also subjected to the pertinent boundary conditions of the problem at hand given as

$$a_1 u(x, t) + \alpha_2 N(x, t) = \alpha_3 \quad (17)$$

$$\beta_1 w(x, t) + \beta_2 V_z(x, t) = \beta_3 \quad \gamma_1 \theta_y(x, t) + \gamma_2 M_y(x, t) = \gamma_3 \quad (18a,b)$$

at the beam ends $x = 0, l$, together with the initial conditions

$$u(x, 0) = \bar{u}_0(x) \quad \dot{u}(x, 0) = \dot{\bar{u}}_0(x) \quad (19a,b)$$

$$w(x, 0) = \bar{w}_0(x) \quad \dot{w}(x, 0) = \dot{\bar{w}}_0(x) \quad (20a,b)$$

where $\bar{u}_0(x)$, $\bar{w}_0(x)$, $\dot{\bar{u}}_0(x)$ and $\dot{\bar{w}}_0(x)$ are prescribed functions. In eqns. (18) V_z and M_y are the reaction and bending moment, which together with the angle of rotation due to bending θ_y are given as

$$V_z = Nw' - EI_y w''' - \frac{EI_y}{GA_z} \left[(Nw')'' + p_z' - p_{sz}' - \rho A \frac{\partial \ddot{w}}{\partial x} \right] - \rho I_y \ddot{\theta}_y \quad (21a)$$

$$M_y = -EI_y w'' - \frac{EI_y}{GA_z} \left[(Nw')' + p_z - p_{sz} - \rho A \ddot{w} \right] \quad (21b)$$

$$\theta_y = \frac{EI_y}{G^2 A_z^2} \left(-p_z' + p_{sz}' - (Nw)'' + \rho A \frac{\partial \ddot{w}}{\partial x} \right) - \frac{1}{GA_z} (EI_y w'''' + \rho I_y \ddot{\theta}_y + GA_z w') \quad (21c)$$

Finally, $\alpha_k, \beta_k, \gamma_k$ ($k = 1, 2, 3$) are functions specified at the beam ends $x = 0, l$. Eqs. (17)-(18) describe the most general nonlinear boundary conditions associated with the problem at hand

and can include elastic support or restraint. It is apparent that all types of the conventional boundary conditions (clamped, simply supported, free or guided edge) can be derived from these equations by specifying appropriately these functions (e.g. for a clamped edge it is $\alpha_1 = \beta_1 = 1, \gamma_1 = 1, \alpha_2 = \alpha_3 = \beta_2 = \beta_3 = \gamma_2 = \gamma_3 = 0$).

The solution of the initial boundary value problem given from eqns (16), subjected to the boundary conditions (17)-(18) and the initial conditions (19)-(20), which represents the nonlinear flexural dynamic analysis of a Timoshenko beam, supported on a tensionless nonlinear three-parameter viscoelastic foundation, presumes the evaluation of the shear deformation coefficient a_z corresponding to the principal centroidal system of axes C_{yz} . This coefficient is established equating the approximate formula of the shear strain energy per unit length [19]

$$U_{appr.} = \frac{a_z Q_z^2}{2AG} \quad (22)$$

with the exact one given from

$$U_{exact} = \int_{\Omega} \frac{(\tau_{xz})^2}{2G} d\Omega \quad (23)$$

and are obtained as [20]

$$a_z = \frac{1}{\kappa_z} = \frac{A}{\Delta^2} \int_{\Omega} [(\nabla\Phi) - \mathbf{d}] \cdot [(\nabla\Phi) - \mathbf{d}] d\Omega \quad (24)$$

where (τ_{xz}) is the transverse (direct) shear stress component, $(\nabla) \equiv \mathbf{i}_y (\partial/\partial y) + \mathbf{i}_z (\partial/\partial z)$ is a symbolic vector with $\mathbf{i}_y, \mathbf{i}_z$ the unit vectors along y and z axes, respectively, Δ is given as

$$\Delta = 2(1+\nu)I_y I_z \quad (25)$$

ν is the Poisson ratio of the cross section material, \mathbf{d} is a vector defined as

$$\mathbf{d} = (\nu I_z y z) \mathbf{i}_y - \left(\nu I_z \frac{y^2 - z^2}{2} \right) \mathbf{i}_z \quad (26)$$

and $\Phi(y, z)$ is a stress function evaluated from the solution of the following Neumann type boundary value problem [20]

$$\nabla^2 \Phi = -2I_z z \quad \text{in } \Omega \quad \frac{\partial \Phi}{\partial n} = \mathbf{n} \cdot \mathbf{d} \quad \text{on } \Gamma = \bigcup_{j=1}^{K+1} \Gamma_j \quad (27a,b)$$

where \mathbf{n} is the outward normal vector to the boundary Γ . In the case of negligible shear deformations $a_z = 0$. It is also worth here noting that the boundary conditions (27b) have been derived from the physical consideration that the traction vector in the direction of the normal vector \mathbf{n} vanishes on the free surface of the beam.

3 INTEGRAL REPRESENTATIONS – NUMERICAL SOLUTION

According to the precedent analysis, the nonlinear flexural dynamic analysis of

Timoshenko beams, supported on a tensionless nonlinear three-parameter viscoelastic foundation, undergoing moderate large deflections reduces in establishing the displacement components $u(x,t)$ and $w(x,t)$ having continuous derivatives up to the second and up to the fourth order with respect to x , respectively, and also having derivatives up to the second order with respect to t (ignoring the inertia terms of the fourth order [21]). These displacement components must satisfy the coupled governing differential equations (16) inside the beam, the boundary conditions (17)-(18) at the beam ends $x=0,l$ and the initial conditions (19)-(20). Eqns (16) are solved using the Analog Equation Method [13] as this is developed for hyperbolic differential equations [22].

4 NUMERICAL EXAMPLES

On the basis of the analytical and numerical procedures presented, a computer program has been written and a representative example has been studied to demonstrate the efficiency of the developed method. In the example, the results have been obtained using $L=21$ nodal points (longitudinal discretization), 400 boundary elements (cross section discretization) and a time step of $\Delta t=1.0 \mu\text{sec}$.

4.1 Example

In order to illustrate the importance of the nonlinear analysis, a simply supported UIC60 rail track, resting on a nonlinear viscoelastic bilateral foundation is examined. The geometric constants of the track and the foundation are given in Table 1. The track is subjected to a concentrated moving harmonic load $p_z(x,t)=P\delta(x-Vt)\sin(\Omega t)$, where P,Ω are the amplitude and the frequency of the harmonic load, respectively and δ is the Dirac's delta function. Moreover, the track is subjected to an either tensile or compressive distributed axial load $p_x(x,t)=\pm 2500(kN/m)$.

In Fig. 2 the time history and the extreme values of the central deflection $w(l/2,t)$ of the track resting on the viscoelastic Winkler foundation and subjected to a concentrated harmonic load at its midpoint ($V=0m/s, \Omega=100rad/s$) is presented, performing either a linear or a nonlinear analysis and taking into account both rotary inertia and shear deformation effect. To illustrate the significant effect of the load frequency, in Table 2 the maximum values of the deflection $w(x,t)$ of the track resting on the nonlinear viscoelastic foundation, subjected to a moving with constant velocity $V=100m/s$ harmonic load are presented for various values of the excitation frequency Ω , performing either linear or nonlinear (for both cases of tensile or compressive axial load) analysis. Moreover, in Table 3 the maximum deflections and bending moments of the track are presented for different types of foundation reaction, for $\Omega=400rad/s, V=100m/s$, while in Fig. 3 the deflection curves $w(x,0.055)$ along the track axis at the time instant $t=0.055s$ as well as their maximum values are also presented for linear and nonlinear analysis for Winkler and Nonlinear viscoelastic foundation. From the obtained results, it is concluded that the discrepancy between the linear and the nonlinear analysis is not negligible and should not be ignored, while the influence of the shear deformation effect (increasing the transverse displacements

and decreasing the bending moments) in both linear and nonlinear analysis is observed. This latter influence is more pronounced as the length of the track becomes smaller.

Table 1: Geometric constants of the UIC60 rail track [8,16] and the foundation of the beam

$l(m)$	10	a_z	2.68
$E(GPa)$	210	$k_L(MPa)$	35
$G(GPa)$	77	$k_{NL}(MN/m^2)$	4×10^8
$I_y(m^4)$	30.55×10^{-6}	$k_p(kN)$	200
$A(m^2)$	76.86×10^{-4}	$c(kNs/m^2)$	145
$\rho(kg/m^3)$	7850	$P(kN)$	100

Table 2: Maximum values of the deflection $w \cdot 10^{-1}(mm)$ of the track, for various values of the excitation frequency Ω

$\Omega(rad/s)$	Linear	Nonlinear – Tensile Load	Nonlinear –Compressive Load
0.1	0.1377	0.1278	0.1511
0.5	0.6738	0.6278	0.7358
1.0	1.2708	1.1970	1.3770
5.0	3.8139	3.7529	3.9737
10	5.2252	5.1603	5.3181
50	5.7224	5.7052	6.7736
100	5.8639	5.8475	7.0310
200	5.7268	5.6447	6.5544
400	5.7828	5.6259	6.8505

Table 3: Maximum deflections w_{max} and bending moments $M_{y,max}$ of the track, for different types of foundation reaction

$w_{max} \cdot 10^{-1}(mm)$ $M_{y,max}(kNm)$	Without Shear Deformation		With Shear Deformation	
	Linear Analysis	Nonlinear Analysis	Linear Analysis	Nonlinear Analysis
Linear Winkler	9.879	14.336	9.973	14.436
	15.449	37.154	15.353	33.345
Linear and Nonlinear Winkler	5.788	6.859	5.923	6.992
	13.468	22.196	13.142	20.608
Pasternak	9.861	14.235	9.937	14.452
	15.422	36.917	15.312	33.123
3-Parameter	5.783	6.851	5.820	6.893
	13.439	22.132	13.127	20.557

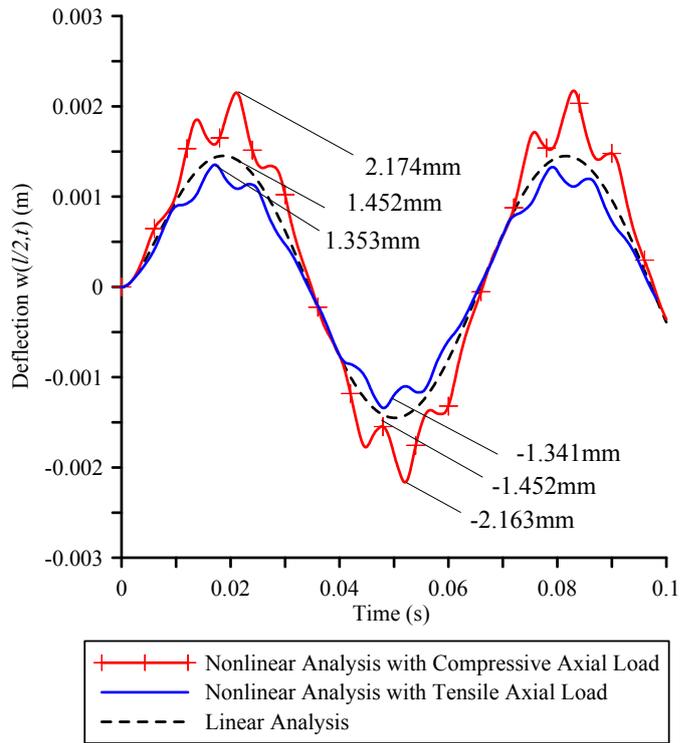


Figure 2: Time history and extreme values of the midpoint deflection $w(l/2, t)$ of the track

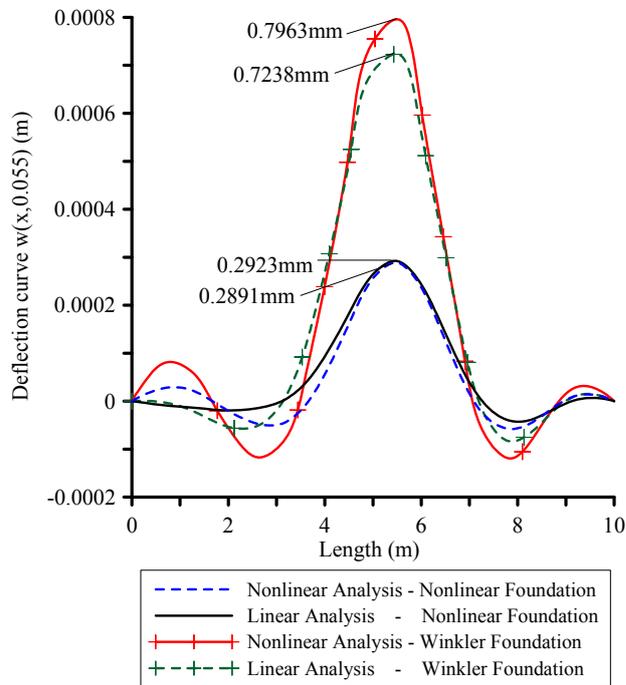


Figure 3: Deflection curves $w(x, 0.055)$, at the time instant $t = 0.055s$ and their maximum values

5 CONCLUDING REMARKS

The main conclusions that can be drawn from this investigation are

- a. The numerical technique presented in this investigation is well suited for computer aided analysis for beams of arbitrary simply or multiply connected cross section having at least one axis of symmetry.
- b. The proposed method is developed for general dynamic moving loading, while the beam is subjected to the most general boundary conditions and rests on a linear or nonlinear viscoelastic foundation.
- c. The lift up of the beam caused by the tensionless character of the foundation is observed, leading to magnification of the consequences of the dynamic response.
- d. In some cases, the effect of shear deformation is significant, especially for low beam slenderness values, increasing the transverse displacements and decreasing the bending moments in both linear and nonlinear analysis.
- e. The discrepancy between the results of the linear and the nonlinear analysis is remarkable.
- f. The response of the beam is strongly influenced by the linear and nonlinear parameters of the foundation reaction.
- g. The damping coefficient is of paramount importance for beams on viscoelastic foundations, as it reduces the vibration amplitude and the consequences of the dynamic response.

ACKNOWLEDGMENTS

The work of this paper was conducted from the “DARE” project, financially supported by a European Research Council (ERC) Advanced Grant under the “Ideas” Programme in Support of Frontier Research [Grant Agreement 228254].

REFERENCES

- [1] Krylov, A.N. Über die erzwungenen schwingungen von gleichformigen elastischen stäben. *Mathematische Annalen* (1905) **61**:211–234 (in German).
- [2] Timoshenko, S.P. Erzwungene schwingungen prismatischer stabe. *Zeitschrift für Mathematik und Physik*(1911) **59** (2): 163–203 (in German).
- [3] Inglis, C.E. *A mathematical treatise on vibration in railway bridges*. The Cambridge University Press, Cambridge, (1934).
- [4] Lowan, A.N. On transverse oscillations of beams under the action of moving variable loads. *Philosophical Magazine, Series 7* (1935) **19** (127): 708–715.
- [5] Koloušek, V. Dynamics of civil engineering structures-parti: general problems, second ed.-partii: continuous beams and frame systems, second ed.-partiii: selected topics, SNTL, Prague,1967, 1956,1961, *Dynamics in engineering structures*, Academia, Prague, Butterworth, London, (inCzech) (1973).
- [6] Frýba, L. *Vibration of solids and structures under moving loads*, Thomas Telford, Londo, (1999).
- [7] Kargarnovin, M.H. and Younesian, D. Dynamics of Timoshenko beams on Pasternak foundation under moving load. *Mechanics Research Communications* (2004) **31**: 713–723.

- [8] Kargarnovin, M.H., Younesian, D., Thompson, D.J and Jones, C.J.C. Response of beams on nonlinear viscoelastic foundations to harmonic moving loads. *Computers & Structures*(2005) **83**: 1865-1877.
- [9] Younesian, D. and Kargarnovin, M.H. Response of the beams on random Pasternak foundations subjected to harmonic moving loads. *Journal of Mechanical Science and Technology* (2009) **23**: 3013-3023.
- [10] Zehsaz, M., Sadeghi, M.H. and Ziaei Asl, A. Dynamic response of railway under a moving load. *Journal of Applied Sciences* (2009) **9**(8): 1474-1481.
- [11] Chen, Y.H., Huang, Y.H. and Shih, C.T. Response of an infinite Timoshenko beam on a viscoelastic foundation to a harmonic moving load. *Journal of Sound and Vibration* (2001) **241** (5): 809-824.
- [12] Kim, S.M. and Cho, Y.H. Vibration and dynamic buckling of shear beam-columns on elastic foundation under moving harmonic loads. *International Journal of Solids and Structures* (2006) **43**: 393-412.
- [13] Katsikadelis, J.T. The Analog Equation Method. A boundary-only integral equation method for nonlinear static and dynamic problems in general bodies. *Theoretical and Applied Mechanics* (2002) **27**: 13-38.
- [14] Timoshenko, S.P. and Goodier, J.N. *Theory of elasticity*. 3rd edn, McGraw-Hill, New York, (1984).
- [15] Cowper, G.R. The shear coefficient in Timoshenko's beam theory. *Journal of Applied Mechanics, ASME* (1966) **33**(2): 335-340.
- [16] Dahlberg, T. *Dynamic interaction between train and nonlinear railway model*. Proc. of Fifth Int. Conf. on Structural Dynamics, Munich, (2002).
- [17] Wu, T.X. and Thompson, D.J. The effects of track non-linearity on wheel/rail impact. *Proc. Inst. Mech. Eng. Part F: J Rail Rapid Transit* (2004) **218**: 1-12.
- [18] Iwnicky, S. *Handbook of railway vehicle dynamics*. Taylor and Francis, New York, (2007).
- [19] Stephen, N.G. Timoshenko's shear coefficient from a beam subjected to gravity loading. *Journal of Applied Mechanics, ASME* (1980) **47**: 121-127.
- [20] Sapountzakis, E.J. and Mokos, V.G. A BEM solution to transverse shear loading of beams. *Computational Mechanics* (2005) **36**: 384-397.
- [21] Thomson, W. T. *Theory of vibration with applications*. EnglewoodCliffs Prentice Hall, (1981).
- [22] Sapountzakis, E.J. and Katsikadelis, J.T. Elastic deformation of ribbed plates under static, transverse and inplane loading. *Computers and Structures* (2000) **74**: 571-581.

SOME CONSEQUENCES OF THE QUALITATIVE ANALYSIS OF THE POINT-SYMMETRIC COUPLED CONSOLIDATION MODELS

Emőke Imre^{*} and Stephen Fityus[†]

^{*} Szent Istvan University, Ybl Miklos Civil Eng. Fac. Budapest, Hungary
and BME, Geotechnical Dep., Budapest, Hungary
e-mail: imreemok@hotmail.com

[†] The School of Engineering, The University of Newcastle Callaghan, Newcastle, Australia
email: stephen.fityus@newcastle.edu.au

Key words: Coupled and uncoupled problems, immediate stress drop, constitutive law.

Summary. *The point-symmetric linear coupled consolidation models, known from the theory of the oedometric testing and from dissipation testing, can be summarized into a single mathematical model in the function of the embedding space dimension m ([1]). When a set of boundary conditions is specified equally for the 1, 2 and 3 dimensional models (i.e. oedometric, cylindrical and spherical models) then a family of related model: a “model-family” is obtained.*

Some inferences of the results of the qualitative analysis of two model-families are presented and discussed in this paper. These are (i) the similarity of the solution within a model-family, (ii) a direct proof that the uncoupled consolidation theories cannot be considered as a special case of the coupled consolidation theories and (iii) the interesting fact that an instantaneous dissipation may be predicted for the uniform initial pore water pressure distribution if the displacement is specified at both boundaries.

1 INTRODUCTION

Consolidation models are commonly employed in the theory of the oedometric testing and dissipation testing. These may differ in terms of the formulation of the constitutive equations and the assumed boundary conditions ([1]). In particular, models often differ in the nature of the condition specified at the outer boundary of the consolidating zone, which can be in terms of either a prescribed radial displacement (v) or a prescribed volumetric strain (ϵ).

In this paper, a single point-symmetric, coupled, linear consolidation constitutive model is considered in a generalised form that can be expressed as a function of the embedding space dimension, m . When the same set of boundary conditions is adopted for the 1, 2 and 3 dimensional models (i.e. oedometric, cylindrical and spherical models), then a family of related models, a “model-family”, is obtained. Two model families (referred to as coupled 1 and coupled 2) are considered here, based on different boundary condition assumptions. A summary of existing consolidation models and their classification within this framework is given in Table 1.

Table 1 Summary and classification of existing consolidation models.

V or ε boundary condition	1D point-symmetric (Oedometric models)	2D point-symmetric (Cylindrical pile models)	3D point-symmetric (Spherical pile models)
no (uncoupled)	Terzaghi [2]	Soderberg [5]	Torstensson [8]
v - v (coupled 1)	Imre [3]	Imre & Rózsa [6]	Imre & Rózsa [9]
v - ε (coupled 2)	Biot [4]	Randolph et al [7]	Imre & Rózsa [10]

2 MODEL

The basic units of the model-families (coupled 1 and 2), differing in one boundary condition, are presented in this section. The models are one dimensional with embedding space of various dimensions, m , as shown in Figure 1.

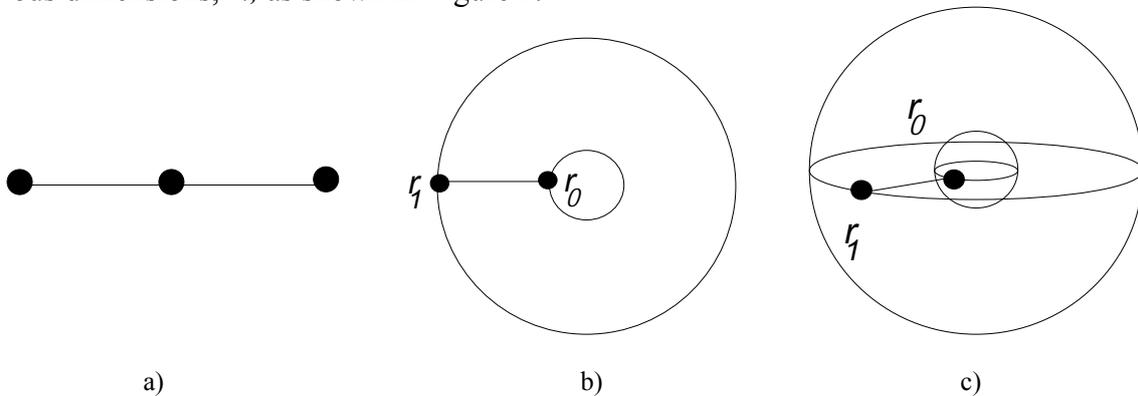


Figure 1. The displacement domain for the point-symmetric models bounded by a (a) 0 dimensional sphere, (b) 1 dimensional sphere, (c) 2 dimensional sphere.

2.1 System of differential equations

Two equations are derived ([1]): one from the equilibrium condition and one from the continuity condition. Equation (1) combines the equilibrium condition, the effective stress equality, the geometrical conditions and the constitutive equations:

$$E_{oed} \frac{\partial \varepsilon}{\partial r} - \frac{\partial u}{\partial r} = 0 \quad (1)$$

Equation (2) compiles the continuity equation, Darcy's law and the geometrical arrangement:

$$-\frac{k}{\gamma_w} \Delta u + \frac{\partial \varepsilon}{\partial t} = 0 \quad (2)$$

In equations (1) and (2), u is the pore water pressure (neglecting the gravitational component of the hydraulic head), ε is the volumetric strain, r and t are the space and the time co-ordinates respectively, k is the coefficient of permeability, γ_w is the unit weight of water, Δ is the Laplacian operator and E_{oed} is the oedometric modulus, which is expressed as:

$$E_{oed} = \frac{2G(1 - \mu)}{1 - 2\mu} = \frac{E(1 - \mu)}{(1 + \mu)(1 - 2\mu)}, \mu < 0.5 \quad (3)$$

In equation (3), G is the shear modulus, E is the Young's modulus, μ is the Poisson's ratio in terms of the effective stress σ' ($\sigma' = \sigma - u$ where σ is the total normal stress). The volumetric strain and the Laplacian operator, containing the dimension m of the embedding space, and expressed in terms of the radial displacement v , are as follows:

$$\varepsilon = \frac{1}{r^{m-1}} \frac{\partial}{\partial r} \left(r^{m-1} v \right) \quad (4)$$

$$\Delta = \frac{1}{r^{m-1}} \frac{\partial}{\partial r} \left(r^{m-1} \frac{\partial}{\partial r} \right) \quad (5)$$

2.2 Boundary conditions

The two model families considered here have three of their boundary conditions in common. These are (for $m=2$)

(1) The (common) *boundary condition #1*: zero pore pressure at distance r_1 .

$$u(t, r) |_{r=r_1} = 0 \quad (6)$$

(2) The (common) *boundary condition #2*: zero gradient in the pore pressure at distance r_0 .

$$\frac{\partial u(t, r)}{\partial r} |_{r=r_0} \equiv 0 \quad (7)$$

(3) The (common) *boundary condition #3*: constant, non-zero radial displacement at distance r_0 .

$$v(t, r) |_{r=r_0} \equiv v_0 > 0 \quad (8)$$

The coupled 1 and coupled 2 model families differ, however, in the nature of their fourth boundary condition. For the coupled 1 models, the condition is

(4a) *Boundary condition #4a* for the coupled 1 models: zero radial displacement at distance r_1

$$v(t, r) |_{r=r_1} \equiv 0 \quad (9)$$

(4b) *Boundary condition #4b* for the coupled 2 models: constant, non-zero volumetric strain at distance r_1

$$\varepsilon(t, r) |_{r=r_1} \equiv \varepsilon_1 > 0. \quad (10)$$

3 QUALITATIVE ANALYSIS

The transient part of the solutions, that is, the solution of the system of differential equations (1) and (2) with the homogeneous form of the boundary conditions, is qualitatively analysed in this section, without actually determining it.

3.1 Analysis of Equation (1)

Explicit expressions are derived for v and u by integrating the equilibrium equation (1) with respect to r subject to boundary condition #1:

$$u(t, y) = E_{oed} \varepsilon(t, r) - E_{oed} \varepsilon(t, r) \Big|_{r=r_1} \quad (11)$$

The boundary condition function can be determined by further integrating the Equation (11) with respect to r using boundary condition #3 and boundary condition #4a for the coupled 1 models:

$$\varepsilon(t, r) \Big|_{r=r_1} = - \left(\frac{\mathbf{u}_{mean}(t)}{E_{oed}} \right) \quad (12)$$

$$\mathbf{u}_{mean}(t) = \frac{r_0 \int_{r_1}^{r_1} r^{n-1} u(t, r) dr}{\int_{r_0}^{r_1} r^{n-1} dr} \quad (13)$$

Equation (11) can also be integrated with respect to r using boundary condition #3 and the homogeneous form of boundary condition #4b for the coupled 2 models, which results in the zero function.

Coupled 1 models

Using Equation (11) and the boundary condition function:

$$\varepsilon(t, r) = \frac{1}{E_{oed}} \left[u(t, r) - \mathbf{u}_{mean}(t) \right] \quad (14)$$

$$v(t, r) = \frac{I}{r^{n-1} E_{oed}} \left(\int_{r_1}^r r^{n-1} u(t, r) dr - \mathbf{u}_{mean}(t) \int_{r_1}^r r^{n-1} dr \right) \quad (15)$$

The initial condition functions for u and v^t have the following relationships:

$$u_0(r) = E_{oed} \mathcal{E}_0^t(r) - E_{oed} \mathcal{E}_0^t(r) \Big|_{r=r_1} \quad (16)$$

$$v_0^t(r) = \frac{I}{r^{n-1} E_{oed}} \left(\int_{r_1}^r r^{n-1} u_0(r) dr - u_{0,mean} \int_{r_1}^r r^{n-1} dr \right) \quad (17)$$

It is apparent that the $v_0^t(r)$ is the zero function when the initial pore water pressure function $u_0(r)$ is uniform.

Coupled 2 models

Using Equation (11) and the boundary condition function:

$$\varepsilon(t, r) = \frac{u(t, r)}{E_{oed}} \quad (18)$$

$$v(t, r) = \frac{1}{E_{oed} r^{n-1}} \int_{r_0}^r r^{n-1} u(t, r) dr \quad (19)$$

The initial condition functions for u and v^t have the following relationships:

$$u_0(r) = E_{oed} \mathcal{E}_0^t(r) \quad (20)$$

$$v_0^t(r) = -\frac{1}{r^{n-1} E_{oed}} \int_{r_0}^r r^{n-1} u_0(r) dr \quad (21)$$

3.2 Analysis of Equation (2)

By integrating Equation (2) twice with respect to r using the homogenous form of the boundary conditions #2 and #1, the following explicit expression is derived for the pore water pressure:

$$u(t, r) = \frac{k}{\gamma_v} \int_{r_1}^r \frac{1}{r^{n-1}} \int_{r_0}^x r^{n-1} \frac{\partial \varepsilon^t}{\partial t} dr dx \quad (22)$$

By further integration with respect to t between 0 and ∞ to give $A(r)$, the area of the subgraph of the dissipation curve $u(t, r)$ can be expressed for any fixed r as follows, for the coupled 1 and 2 models, respectively:

$$A(r) = \int_0^\infty u(t, r) dt = \frac{1}{c} \int_{r_1}^r \frac{1}{t^{n-1}} \int_{r_0}^x x^{n-1} [u_0(x) - u_{0,mean}] dx dt \quad (23)$$

$$A(r) = \int_0^\infty u(t, r) dt = \frac{1}{c} \int_{r_1}^r \frac{1}{x^{n-1}} \int_{r_0}^x x^{n-1} [u_0(x)] dx dt \quad (24)$$

It can be observed that the $A(r)$ is equal to 0 for the coupled 1 models if the initial pore water pressure function $u_0(r)$ is uniform. It follows that dissipation must be instantaneous in this case.

4 DISCUSSION

4.1 Constructing the analytical solution

The analytical solution for the total and effective stress components can be constructed using the analytical solution of the pore water pressure and the displacement or volumetric strain, on the basis of the constitutive equations and the explicit expressions derived from Equation (1). As an example, we consider the cylindrical case here. The constitutive equations valid for embedding space dimension $m=2$ ([7]) are

$$\sigma'_r = -\frac{2G}{1-2\mu} \left[(1-\mu)\varepsilon - (1-2\mu)\frac{\nu}{r} \right] \quad (25)$$

$$\sigma'_\phi = -\frac{2G}{1-2\mu} \left[\mu\varepsilon + (1-2\mu)\frac{\nu}{r} \right] \quad (26)$$

$$\sigma'_z = -\frac{2G\mu}{1-2\mu} \varepsilon \quad (27)$$

For the coupled 1 model, the transient components of the effective stresses are

$$\sigma'_r(t, r) = -\frac{1}{1-\mu} \left[(1-\mu)[u(t, r) - \mathbf{u}_{mean}(t)] - \frac{1-2\mu}{r^2} \left(\int_{r_1}^r ru(t, r) dr - \mathbf{u}_{mean}(t) \int_{r_1}^r r dr \right) \right] \quad (28)$$

$$\sigma'_\phi(t, r) = -\frac{1}{1-\mu} \left[\mu[u(t, r) - \mathbf{u}_{mean}(t)] + \frac{1-2\mu}{r^2} \left(\int_{r_1}^r ru(t, r) dr - \mathbf{u}_{mean}(t) \int_{r_1}^r r dr \right) \right] \quad (29)$$

$$\sigma'_z(t, r) = -\frac{\mu}{1-\mu} [u(t, r) - \mathbf{u}_{mean}(t)] \quad (30)$$

For the coupled 2 model, the transient components of the effective stresses are

$$\sigma'_r(t, r) = -\frac{1}{1-\mu} \left[(1-\mu)u(t, r) - \frac{1-2\mu}{r^2} \int_{r_0}^r ru(t, r) dr \right] \quad (31)$$

$$\sigma'_\phi(t, r) = -\frac{1}{1-\mu} \left[\mu u(t, r) + \frac{1-2\mu}{r^2} \int_{r_0}^r ru(t, r) dr \right] \quad (32)$$

$$\sigma'_z(t, r) = -\frac{\mu}{1-\mu} u(t, r) \quad (33)$$

4.2 Radial stress behavior

For the coupled 1 models, the transient component of the radial effective stress is

$$\sigma'_r(t, r) = -\frac{2G}{1-2\mu} \left[(1-\mu)\varepsilon - (n-1)(1-2\mu)\frac{\nu}{r} \right]$$

$$= -[u(t, r) - \mathbf{u}_{mean}(t)] + \frac{(n-1)}{(1-\mu)} \frac{1}{r^n} \left[\left(\int_{r_1}^r r^{n-1} u(t, r) dr - \mathbf{u}_{mean}(t) \int_{r_1}^r r^{n-1} dr \right) \right] \quad (34)$$

$$\sigma_r^t(t, r_0) = \mathbf{u}_{mean}(t) - u(t, r_0) \quad (35)$$

For the coupled 2 models, the transient component of the radial effective stress is

$$\sigma_r^t(t, r) = -\frac{2G}{1-2\mu} \left[(1-\mu)\varepsilon - (n-1)(1-2\mu)\frac{v}{r} \right] = -u(t, r) + \frac{(n-1)}{(1-\mu)} \frac{1}{r^n} \int_{r_0}^r r^{n-1} u(t, r) dr \quad (36)$$

$$\sigma_r^t(t, r_0) = -u(t, r_0) \quad (37)$$

During dissipation, the radial total stress at r_0 decreases with time by the initial value of the mean pore water pressure for the coupled 1 models, but it is constant for the coupled 2 models. The radial effective normal stress at r_0 increases during dissipation, by the difference between the initial pore water pressure and the initial mean pore water pressure for the coupled 1 models, and by the initial pore water pressure for the coupled 2 models.

4.3 A theoretical consequence of the analysis of Equation (1)

For certain boundary conditions, a coupled model with irrotational displacement may have the same pore water pressure solutions as an uncoupled model (which can be derived under the assumption that the total stress state is constant: see e.g. [11]). The additional boundary condition constraint adopted by [3] can be summarized as follows. The equilibrium equation (1) is integrated to give

$$E_{oed}\varepsilon - u = K(t) \quad (38)$$

Using this, equation (2) can be expressed in terms of u .

$$-\frac{k}{\gamma_v} \Delta u = \frac{\partial u}{E_{oed} \partial t} + \frac{\partial K(t)}{E_{oed} \partial t} \quad (39)$$

Equation (2) reduces to Terzaghi's uncoupled consolidation equation (except that instead of the bulk modulus the oedometric modulus occurs) if the term $K(t)$ is constant for one value of the space coordinate: a condition that is met for the coupled 2 model-family.. However, despite this, the uncoupled model cannot be considered as a special case of the coupled 2 model.

We can derive an explicit expression for the first invariant of the transient part of the total stress tensor for the coupled 2 models using Hooke's law, the effective stress equality and the explicit expression derived for the volumetric strain on the basis of Equation (1) in terms of the pore water pressure. It is

$$\sigma_1^t(t, r) = \sigma_1^d(t, r) + 3u(t, r) = E_{oed} \frac{1+\mu}{1-\mu} \varepsilon(t, r) + 3u(t, r)$$

$$= \left[-\frac{1+\mu}{1-\mu} + 3 \right] u(t,r) = 2 \frac{1-2\mu}{1-\mu} u(t,r) \quad (40)$$

This term is constant if μ (the Poisson's ratio in terms of the effective stress) is equal to 0.5 which is physically impossible in case of consolidation (it implies the soil is incompressible).

4.4 A theoretical consequence of the analysis of Equation (2)

Although the time for consolidation is infinite, using Equation (2), a finite-valued function can be derived in terms of the initial condition for each model-family to characterize the rate of the dissipation at any point r (i.e. as the area of the subgraph of the dissipation curve $u(t,r)$ for fixed r).

According to these expressions, for the coupled 1 models the dissipation time depends on the difference between the initial pore water pressure and the initial mean pore water pressure and it has two zeros. For the coupled 2 models, the dissipation time depends on the initial mean pore water pressure and it has one zero. It follows from the derived expression for the coupled 1 models that the consolidation from a uniform initial pore water pressure distribution is instantaneous since the initial transient volumetric strain is identically equal to zero in this case. By contrast, for the coupled 2 models, the greatest possible dissipation times are predicted for consolidation from a uniform initial pore water pressure distribution.

According to the Zeroth law of Thermodynamics ([12]) The transport of any extensive quantity implies the existence of an intensive quantity, the homogeneous distribution of which is a precondition for equilibrium, and according to the Second law, the movement of an extensive quantity is caused by the inhomogeneous distribution of the intensive quantity, which tends to be eliminated (Theorem 2 of Thermodynamics). In consolidation phenomena, the extensive variable for seepage is the water mass or volume. The intensive variable for seepage is the total hydraulic head of the water phase. These are related through

$$h = z + \frac{u}{\gamma_w} \quad (41)$$

where z is the vertical distance from an arbitrary datum. In the models presented here, the effect of z was neglected assuming that $h=u/\gamma_w$.

The instantaneous dissipation phenomenon indicates that the intensive variable of seepage cannot be the hydraulic head alone. The initial transient effective stress determined from the equilibrium equation gives a precise answer whether a seepage occurs or whether dissipation is instantaneous.

4.5 The stress drop phenomenon in some experimental tests

These results can be considered in the contexts of two types of experimental tests with (basically) constant displacement boundary conditions. One is the oedometric relaxation test where the total stress and the pore water pressure are measured after a fast load imposition. The second is the simple rheological-type cone penetrometer test where the local side friction and cone resistance are measured after the steady penetration ceases. In the latter case the effect of

a stress release, which occurs as the elastic deformation of the rod is recovered when steady penetration is stopped, makes the boundary condition approximate, as the rod diameter slightly decreases and the rod length slightly increases.

Some typical oedometric relaxation test results showing ‘instantaneous stress drop’ can be seen in Figure 2, where t' is a modified time variable indicating the time elapsed since loading ceased ([13]). The rate of the stress decrease is constant, and about equal to the rate of the load imposition, with both being limited by the mechanical characteristics of the constant power servo-system used in the test (which may allow some partial unloading, also).

Figure 3 shows some rheological-type cone penetrometer test results, sampled over a large area and averaged for the various soil types. These also display some instantaneous drop in stress ([14]).

According to the theoretical results, instantaneous dissipation may occur in the tests with constant displacement boundary conditions, (irrespective of the space dimension m) if the initial pore water pressure distribution is constant. The initial pore water pressure distribution may be considered to contain a constant component on the condition that the load imposition is very fast. However, very fast loading rates may invoke a slightly different, more pronounced time dependency in the rheological response that is not described by the constitutive laws derived for well-known models ([15], [16]). Rapid loading may induce a discontinuity when load imposition ceases (i.e. stress drop phenomenon), also. For example, fast initial stress drop during soil relaxation tests with fast, partly-drained load imposition was reported by Whitman [15].

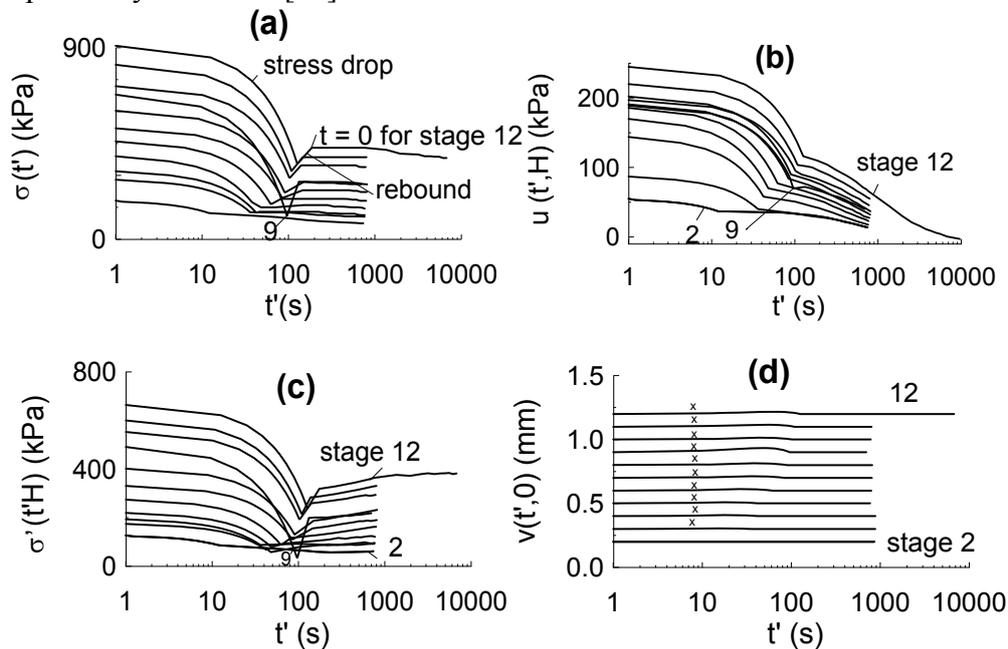


Figure 2. Oedometric relaxation data for a soft clay soil indicating the initial stress drop. (a) Measured total stress, (b) measured pore water pressure, (c) computed effective stress, (d) measured displacement [3]

The instantaneous dissipation response is not easily measured, since pore water pressure measurement is prone to error and the data measured during the stress drop cannot be considered as reliable for the oedometric relaxation test. The measured values are underestimated and there is delay in the measuring system ([17], [18]).

Two time dependent constitutive models that do accommodate some initial stress drop upon relaxation condition are Kelvin's model [20] and Leroueil et al's experimental model [16]. The reason for the sudden stress drop in the case of Kelvin's model is that the dashpot does not store energy, whereas in the case of Leroueil et al's model, it is because the strain rate of the sample changes discontinuously at the end of the loading period from a finite value to zero, while the mean strain is constant. These models fail to simulate the actual stress relaxation, since the stress is constant after the initial stress drop.

The compression curve constructed on the basis of multistage oedometric relaxation test is not the virgin compression curve unless the load imposition is fully drained. If the load imposition consists of a drained and undrained component, it can be assumed that the stress response relates to the drained part. According to some experimental results, the deviation is generally small for clays and may be considerable for freshly deposited loose silts. Larger deviation of the compression curve for loose soils can be experienced probably due to the fact that the effective stress may be negative within the sample (and local hydraulic fracturing may take place) if the preconsolidation pressure of the soil is small [19].

In the case of the rheological-type cone penetrometer test, it is very probable that an amount of stress drop also occurs due to the effect of stress release, as the elastic deformation of the rod is recovered and the rod diameter decreases when steady penetration ceases. However, this is not the case for the measured cone resistance.

In the light of the foregoing comments, the stress drop phenomenon can be attributed to

- the instantaneous dissipation of a constant component of the initial pore water pressure distribution if such component does exist,
- the change in the time dependency of the constitutive law,
- the fact that the soil may store a definite amount of energy,
- the stress release of the rod which modifies the boundary condition (in the case of the rheological type cone penetrometer tests).

5 CONCLUSIONS

The solution of the linear, point-symmetric, coupled consolidation models for two different sets of boundary conditions was qualitatively characterized. The same explicit expression could have been derived for the pore water pressure, strains or the displacements from Equation (1) for a given model-family.

From these, some explicit expressions can be derived for the total and effective stress components which are generally the same within a model family. It follows from this similarity that a point-symmetric problem with a specified embedding space dimension can be studied on the basis of a test related to different embedding space dimension.

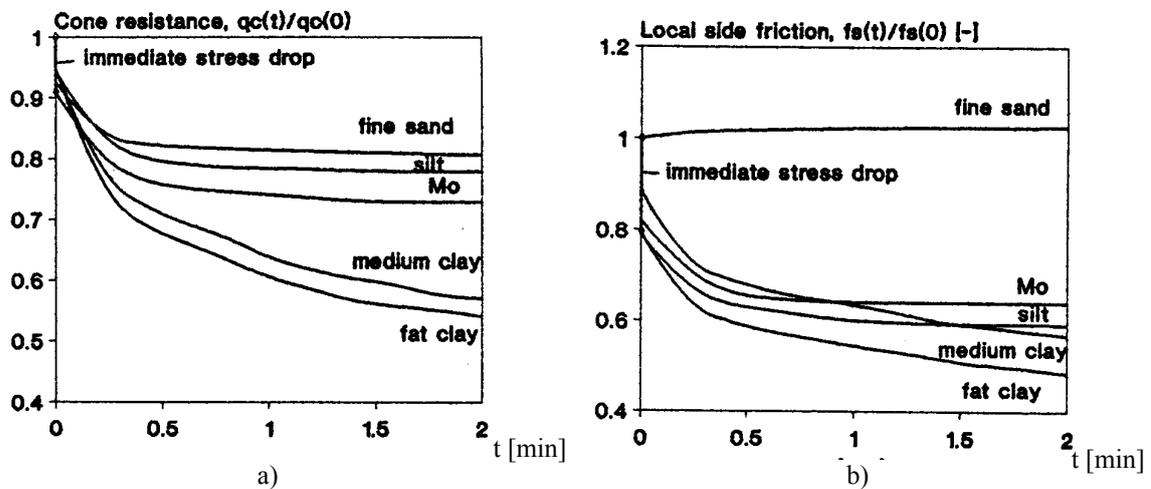


Figure 3. Results of simple rheological type CPT test indicating the immediate stress drop. (a) Local side friction vs. time data. (b) Cone resistance vs time data.

There are two special theoretical inferences coming from the qualitative analysis which can be concluded as follows. The pore water pressure solutions of uncoupled models and, the coupled 2 models are basically the same. However, the former cannot be derived from the latter, in the sense that the assumption of a constant total stress state of the uncoupled models cannot be verified.

In the case of the coupled 1 model-family, if the initial pore water pressure distribution is constant then the transient component of the initial volumetric strain is identically equal to zero, resulting in an instantaneous dissipation. This phenomenon indicates that the intensive variable of seepage, – which is a thermodynamic process – cannot be the total hydraulic head alone. If the distribution of the total hydraulic head is not uniform then the initial transient effective stress function, which can be determined from the equilibrium differential equation, gives a precise answer to whether there is a seepage or whether an instantaneous dissipation takes place.

It follows that some stress drop may theoretically occur for the tests with a constant displacement boundary condition if the initial pore water pressure distribution has a constant component. In a large number of tests with constant displacement boundary conditions, some kind of stress drop has been detected at the beginning of the test, which probably cannot be attributed to measurement errors, and its rate of decrease seems to be limited by the control system. However, further research is needed on the components and the cause of these stress drops since the measured records may entail some measuring errors.

REFERENCES

- [1] Imre, E. Rózsa, P. Bates L., Fityus, S “Evaluation of monotonous and non-monotonous dissipation test results” *Computers and Geotechnics*, (2010) 37:885–904.
- [2] Terzaghi, K. Die Berechnung der Durchlässigkeitsziffer des Tones aus dem Verlauf der hydrodyn. Spannungsercscheinungen, *Sitzber. Ak. Wiss. Wien, Abt.IIa*, (1923). Vol. 123.

- [3] Imre, E. Consolidation models for incremental oedometric tests. *Acta Tech. Acad. Sci. Hung.* (1997-1999) 369-398.
- [4] Biot, M. A. General Theory of Three Dimensional Consolidation. *Journal of Appl. Phys.* (1941). **12**: 155-164.
- [5] Soderberg, L. O. Consolidation Theory Applied to Foundation Pile Time Effects. *Geotechnique*, (1962).**12**: 217-232.
- [6] Imre, E. & Rózsa, P. Consolidation around piles. *Proc. of 3rd Seminar on Deep Foundations on Bored and Auger Piles*. Ghent (1998). 385-391.
- [7] Randolph, M. F. & Wroth, C. P. An analytical solution for the consolidation around displacement piles. *I. J. for Num. Anal. Meth. in Geomechanics*, (1979). **3**: 217-229.
- [8] Torstensson, B. A The pore pressure probe. (1977). Paper No. 34. NGI
- [9] Imre, E. and Rózsa, P. Modelling for consolidation around the pile tip. *Proc. of the 9th Int. Conf. on Piling and Deep Foundations (DFI)*, Nizza. (2002). 513-519.
- [10] Imre, E., Rózsa, P Point-Symmetric Consolidation Models for the Evaluation of the Dissipation Test. *11th IACMAG 2005*, Turin, Italy. (2005). 181-191.
- [11] Sills, G. C. Some conditions under which Biot's Equations of Consolidation Reduce to the Terzaghi's equation. *Geotechnique*, (1975). **25**: 129-132.
- [12] Adkins, C. J. *Equilibrium Thermodynamics*, McGraw-Hill, London. (1968).
- [13] Imre, E. Model discrimination for the oedometric relaxation test. *Proc. of 8-th Baltic Geotechnical Conference*, (1995). 55-60.
- [14] Imre, E., Tarcsai, Gy-né; Györffy, J.; Csizmás, F. Rheological tests with cone penetrometer. *Proc. of the 12th Inter. Conf. on Soil Mech. and Found. Eng.* Rio de Janeiro. (1989). **1**:239-242.
- [15] Whitman, R. V. The Behavior of Soils Under Transient Loading. *Proc. of the 3rd Inter. Conf. on Soil Mech. and Found.* (1957). **1**: 207-210.
- [16] Leroueil, S.; Kabbaj, M.; Tavenas, F.; Bouchard, R. Stress-strain rate relations for the compressibility of sensitive natural clays. *Geotechnique*, (1985). **35**: 159-175.
- [17] Perloff, W. H.; Nair, K.; Smith, J.G. Effects of Measuring System on Pore Water Pressures in the Consolidation Test. (1965). **43**: 338-342.
- [18] Gibson, R. E. An analysis of system flexibility and its effect on time-lag in pore-water pressure measurements. *Geotechnique*. (1963). 1-11.
- [19] Imre, E. Evaluation of quick multistage oedometric relaxation tests. *Proc. of the XIth Danube-European Conference on SMGE, Porec*. (1998). 695-702.
- [20] Findley WN, Lai, JS and Onaran K *Creep and relaxation of nonlinear viscoelastic materials: with an introduction to linear viscoelasticity*. General Publishing Company, Ontario. 369pp. (1989)

WATER RETENTION IN UNSATURATED SOILS SUBJECTED TO WETTING AND DRYING CYCLES

W. ARAIRO^{*}, F. PRUNIER[†] AND I. DJERAN-MAIGRE[†]

^{*} [†] Université de Lyon, INSA de Lyon
Laboratoire de Génie Civil et d'Ingénierie Environnementale (LGCIE)
Bâtiment Coulomb, 20 Avenue Albert Einstein, 69621 Villeurbanne Cedex, France
e-mail: wahib.arairo@insa-lyon.fr
irini.djeran-maigre@insa-lyon.fr
florent.prunier@insa-lyon.fr

Key words: Suction in soil, Hysteresis, Hydraulic cycle, Boundary curve, Scanning curve, Hydro-mechanical coupling.

Abstract. The suction is an essential parameter to describe and understand the behavior of unsaturated soils. The ability of unsaturated soils to retain water is quantified by determining the water retention curves (WRC), which express the hydraulic behavior of porous materials such as soil. These curves are determined by subjecting samples to several drying and wetting cycles. The curve during drying path is located above the wetting curve, developing a hysteresis phenomenon [1], and value of content water at a given suction value depends on the path used to reach this point.

The aim of this paper is to present a study on the hydraulic behavior of soil, water retention capacity due to drying and wetting cycles, pointing out the hydro-mechanical behavior of unsaturated soils.

In the first part, the effect of physical and mechanical properties of soil [32] (initial void ratio, particle size, cohesion, density...) on the water retention is presented.

In the second part, a complete numerical model was developed, based on the empirical model of Van Genuchten [18], to model the two boundary curves, and the experimental scanning data were best-fitted using the same theory of Mualem model [13]. This complete model requires 4 parameters.

This model has been validated with experimental data on different type of soils: sand [10], [34], U.S. Silica F-95 sand [30].

1 INTRODUCTION

The description and prediction of hydro-mechanical behavior of unsaturated soils require good knowledge of their hydraulic properties. One of these basic properties is the water retention curve (WRC), which connects the water content or the degree of saturation with suction. This relationship was used to estimate the hydraulic conductivity, volume change, and to estimate the aqueous diffusion functions of unsaturated soils ([1]; [2]; [3]). For a given matrix suction, water content in the drying curve is always higher than that in the wetting curve. In other words, soil follows different WRCs during a drying and a wetting process. This phenomenon is referred to as hysteresis. This hysteresis is typical of porous media consisting of interconnected pores of varying sizes, in which air is trapped. The contact angle between the water and the pore surface is greater in an advancing meniscus than in a receding meniscus. The hysteresis can also be caused by irregularities in the cross-sections of the void passages or the “ink-bottle” effect [4].

The term suction in geotechnical engineering has been defined in [5] as potential energy similar to the hydraulic head in saturated soils, corresponding to “the applied energy to carry free water from infinity to an unsaturated soil”.

In engineering practice, soil suction is composed of two components: matrix and osmotic suction [1]. The sum of matrix and osmotic suction is called total suction. Matrix suction comes from the capillarity, texture, and surface adsorptive forces of the soil. Osmotic suction arises from the dissolved salts contained in the soil water.

In Geotechnics, matrix suction consist the main part of total suction, and is usually expressed as a negative water pressure in KPa, the difference between air pressure U_a and water pressure U_w .

$$\psi = U_a - U_w \tag{1}$$

2 WATER RETENTION CURVES OF SOIL

The water retention curves are classified into 4 types: boundary drying curve, boundary wetting curve, wetting scanning curves and drying scanning curves.

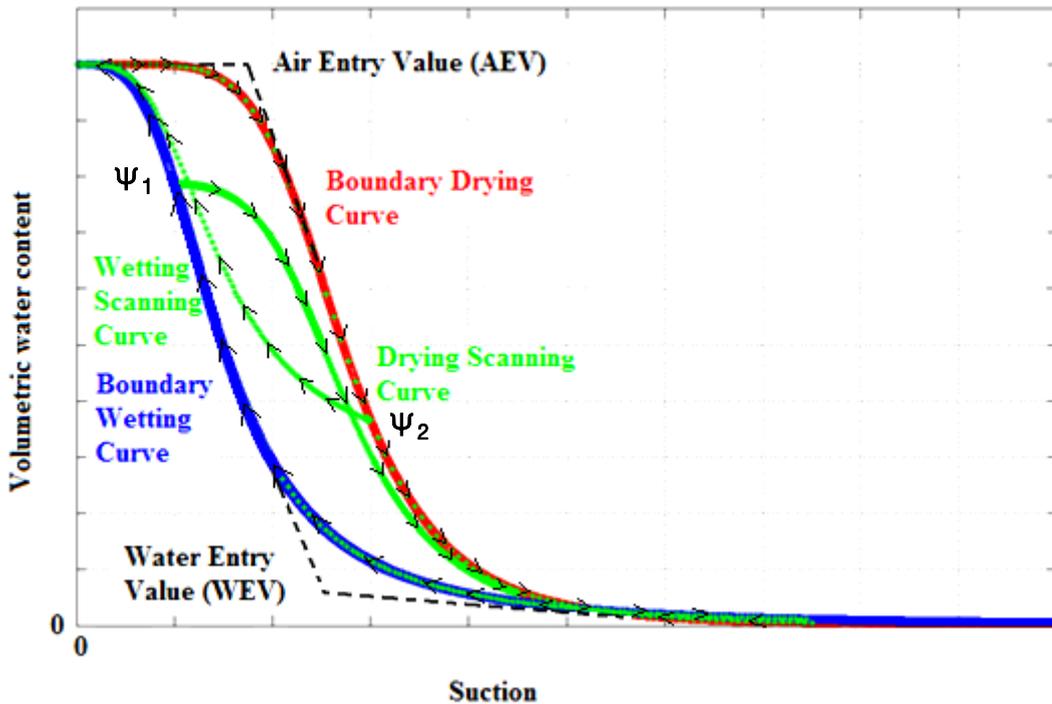


Figure 1: Schematic representation of Water Retention Curve (WRC) with the four curves: the two boundary curves for drying and wetting, and the scanning curves. There are two particular suction values: the air entry value and the water entry value.

On the boundary curves we have:

The air entry value (AEV): it is the higher suction at which the degree of saturation starts to decrease from full saturation at the boundary drying curve.

The Water Entry Value (WEV): it is the lower suction at which the degree of saturation starts to increase from residual water content at the boundary wetting curve.

2.1 Experimental setup for suction

Various experimental tools have been developed to deal with unsaturated soils, or to measure the value of matrix suction (measuring techniques), or to fix it for a test (control techniques) (after [5]). The principal measurement techniques are: Tensiometers, Psychrometers and the filter paper method (for more detail [5]). Most of these instruments have limitations with regard to range of measurement, equilibration times, and cost.

Therefore, there is a need for a method which can cover the practical suction range, be adopted as a basis for routine testing, and inexpensive.

Given the difficulty of measuring the suction, several numerical models have been developed to describe the sigmoid curve of water retention.

2.2 Numerical models

A lot of research has been conducted in the context of WRC; different approaches have been adopted to describe the hysteresis effect in these curves. The various models used to predict hysteretic WRCs can be classified into two categories: conceptual models (domain models) and empirical models.

The first group of conceptual models is based on the theory of independent domains developed by Néel [6] and used by other authors [7]; [8]; [9]. This theory attributes to the water in the soil two areas (group of pores): area without water and water-filled area. The Néel diagram assumes that the soil pore exists in one of two states; full of water or empty. The state of a pore can be characterized by two values of soil suction; namely, (i) drying soil suction, ψ_d ; and (ii) wetting soil suction, ψ_w . When soil suction increases to the drying soil suction of the pore, ψ_d , then the pore is fully drained. When soil suction decreases to the wetting soil suction of the pore, ψ_w , then the pore is filled. The first application of this theory was made by Poulouvasilis [10] and Topp [11], both models require all four curves for calibration. Mualem [12];[13];[14] still used the theory of fields, but with two curves for calibration. Other authors have proposed modifications to the theory of fields, to take into account interactions between domains. Several models, requiring more than two branches for calibration have been developed ([15]; [16]). The final model developed by Mualem [17] requires two branches for calibration: it uses a correction factor for variations in water content calculated on the basis of the independent domain theory.

The second group of models is related to empirical models. These models are based on analysis of the shape and properties of the curve.

A number of empirical models have been developed to describe the non-linear (sigmoid) WRC ([18]; [19]; [21]).

All these empirical models are derived from a general equation of the form:

$$a_1 \theta^{b_1} + a_2 \exp(a_3 \theta^{b_1}) = a_4 \psi^{b_2} + a_5 \exp(a_6 \psi^{b_2}) + a_7 \quad (2)$$

Where $a_1, a_2, a_3, a_4, a_5, a_6, a_7, b_1$ and b_2 are constants, ψ the matrix suction, and θ the volumetric water content.

Most of the researchers try to find equations describing the water retention curve using the simplest set of measurable parameters of soil solid phase such as particle size distribution, bulk density.

2.3 For boundary curves

Leong and Rahardjo [20] and Aubertin and Maksoud [21], Fredlund and Houston, and Fredlund and Sillers ([22]; [23]) found that the equation of Van-Genuchten [18] and that of Fredlund and Xing [19] models are most relevant to a variety of soils to model curves limits of water retention. This choice was made based on the Akaike information criterion.

The Van Genuchten equation can be written as:

$$\theta_v = \theta_r - (\theta_s - \theta_r) \left[1 + \left(\frac{\psi}{a_v} \right)^{n_v} \right]^{-m_v} \quad (3)$$

a_v is a parameter related to the air entry value (AEV), n_v is a parameter related to the variation of water content in the soil once the suction pressure exceeds the AEV, and m_v is related to residual water content θ_r . The index v in (a_v, n_v, m_v) is equal to w in wetting case and $v = d$ in drying case.

n_v and m_v can be related by: $m_v = 1 - \frac{k_m}{n_v}$ for $k_m = 1$ we have the VG-Mualem formulation for $k_m = 2$ we have the VG-Burdine formulation [18]. The Van-Genuchten equation can be expressed in term of degree of saturation:

$$S_{rv} = S_{res} - (S_{rmax} - S_{res}) \left[1 + \left(\frac{\psi}{a_v} \right)^{n_v} \right]^{-m_v} \quad (4)$$

S_{rv} is the degree of saturation corresponding to suction value ψ , S_{res} is the residual degree of saturation, S_{rmax} is the maximum degree of saturation.

Several authors ([24]; [32]) have correlated the parameters of Van-Genuchten with soil bulk density and with the content of soil organic matter, sand, silt and clay. Different numerical techniques were used to find correlations such as linear regression, nonlinear regression, and multivariate nonlinear optimization.

Other authors have estimated the water retention curve using parameters such as bulk density porosity, pore volume, and texture ([32]; [26]).

The differences between the values of Van-Genuchten parameters for the soils studied in [25] may be governed by differences in their physical and chemical properties. It seems that the particle size has the greatest influence on these parameters. Likewise, this study highlights the influence of compactness.

2.4 For scanning curves

Viane et al. [25] compared six models: (1) Mualem (1974), (2) Mualem (1977), (3) Hogarth et al. (1988), (4) Mualem (1984b), (5) Hanks et al. (1969), and (6) Scott et al. (1983), using the

data set of 7 soils. The authors, after a statistical analysis, found that the models of Mualem give comparable results and provide a good prediction of the scanning curve for models with two branches. Pham and Fredlund [31] have used in their comparison of the two models (Mualem (1974) and Mualem (1984)) are especially relevant to Viane et al. (1994) with three other models.

These studies demonstrated the relevance of the model of Mualem [13] for modeling the scanning curves.

Mualem's model can be drawn from scanning curves from the two limit curves.

The scanning drainage curve, which starts from suction ψ_1 on the limit curve of wetting, is given by:

$$\theta_d(\psi_1, \psi) = \theta_w(\psi) + \frac{[\theta_w(\psi_1) - \theta_w(\psi)]}{[\theta_u - \theta_w(\psi)]} [\theta_d(\psi) - \theta_w(\psi)] \quad (5)$$

Where $\theta_w(\psi)$ is the water content on the boundary wetting curve at suction, ψ ; $\theta_d(\psi)$ is the water content on the boundary drying curve at suction, ψ ; and θ_u is the water content at the meeting point of the two boundary curves at zero suction.

Similarly, the wetting scanning curve, which starts from suction ψ_2 on the drying boundary curve, is given by:

$$\theta_w(\psi_2, \psi) = \theta_w(\psi) + \frac{[\theta_u - \theta_w(\psi)]}{[\theta_u - \theta_w(\psi_2)]} [\theta_d(\psi_2) - \theta_w(\psi_2)] \quad (6)$$

The complete model is then formulated with the two models of Van-Genuchten and Mualem. The same formulation has been adopted for Mualem equations in term of degree of saturation.

3 FACTORS AFFECTING THE WATER RETENTION CURVE

A comprehensive description of the behavior of soils partially saturated in water should take into account:

- the hydraulic behavior (retention properties)
- the mechanical behavior (stress/strain relationship)
- the water flow in porous medium.

All three behaviors are known to be coupled together, and especially, the stress strain behavior is explicitly suction-dependent.

So the researches were conducted to show the different physical and mechanical factors affecting the water retention curve.

In this study, we present the main factor: the effect of void ratio, which is the result of mechanical behavior. There are many other factors; in this article we introduce only the main factor related to coupled hydro-mechanical behavior.

Effect of initial void ratio

Kawai et al. [26] studied the effect of the value of void ratio on the boundary Water Retention Curve (WRC). The pore water tends to migrate with increasing suction and when the value of the suction reaches the air entry value (AEV), the water begins to drain. The AEV reflects the magnitude of the capillary zone of saturation in the soil. The more the pores are large, the smaller AEV become. The AEV is inversely proportional to the logarithm of e (e is the void

ratio for the soil). When the soil is wetted, suction decreases and the degree of saturation increases, when the value of sucking decreases beyond a certain value called the Water Entry Value (WEV) (Fig. 1), an increase in the degree of saturation is more remarkable. In this context, Vanapalli et al. [27] published the results of water retention curves of a clay loam under different states of compaction.

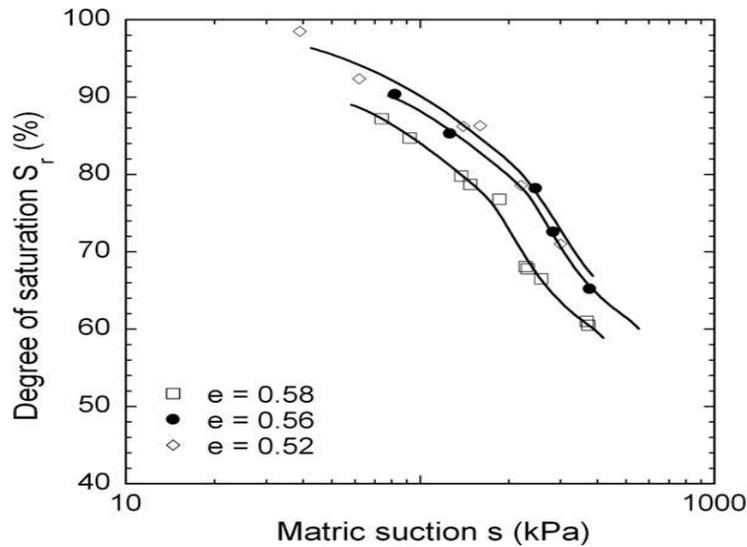


Figure 2: Water retention curves of compacted clay loam under different initial void ratio [27]

To introduce the mechanical coupling, the complete model should take into account the effect of the evolution of soil porosity on the boundary water retention curve.

4 COMPLETE MODEL TAKING INTO ACCOUNT THE VARIATION OF POROSITY

Recent discussions on stress frameworks for unsaturated soils [28]; [29] propose to scale down the analysis to the pore scale to understand the physical implications of saturation. The water retention curve shows that for the same net stress (defined as the difference between total stress and pore air pressure) and same amount of matrix suction, a specimen on a drying path reaches a higher degree of saturation than a sample under a wetting path.

Given that a difference in the degree of saturation induces a different repartition of water in the soil pores, modified proportions of bulk water and meniscus water are encountered according to the suction change following the wetting or the drying path.

It is agreed that suction within the meniscus part of the pore water acts only on the forces at inter-particle contacts whereas suction in bulk water is seen as an overall (isotropic) water pressure deficiency [33]. The skeleton effective stress could thus be identified as a function of external stresses, interstitial fluid pressures or suction, and repartition of pore water or degree of saturation. In other words, whenever a capillary hysteresis is identified in a soil water retention curve, it might significantly affect the state of effective stress within the soil.

In this study, the complete model is based on the Van-Genuchten equation (equation (3)) for boundary curves and on the use of Mualem equations for scanning curves (equations (5), (6)).

The Van-Genuchten equation is improved by adding the D_p parameter to take into account the effect of void ratio evolution on the boundary water retention curves. And D_p can be made a function of the void ratio, according to:

$$D_p = \psi_{ae0} * \lambda(e_0 - e)/e_0 \tag{7}$$

λ is a material coefficient, and ψ_{ae0} is the Air entry value at the referential void ratio e_0 , in the model this value is defined as the intersection of two lines:

$$\theta_d = \theta_s \tag{8}$$

And the tangent line on boundary drying curve at inflection point defined by:

$$\theta_d = (\theta_s + \theta_r)/2 \tag{9}$$

Hence, Eq. (3) will be modified to:

$$\theta_v = \theta_r - (\theta_s - \theta_r) \left[1 + \left(\frac{(\psi - D_p)}{a_v} \right)^{n_v} \right]^{-m_v} \tag{10}$$

In summary, two reference boundary curves (for an initial reference void ratio e_0) are expressed, where θ_w and θ_d are recalculated using the modified Van-Genuchten equation (10) $v = w$ in wetting case and $v = d$ in drying case. The parameter (a_v, n_v) of each curve are determined based on experimental data for a given soil, then the scanning curves are interpolated between the two boundary curves using Mualem equations (equations (5); (6)) depending on the evolution of suction value: the drying case is obtained for positive suction increment whereas the wetting case is obtained for negative suction increment.

Recent studies [28],[29] have adopted the same method of shifting for boundary curve with a linear relation between the air entry value and the void ratio.

This model has four parameters $(a_w, n_w; a_d, n_d)$. By a simple method, these parameters can be determined for each type of soil.

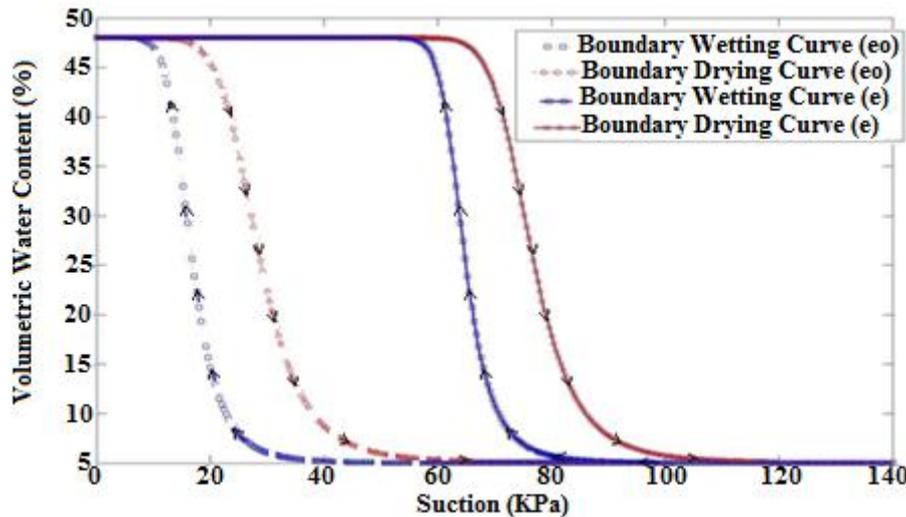


Figure 3: Shape of boundary retention curve with the variation of void ratio: change of void ratio induces a shift of the retention curve (the smaller void ratio, the higher air entry value)

5 VALIDATION OF NUMERICAL MODEL

A comparison of model predictions with experimental data is provided in order to further illustrate the capabilities of the model in simulating real unsaturated soil behavior.

The first experimental data considered are those presented by Lins et al. [34], from test on Hostun sand for which the parameters used in equation (4) are presented in Table 1, and $S_{rmax} = 1, S_{res} = 0$.

The prediction is shown in Fig. 4, from which it can be concluded that the model is reasonably accurate, although the lowest part of the wetting path was not quite exactly predicted. This is due to the number of simplifications adopted, particularly assuming that:

$$m_v = 1 - 1/n_v \tag{11}$$

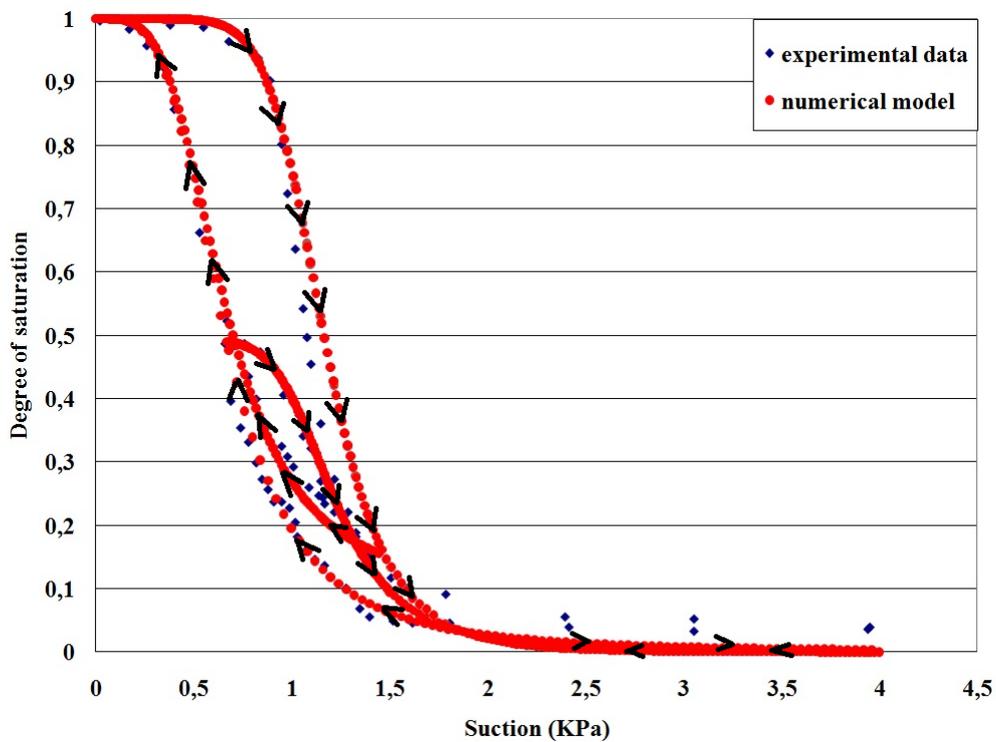


Figure 4: Validation of the proposed WRC model against experimental data from tests on Hostun sand (data after [35]).

The second set of experimental results considered are those presented by Muraleetharan et al. [30] and correspond to a fine poorly-graded sand named US Silica F-95 (Berkeley Springs, West Virginia). The results provided in Muraleetharan et al. [30] are expressed in terms of suction ψ against volumetric water content θ , defined as the volume of fluid divided by the total volume of a representative elementary volume (REV). This volumetric water content θ will be equal to the porosity n when the suction becomes zero. Fig. 5 shows the experimental data points and the model predictions of boundary drying and wetting WRCs and scanning

curves, from which it can be seen that the computed results are a reasonably accurate representation of the experimental data.

Table 1: Van-Genuchten parameters for Hostun sand(data after [34]).

Drying		Wetting	
a_d	n_d	a_w	n_w
1.1376	7.825	0.5963	3.987

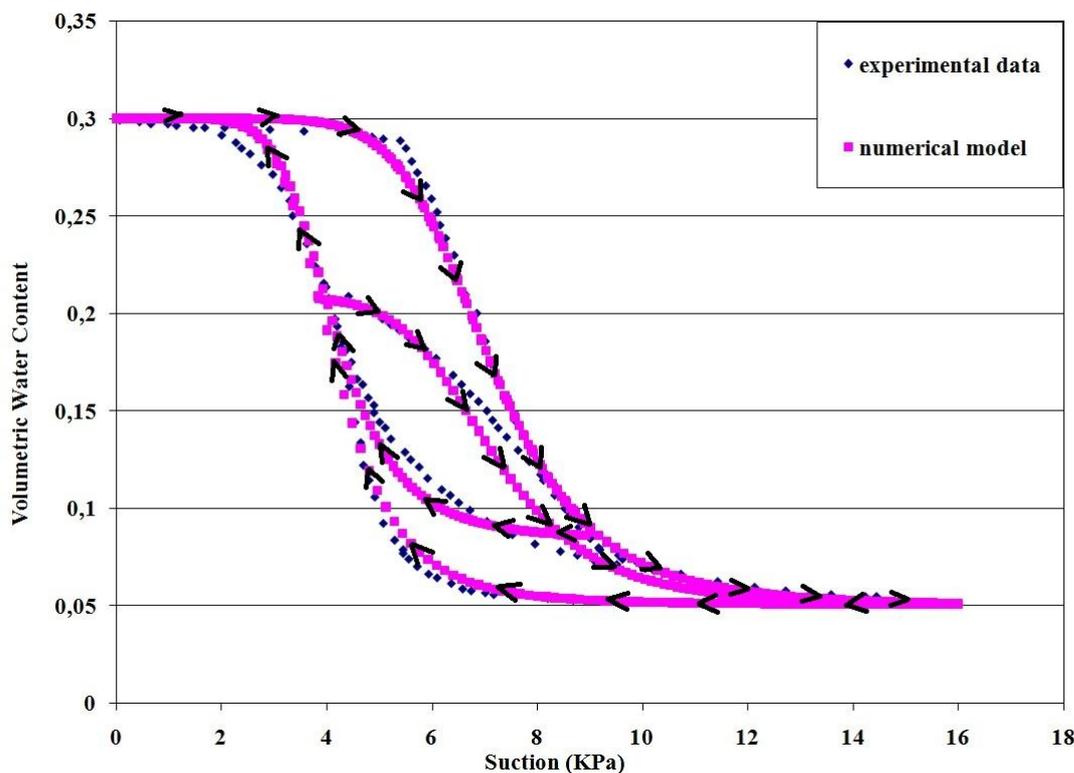


Figure 5: Validation of the proposed WRC model against experimental data from tests on US Silica F-95 fine sand (data after [31])

In order to further illustrate the versatility of the proposed model, for this second validation, the axes will be directly selected as $x = \psi$ and $y = \theta$, as were utilized in the coupled model of Muraleetharan et al. [30]. Now, $\theta_s = 0.3$ and $\theta_r = 0.05$ for the US Silica F-95 sand. The parameters for equation (3) are summarized in Table 2.

Table 2: Van-Genuchten parameters for US Silica F-95 sand (based upon test data after [30])

Drying		Wetting	
a_d	n_d	a_w	n_w
6.7613	8.168	3.9215	7.695

However, in the literature, to the knowledge of the author of this article, there’s no a complete experimental data, for water retention curves with the coupled hydro mechanical effect on those curves, in other terms, there’s no experimental data of hydraulic loading cycles (the four

water retention curves) taking into account the effect of the void ratio. A new experimental setup has been developed, in our laboratory, for measuring the water retention curve and controlling the porosity of the sample subjected to suction cycles.

6 CONCLUSION

A complete representation of the WRC with hysteresis behavior of unsaturated soils is presented, involving a simple series of equations. The great advantage of this technique is its easy application to experimental saturation–suction data for a given soil. The proposed model for water retention curve introduces two reference boundary curves corresponding to a reference void ratio e_0 .

The coupled model has eight parameters: four hydraulic calibration parameters: (a_w ; n_w ; a_d ; n_d), one mechanical parameter (λ) to introduce the effect of void ratio evolution on water retention curve, and three physical parameters (θ_r ; θ_s ; e_0). The first five parameters are calibrated by wetting and drying experimental curves and the last three are directly measured by characterization experiments.

The model presented here deals with the hydro-mechanical behavior and takes into account the deformation characteristics of soils, such as the influence of either the initial or current void ratios. The implementation of the resulting equations into fully hydro-mechanical coupled models, with the stress/strain relation for numerical analyses, is straightforward though, using an effective stress definition, such as Bishop equation.

The model is efficient with two different data sets [31],[35] introducing only the retention behavior. Experiments are in progress to be able to catch the void ratio variation.

REFERENCES

- [1] Fredlund, D.G., and Rahardjo, H. *Soil mechanics for unsaturated soils*. John Wiley & Sons, New York, (1993).
- [2] Vanapalli, S.K., Fredlund, D.G., Pufahl, D.E., and Clifton, A.W. *Model for the prediction of shear strength with respect to soil suction*. Canadian Geotechnical Journal, 33: 379–392 (1996).
- [3] Barbour, S.L. *The soil-water characteristic curve: a historical perspective*. Nineteenth Canadian Geotechnical Colloquium, Canadian Geotechnical Journal, 35: 873–894 (1998).
- [4] Bear, J. *Dynamics of fluids in porous media*. Elsevier, Amsterdam (1969).
- [5] Delage P., Cui Y. *L'eau dans les sols non saturés*. Technique de l'ingénieur C301 (2000).
- [6] Néel, L. *Théorie des lois d'aimantation de Lord Rayleigh*, 1. Cahiers de Physique, Vol.:12, 1–20, (1942).
- [7] Everett, D.H. *A general approach to hysteresis – part 3: a formal treatment of the independent domain model of hysteresis*. Transactions of the Faraday Society, 50: 1077–1096, (1954).
- [8] Everett, D.H. 1955. *A general approach to hysteresis – part 4: an alternative formulation of the domain model*. Transactions of the Faraday Society, 51: 1551–1557, (1955).

- [9] Enderby, A.J. 1955. *The domain model of hysteresis*. Transactions of the Faraday Society, 51: 835–848, (1955).
- [10] Pouloussis, A. *Hysteresis of pore water in granular porous bodies*. Journal of Soil Science, 109: 5–12, (1970).
- [11] Topp, G.C. 1969. *Soil–water hysteresis measured in a sandy loam compared with the hysteretic domain model*. Soil Science Society of America Proceedings, 33: 645–651, (1969).
- [12] Mualem, Y. *Modified approach to capillary hysteresis based on a similarity hypothesis*. Water Resources Research, 9: 1324–1331, (1973).
- [13] Mualem, Y. *A conceptual model of hysteresis*. Water Resources Research, 10: 514–520, (1974).
- [14] Mualem, Y. *Extension of the similarity hypothesis used for modeling the soil water characteristics*. Water Resources Research, 13: 773–780, (1977).
- [15] Pouloussis, A., and Childs, E.C. *The hysteresis of pore water: The non-independence of domains*. Journal of Soil Science, 112: 301–312, (1971).
- [16] Topp, G.C. *Soil water hysteresis in silt loam and clay loam soils*. Water Resources Research, 7: 914–920, (1971b).
- [17] Mualem, Y. *Prediction of the soil boundary wetting curve*. Journal of Soil Science, 137: 379–390 (1984).
- [18] Van Genuchten, M.T. *A closed-form equation for predicting the hydraulic conductivity of unsaturated soils*. Soil Science Society of America Journal, 44: 892–898, (1980).
- [19] Fredlund, D.G., and Xing, A. *Equations for the soil-water characteristic curve*. Canadian Geotechnical Journal, 31: 521–532, (1994).
- [20] Leong E. C. and Rahardjo H. *Review of soil-water characteristic curve equations*. Journal of geotechnical and geoenvironmental engineering, vol.123 No 12:1106-1117, (1997).
- [21] Aubertin, M. and Maksoud, A. *Hysteresis effects on the water retention curve: a comparison between laboratory results and predictive models*. 57th Canadian Geotechnical Conference, Session 3A: 8-15 (2004).
- [22] Fredlund, D.G. and Houston S.L., *Protocol for the assessment of unsaturated soil properties in geotechnical engineering practice*. Can. Geotech. J. 46: 694–707 (2009).
- [23] Fredlund D.G., Sillers W.S. *Statistical assessment of soil-water characteristic curve models for geotechnical engineering*. Canadian Geotechnical Journal. 38: 1297–1313, (2001).
- [24] Porêbska D., Walczak R.T. *Relationship between van Genuchten's parameters of the retention curve equation and physical properties of soil solid phase*. Int. Agrophysics, Vol.: 20, 153-159, (2006).
- [25] Viane, P., Vereecken, H., Diels, J. and Feyen, J. *A statistical analysis of six hysteresis models for the moisture retention characteristic*. Journal of Soil Science, 157: 345–355, (1994).
- [26] Kawai, K., Karube, D. and Kato, S. *The model of water retention curve considering effects of void ratio*. In Proceedings of Asian Conference on Unsaturated Soils, Singapore. Edited by H. Rahardjo, D.G. Toll, and E.C Leong. A.A. Balkema, Rotterdam, The Netherlands. pp. 329–334 (2000).

- [27] Vanapalli, S.K., Pufahl, D.E. and Fredlund, D.G. *The influence of soil structure and stress history on the soil-water characteristic of compacted till*. Geotechnique, 49(2):143-159, (1999).
- [28] Gens, A. and Sheng, D. *On constitutive modeling of unsaturated soils*. Acta Geotechnica 1:137-147, (2006).
- [29] Nuth M, and Laloui L. *Advances in modelling hysteretic water retention curve in deformable soils*. Computers Geotechnics; 35: 835–44, (2008).
- [30] Muraleetharan K.K., Liu C., Wei C., Kibbey T.C.G. and Chen L. *An elastoplastic framework for coupling hydraulic and mechanical behavior of unsaturated soils*. International Journal of Plasticity; 25:473–90, (2009).
- [31] Fredlund D.G. and Pham H.Q. *A study of hysteresis models for soil-water characteristic curves*: Can. Geotech. J. 42: 1548–1568, (2005).
- [32] Rahardjo H., Fredlund D.G. *Factors affecting drying and wetting soil-water characteristic curves of sandy soils*. Can. Geotech. J. 41: 908–920, (2004).
- [33] Wheeler SJ, Sharma RS, Buisson MSR. *Coupling of hydraulic hysteresis and stress-strain behaviour in unsaturated soils*. Geotechnique; 53(1):41–54, (2003).
- [34] Lins Y, Zou Y, Schanz T. *Physical modelling of SWCC for granular materials. Theoretical and numerical unsaturated soil mechanics*, Weimar, Germany; p. 61–74, (2007).

A MULTI-MODEL INCREMENTAL ADAPTIVE STRATEGY TO ACCELERATE PARTITIONED FLUID-STRUCTURE ALGORITHMS USING SPACE-MAPPING

Thomas P. Scholcz*, Alexander H. van Zuijlen and Hester Bijl

*Faculty of Aerospace Engineering
Delft University of Technology
Kluyverweg 1, 2629 HS Delft, The Netherlands
e-mail: t.p.scholcz@tudelft.nl

Key words: Fluid-structure interaction, Partitioned coupling, Space-mapping, Multi-fidelity models, Reduced Order Models

Abstract. High fidelity analysis of fluid-structure interaction systems is often too time-consuming when a large number of model evaluations are required. The choice for a solution procedure depends often on the efficiency of the method and the possibility of reusing existing field solvers. Aggressive Space-Mapping, a technique originally developed for multi-fidelity optimization, is applied to accelerate the partitioned solution procedure of a high fidelity fluid-structure interaction model. The method supports software modularity. Aggressive Space-Mapping (ASM) is applied to an academic testcase and the results are compared with the corresponding Incremental Quasi-Newton (IQN) method. An efficiency metric is defined to facilitate the comparison. The ASM method is found to be more efficient than the corresponding IQN method for the testcases considered. The efficiency of space-mapping increases with increasing fluid-to-structure mass ratio, indicating that the method is especially useful for strongly coupled problems.

1 INTRODUCTION

High-fidelity analysis of fluid-structure interaction systems is often too time-consuming when a large number of model evaluations are required. Examples are found in design, optimization and stochastic analysis of fluid-structure interaction systems [1, 2, 4].

The aim of this contribution is to efficiently obtain transient solutions of a high-fidelity model using a partitioned procedure in combination with a technique from multi-fidelity optimization called Aggressive Space-Mapping (ASM) [1]. Using ASM, information of a cheap low-fidelity model is exploited to accelerate the solution procedure of an expensive high-fidelity model. In the following the cheap low fidelity model is named “the coarse model” and the expensive high-fidelity model “the fine model”. Coarse models can be categorized as [2]

1. Data-fit models: response surfaces, kriging, radial basis functions etc.
2. Reduced order models: Proper Orthogonal Decomposition, modal analysis, Volterra series etc.
3. Hierarchical models: physics-based models of lower fidelity.

In turn, hierarchical models can be categorized as

1. Low-fidelity models that neglect some physics modeled by the high-fidelity models.
2. Low-fidelity models that are the same as the high fidelity model, but converged to a higher residual tolerance.
3. Low-fidelity models that are the same as the high fidelity model, but discretized on a coarser grid or using a lower order discretization method.

Using defect correction with coarse models in the latter category emanates in a wide range of methods known as multi-grid or coarse-grid methods. The application of coarse-grid methods on a fluid-structure interaction problem has been thoroughly investigated in [3]. The observed efficiency gains of coarse-grid methods motivates to explore the application of other defect correction based algorithms such as space-mapping.

An important distinction can be made in the way coarse models are derived from fine models. For the derivation of a large class of coarse models, detailed preliminary problems need to be solved first. Examples are the classical POD and the Volterra series model reduction methods. This is called the a posteriori approach. The a posteriori approach is useful when the coarse model is able to capture variations in model parameters. In that case, it can be used to replace the fine model in design, optimization or uncertainty analysis. On the other hand, there exist the a priori approach [7] which does not assume prior knowledge of the fine model solution but either starts from a model that was initially derived from the fine model and/or is improved during the solution procedure of the fine model.

The advantage of improving the coarse model and solving the fine model simultaneously is that coarse model information can be used to accelerate the solution procedure of the fine model. Examples of such hybrid strategies are found in [5, 6]. Here, a reduced order model is built up during the coupling iterations of a partitioned (implicit) fluid-structure algorithm. The reduced order model is subsequently used to enhance the convergence of fixed point iterations. Model adaptation is performed by enriching the ROM basis from input-output information of the fine model.

Other techniques use reduced order models to provide a better initial guess for the iterative process at the next time step in an implicit time-integration scheme. Examples are found in [8, 9]. Here, a POD-based ROM is built up during the integration of the fine model. Model adaptation is performed by changing the ROM basis using a so-called

incremental eigenspace algorithm. Typically a ROM-prediction is followed by a fine model correction, if necessary.

In this contribution we aim to explore the application of space-mapping to accelerate the solution procedure of the fine model. Model adaptation is performed via an adaptively improved inverse space-mapping function during the coupling iterations at each time step in a partitioned fluid-structure algorithm. Defect correction on the fluid-structure interface is employed to find the new iterate. Using this strategy, any coarse model could be used that shares the interface degrees of freedom with the fine model. The fine and coarse models are considered “black-boxes“ in the space-mapping approach. Similar to the method in [9], the solution of the coarse model can be interpreted as a “shadow” that runs parallel with the solution of the fine model.

First, the general fluid-structure problem is formulated followed by a short discussion of the most common coupling techniques. The concept of space-mapping is explained and the resulting algorithm is applied to a 1-D testcase. Only physics-based models of lower fidelity are used as coarse models in the space-mapping algorithm. Other coarse model types can be used but are not considered in this contribution. Finally, the results are presented and conclusions are drawn.

2 Partitioned coupling techniques

Implicit time integration is often used to ensure a stable numerical scheme. Since implicit schemes require matching conditions at the new time instant t^{n+1} , these schemes lead to coupled problems at each time step in the simulation. Let the vector \mathbf{v} denote the flow variables and vector \mathbf{u} the structure variables at the new time level t^{n+1} . Hiding the dependency on the solution of previous time levels, the coupled problem at time step t^{n+1} is formulated as

$$\mathbf{f}(\mathbf{v}, \mathbf{u}) = \mathbf{0}, \tag{1}$$

$$\mathbf{s}(\mathbf{v}, \mathbf{u}) = \mathbf{0}, \tag{2}$$

where \mathbf{f} represents the discrete fluid equations and \mathbf{s} the discrete structure equations. Solving the discrete fluid equations is often much more expensive than solving the discrete structure equations due to the large range of important scales present in the fluid.

2.1 Problem formulation

The problem can be formulated as a problem in the degrees of freedom of the fluid-structure interface only, with the flow variables \mathbf{v} and structure variables \mathbf{u} treated as internal variables of the residual operator [5, 6]. A Dirichlet-Neumann decomposition is normally employed to solve the system in a partitioned fashion. Introducing the fluid operator $F()$ and structure operator $S()$:

$$\mathbf{y} = F(\mathbf{x})$$

$$\mathbf{x} = S(\mathbf{y})$$

- | | |
|---|--|
| <ol style="list-style-type: none"> 1. Apply the interface displacement \mathbf{x} to the boundary of the fluid domain. 2. Deform the grid in the fluid domain to the displacement of the boundary. 3. Calculate the flow variables \mathbf{v} in the fluid domain. 4. Obtain the stress \mathbf{y} on the boundary | <ol style="list-style-type: none"> 1. Apply the stress \mathbf{y} on the boundary of the structure. 2. Calculate the structural variables \mathbf{u} in the domain of the structure. 3. Obtain the displacement \mathbf{x} on the boundary of the structure, |
|---|--|

It follows by substitution of $\mathbf{y} = F(\mathbf{x})$ in $\mathbf{x} = S(\mathbf{y})$ that

$$\underbrace{S \circ F(\mathbf{x}) - \mathbf{x}}_{\mathbf{R}(\mathbf{x})} = \mathbf{0}. \quad (3)$$

Strong coupling algorithms aim to minimize residual \mathbf{R} as far as possible using a minimal number of residual evaluations at each time step in the sequential time integration process. The kinematic and dynamic interface conditions at the fluid-structure interface are satisfied when equation (3) holds. For an overview of strong coupling procedures, see [10, 11]. As mentioned in [11], the choice of partitioned approach often depends on the possibility of reusing existing field solvers. The strategy presented in this work supports software modularity and is in addition modular with respect to the coarse models used to accelerate the solution procedure of the fine model.

2.2 Classical coupling algorithms

Classical algorithms aim to reduce residual \mathbf{R} directly without the use of a coarse model space. The most common coupling algorithms in this category that support software modularity are found in [10, 11]. These are: the fixed-point iteration method, fixed-point iteration method with Aitken acceleration and Incremental Quasi-Newton (IQN) methods. Space-mapping is introduced in section 2.3. It is shown in section 2.3.2 that the resulting ASM algorithm is related to the IQN algorithm, which motivates a comparison between the two algorithms. This is done by performing numerical experiments in section 3.

2.3 Space-mapping

Space-mapping relies on the availability of an expensive fine model and a cheap coarse model that model the same physical phenomena. The use of the fluid operator is the most expensive operation in most fluid-structure interaction problems. For this reason we assume the availability of a fine and coarse fluid operator, resulting in the definition of the fine and coarse residual \mathbf{R} and $\tilde{\mathbf{R}}$:

Fine fluid model :

1. Evaluation of $F()$ expensive.
2. Fine residual: $\mathbf{R}(\mathbf{x}) = S \circ F(\mathbf{x}) - \mathbf{x}$
3. Solution: $\mathbf{x}^* = \arg \min \|\mathbf{R}(\mathbf{x})\|$
4. Accurate solution.

Coarse fluid model :

1. Evaluation of $\tilde{F}()$ cheap.
2. Coarse residual $\tilde{\mathbf{R}}(\mathbf{z}) = S \circ \tilde{F}(\mathbf{z}) - \mathbf{z}$
3. Solution $\mathbf{z}^* = \arg \min \|\tilde{\mathbf{R}}(\mathbf{z})\|$
4. Inaccurate solution.

The solution of the fine and coarse model at t^{n+1} are denoted by \mathbf{x}^* and \mathbf{z}^* .

2.3.1 Space-mapping function

Let \mathbf{X} be the space of all interface displacements that can be reached by the fine model and let \mathbf{Z} be all interface displacements that can be reached by the coarse model. The concept of space-mapping is shown in figure 1.

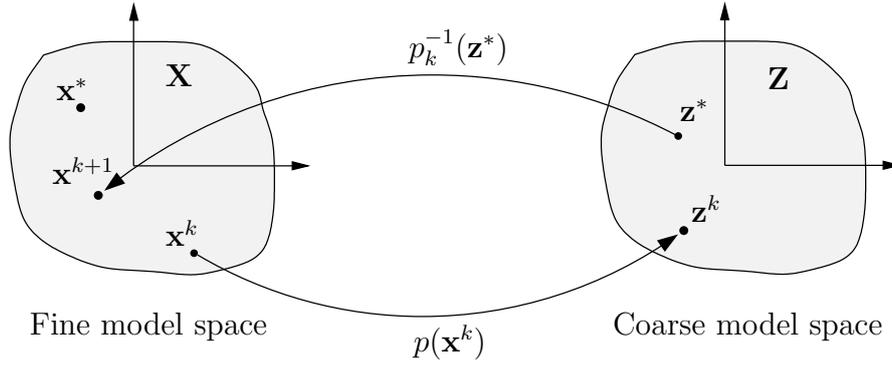


Figure 1: Concept of space-mapping

A choice for the space-mapping function could be

$$\mathbf{z} = p(\mathbf{x}) = \arg \min_{\mathbf{z} \in \mathbf{Z}} \|\tilde{\mathbf{R}}(\mathbf{z}) - \mathbf{R}(\mathbf{x})\|, \quad (4)$$

which can be evaluated using coupling iterations with the coarse model, e.g by fixed-point iterations

$$\mathbf{z}^{q+1} = S \circ \tilde{F}(\mathbf{z}^q) - \mathbf{R}(\mathbf{x}). \quad (5)$$

The space-mapping function has the following property:

$$p(\mathbf{x}^*) = \mathbf{z}^*, \quad (6)$$

which states that the mapping function p is perfect [1]. In words: the space-mapping function maps the fine model solution \mathbf{x}^* to the solution of the coarse model, \mathbf{z}^* . The inverse of the space-mapping function is given by

$$\mathbf{x} = p^{-1}(\mathbf{z}) = \arg \min_{\mathbf{x} \in \mathbf{X}} \|\mathbf{R}(\mathbf{x}) - \tilde{\mathbf{R}}(\mathbf{z})\|. \quad (7)$$

Evaluation of the inverse space-mapping function in (7) is as hard as solving the fine model directly. Therefore, an approximation is used at iterate number k :

$$\mathbf{x}^{k+1} = p_k^{-1}(\mathbf{z}^*) \quad (8)$$

Due to the fact that $p_k^{-1} \approx p^{-1}$ we have $\mathbf{x}^{k+1} \neq \mathbf{x}^*$ but as $p_k \rightarrow p^{-1}$ for increasing k it holds that $\mathbf{x}^{k+1} \rightarrow \mathbf{x}^*$ upon convergence. When a Taylor series approximation is used for the inverse space-mapping function, the ASM method results. This is the topic of the next section.

2.3.2 Aggressive Space-Mapping method

The space-mapping function is expanded in a Taylor series

$$p_k(\mathbf{x}) \approx p(\mathbf{x}^k) + \nabla_{\mathbf{x}}^k p(\mathbf{x} - \mathbf{x}^k). \quad (9)$$

Using the relations $\mathbf{z}^k = p(\mathbf{x}^k)$ and $\mathbf{z} \approx p_k(\mathbf{x})$, the approximation of the inverse space-mapping function is found from equation (9)

$$p_k^{-1}(\mathbf{z}) \approx \mathbf{x}^k + (\nabla_{\mathbf{x}}^k p)^{-1}(\mathbf{z} - \mathbf{z}^k). \quad (10)$$

The new iterate is found by substitution of the approximated inverse space-mapping function from equation (10) in equation (8)

$$\mathbf{x}^{k+1} = \mathbf{x}^k + (\nabla_{\mathbf{x}}^k p)^{-1}(\mathbf{z}^* - \mathbf{z}^k). \quad (11)$$

The ASM method is summarized in algorithm 1 and compared with the IQN method in algorithm 2. Both algorithms use Broyden's first method to update the Jacobian. Other Jacobian approximations could be used as long as input/output information is sufficient in order to obtain the approximation, e.g. the method of Vierendeels and Degroote [5]. Line number 4 and 15 are left blank in algorithm 2 to emphasize that additional work is performed by the ASM algorithm compared to the IQN algorithm. In addition to the work performed to evaluate the space-mapping functions on line 4 and 15, work need to be performed to obtain the coarse model solution \mathbf{z}^* . From the comparison of algorithm 1 and 2 it becomes clear that there is a strong connection between the ASM method and the IQN method. The ASM method reduces to the IQN method when the Taylor expansion of the fine model is taken as the coarse model $\tilde{\mathbf{R}}(\mathbf{z}) = \mathbf{R}^k + \nabla_{\mathbf{x}}^k \mathbf{R}(\mathbf{z} - \mathbf{x}^k)$ in the space mapping procedure. The IQN algorithm is therefore a special case of the ASM algorithm.

On the other hand, we can think of the ASM method as a quasi-Newton method applied to the (nonlinear) system of equations $p(\mathbf{x}) - \mathbf{z}^* = \mathbf{0}$ instead of $\mathbf{R}(\mathbf{x}) = \mathbf{0}$.

Algorithm 1 Aggressive Space-Mapping

```

1:  $k = 0$ 
2:  $\nabla_{\mathbf{x}}^0 p = \mathbf{I}$ 
3:  $\mathbf{R}^0 = S \circ F(\mathbf{x}^0) - \mathbf{x}^0$ 
4:  $\mathbf{z}^0 = \operatorname{argmin}_{\mathbf{z} \in \mathbf{Z}} \|\tilde{\mathbf{R}}(\mathbf{z}) - \mathbf{R}^0\|$ 
5: while  $\|\mathbf{R}^k\| > \epsilon$  do
6:   if  $k = 0$  then
7:      $\mathbf{x}^{k+1} = \mathbf{x}^k + (\nabla_{\mathbf{x}}^k p)^{-1}(\mathbf{z}^* - \mathbf{z}^k)$ 
8:   else
9:      $\Delta \mathbf{x} = \mathbf{x}^k - \mathbf{x}^{k-1}$ 
10:     $\nabla_{\mathbf{x}}^k p = \nabla_{\mathbf{x}}^{k-1} p + \frac{\mathbf{z}^k - \mathbf{z}^{k-1} - \nabla_{\mathbf{x}}^{k-1} p \Delta \mathbf{x}}{\Delta \mathbf{x}^T \Delta \mathbf{x}} \Delta \mathbf{x}^T$ 
11:     $\mathbf{x}^{k+1} = \mathbf{x}^k + (\nabla_{\mathbf{x}}^k p)^{-1}(\mathbf{z}^* - \mathbf{z}^k)$ 
12:   end if
13:    $k = k + 1$ 
14:    $\mathbf{R}^k = S \circ F(\mathbf{x}^k) - \mathbf{x}^k$ 
15:    $\mathbf{z}^k = \operatorname{argmin}_{\mathbf{z} \in \mathbf{Z}} \|\tilde{\mathbf{R}}(\mathbf{z}) - \mathbf{R}^k\|$ 
16: end while
    
```

Algorithm 2 Incremental Quasi-Newton

```

1:  $k = 0$ 
2:  $\nabla_{\mathbf{x}}^0 \mathbf{R} = -\mathbf{I}$ 
3:  $\mathbf{R}^0 = S \circ F(\mathbf{x}^0) - \mathbf{x}^0$ 
4:
5: while  $\|\mathbf{R}^k\| > \epsilon$  do
6:   if  $k = 0$  then
7:      $\mathbf{x}^{k+1} = \mathbf{x}^k - (\nabla_{\mathbf{x}}^k \mathbf{R})^{-1} \mathbf{R}^k$ 
8:   else
9:      $\Delta \mathbf{x} = \mathbf{x}^k - \mathbf{x}^{k-1}$ 
10:     $\nabla_{\mathbf{x}}^k \mathbf{R} = \nabla_{\mathbf{x}}^{k-1} \mathbf{R} + \frac{\mathbf{R}^k - \mathbf{R}^{k-1} - \nabla_{\mathbf{x}}^{k-1} \mathbf{R} \Delta \mathbf{x}}{\Delta \mathbf{x}^T \Delta \mathbf{x}} \Delta \mathbf{x}^T$ 
11:     $\mathbf{x}^{k+1} = \mathbf{x}^k - (\nabla_{\mathbf{x}}^k \mathbf{R})^{-1} \mathbf{R}^k$ 
12:   end if
13:    $k = k + 1$ 
14:    $\mathbf{R}^k = S \circ F(\mathbf{x}^k) - \mathbf{x}^k$ 
15:
16: end while
    
```

More computational work per fine model evaluation is performed in the ASM algorithm than in the IQN algorithm. The number of fine residual evaluations should therefore decrease significantly to gain efficiency with respect to the IQN method. The efficiency of the ASM method depends on the kind of coarse model used, the alignment of the coarse and fine model in time and model parameters. For a certain choice of coarse model, the efficiency should be determined by performing numerical experiments. To ease comparison of different techniques, an efficiency metric is defined in the following section.

2.4 Aggressive Space-Mapping versus Incremental Quasi-Newton method

The efficiency of the ASM method and the IQN method is determined by the cost per time step for a fixed tolerance ϵ in algorithm 1 and 2. Let $W_f^{k,n}$ and $W_c^{k,n}$ be the cost (flops or CPU time) of the k^{th} fine and coarse model iteration respectively at the n^{th} time step in the sequential time integration process.

$$W_f^{k,n} = \text{CPU time } S \circ F(\mathbf{x}^k) - \mathbf{x}^k, \quad (12)$$

$$W_c^{k,n} = \text{CPU time } S \circ \tilde{F}(\mathbf{z}^k) - \mathbf{z}^k. \quad (13)$$

The average cost per time step t^n of a fine and coarse model residual evaluation is then found from

$$\bar{W}_f(t^n) = \frac{1}{N_f} \sum_{k=1}^{k=N_f} W_f^{k,n} \quad \text{and} \quad \bar{W}_c(t^n) = \frac{1}{N_c} \sum_{k=1}^{k=N_c} W_c^{k,n} \quad (14)$$

The total cost per time step t^n of the ASM method and the IQN method is then estimated by

$$W^{\text{sm}}(t^n) = \bar{W}_f^{\text{sm}} N_f^{\text{sm}} + \bar{W}_c^{\text{sm}} N_c^{\text{sm}} \quad \text{and} \quad W^{\text{qn}}(t^n) = \bar{W}_f^{\text{qn}} N_f^{\text{qn}}, \quad (15)$$

where \bar{W}_f^{sm} and \bar{W}_f^{qn} are the average work of fine model residual evaluations per time step in the ASM and IQN method respectively. The time dependent efficiency of the ASM algorithm relative to the IQN algorithm is subsequently found from the ratio of work per time step

$$\eta(t^n) = 1 - \frac{W^{\text{sm}}(t^n)}{W^{\text{qn}}(t^n)} = 1 - \underbrace{\frac{N_f^{\text{sm}}}{N_f^{\text{qn}}}}_{\kappa_f} - \underbrace{\frac{\bar{W}_c^{\text{sm}}}{\bar{W}_f^{\text{qn}}}}_{\gamma_c} \underbrace{\frac{N_c^{\text{sm}}}{N_f^{\text{qn}}}}_{\kappa_c}. \quad (16)$$

Equation (16) is valid since $\bar{W}_f^{\text{sm}} \approx \bar{W}_f^{\text{qn}}$. The relative efficiency at t^n depends on three ratios: the ratio of fine model residual evaluations κ_f , the ratio of coarse to fine work γ_c and the ratio of the coarse to fine residual evaluations κ_c . The ASM method is more efficient than the IQN method if $\eta > 0$. From (16) it becomes clear that the success of space-mapping is not only determined by the reduction of fine model evaluations but also by the cost to achieve this reduction. For a constant κ_f , the product $\gamma_c \kappa_c$ determines the effect on the relative efficiency. It follows that a large number of coarse model evaluations with a very cheap coarse model can be as successful as a few iterations with a more expensive coarse model as long as the same reduction in the number of fine model evaluations is achieved.

3 1-D test case: Piston problem

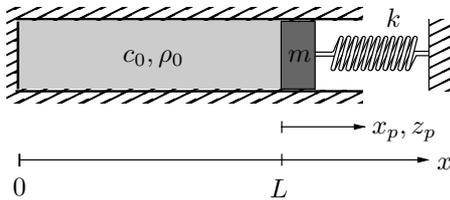


Figure 2: 1D testcase: Conceptual domain

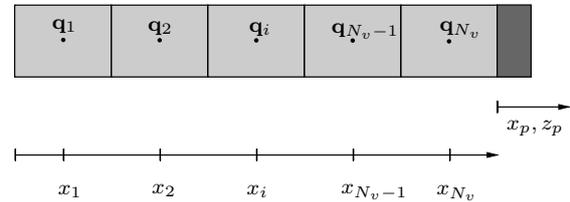


Figure 3: 1D testcase: Computational domain

3.1 Fine fluid model

The fluid in the piston is governed by the Euler equations of gas dynamics [3, 12]

$$\partial_t \mathbf{q} + \partial_x \mathbf{f}(\mathbf{q}) = \mathbf{0}. \quad (17)$$

Since a one-dimensional compressible inviscid fluid is assumed with $p = \frac{p_0}{\rho_0^\gamma} \rho^\gamma$, the energy equation in (17) becomes redundant and only the mass and momentum balance are necessary to describe the physics of the fluid. The state and flux vector are therefore given by

$$\mathbf{q} = (\rho \ \rho u)^T \quad \text{and} \quad \mathbf{f}(\mathbf{q}) = (\rho u \ \rho u^2 + p)^T, \quad (18)$$

with ρ the fluid density and u the horizontal fluid velocity. The Finite Volume technique is used to transform the integral form of equation (17) to a semi-discrete nonlinear system of equations

$$\partial_t \mathbf{w}_f + \mathbf{A}(\mathbf{w}_f) \mathbf{w}_f + \mathbf{A}_{fs} \mathbf{w}_s = \mathbf{0}. \quad (19)$$

Here, $\mathbf{w}_f = (\mathbf{q}_1^T \ \mathbf{q}_2^T \dots \mathbf{q}_{N_v}^T)^T$ is the discrete state vector of the fluid, \mathbf{A}_{fs} is the structure-to-fluid coupling matrix and $\mathbf{w}_s = [x_p \ \dot{x}_p]^T$ the state vector of the structure. The computational domain is shown in figure 3. To perform the coupling, a transpiration boundary condition is used on the fluid-structure interface, see [3]. From (19) it becomes clear that a fine model residual evaluation requires the solution of a nonlinear system of equations. The nonlinear system of equations is solved using simple Picard-iterations.

3.2 Coarse fluid model

If the flux vector \mathbf{f} in (18) is linearized around the equilibrium state of the fluid: $\mathbf{q} = (\rho_0 \ 0)^T$ we obtain

$$\partial_t \mathbf{q}' + \partial_{\mathbf{q}} \mathbf{f}|_{\mathbf{q}=\mathbf{q}_0} \partial_x \mathbf{q}' = \mathbf{0}, \quad (20)$$

with $\partial_{\mathbf{q}} \mathbf{f}$ the Jacobian of the nonlinear flux \mathbf{f} as found in [3, 12] and \mathbf{q}' a perturbation with respect to the equilibrium state vector \mathbf{q}_0 . If the Finite Volume method is applied to the integral form of (20), a discrete linear system of equations results

$$\partial_t \tilde{\mathbf{w}}'_f + \mathbf{A} \tilde{\mathbf{w}}'_f + \mathbf{A}_{fs} \tilde{\mathbf{w}}_s = \mathbf{0}. \quad (21)$$

Here, the state vector is denoted $\tilde{w}_s = [z_p \ \dot{z}_p]^T$. A coarse residual evaluation requires only the solution of a linear system of equations.

3.3 Testcases

The fluid-to-structure mass ratio ζ and ratio of characteristic time-scales λ are defined by

$$\zeta = \rho_0 L / m \quad \text{and} \quad \lambda = L \omega / c_0, \quad (22)$$

with $\omega = \sqrt{\frac{k}{m}}$ the natural frequency of the mass-spring system. It is well known that the convergence of fixed-point iterations depends on the ratio $\zeta/\lambda = \frac{\rho_0 c_0}{\sqrt{km}}$ and the time step used in the sequential integration process, see [12]. In order to study the performance of the ASM-algorithm for various levels of coupling strength, we fix the ratio of characteristic time scales and increase the fluid-to-structure mass ratio.

	Similarity parameters		Structural parameters		Fluid parameters		
	ζ [-]	λ [-]	m [kg]	k [$\frac{kg}{ms^2}$]	L [m]	ρ_0 [$\frac{kg}{m^3}$]	c_0 [$\frac{m}{s}$]
FSI-1	1/2	0.85	4	64300	2	1	300
FSI-2	2/3	0.85	3	48225	2	1	300
FSI-3	2	0.85	1	16075	2	1	300

Table 1: Physical parameters and similarity parameters of the 1-D FSI test cases

The testcases are collected in table 1. For each testcase we are interested in the relative time-dependent efficiency η as defined in section 2.4. The coarse fluid-structure model has a (non-dimensional) coupled period of $P_1 = 6.19$, $P_2 = 5.96$ and $P_3 = 5.04$ for the testcases FSI-1, FSI-2 and FSI-3 respectively. Time steps in the simulation are given with respect to the coarse coupled period as will become clear in section 3.4.

3.4 Numerical experiments

The initial conditions and numerical parameters are collected in table 2. The piston is released from an initial displacement. For the nondimensionalization the same convention is used as in [3]. The coarse and fine model use the same discretization.

Description	fine model		coarse model	
Initial piston displacement	\bar{x}_p^0	0.5	\bar{z}_p^0	0.5
Initial fluid density	$\bar{\rho}^0$	$\frac{2}{3}$	$\bar{\rho}^0$	$-\frac{1}{3}$
Number of finite volume cells	N_v	64	N_v	64
Number of time steps	N_t	500	N_t	500
Time step	Δt	$P/500$	Δt	$P/500$

Table 2: Nondimensional initial conditions and numerical parameters

The BDF2 time-integration scheme is used to integrate the semi-discrete coupled system in a partitioned way. The tolerance in algorithm 1 and 2 is set to a small value: $\epsilon = 1 \times 10^{-12}$, resulting in a strongly coupled solution. The fine and coarse model fluid density as a function of (nondimensional) space and time is given in figure 4 and 5 respectively for testcase FSI-2. A pressure wave is present in the fine model density response, hitting the piston at $\bar{t}/P_2 \approx 0.4$ and $\bar{t}/P_2 \approx 0.7$. The effect of the pressure wave is also present in

the solution of the piston in figure 6. There is no pressure wave present in the solution of the linear coarse model as can be seen in figure 5 and 7 respectively.

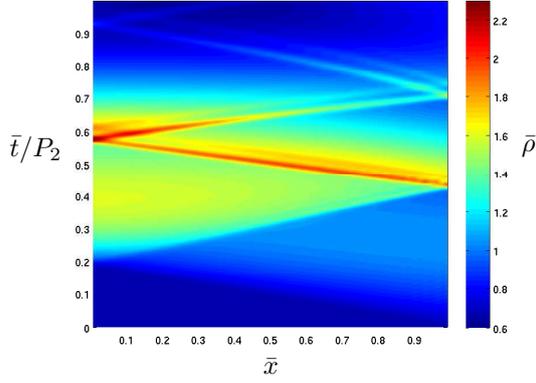


Figure 4: Fine density response of FSI-2

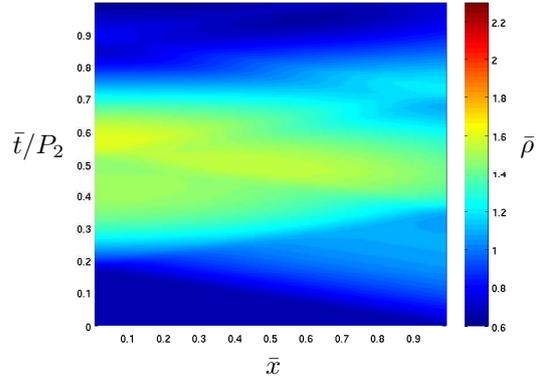


Figure 5: Coarse density response of FSI-2

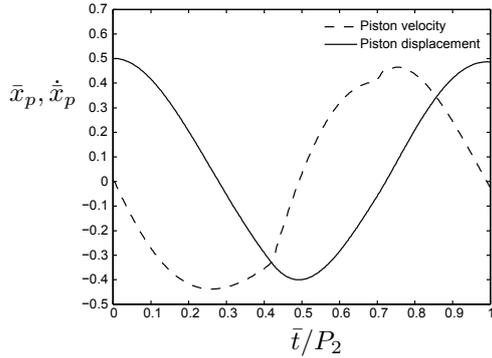


Figure 6: Fine piston response of FSI-2

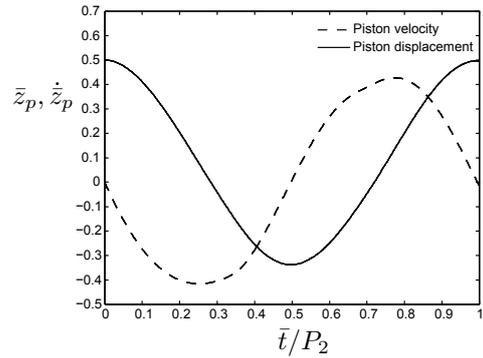


Figure 7: Coarse piston response of FSI-2

The number of fine model iterations used in the ASM method, the IQN method and the fixed-point iteration method as a function of time are plotted in figure 8. As expected, both ASM and IQN are more efficient than the fixed-point iteration method. The time dependent efficiency of the ASM algorithm relative to the IQN algorithm is shown in figure 9. It can be seen from figure 9 that the efficiency of ASM and IQN are comparable ($\eta \approx 0$), except at the moments when the pressure wave hits the piston. The ASM method is more efficient than the IQN method at these instances. This indicates that the efficiency of ASM increases when a strong interaction between the structure and the fluid is present. The efficiency η is equal to a small negative number $-\gamma_c \kappa_c$ ($O(10^{-3})$) when $N_f^{sm} = N_f^{qn}$. Hence, the coarse model work is negligible for this particular problem. The total relative efficiency of a single period is given by $\hat{\eta} = 1 - \frac{\sum W^{sm}(t^n)}{\sum W^{qn}(t^n)}$. The total efficiencies are

$\hat{\eta}_1 = 6.1\%$, $\hat{\eta}_2 = 6.2\%$ and $\hat{\eta}_3 = 8.3\%$ for FSI-1, FSI-2 and FSI-3 respectively. Hence, the relative efficiency increases when the algorithm is applied to strongly coupled fluid-structure interaction problems.

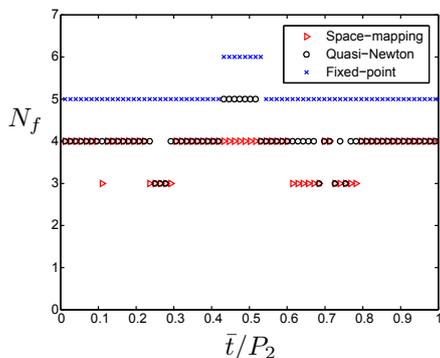


Figure 8: Fine model iterations N_f for FSI-2

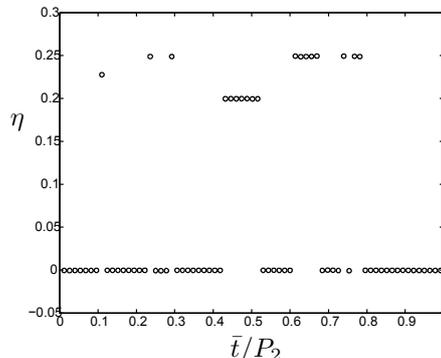


Figure 9: Efficiency η for FSI-2

4 Conclusions

- Aggressive Space-Mapping is successfully applied to obtain the transient solution of an academic fluid-structure problem in a partitioned way.
- Aggressive Space-Mapping is found to be more efficient ($\hat{\eta} = 6\%$ to 8%) than the corresponding Incremental Quasi-Newton method for the considered testcases and time intervals. The time-dependent efficiency is high at the moments of strong interaction ($\eta = 20\%$ to 25%), e.g. when a pressure wave hits the structure.
- The efficiency of Aggressive Space-Mapping increases with increasing fluid-to-structure mass ratio, keeping the ratio of characteristic time scales fixed. The efficiency is therefore higher for strongly coupled problems with large fluid densities and/or flexible and light structures.

5 Acknowledgement

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7 / 2007-2013) under a grant agreement number 233665. FFAST (Future Fast Aeroelastic Simulation Technologies) is a collaborative research project aimed at developing, implementing and assessing a range of numerical simulation technologies to accelerate future aircraft design. Advances in critical load identification and reduced order modelling methods will potentially provide a step change in the efficiency and accuracy of the dynamic aeroelastic loads process. The partners in FFAST are: University of Bristol, INRIA, CSIR, TU Delft, DLR, IRIAS, University of Liverpool, Politecnico di Milano, NUMECA, Optimad Engineering, Airbus-UK, EADS-MS and IITP.

REFERENCES

- [1] Echeverría, D. and Hemker, P.W. Space-mapping and defect correction, *CMAM J*(2005) **5**(2):107-136.
- [2] Robinson, T.D. and Eldred, M.S. and Willcox K.E. and Haines, R. Surrogate-based optimization using multifidelity models with variable parameterization and corrected space mapping *AIAA J* (2008) **46**(11):2814–2822
- [3] Zuijlen, A.H. and Bosscher, S. and Bijl, H. Two level algorithms for partitioned fluid-structure interaction computations *Comput. Methods Appl. Mech. Engrg.* (2007) **196**(8):1458–1470
- [4] Verhoosel, C.V. and Scholcz, T.P. Uncertainty and reliability analysis of fluid-structure stability boundaries, *AIAA J* (2009) **47**(1):91–104.
- [5] Degroote, J. and Bathe, K. and Vierendeels, J. Performance of a new partitioned procedure versus a monolithic procedure in fluid-structure interaction *Computers and Structures* (2009) **87**:793–801
- [6] Vierendeels, J and Lanoye, L. and Degroote, J. and Verdonk, P Implicit coupling of partitioned fluid-structure interaction problems with reduced order models *Computers and Structures* (2007) **85**:970–976
- [7] Ryckelynck, D. and Chinesta, F. On the *a priori* Model Reduction: Overview and Recent Developments *Arch. Comput. Meth. Engrg.* (2006) **13**(1):91–128
- [8] Ryckelynck, D. and Missoum Benziane, D. Multi-level a priori hyper reduction of mechanical models involving internal variables *Comput. Methods Appl. Mech. Engrg.* (2010) **199**:1134–1142
- [9] Markovinović, R. and Jansen, J.D. Accelerating iterative methods using reduced-order models as solution predictors *Int. J. Numer. Methods Engrg.* (2006) **68**:525–541
- [10] Matthies, H.G. and Niekamp, R. and Steindorf, J. Algorithms for strong coupling procedures *Comput. Methods Appl. Mech. Engrg.* (2006) **195**:2028–2049
- [11] Gallinger, T and Bletzinger, K Comparison of algorithms for strongly coupled partitioned fluid-structure interaction *ECCOMAS* Lisbon, June 2010
- [12] Brummelen, E.H. and R. de Borst On the nonnormality of subiteration for a fluid-structure-interaction problem *SIAM J. Sci. Comput.* (2005) **2**:599–621

A NEW NUMERICAL SCHEME FOR A LINEAR FLUID–STRUCTURE INTERACTION PROBLEM

MARÍA GONZÁLEZ* AND VIRGINIA SELGAS*

*Departamento de Matemáticas
Facultad de Informática, Universidad de A Coruña
Campus de Elviña, 15071 A Coruña, Spain
e-mails: mgtaboad@udc.es and vselgas@udc.es

Key words: Fluid-structure interaction, incompressible fluid, linear elasticity, finite element method

Abstract. We consider a linear fluid–structure interaction problem consisting of the time-dependent Stokes equations coupled with those of linear elastodynamics. We assume that the fluid and the solid interact through a fixed interface. Then, we reformulate the problem following the ideas of [6], and propose a new monolithic method in terms of the velocity (both in the fluid and the solid) and the fluid pressure. We discretize the problem using the implicit Euler method for the time variable, piecewise linear elements in the solid and the mini-element in the fluid domain. Displacements in the structure can be recovered by means of a quadrature formula. Our numerical results confirm the robustness and good convergence properties of the proposed scheme. Moreover, our approach is easy to implement as compared with other methods available in the literature.

1 INTRODUCTION

We consider a time-dependent system modeling the interaction between a Stokes fluid and an elastic structure in two or three dimensional bounded domains, and assume that the interface between the fluid and the solid is fixed. This model was studied in [2, 3] and can be used when the solid undergoes only infinitesimal elastic displacements but its velocity is large enough so that the fluid and the structure remain fully coupled; see [2] for more details.

In [2], a divergence-free weak formulation of this problem, that does not involve the fluid pressure field, was introduced and analyzed. The existence and uniqueness of a weak solution was proved and, under some additional assumptions on the data, strong energy estimates and the existence of a L^2 -integrable pressure field were derived. Semidiscrete finite element approximations were defined and studied in [3]. The existence of finite element solutions is proved there using an auxiliary discretely divergence-free formulation

and a discrete inf-sup condition, that allows to establish the existence of a finite element pressure. Strong a priori estimates for the finite element solutions and semidiscrete error estimates were also derived in [3]. Previous work concerning this model include, besides [2, 3], eigenmode analysis [7], homogenization [1], the one-dimensional case [4] and a numerical algorithm [5].

In this work, we reformulate the problem following the ideas of [6], and propose a new fully discrete scheme in terms of the velocity (both in the fluid and the solid) and the fluid pressure. We then discretize the problem using the implicit Euler method for the time variable, the mini-element for the Stokes problem and piecewise linear elements in the solid domain. As we will see, displacements in the structure can be recovered using a quadrature formula. We remark that this new approach is easy to implement and the numerical experiments carried out confirm its robustness and good convergence properties.

2 MODEL PROBLEM

We assume that the fluid and the solid occupy two adjacent Lipschitz domains, $\Omega_F \subset \mathbb{R}^d$ and $\Omega_S \subset \mathbb{R}^d$, respectively, where $d = 2$ or 3 is the space dimension. We let $\Sigma := \partial\Omega_F \cap \partial\Omega_S$ denote the interface between the fluid and the solid, and let $\Gamma_F := \partial\Omega_F \setminus \Sigma$ and $\Gamma_S := \partial\Omega_S \setminus \Sigma$ denote, respectively, the parts of the fluid and solid boundaries excluding the interface Σ ; we assume that $\text{meas}(\Gamma_F \cup \Gamma_S) \neq 0$. Finally, we denote by Ω the entire fluid–solid region, that is, $\Omega := \Omega_F \cup \Sigma \cup \Omega_S$. In the figure below, we represent from left to right the situations where $\Gamma_F = \emptyset, \Gamma_S \neq \emptyset$ or $\Gamma_F \neq \emptyset \neq \Gamma_S$.

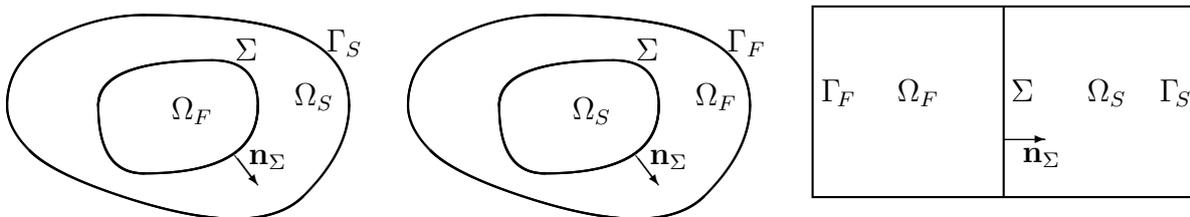


Figure 1: Domain of the problem. Cases $\Gamma_F = \emptyset, \Gamma_S \neq \emptyset$ and $\Gamma_F \neq \emptyset \neq \Gamma_S$.

Given $T > 0$, we consider a time-dependent Stokes fluid in Ω_F :

$$\left\{ \begin{array}{ll} \rho_F \partial_t \mathbf{v}_F - \mathbf{div}(\boldsymbol{\sigma}_F) = \mathbf{f}_F & \text{in } (0, T) \times \Omega_F, \\ \boldsymbol{\sigma}_F = -pI + 2\nu \boldsymbol{\varepsilon}(\mathbf{v}_F) & \text{in } (0, T) \times \Omega_F, \\ \mathbf{div}(\mathbf{v}_F) = 0 & \text{in } (0, T) \times \Omega_F, \\ \mathbf{v}_F = \mathbf{0} & \text{on } (0, T) \times \Gamma_F, \\ \mathbf{v}_F(0) = \mathbf{v}_F^0 & \text{in } \Omega_F, \end{array} \right. \quad (1)$$

where \mathbf{v}_F denotes the fluid velocity, $\boldsymbol{\sigma}_F$ the fluid stress tensor, \mathbf{f}_F the given fluid body force, p the fluid pressure, ρ_F the constant fluid density, ν the kinematic viscosity and \mathbf{v}_F^0 is the given initial fluid velocity. We recall that $\boldsymbol{\varepsilon}(\mathbf{v}) := \frac{1}{2}(\nabla\mathbf{v} + (\nabla\mathbf{v})^\dagger)$ denotes the strain tensor of small deformations.

In the solid region Ω_S , we consider the equations of linear elasticity:

$$\left\{ \begin{array}{ll} \rho_S \partial_{tt} \mathbf{u}_S - \operatorname{div}(\boldsymbol{\sigma}_S) = \mathbf{f}_S & \text{in } (0, T) \times \Omega_S, \\ \boldsymbol{\sigma}_S = \lambda \operatorname{div}(\mathbf{u}_S) I + 2\mu \boldsymbol{\varepsilon}(\mathbf{u}_S) & \text{in } (0, T) \times \Omega_S, \\ \mathbf{u}_S = \mathbf{0} & \text{on } (0, T) \times \Gamma_S, \\ \mathbf{u}_S(0) = \mathbf{u}_S^0 & \text{in } \Omega_S, \\ \partial_t \mathbf{u}_S(0) = \mathbf{v}_S^0 & \text{in } \Omega_S, \end{array} \right. \quad (2)$$

where \mathbf{u}_S denotes the displacement of the solid, $\boldsymbol{\sigma}_S$ the Cauchy stress tensor, \mathbf{f}_S the given loading force, ρ_S the constant solid density and λ and μ are the Lamé constants. The given initial data, \mathbf{u}_S^0 and \mathbf{v}_S^0 , represent, respectively, the initial displacement and the initial structural velocity.

We assume further that the velocity and the normal stresses are continuous across the interface Σ :

$$\mathbf{v}_F = \partial_t \mathbf{u}_S \quad \text{and} \quad \boldsymbol{\sigma}_F \mathbf{n}_\Sigma = \boldsymbol{\sigma}_S \mathbf{n}_\Sigma \quad \text{on } (0, T) \times \Sigma, \quad (3)$$

where \mathbf{n}_Σ is the unit normal vector along Σ pointing to Ω_S ; see Figure 1.

3 A WEAK FORMULATION

In order to propose a new numerical scheme to solve problem (1)–(3), we first derive a weak formulation for the problem in terms of the the fluid velocity \mathbf{v}_F , the structural velocity $\mathbf{v}_S := \partial_t \mathbf{u}_S$, and the pressure field p . In what follows, given a scalar or vector field $\xi \equiv \xi(t, \mathbf{x})$, we denote $\xi(t) := \xi(t, \cdot)$.

Multiplying the first equation of (1) by a test function $\mathbf{w}: \Omega_F \rightarrow \mathbb{R}^d$ such that $\mathbf{w} = \mathbf{0}$ on Γ_F , and integrating by parts, we obtain:

$$\rho_F \frac{d}{dt} \int_{\Omega_F} \mathbf{v}_F \cdot \mathbf{w} + \int_{\Omega_F} \boldsymbol{\sigma}_F : \boldsymbol{\varepsilon}(\mathbf{w}) - \int_{\Sigma} \boldsymbol{\sigma}_F \mathbf{n}_\Sigma \cdot \mathbf{w} = \int_{\Omega_F} \mathbf{f}_F \cdot \mathbf{w}. \quad (4)$$

Then, using the definition of $\boldsymbol{\sigma}_F$, we can write

$$\rho_F \frac{d}{dt} \int_{\Omega_F} \mathbf{v}_F \cdot \mathbf{w} - \int_{\Omega_F} p \operatorname{div}(\mathbf{w}) + 2\nu \int_{\Omega_F} \boldsymbol{\varepsilon}(\mathbf{v}_F) : \boldsymbol{\varepsilon}(\mathbf{w}) - \int_{\Sigma} \boldsymbol{\sigma}_F \mathbf{n}_\Sigma \cdot \mathbf{w} = \int_{\Omega_F} \mathbf{f}_F \cdot \mathbf{w}. \quad (5)$$

On the other hand, the weak formulation of the third equation in (1) is

$$\int_{\Omega_F} q \operatorname{div}(\mathbf{v}_F) = 0 \quad \forall q \in L^2(\Omega_F). \quad (6)$$

In the solid region Ω_S , we rewrite the equations of linear elasticity in terms of the structural velocity, $\mathbf{v}_S = \partial_t \mathbf{u}_S$, and the stress tensor, $\boldsymbol{\sigma}_S$:

$$\left\{ \begin{array}{ll} \rho_S \partial_t \mathbf{v}_S - \mathbf{div}(\boldsymbol{\sigma}_S) = \mathbf{f}_S & \text{in } (0, T) \times \Omega_S, \\ \partial_t \boldsymbol{\sigma}_S = \lambda \operatorname{div}(\mathbf{v}_S) I + 2\mu \boldsymbol{\varepsilon}(\mathbf{v}_S) & \text{in } (0, T) \times \Omega_S, \\ \mathbf{v}_S = \mathbf{0} & \text{on } (0, T) \times \Gamma_S, \\ \mathbf{v}_S(0) = \mathbf{v}_S^0 & \text{in } \Omega_S, \\ \boldsymbol{\sigma}_S(0) = \boldsymbol{\sigma}_S^0 & \text{in } \Omega_S, \end{array} \right. \quad (7)$$

where $\boldsymbol{\sigma}_S^0 := \lambda \operatorname{div}(\mathbf{u}_S^0) I + 2\mu \boldsymbol{\varepsilon}(\mathbf{u}_S^0)$. We remark that the displacement of the solid \mathbf{u}_S can then be recovered as

$$\mathbf{u}_S(t) = \mathbf{u}_S^0 + \int_0^t \mathbf{v}_S(s) ds, \quad (8)$$

and that the first coupling condition in (3) can be written as

$$\mathbf{v}_F = \mathbf{v}_S \quad \text{on } (0, T) \times \Sigma. \quad (9)$$

Multiplying the first equation of (7) by a test function $\mathbf{w}: \Omega_S \rightarrow \mathbb{R}^d$ such that $\mathbf{w} = \mathbf{0}$ on Γ_S and integrating by parts, we obtain:

$$\rho_S \frac{d}{dt} \int_{\Omega_S} \mathbf{v}_S \cdot \mathbf{w} + \int_{\Omega_S} \boldsymbol{\sigma}_S : \boldsymbol{\varepsilon}(\mathbf{w}) + \int_{\Sigma} \boldsymbol{\sigma}_S \mathbf{n}_{\Sigma} \cdot \mathbf{w} = \int_{\Omega_S} \mathbf{f}_S \cdot \mathbf{w}. \quad (10)$$

Then, multiplying the second equation of (7) by a test function $\boldsymbol{\tau}$ defined in Ω_S , and integrating in Ω_S , we get

$$\frac{d}{dt} \int_{\Omega_S} \boldsymbol{\sigma}_S : \boldsymbol{\tau} = \lambda \int_{\Omega_S} \operatorname{div}(\mathbf{v}_S) \operatorname{tr}(\boldsymbol{\tau}) + 2\mu \int_{\Omega_S} \boldsymbol{\varepsilon}(\mathbf{v}_S) : \boldsymbol{\tau}. \quad (11)$$

Taking into account the coupling condition (9), it is reasonable to look for a global unknown $\mathbf{v}(t) \in \mathbf{H}_0^1(\Omega)$ defined by

$$\mathbf{v} := \begin{cases} \mathbf{v}_F & \text{in } (0, T) \times \Omega_F, \\ \mathbf{v}_S & \text{in } (0, T) \times \Omega_S, \end{cases} \quad (12)$$

and consider global test functions $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$. We also define

$$\rho := \begin{cases} \rho_F & \text{in } \Omega_F, \\ \rho_S & \text{in } \Omega_S, \end{cases} \quad \mathbf{f} := \begin{cases} \mathbf{f}_F & \text{in } (0, T) \times \Omega_F, \\ \mathbf{f}_S & \text{in } (0, T) \times \Omega_S, \end{cases} \quad (13)$$

and introduce a global version of the initial data,

$$\mathbf{v}^0 := \begin{cases} \mathbf{v}_F^0 & \text{in } \Omega_F, \\ \mathbf{v}_S^0 & \text{in } \Omega_S. \end{cases} \quad (14)$$

Then, summing up equations (5) and (10), and using the second coupling condition in (3), we obtain

$$\frac{d}{dt} \int_{\Omega} \rho \mathbf{v} \cdot \mathbf{w} + 2\nu \int_{\Omega_F} \boldsymbol{\varepsilon}(\mathbf{v}_F) : \boldsymbol{\varepsilon}(\mathbf{w}) + \int_{\Omega_S} \boldsymbol{\sigma}_S : \boldsymbol{\varepsilon}(\mathbf{w}) - \int_{\Omega_F} p \operatorname{div}(\mathbf{w}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{w}, \quad (15)$$

for any $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$.

Therefore, given \mathbf{v}^0 and $\boldsymbol{\sigma}_S^0$, a weak formulation for the fluid–structure interaction problem (1)–(3) reads:

For each $t \in (0, T]$, find $\mathbf{v}(t) \in \mathbf{H}_0^1(\Omega)$, $p(t) \in L^2(\Omega_F)$ and $\boldsymbol{\sigma}_S(t) \in \mathbf{L}_{\text{sym}}^2(\Omega_S)$ such that

$$\begin{cases} \frac{d}{dt} \int_{\Omega} \rho \mathbf{v} \cdot \mathbf{w} + 2\nu \int_{\Omega_F} \boldsymbol{\varepsilon}(\mathbf{v}_F) : \boldsymbol{\varepsilon}(\mathbf{w}) + \int_{\Omega_S} \boldsymbol{\sigma}_S : \boldsymbol{\varepsilon}(\mathbf{w}) - \int_{\Omega_F} p \operatorname{div}(\mathbf{w}) = \int_{\Omega} \mathbf{f} \cdot \mathbf{w}, \\ \frac{d}{dt} \int_{\Omega_S} \boldsymbol{\sigma}_S : \boldsymbol{\tau} = \lambda \int_{\Omega_S} \operatorname{div}(\mathbf{v}_S) \operatorname{tr}(\boldsymbol{\tau}) + 2\mu \int_{\Omega_S} \boldsymbol{\varepsilon}(\mathbf{v}_S) : \boldsymbol{\tau}, \\ \int_{\Omega_F} q \operatorname{div}(\mathbf{v}_F) = 0, \\ \mathbf{v}(0) = \mathbf{v}^0 \quad \text{in } \Omega, \quad \boldsymbol{\sigma}_S(0) = \boldsymbol{\sigma}_S^0 \quad \text{in } \Omega_S, \end{cases} \quad (16)$$

for all $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$, $\boldsymbol{\tau} \in \mathbf{L}_{\text{sym}}^2(\Omega_S)$ and $q \in L^2(\Omega_F)$, where

$$\mathbf{L}_{\text{sym}}^2(\Omega_S) := \{ \boldsymbol{\sigma} \in [L^2(\Omega_S)]^{d \times d} : \boldsymbol{\sigma}^t = \boldsymbol{\sigma} \text{ in } \Omega_S \}. \quad (17)$$

4 FULLY DISCRETE SCHEME

4.1 Time discretization

Let N be a given nonnegative integer. We first consider a uniform partition of $[0, T]$, $\{t_n\}_{n=0}^N$, and denote $\Delta t := T/N$. For each time step t_n , $n = 1, 2, \dots$, we approximate

$$\frac{d}{dt} \mathbf{v}(t_n) \approx \frac{\mathbf{v}(t_n) - \mathbf{v}(t_{n-1})}{\Delta t} \quad \text{and} \quad \frac{d}{dt} \boldsymbol{\sigma}_S(t_n) \approx \frac{\boldsymbol{\sigma}_S(t_n) - \boldsymbol{\sigma}_S(t_{n-1})}{\Delta t}, \quad (18)$$

and, for any scalar or vector field $\xi = \xi(t, \mathbf{x})$, we denote $\xi^n \approx \xi(t_n)$.

Then, given \mathbf{v}^0 and $\boldsymbol{\sigma}_S^0$, for $n = 1, 2, \dots$, we solve for $\mathbf{v}^n \in \mathbf{H}_0^1(\Omega)$, $p^n \in L^2(\Omega_F)$ and $\boldsymbol{\sigma}_S^n \in \mathbf{L}_{\text{sym}}^2(\Omega_S)$ such that

$$\begin{cases} \int_{\Omega} \rho \mathbf{v}^n \cdot \mathbf{w} + 2\nu \Delta t \int_{\Omega_F} \boldsymbol{\varepsilon}(\mathbf{v}_F^n) : \boldsymbol{\varepsilon}(\mathbf{w}) + \Delta t \int_{\Omega_S} \boldsymbol{\sigma}_S^n : \boldsymbol{\varepsilon}(\mathbf{w}) - \Delta t \int_{\Omega_F} p^n \operatorname{div}(\mathbf{w}) = \\ \quad = \Delta t \int_{\Omega} \mathbf{f}(t_n) \cdot \mathbf{w} + \int_{\Omega} \rho \mathbf{v}^{n-1} \cdot \mathbf{w}, \\ \int_{\Omega_S} \boldsymbol{\sigma}_S^n : \boldsymbol{\tau} = \lambda \Delta t \int_{\Omega_S} \operatorname{div}(\mathbf{v}_S^n) \operatorname{tr}(\boldsymbol{\tau}) + 2\mu \Delta t \int_{\Omega_S} \boldsymbol{\varepsilon}(\mathbf{v}_S^n) : \boldsymbol{\tau} + \int_{\Omega_S} \boldsymbol{\sigma}_S^{n-1} : \boldsymbol{\tau}, \\ \int_{\Omega_F} q \operatorname{div}(\mathbf{v}_F^n) = 0, \end{cases} \quad (19)$$

for all $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$, $\boldsymbol{\tau} \in \mathbf{L}_{\text{sym}}^2(\Omega_S)$ and $q \in L^2(\Omega_F)$.

Now we remark that taking $\boldsymbol{\tau} = \boldsymbol{\varepsilon}(\mathbf{w})$ in the second equation of (19), we have that

$$\begin{aligned} \int_{\Omega_S} \boldsymbol{\sigma}_S^n : \boldsymbol{\varepsilon}(\mathbf{w}) &= \lambda \Delta t \int_{\Omega_S} \operatorname{div}(\mathbf{v}_S^n) \operatorname{div}(\mathbf{w}) + 2\mu \Delta t \int_{\Omega_S} \boldsymbol{\varepsilon}(\mathbf{v}_S^n) : \boldsymbol{\varepsilon}(\mathbf{w}) \\ &+ \int_{\Omega_S} \boldsymbol{\sigma}_S^{n-1} : \boldsymbol{\varepsilon}(\mathbf{w}) \end{aligned} \quad (20)$$

Then, substituting (20) in the first equation of (19), we derive the following semidiscrete in time scheme to solve problem (1)–(3):

Given \mathbf{v}^0 and $\boldsymbol{\sigma}_S^0$, for $n = 1, 2, \dots$, we look for $\mathbf{v}^n \in \mathbf{H}_0^1(\Omega)$, $p^n \in L^2(\Omega_F)$ such that

$$\left\{ \begin{aligned} \int_{\Omega} \rho \mathbf{v}^n \cdot \mathbf{w} + \int_{\Omega} \kappa \boldsymbol{\varepsilon}(\mathbf{v}^n) : \boldsymbol{\varepsilon}(\mathbf{w}) + \lambda (\Delta t)^2 \int_{\Omega_S} \operatorname{div}(\mathbf{v}_S^n) \operatorname{div}(\mathbf{w}) - \Delta t \int_{\Omega_F} p^n \operatorname{div}(\mathbf{w}) &= \\ = \Delta t \int_{\Omega} \mathbf{f}(t_n) \cdot \mathbf{w} + \int_{\Omega} \rho \mathbf{v}^{n-1} \cdot \mathbf{w} - \Delta t \int_{\Omega_S} \boldsymbol{\sigma}_S^{n-1} : \boldsymbol{\varepsilon}(\mathbf{w}), & \\ \int_{\Omega_F} q \operatorname{div}(\mathbf{v}_F^n) = 0, & \end{aligned} \right. \quad (21)$$

for all $\mathbf{w} \in \mathbf{H}_0^1(\Omega)$ and $q \in L^2(\Omega_F)$, where $\kappa := 2\nu \Delta t$ in Ω_F and $\kappa := 2\mu (\Delta t)^2$ in Ω_S . We then approximate

$$\boldsymbol{\sigma}_S(t_n) \approx \boldsymbol{\sigma}_S^n := \boldsymbol{\sigma}_S^{n-1} + \Delta t (\boldsymbol{\sigma}_S^n)', \quad (22)$$

where $(\boldsymbol{\sigma}_S^n)' := \lambda \operatorname{div}(\mathbf{v}_S^n) I + 2\mu \boldsymbol{\varepsilon}(\mathbf{v}_S^n)$ in Ω_S .

Let us define the bilinear forms $a : \mathbf{H}_0^1(\Omega) \times \mathbf{H}_0^1(\Omega) \rightarrow \mathbb{R}$ and $b : L^2(\Omega_F) \times \mathbf{H}_0^1(\Omega) \rightarrow \mathbb{R}$, and the linear functional $l^n : \mathbf{H}_0^1(\Omega) \rightarrow \mathbb{R}$:

$$\begin{aligned} a(\mathbf{v}, \mathbf{w}) &:= \int_{\Omega} \rho \mathbf{v} \cdot \mathbf{w} + \int_{\Omega} \kappa \boldsymbol{\varepsilon}(\mathbf{v}) : \boldsymbol{\varepsilon}(\mathbf{w}) + \lambda (\Delta t)^2 \int_{\Omega_S} \operatorname{div}(\mathbf{v}) \operatorname{div}(\mathbf{w}), \\ b(q, \mathbf{w}) &:= -\Delta t \int_{\Omega_F} q \operatorname{div}(\mathbf{w}), \quad \forall q \in L^2(\Omega_F), \\ l^n(\mathbf{w}) &:= \Delta t \int_{\Omega} \mathbf{f}(t_n) \cdot \mathbf{w} + \int_{\Omega} \rho \mathbf{v}^{n-1} \cdot \mathbf{w} - \Delta t \int_{\Omega_S} \boldsymbol{\sigma}_S^{n-1} : \boldsymbol{\varepsilon}(\mathbf{w}), \end{aligned} \quad (23)$$

for any $\mathbf{v}, \mathbf{w} \in \mathbf{H}_0^1(\Omega)$ and any $q \in L^2(\Omega_F)$. With these notations, problem (21) can be written as follows:

Given \mathbf{v}^0 and $\boldsymbol{\sigma}_S^0$, for $n = 1, 2, \dots$, we look for $\mathbf{v}^n \in \mathbf{H}_0^1(\Omega)$, $p^n \in L^2(\Omega_F)$ such that

$$\left\{ \begin{aligned} a(\mathbf{v}^n, \mathbf{w}) + b(p^n, \mathbf{w}) &= l^n(\mathbf{w}), \quad \forall \mathbf{w} \in \mathbf{H}_0^1(\Omega), \\ b(q, \mathbf{v}^n) &= 0, \quad \forall q \in L^2(\Omega_F). \end{aligned} \right. \quad (24)$$

4.2 Finite element discretization

We consider now finite element subspaces $\mathbf{V}_h \subset \mathbf{H}_0^1(\Omega)$ and $Q_h \subset L^2(\Omega_F)$, such that $(\mathbf{V}_h|_{\Omega_F}, Q_h)$ is a stable pair for the Stokes problem. The corresponding fully discrete scheme to solve problem (1)–(3) reads as follows:

Given \mathbf{v}^0 and $\boldsymbol{\sigma}_S^0$, for $n = 1, 2, \dots$, we look for $\mathbf{v}_h^n \in \mathbf{V}_h$, $p_h^n \in Q_h$ such that

$$\begin{cases} a(\mathbf{v}_h^n, \mathbf{w}_h) + b(p_h^n, \mathbf{w}_h) = l^n(\mathbf{w}_h), & \forall \mathbf{w}_h \in \mathbf{V}_h, \\ b(q_h, \mathbf{v}_h^n) = 0, & \forall q_h \in Q_h, \end{cases} \quad (25)$$

and then compute

$$\boldsymbol{\sigma}_S^n = \boldsymbol{\sigma}_S^{n-1} + \Delta t (\lambda \operatorname{div}(\mathbf{v}_h^n) I + 2\mu \boldsymbol{\varepsilon}(\mathbf{v}_h^n)).$$

5 NUMERICAL RESULTS

We implemented the fully discrete scheme (25) in a MATLAB code in the case $d = 2$, choosing the mini-element to approximate the solution in the fluid domain and piecewise linear elements in the solid domain.

In this section, we show some results for a particular test problem with the known solution

$$\begin{aligned} v_1(t, \mathbf{x}) &= \begin{cases} \cos x_2 e^t & \text{in } (0, 1) \times \Omega_F, \\ (\cos x_2 + \sin x_1) e^t & \text{in } (0, 1) \times \Omega_S, \end{cases} \\ v_2(t, \mathbf{x}) &= \sin x_1 e^t \quad \text{in } (0, 1) \times \Omega, \\ p(t, \mathbf{x}) &= -2 \cos x_1 e^t \quad \text{in } (0, 1) \times \Omega_F, \end{aligned} \quad (26)$$

where $\Omega_F = (-1, 0) \times (-1, 1)$ and $\Omega_S = (0, 1) \times (-1, 1)$. We consider the case in which the fluid and the solid have similar mass densities; more precisely, we take $\rho_F = \rho_S = 1$. The kinematic viscosity is $\nu = \frac{1}{2}$ and the Lamé parameters are $\mu = \frac{1}{2}$ and $\lambda = 1$.

5.1 Validation of the spatial discretization

To start with, we test the choice of finite element spaces, that is, the mini-element in the fluid domain and piecewise linear elements in the solid domain. To this end, we solved the test problem for a fixed time. In particular, we took the time step $\Delta t = 1$ and used exact Dirichlet boundary conditions for the fixed time. We considered a sequence of meshes obtained by uniformly refining the initial mesh; we denote by $h = l, l/2, \dots$ the corresponding mesh-sizes, N_h is the associated number of vertices and *dof* denotes the corresponding degrees of freedom (d.o.f.). The computed solutions, $\mathbf{v}_h = (v_{1,h}, v_{2,h})$, p_h , were compared with the exact one, $\mathbf{v} = (v_1, v_2)$, p , to determine the order of convergence.

In Table 1, we show the total number of vertices, the corresponding d.o.f., and the absolute errors for each unknown, that we denote $e_h(v_j) := \|v_j - v_{j,h}\|_{H^1(\Omega)}$ ($j = 1, 2$) and $e_h(p) := \|p - p_h\|_{L^2(\Omega_F)}$. The corresponding experimental convergence rates for each unknown are shown in Table 2. The experimental convergence rate is computed for each pair of consecutive meshes as the slope in log-log scale of the error versus the number of vertices of the associated mesh:

$$r_h := -2 \frac{\log e_{h/2} - \log e_h}{\log N_{h/2} - \log N_h}. \quad (27)$$

Table 1: Absolute errors for different meshes at a fixed time.

h	N_h	dof	$e_h(v_1)$	$e_h(v_2)$	$e_h(p)$
l	195	772	0.075514	0.043627	0.022462
$l/2$	737	3127	0.037705	0.021663	0.007871
$l/4$	2865	12595	0.018833	0.010744	0.002907
$l/8$	11297	50563	0.009400	0.005344	0.001071
$l/16$	44865	202627	0.004695	0.002666	0.000385

Table 2: Experimental convergence rates for successive meshes at a fixed time.

$r_h(v_1)$	1.044696	1.022568	1.012914	1.006846
$r_h(v_2)$	1.053087	1.032904	1.017875	1.008713
$r_h(p)$	1.577438	1.466830	1.455395	1.482009
$r_h(\mathbf{u}, p)$	1.070852	1.035734	1.020001	1.010696

We can observe that, as the mesh becomes finer, the experimental rate of convergence approaches 1 for both velocity components, whereas the method is superconvergent in the pressure variable, with rates of convergence that approach 1.5. It is important to remark that we have observed a similar behavior in other tests.

5.2 Validation of the time discretization

The main aim of this example is to study the stability of the numerical scheme (25) as the time step Δt becomes smaller. To this end, we chose the mesh of size $h = l/8$, and then decreased Δt as shown in Table 3. We remark that the associated errors, $e_h^{\Delta t}(v_j) := \|v_j - v_{j,h}^{\Delta t}\|_{L^2((0,1), H^1(\Omega))}$ ($j = 1, 2$) and $e_h^{\Delta t}(p) := \|p - p_h^{\Delta t}\|_{L^2((0,1), L^2(\Omega_F))}$, stagnate faster for the velocity than for the pressure.

Table 3: Absolute errors for the mesh-size $l/8$ and different time steps.

Δt	$e_h^{\Delta t}(v_1)$	$e_h^{\Delta t}(v_2)$	$e_h^{\Delta t}(p)$
10^{-1}	0.039295	0.041494	0.238824
10^{-2}	0.017876	0.010804	0.023515
10^{-3}	0.018346	0.011303	0.003186
10^{-4}	0.022725	0.016331	0.002571

5.3 Numerical results for the fully discrete scheme

Finally, we present the numerical results obtained taking a fixed time step and successively refined meshes. In Tables 4 and 6 we show the results obtained for $\Delta t = 10^{-2}$ and $\Delta t = 10^{-3}$, respectively. The corresponding experimental convergence rates are shown in Tables 5 and 7.

Table 4: Absolute errors for successive meshes and the fixed time step $\Delta t = 10^{-2}$.

h	$e_h^{\Delta t}(v_1)$	$e_h^{\Delta t}(v_2)$	$e_h^{\Delta t}(p)$
l	0.152335	0.098009	0.053320
$l/2$	0.073220	0.045141	0.029586
$l/4$	0.035434	0.021105	0.024505
$l/8$	0.017876	0.010804	0.023515

Table 5: Experimental convergence rates for successive meshes and the fixed time step $\Delta t = 10^{-2}$.

$r_h(v_1)$	1.102011	1.069099	0.997426
$r_h(v_2)$	1.166158	1.119933	0.976049
$r_h(p)$	0.885995	0.277532	0.060108
$r_h(\mathbf{v}, p)$	1.098631	0.942423	0.615441

For $\Delta t = 10^{-2}$, the experimental convergence rates are around 1 for the two components of the velocity, but this time step appears to be too rough to obtain good convergence behavior in the pressure variable. This problem can be solved using a smaller Δt . Indeed, we can observe in Table 7 that for $\Delta t = 10^{-3}$ a good convergence behavior is obtained for the pressure too.

Acknowledgements. The research of the authors was partially supported by the MEC research project MTM2010-21135-C02-01.

Table 6: Absolute errors for successive meshes and the fixed time step $\Delta t = 10^{-3}$.

h	$e_h^{\Delta t}(v_1)$	$e_h^{\Delta t}(v_2)$	$e_h^{\Delta t}(p)$
l	0.194069	0.147480	0.054495
$l/2$	0.092336	0.067421	0.017108
$l/4$	0.040903	0.027274	0.006349
$l/8$	0.018346	0.011303	0.003186

Table 7: Experimental convergence rates for successive meshes and the fixed time step $\Delta t = 10^{-3}$.

$r_h(v_1)$	1.117301	1.199388	1.1687576
$r_h(v_2)$	1.177398	1.333136	1.284060
$r_h(p)$	1.742736	1.460054	1.005267
$r_h(\mathbf{v}, p)$	1.158779	1.247323	1.198638

REFERENCES

- [1] S. Dasser, A penalization method for the homogenization of a mixed fluid–structure problem, *C.R. Acad. Sci. Paris Sér. I Math.* (1995) **320**:759–764.
- [2] Q. Du, M.D. Gunzburger, L.S. Hou and J. Lee, Analysis of a linear fluid–structure interaction problem, *Discrete and Continuous Dynamical Systems* (2003) **9(3)**:633–650.
- [3] Q. Du, M.D. Gunzburger, L.S. Hou and J. Lee, Semidiscrete finite element approximations of a linear fluid–structure interaction problem, *SIAM J. Numer. Anal.* (2004) **42**:1–29.
- [4] D. Errate, M.J. Esteban and Y. Maday, Couplage fluide–structure: Un modèle simplifié en dimension 1, *C.R. Acad. Sci. Paris Sér. I Math.* (1994) **318**:275–281.
- [5] C. Farhat, M. Lesoinne and P. LeTallec, Load and motion transfer algorithms for fluid/structure interaction problems with non–matching discrete interfaces: Momentum and energy conservation, optimal discretization and application to aeroelasticity, *Comput. Meth. Appl. Mech. Engrg.* (1998) **157**:95–114.
- [6] O. Kayser–Herold and H.G. Matthies, A unified least–squares formulation for fluid–structure interaction problems, *Computers and Structures* (2007) **85**:998–1011.
- [7] R. Schulkes, Interactions of an elastic solid with a viscous fluid: Eigenmode analysis, *J. Comput. Phys.* (1992) **100**:270–283.

A SUBSTRUCTURING FE MODEL FOR STRUCTURAL-ACOUSTIC PROBLEMS WITH MODAL-BASED REDUCTION OF POROELASTIC INTERFACE

Romain Rumpler^{*†}, Jean-François Deü^{*} and Peter Göransson[†]

^{*}Structural Mechanics and Coupled Systems Laboratory (LMSSC)
Conservatoire National des Arts et Métiers
292 rue Saint-Martin, case 353, 75141 Paris Cedex 03, France
e-mail: deu@cnam.fr - www.cnam.fr/lmssc/

[†]The Marcus Wallenberg Laboratory for Sound and Vibration Research (MWL)
Department of Aeronautical and Vehicle Engineering, KTH
KTH School of Engineering Sciences, SE-100 44 Stockholm, Sweden
e-mail: pege@kth.se, rumpler@kth.se - www.kth.se/en/sci/institutioner/ave/avd/mwl

Key words: Poroelastic materials; Noise reduction; Reduced model; Structural-acoustics.

Abstract. In this work, a component mode synthesis technique is proposed to improve the computational efficiency of Finite Element problems including 3D modelling of poroelastic materials. The modal reduction relies on real-valued modes, solution of a standard eigenvalue problem, based on a classical solid and fluid displacements formulation of the porous media. Efficiency in terms of degrees of freedom, convergence, sparsity and computation time is presented.

1 INTRODUCTION

Modelling poroelastic materials for interior noise reduction, extensively used in the transport industry, can lead to rather expensive Finite Element (FE) models. Therefore, efforts have been made in the last decade to propose efficient solution strategies for the Biot-Allard theory [1]. Use of a mixed displacement-pressure formulation for the solid and fluid phases respectively [2] downsized the number of degrees of freedom (dofs) per node from 6, when using a standard solid and fluid phases displacement formulation, to 4 dofs. Hierarchical elements also proved to reduced the number of dofs needed to model the porous media [3]. The use of equivalent acoustic impedances [4, 5], while implying low computational cost, is limited by strong assumptions, or subject to a preliminary identification step thus hampering computational efficiency. Alternatively, modal reduction techniques have been proposed and applied to poroelastic FE formulations, in an attempt

to keep a fine and complex 3D modelling of the problem in the scope of low frequency applications [6–8].

In this work, a component mode synthesis is applied to the dissipative part of a 3D poro-acoustic FE problem. The standard solid and fluid displacements formulation is used to model poroelastic media. A direct computation scheme at each frequency increment is used to solve the frequency-dependent problem, and frequency response of the mean quadratic pressure in the acoustic domain is computed as an indicator of the sound level. Real-valued modes based on the bi-phase poroelastic media are used to define a transformation applied once at the initial increment, and suitable for the frequency range of interest. After a presentation of the formulation as well as the modal method used, the proposed reduction is tested on a rigid cavity treated with a porous layer on one wall. While showing substantial computation time speed-up, due to both the reduced number of dofs and the good sparsity of the reduced system, the relatively low convergence rate suggests further possible improvements.

2 FE FORMULATION FOR THE PORO-ACOUSTIC PROBLEM

A poro-acoustic problem is considered, which description and notations are presented on Fig. 1. The acoustic fluid and the porous media occupy the domains Ω_F and Ω_P respectively. The compressible fluid is described using pressure fluctuation (p) as primary variable (Subsection 2.1.1), while fluid and solid phases homogenized displacements ($\mathbf{u}_s, \mathbf{u}_f$) are retained as primary variables for the porous media (Subsection 2.1.2). The domains boundaries are separated into contours of:

- imposed Dirichlet boundary conditions denoted $\partial_1\Omega_F$ and $\partial_1\Omega_P$,
- prescribed Neumann boundary conditions denoted $\partial_2\Omega_F$ and $\partial_2\Omega_P$,
- coupling interface between acoustic fluid and porous media (Γ_{FP}).

The FE formulation is presented for a permanent harmonic response at angular frequency ω .

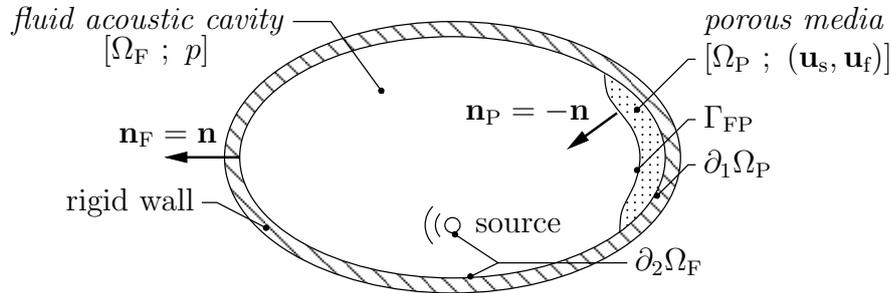


Figure 1: Description and notations of the poro-acoustic interaction problem

2.1 Dynamic equations and constitutive laws

2.1.1 Compressible fluid (p)

The internal fluid within cavities is assumed compressible and inviscid, satisfying the Helmholtz equation derived from the motion, continuity, and constitutive equations:

$$\Delta p + \frac{\omega^2}{c_0^2} p = 0 \quad \text{in } \Omega_F \quad (1)$$

where c_0 is the constant speed of sound in the fluid, and p the pressure fluctuation scalar field. The limit case $\omega = 0$ is not given by Equation (1). Though not considered in this work, the solution is given by the static solution of the coupled fluid-structure problem [9].

2.1.2 Porous media Biot theory ($\mathbf{u}_s, \mathbf{u}_f$)

Notation	Description
ρ_s	Solid frame density
$(\lambda; \mu)$	Lamé parameters for the solid frame
ρ_f	Ambient fluid density
η	Ambient fluid viscosity
P_0	Ambient fluid standard pressure
γ	Heat capacity ratio for the ambient fluid
Pr	Prandtl number for the ambient fluid
ϕ	Porosity
α_∞	Tortuosity
σ	Static flow resistivity
Λ	Viscous characteristic length
Λ'	Thermal characteristic length

Table 1: List of material parameters

At angular frequency ω , the proelastic media satisfies the following elastodynamic linearized equations, derived in the Biot-Allard theory [1], taking into account inertia and viscous coupling effects between solid and fluid phases:

$$\operatorname{div} \boldsymbol{\sigma}_s - i\omega \tilde{b}(\omega)(\mathbf{u}_s - \mathbf{u}_f) + \omega^2 [(\rho_s + \rho_a)\mathbf{u}_s - \rho_a \mathbf{u}_f] = \mathbf{0} \quad \text{in } \Omega_P \quad (2a)$$

$$\operatorname{div} \boldsymbol{\sigma}_f - i\omega \tilde{b}(\omega)(\mathbf{u}_f - \mathbf{u}_s) + \omega^2 [-\rho_a \mathbf{u}_s + (\phi \rho_f + \rho_a)\mathbf{u}_f] = \mathbf{0} \quad \text{in } \Omega_P \quad (2b)$$

where \mathbf{u}_s and \mathbf{u}_f are respectively the solid phase and fluid phase averaged displacements in the sense of Biot theory. $\tilde{b}(\omega)$ (henceforth denoted \tilde{b} , where $\tilde{\cdot}$ refers to a complex-valued quantity) and ρ_a are respectively the complex frequency-dependent viscous drag and the inertia coupling parameter, based on the standard notations of material parameters introduced in Table 1 [1], and given by:

$$\tilde{b} = \sigma \phi^2 \left[1 + \frac{4i\omega \alpha_\infty^2 \eta \rho_f}{\sigma^2 \Lambda^2 \phi^2} \right]^{\frac{1}{2}} \quad (3)$$

$$\rho_a = \phi \rho_f (\alpha_\infty - 1) \quad (4)$$

$\boldsymbol{\sigma}_s$ and $\boldsymbol{\sigma}_f$ are the averaged stress tensors for the solid and fluid phases respectively. They satisfy the Lagrangian stress-strain relations developed by Biot:

$$\boldsymbol{\sigma}_s = \left(\lambda + \frac{(1-\phi)^2}{\phi} \tilde{K}_f(\omega) \right) \text{tr}[\boldsymbol{\varepsilon}(\mathbf{u}_s)] \mathbf{I} + 2\mu \boldsymbol{\varepsilon}(\mathbf{u}_s) + (1-\phi) \tilde{K}_f(\omega) \text{tr}[\boldsymbol{\varepsilon}(\mathbf{u}_f)] \mathbf{I} \quad (5a)$$

$$\boldsymbol{\sigma}_f = (1-\phi) \tilde{K}_f(\omega) \text{tr}[\boldsymbol{\varepsilon}(\mathbf{u}_s)] \mathbf{I} + \phi \tilde{K}_f(\omega) \text{tr}[\boldsymbol{\varepsilon}(\mathbf{u}_f)] \mathbf{I} \quad (5b)$$

where $\boldsymbol{\varepsilon}(\mathbf{u}_s)$ and $\boldsymbol{\varepsilon}(\mathbf{u}_f)$ are the strain tensors associated to the averaged displacements vector fields \mathbf{u}_s and \mathbf{u}_f , defined by:

$$\boldsymbol{\varepsilon}(\mathbf{v}) = \frac{1}{2} (\mathbf{grad} \mathbf{v} + \mathbf{grad}^T \mathbf{v}) \quad (6)$$

Beside the standard material parameters presented in Table 1, the effective bulk modulus of the fluid phase $\tilde{K}_f(\omega)$ (henceforth denoted \tilde{K}_f) is also introduced. For reasons to be presented in Subsection 3.2, its expression [1] is separated into its zero-frequency limit and complex frequency-dependent behaviour:

$$\tilde{K}_f = \frac{\gamma P_0}{\gamma - (\gamma - 1) \left[1 + \frac{8\eta}{i\omega Pr \Lambda'^2 \rho_f} \left(1 + \frac{i\omega Pr \Lambda'^2 \rho_f}{16\eta} \right)^{\frac{1}{2}} \right]^{-1}} = P_0 + \left(\tilde{K}_f - P_0 \right) \quad (7)$$

which, when introduced in Equations (5), leads to the following expressions of the stress-strain relations using Voigt notation:

$$\boldsymbol{\sigma}_s = \mathbf{D}_s^{(1)} \boldsymbol{\varepsilon}(\mathbf{u}_s) + \tilde{\mathbf{D}}_s^{(2)}(\omega) \boldsymbol{\varepsilon}(\mathbf{u}_s) + \mathbf{D}_{sf}^{(1)} \boldsymbol{\varepsilon}(\mathbf{u}_f) + \tilde{\mathbf{D}}_{sf}^{(2)}(\omega) \boldsymbol{\varepsilon}(\mathbf{u}_f) \quad (8a)$$

$$\boldsymbol{\sigma}_f = \mathbf{D}_{sf}^{(1)} \boldsymbol{\varepsilon}(\mathbf{u}_s) + \tilde{\mathbf{D}}_{sf}^{(2)}(\omega) \boldsymbol{\varepsilon}(\mathbf{u}_s) + \mathbf{D}_f^{(1)} \boldsymbol{\varepsilon}(\mathbf{u}_f) + \tilde{\mathbf{D}}_f^{(2)}(\omega) \boldsymbol{\varepsilon}(\mathbf{u}_f) \quad (8b)$$

with:

$$\begin{aligned} \mathbf{D}_s^{(1)} &= \left(\lambda + \frac{(1-\phi)^2}{\phi} P_0 \right) \mathbf{D} + \mu \mathbf{diag} \begin{pmatrix} 2 & 2 & 2 & 1 & 1 & 1 \end{pmatrix} \\ \mathbf{D}_{sf}^{(1)} &= (1-\phi) P_0 \mathbf{D} \\ \mathbf{D}_f^{(1)} &= \phi P_0 \mathbf{D} \\ \tilde{\mathbf{D}}_s^{(2)}(\omega) &= \left(\tilde{K}_f - P_0 \right) \mathbf{D}_s^{(2)} = \frac{(1-\phi)^2}{\phi} \left(\tilde{K}_f - P_0 \right) \mathbf{D} \\ \tilde{\mathbf{D}}_{sf}^{(2)}(\omega) &= \left(\tilde{K}_f - P_0 \right) \mathbf{D}_{sf}^{(2)} = (1-\phi) \left(\tilde{K}_f - P_0 \right) \mathbf{D} \\ \tilde{\mathbf{D}}_f^{(2)}(\omega) &= \left(\tilde{K}_f - P_0 \right) \mathbf{D}_f^{(2)} = \phi \left(\tilde{K}_f - P_0 \right) \mathbf{D} \end{aligned} \quad \text{where } \mathbf{D} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

In this contribution, the Lamé parameters for the solid frame are considered real and frequency-independent, so that no structural damping is taken into account in the porous media behaviour. However, the method presented is also valid and straightforward to establish when structural damping is taken into account.

2.2 Fluid-structure interaction problem

In this subsection, boundary and coupling conditions are recalled for the poro-acoustic coupled problem presented in Fig. 1, in order to establish the discretized FE problem.

2.2.1 Poro-acoustic coupling and boundary conditions

At external boundary of the acoustic fluid domain, rigid cavity conditions are classically imposed by setting a free pressure field ($\partial_1\Omega_F = \{\emptyset\}$). A harmonic excitation is prescribed via an acoustic source:

$$\mathbf{grad} p \cdot \mathbf{n} = \omega^2 \rho_F u_{Fb} \quad \text{on } \partial_2\Omega_F \quad (9)$$

where u_{Fb} is set to zero out of the acoustic source included in $\partial_2\Omega_F$.

Coupling at interface Γ_{FP} is given by normal stress and normal displacement continuity conditions between acoustic fluid and both fluid and solid phases of porous media:

$$\boldsymbol{\sigma}_s \mathbf{n} + (1 - \phi) p \mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_{FP} \quad (10a)$$

$$\boldsymbol{\sigma}_f \mathbf{n} + \phi p \mathbf{n} = \mathbf{0} \quad \text{on } \Gamma_{FP} \quad (10b)$$

$$\mathbf{u}_F \cdot \mathbf{n} - (1 - \phi) \mathbf{u}_s \cdot \mathbf{n} - \phi \mathbf{u}_f \cdot \mathbf{n} = 0 \quad \text{on } \Gamma_{FP} \quad (11)$$

where ϕ is the porosity of the porous material, i.e. the volume fraction of fluid.

No external force is applied to the outer boundary of the porous media beside at interface Γ_{FP} . Therefore, $\partial_2\Omega_P = \{\emptyset\}$ in the considered problem. Finally, at external boundary $\partial_1\Omega_P$, two types of boundary conditions can be prescribed, the porous material being considered either as sliding or bounded to a rigid wall (Table 2).

<i>Bounded layer</i>	<i>Sliding layer</i>
$\mathbf{u}_s = \mathbf{0}$	$\mathbf{u}_s \cdot \mathbf{n}_P = 0$
$\mathbf{u}_f \cdot \mathbf{n}_P = 0$	$\mathbf{u}_f \cdot \mathbf{n}_P = 0$

Table 2: Boundary conditions for porous layer on $\partial_1\Omega_P$

2.2.2 Finite element discretized problem

The test-function method is used to derive the variational formulation of the coupled problem. For this purpose, the spaces of sufficiently smooth functions C_p , $C_{\mathbf{u}_s}$ and $C_{\mathbf{u}_f}$ are introduced, associated to the field variables p , \mathbf{u}_s and \mathbf{u}_f respectively. Let δp , $\delta \mathbf{u}_s$, $\delta \mathbf{u}_f$ be the frequency-independent test functions, associated to p , \mathbf{u}_s , \mathbf{u}_f respectively, and belonging to their respective admissible spaces C_p , $C_{\mathbf{u}_s}^* = \{\delta \mathbf{u}_s \in C_{\mathbf{u}_s} \mid \delta \mathbf{u}_s = \mathbf{0} \text{ on } \partial_1\Omega_P\}$, and $C_{\mathbf{u}_f}^* = \{\delta \mathbf{u}_f \in C_{\mathbf{u}_f} \mid \delta \mathbf{u}_f = \mathbf{0} \text{ on } \partial_1\Omega_P\}$.

Equations (1), (9), and (11) lead to:

$$\begin{aligned} \int_{\Omega_F} \mathbf{grad} p \cdot \mathbf{grad} \delta p \, dV - \frac{\omega^2}{c_0^2} \int_{\Omega_F} p \delta p \, dV - \omega^2 \rho_F (1 - \phi) \int_{\Gamma_{FP}} \mathbf{u}_s \cdot \mathbf{n} \delta p \, d\Sigma \\ - \omega^2 \rho_F \phi \int_{\Gamma_{FP}} \mathbf{u}_f \cdot \mathbf{n} \delta p \, d\Sigma = \omega^2 \rho_F \int_{\partial_2\Omega_F} u_{Fb} \delta p \, d\Sigma \end{aligned} \quad (12)$$

Equations (2a), (8a), and (10a) lead to:

$$\begin{aligned}
 & \int_{\Omega_P} \text{tr} \left[\mathbf{D}_s^{(1)} \boldsymbol{\varepsilon}(\mathbf{u}_s) \boldsymbol{\varepsilon}(\delta \mathbf{u}_s) \right] dV + \int_{\Omega_P} \text{tr} \left[\mathbf{D}_{sf}^{(1)} \boldsymbol{\varepsilon}(\mathbf{u}_f) \boldsymbol{\varepsilon}(\delta \mathbf{u}_s) \right] dV \\
 & + \left(\tilde{K}_f - P_0 \right) \left[\int_{\Omega_P} \text{tr} \left[\mathbf{D}_s^{(2)} \boldsymbol{\varepsilon}(\mathbf{u}_s) \boldsymbol{\varepsilon}(\delta \mathbf{u}_s) \right] dV + \int_{\Omega_P} \text{tr} \left[\mathbf{D}_{sf}^{(2)} \boldsymbol{\varepsilon}(\mathbf{u}_f) \boldsymbol{\varepsilon}(\delta \mathbf{u}_s) \right] dV \right] \\
 & + i\omega \tilde{b} \left[\int_{\Omega_P} \mathbf{u}_s \cdot \delta \mathbf{u}_s dV - \int_{\Omega_P} \mathbf{u}_f \cdot \delta \mathbf{u}_s dV \right] \\
 & - \omega^2 \left[\int_{\Omega_P} (\rho_s + \rho_a) \mathbf{u}_s \cdot \delta \mathbf{u}_s dV - \int_{\Omega_P} \rho_a \mathbf{u}_f \cdot \delta \mathbf{u}_s dV \right] - (1 - \phi) \int_{\Gamma_{FP}} p \mathbf{n} \cdot \delta \mathbf{u}_s d\Sigma = 0
 \end{aligned} \tag{13}$$

Equations (2b), (8b), and (10b) lead to:

$$\begin{aligned}
 & \int_{\Omega_P} \text{tr} \left[\mathbf{D}_f^{(1)} \boldsymbol{\varepsilon}(\mathbf{u}_f) \boldsymbol{\varepsilon}(\delta \mathbf{u}_f) \right] dV + \int_{\Omega_P} \text{tr} \left[\mathbf{D}_{sf}^{(1)} \boldsymbol{\varepsilon}(\mathbf{u}_s) \boldsymbol{\varepsilon}(\delta \mathbf{u}_f) \right] dV \\
 & + \left(\tilde{K}_f - P_0 \right) \left[\int_{\Omega_P} \text{tr} \left[\mathbf{D}_f^{(2)} \boldsymbol{\varepsilon}(\mathbf{u}_f) \boldsymbol{\varepsilon}(\delta \mathbf{u}_f) \right] dV + \int_{\Omega_P} \text{tr} \left[\mathbf{D}_{sf}^{(2)} \boldsymbol{\varepsilon}(\mathbf{u}_s) \boldsymbol{\varepsilon}(\delta \mathbf{u}_f) \right] dV \right] \\
 & + i\omega \tilde{b} \left[\int_{\Omega_P} \mathbf{u}_f \cdot \delta \mathbf{u}_f dV - \int_{\Omega_P} \mathbf{u}_s \cdot \delta \mathbf{u}_f dV \right] \\
 & - \omega^2 \left[\int_{\Omega_P} (\phi \rho_f + \rho_a) \mathbf{u}_f \cdot \delta \mathbf{u}_f dV - \int_{\Omega_P} \rho_a \mathbf{u}_s \cdot \delta \mathbf{u}_f dV \right] - \phi \int_{\Gamma_{FP}} p \mathbf{n} \cdot \delta \mathbf{u}_f d\Sigma = 0
 \end{aligned} \tag{14}$$

After discretization of the various terms in Eqs. (12)-(14) by the FE method and dividing Eq. (12) by ρ_F , the following matrix equation for the coupled problem is obtained:

$$\begin{aligned}
 & \left(\begin{bmatrix} \mathbf{K}_F & \mathbf{0} & \mathbf{0} \\ -(1 - \phi) \mathbf{A}_{Fs}^T & \mathbf{K}_{ss}^{(1)} & \mathbf{K}_{sf}^{(1)} \\ -\phi \mathbf{A}_{Ff}^T & \mathbf{K}_{sf}^{(1)T} & \mathbf{K}_{ff}^{(1)} \end{bmatrix} + \left(\tilde{K}_f - P_0 \right) \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{ss}^{(2)} & \mathbf{K}_{sf}^{(2)} \\ \mathbf{0} & \mathbf{K}_{sf}^{(2)T} & \mathbf{K}_{ff}^{(2)} \end{bmatrix} \right. \\
 & \left. + i\omega \tilde{b} \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{ss} & \mathbf{C}_{sf} \\ \mathbf{0} & \mathbf{C}_{sf}^T & \mathbf{C}_{ff} \end{bmatrix} - \omega^2 \begin{bmatrix} \mathbf{M}_F & (1 - \phi) \mathbf{A}_{Fs} & \phi \mathbf{A}_{Ff} \\ \mathbf{0} & \mathbf{M}_{ss} & \mathbf{M}_{sf} \\ \mathbf{0} & \mathbf{M}_{sf}^T & \mathbf{M}_{ff} \end{bmatrix} \right) \begin{bmatrix} \mathbf{P} \\ \mathbf{U}_s \\ \mathbf{U}_f \end{bmatrix} = \begin{bmatrix} \omega^2 \mathbf{U}_{Fb} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}
 \end{aligned} \tag{15}$$

This non-symmetric formulation can be symmetrized for a resolution in the frequency domain by dividing the acoustic equation by ω^2 ($\omega \neq 0$). The interest of rewriting the porous media formulation into four matrices ($\mathbf{K}_{ii}^{(1)}$, $\mathbf{K}_{ii}^{(2)}$, \mathbf{C}_{ii} , and \mathbf{M}_{ii}) is already partly visible. In fact, it involves constant real-valued matrices which can be assembled once, while only the complex and frequency-dependent factors \tilde{K}_f and \tilde{b} are recomputed at each frequency increment. In addition to that, the amount of memory used is the same as using two complex-valued and frequency-dependent matrices, as the sparsity is unchanged. More importantly, the main interest underlined in this work is the possibility to use such a formulation in the context of modal reduction techniques.

3 MODAL REDUCTION OF THE POROUS MEDIA

3.1 Presentation of the proposed solution strategy

The proposed reduction method is applied to the dissipative porous media of a poro-acoustic coupled problem, which is the costly part of the model. For the sake of conciseness, the case of a rigid acoustic cavity with a single porous layer on one wall is considered. Notations used are presented in Fig. 2.



Figure 2: Problem description for modal reduction of porous media

The acoustic degrees of freedom (dofs) are separated into internal ones (subscript \bar{I}), and those at interface with the porous media (subscript I). These notations allow easy extension of the method to problems with multiple interfaces [10]. The coupled porous media matrices are now considered, involving four matrices $\mathbf{K}_P^{(1)}$, $\mathbf{K}_P^{(2)}$, \mathbf{C}_P , and \mathbf{M}_P corresponding to the set of unknowns \mathbf{U}_P such that, for each matrix indexed by P , i.e. $\mathbf{B}_P \in \{\mathbf{K}_P^{(1)}, \mathbf{K}_P^{(2)}, \mathbf{C}_P, \mathbf{M}_P\}$:

$$\mathbf{B}_P = \begin{bmatrix} \mathbf{B}_{ss} & \mathbf{B}_{sf} \\ \mathbf{B}_{sf} & \mathbf{B}_{ff} \end{bmatrix} \quad \text{and} \quad \mathbf{U}_P = \begin{bmatrix} \mathbf{U}_s \\ \mathbf{U}_f \end{bmatrix} \quad (16)$$

Similarly, the coupling between the interface acoustic dofs (subscript I) and the porous dofs (subscript P) is denoted:

$$\mathbf{A}_{IP} = [(1 - \phi)\mathbf{A}_{Is} \quad \phi\mathbf{A}_{If}] \quad (17)$$

Consequently, for modal reduction purposes, matrix set of equations (15) can be written:

$$\begin{bmatrix} \mathbf{K}_{\bar{I}\bar{I}} - \omega^2\mathbf{M}_{\bar{I}\bar{I}} & \mathbf{K}_{\bar{I}I} - \omega^2\mathbf{M}_{\bar{I}I} & \mathbf{0} \\ \mathbf{K}_{I\bar{I}} - \omega^2\mathbf{M}_{I\bar{I}} & \mathbf{K}_{II} - \omega^2\mathbf{M}_{II} & -\omega^2\mathbf{A}_{IP} \\ \mathbf{0} & -\mathbf{A}_{IP}^T & \mathbf{K}_P^{(1)} + (\tilde{K}_f - P_0)\mathbf{K}_P^{(2)} + i\omega\tilde{b}\mathbf{C}_P - \omega^2\mathbf{M}_P \end{bmatrix} \begin{bmatrix} \mathbf{P}_{\bar{I}} \\ \mathbf{P}_I \\ \mathbf{U}_P \end{bmatrix} = \begin{bmatrix} \omega^2\mathbf{U}_{\bar{I}b} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \quad (18)$$

and can be symmetrized by dividing the acoustic equations (lines 1 and 2) by ω^2 ($\omega \neq 0$).

3.2 Modal reduction

From the proposed expression of the porous media FE problem, real-valued normal modes can be computed associated to the conservative proelastic eigenvalue problem:

$$\left(\mathbf{K}_P^{(1)} - \omega^2\mathbf{M}_P \right) \phi = \mathbf{0} \quad (19)$$

It is supposed that the Dirichlet boundary conditions imposed to the porous media result in a nonsingular $\mathbf{K}_P^{(1)}$ matrix, therefore removing zero-frequency modes. A modal reduction basis Φ_{Pm} is built, selecting m low frequency modes. They are normalized with respect to the porous mass matrix \mathbf{M}_P so that:

$$\Phi_{Pm}^T \mathbf{M}_P \Phi_{Pm} = \mathbf{I}_m \quad (20a)$$

$$\Phi_{Pm}^T \mathbf{K}_P^{(1)} \Phi_{Pm} = \mathbf{\Omega}_m \quad (20b)$$

where \mathbf{I}_m is a unit matrix of dimension m , and $\mathbf{\Omega}_m$ a diagonal matrix of same size, with the m lowest eigenvalues of (19) on its diagonal.

There are two key points that make a reduction method computationally efficient, which are its ability to:

- converge rapidly to the expected solution when adding modes in the basis, thus allowing a subsequent reduction in the number of dofs, as well as a reasonable time allocated to the computation of the modes,
- preserve or improve the sparsity of the matrices after projection, and ideally produce diagonal submatrices.

The former aspect will be examined on an example in Section 4. Regarding the second aspect, the sparsity of matrices $\mathbf{K}_P^{(2)}$ and \mathbf{C}_P after projection on the modal basis is fundamental to take advantage of the diagonal form of projected $\mathbf{K}_P^{(1)}$ and \mathbf{M}_P . The choices made for the discretization of porous media, among which the separation of the “static” and “dynamic” parts of the effective bulk modulus, seem to fulfill this requirement. In fact, as will be shown in Section 4, it results in sparse reduced $\mathbf{K}_P^{(2)}$ and \mathbf{C}_P , and even orthogonality of some modes with respect to these matrices. Therefore, after testing the m retained modes for their orthogonal properties with respect to $\mathbf{K}_P^{(2)}$ and \mathbf{C}_P , they are separated into o “orthogonal” (Φ_{Po}) and n “non-orthogonal” (Φ_{Pn}) ones, so that:

$$\begin{bmatrix} \Phi_{Pn} & \Phi_{Po} \end{bmatrix}^T \mathbf{M}_P \begin{bmatrix} \Phi_{Pn} & \Phi_{Po} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_o \end{bmatrix} \quad (21a)$$

$$\begin{bmatrix} \Phi_{Pn} & \Phi_{Po} \end{bmatrix}^T \mathbf{K}_P^{(1)} \begin{bmatrix} \Phi_{Pn} & \Phi_{Po} \end{bmatrix} = \begin{bmatrix} \mathbf{\Omega}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{\Omega}_o \end{bmatrix} \quad (21b)$$

$$\begin{bmatrix} \Phi_{Pn} & \Phi_{Po} \end{bmatrix}^T \mathbf{K}_P^{(2)} \begin{bmatrix} \Phi_{Pn} & \Phi_{Po} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\kappa}_n & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\kappa}_o \end{bmatrix} \quad (21c)$$

$$\begin{bmatrix} \Phi_{Pn} & \Phi_{Po} \end{bmatrix}^T \mathbf{C}_P \begin{bmatrix} \Phi_{Pn} & \Phi_{Po} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\zeta}_n & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\zeta}_o \end{bmatrix} \quad (21d)$$

where \mathbf{I}_n , $\mathbf{\Omega}_n$ and \mathbf{I}_o , $\mathbf{\Omega}_o$, $\boldsymbol{\kappa}_o$, $\boldsymbol{\zeta}_o$ are diagonal matrices of respective dimensions n and o , while $\boldsymbol{\kappa}_n$ and $\boldsymbol{\zeta}_n$ are non-diagonal sparse square matrices of dimension n .

There are several options for the choice of attachment functions, but in this work, the single degree of freedom (dof) per node associated to the acoustic domain is put

to advantage. Attachment functions are computed as the $\mathbf{K}_P^{(1)}$ -static responses of the porous media to unit pressure successively imposed at each interface acoustic dof:

$$\begin{bmatrix} -\mathbf{A}_{IP}^T & \mathbf{K}_P^{(1)} \end{bmatrix} \begin{bmatrix} \mathbf{I}_I \\ \boldsymbol{\Psi}_{PI} \end{bmatrix} = [\mathbf{0}] \Rightarrow \boldsymbol{\Psi}_{PI} = \mathbf{K}_P^{(1)-1} \mathbf{A}_{IP}^T \quad (22)$$

Again, assumption is made that $\mathbf{K}_P^{(1)}$ is not singular. If otherwise, a shift in frequency using the mass matrix \mathbf{M}_P can be set instead, to define pseudo-static attachment functions, but is not considered in this work.

The corresponding change of basis, leaving acoustic dofs uncondensed, is then:

$$\begin{bmatrix} \mathbf{P}_{\bar{I}} \\ \mathbf{P}_I \\ \mathbf{U}_P \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{\bar{I}} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_I & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}_{PI} & \boldsymbol{\Phi}_{Pn} & \boldsymbol{\Phi}_{Po} \end{bmatrix} \begin{bmatrix} \mathbf{P}_{\bar{I}} \\ \mathbf{P}_I \\ \boldsymbol{\alpha}_n \\ \boldsymbol{\alpha}_o \end{bmatrix} \quad (23)$$

where $\boldsymbol{\alpha}_n$ and $\boldsymbol{\alpha}_o$ are the modal coordinates vectors associated to the selected “non-orthogonal” and “orthogonal” modes respectively.

Applying change of basis (23) to symmetrized Eq. (18) leads to the following reduced set of equations:

$$\left(\begin{bmatrix} \frac{1}{\omega^2} \mathbf{K}_{I\bar{I}} - \mathbf{M}_{I\bar{I}} & \frac{1}{\omega^2} \mathbf{K}_{II} - \mathbf{M}_{II} & \mathbf{0} & \mathbf{0} \\ \frac{1}{\omega^2} \mathbf{K}_{I\bar{I}} - \mathbf{M}_{I\bar{I}} & \frac{1}{\omega^2} \mathbf{K}_{II} - \mathbf{M}_{II} - \mathbf{K}_{PI}^{(1)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Omega}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{\Omega}_o \end{bmatrix} + (\tilde{K}_f - P_0) \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{PI}^{(2)} & \mathbf{K}_{PIn}^{(2)} & \mathbf{K}_{PIo}^{(2)} \\ \mathbf{0} & \mathbf{K}_{PnI}^{(2)} & \boldsymbol{\kappa}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{PoI}^{(2)} & \mathbf{0} & \boldsymbol{\kappa}_o \end{bmatrix} \right. \quad (24)$$

$$\left. + i\omega \tilde{b} \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{PII} & \mathbf{C}_{PIn} & \mathbf{C}_{PIo} \\ \mathbf{0} & \mathbf{C}_{PnI} & \boldsymbol{\zeta}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{PoI} & \mathbf{0} & \boldsymbol{\zeta}_o \end{bmatrix} - \omega^2 \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{PII} & \mathbf{M}_{PIn} & \mathbf{M}_{PIo} \\ \mathbf{0} & \mathbf{M}_{PnI} & \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{M}_{PoI} & \mathbf{0} & \mathbf{I}_o \end{bmatrix} \right) \begin{bmatrix} \mathbf{P}_{\bar{I}} \\ \mathbf{P}_I \\ \boldsymbol{\alpha}_n \\ \boldsymbol{\alpha}_o \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{Fb} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}$$

where for $\mathbf{B}_P \in \{\mathbf{K}_P^{(1)}, \mathbf{K}_P^{(2)}, \mathbf{C}_P, \mathbf{M}_P\}$:

$$\begin{aligned} \mathbf{B}_{P_{II}} &= \boldsymbol{\Psi}_{PI}^T \mathbf{B}_P \boldsymbol{\Psi}_{PI} \\ \mathbf{B}_{P_{In}} &= \boldsymbol{\Psi}_{PI}^T \mathbf{B}_P \boldsymbol{\Phi}_{Pn} = \mathbf{B}_{P_{nI}}^T \\ \mathbf{B}_{P_{Io}} &= \boldsymbol{\Psi}_{PI}^T \mathbf{B}_P \boldsymbol{\Phi}_{Po} = \mathbf{B}_{P_{oI}}^T \end{aligned}$$

This reduction can be further improved using dynamic condensation of the “orthogonal” modal coordinates, which is rather straightforward, and not presented in this contribution.

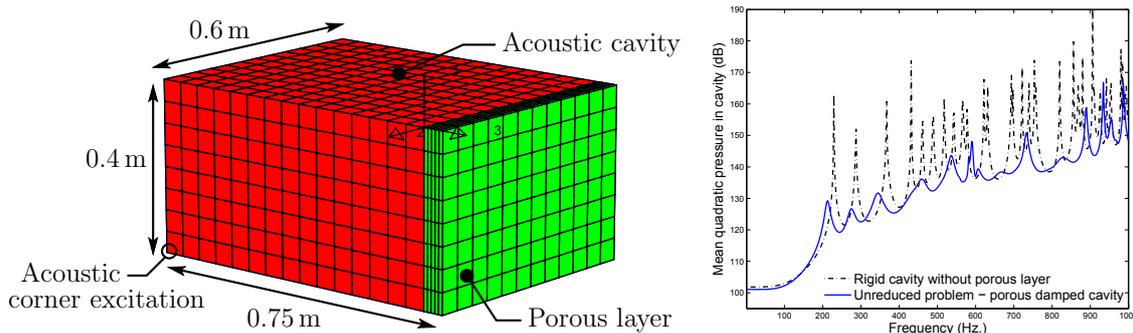
4 APPLICATION AND RESULTS

The proposed reduction of porous media is tested on a dissipative poro-acoustic example initially proposed in [6]. It consists of a 3D hexahedric acoustic cavity of dimensions $0.4 \times 0.6 \times 0.75 \text{ m}^3$ (see Fig. 3), with rigid walls, and filled with air. One wall is covered with a 5 cm-thick porous layer. The low frequency behaviour is tested applying a harmonic volume velocity source (Eq. (9)) at a corner of the cavity opposite the layer.

<i>Frame</i>	<i>Fluid</i>	<i>Porous</i>
$\lambda = 905357 \text{ Pa}$	$c_0 = 343 \text{ m/s}$	$\phi = 0.96$
$\mu = 264062 \text{ Pa}$	$\gamma = 1.4$	$\sigma = 32 \text{ kNs/m}^4$
$\rho_s = 30 \text{ kg/m}^3$	$Pr = 0.71$	$\alpha_\infty = 1.7$
	$\rho_f = 1.21 \text{ kg/m}^3$	$\Lambda = 90 \mu\text{m}$
	$\eta = 1.84 \cdot 10^{-5} \text{ Ns/m}^2$	$\Lambda' = 165 \mu\text{m}$

Table 3: Air and porous material parameters

The cavity is discretized by a $8 \times 12 \times 15$ mesh of 8-node hexahedric elements with pressure as single degree of freedom per node. The porous material, described by the Biot-Allard theory, and which material parameters are given in Table 3, is discretized by a $8 \times 12 \times 5$ mesh of 8-node hexahedric elements (Fig. 3), with 6 dofs per node corresponding to the fluid and solid phase displacements. Sticking Dirichlet boundary conditions are applied to the porous foam face in contact with the covered wall, and sliding conditions are prescribed on the side faces (see Table 2). This leads to a FE model with 1872 acoustic dofs, and 3070 porous dofs.


Figure 3: Acoustic cavity mesh and dimensions - Mean quadratic pressure reference FRF

The frequency response of the mean quadratic pressure in the acoustic cavity is given as an output (Fig. 3). The convergence is first checked, increasing the number of modes included in the basis, for a response in the range $[0 - 1000]$ Hz. Although many modes are needed in order to capture the dynamic behaviour of the porous media in the considered frequency range, the solution eventually converges toward the original solution (Fig. 4).

However, the interest of the proposed method specifically lies in the fact that real-valued modes are computed directly from the coupled porous problem, leading to good sparsity properties, as illustrated in Fig. 5. Thus, the 3070 porous media dofs are downsized to 800 modal unknowns, of which 414 correspond to $\mathbf{K}_p^{(2)}$ - and \mathbf{C}_p -orthogonal modes, as introduced in Subsection 3.2. The sparsity is mostly affected by the use of attachment functions which fully couple interface dofs to modal unknowns. From a storage perspective, using $\mathbf{K}_p^{(1)}$ -static response for attachment function leads to uncoupled interface and modal unknowns in the reduction of $\mathbf{K}_p^{(1)}$ (See Eq. 24).

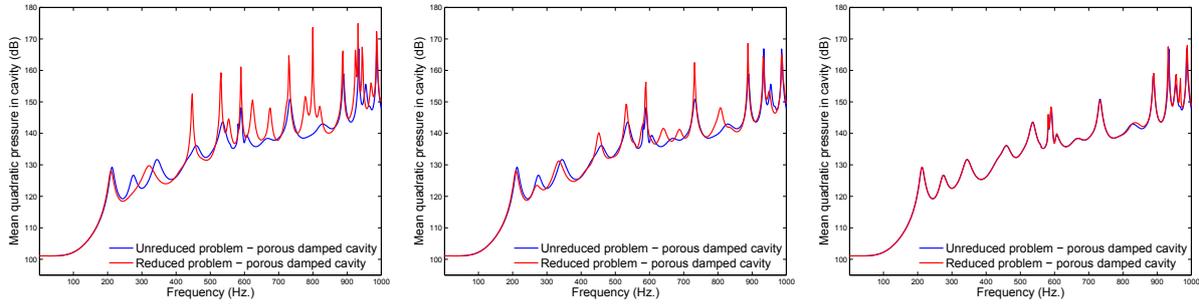


Figure 4: Mean quadratic pressure FRF. Convergence of the reduction: 100, 500, 800 modes

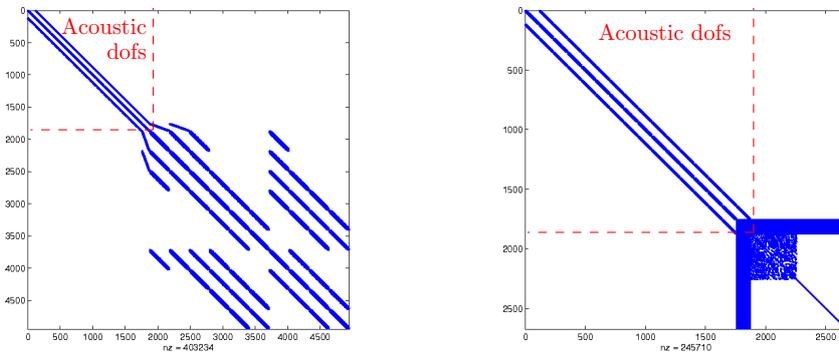


Figure 5: Sparsity of the system matrix for unreduced and reduced porous media

Regarding the computation time, a reduction including 800 modes in the basis leads to a factor 2.6 to 3.5 for the CPU time (Fig. 6), with 500 increments computed, depending whether the offset due to modes computation is taken into account or not.

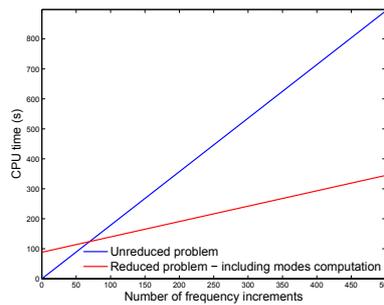


Figure 6: CPU time comparison for FRF computation

5 CONCLUSION

In this communication, the variational formulation of a harmonic poro-acoustic problem was presented. In order to improve the computational efficiency of the FE model, a modal-

based reduction of the poroelastic material was proposed, and tested on a rigid cavity treated with a porous layer. On this test application, the porous dofs were downsized from 3070 to 800 dofs, thus downsizing the coupled problem from 4942 to 2672 dofs, while preserving good sparsity properties. The computation time is therefore greatly improved. Regarding accuracy, the reduced model showed excellent match with the original problem up to 800 Hz, and good approximation up to 1000 Hz. This method can be easily combined with a similar reduction for the acoustic part evaluated by the authors [10]. Ongoing works are focusing on improvements for faster convergence to a good approximation of the original solution, as well as predictive criterion for modal truncation of the basis.

References

- [1] J. F. Allard. *Sound propagation in porous media: modelling sound absorbing materials*. Elsevier, London, 1993.
- [2] N. Atalla, M. A. Hamdi, and R. Panneton. Enhanced weak integral formulation for the mixed (u,p) poroelastic equations. *The Journal of the Acoustical Society of America*, 109(6):3065–3068, 2001.
- [3] N. E Hörlin, M. Nordström, and P. Göransson. A 3-D hierarchical FE formulation of biot’s equations for elasto-acoustic modelling of porous media. *Journal of Sound and Vibration*, 245(4):633–652, 2001.
- [4] A. Bermúdez, L. Hervella-Nieto, and R. Rodríguez. Finite element computation of the vibrations of a plate-fluid system with interface damping. *Computer Methods in Applied Mechanics and Engineering*, 190(24-25):3021–3038, 2001.
- [5] J. F Deü, W. Larbi, and R. Ohayon. Vibration and transient response of structural-acoustic interior coupled systems with dissipative interface. *Computer Methods in Applied Mechanics and Engineering*, 197(51-52):4894–4905, 2008.
- [6] P. Davidsson and G. Sandberg. A reduction method for structure-acoustic and poroelastic-acoustic problems using interface-dependent lanczos vectors. *Computer Methods in Applied Mechanics and Engineering*, 195(17-18):1933–1945, 2006.
- [7] C. Batifol, M. N. Ichchou, and M. A. Galland. Hybrid modal reduction for poroelastic materials. *Comptes Rendus Mécanique*, 336(10):757–765, 2008.
- [8] O. Dazel, B. Brouard, N. Dauchez, A. Geslain, and C. H Lamarque. A free interface CMS technique to the resolution of coupled problem involving porous materials, application to a monodimensional problem. *Acta Acustica united with Acustica*, 96(2):247–257, 2010.
- [9] H. J.-P. Morand and R. Ohayon. *Fluid Structure Interaction*. Wiley, 1995.
- [10] R. Rumpler, A. Legay, and J. F Deü. Performance of a restrained-interface substructuring FE model for reduction of structural-acoustic problems with poroelastic damping. *Manuscript submitted for publication*.

COARSE LEVEL NEWTON-KRYLOV ACCELERATION OF SUB-ITERATIONS IN PARTITIONED FLUID-STRUCTURE INTERACTION

Alexander H. van Zuijlen* and Hester Bijl†

*†Faculty of Aerospace Engineering
Delft University of Technology
P.O. Box 5058, 2600 GB Delft, The Netherlands
e-mail: A.H.vanZuijlen@tudelft.nl

Key words: High order time integration, fluid-structure interaction, partitioned integration

Abstract. Computational fluid-structure interaction is commonly performed using a partitioned approach. For strongly coupled problems sub-iterations are required, increasing computational time as flow and structure have to be resolved multiple times every time step. Reductions in computing times can be achieved by e.g. improving the convergence of the sub-iteration technique and/or performing sub-iterations on a coarse level, but also by improving the iterative solver used in the flow solver.

In this paper we investigate the combination of a multilevel acceleration technique for sub-iterations which employs a Newton-Krylov solver on the coarse level to obtain high convergence for the correction term and a multigrid solver which performs only a limited amount of iterations on the fine level to reduce memory and computing requirements. For switching between a coarse grid correction and fine grid solve, an automated coarse grid ACG(r) selection algorithm is proposed. The algorithm is applied to an academic, two dimensional test case with incompressible flow. Compared to sub-iterating with a (memory intensive) JFNK algorithm on the fine mesh, the hybrid algorithm already requires 10% less computing time. When compared to sub-iterating with multigrid, the performance increase for the hybrid scheme is a factor 3.

1 INTRODUCTION

Fluid-structure interaction in the incompressible flow regime is often encountered in engineering problems, e.g. aeroelasticity of wind turbine blades, flapping wing flight of micro aerial vehicles or underwater propulsion using deformable bodies. Especially in the field of bio inspired propulsion/control, one encounters strongly coupled fluid-structure interaction in incompressible media. These interactions are challenging to simulate in

a partitioned procedure, which is a common way of simulating complex fluid-structure interaction problems as it allows the usage of sophisticated mono-disciplinary solvers. Every time step sub-iterations are required between flow and structure solver to resolve the fully coupled solution, reducing computational efficiency.

A robust and effective sub-iteration technique with incompressible flow domains is by no means a trivial task as the coupling gets stronger as the time step is reduced [1] and straight forward sub-iterations such as block Gauss-Seidel iterations diverge. More sophisticated algorithms such as the Aitken [2] and interface quasi-Newton [3] method are required. Still, a number of sub-iterations need to be performed at every time step.

To reduce the computational work for the sub-iterations, we previously proposed a multilevel algorithm [4], which resolves a defect-correction on a coarse fluid mesh for the initial sub-iterations when still a large part of the partitioning error has a large wave number component. This part of the partitioning error can be effectively reduced on a coarse fluid mesh at a low computational expense compared to solving the fluid equations on the fine mesh. The algorithm was successfully applied for compressible flow using a density based finite volume code with nonlinear multigrid. For these cases, the flow solution of the previous sub-iteration is continuously closer to the fully coupled solutions as the sub-iterations converge. Therefore the flow does not have to be resolved to a very high precision at every sub-iteration, which limits the computational cost per sub-iteration. However, when using the same density based nonlinear multigrid solver with incompressible flow, the number of multigrid iterations required per sub-iteration was much higher [5] for the accurate representation of the pressure field, increasing computational cost per sub-iteration.

In order to overcome the poor convergence of the nonlinear multigrid solver for incompressible flow problems using a density based finite volume code, preconditioning of the equations can be applied [6]. Another approach is to use a different iterative solver, e.g. a Jacobian Free Newton Krylov (JFNK) solver [7]. This solver has the advantage that flow field (pressure field) is solved globally and the method therefore does not experience the same reduction in convergence rate as the nonlinear multigrid solver for incompressible flow. On the other hand JFNK has the disadvantage that a preconditioner is required, e.g. an ILU (incomplete LU decomposition), for fast convergence and that this preconditioner is expensive both in terms of memory and computing time. For unsteady flow the preconditioner can be reused for a number of time steps so that computational costs are reduced, but for sub-iterations in a fluid-structure interaction setting, the intermediate solutions during the sub-iterations may differ too much to use the same preconditioner especially during the initial sub-iterations [5].

In this paper we propose a hybrid algorithm which uses a multilevel acceleration for the sub-iterations, using a fixed number of nonlinear multigrid iterations on the finest mesh and a JFNK algorithm for the coarse grid correction computation. By choosing a fixed number of multigrid iterations on the finest level, the computational costs of a fine level sub-iteration is reduced at the expense of a lower convergence. This way both

a high level of convergence can be reached on the coarse level for limited computational effort and the memory requirements stay within limits as the ILU preconditioner is only required for the coarse mesh.

The question is, however, whether the relaxation of the convergence for the fine level sub-iterations, reduces the effectiveness of sub-iterating to such an extent that the computational speed gained by not iterating the solution to a high convergence on the fine mesh level is eliminated by the increase in sub-iterations to converge to the fully coupled solution. To answer this question, the hybrid algorithm is applied to an incompressible, two-dimensional test case [8] and its efficiency compared to only sub-iterating on the fine level and multilevel acceleration without JFNK.

First the two dimensional incompressible test case is given in Section 2. Next the hybrid multilevel algorithm is introduced in Section 3, where after the algorithm is used to solve the coupled problem in Section 4. Finally conclusions and recommendations are given in Section 5.

2 COUPLED PROBLEM

In this paper a fluid-structure interaction problem is addressed which is based on the benchmark problem by Turek [8]. The original problem consists of an incompressible fluid around a circular cylinder with a flexible trailing flap. The coupled problem consists of a fluid domain Ω_f , see Fig. 1, which is modeled as an incompressible fluid, and a structure

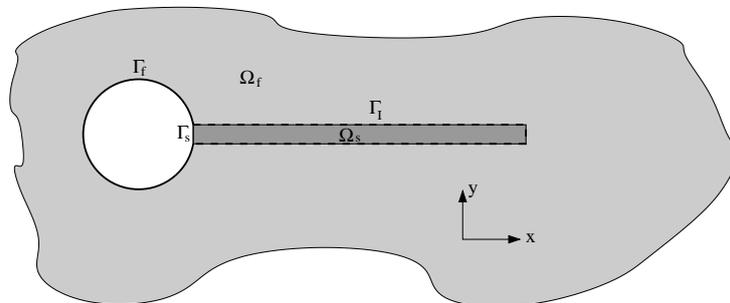


Figure 1: Computational domain.

domain Ω_s which is here modeled as a linear elastic body. The boundaries of the domain are given by Γ_s , Γ_f , which result in boundary conditions for the structure and fluid dynamics respectively. In the partitioned approach the interface boundary between the fluid and the structure domains Γ_I is denoted by two boundaries Γ_{sf} and Γ_{fs} which close the structure and fluid domains so that each domain can be treated separately from the other. The coupling between flow and structure is introduced in the boundary conditions that are imposed on Γ_{sf} and Γ_{fs} and that should yield continuity of displacement (of the interface) and stresses. Since the flap has the highest flexibility in y -direction and the shear stresses mainly act in x -direction, we simplify the continuity of stresses to a

continuity in pressure so that the conditions at the interface are

$$d_{\Gamma_{fs}} = d_{\Gamma_{sf}}, \quad (1)$$

$$p_{\Gamma_{sf}} = p_{\Gamma_{fs}}, \quad (2)$$

where d denotes the displacement of the interface boundary and p the pressure. At the moment it is still assumed that the spatial coupling is continuous and the temporal coupling instantaneous. Since the domains have been split and the coupling is performed using boundary conditions, readily available flow and structure solvers can be used to discretize and resolve their own dynamics on their own domains. Therefore, we do not address the specific spatial discretization of the solvers and simply write

$$\frac{d\mathbf{w}_s}{dt} + \mathbf{D}_s(\mathbf{w}_s, \mathbf{p}_{\Gamma_{sf}}) = \mathbf{S}_s, \quad (3)$$

$$\frac{d\mathbf{w}_f}{dt} + \mathbf{D}_f(\mathbf{w}_f, \mathbf{d}_{\Gamma_{fs}}) = \mathbf{S}_f, \quad (4)$$

wherein \mathbf{w}_s and \mathbf{w}_f are the discrete state vectors for the structure state and fluid state respectively. They contain e.g. the structural displacement or the fluid density. The spatial discretization of the governing equation is simplified by the operator \mathbf{D} , which depends both on the state \mathbf{w} and on the fluid-structure interface conditions $\mathbf{p}_{\Gamma_{sf}}$ (the discrete pressure acting on the structure) and $\mathbf{d}_{\Gamma_{fs}}$ (the displacement of the fluid domain boundary). The right hand side may contain terms \mathbf{S} that may arise from boundary conditions on Γ_s and Γ_f . Equations (3) and (4) are in semi-discrete form. We assume that the time integration is performed by the same implicit scheme in both domains. The structure and flow solver programs can then be described as solution techniques that can find solutions \mathbf{w}_s^{n+1} and \mathbf{w}_f^{n+1} under the boundary conditions $\mathbf{p}_{\Gamma_{sf}}$ and $\mathbf{d}_{\Gamma_{fs}}$ such that they satisfy (or minimize)

$$\mathbf{r}_s(\mathbf{w}_s^{n+1}, \mathbf{p}_{\Gamma_{sf}}) - \mathbf{s}_s = \mathbf{0}, \quad (5)$$

$$\mathbf{r}_f(\mathbf{w}_f^{n+1}, \mathbf{d}_{\Gamma_{fs}}) - \mathbf{s}_f = \mathbf{0}, \quad (6)$$

wherein \mathbf{r} the residual function (discretized representation of the governing equations), \mathbf{s} a constant source term within the time step that can depend on e.g. previous solutions or boundary conditions, \mathbf{p}_{Γ} the discrete pressures in the boundary nodes and \mathbf{d}_{Γ} the discrete displacements of the boundary nodes. The subscript s, f denotes that the discrete quantities belong to the structure and fluid domains respectively. A Computational Structure Dynamics (CSD) package is capable of finding a \mathbf{w}_s^{n+1} such that (5) is satisfied for a given pressure load $\mathbf{p}_{\Gamma_{sf}}$. A Computational Fluid Dynamics (CFD) package is able to find a \mathbf{w}_f^{n+1} such that (6) is satisfied for a given boundary displacement $\mathbf{d}_{\Gamma_{fs}}$. A fully implicit (or fully coupled) solution would require

$$\mathbf{r}_s(\mathbf{w}_s^{n+1}, \mathbf{p}_{\Gamma_{sf}}^{n+1}) - \mathbf{s}_s = \mathbf{0}, \quad (7)$$

$$\mathbf{r}_f(\mathbf{w}_f^{n+1}, \mathbf{d}_{\Gamma_{fs}}^{n+1}) - \mathbf{s}_f = \mathbf{0}, \quad (8)$$

wherein the superscript $n+1$ denotes the discrete solution at the new time level t_{n+1} . The coupling between (7) and (8) now poses a problem in a partitioned approach as the pressure acting on the structure interface $\mathbf{p}_{\Gamma_{sf}}^{n+1}$ depends on the fluid state \mathbf{w}_f^{n+1} and the displacement of the fluid boundary $\mathbf{d}_{\Gamma_{fs}}^{n+1}$ depends on the structure state \mathbf{w}_s^{n+1} . Both the spatial coupling (transferring data from the flow to the structure mesh and vice versa) and the temporal coupling (obtaining an implicitly coupled solution) are addressed in the next sections.

2.1 Spatial coupling

The coupling between flow and structure takes place at the fluid-structure boundary. In the continuous case, this boundary Γ_I would be identical for both fluid and structure domains, however, at the discrete level, the boundary Γ_{sf} and Γ_{fs} do not have to be matching and gaps or overlaps may occur. In this paper a radial basis function interpolation is used [9], both for the interpolation of displacements from structure to fluid mesh and for the interpolation of surface pressure from fluid to structure mesh. The interpolations \mathcal{I}_{fs} and \mathcal{I}_{sf} respectively denote the interpolation from structure to flow and vice versa.

2.2 Temporal coupling

In partitioned fluid-structure interaction, obtaining the coupled solution described by (7) and (8) would require sub-iterating, e.g. when a sequential algorithm is used

$$\mathbf{r}_s(\mathbf{w}_s^i, \hat{\mathbf{p}}_{\Gamma_{sf}}^i) - \mathbf{s}_s = \mathbf{0}, \quad (9)$$

$$\mathbf{r}_f(\mathbf{w}_f^i, \mathbf{d}_{\Gamma_{fs}}^i) - \mathbf{s}_f = \mathbf{0}, \quad (10)$$

wherein the superscript i denotes the i -th sub-iteration and $\hat{\mathbf{p}}_{\Gamma_{sf}}^i$ is the *estimation* of the fluid pressure acting on the structure for the i -th sub-iteration and $\mathbf{d}_{\Gamma_{fs}}^i = \mathcal{I}_{fs}(\mathbf{d}_{\Gamma_{fs}}^i)$ is the fluid boundary displacement obtained by interpolation of the structure boundary displacement. The simplest choice for the pressure estimation is

$$\hat{\mathbf{p}}_{\Gamma_{sf}}^i = \mathcal{I}_{sf}(\mathbf{p}_{\Gamma_{fs}}^{i-1}), \quad (11)$$

which results in a block-Gauss-Seidel type of iteration, but which is not guaranteed to be stable. To increase robustness under-relaxation can be applied, but generally at the expense of slower convergence rate. In this paper we focus on the widely applied Aitken method [2], which applies an adaptive under-relaxation to the estimation for the next time step

$$\hat{\mathbf{p}}_{\Gamma_{sf}}^{i+1} = \hat{\mathbf{p}}_{\Gamma_{sf}}^i + \theta^{i+1}(\mathbf{p}_{\Gamma_{sf}}^i - \hat{\mathbf{p}}_{\Gamma_{sf}}^i), \quad (12)$$

for which the under-relaxation parameter θ^{i+1} is obtained from

$$\theta^{i+1} = \theta^i \left(1 - \frac{(\Delta \mathbf{e}^i)^T (\mathbf{e}^i)}{(\Delta \mathbf{e}^i)^T (\Delta \mathbf{e}^i)} \right), \quad (13)$$

with $\mathbf{e}^i = \mathbf{p}_{\Gamma_{sf}}^i - \hat{\mathbf{p}}_{\Gamma_{sf}}^i$ the error between the estimated and the resulting pressure after solving (9) and (10) for iteration i and $\Delta \mathbf{e}^i = \mathbf{e}^i - \mathbf{e}^{i-1}$. For the first under-relaxation step a θ has to be chosen as \mathbf{e}^{i-1} is not available yet. One can use last known value and at the very start of the computation any (sufficiently small) value can be taken.

3 HYBRID MULTILEVEL ACCELERATION

In order to reduce the computational time for updating the fluid solution during sub-iterations, the multilevel acceleration scheme solves for a correction of the fluid solution on a coarse mesh. In the following we introduce the subscripts h and H to denote that variables or operators are defined on the fine or coarse fluid mesh respectively. In order to obtain the coarse level correction, it is first assumed that a fluid solution $\mathbf{w}_{h,f}^{i-1}$ is known (e.g. from the previous sub-iteration or previous time step). When this solution is substituted in (10) a residual is identified

$$\mathbf{r}_{h,f}(\mathbf{w}_{h,f}^{i-1}, \mathbf{d}_{h,\Gamma_{fs}}^i) - \mathbf{s}_{h,f} = \hat{\mathbf{r}}_{h,f}, \quad (14)$$

wherein $\hat{\mathbf{r}}_{h,f}$ denotes the residual in the fluid domain on the fine fluid mesh h . Subtraction of (14) from (10) gives a relation for the change in the solution $\Delta \mathbf{w}_{h,f}^i = \mathbf{w}_{h,f}^i - \mathbf{w}_{h,f}^{i-1}$ that is desired for the current sub-iteration

$$\mathbf{r}_{h,f}(\mathbf{w}_{h,f}^{i-1} + \Delta \mathbf{w}_{h,f}^i, \mathbf{d}_{h,\Gamma_{fs}}^i) - \mathbf{r}_{h,f}(\mathbf{w}_{h,f}^{i-1}, \mathbf{d}_{h,\Gamma_{fs}}^i) = -\hat{\mathbf{r}}_{h,f}. \quad (15)$$

Solving the correction $\Delta \mathbf{w}_{h,f}^i$ with an iterative solution technique is as expensive as iteratively solving (10) directly for $\mathbf{w}_{h,f}^i$. Therefore (15) is approximated on a coarse grid, reducing computational costs. It should be noted that (15) does not include the source terms $\mathbf{s}_{h,f}$ derived from boundary conditions and that the smaller the right-hand-side term (in general) the smaller the correction term that is computed, which is fundamentally different from solving (10) on a coarse mesh.

3.1 Coarse grid correction

The coarse grid correction is obtained by approximation of (15) on a coarse mesh

$$\mathbf{r}_{H,f}(\mathbf{w}_{H,f}^{i-1} + \Delta \mathbf{w}_{H,f}^i, \dot{\mathbf{x}}_{H,f}^i) - \mathbf{r}_{H,f}(\mathbf{w}_{H,f}^{i-1}, \dot{\mathbf{x}}_{H,f}^i) = -\hat{\mathbf{r}}_{H,f}, \quad (16)$$

wherein, instead of displacement of the boundary, the mesh velocity $\dot{\mathbf{x}}$ caused by the displacement of the mesh is denoted to represent the coupling between structure and flow. The mesh velocity is obtained after the displacement of the mesh by imposing the Discrete Geometric Conservation Law (DGCL) [10]. In (16) we use geometric restriction operators to obtain coarse level variables

$$\mathbf{w}_{H,f}^{i-1} = R_{\mathbf{w}} \mathbf{w}_{h,f}^{i-1}, \quad (17)$$

$$\hat{\mathbf{r}}_{H,f}^i = R_{\mathbf{r}} \hat{\mathbf{r}}_{h,f}^i, \quad (18)$$

wherein $R_{\mathbf{w}}$ and $R_{\mathbf{r}}$ are the restriction operators for the solution and residual respectively. The restriction operators can differ, depending on the variables that are stored in \mathbf{w} and $\hat{\mathbf{r}}$ are already (cell) volume scaled. If the variables are volume scaled, the restriction operator is simply summation, otherwise the restriction computes a volume weighted average on the coarse level. The coarse level mesh velocity follows from imposing the DGCL on the coarse level. When $\Delta\mathbf{w}_{H,f}^i$ is obtained on the coarse level, the solution on the fine level is corrected by prolongation of the correction term to the fine grid

$$\mathbf{w}_{h,f}^{i+1} = \mathbf{w}_{h,f}^i + P_{\mathbf{w}}\Delta\mathbf{w}_{H,f}^i, \quad (19)$$

wherein $P_{\mathbf{w}}$ is the prolongation operator for the fluid state variables \mathbf{w} . In this paper a linear interpolation is used as prolongation operator. Since the coarse grid operator $\mathbf{r}_{H,f}$ is a nonlinear operator, obtaining $\Delta\mathbf{w}_{H,f}^i$ requires an iterative solution technique.

3.2 Coarse level iterative solver

In order to reuse the iterative solvers already available for the fine mesh, the coarse grid correction (16) can be cast into the following form

$$\mathbf{r}_{H,f}(\mathbf{w}_{H,f}^i, \dot{\mathbf{x}}_{H,f}^i) + \mathbf{s}_{H,f} + \hat{\mathbf{s}}_{H,s} = \mathbf{0}, \quad (20)$$

which is equal to the original problem (10) on a coarse mesh with an additional source term

$$\hat{\mathbf{s}}_{H,s} = \hat{\mathbf{r}}_{H,f} - \mathbf{r}_{H,f}(\mathbf{w}_{H,f}^{i-1}, \dot{\mathbf{x}}_{H,f}^i) - \mathbf{s}_{H,f}. \quad (21)$$

The source term is constant during the coarse level iterations. If we denote $\mathbf{w}_{H,f}^{i,j}$ the j -th iterative approximation of $\mathbf{w}_{H,f}^i$, then the fluid residual during the coarse level iterations is defined by

$$\epsilon_{it}^j = \mathbf{r}_{H,f}(\mathbf{w}_{H,f}^{i,j}, \dot{\mathbf{x}}_{H,f}^i) + \mathbf{s}_{H,f} + \hat{\mathbf{s}}_{H,s}. \quad (22)$$

For incompressible flow with the current density based finite volume flow solver, an accurate representation of the pressure field (and thus the loads on the structure) requires a high level of iterative convergence (very low value of ϵ_{it}). E.g. when for a compressible medium only 20 multigrid iterations were required to obtain a sufficiently accurate pressure field, in the incompressible case it is roughly ten times larger. As explained in the introduction, either a nonlinear multigrid algorithm is used to minimize ϵ_{it}^j or a Jacobian Free Newton-Krylov algorithm. The JFNK algorithm initially needs more computing time to build the ILU preconditioner. However, once the preconditioner is built, the Newton updates can be efficiently evaluated using a GMRES algorithm and only few Newton updates are required to obtain a high iterative convergence.

3.3 Automated level selection

The steps above describe how a sub-iteration can be performed by computing a coarse grid correction. However, still a criterion is needed to determine whether to perform a sub-iteration on the coarse level or that it should be performed on the fine level. Previously

either a fixed number of coarse level iterations was performed followed by one fine level iteration, or the coupling was first fully resolved using coarse level iterations after which only fine level iterations were performed. The latter option has the drawback that if the coupling is resolved fully on the coarse level, still a partitioning error is present on the fine level that will not be reduced by further sub-iterating on the coarse level. The first option is more efficient, but it had to be determined manually how many coarse level sub-iterations gave the highest efficiency. Therefore a ratio is proposed to automatically determine the mesh level on which the sub-iteration is performed. The idea is that when the residual from the fine fluid mesh is restricted to the coarse mesh, any high wave number mode is filtered out by the averaging (or summation) of neighboring cells. This also means that the norm of the residual on the coarse level is reduced compared to the norm of the residual on the fine level. When it is assumed that the correction that is computed scales approximately with the norm of the residual, it can be estimated that when the residual on the coarse level is much smaller than the residual on the fine level, higher wave number errors remain dominant and are not reduced by the coarse level correction and a fine level solve is desired. Using the L_2 norms of the residual on the fine and coarse levels

$$L_{2,h} = \sqrt{\sum_{i=1}^{N_h} \frac{\hat{\mathbf{r}}_{h,i}^2}{\Omega_{h,i}}}, \quad (23)$$

$$L_{2,H} = \sqrt{\sum_{i=1}^{N_H} \frac{\hat{\mathbf{r}}_{H,i}^2}{\Omega_{H,i}}}, \quad (24)$$

wherein it is assumed that $\hat{\mathbf{r}}$ is volume scaled, so that taking the square of the residual and dividing it by the cell volume Ω is effectively multiplying the unscaled residual squared by the cell volume. This way the norms on the fine and coarse level give exactly the same value when the residual can be perfectly represented on the coarse level (e.g. residual is constant). These norms can be determined for each governing equation separately to avoid an inaccurate estimation when the equations are not solved dimensionless. The high wave number residual on the fine grid that cannot be represented on the coarse level (and hence no correction computed), can be estimated as the difference $L_{2,h} - L_{2,H}$. We now define the coarse grid residual ratio as

$$r_{CG} = \frac{L_{2,H}}{L_{2,h} - L_{2,H}}. \quad (25)$$

This ratio approaches zero when more sub-iterations are performed on the coarse level. When it is equal to one, it is estimated that the coarse level correction is of the same order as the partitioning error still remaining on the fine level due to higher wave number errors. The Automated Coarse Grid selection algorithm is now denoted by $ACG(r)$ and

selects the coarse level correction when

$$r_{CG} > r. \quad (26)$$

The smaller the value for r , the more sub-iterations are performed on the coarse level and the number of fine level sub-iterations is reduced. However, the lower the value of r more coarse grid corrections are performed that are hardly effective. Especially depending on the difference in computing time for a fine level and coarse level solve, the value of r can be typically chosen between 0.2 and 1. The faster the coarse grid correction compared to a fine level sub-iteration, the lower the value for r can be.

4 RESULTS

The proposed multilevel acceleration technique is applied to a strongly coupled two-dimensional test problem based on [8]. The test case consists of a circular cylinder of diameter 0.1m in a channel with height $H = 0.41\text{m}$, length $L = 2.5\text{m}$, with an elastic flap behind it of length $l = 0.35\text{m}$ and thickness $h = 0.02\text{m}$, see Fig. 2. The inflow is a parabolic

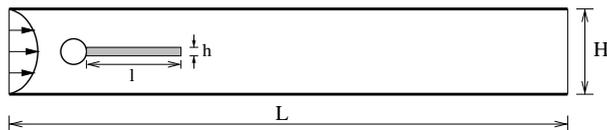


Figure 2: Two-dimensional laminar testcase.

velocity profile (see [8] for details) with a mean velocity of 2 and a maximum velocity of 3m/s. The Reynolds number is based on the mean velocity and cylinder diameter and is $Re = 200$. The structure is modeled as a linear elastic structure with a density equal to the flow density $\rho = 1000\text{kg/m}^3$ and a Young's modulus of $E = 5.6 \cdot 10^6\text{kg}/(\text{m}\cdot\text{s}^2)$. Time integration is performed by an implicit, third-order accurate, multistage Runge-Kutta scheme with a time step $\Delta t = 0.01\text{s}$ (for details on time integration see [10]). Each implicit Runge-Kutta stage is subiterated until the discrete pressure forces on the structure mesh $\|\mathbf{p}_{\Gamma_{sf}}^i - \hat{\mathbf{p}}_{\Gamma_{sf}}^i\|_2 \leq 10^{-3}$. The multigrid levels in the fluid are obtained by agglomeration of fine level cells such that the coarsening ratio is close to the theoretical optimal value of 4 (for two-dimensional applications). However, as the mesh is deforming, the lay-out of the coarse grid may change and even the amount of cells on the coarse level may differ. Therefore, the ILU preconditioner for the JFNK algorithm on the coarse level cannot be reused and has to be rebuilt for every new coarse grid sub-iteration. In this paper only the fine mesh (21307 cells) and the medium mesh (5859 cells) are used in the multilevel acceleration. However, the $ACG(r)$ can just as well be used to identify even coarser meshes to perform sub-iterations on.

First the computational work is determined for each part of the algorithm; a distinction is made between solving the fluid domain on the fine grid (FG) or solving the coarse

Table 1: Computational work (scaled by FG/MG(50)) and convergence for solving single sub-iteration.

<i>Algorithm</i>	<i>CPU time</i> [-]	<i>Overhead</i>	<i>Convergence</i> [$^{10} \log$]
FG/MG(50)	1	15%	-0.01
FG/MG(500)	8.8	2%	-0.1
FG/JFNK	1.75	9%	< -6
CG/MG(50)	0.35	43%	-0.005
CG/MG(500)	1.6	9%	-0.05
CG/JFNK	0.6	25%	< -6

Table 2: Number of sub-iterations and computational work (scaled by the work for FG/JFNK) for different fine grid and ACG(0.5) solver settings for solving one time step.

<i>Algorithm</i>	<i># CG/stage</i>	<i># FG/stage</i>	<i>CPU time</i> [-]
FG/JFNK	-	17	1
FG/MG(50)	-	220.7	6.7
FG/MG(500)	-	31	7.8
ACG(0.5)-CG/MG(50)-FG/MG(50)	306	46.7	4.9
ACG(0.5)-CG/MG(500)-FG/MG(50)	45.7	14	2.6
ACG(0.5)-CG/MG(500)-FG/MG(500)	33	9	3.9

grid correction (CG) and whether the multigrid solver is used with x iterations (MG(x)) or that the JFNK algorithm is applied (JFNK). The computational work includes any overhead costs such as the agglomeration of the mesh and file I/O. The computational work with respect to 50 multigrid iterations on the fine grid, MG(50), is given in Table 1, including the percentage of overhead and the (average) reduction in residual reached. To investigate the influence of not reaching the required convergence level before starting a new sub-iteration, the number of sub-iterations (per implicit Runge-Kutta stage) and computational work is compared for iterating only on the fine level with either JFNK, FG/MG(500) and FG/MG(50) and compared to the ACG(0.5) algorithm with multigrid solvers used on both the fine and the coarse level, see Table 2. The first part of the table shows that the FG/JFNK algorithm, with full convergence of the flow solver, requires the least amount of sub-iterations. The multigrid solver performs poorly as the computational time is increased by at least a factor 6.7. The second part of the table shows that when an automated multilevel algorithm with multigrid solves on both levels is used, an increased convergence on the coarse mesh strongly reduces the number of sub-iterations on both the coarse level (reduction of 85%) and even fine grid sub-iterations (reduction of 70%). The additional work performed on the coarse level results in a net reduction in computing time of 47%. However, when the convergence level on the fine grid is increased as well, the additional work performed for the fine grid solver is not compensated by the further reduction of sub-iterations. Therefore we choose the hybrid scheme to work with JFNK on the coarse level to obtain high convergence and MG(50) on the fine level to reduce

Table 3: Number of sub-iterations and computational work (scaled by the work for FG/JFNK) for the hybrid ACG(r) scheme for solving one time step.

<i>Algorithm</i>	<i># CG/stage</i>	<i># FG/stage</i>	<i>CPU time [-]</i>
ACG(0.25)-CG/JFNK-FG/MG(50)	41.3	4.7	0.98
ACG(0.5)-CG/JFNK-FG/MG(50)	32.3	8	0.90
ACG(1.0)-CG/JFNK-FG/MG(50)	24.3	13.7	0.93
ACG(0.5)-CG/JFNK-FG/JFNK	15.3	6	0.67

computational costs and memory requirements. The results for the hybrid scheme are presented in Table 3. The influence of the residual ratio in the ACG(r) algorithm is as expected: the smaller the value for r , the more the reduction in fine level sub-iterations at the expense of an increased number of coarse level sub-iterations. The smaller the amount of work for a coarse level sub-iteration compared to the fine level iteration, the smaller the value of r can be. In our case ACG(0.5) has the lowest computational time and although the fine grid sub-iterations only use 50 multigrid cycles and do not reduce the residual as much as JFNK, it is even 10% faster than FG/JFNK. When compared to the multilevel algorithms that use the multigrid solver on both fine and coarse grids, the hybrid scheme has a reduction in computing time by a factor of 3. Finally, the ACG(0.5) using JFNK on both the fine and coarse level shows the highest reduction in computing time of 33% compared to FG/JFNK.

5 CONCLUSIONS

In this paper a hybrid multilevel algorithm is proposed which uses JFNK to solve the coarse grid correction term and nonlinear multigrid using only a specified number of iterations to solve the fine grid fluid equations. For incompressible flow the multigrid solver has a low convergence rate and the number of sub-iterations to obtain the fully converged solution increase by a factor of 12 when only 50 multigrid iterations are performed compared to converging every sub-iteration to a high precision. However, sub-iterations can be effectively reduced by imposing a high level of convergence only for the coarse grid correction term. With the proposed automated coarse grid selection algorithm with a ratio $r = 0.5$, a reduction of 10% in computing time was obtained with the hybrid scheme compared to sub-iterating only on the fine mesh with the JFNK solver. When the hybrid scheme is compared to a multilevel algorithm which only uses the multigrid algorithm, a reduction in computing time of a factor of 3 is achieved. Although the multilevel algorithm with JFNK on both the fine and the coarse level shows even a reduction of 33% in computing time, the advantage of the hybrid algorithm is that the memory intensive JFNK is not required on the fine mesh.

Currently the fine level JFNK solver reuses the ILU preconditioner during sub-iterating, reducing the computational cost of updating the preconditioner continuously. Currently the same strategy is not employed for the coarse level JFNK solver, as the coarse mesh

is rebuilt at every sub-iteration and the agglomeration procedure can result in different coarse level meshes. When the agglomeration would remain fixed, an additional performance increase could be expected for the coarse level JFNK as the ILU preconditioner could then be reused.

REFERENCES

- [1] C. Forster, W. Wall, E. Ramm, Artificial added mass instabilities in sequential staggered coupling of nonlinear structures and incompressible viscous flows, *Comput. Methods Appl. Mech. Engrg.* 196 (7) (2007) 1278–1293.
- [2] D. Mok, W. Wall, E. Ramm, Accelerated iterative substructure schemes for stationary fluid-structure interaction, in: *First MIT Conference on Computational Fluid and Solid Mechanics*, 2001, pp. 1325–1328.
- [3] R. Haelterman, J. Degroote, D. v. Heule, J. Vierendeels, The quasi-newton least squares method: A new and fast secant method analyzed for linear systems, *SIAM J. Numer. Anal.* 47 (2009) 2347–2368.
- [4] A. v. Zuijlen, S. Bosscher, H. Bijl, Two level algorithms for partitioned fluid-structure interaction computations, *Comput. Methods Appl. Mech. Engrg.* 196 (2007) 1458–1470.
- [5] P. Lucas, A. v. Zuijlen, H. Bijl, A preconditioned jfnk algorithm applied to unsteady incompressible flow and fluid structure interaction problems, *Comput. Model. in Engrg. & Sci.* 59 (2010) 79–105.
- [6] A. Chorin, A numerical method for solving incompressible viscous flow problems, *J. Comput. Phys.* 2 (1967) 12–26.
- [7] P. Lucas, A. v. Zuijlen, H. Bijl, Fast unsteady flow computations with a jacobian-free newton krylov algorithm, *J. Comput. Phys.* 229 (2010) 9201–9215.
- [8] S. Turek, J. Hron, Proposal for numerical benchmarking of fluid-structure interaction between an elastic object and laminar incompressible flow, in: H.-J. Bungartz, M. Schaefer (Eds.), *Fluid-Structure Interaction: Modelling, Simulation, Optimisation*, Springer, 2006.
- [9] A. d. Boer, A. v. Zuijlen, H. Bijl, Review of coupling methods for non-matching meshes, *Comput. Methods Appl. Mech. Engrg.* 196 (2007) 1515–1525.
- [10] A. v. Zuijlen, A. d. Boer, H. Bijl, Higher order time integration through smooth mesh deformation for 3d fluid-structure interaction simulations, *J. Comput. Phys.* 224 (2007) 414–430.

COMPARISONS OF COUPLING STRATEGIES FOR MASSIVELY PARALLEL CONJUGATE HEAT TRANSFER WITH LARGE EDDY SIMULATION

S. Jaure*, F. Duchaine* L.Y.M.Gicquel*

*European Centre for Research and Advanced Training in Scientific Computation (CERFACS)
CERFACS, 42 Avenue Gaspard Coriolis, 31057 Toulouse Cedex 01, France
Tel : (33).(0)5.61.19.31.31 - Fax : (33).(0)5.61.19.30.00 e-mail: jaure@cerfacs.fr

Key words: Multiphysics, High Performance Computing, Coupling, Massively Parallel, Conjugate Heat Transfer

Abstract. The optimization of gas turbines is a complex multi-physic and multi-component problem that has long been based on engineer intuitions and expensive experiments or trial and error tests. Today, turbine experts commonly acknowledge that computer simulation is a very promising path for optimization, which can reduce costs and diminish the duration of the design process. Computations however remain a great challenge essentially because of the High Performance Computing (HPC) context, which is necessary for accurate estimates of real-life type of problems. Despite this difficulty, current high-fidelity computer simulations become accessible for specific components of gas turbines [5]. These stand-alone simulations and solutions now face a new challenge: to improve the quality of the results, new physics must be introduced with specific and distinct numerical models. For example, in the context of multi-component simulations, further improving the accuracy of turbine wall temperature is of limited interest if wall temperature boundary conditions are still set approximately. Dealing with multi-physics, recent studies have shown interesting results by taking into account reactive flow as well as radiative and conductive heat transfers to predict wall temperature of a helicopter combustion chamber [2, 1].

Based on the simulation of conjugate heat transfer within an industrial combustor, the current study aims at comparing different strategies of code coupling on HPC architectures. The flow solver is the Large Eddy Simulation (LES) code AVBP already ported on massively parallel architectures [5]. The conduction solver is based on the same data structure and thus has the same performances in term of parallelism. Coupling these two codes although possible requires exchanging and treating information based on two different computational grids and time evolutions. Such transfers have to be thought to maintain code scalability while maintaining numerical accuracy, thus raising communication and HPC issues: transferring data from a distributed interface to an other distributed

interface in a parallel way and on a very large number of processors is challenging and the solutions are not yet clear.

The strategies investigated in this work go from standard client/server couplers to fully distributed couplers. Although the standard client/server couplers are easier to implement, they appear to have scalability issues which fully distributed methods do not share.

1 Introduction

Multiphysical simulation is a relatively new research field within the Computational Fluid Dynamics (CFD) community, and more generally in computational physics. There are two main strategies to perform multiphysics: on the first hand one can write an all in one solver handling all of the different physics together, on the other hand several individual solvers can be interfaced together forming a new aggregated solver. The first strategy requires developing new solvers capable of handling different physics, and thus different spacial and time scales, different types of equations, etc... Whereas the second solution has the advantage of focusing the development effort on the interface between the codes which is much easier and thus the preferred solution. The latter is called code coupling.

Generic couplers already exist (MpCCI, PALM, etc...) enabling users to interface different solvers in a simple way, producing an aggregate multiphysical solver. However the structure of these tools is generally based on one or a few sequential operators providing the services to interface the different solvers. As long as the solvers remain sequential or the applications do not require an important amount of parallelism, fairly good results can be obtained without a significant impact on the performance. However it is by design incompatible with massively parallel applications: in massively parallel applications scalability is synonym with data distribution whereas centralization is synonym with bottle neck. Thus for massively parallel cases these *sequential* "couplers" could become potential bottlenecks, specially if there is a lot of inter-code exchanges.

The goal of this study is building a scalable conjugate heat transfer simulation by coupling a CFD code with a thermal conduction simulation using two CERFACS codes: AVBP for the LES and AVTP for the thermal conduction. The scalability of AVBP has already been studied [9, 5]. Even though the demonstrator is built using those codes, all the concepts presented here remain general. Because LES is used for the fluid domain, the codes must be more tightly coupled than if RANS was used. Discussing the usage of RANS over LES for coupling applications is not the goal of this paper. Though as said in [3], RANS accuracy is more limited by closure models than by mesh resolution, meaning that RANS is not necessarily able to take advantage of high computing power, unlike LES.

In this multi-code coupling environment there are two phases: the initialization phase and the run-time phase. During the initialization phase the different codes are to be

connected to each other, this implies performing geometric searches on the partitioned geometry. Then follows the run-time phase, which corresponds to the actual physical computation with the inter-code information exchanges.

In the first part of this work we will demonstrate how a sequential coupler can become a bottleneck for a massively parallel multiphysical application, either during the initialization or the run-time of the simulation. Then we will propose and assess a methodology designed for massively parallel applications.

2 Sequential coupling and massively parallel applications

The main service provided by a coupler is to enable codes to communicate at their interfaces. Here an “interface” stands for a portion of the geometrical domain which is shared between two solvers: for conjugate heat transfer it is just surfaces, but for other types of applications such as radiative coupling the “coupling interface” can be the entire domain. Because different codes operate on different physics, they may have different meshes, different processor counts and thus different domain partitioning. In order to communicate physical fields from a solver to an other, the data must be projected from the source code’s interface to the destination code’s interface. Though this operation is a simple interpolation process, in massively parallel codes (especially for CFD codes), the source and destination interfaces are distributed on different sets of the processors, making the task more complex.

How the geometry is distributed depends on the type of code, whether structured or unstructured, and on the partitioning algorithm (recursive inertial bisection, recursive coordinate bisection, etc...). To remain general we will assume that both coupled codes are unstructured and we do not make any assumption on the partitioning algorithm used. All communications are assumed to be performed using an Application Programming Interface (API) such as Message Passing Interface (MPI), which is the actual standard for inter process communications in computational physics. The conclusions could be generalized to other network type mediums, e.g., ordinary IP network. However assessing coupling strategies performed with slow mediums like temporary files is not considered here.

The most straight forward method to solve this two code conjugate heat transfer problem, demonstrated by Duchaine et al. [2] and then reused by Amaya, Poitou et al. [1], is to gather the entire distributed interface on to a service processor, perform the remapping or interpolation and then scatter this interface information on to the destination solver’s interface. This follows a many-to-one and one-to-many communication scheme (Fig. 1). In this type of scheme the entire stress is focused on the service processor:

- receiving data from all the processors can lead to network collisions and thus degraded message passing performance.
- enough memory to gather the entire interface on the service processor is mandatory, and in some cases may not be possible.

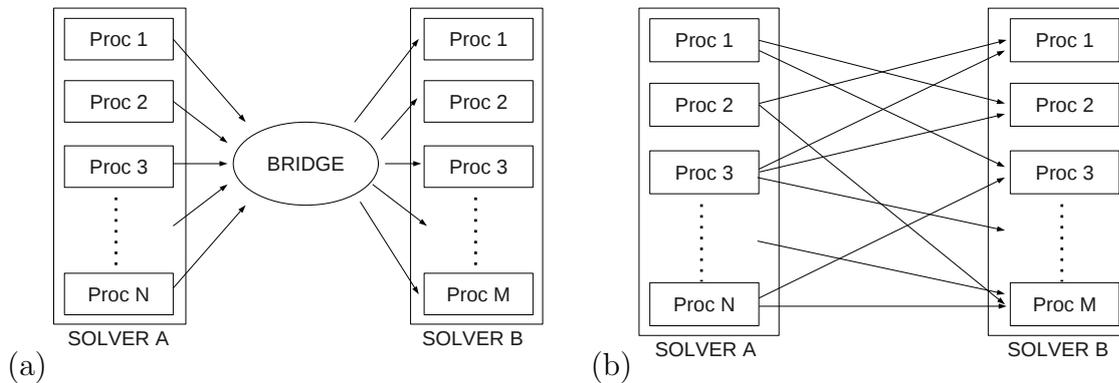


Figure 1: (a) Centralized Communication scheme (CCS), (b) Direct Communication scheme (DCS)

- the service processor has to perform all the interpolation process in a sequential way.

We can also see that in this communication algorithm there are two levels of communications introducing a high latency. This scheme will be called a *Centralized Communication Scheme (CCS)*.

A direct communication scheme (Fig. 1) can be used to transfer the data, this however requires knowing *a priori* the adjacency information between both solver's processors. For each vertex of a given code, its dependencies, i.e. the source vertices used to interpolate the value of the given vertex, must be identified (their processor and interpolation weights).

Here we will assume this information is known, we will focus on the communication scheme, a methodology to obtain this information will be proposed in section 3. We will call this scheme a *Direct Communication Scheme (DCS)*. In DCS only one level of communication is needed, thus leading to lower latency than in CCS. Locality of data is also respected and the quantity of data processed by each worker remains small leading to low memory, communication and CPU stress. Most of all, in this scheme the stress is distributed over the entire set of processes handling the interfaces instead of centralizing it on one service processor.

2.1 Scalability evaluation of the two coupling strategies

These two schemes have been compared on a SGI-ALTIX ICE super computer. For the tests, a toy application implementing both communications schemes was timed and tests were performed for small to large processor partitions: from 64 to 8192 processors.

Figure 2 shows for two size of messages that the transfer time on the CCS scales with the number of processors used, whereas the DCS maintains almost constant timings regardless of the processor count. The differences in timings have to be evaluated in CPU cost, hence multiplied by the number of processors used. On massively parallel applications this can lead to huge differences, e.g., using the data of the "small messages" (Fig. 2) on 8192 processors performing 10000 coupling iterations the CCS consumes 4200

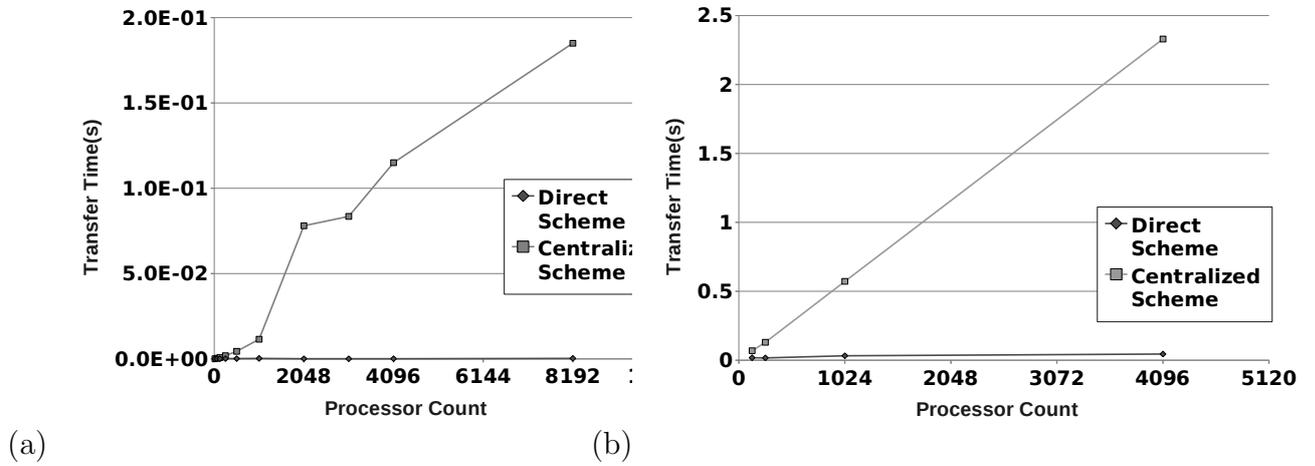


Figure 2: Coupler transfer time : (a) small messages 1Kb, (b) large messages 100Kb

CPU hours whereas less than 10 hours for the DCS (Fig. 3).

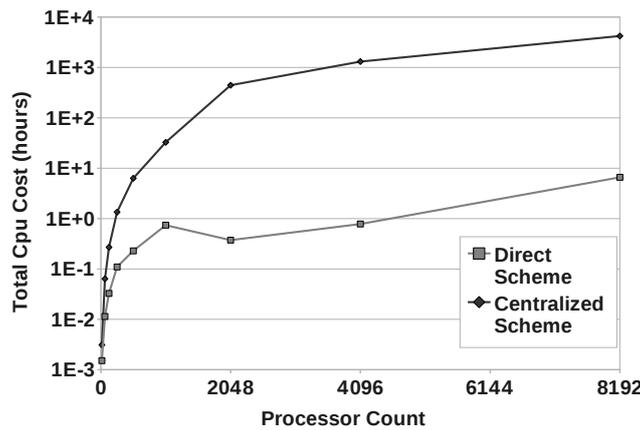


Figure 3: Coupler price in CPU Hours

The conclusion is obvious, in order to perform massively parallel coupling simulations, the communication scheme must stay direct. The result is not surprising since all scalable massively parallel codes use this communication scheme for their internal communications.

2.2 Memory requirements of the two coupling strategies

As said earlier, CCS can be limited by the amount of memory on the service processor. This is unfortunate since to take advantage of higher computing power, simulation sizes have to increase. DCS does not explicitly share that limitation. However implementing such a scheme requires being able to connect the interfaces together.

Connecting two interfaces together is fairly simple, either use a brute force nearest neighbor technique [2], or a more efficient tree based approach (Oct-Trees, balanced Kd Trees [4], etc...) if interface node count becomes high. Unfortunately these methods have a drawback: constructing trees requires memory, and such algorithms are not directly suitable for distributed geometry. The peak memory consumption criterion during the initialization phase is critical even for DCS. If the connection algorithm requires more memory locally than available on each node, no simulation can run at all. Typical examples of such constraints are some IBM Blugene machines where the available memory per core is as low as 256Mb.

3 Mesh Interface Connection

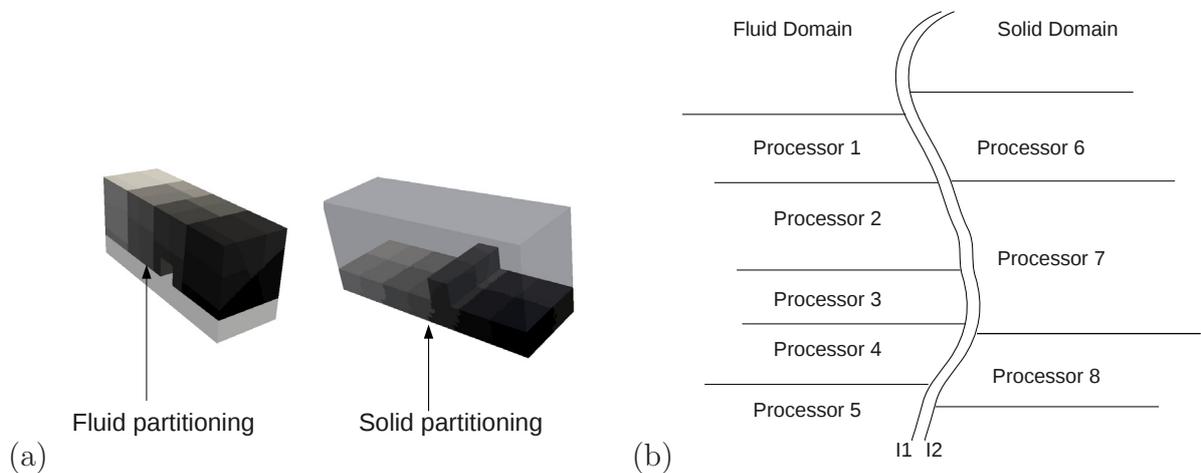


Figure 4: (a) An example of a partitioned geometry (each color group corresponds to a partition), (b) Interface partitioning

As we have seen, coupling massively parallel applications has to remain a distributed process to ensure scalability: not only during the run-time part, but also the initialization part. On one hand, distributing the work load during the run-time minimizes transfer time from a solver to an other. On the other hand, distributing the workload in the initialization improves the capacity of the coupled application to handle large simulations. Thus a fully distributed methodology to perform interface connection is proposed in the following by addressing the processors adjacency first and the interpolation coefficients in a second step.

Figure 4 illustrates the typical environment to be dealt with in massively parallel CFD applications where the meshes are partitioned into sub-domains each processed by a different processor. This partitioning also applies to the coupling interfaces (I1,I2), as represented on Fig. 4. As the partitioning algorithm is usually not aware of the coupling process, the different distributions have no reason to match, leading to complex associ-

ations between interface processors of both solvers. As the communications are bidirectional, the coupling interface connection process is separated into two phases where each solver is either server or client.

3.1 Computing coupling interface’s processors adjacency

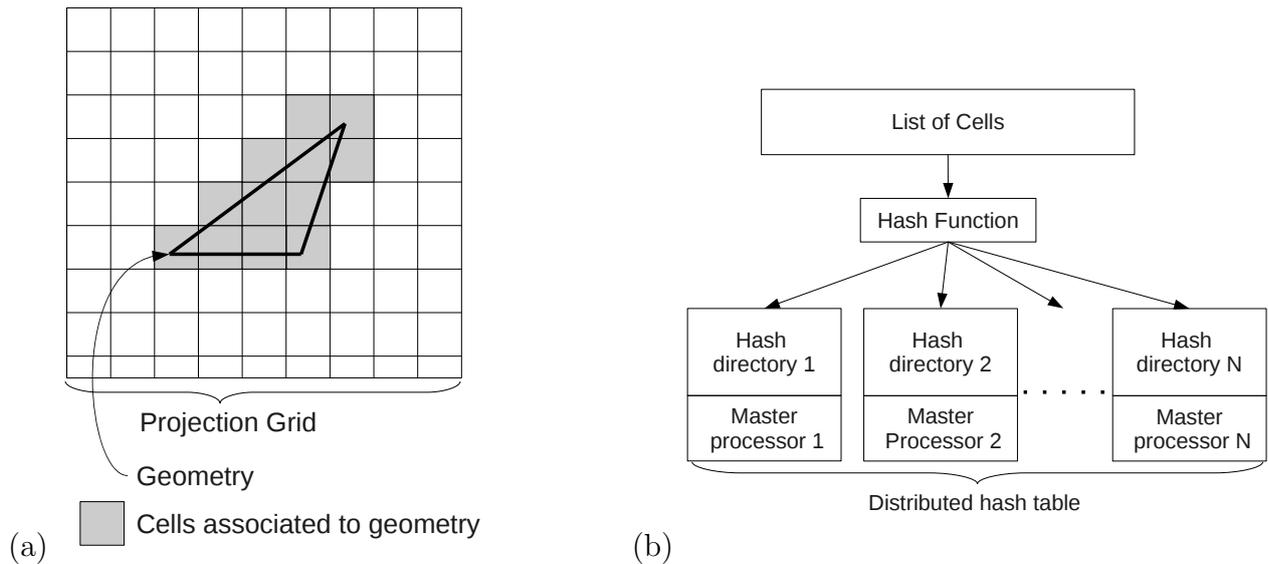


Figure 5: (a) Geometry projection to cell list,(b) Hashing of the list of cells

In order to locate the processor having parts of the interface involved in the coupling scheme, the server solver has to map its entire geometry into a distributed database. To do so each processor of the server projects its geometry on to a coarse uniform grid, associating 3D scalar vertices (x, y, z) to cell coordinates (i, j, k) (Fig. 5), each processor ending up with a list of cells to be mapped into a *Distributed Hash Table* (DHT). A DHT is a hash table ([7], [6]) for which the hash directories have been distributed over a network (Fig. 5). In this implementation each hash directory is assigned to a unique processor within the solver, the processors managing hash directories are called *master processors*. The *number of master processors* can be chosen between one up to the total number of processors available to the solver. Choosing a unique master gives a centralized scheme, and, if all processors are masters then the system becomes a peer-to-peer scheme. This entire process is a distributed version of a technique called spatial hashing [10] where the distributed hash table stores the associations between cells and processors. Knowing all this information, one can compute the list of processors associated to each cell (i, j, k) .

Then each processor of the client solver interrogates the distributed hash table stored on the server solver (using the same projection and hashing mechanism) to obtain the list of potential processors which share a portion of the coupled interface.

3.2 Computing interpolation coefficients

A first identification of the communication paths being available, each processor communicates with its potential neighbors the portion of interface they could share. Whenever possible, the neighbor processor then computes interpolation coefficients (based on an efficient Kd-Tree nearest neighbor search algorithm), and affects to those coefficients a *quality of interpolation* scalar (based on the distance between the elements to interpolate). This information is then received by the interrogating processor, which selects for each interpolated element the source providing the best interpolation quality. This information is then reorganized to produce an interpolation sparse matrix which will be used whenever a solver receives data. To conclude, each processor communicates to its neighbors the list of vertices it depends on.

3.3 Scalability and memory tests of the proposed algorithm

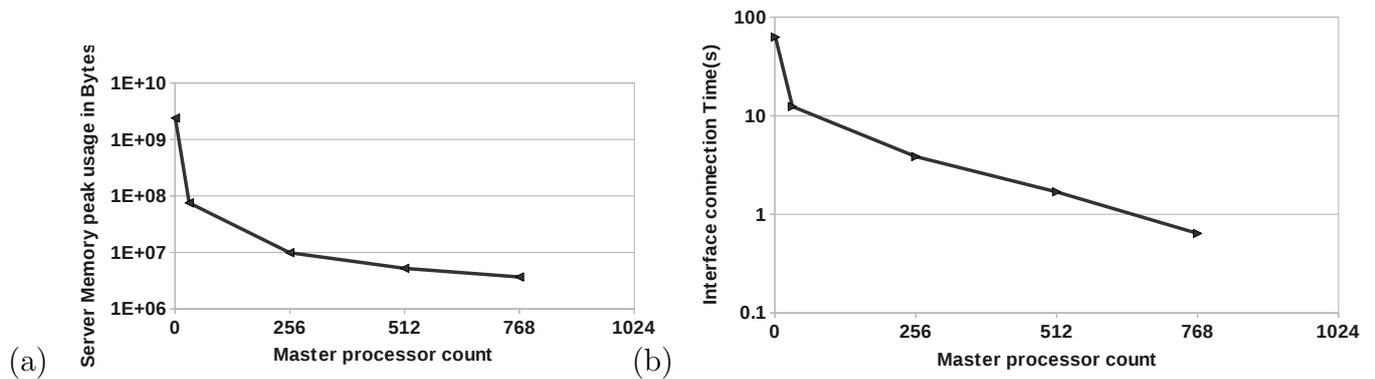


Figure 6: (a) Peak memory usage on server solver, (b) Interface connection time

As explained before, scalability must be analyzed from a general point of view, that is not only considering *CPU* stress, but also *memory* stress. The results discussed below are obtained by instrumenting the code, that is adding timers at strategic points and instrumenting the dynamic allocator.

In this first illustration, a key functionality is tested: the distributed hash table mechanism. The test is carried out in the following way:

- Two pseudo solvers are started, one will act as a server the other one as a client. The solvers do not have an equal processor count.
- A global set of nodes is distributed throughout each solver's processors.
- The server solver maps its nodes using the distributed hash table.
- Each processor of the client solver asks the server through the distributed hash table the location of its local nodes (several locations are possible).

After the first validation phase with small amounts of nodes, the real size tests are carried out on partitions of a SGI ALTIX ICE super computer. The results for 1024 cores partitioned into a server solver with 768 computing cores and a client with 256 computing cores exchanging the location of 100 million nodes are shown on Fig. 6. Local peak memory consumption used by the method for different count of master processors is specifically illustrated. The main goal has been achieved: by increasing the number of master processors, the memory load can be distributed over the solver's processors. An added benefit seen on Fig. 6 is a global application speed that is also increasing with the number of master processors. This is explained by the reduction of the quantity of data that each master processor has to process.

3.3.1 Full process tests: application on an industrial combustor

The entire interface processor adjacency computation and interpolation process has been tested on an industrial combustor configuration. The combustion chamber walls were coupled to a LES reacting flow in the burner. The test have been carried out on a 9.2 million cell fluid mesh involving 170 thousand coupled nodes and a 6.7 million cell solid mesh involving 226 thousand coupled nodes. To ensure full execution of the conjugate heat transfer problems, tests were executed in pure *peer-to-peer* mode, i.e., all processors are master processors. The test results have been summarized in table 1.

Processors		Memory peak		Connection wall clock time
AVBP	AVTP	AVBP	AVTP	
128	32	12Mb	26Mb	15.5s
256	64	9Mb	17Mb	11s
512	128	7Mb	15Mb	9s

Table 1: Full connection process test results

The primary objective which is maintaining a reasonable memory consumption has been clearly obtained. Note also that the connection timings remain relatively low and do decrease with processor count. These values do not decrease linearly with processor count: due to partitioning the processors handling the boundaries do not scale linearly with processor count. As for the transfer timings, when artificially synchronized (this would be the case for a perfect load balancing), the DCS maintained very low communication latencies: total transfer time (communication+interpolation) of the order of 1.10-4s. Since in this case, the iteration time for each code is of the order of 1s, the transfer timings can be neglected. Without artificial synchronization, because both codes execute at different speeds, they have to wait for each other during the inter-code transfers. These synchronization times depend on the load balancing. Obtaining a perfect load balancing is a difficult task, because it depends on the mesh, the solver type and the number of solver iterations between two inter-code exchanges. In the these tests, the time spent during

these inter-code synchronizations went up to 5s. The solutions to this problem are not clear yet.

4 Conclusion

This paper compares two different code coupling strategies for multi-physical massively parallel applications in CFD. Massively parallel codes in CFD obtain good scalability by distributing their geometry over large sets of processors and because each code has its own geometry distribution, connecting coupling interfaces and transferring data from a code to another is a fairly complex operation. Connecting interfaces and transferring data can be done by either using a centralized scheme or a fully distributed scheme. As discussed in this work, the centralized scheme has poor scalability and is incompatible with large simulations due to its lack of memory distribution. To circumvent this major drawback a fully distributed methodology, in terms of interface connection and data transfer, has been proposed. As demonstrated here using this methodology, very large simulations can be tightly coupled on massively parallel machines without severely degrading the performances of the codes. Other fully distributed mesh connection methods must also be stated here, notably the methodology developed in the PUNDIT program [8].

Note that the load balancing issues still have to be investigated. In practice, obtaining a good load balancing between the different codes is difficult and thus a great deal of CPU power is wasted in inter-code waits which need to be reduced. Finally this work naturally leads to future large conjugate heat transfer simulations where more physical topics will be investigated, notably coupling frequency and multiphysical solution convergence.

REFERENCES

- [1] J. Amaya, E. Collado, B. Cuenot, and T. Poinsot. Coupling LES, radiation and structure in gasturbine simulations. *Proceedings of the CTR Summer Program*, 2010.
- [2] F. Duchaine, A. Corpron, V. Moureau, F. Nicoud, and T. Poinsot. Development and assessment of a coupled strategy for conjugate heat transfer with large eddy simulation. application to a cooled turbine blade. *International Journal of Heat and Fluid Flow*, pages 1129–1141, 2009.
- [3] U.S. Department Of Energy. The opportunities and challenges of exascale computing. pages 12–13, 2010.
- [4] J. E. Goodman, J. O'Rourke, and P. Indyk. *Handbook of Discrete and Computational Geometry (2nd ed.)*, chapter Chapter 39 : Nearest neighbors in high-dimensional spaces. Chapman & Hall/CRC, 2004.
- [5] N. Gourdain, L. Gicquel, G. Staffelbach, O. Vermorel, F. Duchaine, J-F Boussuge, and T. Poinsot. High performance parallel computing of flows in complex geometries: II. applications. *Computational Science and Discovery*, 2, 2009.

- [6] R. Jenkins. *Dr Dobbs Journal*, 1997.
- [7] Donald Knuth. *The Art of Computer Programming, volume 3, Sorting and Searching*, page 506542. 1973.
- [8] J. Sitaraman, M. Floros, A. M. Wissink, and M. Potsdam. Parallel unsteady overset mesh methodology for a multi-solver paradigm with adaptive cartesian grids. *26th AIAA Applied Aerodynamics Conference*, 2008.
- [9] G. Staffelbach, J.M. Senoner, L.Y.M. Gicquel, and T. Poinsot. Large eddy simulation of combustion on massively parallel machines. *In 8th International meeting High performance computing for computational science*, pages 444–464, 2008.
- [10] M. Teschner and B. Heidelberger. Optimized spatial hashing for collision detection of deformable objects. *Proc. VMV, Munich, Germany*, 2003.

FINITE ELEMENT METHOD FOR STRONGLY-COUPLED SYSTEMS OF FLUID-STRUCTURE INTERACTION WITH APPLICATION TO GRANULAR FLOW IN SILOS

SVEN REINSTÄDLER *, ANDREAS ZILIAN * AND DIETER DINKLER *

*Institute for Structural Analysis
TU Braunschweig
Beethovenstr. 51, 38106 Braunschweig, Germany
e-mail: s.reinstaedler@tu-bs.de, www.statik.tu-bs.de

Key words: Thin-walled Structure, Non-Newtonian Fluid, Space-Time Finite Elements, Level-Set-Method, XFEM

Abstract. A monolithic approach to fluid-structure interactions based on the space-time finite element method (STFEM) is presented. The method is applied to the investigation of stress states in silos filled with granular material during discharge. The thin-walled silo-shell is modeled in a continuum approach as elastic solid material, whereas the flowing granular material is described by an enhanced viscoplastic non-Newtonian fluid model. The weak forms of the governing equations are discretized by STFEM for both solid and fluid domain. To adapt the matching mesh nodes of the fluid domain to the structural deformations, a mesh-moving scheme using a neo-Hookean pseudo-solid is applied. The finite element approximation of non-smooth solution characteristics is enhanced by the extended finite element method (XFEM). The proposed methodology is applied to the 4D (space-time) investigation of deformation-dependent loading conditions during silo discharge.

1 MOTIVATION

While the civil engineer of a silo-shell is interested in the pressure affecting the structure, the process engineer must ensure that the granular material discharges as a mass flow. If the funnel-shaped discharge zone is not well-suited for mass flow, consolidated granulars may build inner slots and so-called “dead zones” may occur. In case of kernel flow, which is also provided by an excentric discharge opening, the stress states inside the silo are unknown. If inner slots appear close to the silo wall, the structure is not loaded symmetrically anymore, which may lead to overload and failure.

2 INTRODUCTION

This contribution discusses recent developments in simultaneous (monolithic) analysis of fluid-structure interaction based on the space-time finite element method. A weighted residual-based approach to numerical analysis of fluid flow along flexible thin-walled structures, enabling the investigation of flow-induced vibrations of strongly coupled systems, is presented. Within the simultaneous solution procedure, velocity variables are used for both fluid and solid, and the whole set of model equations is discretized by a stabilized time-discontinuous space-time finite element method (TD-STFEM). The flexible structure is modeled using a three-dimensional continuum approach in a total Lagrangian setting considering large displacements. In the flow domain the incompressible Navier-Stokes equations describe the characteristics of flowing granular material as non-Newtonian fluid. A continuous finite element mesh is applied to the entire spatial domain, and the discretized model equations are assembled in a single set of algebraic equations, considering the two-field problem as a whole. The continuous fluid-structure mesh with identical orders of approximation for both solid and fluid in space and time automatically yields conservation of mass and momentum at the fluid-structure interface. A mesh-moving scheme is used to adapt the nodal coordinates of the fluid space-time finite element mesh to the structural deformation.

3 STRONG FORM OF MODEL EQUATIONS

3.1 Elasto-dynamics of the continuum-based structural model

The state of motion of deformable continua is characterized by displacements \mathbf{u} and velocities $\mathbf{v} = \dot{\mathbf{u}}$. The rate of deformation is described by the nonlinear Green-Lagrangian strain rate tensor $\dot{\mathbf{E}} = \frac{1}{2}(\dot{\mathbf{F}}^T \mathbf{F} + \mathbf{F}^T \dot{\mathbf{F}})$, where $\mathbf{F} = \mathbf{I} + \nabla_0 \mathbf{u}$ denotes the deformation gradient. The balance equation of momentum

$$\rho_0 (\dot{\mathbf{v}} - \mathbf{b}) - \nabla_0 \cdot (\mathbf{F}\mathbf{S}) = \mathbf{0} \quad (1)$$

has to be fulfilled in the reference space-time domain $Q_0 = \Omega_0 \times [0, T]$ (subscript 0). ρ_0 indicates structural density and $\rho_0 \mathbf{b}$ the body forces. The material behavior is assumed to be elastic and is described by the second Piola-Kirchhoff stress tensor \mathbf{S} , the Green Lagrange strain tensor and the tensor \mathbb{C} of elasticity, which is deduced from the strain-energy function Ψ

$$\dot{\mathbf{S}} = \mathbb{C} : \dot{\mathbf{E}} \quad \text{with} \quad \mathbb{C} = \frac{\partial^2 \Psi}{\partial^2 \mathbf{E}} \quad (2)$$

shown in rate formulation. Close to critical loads, the thin-walled shell structure may be subjected to large deformations at small strains. Therefore, the elasticity tensor of the St. Venant-Kirchhoff model can be applied, which is defined by the strain-energy function $\Psi = \frac{1}{2} \lambda (\text{tr} \mathbf{E})^2 + \mu \mathbf{E} : \mathbf{E}$ depending on the Lamé coefficients λ and μ as model parameters.

3.2 Fluid dynamics

The motion of viscous and incompressible fluids is governed by momentum balance and continuity equation

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} - \mathbf{b} \right) - \nabla \cdot \boldsymbol{\sigma} = \mathbf{0} \quad (3)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (4)$$

which are defined in the deforming space-time domain $Q = \Omega \times [0, T]$ of the current configuration. The flow characteristics of the granular material is described by a phenomenological material model for incompressible viscoplastic fluids applying a power law as a first approach, where

$$\boldsymbol{\sigma} = -p \mathbf{I} + \phi \mathbf{D} \quad \text{with} \quad \phi = \frac{p \cdot \sin(\varphi) + c}{\sqrt{J_2^{\mathbf{D}}} + \epsilon} + 2\eta \left(4 J_2^{\mathbf{D}} \right)^{\frac{\nu-1}{2}} \quad (5)$$

denotes the Cauchy stress tensor [4]. The viscosity ϕ depends on the pressure p and the second invariant $J_2^{\mathbf{D}}$ of the strain rate tensor $\mathbf{D} = \frac{1}{2} (\nabla \mathbf{v} + (\nabla \mathbf{v})^T)$, describing the kinematics. The limit state between material at rest and flow is characterized by the angle of friction φ and cohesion c as model parameters. In case of increasing strain rates the mechanical behaviour is further described by a power law, depending on the viscosity η and ν . The regularisation parameter $\epsilon \ll 1$ ensures the evaluation of stress states for disappearing strain rates, see Figure 1 .

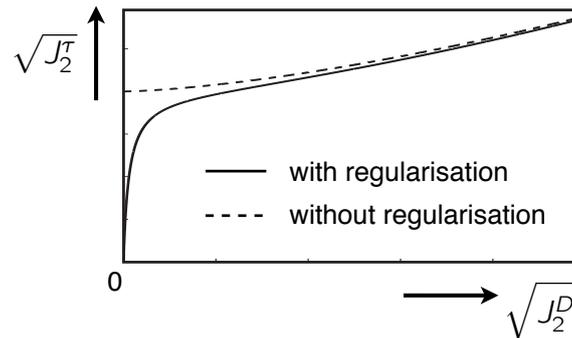


Figure 1: Stress-strain-relation [4]

4 SPACE-TIME WEAK FORM AND DISCRETIZATION

A monolithic approach to fluid-structure interaction is chosen in order to develop an accurate and uniform solution procedure for coupled systems exhibiting strong interactions [3]. The governing equations for both solid and fluid are formulated in terms of velocity variables and are discretized with the TD-STFEM [1]. A continuous finite element mesh is applied to the entire spatial domain, and the discretized model equations are assembled into a single set of algebraic equations, considering the multi-field problem as a whole.

The space-time finite element method provides a consistent discretization of both space and time, avoiding semi-discrete formulations. For the sake of efficiency, the space-time domain Q is subdivided into a sequence of space-time slabs $Q_n = \Omega \times [t_n, t_{n+1}]$ which are solved successively.

4.1 Space-time finite element formulation of geometric nonlinear elastic solids

The TD-STFEM formulation of the geometric nonlinear elastic solid within the time slab $Q_{0,n}$ reads

$$\int_{Q_{0,n}} \delta \mathbf{v} \cdot \rho_0 (\dot{\mathbf{v}} - \mathbf{b}) dQ_0 + \int_{Q_{0,n}} \dot{\mathbf{E}}(\delta \mathbf{v}, \mathbf{u}) : \mathbf{S} dQ_0 \quad (6)$$

$$+ \sum_e \int_{Q_{0,n}} \delta \mathbf{S} : \left(\mathbb{C}^{-1} : \dot{\mathbf{S}} - \dot{\mathbf{E}}(\mathbf{v}, \mathbf{u}) \right) dQ_0 \quad (7)$$

$$+ \int_{\Omega_0} \delta \mathbf{v}^T(t_n^+) \cdot \rho_0 (\mathbf{v}(t_n^+) - \mathbf{v}(t_n^-)) d\Omega_0 \quad (8)$$

$$+ \sum_e \int_{\Omega_0} \delta \mathbf{S}(t_n^+) : \mathbb{C}^{-1} : (\mathbf{S}(t_n^+) - \mathbf{S}(t_n^-)) d\Omega_0 \quad (9)$$

$$+ \sum_e \int_{\Omega_0} \boldsymbol{\tau}^S \cdot \delta \dot{\mathbf{v}} \cdot (\rho_0 (\dot{\mathbf{v}} - \mathbf{b}) - \nabla_0 \cdot (\mathbf{F} \mathbf{S})) dQ_0 = 0 \quad (10)$$

where e runs over all elements.

The resulting displacement field \mathbf{u} determined by time integration of the velocities and continuous in both space and time, is needed for computing the Green-Lagrange rate of strain tensor $\dot{\mathbf{E}}(\mathbf{u}, \mathbf{v})$ respectively the deformation gradient $\mathbf{F}(\mathbf{u})$. In the finite element formulation above, line (6) represents the weak form of momentum balance, line (7) fulfills the constitutive law for the state of thermodynamic equilibrium in weighted residual form at element level. The jump terms for velocities in (8) and equilibrium stress in (9) satisfy

the initial conditions of the time slab in integral form. The line (10) represents a Petrov-Galerkin stabilization of the momentum equation with the stabilization parameter

$$\tau_{ij}^S = \begin{cases} 0 & i \neq j \\ \frac{1}{\sqrt{\left(\frac{2}{\Delta t}\right)^2 + \left(\frac{2c_p}{\Delta x_i}\right)^2}} & i = j \end{cases} \quad (11)$$

following HUGHES and HULBERT [1] depending on the time slab width Δt , the velocity $c_p = \sqrt{\frac{\lambda+2\mu}{\rho}}$ of compression waves and the characteristic element length Δx_i in the i -th direction.

The Dirichlet boundary condition is fulfilled exactly by the approximation space for the velocity field. While homogenous Neumann boundary condition are fulfilled in weak form, surface loads can be included into the formulation by the integral $\int \delta \mathbf{v} \mathbf{h}_0 dP_0^N$ over the Neumann boundary P_0^N .

To represent the curvature of thin-walled shell-structures, second order polynomials of the serendipity family are used for geometry description of the finite elements. The same approximation is used for the velocities, which are continuous in space and discontinuous and piecewise linear in time.

In contrast, the ansatz functions for the stresses are discontinuous in both space and time, leading to a mixed-hybrid formulation in which stress variables are defined only inside finite elements. The approximation of the normal stresses is chosen with respect to the constitutive equation, assuming linear kinematics and Poisson's ratio equal to zero. In this case the normal stresses are proportional to the velocity gradient in the appropriate direction ($\sigma_{ii} \approx v_{i,i}$). Thus the approximation is trilinear in the particular direction of the stress component and perpendicular quadratic. To fulfill the LBB condition and suppress numerical phenomena like shear locking the shear stresses are approximated by lower order shape functions. To improve the performance of distorted finite elements, the locally defined functions for the stresses (Fig. 2) are mapped into the global coordinate system by the Jacobi matrix, evaluated at the center of the element (Fig. 3).

The balanced shape functions for the stresses according to a quadratic approach are deduced from the divergence-free shape functions for linear mixed finite elements developed by PIAN [5]. Comparisons have shown, that already a linear approximation of geometry and physics allow analysis of the structural behaviour of thin-walled shells.

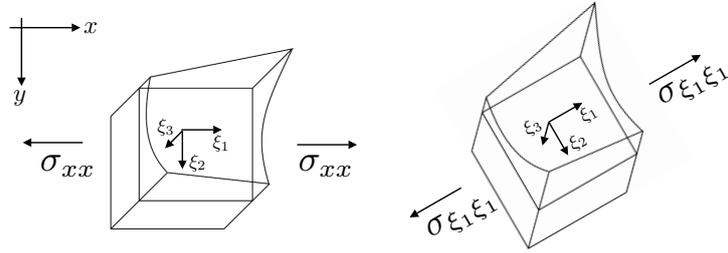


Figure 2: Locally defined ansatz functions

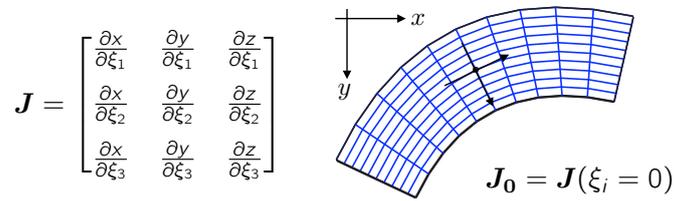


Figure 3: Evaluation of the Jacobian

4.2 Space-time finite element formulation of incompressible fluids

The time-discontinuous, stabilized space-time finite element formulation of the non-Newtonian fluid model within the time slab Q_n reads

$$\int \delta \mathbf{v} \rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} - \mathbf{b} \right) dQ \quad (12)$$

$$+ \int \mathbf{D}(\delta \mathbf{v}) : \phi(p, \mathbf{v}) \mathbf{D}(\mathbf{v}) dQ - \int \nabla \cdot \delta \mathbf{v} p dQ \quad (13)$$

$$+ \int \delta p \nabla \cdot \mathbf{v} dQ \quad (14)$$

$$+ \int \delta \mathbf{v}(t_n^+) \cdot \rho (\mathbf{v}(t_n^+) - \mathbf{v}(t_n^-)) d\Omega \quad (15)$$

$$+ \sum_e \int_{Q^e} \tau_M \frac{1}{\rho} \mathcal{L}(\delta \mathbf{v}, \delta p) \cdot (\mathcal{L}(\mathbf{v}, p) - \rho \mathbf{b}) dQ \quad (16)$$

$$+ \sum_e \int_{Q^e} \tau_C \rho \nabla \cdot \delta \mathbf{v} \nabla \cdot \mathbf{v} dQ = 0, \quad (17)$$

where line (12) and (13) represents the weak form of momentum balance and line (14) the incompressibility constraint of the continuity equation. The velocity jump term of

the time-discontinuous formulation is shown in line (15). The GLS-stabilization of the momentum equation (16) contains the residual of the momentum balance in strong form, defined as

$$\mathcal{L}(\mathbf{v}, p) = \rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) - \nabla \cdot (-p \mathbf{I} + \phi \mathbf{D}).$$

The appropriate stabilization parameter

$$\tau_M = \frac{1}{\sqrt{\left(\frac{2}{\Delta t}\right)^2 + \left(\frac{2|\tilde{\mathbf{v}}_e|}{h}\right)^2 + \left(\frac{4\nu}{h^2}\right)^2}} \quad (18)$$

derived by TEZDUYAR ET AL. [7], depends on the characteristic element length h , the time slab width Δt , the kinematic viscosity $\nu = \mu/\rho$ and the velocities $\tilde{\mathbf{v}}_e = \mathbf{v}_e - \mathbf{v}_{N,e}$ between the fluid and its moving mesh. The GLS-stabilization of the continuity equation in line (17), containing the stabilization parameter

$$\tau_C = h |\tilde{\mathbf{v}}_e| \zeta(\text{Re}_e) \quad \text{with} \quad \zeta = \begin{cases} \text{Re}_e/3 & \text{Re}_e < 3 \\ 1 & \text{Re}_e \geq 3 \end{cases} \quad \text{and} \quad \text{Re}_e = \frac{|\tilde{\mathbf{v}}_e| h}{2\nu} \quad (19)$$

completes the setup of the finite element formulation.

Finite elements with linear and quadratic approximation of the velocities in space have been investigated. The pressure is described linear in both cases, such that the stabilization of the pressure-field is needed only if an equal order (linear) interpolation of the velocities is used. The evolution of the variables \mathbf{v} and p in time is described by linear functions independently of the discretization in space.

4.3 Mesh motion approach based on a neo-Hooke pseudo-structure

The continuous fluid-structure mesh with identical orders of approximation for both solid and fluid in space and time automatically yields integral conservation of mass and momentum at the fluid-structure interface.

For large structural displacements updated coordinates have to be found for the mesh nodes inside the fluid domain. In order to avoid overlapping nodes and to reduce distortions of extremely stretched elements a pseudo-elastic continuum following a neo-Hookean material law is used as a mesh-moving scheme.

The elasticity tensor associated to the neo-Hookean material can be specified by

$$\mathbb{C} = \lambda \mathbf{C}^{-1} \otimes \mathbf{C}^{-1} + 2(\mu - \lambda \ln(J)) \mathbf{C}^{-1} \odot \mathbf{C}^{-1}, \quad (20)$$

where $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ is the right Cauchy-Green tensor and J^2 its third invariant.

5 LEVEL-SET-METHOD AND EXTENDED SPACE-TIME FINITE ELEMENT METHOD

In order to describe the motion of the free surface of the discharging granular material the level-set-method is applied [8]. The level-set-method implies the solution of the transport equation

$$\mathcal{L}_\phi = \frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi = 0 \quad (21)$$

for the distance function ϕ moved by the velocities within the fluid domain. The distance function defines the position of a material point in space and time to an interface separating two domains at zero level ($\phi = 0$). To evaluate the current position of the interface in a space-time slab $Q_n = \Omega_t \times [t_n, t_{n+1}]$, the transport equation is transferred into a time-discontinuous space-time finite element formulation with GLS-stabilization for the convective part of the material derivative as follows

$$\int_Q \delta \phi \left(\frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi \right) dQ \quad (22)$$

$$+ \int_\Omega \delta \phi(t_n^+) (\phi(t_n^+) - \phi(t_n^-)) d\Omega \quad (23)$$

$$+ \sum_e \int_Q \mathcal{L}_\phi(\delta \phi) \tau_L \mathcal{L}_\phi(\phi) dQ = 0. \quad (24)$$

In convection dominated problems the solution of the level-set-equation loses accuracy during the transport process, in particular in the vicinity of inflow-boundaries. To improve the solution of subsequent time-slabs, the level-set-function needs to be reinitialized. Therefore, the fast-marching-method [6] is used as time-efficient algorithm.

Applying the XFEM, kinks in the velocity and jumps in the pressure field between the two fluid domains can be taken into account, extending the approximation of variables u by additional shape functions M in \mathcal{N}_{ext} .

$$u_{\text{ext}}(\mathbf{x}, t) = \sum_{k \in \mathcal{N}_{\text{std}}} N_k(\mathbf{x}, t) \hat{u}_k + \sum_{j \in \mathcal{N}_{\text{ext}}} M_j(\mathbf{x}, t) \tilde{u}_j \quad (25)$$

In case of linear approximation of the physics in \mathcal{N}_{std} additional shape functions

$$M_j = N_j(\mathbf{x}, t) \psi_j^{\text{sgn}}(\mathbf{x}, t) \quad \text{with} \quad \psi_j^{\text{sgn}}(\mathbf{x}, t) = \frac{1}{2} \left(1 - \text{sgn}(\phi(\mathbf{x}, t)) \cdot \text{sgn}(\phi(\mathbf{x}_j, t_j)) \right) \quad (26)$$

based on sgn-enrichment are used to represent discontinuities in the pressure field. On the contrary, the enrichment functions

$$M_j = N_j(\mathbf{x}, t) R(\mathbf{x}, t) \quad \text{with} \quad R(\mathbf{x}, t) = \sum_{j \in \mathcal{N}^{\text{ext}}} |\phi_j| N_j - \left| \sum_{j \in \mathcal{N}^{\text{ext}}} \phi_j N_j \right| \quad (27)$$

for continuous velocities contain a ridge function R for the specification of kinks. In case of quadratic finite elements the pressure approximation is still extended by shape functions based on sgn-enrichment, whereas for the quadratic approximation space of the velocities $\mathcal{N}_{\text{std}}^{\text{quad}}$ linear abs-enrichments $N_j^{\text{lin}} \psi_j^{\text{abs}}$ are used following the recommendation of ZILIAN [9].

For the integration over the fluid domain, XFEM elements have to be subdivided into subelements, where each subelement belongs to a single fluid domain (Fig. 4). With the developed split-algorithm for 4D space-time finite elements, the velocity-field of flowing granular material in silos with kernel flow and predefined inner slots can be investigated.

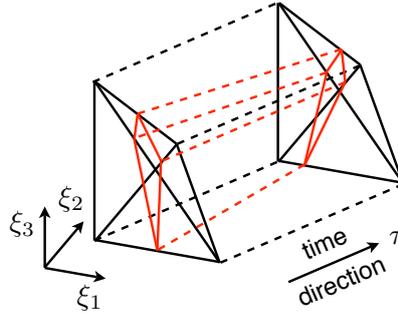


Figure 4: Intersected 4D space-time finite element

6 SOLUTION STRATEGY

To solve the transport equation for the interface the values of the velocities in the fluid domain are needed. On the other hand the enriched approximation space for the physics depends on the position of the interface, which leads to a strongly-coupled problem. To ensure the convergence of the boundary-sensitive problem, while taking the physical and geometrical nonlinearities into account, a staggered solution scheme is used. The outer loop updates the solution of the transport equation with the current velocities, whereas in the inner loop the fluid-structure-interaction problem is solved for a given approximation space by the Newton-Raphson method.

Given the position of the interface separating the two fluid domains, the nonlinear system of discretized model equations for solid (S), fluid (F) and fluid mesh dynamics (M) have to be solved iteratively. To enable a computation in acceptable time steps, the Newton-Raphson method is applied. Therefore the system of algebraic equations has to be differentiated with respect to the unknowns $\hat{\mathbf{x}}$. The updated solution to the multi-field problem may be established by computing

$$\hat{\mathbf{x}}^{i+1} = \hat{\mathbf{x}}^i - ((\mathbf{K}|_{\hat{\mathbf{x}}^i})^{-1} \mathbf{r}|_{\hat{\mathbf{x}}^i}), \quad (28)$$

where \mathbf{r} is the right hand side. The structure of the linearized system of equations, assembling matrix \mathbf{K} , is shown in Fig. 5. In the present case of fluid-structure-interaction the unknowns are given by the physical quantities $\hat{\mathbf{v}}$ and \hat{p} and the mesh motion $\hat{\mathbf{d}}$ affecting the evaluation of the discretized model equations.

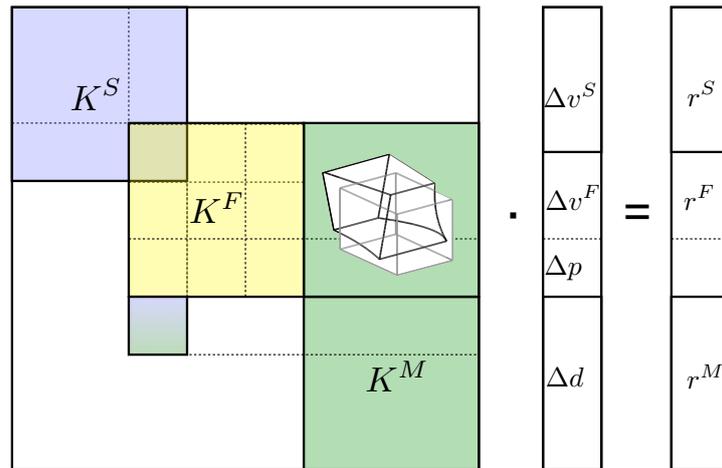


Figure 5: Structure of the strong-coupled linearized system of equations

Velocities as primary variable for both fluid and solid enable a direct coupling that implicates no-slip boundary conditions in between. Slip boundary conditions can be realized by interface elements using Lagrange multipliers to couple the velocities of the two domains at the interface in normal direction \mathbf{n} . The weak form of the kinematic condition $\mathbf{n}(\mathbf{v}_F - \mathbf{v}_S) = 0$ and equilibrium of the normal stress t at the boundary reads

$$\int \delta t (\mathbf{v}_F - \mathbf{v}_S) \cdot \mathbf{n} dP^D - \int \delta (\mathbf{v}_F \cdot \mathbf{n}) t dP^N + \int \delta (\mathbf{v}_S \cdot \mathbf{n}) t dP^N, \quad (29)$$

leading to the structure of the weak-coupled system of equations shown in Fig. 6.

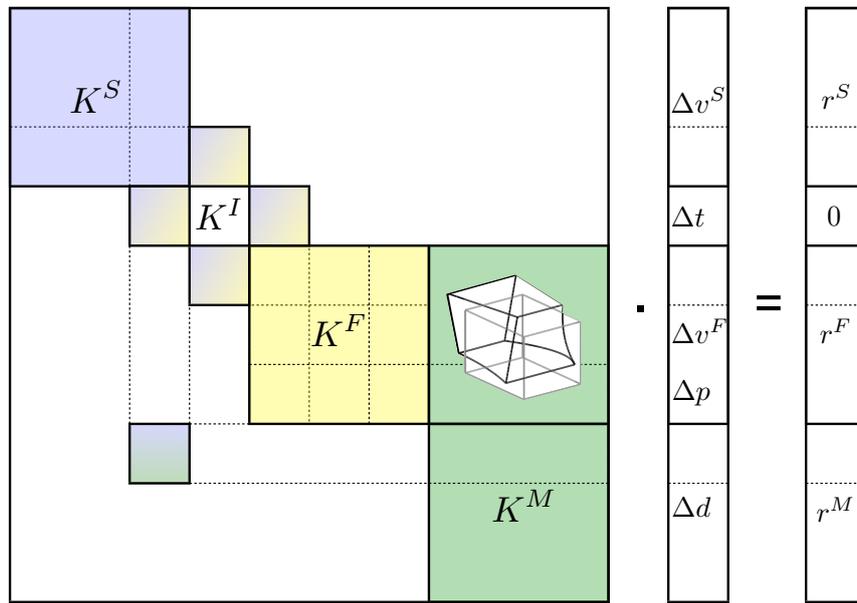


Figure 6: Structure of the weak-coupled linearized system of equations

7 VERIFICATION EXAMPLES

For verification the presented approach is applied to two types of 3D silo discharge. The first example shows a silo with a funnel-shaped discharge zone and a predefined eccentric inner slot, streamed by granular material. The dead zone is characterized by nearly inactive cohesive granular material around the inner slot. The velocity of the granular material and flow direction in the inner slot is shown in Fig. 7. Furthermore, the reaction forces around the support and symmetry plane are plotted on the superelevated deformed configuration.

As further example a silo discharge by mass flow caused by gravitation has been computed. The fluid domain inside is connected to the circumferential 3d silo structure by slip

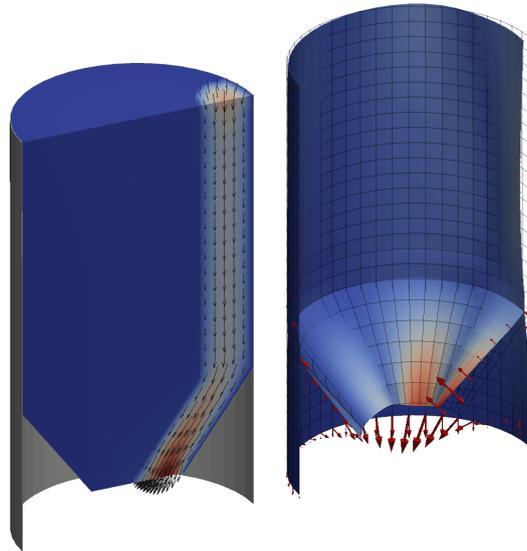


Figure 7: Velocity-field and reaction forces on the deformed configuration

boundary conditions connecting only the velocities of the two domains along the interface in normal direction. An extruded model of the finite element mesh showing the initial state of the filled silo is shown in Fig. 8 on the left. The finite element mesh consists of the shell-structure, the interface and the splitted fluid domain containing a pseudo-solid for mesh movements. The lowering of the free surface between the granular material and air above is depicted in Fig. 8.

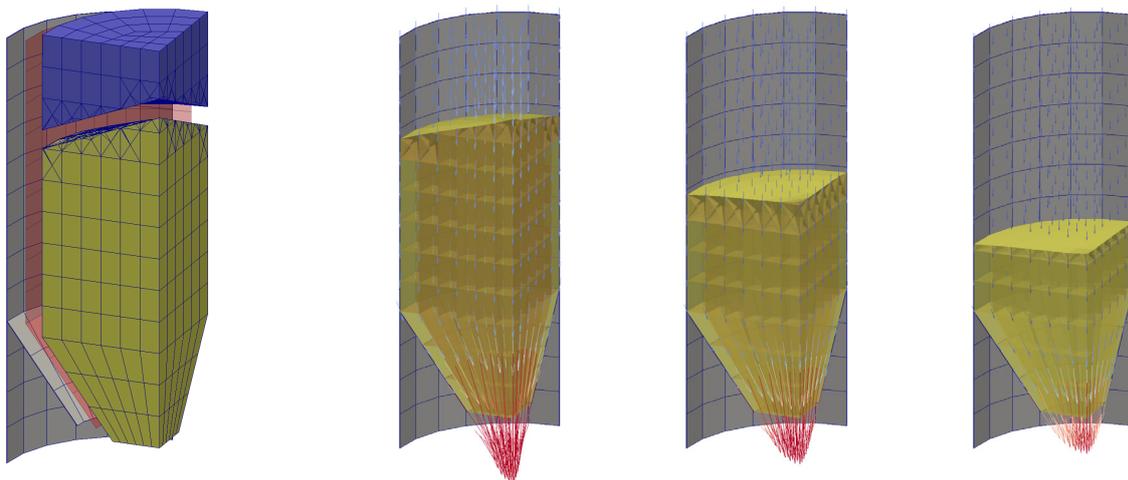


Figure 8: Extruded mesh and lowering of the free surface during emptying

REFERENCES

- [1] T. Hughes; G. Hulbert: Space-time finite element methods for elastodynamics: Formulations and error estimates. *Computer Methods in Applied Mechanics and Engineering*, 66, 3, (1988), 339–363.
- [2] B. Hübner; D. Dinkler: A simultaneous solution procedure for strong interactions of generalized Newtonian fluids and viscoelastic solids at large strains, *International Journal for Numerical Methods in Engineering* 64 (2005) 920–939.
- [3] B. Hübner; E. Walhorn; D. Dinkler: A Monolithic Approach to Fluid-Structure Interaction using Space-Time Finite Elements. *Computer Methods in Applied Mechanics and Engineering*, 193, 23–26, (2004), 2069–2086.
- [4] D. Perić; S. Slijepčević: Computational modelling of viscoplastic fluids based on a stabilised finite element method, *Engineering Computations*, Vol.18 No.3/4 (2001) 577-591.
- [5] T. H. H. Pian; K. Sumihara: Rational approach for assumed stress finite elements, *Int. J. Numer. Meth. Eng.*, 20:1685-1695, 1985.
- [6] J. A. Sethian; A. Vladimirov: Fast methods for the Eikonal and related Hamilton-Jacobi equations on unstructured meshes, *Proc. Natl. Acad. Sci. USA*, 97 (2000) 5699-5703
- [7] T. E. Tezduyar; M. Behr: Finite element solution strategies for large-scale flow simulations, *Computer Methods in Applied Mechanics and Engineering* (112) (1994) 3 – 24.
- [8] A. Zilian; A. Legay: The enriched space-time finite element method (EST) for simultaneous solution of fluid-structure interaction, *International Journal for Numerical Methods in Engineering* 75 (2008) 305–334.
- [9] A. Zilian; A. Legay: Enriched space-time finite elements for fluid-thin structure interaction, *European Journal of Computational Mechanics*, 17:725-736, 2008.

HIGHER ORDER TIME INTEGRATION SCHEMES FOR THERMAL COUPLING OF FLOWS AND STRUCTURES

V. Kazemi-Kamyab * [†], A.H. van Zuijlen[†] AND H. Bijl[†]

[†]Aerodynamics Section
Faculty of Aerospace Engineering
Delft University of Technology
P.O. Box 5058, 2600 GB Delft, The Netherlands
* Email: v.kazemikamyab@tudelft.nl

Key words: Higher Order Implicit Time Integration, IMEX, Partitioned Method, Conjugate Heat Transfer

Abstract. The application of higher order implicit time integration schemes to conjugate heat transfer problems is analyzed with Dirichlet-Neumann as the decomposition method. In the literature, only up to second order implicit time integration schemes have been reported while there is a potential for gaining computational efficiency using higher orders. For loose coupling of the domains, the IMEX scheme consisting of the ESDIRK scheme for integrating the governing equations within the subdomains and an ERK scheme for explicit integration of the explicit coupling terms is utilized. The IMEX scheme is analyzed for two cases. In one, the material properties of the coupled domains are the same and in the other they are different. While for both cases, the IMEX scheme preserves the design order of the time integration scheme, different stability and accuracy properties are observed for the two. Finally, the computational efficiency of the higher order IMEX schemes relative to the second order θ scheme is demonstrated using a test case in 2-D involving coupled conduction problem of three domains.

1 INTRODUCTION

Thermal interaction of flows and structures, also known as conjugate heat transfer, arises in many engineering disciplines such as aerospace (turbine blades in jet engines), manufacturing (continuous casting), and MEMS (cooling of electronic chips). In order to obtain a better understanding of the physics of the coupled problem and hence to increase the efficiency and/or safety of the designs, numerical simulation serves as a viable tool. However, accurate numerical solution of transient conjugate heat transfer problems (CHT) can be time consuming and thus methods to achieve the desired accuracy with less computational work is of great importance.

To improve computational efficiency, higher order implicit time integration schemes as opposed to the traditionally used 1st and 2nd order implicit schemes are investigated. For this purpose, a family of multi-stage implicit Runge-Kutta schemes (IRK) is analyzed which can be made of arbitrary higher order while retaining robustness (stability) and efficiency. For partitioned solve of the domains, the Dirichlet-Neumann decomposition is used to tackle the spatial coupling of the domains and for time integration a higher order mixed implicit-explicit (IMEX) scheme similar to one in [1] (used for partitioned solve of the mechanical coupling of flows and structures) is considered. In the application of the higher order IMEX schemes to the loose partitioned solve of the CHT problem, the time integration's design order preservation, their efficiency relative to lower order schemes, and their stability are investigated. In particular, their behavior (stability and accuracy) is analyzed for two cases. In one, the material properties of the coupled domains are the same and in the other they are different.

2 MODEL PROBLEM

In order to investigate the properties of the numerical algorithm for thermal coupling of the domains, a model problem consisting of the one dimensional transient heat conduction in two sub-domains $\Omega_1 = [-1, 0]$ and $\Omega_2 = [0, 1]$ which are separated by the common interface Γ at $x = 0$ is considered. The two subdomains can have the same or different material properties. For simplicity, it is assumed that the material properties of each sub-domain (k thermal conductivity, c_p heat capacity, and ρ density) are constant. The governing equation within each subdomain is given by:

$$(\rho c_p)_1 \frac{\partial T_1}{\partial t} = -\frac{\partial}{\partial x} \left(-k_1 \frac{\partial T_1}{\partial x} \right) \quad -1 \leq x \leq 0, \quad (1)$$

$$(\rho c_p)_2 \frac{\partial T_2}{\partial t} = -\frac{\partial}{\partial x} \left(-k_2 \frac{\partial T_2}{\partial x} \right) \quad 0 \leq x \leq 1. \quad (2)$$

The non-interface boundaries at $x = \pm 1$ are insulated. For a well-posed problem, the continuity of the temperature and heat flux is imposed at the common interface of the domains, $x = 0$.

$$T_1(\Gamma) = T_2(\Gamma), \quad (3)$$

$$-k_1 \frac{\partial T_1}{\partial x} \Big|_{\Gamma} = -k_2 \frac{\partial T_2}{\partial x} \Big|_{\Gamma}. \quad (4)$$

3 IMPOSING INTERFACE BOUNDARY CONDITIONS

To solve the coupled problem, domain decomposition method, in particular the Dirichlet-Neumann (D-N), is used where the global domain $\Omega = [-1, 1]$ is split into Ω_1 and Ω_2 . The Dirichlet (temperature) condition is assigned to Ω_2 with the interface temperature

of Ω_1 being prescribed as its value and the Neumann (the flux) condition to Ω_1 with the interface heat flux of Ω_2 being prescribed as its value thus satisfying the two interface conditions in Eqn.(3) and Eqn.(4).

In the partitioned approach, the stability of the coupling algorithm depends on the correct assignment of the interface conditions to the subdomains. Henshaw [2] analyzed the stability and rate of convergence of the interface iterations with the space kept continuous, and with the θ scheme for time integration and arrived at a criteria for imposing the interface conditions. With reference to the assigned interface boundary conditions in the model problem considered here, the criteria is:

$$|R| \approx \frac{k_2}{k_1} \sqrt{\frac{\alpha_1}{\alpha_2}} \quad (5)$$

The assigned interface boundary conditions are stable if $R \leq 1$, otherwise the two conditions must be interchanged. Here, it is assumed that, based on the material properties of the two subdomains, the imposed interface boundary conditions (Dirichlet condition imposed on Ω_2 and Neumann condition on Ω_1) satisfy the criteria. Based on Eqn.(5), when the sub-domains have similar material properties $R \approx 1$, the Dirichlet-Neumann formulation will encounter difficulties and convergence rate of the interface iterations is slow. However, when the material properties of the two subdomains are far apart, $R \ll 1$, the interface iterations converge rapidly.

4 SEMI-DISCRETE FORM

To solve the coupled PDE, space and time are discretized separately, with the space being discretized first in order to arrive at the semi-discrete form of the coupled problem. With reference to the imposed interface boundary conditions, in the case of coupling a fluid and a solid, typically $\Omega_s = \Omega_1$ and $\Omega_f = \Omega_2$. Therefore, following a common practice, Ω_1 is discretized using linear-basis finite elements (FEM) and Ω_2 using cell centered finite volumes (CFV) (see Fig.(1)). This combination of space discretizations for thermal coupling of the domains has been in used in the literature (for example see [3]). With the space discretized, the interface boundary conditions for the subdomains are evaluated, Eqn.(6) and Eqn.(7).

$$T_{2,1/2} = T_{1,0}, \quad (6)$$

$$q_{1,0} = -\frac{k_{2,1}}{\frac{\Delta x_{2,1}}{2}} (T_{2,1} - T_{2,1/2}). \quad (7)$$

The two equations are essentially the discrete form of the interface conditions Eqn.(3) and Eqn.(4). In order to present the model problem in matrix form, for simplicity, the

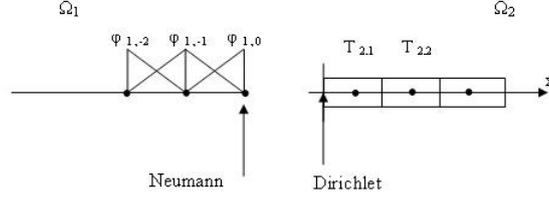


Figure 1: Discretization of the computational domain using FEM-FVM

interface temperature of Ω_2 is directly expressed by the interface node of Ω_1 . Therefore, using Eqn.(6), Eqn.(7) can be expressed by:

$$q_{1,0} = -\frac{k_{2,1}}{\frac{\Delta x_{2,1}}{2}}(T_{2,1} - T_{1,0}) = S'_1 T_1 + S_{12} T_2. \quad (8)$$

The expression on RHS of Eqn.(8) is the matrix representation of $q_{1,0}$ based on the nodal temperatures of the two domains. Since $q_{1,0}$ is defined at the interface, the matrices S'_1 and S_{12} contain mainly zero entries (for this 1-D problem, each has only one non-zero entry). For the two sub-domains after discretization in space, the following two coupled semi-discrete forms are obtained:

$$\frac{d}{dt}(M_1 T_1) = S_1 T_1 + S'_1 T_1 + S_{12} T_2 + b_1 \quad , \quad (9)$$

$$\frac{d}{dt}(M_2 T_2) = S_2 T_2 + S_{21} T_1 + b_2 \quad . \quad (10)$$

T_1 and T_2 are vectors containing the unknown temperature nodes, M_1 and M_2 diagonal matrices (mass-matrix), S_1 and S_2 are sparse tridiagonal matrices (stiffness matrix) and b_1 are b_2 are vectors containing the non-interface boundary terms. $S_{21} T_1$ represents the interface temperature assigned to Ω_2 . Similar to S'_1 and S_{12} , S_{21} contains only one non-zero entry for this 1-D test case.

5 TIME INTEGRATION

By applying a time integration scheme to the semi-discrete systems Eqn.(9) and Eqn.(10), a fully discrete system of the coupled problem can be obtained. Here, the higher order implicit ESDIRK schemes which are L-stable are considered. The L-stability property is desired since it allows for robust and stable treatment of stiffness within the system and thus the time step is restricted by accuracy rather than stability. For an ODE system of the form $\frac{dT}{dt} = \mathcal{F}(T, t)$, the solution at each stage of the ESDIRK scheme is evaluated by:

$$T^{(k)} = T^n + \Delta t \sum_{i=1}^k a_{ki} \mathcal{F}^{(i)} \quad , \quad (11)$$

with a_{ki} the coefficients of the corresponding stage. High order solution at the next time level can be achieved by the weighted sum of the residual functions such that the lower order errors cancel out:

$$T^{n+1} = T^n + \Delta t \sum_{i=1}^s b_i \mathcal{F}^{(i)} \quad , \quad (12)$$

where b_i are the weight factor with $\sum_i b_i = 1$.

5.1 Partitioning

In the monolithic approach all the interface terms are evaluated implicitly, however, in the partitioned method some or all of the interface terms are treated explicitly, depending on the arrangement with which the two solvers are solved. Following [1], an Additive Runge-Kutta scheme is used which consists of an Explicit first stage, Singly Diagonally Implicit Runge-Kutta (ESDIRK) scheme for integrating the governing equations within the subdomains and an ERK scheme for explicit integration of explicit coupling terms and hence referred to as an IMEX scheme. For the IMEX scheme considered here, the stage coefficients given in [4] are used. The coefficients of the implicit ESDIRK scheme are denoted by a_{ki}^I and that of the ERK scheme by a_{ki}^E and both the schemes have the same weight factors b_i . Here, the coupled system is solved using the block Gauss-Seidel with integrating first Ω_2 (GS-21). The solution to the temperature field in Ω_2 at stage k , $T_2^{(k)}$, using the IMEX scheme is given by:

$$(M_2 - \Delta t a_{kk}^I S_2) T_2^{(k)} = E_2^{(k)} + \Delta t a_{kk}^I (S_{21} T_1)^* + \Delta t a_{kk}^I f_2^{(k)} \quad , \quad (13)$$

with

$$E_2^{(k)} = M_2 T_2^n + \Delta t \sum_{i=1}^{k-1} a_{ki}^I (S_2 T_2 + S_{21} T_1 + f_2)^{(i)} \quad . \quad (14)$$

Now using the updated solution field in Ω_2 , $T_1^{(k)}$ is computed:

$$(M_1 - \Delta t a_{kk}^I S_1) T_1^{(k)} = E_1^k + \Delta t a_{kk}^I (S_{12} T_2)^{(k)} + \Delta t a_{kk}^I (S_1' T_1)^* + \Delta t a_{kk}^I f_1^{(k)} \quad , \quad (15)$$

with

$$E_1^{(k)} = M_1 T_1^n + \Delta t \sum_{i=1}^{k-1} a_{ki}^I (S_{12} T_2 + S_1' T_1 + S_1 T_1 + f_1)^{(i)} \quad . \quad (16)$$

E_1 and E_2 represent the known contributions from the previous stages and/or step and $(S_{21} T_1)^*$ and $(S_1' T_1)^*$ are the coupling terms that need to be predicted. The choice of the predictor must be such that the design order of the time integration scheme is preserved -without the use of sub-iterations. Since the coupled problem is linear, only T_1^* needs to be predicted which following [1] is given by:

$$T_1^* = \sum_{i=1}^{k-1} \frac{a_{ki}^E - a_{ki}^I}{a_{kk}^I} T_1^{(i)} \quad , \quad (17)$$

The solution at the next time step is obtained by applying Eqn.(12) to each subdomain (noting that for this 1-D test case, M_1 and M_2 are constant):

$$T_1^{n+1} = T_1^n + \Delta t \sum_{i=1}^s b_i M_1^{-1} (S_{12} T_2 + S_1' T_1 + S_1 T_1 + f_1)^{(i)} \quad (18)$$

$$T_2^{n+1} = T_2^n + \Delta t \sum_{i=1}^s b_i M_2^{-1} (S_{21} T_1 + S_2 T_2 + f_2)^{(i)}. \quad (19)$$

In practice where two separate solvers are used, in Eqn.(17), T_1 is replaced with the interface nodes (for the 1-D problem considered here with $T_{1,0}$). In addition, to use the IMEX scheme, the solution field of the computed stage must be stored for evaluating the later stages and the solution to the time step.

The IMEX scheme introduced will be applied to two cases of $R \ll 1$ and $R = 1$.

6 RESULTS

The model problem described is used to assess the the accuracy and stability of the higher order IMEX scheme. The length of each subdomain is unit meter. Ω_1 is discretized using 256 elements, and Ω_2 with 256 cells. A step initial condition is imposed on the global domain: $T_i \in \Omega_1 = -1$ and $T_i \in \Omega_2 = 1$. Block Gauss-Seidel integrating Ω_2 first, GS-21, is used to solve the coupled system. As Table.(1) shows two sets of material properties are used. Even though Eqn.(5) is used to approximate the rate of convergence of the interface iterations for the θ scheme, as it will be demonstrated, it can also be used as means to estimate whether loose coupling of the domains using IMEX will encounter any difficulties.

Table 1: Model problem material properties

Set	Subdomain	$k(W/mK)$	$\rho c_p(J/m^3K)$	R
1	1	25	5000*500	.05
	2	0.025	1000*1	
2	1	25	5000*500	1
	2	25	5000*500	

To evaluate the differences between the two cases, the partitioning error is investigated. In this model problem, the two sources of error in the partitioned solution are the

partitioning and time integration errors. The partitioning error is defined as the difference between the partitioned and monolithic solutions both computed with the same time step. The time integration error is evaluated by computing the difference between the monolithic solution obtained using a certain time step and the temporally exact solution. The temporally exact solution was acquired by solving the coupled problem monolithically using a fine time step, here $\Delta t = 0.01$ sec. When the partitioned error is below the time integration error, the dominant source of error in the partitioned solution is the time integration error.

6.1 Case of $R \ll 1$

Fig.(2(a)) shows the convergence plot of the partitioning error of the third and fourth order IMEX schemes (denoted by IMEX3 and IMEX4). In addition, the convergence plot of time integration errors are shown. The partitioning error of the IMEX schemes, in the asymptotic range, have the same order as their corresponding time integration errors, thus allowing the partitioning solution to have the design order of the time integration scheme. For IMEX3, the 3^d order slope is observed more clearly if smaller time steps are considered. The figure also shows that when previous stage solution is used as predictor for the explicit coupling terms, reduction to second order (the stage order) is observed.

For IMEX3, the partitioning error is below the time integration error for all the time steps considered here. For IMEX4 it is above the time integration error in the asymptotic range. It should be noted that for both schemes, as the grid is further refined the partitioning error increases and for IMEX3 eventually moves above the time integration error. As it will shown later, for a 2-D test case the partitioning errors for both schemes are above the time integration error at rather coarse grids. In such cases, subiterations (or interface iterations) might need to be performed in order to increase the accuracy of the partitioned solution. A question that arises is whether gain in computational efficiency relative to second order time integration schemes is observed when the higher order ESDIRK schemes with subiterations are used for the partitioned solve of the domains.

An important behavior of the higher order IMEX schemes that is observed in their application to cases where $R \ll 1$ is that the solution remains stable even for time steps on the order of the time scale of diffusion within the problem (for this test case, estimated by $t_d \propto \frac{L_1^2}{\frac{k_1}{\rho_1 c_{p,1}}} = 10^5$), even though they are partly explicit.

6.2 Case of $R = 1$

Fig.(2(b)) shows the convergence plot of the partitioning error for the third and fourth order IMEX schemes for $R = 1$. In addition, the time integration errors of the third and fourth order ESDIRK schemes are also plotted. Just as in the case of $R \ll 1$, the IMEX schemes preserve the design order of the time integration for $R = 1$ in the asymptotic range. For both the IMEX schemes, the partitioning errors are above their respective time integration errors and in comparison to their counterparts in the $R \ll 1$, they

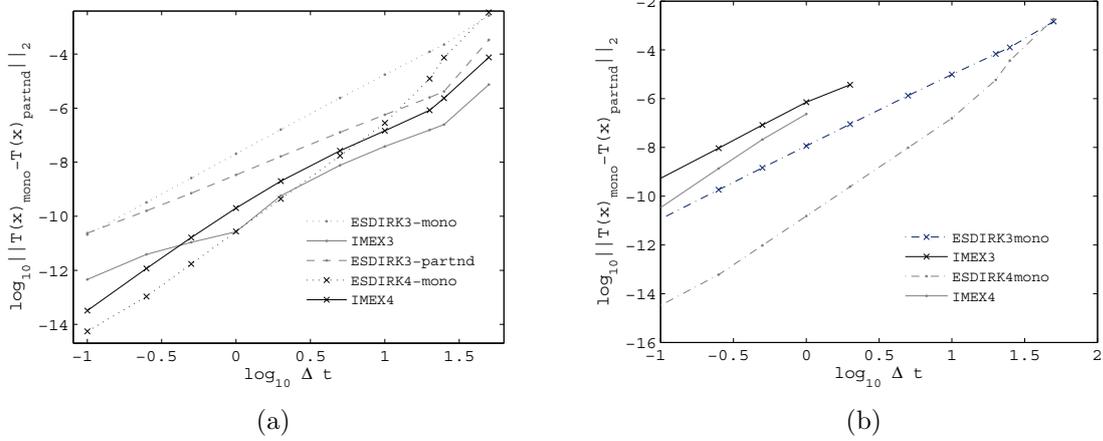


Figure 2: Convergence plot of the partitioning and time integration errors for 3rd and 4th order schemes for case of (a) $R \ll 1$ and (b) $R = 1$.

are larger. Thus, by comparing the behavior of the schemes in the asymptotic range, one can conclude that as R approaches unity, the influence of the partitioning error on the partitioned solution becomes more noticeable, causing the larger deviation of the partitioned solution from the monolithic one.

Referring to Fig.(2(b)), it is noted that the partitioning errors of the IMEX schemes are shown only up to a maximum time step of $\Delta t = 2$, since above it the partitioned schemes are unstable. This is another important difference in the application of the IMEX schemes to the two cases. Such instabilities are not present in the monolithic solution- by observing the time integration error, and their appearance in the IMEX schemes are due to IMEX schemes' partly explicit nature. For this test case, where the material properties and the mesh spacing of the two domains are the same, instability occurs when the Fourier number of the domains exceeds $Fo = Fo_1 = Fo_2 = 1$. Thus, when the loosely coupled higher order IMEX schemes are used for time integration of cases where $R \approx 1$, the time step size choice is severely reduced, compared to the time step that can be selected independently for each domain, due to stability reasons. In order to stabilize the algorithm, subiterations must be performed.

6.3 Efficiency of the higher order IMEX scheme

When analyzing the thermal coupling of typical fluids and solids, it is observed that $R \ll 1$. For example for steel-air $R \approx 0.07$, and for water-copper $R \approx .0013$. Thus, the loosely-coupled IMEX scheme can be used for time-stepping the coupled problem without encountering any stability issues. However, in comparison to first order BDF1 and second order time integration schemes such as the θ scheme or BDF2 where only one implicit solve at each time step is performed, the higher order IMEX schemes, with s number of stages, computes $s - 1$ implicit solves at each time step. Therefore, the issue that

Table 2: Set of parameters used in analyzing the model problem, N : number of cells

Subdomain	L_x	L_y	N_x	N_y	k	ρc_p
Ω_1	.5	.5	20	10	54	7400×475
Ω_2	.5	.5	20	20	.06	1000×1
Ω_3	.5	.5	20	10	54	7400×475

needs to be considered is whether the higher order IMEX schemes are computationally more efficient than the lower order schemes in acquiring a certain accuracy. In that respect, the computational work of third and fourth order IMEX schemes are compared to that of the θ scheme using the model problem depicted in Fig.(3(a)) where thermal coupling of 3 domains in 2-D is considered. The governing equation within each domain is transient conduction. The outer boundaries are insulated. In the previous sections, the application of the IMEX schemes to the combination of FEM-CFV grids was shown. In this section, the IMEX schemes is used for time-stepping the coupled subdomains where all the subdomains are discretized using cell-centered finite volumes. In order to use the D-N formulation, an approximation to the interface temperature is then required. Here, this is obtained based on equating one-sided differences to compute the interface heat flux (see [5] for derivation). The material properties, the size of each domain, and the number of cells used to discretize each domain, are shown in Table.(2). With reference to the material properties shown in Table.(2), for a stable time-stepping of the coupled problem, Ω_2 takes the Dirichlet condition at the two interfaces, while the other two subdomains take the flux condition. The following initial condition was imposed on the global domain:

$$T_i(\mathbf{x}) = 0 \text{ for } (\mathbf{x}) \in \Omega_i \text{ with } i = 1, 3 \text{ and } T_2(\mathbf{x}) = \begin{cases} 1 & \text{if } x \leq 0.25 \\ 0 & \text{if } x > 0 \end{cases}. \text{ The simulation were carried out to } t_{final} = 1000 \text{ sec.}$$

Figure 3(b) shows the total error of the partitioned solution using the IMEX scheme against the work, where work is defined as the total number of implicit stage calculations during the simulation ($W = (s - 1) * \frac{t_{final}}{\Delta t}$). The total error of the partitioned scheme is obtained by comparing the partitioned solution field to its corresponding temporally exact solution field and taking the L_2 norm of the difference. The temporally exact solution was obtained using the fourth order ESDIRK and $\Delta t = 0.01$ sec. Furthermore, the results are depicted for two possible sequences with which coupled problem can be solved. The work of the second order θ scheme is also plotted as reference.

As Fig.(3(b)) shows, even in the case where an approximation to the interface equation has to be introduced to solve the coupled problem, the IMEX schemes preserve their respective design order in the asymptotic range. This result is irrespective of the sequence with which the subdomains are integrated. However, comparing the total error convergence curves for the two different integration sequences it is noted that the inte-

gration sequence does influence the magnitude of the partition error and hence the total error. After performing some numerical tests, it was observed that the magnitude of the difference depends on parameters such as the initial condition and the outer boundaries imposed on the global domain, the mesh ratio of the coupled domains, and the duration of the computations, t_{final} , and therefore a general conclusion in selecting the sequence which will always give the lowest partition error cannot be made.

For this test case, since the results, for both the θ and IMEX schemes, are not influenced noticeably by the order with which the domains are integrated, the following is true for both the integration sequences considered here. Since the governing equation within all the subdomains is transient conduction, a rough estimate of the time step to accurately capture the transients based on a Fourier number of unity is $\Delta t \approx \frac{\Delta x^2}{\alpha} = \frac{(0.5/20)^2}{.06/(1000)} \approx 10sec$. For the θ scheme with $\Delta t = 10$, the computational work performed to arrive at $t_{final} = 1000sec$ is approximately 100 and the accuracy of the solution is $\approx 10^{-5.5}$. Even though the 3rd and 4th order IMEX schemes require respectively 3 and 5 implicit solves per time step compared to the partitioned θ scheme, to arrive at the same accuracy the amount of work required by them are roughly 1.6 and 2.5 times less than θ scheme (for this particular test case).

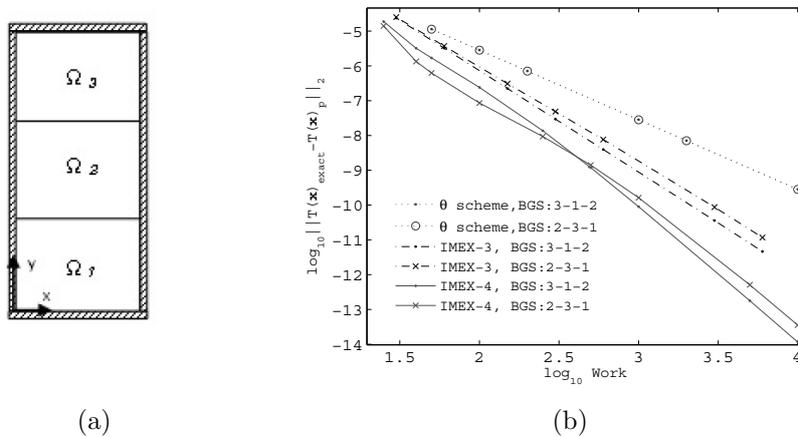


Figure 3: (a) Three domains thermally coupled, (b) Comparison of the computational work of the higher order IMEX and the 2nd order θ scheme.

7 CONCLUSIONS

- The higher order implicit ESDIRK schemes were used to solve thermally coupled domains with Dirichlet-Neumann as the decomposition method. For loose coupling of the domains, the IMEX scheme consisting of the ESDIRK scheme for integrating the governing equations within the subdomains and an ERK scheme for explicit integration of the explicit coupling terms was utilized. The stability and accuracy

of the IMEX scheme was evaluated for two cases: In one, the material properties of the coupled domains are quite different ($R \ll 1$) and in the other they are the same ($R = 1$). For both cases, in the asymptotic range, the IMEX schemes preserved the design order of the time integration scheme. However, it was observed that in comparison to $R \ll 1$, for $R = 1$ the influence of the partitioning error on the partitioned solution becomes more noticeable, causing the larger deviation of partition solution from the monolithic one. Furthermore, while for $R \ll 1$, the IMEX schemes did not encounter any stability issues for relatively small time steps (compared to the time scale of diffusion in the problem), for $R = 1$ instabilities do appear at such time steps. For the considered test case, instability occurred when the Fourier number of the domains exceeded unity, $Fo = Fo_1 = Fo_2 > 1$. To stabilize the algorithm, subiterations must be performed.

- The computational efficiency of the higher order IMEX schemes relative to the partitioned second order θ scheme was considered using a test case in 2-D involving coupled conduction problem of three domains. Even though greater number of implicit solved per time step is required for the IMEX schemes compared to the θ scheme, it was demonstrated that to arrive at the same accuracy, the IMEX schemes required less computational effort.

REFERENCES

- [1] van Zuijlen, A.H. and Bijl, H. Implicit and explicit higher order time integration schemes for structural dynamics and fluid-structure interaction computations. *Computers and Structures*. (2005) **83**:93-105.
- [2] Henshaw, W.D. and Chand, K.K. A composite grid solver for conjugate heat transfer in fluid-structure systems. *Journal of Computational Physics*. (2009) **228**:3708-3741.
- [3] Roe, B. Haselbacher, A. and Geubelle, P.H. Stability of fluid-structure thermal simulations on moving grids. *International Journal for Numerical Methods in Fluids*. (2007) **54**:1097-1117.
- [4] Kennedy, C.A. and Carpenter, M.H. Additive Runge-Kutta schemes for convection-diffusion-reaction equations. *Applied Numerical Mathematics*. (2003) **44**:139-181.
- [5] Carlson, K.D. Lin, W.L. and Chen, C.J. Pressure Boundary Conditions of Incompressible Flows with Conjugate Heat Transfer on Nonstaggered Grids Part II: Applications. *Numerical Heat Transfer, Part A*. (1997) **32**:481-501.

MIXING SNAPSHOTS AND FAST TIME INTEGRATION OF PDEs

MARÍA-LUISA RAPÚN[†], FILIPPO TERRAGNI[†] AND JOSÉ M. VEGA[†]

[†] Dpto. Fundamentos Matemáticos, E.T.S.I. Aeronáuticos
Universidad Politécnica de Madrid
Plaza Cardenal Cisneros 3, 28040 Madrid, Spain

Key words: Reduced Order Models, Proper Orthogonal Decomposition, Galerkin Projection, Partial Differential Equations

Abstract. *A local proper orthogonal decomposition (POD) plus Galerkin projection method was recently developed to accelerate time dependent numerical solvers of PDEs. This method is based on the combined use of a numerical code (NC) and a Galerkin system (GS) in a sequence of interspersed time intervals, I_{NC} and I_{GS} , respectively. POD is performed on some sets of snapshots calculated by the numerical solver in the I_{NC} intervals. The governing equations are Galerkin projected onto the most energetic POD modes and the resulting GS is time integrated in the next I_{GS} interval. The major computational effort is associated with the snapshots calculation in the first I_{NC} interval, where the POD manifold needs to be completely constructed (it is only updated in subsequent I_{NC} intervals, which can thus be quite small). As the POD manifold depends only weakly on the particular values of the parameters of the problem, a suitable library can be constructed adapting the snapshots calculated in other runs to drastically reduce the size of the first I_{NC} interval and thus the involved computational cost. The strategy is successfully tested in (i) the one-dimensional complex Ginzburg-Landau equation, including the case in which it exhibits transient chaos, and (ii) the two-dimensional unsteady lid-driven cavity problem.*

1 INTRODUCTION

Reduced order models (ROMs) have become an increasingly active research field along the last twenty years. This is due to their interest in both understanding basic mechanisms of fluid systems [7, 15] and improving prediction and design in industrial processes [4, 11]. Concerning the latter, in mature sectors such as automotive and aeronautics, a trend is observed to promote the use of computational fluid dynamics (CFD) in aerodynamics design, intending to decrease the huge cost of wind tunnel tests. The main difficulty is that traditional CFD approaches (such as direct numerical simulation or turbulence models)

still require huge computational resources and CPU time, especially in multi-parameter problems. As a consequence, reducing computational effort of CFD solvers is becoming a crucial step to facilitate their industrial use.

Proper orthogonal decomposition (POD) combined with projection of the governing equations onto a POD manifold has been seen to produce ROMs of both steady [1, 10] and evolution problems [3, 12] which allow to drastically decrease computational cost. In evolution problems, these models consist in (i) identifying a low dimensional POD manifold that contains a good approximation of the relevant dynamics and (ii) using as reduced order model the Galerkin system (GS) obtained by projecting the governing equations onto this manifold. A basis of the POD manifold is constructed by the method of snapshots [17], which consists in applying POD methodology to a set of numerically calculated snapshots spanning a portion of the phase space of the dynamical system that contains all relevant orbits. Time integration of the GS turns out to involve a much smaller computational effort than time integration of the original system.

A major drawback is that the resulting GS may exhibit spurious dynamics in a somewhat unpredictable way. The reason for that is still controversial, but seems to be due to the non-invariance of the POD manifold under the true dynamics [14]. Thus, intended solutions to this difficulty somehow correct either the GS or the POD manifold in order to make the latter invariant [6, 16]. In all these cases, the GS is intended to approach the system dynamics in a particular attractor, which can be periodic, quasi-periodic, or chaotic, and it is not suitable to reproduce transient behaviors.

A somewhat different approach, called *local POD plus Galerkin projection* (LPOD+GP) method, was presented in [13] for one-dimensional parabolic equations and extended in [18] for two-dimensional fluid dynamics problems. In both works, the LPOD+GP method turned out to be both quite computationally efficient and robust, providing a good approximation of the considered dynamics (either transients or a given attractor), eliminating spurious behaviors. The main idea (see the sketch in Fig. 1) is to combine a numerical code (NC) and a GS in interspersed time intervals, I_{NC} and I_{GS} , respectively. Snapshots are computed by the NC in each I_{NC} interval and are used to either calculate (in the first I_{NC} interval) or update (in subsequent I_{NC} intervals) the relevant POD modes. The governing equations are then Galerkin projected onto the most energetic POD modes to obtain a GS that is integrated in the next I_{GS} interval. Of course, the key point of the method is to decide when each I_{GS} interval must be terminated because the GS approximation is no longer acceptable. This is accomplished by means of an a priori error estimate based on the amplitude of some additional higher order modes. In addition, a second GS that retains a few more modes than necessary is integrated and (in conjunction with the above mentioned a priori error estimate) provides a safe criterion for switching between I_{GS} and I_{NC} intervals, even in cases in which the involved dynamics are really complex.

Since the first I_{NC} interval is much longer than the remaining I_{NC} intervals and the GS is much computationally inexpensive than the NC, the crucial step to further improve the

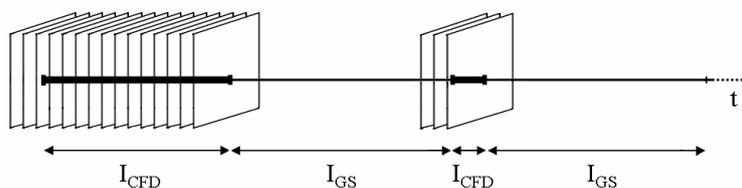


Figure 1: The local POD plus Galerkin projection method.

performance of the LPOD+GP method is to reduce the computational effort associated with the snapshots calculation in the first I_{NC} interval. A key observation is that the POD manifold depends only weakly on the particular values of the parameters of the problem. In this work we will show how a suitable *snapshots library* can be constructed from the POD manifolds calculated in other runs (for other parameter values). The use of this library allows for drastically reducing the size of the first I_{NC} interval and thus the involved computational cost. In other words, the library can be constructed using the POD modes obtained from a set of snapshots calculated by the NC for other parameter values (in, e.g., former applications of the LPOD+GP method). These POD modes may (possibly) not contain a good approximation of the true dynamics for the actual set of parameters, but they can be updated by adding a few snapshots for the true parameter values. The computation of these new snapshots only requires to use the NC in a small I_{NC} interval.

In order to explain and apply the introduced ideas, we will consider two test problems.

A. The *complex Ginzburg-Landau (CGL) equation*

$$\partial_t u = (1 + i\alpha)\partial_{xx}u + \mu u - (1 + i\beta)|u|^2 u, \quad \text{with } u = 0 \text{ at } x = 0, 1, \quad (1)$$

with initial condition $u(x, 0) = i \sin(2\pi x) + (1 + i) \sin(3\pi x)$, which is a fairly simple equation that exhibits intrinsically complex dynamics [2]. The state variable u is complex and the parameters μ , α , and β are real. This equation exhibits the modulational instability if $\alpha\beta < 1$ and μ is large. Increasing μ beyond the modulational instability usually yields chaotic dynamics.

B. The *unsteady lid-driven cavity (ULDC) problem* [5, 8, 9], which describes the motion of a liquid in an enclosed cavity whose upper wall is moving back and forth. The governing equations are

$$\nabla \cdot \mathbf{v} = 0, \quad \partial_t \mathbf{v} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\nabla p + \text{Re}^{-1} \Delta \mathbf{v}, \quad (2)$$

in the spatial domain $0 < x < 1$, $0 < y < 1$, with boundary conditions

$$\mathbf{v} = \mathbf{0} \text{ at } x = 0, 1 \text{ and } y = 0, \quad \mathbf{v} = (16 h(t)x^2(1-x)^2, 0) \text{ at } y = 1. \quad (3)$$

Here, $\mathbf{v} = (v_x, v_y)$ and p are the dimensionless velocity and pressure, and the Reynolds number is defined as $\text{Re} = u^*L/\nu$, where u^* is the maximum lid forcing (horizontal) velocity, L is the width of the cavity, and ν is the kinematic viscosity. The function h accounts for temporal oscillations (either periodic or quasi-periodic). Such time dependence of the driving velocity permits nontrivial dynamics at large time for moderate Reynolds number.

The remainder of the paper is organized as follows. The basic LPOD+GP method is briefly recalled and its improvement using POD modes libraries is developed in section 2, where notation is established. The results of the paper are presented and discussed in section 3. The paper ends with some concluding remarks, in section 4.

2 THE LOCAL POD PLUS GALERKIN PROJECTION METHOD

Let us consider a real or complex system of semilinear parabolic equations of the type

$$\mathcal{M}\partial_t\mathbf{q} = \mathcal{L}\mathbf{q} + \mathbf{f}(\mathbf{q}, t), \quad (4)$$

where \mathbf{q} is a state vector, \mathcal{M} and \mathcal{L} are linear operators, with the highest order derivatives accounted for in \mathcal{L} , and \mathbf{f} is a nonlinear operator. In the I_{GS} intervals, \mathbf{q} is approximated by a linear combination of POD modes \mathbf{Q}_i as

$$\mathbf{q} \simeq \mathbf{q}_{GS}^n = \sum_{i=1}^n A_i(t)\mathbf{Q}_i, \quad (5)$$

for certain amplitudes A_i ¹. Replacing (5) into (4) and projecting the resulting equations onto the POD modes, yields the following GS

$$\sum_{j=1}^n \mathcal{M}_{ij}^{GS} \frac{dA_j}{dt} = \sum_{j=1}^n \mathcal{L}_{ij}^{GS} A_j + f_i^{GS}(A_1, \dots, A_n, t), \quad (6)$$

where the matrices \mathcal{M}^{GS} and \mathcal{L}^{GS} , and the nonlinear functions f_i^{GS} are defined as

$$\mathcal{M}_{ij}^{GS} = \langle \mathbf{Q}_i, \mathcal{M}\mathbf{Q}_j \rangle, \quad \mathcal{L}_{ij}^{GS} = \langle \mathbf{Q}_i, \mathcal{L}\mathbf{Q}_j \rangle, \quad f_i^{GS} = \langle \mathbf{Q}_i, \mathbf{f}\left(\sum_{k=1}^n A_k\mathbf{Q}_k, t\right) \rangle. \quad (7)$$

POD and Galerkin projection are performed in terms of a suitable inner product $\langle \cdot, \cdot \rangle$. Computational efficiency of the LPOD+GP method is enhanced by using an inner product based on a limited number of discretization mesh points [13, 18]. Note that POD modes are orthonormal with respect to the considered inner product.

¹When the boundary conditions are nonhomogeneous \mathbf{q} is usually replaced by $\mathbf{q} - \mathbf{q}_0$ in the expansion (5), where \mathbf{q}_0 satisfies the nonhomogeneous boundary conditions.

In the first I_{NC} interval, POD modes are calculated from a set of N snapshots, namely N instantaneous distributions of \mathbf{q} . After truncation to $n \leq N$ modes, the relative root mean square error when reconstructing the N snapshots is given by

$$\text{RRMSE}_n^N = \sqrt{\frac{\sum_{j=n+1}^N (\sigma_j)^2}{\sum_{j=1}^N (\sigma_j)^2}}, \quad (8)$$

where σ_j is the singular value associated with the mode \mathbf{Q}_j . This formula is used to select the number of retained modes. In subsequent I_{NC} intervals, the POD manifold is only updated by applying POD to the following set of vectors

$$\hat{\nu}_1 \hat{\mathbf{Q}}_1, \dots, \hat{\nu}_n \hat{\mathbf{Q}}_n, \nu_1 \mathbf{Q}_1, \dots, \nu_N \mathbf{Q}_N. \quad (9)$$

Here, $\hat{\mathbf{Q}}_1, \dots, \hat{\mathbf{Q}}_n$ are the POD modes used in the last I_{GS} interval, while the weights $\hat{\nu}_1, \dots, \hat{\nu}_n, \nu_1, \dots, \nu_N$ are defined as

$$\hat{\nu}_j = \min \left\{ \frac{\hat{\sigma}_j}{\sqrt{\sum_{k=1}^n (\hat{\sigma}_k)^2}}, \frac{\langle |A_j| \rangle}{\sqrt{\sum_{k=1}^n \langle |A_k| \rangle^2}} \right\}, \quad \nu_j = \frac{\sigma_j}{\sqrt{\sum_{k=1}^N (\sigma_k)^2}}, \quad (10)$$

where, for each j , $\hat{\sigma}_j$ is the singular value associated with $\hat{\mathbf{Q}}_j$ calculated in the last I_{NC} interval and $\langle |A_j| \rangle$ is the temporal mean value of $|A_j|$ in the last I_{GS} interval; $\mathbf{Q}_1, \dots, \mathbf{Q}_N$ are the POD modes calculated from the new snapshots in the new I_{NC} interval, and $\sigma_1, \dots, \sigma_N$ are the corresponding singular values. By defining the weights of old and new POD modes as in (10), we appropriately update the POD manifold avoiding its contamination and making it dependent on the local dynamics [13, 18].

The instantaneous, spatial, relative error associated with the Galerkin approximation (5) in each I_{GS} interval is measured by

$$E^n = \|\mathbf{q} - \mathbf{q}_{GS}^n\| / \|\mathbf{q}\|. \quad (11)$$

If E^{n_1} is sufficiently small for some $n_1 > n$, then the quantity

$$E_n^{n_1} = \sqrt{\frac{\sum_{j=n+1}^{n_1} (A_j)^2}{\sum_{j=1}^{n_1} (A_j)^2}} \quad (12)$$

is a good estimate of E^n [13, 18]. This error estimate is fairly standard and plays an essential role in the LPOD+GP method.

Now, the LPOD+GP method intends to approximate the solution of (4) within an error bound ε in each I_{GS} interval. This goal is achieved by retaining at the beginning of each I_{GS} interval a few more modes than necessary, which also provides an error estimate to monitor the error [13, 18]. The method also involves a second GS which allows to deal with highly unstable dynamics [13, 18].

2.1 The basic method

As developed in [13, 18], the basic LPOD+GP method proceeds as follows. A previous selection is made of the various parameters of the method appearing below, namely the RRMS error bound, ε , the constant K , the time interval between snapshots, δ_{snaps} , the number of snapshots in the first I_{NC} interval, and the minimum length of the I_{GS} intervals, $\delta_{GS,\min}$. The LPOD+GP method can then be summarized in four steps, as follows.

- i. In the first I_{NC} interval, POD modes are calculated taking as snapshots the selected δ_{snaps} -equispaced portraits of \mathbf{q} . In the remaining I_{NC} intervals, POD modes are calculated from the modified snapshots defined in (9).
- ii. Three numbers of modes, n , n_1 , and n_2 , are defined as the smallest integers satisfying $\text{RRMSE}_n^N < \varepsilon_1 = \varepsilon/K$, $\text{RRMSE}_{n_1}^N < \varepsilon_1/K$, $\text{RRMSE}_{n_2}^N < \varepsilon_1/K^2$, where the RRMSE is defined in terms of the singular values as in equation (8) and the parameter K needs to be calibrated (see [13, 18] and §2.2).
- iii. Two GSs are constructed, retaining n_1 and n_2 modes, to calculate $\mathbf{q}_{GS}^{n_1} = \sum_{j=1}^{n_1} A_j \mathbf{Q}_j$ and $\tilde{\mathbf{q}}_{GS}^{n_2} = \sum_{j=1}^{n_2} \tilde{A}_j \mathbf{Q}_j$, respectively, taking as initial condition at $t = t_0 + \delta_{NC}$ the projections onto the POD manifolds of the NC solution calculated in step (i). Both GSs are integrated monitoring the error estimate $E_n^{n_1}$ defined in (12) and the following estimate of E^{n_1} , $\hat{E}_{n_1}^{n_2} = \|\|\mathbf{q}_{GS}^{n_1} - \tilde{\mathbf{q}}_{GS}^{n_2}\|\| / \|\|\tilde{\mathbf{q}}_{GS}^{n_2}\|\| - E_n^{n_1}$. Integration proceeds until the last value of t , t_1 , such that

$$E_n^{n_1} \leq \varepsilon, \quad \hat{E}_{n_1}^{n_2} \leq \varepsilon_1. \quad (13)$$

The second restriction in (13) is used to impose consistency between the two GSs in connection with higher order modes. Now, there are two alternatives.

1. If the resulting value of $\delta_{GS} < \delta_{GS,\min}$, then a new value of δ_{NC} is defined as $\delta_{NC,\text{new}} = \min\{\delta_{NC,\text{estimated}}, 2\delta_{NC,\text{old}}\}$, where

$$\delta_{NC,\text{estimated}} = \delta_{NC,\text{old}} + \max\left\{\delta_{\text{snaps}}, \frac{\delta_{GS,\min} - \delta_{GS}}{\delta_{GS,\min}} \delta_{NC,\text{old}}\right\}.$$

The NC solution is completed in the new part added to the I_{NC} interval and step (ii) is repeated.

2. Otherwise, the method proceeds to next step.

- iv. If $t_1 < T$ (final value of t), then the value of \mathbf{q} at t_2 reconstructed from the last Galerkin state with n_2 modes is taken as initial condition, and step (i) is repeated. Otherwise, the procedure ends.

The effectiveness of the LPOD+GP method can be measured in terms of the ratio of the total time span to the total length of the I_{NC} intervals, namely

$$\text{Theoretical Compression} = T / \sum \delta_{NC}. \quad (14)$$

The actual CPU compression factor may also be defined [18], which obviously depends on the CPU unit and the software used to construct the ROM.

As thoroughly shown in [13, 18], the method produces fairly good results (namely, the method provides the solution within the required precision, with large theoretical compression factors), which is due to the fact that the a priori error estimate defined above works quite well.

2.2 Improved method using weighted POD modes libraries

The POD modes library can be constructed in various ways, including:

- The POD manifold resulting from applying POD to a set of generic functions, such as Fourier and orthogonal polynomials, depending on the boundary conditions.
- The POD manifold resulting from other runs of the method for other parameter values. In this case, the weights of the POD modes are those appearing in the last I_{NC} interval, defined in eq.(10).
- Different libraries can be mixed up by just applying POD to the joint sets of modes, after appropriately weighting them as explained in the previous item.

Once the POD modes library is defined, it is used in the first I_{NC} interval as done in the basic method with the old POD manifold in subsequent I_{NC} intervals. In other words, the POD manifold in the first I_{NC} interval is calculated by applying POD to the set (9), where $\hat{\nu}_j \hat{\mathbf{Q}}_j$ are the weighted modes from the library, while $\nu_j \mathbf{Q}_j$ are as in eq.(9).

The idea is that the required number of snapshots, N , will be small provided that the POD modes library includes some of the directions in the required POD manifold. In fact, the selection of the library is not critical.

3 RESULTS

For illustration, the basic LPOD+GP method and its improvement described above are now applied to the two test problems introduced in §1.

3.1 The complex Ginzburg-Landau equation

In the CGL equation, the numerical solver results from discretizing the spatial derivatives with centered, second order finite differences, and integrating the resulting system of ODEs using Matlab ode15s, which is also used to time integrate the Galerkin system.

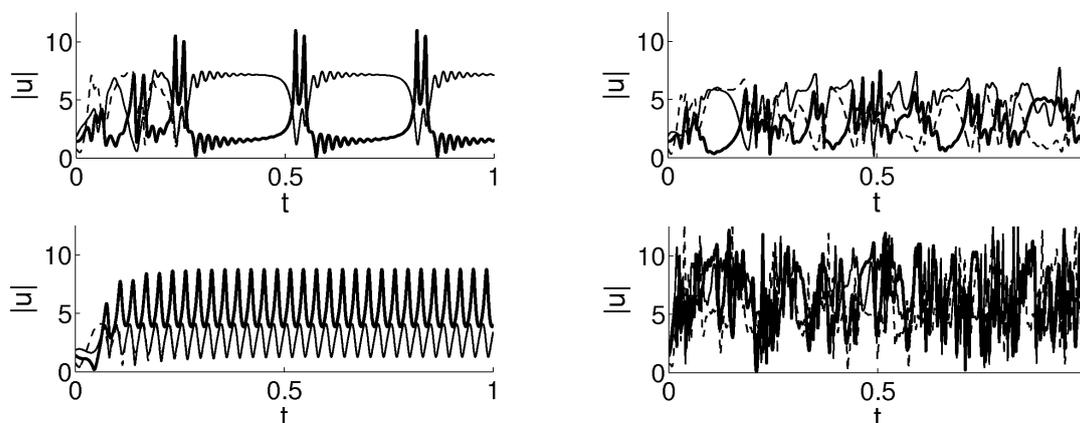


Figure 2: CGL equation: evolution at the points $x = 1/2$ (—), $x = 1/4$ (—), and $x = 3/4$ (---) in the test cases TC1 (top, left), TC2 (top, right), TC3 (bottom, left), and TC4 (bottom, right).

Four test cases are considered, which are defined for the following parameter values

$$(\mu, \alpha, \beta) = (95, -1.5, 12) \text{ (TC1)}, \quad (\mu, \alpha, \beta) = (85, -2, 19) \text{ (TC2)}, \quad (15)$$

$$(\mu, \alpha, \beta) = (65, -1.5, 10) \text{ (TC3)}, \quad (\mu, \alpha, \beta) = (180, -2, 15) \text{ (TC4)}. \quad (16)$$

These yield representative dynamics of the equation (see Fig.2) and will be used to construct POD modes libraries. The first two of them will also be used to check the performance of the method. Note that after a transient the system shows reflection symmetric relaxation oscillations in TC1 and even simpler oscillations in TC3, while TC2 and TC4 show representative, non-reflection symmetric chaotic dynamics.

The performance of the basic and improved versions of the method is illustrated in Fig.3, where the lengths of the interspersed I_{NC} and I_{GS} intervals is indicated in terms of the time steps required in each of them, both using the basic version of the method (labelled as LPOD+GP) and the improved version of the method, with various libraries, as indicated. The label L_F stands for the library resulting from applying POD to the set of Fourier modes $\sin(\ell\pi x)$, for $\ell = 1, \dots, 50$, and the labels L_{TCk} , for $k = 1, \dots, 4$, denote the last POD manifold resulting from the application of the basic LPOD+GP method to the test case TCk in the interval $0 \leq t \leq 1$; $L_F + L_{TC3}$ and $L_{TC1} + L_{TC3}$ denote the libraries resulting from mixing (as explained above) the two indicated libraries. The parameters of the method are $\varepsilon = 0.005$, $K = 100$, $\delta_{\text{snaps}} = 0.0005$, and $\delta_{GS, \text{min}} = 0.06$; the numbers of required POD modes are of the order of $(n, n_1, n_2) = (25, 30, 40)$ in both test cases TC1 and TC2. As a general comment, both the basic and the improved method work as well as the basic method did in [13].

Some remarks are now in order. Concerning the test case TC1 (left plots in Fig.3):

- The basic LPOD+GP method requires only two I_{NC} intervals, whose lengths are 403 and 1 time steps. The theoretical compression factor is 4.95.

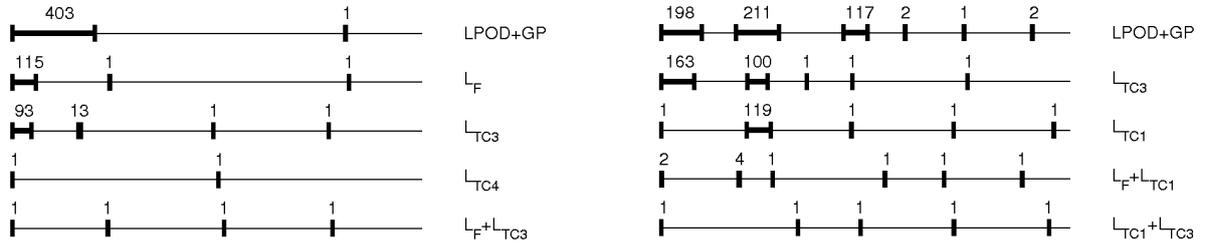


Figure 3: CGL equation: the I_{NC} (—) and I_{GS} (—) intervals in $0 \leq t \leq 1$ for the test cases TC1 (left) and TC2 (right). The number shown above each I_{NC} interval is the number of required time steps.

- Using both the generic library L_F and the customized library L_{TC3} produce similar results, namely they divide by 3.5 the length of the first I_{NC} interval (the new compression factor is 18.39). This is because the dynamics of the test case TC3 are much simpler than those of the test case that is simulated (see Fig.2).
- Using the library L_{TC4} produces optimal results since it reduces the length of the first I_{NC} interval to its minimum possible value, namely just one time step (the resulting compression factor is 1000). This is because the new library results from the test case TC4, which exhibits more complex dynamics than the dynamics that are being simulated. It is remarkable that the parameter values of the test cases TC1 and TC4 are quite different from each other (see eqs.(15)-(16)).
- If the (somewhat simple) libraries L_F and L_{TC3} are mixed up, the performance of the method is as optimal as when using the (more complex) library L_{TC4} . This seems to be due to the fact that each one of the libraries L_F and L_{TC3} spans a limited part of the phase space of the equation, but when these two libraries are mixed up, the resulting library spans a larger part of the phase space. The latter is in fact large enough as to allow for completing the POD manifold calculation with a slight updating in the first I_{NC} interval.

The last two remarks illustrate well the robustness of the method. The performance of the basic and improved LPOD+GP methods for the test case TC2 (right plots in Fig.3) exhibits similar trends. The main difference is that the dynamics of TC2 are more complex and the basic LPOD+GP method requires larger I_{NC} intervals. The improved method instead highly decreases the I_{NC} intervals, especially when the L_{TC1} and L_{TC3} libraries (both resulting from much simpler dynamics than the dynamics of TC2) are mixed up, which increases the compression factor from 3.76 to 400.

Finally, to further illustrate the robustness of the method, a random generation of the parameters of the CGL equation (1) in the intervals $\mu \in [50, 100]$, $\alpha \in [-2.5, -1.5]$, $\beta \in [10, 20]$ produced the values $(\mu, \alpha, \beta) = (81.87, -1.77, 15.45)$. In this case the method behaves as before: without libraries the compression factor is 3.01, while using the libraries

L_{TC3} , L_{TC1} , $L_F + L_{TC1}$, and $L_{TC1} + L_{TC3}$ the compression factors are 8.24, 15.06, 333, and 400, respectively.

3.2 The unsteady lid-driven cavity problem for $Re=100$

The numerical solver for the ULDC problem (2)–(3) is a rough industrial-like code, which includes some artificial tricks to accelerate its performance. The Galerkin projection instead is based on the exact equations, discretized using a classical Crank-Nicolson scheme; see [18] for details. Five test cases are now considered to build up the POD modes libraries. These test cases are obtained using the following lid driving function h (appearing in the boundary condition (3)):

$$\begin{aligned} h &= \sin(2\pi t/7) \cos(t/14) \quad (\text{TC1}), & h &= \sin(\pi t/10) \cos(5t/4) \quad (\text{TC2}), \\ h &= \sin(2t/\pi) + 0.5 \cos t \quad (\text{TC3}), & h &= \sin t \quad (\text{TC4}), & h &= 1 \quad (\text{TC5}). \end{aligned}$$

Note that the first three of them are quasi-periodic, the fourth one is periodic, and the last one is steady. The initial condition is always the quiescent state, meaning that all of them exhibit unsteady behaviors, even under steady forcing.

The counterpart of Fig.3 is Fig.4. Note that the compression factors are now smaller, which is due to the nature of the numerical solver that is being used. Otherwise, the performance of the basic and improved LPOD+GP method exhibits the same trends as in the CGL equation. In particular:

- The libraries always shorten the length of the first I_{NC} interval, and this effect is stronger when the library results from a test case that exhibits more complex dynamics.
- Mixing libraries always produces a larger benefit than when each library is used alone. This is true even in cases in which the dynamics implicit in the libraries are simpler than the dynamics that are being simulated.
- The reduction of the first I_{NC} interval (which was already somewhat small as resulting from the basic LPOD+GP method) is only moderate, but the libraries also succeed in reducing the length of the subsequent I_{NC} intervals, thus improving the overall performance of the method.

4 CONCLUSIONS

A method based on POD modes libraries has been developed that highly improves the performance of the basic LPOD+GP method. The improvement has been illustrated and checked in two paradigmatic examples, namely the complex Ginzburg-Landau equation and the unsteady lid-driven cavity problem. The following remarks are in order:

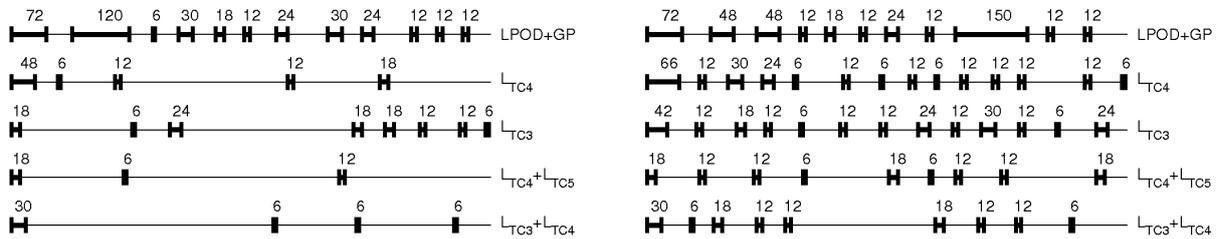


Figure 4: Counterpart of Fig.3 for the ULDC problem.

- The selection of the POD modes libraries is not critical. They can be constructed either out of a set of generic functions or using POD manifolds resulting from previous runs of the method for different parameter values. This is quite interesting envisaging applications in industrial environments, where solvers are usually run for a large amount of sets of parameter values.
- The libraries resulting from more complex dynamics than those that are being simulated usually work better than the libraries resulting from simpler dynamics.
- The combination of POD modes libraries usually produces much better results than when each library is used alone.

All these suggest that many parabolic equations and systems might exhibit a POD manifold that approximately contains (or almost contains) not only the attractors but also a significant part of the most relevant transient behaviors.

REFERENCES

- [1] D. ALONSO, A. VELAZQUEZ, AND J. M. VEGA, *A method to generate computationally efficient reduced order models*, *Comput. Meth. Appl. Mech. Eng.*, 198 (2009), pp. 2683–2691.
- [2] I. S. ARANSON AND L. KRAMER, *The world of the complex Ginzburg-Landau equation*, *Rev. Mod. Phys.*, 74 (2002), pp. 100–142.
- [3] G. BERKOOZ, P. HOLMES, AND J. L. LUMLEY, *The proper orthogonal decomposition in the analysis of turbulent flows*, *Annu. Rev. Fluid Mech.*, 25 (1993), pp. 539–575.
- [4] K. BIZON, G. CONTINILLO, L. RUSSO, AND J. SMULA, *On POD reduced models of tubular reactor with periodic regimes*, *Comput. Chem. Eng.*, 32 (2008), pp. 1305–1315.

- [5] W. CAZEMIER, R. W. C. P. VERSTAPPEN, AND A. E. P. VELDMAN, *Proper orthogonal decomposition and low-dimensional models for driven cavity flows*, Phys. Fluids, 10 (1998), pp. 1685–1699.
- [6] M. COUPLET, C. BASDEVANT, AND P. SAGAUT, *Calibrated reduced-order POD-Galerkin system for fluid flow modelling*, J. Comput. Phys., 207 (2005), pp. 192–220.
- [7] E. H. DOWELL AND K. C. HALL, *Modeling of fluid-structure interaction*, Annu. Rev. Fluid Mech., 33 (2001), pp. 445–490.
- [8] P. W. DUCK, *Oscillatory flow inside a square cavity*, J. Fluid Mech., 122 (1982), pp. 215–234.
- [9] U. GHIA, K. N. GHIA, AND C. T. SHIN, *High-Re solutions for incompressible flow using the Navier-Stokes equations and a multigrid method*, J. Comput. Phys., 48 (1982), pp. 387–411.
- [10] P. LEGRESLEY AND J. ALONSO, *Investigation of non-linear projection for POD based reduced order models for Aerodynamics*, AIAA paper 2001-0926, Jan. 2001.
- [11] T. LIEU, C. FARHAT, AND M. LESOINNE, *Reduced-order fluid/structure modeling of a complete aircraft configuration*, Comput. Meth. Appl. Mech. Eng., 195 (2006), pp. 5730–5742.
- [12] D. J. LUCIA, P. S. BERAN, AND W. A. SILVA, *Reduced-order modelling: new approaches for computations physics*, Prog. Aeosp. Sci., 40 (2004), pp. 51–117.
- [13] M. L. RAPÚN AND J. M. VEGA, *Reduced order models based on local POD plus Galerkin projection*, J. Comput. Phys., 229 (2010), pp. 3046–3063.
- [14] D. REMPFER, *On low-dimensional Galerkin models for fluid flow*, Theor. Comput. Fluid Dyn., 14 (2000), pp. 75–88.
- [15] D. REMPFER, *Low-dimensional modeling and numerical simulation of transition in simple shear flows*, Annu. Rev. Fluid Mech., 35 (2003), pp. 229–265.
- [16] S. SIRISUP, G. E. KARNIADAKIS, D. XIU, AND I. G. KEVREKIDIS, *Equations-free/Galerkin-free POD assisted computation of incompressible flows*, J. Comput. Phys., 207 (2005), pp. 568–587.
- [17] L. SIROVICH, *Turbulence and the dynamics of coherent structures*, Q. Appl. Math., XLV (1987), pp. 561–590.
- [18] F. TERRAGNI, E. VALERO, AND J. M. VEGA, *Local POD plus Galerkin projection in the unsteady lid-driven cavity problem*, preprint, 2010.

MULTI-LEVEL AND QUASI-NEWTON ACCELERATION FOR STRONGLY COUPLED PARTITIONED FLUID-STRUCTURE INTERACTION

J.J. KREEFT*, M. WEGHS, A.H. VAN ZUIJLEN AND H. BIJL

*Faculty of Aerospace Engineering, Delft University of Technology
Kluyverweg 1, 2629 HS, Delft, The Netherlands
e-mail: j.j.kreeft@tudelft.nl

Key words: Fluid-structure interaction, reduced order modeling, quasi-Newton coupling, multi-level acceleration, adaptivity.

Abstract. Two reduced order models are presented for the simulation of physically strong coupled fluid-structure interaction problem, based on computationally partitioned flow and structure solvers. The reduced order models used are a class of quasi-Newton coupling methods to obtain a stable solution and to reduce the number of subiterations. The second reduced order model is a multi-level acceleration in with coarse grid computations are performed in order to reduce computational costs. Finally an adaptive multi-level strategy is described, that contains an indicator for when to switch from coarse to fine grid level and vice versa.

1 INTRODUCTION

Fluid-structure interaction (FSI) is the interaction between a moving or deformable structure with a fluid flow. Due to the advances in computer power and numerical algorithms we are able to simulate increasingly complex problems. Despite these advances, issues in efficiency and stability of certain FSI problems remain. This is especially apparent in the simulation of physically strong interaction between the flow and structure, using a computationally partitioned coupling method.

The main advantages of the partitioned approach above the monolithic approach is that it allows reusing of existing flow and structure solvers, that are often already developed separately in the past. The main drawback of the partitioned approach over the monolithic approach is that the interface conditions,

$$\vec{v}_f = \frac{\partial \vec{d}_s}{\partial t}, \quad \sigma_f \cdot \vec{n} = \sigma_s \cdot \vec{n}, \quad \text{on } \Gamma_I, \quad (1)$$

are not satisfied automatically. A mismatch in interface conditions in the partitioned approach will eventually lead to stability problems in case of strong physical interaction.

Subiterations are therefore introduced in order to satisfy conditions (1). Because in the partitioned approach the solvers are considered to be blackboxes, a class of reduced order models will be used, called quasi-Newton subiterations, to stabilize the subiteration routine. Examples of these reduced order models that will be summarized here are, in order of increasing complexity, stability and efficiency: the classical Gauss-Seidel method, Aitken's adaptive underrelaxation method [6] and the least-squares method [2]. Other methods that belong to the class mentioned before, that are not discussed in this work, are Interface-GMRES(R) [1] or Newton-Krylov method [7]. It is shown that these are usually outperformed by the least-squares method [2, 5].

Obtaining stability by means of subiterations comes at the cost of a substantial increase in computation times, which depends heavily on the type of quasi-Newton method chosen. A second reduced order model will be introduced to suppress computational costs, i.e. a multi-level acceleration technique, [10], for reducing the costs of the solvers individually. This is especially of interest for the flow solver, which usually takes up most of the computation time. The multi-level acceleration technique discussed here is a coarse grid prediction method. It relies on a solve at the coarse grid level to predict the state at the fine grid level. Coarse grid levels are sufficient for stability, because it are the lower modes that might cause instabilities, and therefore need to be suppressed by the subiterations, [3]. Several strategies are possible with this multi-level acceleration method. One is to first subiterate on the coarse grid until certain convergence criteria is met, and then subiterate on the fine grid level. Alternatively, we can alternate coarse grid and fine grid level solves using a given pattern, e.g. two coarse grid solves followed by one fine grid solve. Finally we present an adaptive multi-level method, which automatically switches between coarse and fine grid level, based on a multi-level ratio criteria. It will be shown that this method significantly reduces the number of fine grid solves, at the cost of a small increase in total number of subiterations.

2 QUASI-NEWTON METHODS

In case of a Dirichlet-Neumann decomposition of the FSI problem, the stress distribution \mathbf{p} on the fluid-structure interface is passed from the flow solver \mathcal{F} to the structure solver \mathcal{S} and the displacement of the interface \mathbf{d} is transferred the other way around. We consider the flow and structure solver as two black boxes that must be coupled in some way. The flow solver gets as an input the structural displacement and has as an output the fluid stresses acting on the interface. For the structure solver we get vice versa. In equations this becomes

$$\mathbf{d} = \mathcal{S}(\mathbf{p}), \tag{2}$$

$$\mathbf{p} = \mathcal{F}(\mathbf{d}). \tag{3}$$

2.1 Quasi-Newton formulation

In many applications the interaction between fluid and structure is weak, e.g. in aeroelastic problems. These problems can be solved with so-called loosely coupled methods, [4, 8, 11]. Loosely coupled methods consist of a single solve per time-step. These algorithms do not enforce equilibrium on the fluid-structure interface and are therefore often unstable in cases with strong interaction (where density and stiffness ratios are of the same order). The errors made can diverge for more strong interactions. One way to stabilize the coupling of the partitioned system is to make use of subiterations.

In case we solve (2) and (3) sequentially, it can be written in a fixed-point formulation,

$$\mathbf{p} = \mathcal{F} \circ \mathcal{S}(\mathbf{p}), \quad (4)$$

so what goes in must come out in order to have equilibrium at the interface. Alternatively this can be written as a residual problem,

$$\mathcal{R}(\mathbf{p}) \equiv \mathcal{F} \circ \mathcal{S}(\mathbf{p}) - \mathbf{p} = 0, \quad (5)$$

where \mathcal{R} is the residual operator of the coupled problem. A popular method to solve a nonlinear system, $\mathcal{R}(\mathbf{p}) = 0$, in dynamical problems is to use Newton's method, so solve the system (with $\mathbf{r}^k = \mathcal{R}(\mathbf{p}^k)$):

$$\left. \frac{d\mathcal{R}}{d\mathbf{p}} \right|_{\mathbf{p}^k} \Delta\mathbf{p}^k = -\mathbf{r}^k, \quad (6a)$$

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \Delta\mathbf{p}^k. \quad (6b)$$

However, when using blackbox solvers the exact Jacobian of \mathcal{R} is in general unknown as the derivatives of \mathcal{F} and \mathcal{S} w.r.t. \mathbf{d} and \mathbf{p} are often unavailable, especially the cross derivatives $\partial\mathcal{F}/\partial\mathbf{d}$ and $\partial\mathcal{S}/\partial\mathbf{p}$. The Jacobian must be approximated instead.

First note however that instead of solving (6), using an approximate Jacobian, it is also possible, and much cheaper, to use the following explicit expression,

$$\mathbf{p}^{k+1} = \mathbf{p}^k + \left(\left. \frac{d\mathcal{R}}{d\mathbf{p}} \right|_{\mathbf{p}^k} \right)^{-1} (-\mathbf{r}^k), \quad (7)$$

where we need an approximation for the inverse of the Jacobian instead. Equation (7) is much cheaper than (6), because (6) requires to solve a dense linear system, where as (7) only uses a matrix vector product. The residual in (7) is calculated as

$$\mathbf{r}^k = \mathcal{R}(\mathbf{p}^k) = \mathcal{F} \circ \mathcal{S}(\mathbf{p}^k) - \mathbf{p}^k = \tilde{\mathbf{p}}^{k+1} - \mathbf{p}^k, \quad (8)$$

with convergence criteria $\|\mathbf{r}^k\|_2 \leq \epsilon$. This equation shows that $\tilde{\mathbf{p}}^{k+1} = \mathcal{F} \circ \mathcal{S}(\mathbf{p}^k)$.

2.2 Approximations for the inverse of the Jacobian

Resume by listing three suitable reduced order models for the approximation for the inverse of the Jacobian. In this section we will discuss the Gauss-Seidel method, Aitken's adaptive underrelaxation method, and the Least-Squares method.

The Gauss-Seidel method with subiterations, where the structure is solved first, is given by

$$\mathbf{d}^{k+1} = \mathcal{S}(\mathbf{p}^k), \quad (9a)$$

$$\mathbf{p}^{k+1} = \mathcal{F}(\mathbf{d}^{k+1}). \quad (9b)$$

Since we measure convergence of the pressure, it is useful to write it as

$$\mathbf{p}^{k+1} = \mathcal{F} \circ \mathcal{S}(\mathbf{p}^k), \quad (10)$$

and so the residual for the Gauss-Seidel method becomes

$$\mathbf{r}^k = \mathcal{R}(\mathbf{p}^k) = \mathcal{F} \circ \mathcal{S}(\mathbf{p}^k) - \mathbf{p}^k = \mathbf{p}^{k+1} - \mathbf{p}^k. \quad (11)$$

Combine (10) and (11) with (7) to find the expression for the approximation of the inverse of the Jacobian for the Gauss-Seidel method,

$$\overline{\left(\frac{d\mathcal{R}}{d\mathbf{p}} \Big|_{\mathbf{p}^k} \right)^{-1}}_{\text{GS}} = -I. \quad (12)$$

Unfortunately for strong coupling problems, sub-iterations for the Gauss-Seidel method converge slow or even do not converge at all. A first step to overcome these problems is to introduce underrelaxation.

The Gauss-Seidel method often appears to be either very slow (needs many sub-iterations) or even unstable. A simple and already very effective way is to introduce some underrelaxation. First perform a structure and flow solve as was done in the previous section (10). Now the output will not directly be used as an input for the next subiteration, but an underrelaxation will be performed, which is a linear combination of the output and the state of the previous subiteration,

$$\mathbf{p}^{k+1} = \omega \tilde{\mathbf{p}}^{k+1} + (1 - \omega) \mathbf{p}^k = \mathbf{p}^k + \omega \mathbf{r}^k. \quad (13)$$

The underrelaxation has a stabilizing effect on the convergence process. Aitken's underrelaxation method goes one step further, it uses an adaptive underrelaxation parameter ω^k , that is based on the one-dimensional Secant method. This Secant method uses a finite-difference approximation of the Jacobian. Suppose we want to solve the scalar equation

$\mathcal{R}(p) = 0$ using Secant's method, then

$$\begin{aligned} p^{k+1} &= p^k + \frac{p^k - p^{k-1}}{\mathcal{R}(p^k) - \mathcal{R}(p^{k-1})} (-\mathcal{R}(p^k)) \\ &= p^k + \frac{\omega^{k-1} r^{k-1}}{r^k - r^{k-1}} (-r^k) \\ &= p^k + \omega^k r^k. \end{aligned}$$

For a system with vectors the idea is the same except that it is not possible to divide by a vector. An innerproduct multiplication with $\mathbf{r}^k - \mathbf{r}^{k-1}$ is used instead. Aitken's approximation of the inverse of the Jacobian becomes

$$\overline{\left(\frac{d\mathcal{R}}{d\mathbf{p}} \Big|_{\mathbf{p}^k} \right)^{-1}}_{\text{Aitken}} = -\omega^k I, \quad \text{with} \quad \omega^k = -\omega^{k-1} \frac{\langle (\mathbf{r}^{k-1}), (\mathbf{r}^k - \mathbf{r}^{k-1}) \rangle}{\langle (\mathbf{r}^k - \mathbf{r}^{k-1}), (\mathbf{r}^k - \mathbf{r}^{k-1}) \rangle}. \quad (14)$$

The Aitken method results in a recursive, single-diagonal, single-value approximation for the inverse of the Jacobian. The Aitken method is known to be very cheap, while a big increase in stability and a reduction in the number of subiterations are noticed.

Despite the capabilities of the Aitken method, it also has its limits. It weighs the residual of every node with the same factor, which results in a sub-optimal convergence. A method that weights each node differently is the least-squares¹ method [2]. The method is based on the reuse of the state and residual from previous subiterations, including that from previous timesteps, to make a linear extrapolation for the new state.

The method relies on a least-squares approximation of the difference between the newly retrieved data and residual, $\tilde{\mathbf{p}}^{k+1}$, $\tilde{\mathbf{r}}^k$ and a (large) number of previously obtained data and residual vectors, $\tilde{\mathbf{p}}^{i+1}$, \mathbf{r}^i , given by

$$\Delta \tilde{\mathbf{p}}^{i+1} = \tilde{\mathbf{p}}^{i+1} - \tilde{\mathbf{p}}^{k+1}, \quad (15a)$$

$$\Delta \mathbf{r}^i = \mathbf{r}^i - \mathbf{r}^k, \quad (15b)$$

for $i = 0, \dots, k-1$. The desired change in residual vector, $\Delta \mathbf{r}^k = 0 - \mathbf{r}^k$, can be approximated by a linear combination of the known residual changes $\Delta \mathbf{r}^i$, as

$$\Delta \mathbf{r}^k \approx \sum_{i=1}^{k-1} c_i^k \Delta \mathbf{r}^i. \quad (16)$$

Because, in general, this gives an overdetermined system, for the coefficients \mathbf{c}^k , the system is solved as a minimization problem using the least-squares method,

$$\mathbf{c}^k = \arg \min \left\| \mathbf{r}^k + \sum_{i=1}^{k-1} c_i^k \Delta \mathbf{r}^i \right\|_2. \quad (17)$$

¹also called IQN-ILS, Interface Quasi-Newton method with and approximation for the Inverse of the Jacobian using Least-Squares method.

Because $\Delta\tilde{\mathbf{p}}^k$ corresponds to $\Delta\mathbf{r}^k$, a similar linear combination as in (16) gives an approximation for $\Delta\tilde{\mathbf{p}}^k$,

$$\Delta\tilde{\mathbf{p}}^k \approx \sum_{i=1}^{k-1} c_i^k \Delta\tilde{\mathbf{p}}^i. \quad (18)$$

Finally, from (8) and (15) we find

$$\mathbf{p}^{k+1} - \mathbf{p}^k = \Delta\mathbf{p}^k = \Delta\tilde{\mathbf{p}}^k - \Delta\mathbf{r}^k \quad (19)$$

which gives us our new pressure state \mathbf{p}^{k+1} , by substituting (16) and (18), with coefficients found in (17).

Note that (17) may result in a QR-decomposition, applied to a matrix $\mathbf{V}^k = \mathbf{Q}^k \mathbf{R}^k$, whose columns are the vectors $\Delta\mathbf{r}^i$. Then (17) becomes

$$\mathbf{R}^k \mathbf{c}^k = \mathbf{Q}^{kT} \Delta\tilde{\mathbf{r}}^k,$$

and so the approximation of the inverse of the Jacobian using the Least-Squares method is given by

$$\overline{\left(\frac{d\mathcal{R}}{d\mathbf{p}} \Big|_{\mathbf{p}^k} \right)^{-1}}_{\text{LS}} = \mathbf{W}^k \mathbf{R}^{k-1} \mathbf{Q}^{kT} - I, \quad (20)$$

with \mathbf{W}^k a matrix whose columns are the vectors $\Delta\tilde{\mathbf{p}}^i$. This gives a full matrix for the approximated inverse Jacobian, that is diagonally dominated.

3 MULTI-LEVEL ACCELERATION IN FSI

3.1 Derivation of coarse grid prediction method

Let \mathbf{v} be a vector containing all structure variables, and let \mathbf{w} be a vector containing all flow variables (i.e. (u, p) for incompressible and (ρ, u, p) for compressible flows). Let \mathbf{d} be the displacement vector at the fluid-structure interface, Γ_I , and \mathbf{p} be the pressure state on Γ_I . These are obtained by interpolating \mathbf{v} , respectively \mathbf{w} , towards the fluid-structure interface, using an interpolation function ϕ_{Γ_I} , i.e. $\mathbf{d} = \phi_{\Gamma_I}(\mathbf{v})$, $\mathbf{p} = \phi_{\Gamma_I}(\mathbf{w})$. Now define the following two (black-box) solvers:

$$\mathcal{S}(\mathbf{v}; \mathbf{p}) - \mathbf{s} = 0, \quad (21a)$$

$$\mathcal{F}(\mathbf{w}; \mathbf{d}) - \mathbf{f} = 0, \quad (21b)$$

where \mathcal{S} and \mathcal{F} are the implicit parts of the structure and flow solver, with $(a; b)$ their arguments, where a are the variables to be solved and b the input variables from the other solver(s). The vectors \mathbf{s} and \mathbf{f} are the explicit parts of the solvers. These do not change during coupling iterations. We make a distinction between quantities corresponding to a fine and coarse grid using the subscripts h and H , respectively.

The fine grid solution is obtained by solving equations (21a) and (21b) in series. First solve for \mathbf{v}_h^{k+1} using \mathbf{p}_h^k as input, and obtain from that \mathbf{d}_h^{k+1} . Then use this interface displacement as input to obtain the flow state, $\tilde{\mathbf{w}}_h^{k+1}$ and corresponding pressure at the interface, \mathbf{p}_h^{k+1} .

$$\begin{cases} \mathcal{S}(\mathbf{v}_h^{k+1}; \mathbf{p}_h^k) - \mathbf{s}_h^n = 0, \\ \mathcal{F}(\tilde{\mathbf{w}}_h^{k+1}; \mathbf{d}_h^{k+1}) - \mathbf{f}_h^n = 0. \end{cases} \quad (22)$$

Equations (21) are not only used to solve for structure and flow variables, but can also be used in an evaluation to obtain structure and flow residuals. For that, substitute the solutions of the previous iteration,

$$\begin{cases} \mathbf{r}_{s,h}^k = \mathcal{S}(\mathbf{v}_h^k; \mathbf{p}_h^k) - \mathbf{s}_h^n, \\ \mathbf{r}_{f,h}^k = \mathcal{F}(\tilde{\mathbf{w}}_h^k; \mathbf{d}_h^{k+1}) - \mathbf{f}_h^n. \end{cases} \quad (23)$$

Now we can define the error of each quantity, $\epsilon_{\mathbf{v},h}^k = \mathbf{v}_h^{k+1} - \mathbf{v}_h^k$ and $\epsilon_{\tilde{\mathbf{w}},h}^k = \tilde{\mathbf{w}}_h^{k+1} - \tilde{\mathbf{w}}_h^k$. The system to solve becomes

$$\begin{cases} \mathcal{S}(\epsilon_{\mathbf{v},h}^k; \mathbf{0}) = -\mathbf{r}_{s,h}^k, \\ \mathcal{F}(\epsilon_{\tilde{\mathbf{w}},h}^k; \mathbf{0}) = -\mathbf{r}_{f,h}^k, \end{cases} \quad (24)$$

A similar system can be defined for the coarse grid,

$$\begin{cases} \mathcal{S}(\epsilon_{\mathbf{v},H}^k; \mathbf{0}) = -\mathbf{R}\mathbf{r}_{s,h}^k, \\ \mathcal{F}(\epsilon_{\tilde{\mathbf{w}},H}^k; \mathbf{0}) = -\mathbf{R}\mathbf{r}_{f,h}^k, \end{cases} \quad (25)$$

except that we use the reduction, performed by matrix \mathbf{R} , of the fine grid residual to the coarse grid. Indeed the solution is fully coupled when the residuals $\mathbf{r}_{s,h}^k$ and $\mathbf{r}_{f,h}^k$ are zero and not when $\mathbf{r}_{s,H}^k$ and $\mathbf{r}_{f,H}^k$ equal zero. The reduction matrix is also used to obtain the states on the coarse grid,

$$\mathbf{v}_H^k = \mathbf{R}\mathbf{v}_h^k \quad \text{and} \quad \mathbf{w}_H^k = \mathbf{R}\mathbf{w}_h^k. \quad (26)$$

We can replace the errors by the difference between two solves,

$$\begin{cases} \mathcal{S}(\mathbf{v}_H^{k+1}; \mathbf{p}_H^k) - \mathcal{S}(\mathbf{v}_H^k; \mathbf{p}_H^k) = -\mathbf{R}\mathbf{r}_{s,h}^k, \\ \mathcal{F}(\tilde{\mathbf{w}}_H^{k+1}; \mathbf{d}_H^{k+1}) - \mathcal{F}(\tilde{\mathbf{w}}_H^k; \mathbf{d}_H^{k+1}) = -\mathbf{R}\mathbf{r}_{f,h}^k. \end{cases} \quad (27)$$

Finally write it in the form of (22), to obtain the method on the coarse grid level,

$$\begin{cases} \mathcal{S}(\mathbf{v}_H^{k+1}; \mathbf{p}_H^k) - \mathbf{s}_H^{k+1} = 0 & \text{with} & \mathbf{s}_H^{k+1} = -\mathbf{R}\mathbf{r}_{s,h}^k + \mathbf{r}_{s,H}^k + \mathbf{s}_H^k \\ \mathcal{F}(\tilde{\mathbf{w}}_H^{k+1}; \mathbf{d}_H^{k+1}) - \mathbf{f}_H^{k+1} = 0 & \text{with} & \mathbf{f}_H^{k+1} = -\mathbf{R}\mathbf{r}_{f,h}^k + \mathbf{r}_{f,H}^k + \mathbf{f}_H^k \end{cases} \quad (28)$$

Note that in the coarse grid formulation the terms \mathbf{s}_H^{k+1} and \mathbf{f}_H^{k+1} are no longer constant, but that there appears a recurrence relation.

The new structure state is obtained by a prolongation of the coarse grid correction,

$$\begin{aligned}\mathbf{v}_h^{k+1} &= \mathbf{v}_h^k + \mathbf{P}\epsilon_{\mathbf{v},H}^k \\ &= \mathbf{v}_h^k + \mathbf{P}(\mathbf{v}_H^{k+1} - \mathbf{v}_H^k)\end{aligned}\quad (29)$$

and the new fine grid states of the flow states are given using the following prolongation,

$$\tilde{\mathbf{w}}_h^{k+1} = \mathbf{w}_h^k + \mathbf{P}(\tilde{\mathbf{w}}_H^{k+1} - \mathbf{w}_H^k). \quad (30)$$

From $\tilde{\mathbf{w}}_h^{k+1}$ we can extract the pressure state at the fluid-structure interface, $\tilde{\mathbf{p}}_h^{k+1}$, using the map ϕ_{Γ_I} . Now calculate the approximate inverse Jacobian, \mathcal{J}^{-1} , with one of the methods discussed in the previous section, and apply the quasi-Newton step,

$$\mathbf{p}_h^{k+1} = \mathbf{p}_h^k + \mathcal{J}^{-1}(-\mathbf{r}^k), \quad (31)$$

where $\mathbf{r}^k = \tilde{\mathbf{p}}_h^{k+1} - \mathbf{p}_h^k$, so

$$\begin{aligned}\mathbf{p}_h^{k+1} &= \mathbf{p}_h^k + \mathcal{J}^{-1}(-\mathbf{P}\epsilon_H^k) \\ &= \mathbf{p}_h^k + \mathcal{J}^{-1}(-\mathbf{P}(\tilde{\mathbf{p}}_H^{k+1} - \mathbf{p}_H^k)).\end{aligned}\quad (32)$$

3.2 Multi-level convergence criteria

As mentioned in the introduction, a multi-level convergence criteria will be derived. The fine grid residual, \mathbf{r}_F , can be decomposed in a part corresponding to the coarse grid residual on the fine grid, \mathbf{r}_C , and high wave number residual term, $\boldsymbol{\delta}$, for which holds that it vanishes when restricting it to the coarse grid, $\mathbf{R}\boldsymbol{\delta} = 0$,

$$\mathbf{r}_F = \mathbf{r}_C + \boldsymbol{\delta}. \quad (33)$$

Note that compared to the previous section, we have $\mathbf{r}_F = \mathbf{r}_h$ and $\mathbf{r}_C = \mathbf{P}\mathbf{r}_H$, so both are defined w.r.t. the fine grid. As we subiterate on the coarse grid, we will notice that $\mathbf{r}_C \rightarrow 0$, while $\mathbf{r}_F \rightarrow \boldsymbol{\delta}$. The new multi-level convergence criteria is related to the ratio between the coarse grid residual and the high wave number residual,

$$\frac{\|\mathbf{r}_C\|}{\|\boldsymbol{\delta}\|} = \frac{\|\mathbf{r}_C\|}{\|\mathbf{r}_F - \mathbf{r}_C\|} \leq C. \quad (34)$$

A reasonable criteria is to set $C = 1$, in that case $\|\mathbf{r}_C\| \leq \boldsymbol{\delta}$. Then the remaining coarse grid residual is smaller than the high wave number residual, that cannot be removed on the coarse level. Therefore a prolongation to a finer grid seems reasonable.

Note that in practice the high wave number residual $\|\mathbf{r}_F - \mathbf{r}_C\|$ is not known, but the individual residuals, $\|\mathbf{r}_F\|$ and $\|\mathbf{r}_C\|$, are. Instead, we will measure a slightly higher value,

$$\frac{\|\mathbf{r}_C\|}{\|\mathbf{r}_F\| - \|\mathbf{r}_C\|} \geq \frac{\|\mathbf{r}_C\|}{\|\mathbf{r}_F - \mathbf{r}_C\|}. \quad (35)$$

4 NUMERICAL INVESTIGATION

The discussed reduced order models are applied to the cylinder with flapping beam benchmark case, version FSI3, introduced in [9]. The case shows transient behavior with large displacement and is strongly coupled, since the densities of the flow and structure are equal. Details of the used solvers can be found in [12]. Most important to mention here is that time integration is performed using the ESDIRK3 method, an implicit, third-order accurate, multistage Runge-Kutta scheme, see [11]. Each implicit Runge-Kutta stage is subiterated until $\|\mathbf{r}_h\| \leq 10^{-3}$. The coarse and fine grids are shown in figure 1. The ratio of fine to coarse grid is 20737 to 5442 cells, i.e. about 3.8.

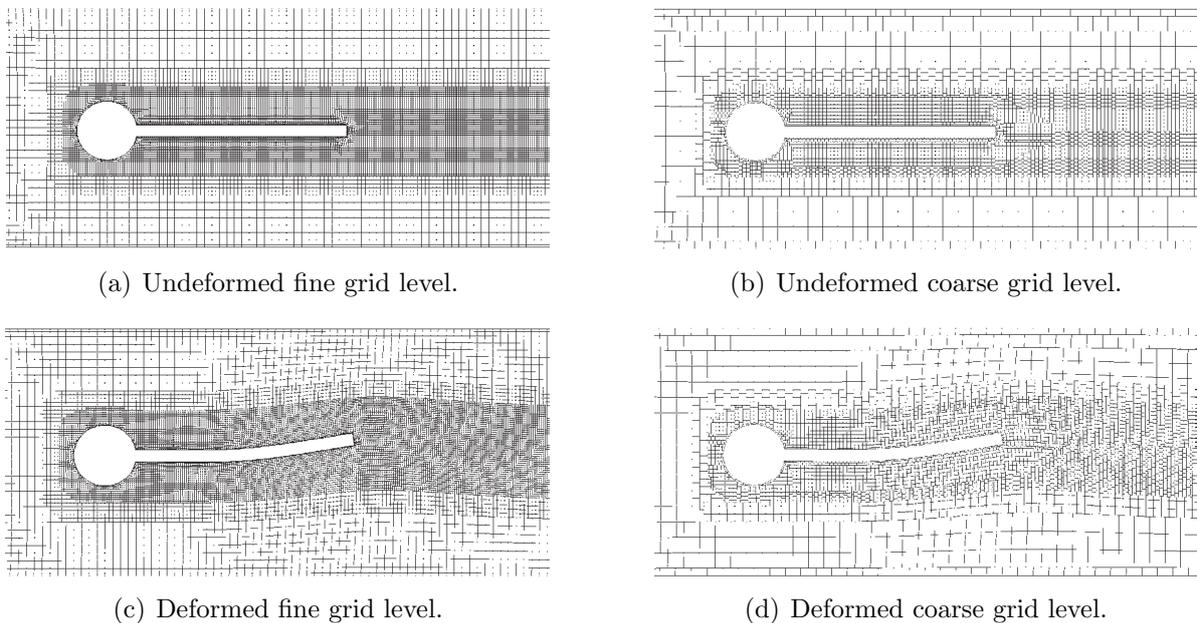


Figure 1: Fine and coarse level meshes for undeformed and deformed situation.

4.1 Results

A list of different multi-level strategies are tested in combination with Aitken and Least-squares quasi-Newton method:

- **FG:** Fine grid solves only.
- **∞ CG- ∞ FG:** First iterate on coarse grid level until convergence, thereafter iterate on fine grid level until convergence.
- **2CG-1FG:** Alternate between two coarse grid solves with one fine grid solve. Final solve is always a fine grid solve.

Method		Total	Fine	Coarse	
Aitken	FG	30.7 10.2	30.7 10.2	0.0	0.0
Least-squares	FG	20.7 6.9	20.7 6.9	0.0	0.0
Aitken	∞ CG- ∞ FG	56.6 18.9	21.2 7.1	35.4	11.8
Least-squares	∞ CG- ∞ FG	42.7 14.2	18.4 6.1	24.3	8.1
Aitken	2CG-1FG	36.4 12.1	12.5 4.2	23.9	8.0
Least-squares	2CG-1FG	30.5 10.2	10.7 3.6	19.8	6.6
Aitken	3CG-1FG	38.2 12.7	10.3 3.4	27.9	9.3
Least-squares	3CG-1FG	35.1 11.7	9.5 3.2	25.6	8.5
Aitken	adaptive $C = 1.0$	36.4 12.1	12.2 4.1	24.2	8.1
Least-squares	adaptive $C = 1.0$	27.4 9.1	10.8 3.6	16.6	5.5
Aitken	adaptive $C = 0.5$	39.2 13.1	10.9 3.6	28.3	9.4
Least-squares	adaptive $C = 0.5$	29.4 9.8	9.2 3.1	20.2	6.7

Table 1: Average number of subiterations for 100 timesteps. In black: average per timestep. In blue: average per implicit RK stage.

- **Adaptive $C = 0.5$:** Adaptively alternate between coarse grid solves with fine grid solves. Final solve is always a fine grid solve.

The results are given in table 1. It shows the average number of subiterations taken over 100 timesteps. It show both the average number per timestep as well as per implicit Runge-Kutta stage. With respect to fine grid solves only, the ∞ CG- ∞ FG method shows a major increase in total number of subiterations and only a minor decrease in the expensive fine grid solves. The alternating method 2CG-1FG also has an increase in total number of subiterations, but the number of fine grid subiterations is already halved. Finally the adaptive multi-level method with $C = 0.5$ shows the best result. With respect to fine grid level only, a minor increase in total number of subiterations is noticed, while there is a major decrease in fine grid solves. Per implicit RK stage, the number of fine grid solves reduces from 10.2 to 3.6 when using Aitken’s method, and from 6.9 to 3.1 when using least-squares method.

4.2 Convergence behavior

Figure 2 shows the convergence behavior for a full timestep of Aitken’s and least-squares method using fine grid level only, and the least-squares method with adaptive multi-level strategy. Most important to see is that for the first five subiterations each stage, the multi-level approach has the same convergence behavior as the fine grid level only approach. Only for the last few subiterations the convergence ratio is less for the multi-level approach than for the fine grid level only approach.

To indicate the existence of δ , we compare the convergence behavior of the ∞ CG- ∞ FG strategy with both Aitken and least-squares, with the convergence of least-squares adap-

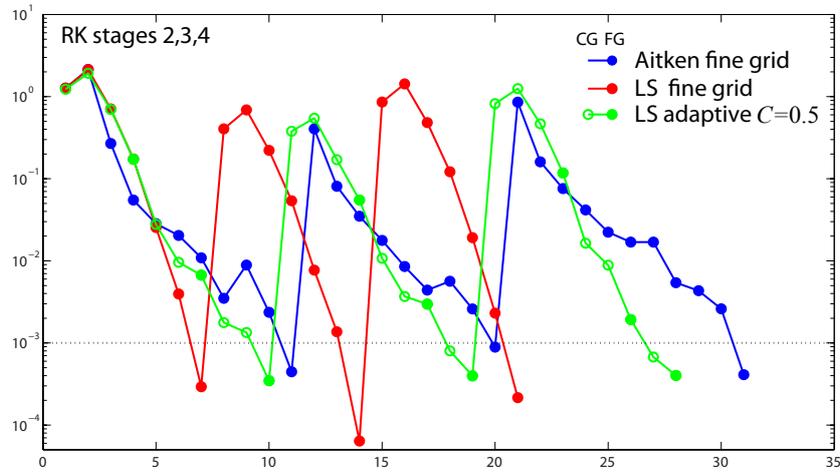


Figure 2: Convergence behavior of a typical timestep.

tive multi-level $C = 0.5$, see figure 3. The ∞ CG- ∞ FG methods show a clear jump when switching from coarse to fine level. Both Aitken and Least-squares method jump almost to the same level. This clearly indicates that although convergence on the coarse grid has been reached, there remains to exist a high frequency mode, δ , that cannot be suppressed by the coarse level solves. This mode is indeed recognized by the adaptive multi-level strategie, which therefore uses a finegrid solve just above the δ -level, to suppress this mode.

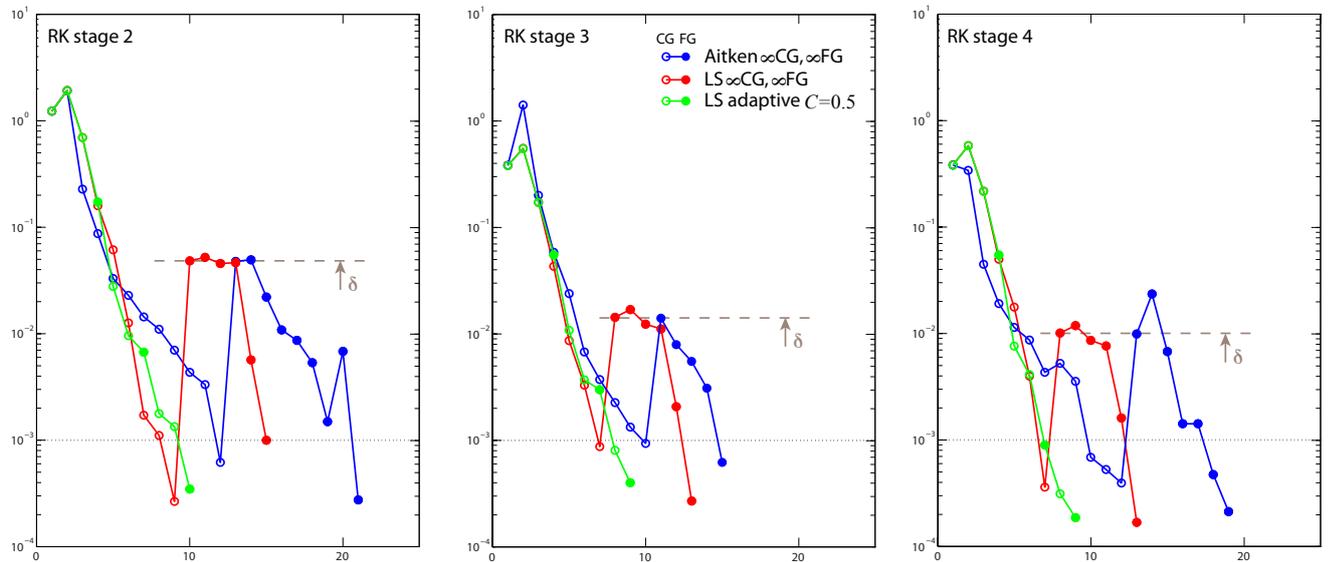


Figure 3: Convergence behavior of a typical timestep, that indicates the existence of the high frequency mode, δ .

REFERENCES

- [1] E.H. van Brummelen, C. Michler and R. de Borst, “Interface-GMRES(R) acceleration of subiteration for fluid-structure-interaction problems”, Technical Report *Delft Aerospace Computational Science*, (2005), DACS-05-001.
- [2] J. Degroote, R. Haelterman, S. Annerel, P. Bruggeman, J. Vierendeels, “Performance of partitioned procedures in fluid-structure interaction”, *Computers & Structures*, (2010), **88**, 446-457.
- [3] J. Degroote, S. Annerel and J. Vierendeels, “Stability analysis of Gauss-Seidel iterations in a partitioned simulation of fluid-structure interaction”, *Computers & Structures*, (2010), **88**, 5-6, 263-271.
- [4] C. Farhat, K. van der Zee and P. Geuzaine, “Provably second-order time-accurate loosely coupled solution algorithms for transient nonlinear computational aeroelasticity”, *Comp. Meth. in Appl. Mech. and Eng.*, (2006), **195**, 17-18, 1973-2001.
- [5] T. Gallinger and K. Bletzinger, “Comparison of algorithms for strongly coupled partitioned fluid-structure interaction - Efficiency versus simplicity”, *Proceeding ECCOMAS CFD 2010*.
- [6] U. Küttler and W. Wall, “Fixed-point fluid-structure interaction solvers with dynamic relaxation”, *Computational Mechanics*, (2008), **43**, 1, 61-72.
- [7] C. Michler, E.H. van Brummelen and R. de Borst, “An interface Newton-Krylov solver for fluid-structure interaction, *International Journal for Numerical Methods in Fluids*, (2005), **47**, 10-11, 1189-1195.
- [8] S. Piperno, C. Farhat, B. Larrouturou, “Partitioned procedures for the transient solution of coupled aeroelastic problems. Part I: model problem, theory and two-dimensional application”, *Comp. Meth. Appl. Mech. Eng.*, (1995), **124**, 1-2, 79-112.
- [9] S. Turek, J. Hron, in *Fluid-Structure Interaction: Modelling, Simulation, Optimisation*, ed. by H.J. Bungartz, M. Schaefer, Springer, (2006).
- [10] A.H. van Zuijlen, S. Bosscher, H. Bijl, “Two level algorithms for partitioned fluid-structure interaction computations”, *Comp. Meth. Appl. Mech. Eng.*, (2007), **196**, 1458-1470.
- [11] A.H. van Zuijlen, A. de Boer and H. Bijl, “Higher-order time integration through smooth mesh deformation for 3D fluid-structure interaction simulations”, *Journal of Computational Physics*, (2007), **224**, 1, 414-430.
- [12] A.H. van Zuijlen and H. Bijl, “Multi-level accelerated sub-iterations for fluid-structure interaction”, in *FSI II*, Lecture notes Comp. Sc. Eng., (2010), **73**, 1-25.

2 DESIGN

O-PALM is a dynamic coupler in the sense that it can be used to implement a coupling where the component execution scheduling and the data exchange patterns cannot be entirely defined before execution.

The historical reason for this feature dates back to 1996 when the MERCATOR operational oceanography project faced the problem to set-up a new operational suite with Data Assimilation for research and operations in an evolving configuration.

Data assimilation can be roughly defined as a collection of techniques aiming to improve numerical models skills by the use of observational data. They are based on computationally expensive algebraic algorithms involving a model, observation treatments, and the statistical characterisation of the errors on both sides. At that time, in MERCATOR, the choice of the model configuration was not yet finalized, there were several candidate assimilation methods to test, there were different kinds of observations to handle and the same system should have been used for research and operations. All these needs of flexibility lead to the implementation of the assimilation suite as a coupling between model, observations handling, error statistics and algebra instead of hard-coding data assimilation routines in the model, or vice-versa [5].

Some data assimilation algorithms are based on an iterative minimization: this implies the repeated execution of the tasks and the total number of iterations is not necessarily known beforehand. Moreover, in some configurations some tasks are activated only if some observations are available at run-time.

This specific requirement imposed to conceive a coupler of independent parallel codes capable to deal with complex coupling algorithms allowing for the conditional and/or repeated execution of the coupled components. The main goals and constraints were user friendliness, modularity, portability and high performances on parallel computers.

Existing couplers were not a suitable choice for the lack of the dynamic aspects. This lead, after a thorough feasibility study [6], to the design of a new MPMD dynamic parallel coupler, based on the MPI message passing and process management standard library [7,8].

In our definition, a dynamic coupler has to fulfil three main requirements:

- process management: this means that the coupler has to be able to start and synchronise the tasks and to handle algorithms with loops and conditional switches
- buffered communications: in order to grant full flexibility, avoiding deadlocks dependencies on the production and reception order, at least the production side of a communication has to be non blocking. This requires the explicit handling of a storage space for pending communications. This feature allows for some extra possibilities, such as the linear combination of cumulated fields and the explicit permanent storage of objects that are to be repeatedly received
- object versioning: the flexible use of a temporary storage space for parallel communications requires special care to grant the coherency of the stored global objects. The Last In Only Out paradigm is adopted: every new version of an object replaces the previous ones. Nevertheless, for parallel communications, we count a new version of an object only when all the processes of the producing code have provided

their contributions. For loosely synchronised codes, it implies the introduction of stamps to keep track of what version of an object new contributions belong to.

The same way, in a parallel coupling, a coupler has to deal with two levels of parallelism:

- concurrent tasks parallelism: independent tasks can run concurrently on separate sets of processes. The coupler has to deal with the concurrent execution, to establish all the needed intercommunication contexts and to grant synchronisation.
- distributed coupled codes: as a second level of parallelism we account for the inner parallelism of the coupled codes, mostly related to data distribution. The coupler has to grant private and robust intracommunication contexts and, most important, to be able to manage the data exchanges between sets of processes, including the remapping between codes with different distributions of the same physical objects.

Since one of the main aims of coupling is the reuse of legacy codes, we tried to reduce the intrusiveness of the coupling instructions in the source codes. For this reason we adopted the so-called end point communication paradigm: the producer of an object does not know anything about the recipients (if any) and the other way round. The coupler makes the matching.

In order, once more, to minimize the interventions in the codes, we defined a reduced set of multi-language API calls, complemented by a very detailed Graphic User Interface: most information - such as the coupling algorithm, the communication patterns and the parallel distributions - is easily described in the graphic interface. The changes in the code have minimal impact and, because of the use of the one-sided communications paradigm, they are independent of the specific coupling algorithm.

O-PALM can thus be interfaced to F77, F90, C or C++ compiled codes or to the main interpreted languages such as Python, Perl, Java, Tcl/Tk, Octave and also to black-box pre-compiled executables that can be linked at run-time against dynamic libraries.

The last constraints that drove the design of O-PALM are related to its uses and diffusion. The operational usage imposes robustness and high performances. This not only determined some implementation choices, but it also lead to the integration in O-PALM of a real-time monitor, allowing to display in the graphic user interface the status of the execution while running and of a performance analyser that works on trace files and helps tuning and optimising the coupled application. For research applications there are other criteria, such as portability, that imposed to rely on standard coding and message passing techniques) and user friendliness. The latter not only drove the design of the Graphic User Interface, but it lead also the the introduction of an algebra toolbox providing a palette of predefined generic algebraic operations ranging from BLAS to parallel linear algebra solvers and minimisers that can be coupled to any other user defined code.

Since then, the range of O-PALM applications has largely extended beyond data assimilation and it is the coupler of choice in a number of multi-physics coupling when dealing with flexible configurations [9]. Most relevant applications are in computational fluid dynamics for fluid-structure interactions modelling or for automatic shape optimisation computations or in event driven couplings in soil surface modelling [10]. New applications

carried new requirements for the design of O-PALM future versions, but also some contributed developments. For this reason, the coupler has become open-source at the beginning of 2011. This is the most suitable environment to deal with the new challenges that massively parallel machines and exascale computing will carry into high performances scientific and engineering computing.

3 IMPLEMENTATION

The coupler implementation went through several steps.

At the very beginning of the project, when the MPI2 standard [11] was recently published, but hardly any complete and robust implementation was available, we implemented an MPI1 emulation based on a pool of idle processes, released under the name of PALM_RESEARCH, later changed into PALM_SP. It was dedicated to functional tests, but in practice it proved to be very effective in some cases and it still used for some full size applications. Some interesting features of this first implementation could now be seen as possible optimisations under some conditions and will be hopefully reintroduced in the current O-PALM version in a near future.

In 2003 we released the first fully MPMD version of O-PALM under the name PALM_MP. It was based on the MPI2 process management and communication layer. The main components of PALM_MP are

- the scheduler that handles the process management and the execution of the coupled components accordingly to the algorithm described in the user interface. PALM can schedule several parallel codes to run concurrently to perform independent tasks if enough resources are available. Since starting an independent executable always causes a overhead, PALM offers the option to merge into a single executable the coupled components that are started in a sequence.
- the optimised communication scheme managed by a driver that takes care of the data transfer between parallel programs. This is one of the most evolved components of PALM and handles very complex communication patterns with some very practical features, such as the remapping of objects exchanged by parallel codes with different distributions, the selection of object subsets entirely from the user interface, the presence of an explicitly managed permanent repository for objects to be repeatedly received.

Since then the coupler has been constantly enhanced and optimised.

With respect to PALM_MP, the current O-PALM release offers

- the possibility to interface commercial black-box codes (such as Fluent, Abaqus MSC/MARC) by the use of external dynamic libraries and/or a socket based layer
- the possibility to interface interpreted languages such as
- the implementation of a simplified working mode entirely compliant with the MPI-1 library
- the optimisation of repeated well synchronised communications that don't require the intervention of the driver
- the enhancement of the parallel algebra toolbox that is soon going to include the

CWIPI interpolation library from ONERA for the grid to grid remapping.

More technical details on the implementation go beyond the scope of this paper. A detailed source of information is the web site dedicated to the O-PALM coupler http://www.cerfacs.fr/globc/PALM_WEB.

The Graphic User Interface, called PrePALM, is a portable Tcl/Tk application (Fig. 1). It deserves some words because it is the part of the coupler that users interact with most of the time. The relevant features of the coupled components are described in identity cards that do not depend on the specific coupling algorithms. The user describe the execution scheduling, the parallel sections, the data exchange patterns and the algebraic treatments, entirely from within the user interface. It ends up providing the input file for the coupler executable itself and the source code for the wrappers of the coupled component that take entirely care of the set-up of the communication context with no need of change in the components sources.

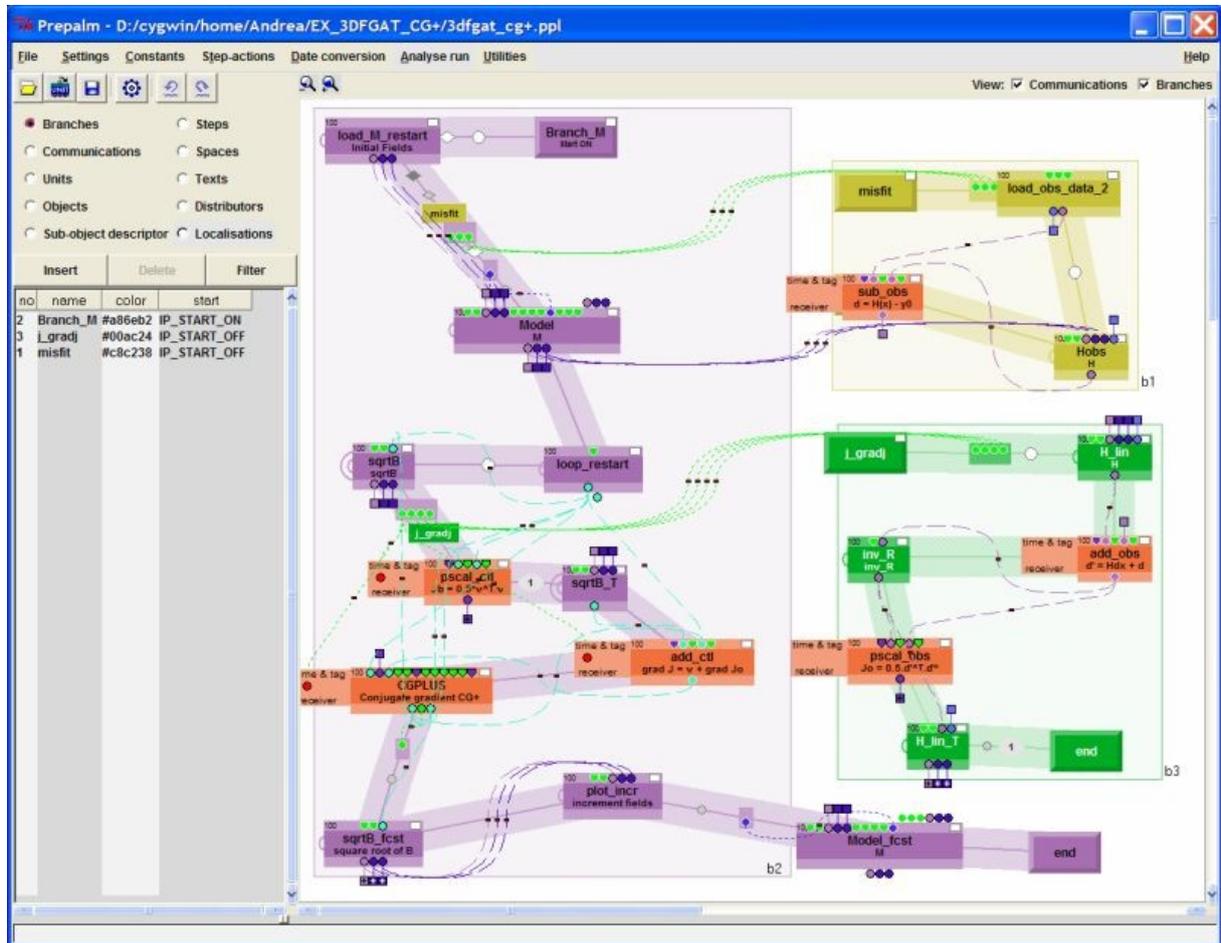


Figure 1: The PrePALM Graphic User Interface

The same graphic tool can be used at run-time to monitor the simulation status, with colour codes to distinguish between components running, waiting for resources or not yet scheduled

and with counters for the completed executions and the performed data exchanges. The same tool again provides post-mortem some statistics on the memory and CPU time resources used by the components (Fig. 2).

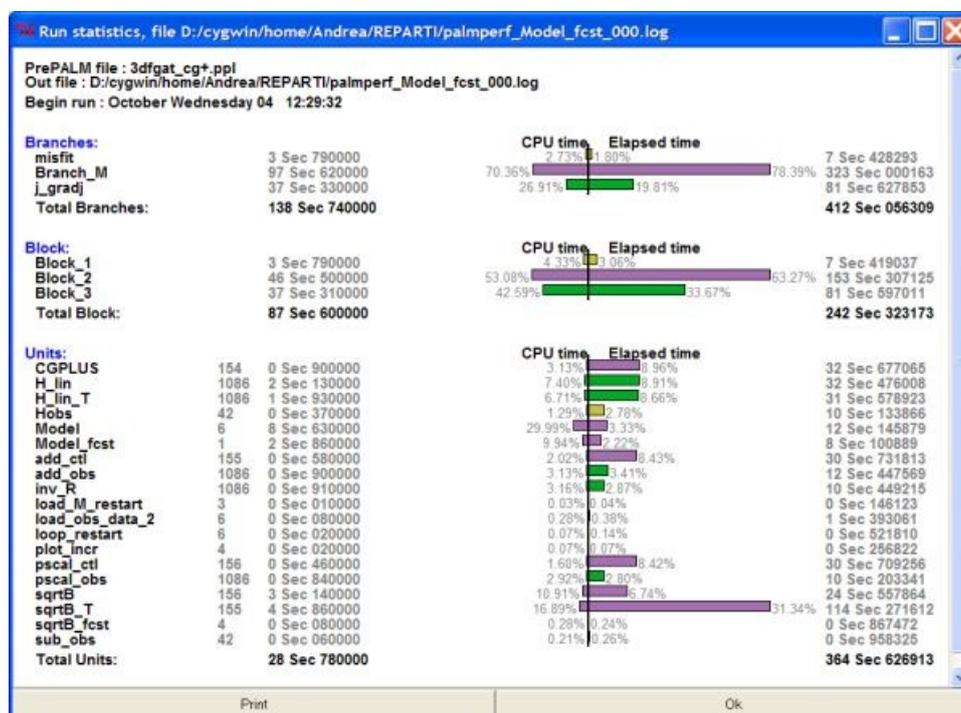


Figure 2: The PrePALM Graphic User Interface

4 FUTURE DIRECTIONS

Among the most important technical challenges for the evolution of O-PALM towards exascale applications, there is the search for the best trade-off between a centralised and a fully distributed approach. If on one side process management and monitoring is a key issue for dynamic coupling, it imposes an overhead and a risk of bottlenecks for massively parallel applications. Optimisations bypassing the O-PALM scheduler and launcher are under study, but they have to keep the capability of reorganizing the layout of the application in case, for instance, of automatic load balancing or adaptive meshes.

The same considerations apply for the communication handling: to obtain very effective parallel communications on massively parallel configurations, we'll have to look the best compromise between flexibility and monitoring on one side and performances on the other.

Furthermore the new directions in high performances computing introduce some further constraints, such as the capability to reorganise the coupling layout around self-tuning applications, changing their configuration at run-time or the capability to implement a resilient parallel coupling on a very high number of processors.

Finally, some application domains require specific treatments. It is the case of the physical

constraints (e.g. mass or energy conservation) of the grid-to-grid interpolations for climate modelling applications. In such a case, the use of O-PALM and the CWIPI interpolation library (with specific enhancements) has to be thoroughly studied and evaluated on test cases of increasing complexity and size.

The open source distribution of O-PALM is the most suitable environment to accept collaborations and contributions on the coupler development on all these topics.

5 SOME APPLICATIONS

Current O-PALM applications largely go beyond data assimilation and cover many fields of multi-physics coupling ranging from oceanography to hydraulics, from hydrology to agronomy, from aeronautics to space engineering and so on.

Some of them are particularly representative of the advantages coming from the dynamic coupling and from the user friendliness of the API's and of the user interface.

For instance we could mention the use of O-PALM for the coupling of an adaptive 2D surface biosphere model to different parallel atmosphere circulation limited area models. In such a case, not only the dynamic features of O-PALM can easily take into account the adaptive model, but also the compact syntax used to describe data exchanges allows for a quite generic implementation with different atmosphere models [12].

Some full size, near real time applications, like the operational ocean data assimilation and forecast suite of the MERCATOR operational oceanography centre or the air quality data assimilation and forecasting system Valentina, based on the MOCAGE chemistry and transport model [13], provide a very satisfactory test bench for the O-PALM performances in large scale parallel applications.

We should also mention the recent use of O-PALM for the implementation of a demonstrative data assimilation suite based on a 1D hydraulic model used in flood forecasting. The graphic algorithm representation proves to be a very useful pedagogical tool. Furthermore, the generic formalism allows for the application of the demo suite to real life applications with no changes in the code lines [14].

To give a detailed, application oriented example of the use of O-PALM in real size problems, we present here a study in a relevant computational fluid dynamics research domain: the determination of heat loads that is a key issue in gas turbines conception, because wall temperatures and heat fluxes are a major constraint in the design of combustors and turbine blades. Indeed, the life duration of turbine components directly depends on the wall temperature and therefore designers imperatively need an accurate prediction tool. Numerical simulations of the thermal interaction between fluid flows and solids is therefore of primary interest. The complex flows observed in the turbine environment are much better predicted in the Large Eddy Simulations (LES) framework, especially when considering thermal effects.

To simulate a cooled turbine blade it is necessary to couple a LES solver and a heat transfer code within solids [15,16]. The LES solver used is the AVBP code [17,18,19,20], which solves the full compressible Navier-Stokes equations on unstructured meshes, using a

cell-vertex/finite element approximation and a Taylor-Galerkin weighted residual central distribution scheme. The calculation of thermal diffusion in solids is performed with the code AVTP, solving the classical heat equation [16]. AVTP also uses unstructured meshes and is advanced with a first-order explicit forward Euler scheme.

The study simulates a cooled blade of the T120 cascade, which was designed by Rolls Royce Deutschland for the European project AITEB [21]. Experiments were conducted in the High-speed Cascade Wind Tunnel of the Institute of Jet Propulsion of Aachen [22]. The highly-loaded high-pressure turbine airfoil of the T120 cascade was designed to have a large separation on the pressure side. The blade is operated at a Reynolds number of $3.8 \cdot 10^5$ and a Mach number of 0.87, based on the exit velocity and the chord. The film cooling device of the T120D blade is composed of three holes located on the pressure side, repeated in the spanwise direction to form a pattern of jet rows (Fig. 3-a). The computational domain covers one cooling hole pattern in the spanwise direction, with periodicity boundary conditions. The unstructured mesh is composed of 6.5 millions of tetrahedral elements for the fluid zone, and 600 000 elements within the solid. The skin meshes are the same for the fluid and the solid so that no interpolation error is introduced at this level when CHT is simulated. The converged thermal state is obtained in 10 characteristic solid time scale and requires about 4800 CPU hours. At the converged state, the net heat flux through the blade reaches zero. Figure 3-b shows an instantaneous snapshot of temperature distribution in the fluid and solid domains.

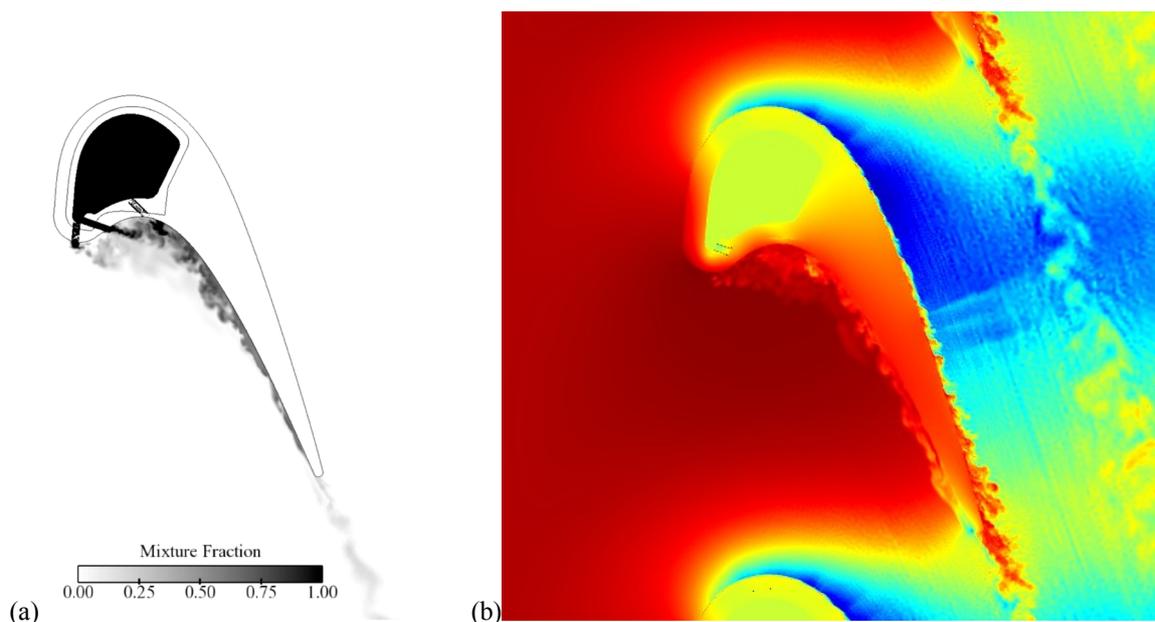


Figure 3: Instantaneous snapshot of cooling air distribution (a) and temperature distribution in the fluid and solid domains (b).

Figure 4 compares experimental and numerical cooling efficiency fields on a 2D plot over the pressure side.

The computation matches the experimental visualization fairly well and evidences the thermal effects of the cooling jets on the vane.

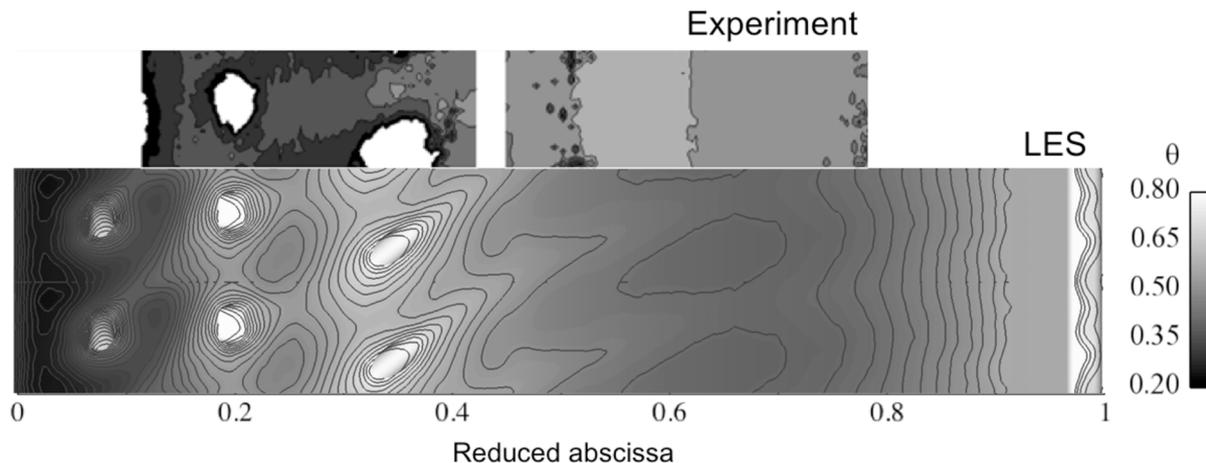


Figure 4: 2D plot of time-averaged cooling efficiency on the pressure side: comparison of experimental results and coupled simulation. The scale of θ corresponds only to the LES field.

6 CONCLUSIONS

The O-PALM coupler has reached a high degree of maturity and stability. It is currently used for more than 50 research and industrial projects ranging from operational data assimilation to multi-physics modelling, from climate change impact assessment to fluid and structure interactions. It implements a dynamic coupling paradigm based on portable and effective technical solutions. It is well suited for the evolution of the current coupling technology towards the exaflop machines of next generations and several research centres and engineering companies are attentively considering this issue. Having become open-source since the beginning of 2011, O-PALM is aiming to be a tool of reference for dynamic and high performances coupling applications and ongoing researches in coupling technology.

REFERENCES

- [1] S. Valcke, S. and Redler, R. *OASIS4 User Guide* PRISM–Support Initiative Technical Report No 4 (2006).
- [2] Valcke, S., Guilyardi, E. and Larsson, C. PRISM and ENES: A European approach to Earth system modelling. *Concurrency Computat.: Pract. Exper.* (2006) **18**(2):231-245.
- [3] Valcke, S., Déclat, D., Redler, R., Ritzdorf, H., Schoenemeyer, T. and Vogelsang, R. The PRISM Coupling and I/O System. VECPAR'04, Proceedings of the 6th International Meeting, vol. 1 *High performance computing for computational science* (2004).
- [4] Bahrel, P. Mercator Océan. Global to regional ocean monitoring and forecasting. In *Ocean weather forecasting: An integrated view of operational oceanography*. Springer (2006):381-395.
- [5] André, J.C., Blanchet, I. and Piacentini, A. Besoins algorithmiques et informatiques en océanographie opérationnelle. Le programme Mercator. *Calculateurs Parallèles, Réseaux et Systèmes répartis* (2000) **11**(3):275-294.
- [6] The PALM group Etude de faisabilité du projet PALM, *CERFACS Technical Report TR/CMGC/98-50* (1998).
- [7] Fouilloux, A. and Piacentini, A. The PALM Project: MPMD Paradigm for an Oceanic

- Data Assimilation Software, *Lecture Notes In Computer Science* (1999) **1685**:1423-1430.
- [8] Piacentini, A. and The PALM Group, PALM: A Dynamic Parallel Coupler, *Lecture Notes In Computer Science* (2003) **2565**:479-492.
- [9] Buis, S., Piacentini, A. and Déclat, D. PALM: A Computational framework for assembling high performance computing applications, *Concurrency Computat.: Pract. Exper.* (2006) **18**(2):247-262.
- [10] Duret, S., Drouet, J.L., Hutchings, N.J., Durand, P., Theobald, M.R., Salmon-Monviola, J., Dragosits, U., Maury, O., Sutton, M.A. and Cellier, P. NitroScape: an integrated model of nitrogen fluxes and transformation at the landscape scale. *Environmental Pollution*. Submitted.
- [11] *MPI: A Message-Passing Interface Standard, Version 2.2*, High Performance Computing Center Stuttgart (2009)
- [12] Haas, E. and Piacentini, A. Integrated Earth System Modeling - A Coupled Multiscale Bio-, Hydro- and Atmosphere Model for Numerical Process Simulations and Regional Climate Impact Analysis, *Mathematics and Computers in Simulation* (2010) **80**(8) Special Issue ESCO 2008 Conference.
- [13] Massart, S., Clerbaux, C., Piacentini, A., Cariolle, D., Turquety, S. and Hadj-Lazaro, J. First steps toward the assimilation of IASI ozone data into the MOCAGE-PALM system, *Atmos. Chem. Phys.* (2009) **9**:5073-5091.
- [14] Ricci, S., Piacentini, A., Thual, O., Le Pape, E. and Jonville G. Correction of upstream flow and hydraulic state with data assimilation for flood forecasting, *Hydrol. Earth Syst. Sci. Discuss.* (2010) **7**:1-55.
- [15] Duchaine, F., Mendez, S., Nicoud, F., Corpron, A., Moureau, V. and Poinso, T. Conjugate heat transfer with large eddy simulation application to gas turbine components. *C. R. Acad. Sci. Mécanique* (2009) **337**(6-7):550-561.
- [16] Duchaine, F., Corpron, A., Pons, L., Moureau, V., Nicoud, F. and Poinso, T. Development and assessment of a coupled strategy for conjugate heat transfer with large eddy simulation: Application to a cooled turbine blade. *Int. J. Heat Fluid Flow* (2009) **30**:1129-1141.
- [17] Schönfeld, T. and Poinso, T. Influence of boundary conditions in LES of premixed combustion instabilities. In *Annual Research Briefs of the Center for Turbulence Research, NASA Ames, Stanford Univ, USA* (1999):73-84.
- [18] Moureau, V., Lartigue, G., Sommerer, Y., Angelberger, C., Colin, O. and Poinso, T. Numerical methods for unsteady compressible multi-component reacting flows on fixed and moving grids. *J. Comput. Phys.* (2005) **202**(2):710-736.
- [19] Mendez, S. and Nicoud, F. Large-eddy simulation of a bi-periodic turbulent flow with effusion. *J. Fluid Mech.* (2008) **598**:27-65.
- [20] Roux, A., Gicquel, L. Y. M., Reichstadt, S., Bertier, N., Staffelbach, G., Vuillot, F. and Poinso, T. Analysis of unsteady reacting flows and impact of chemistry description in large eddy simulations of side-dump ramjet combustors. *Combust. Flame* (2010) **157**:176-191.
- [21] Haselbach, F. and Schiffer, P. Aiteb - an european research project on aerothermodynamics of turbine endwalls and blades. *Int. J. of Therm. Sci.* (2007) **13**(2):97-108.
- [22] Gomes, R. A. and Niehuis, R. Film cooling effectiveness measurements with periodic

unsteady iflow on highly loaded blades with main flow separation. In *Proceedings of ASME Turbo Expo 2009: Power for Sea, Land and Air* . Orlando, Florida, USA (2009).

PARTITIONED TIME INTEGRATION METHODS FOR HARDWARE IN THE LOOP BASED ON LINEARLY IMPLICIT L-STABLE ROSENBROCK METHODS

ORESTE S. BURSI^{*}, CHUANGUO JIA[†], ZHEN WANG^{*}

^{*} Department of Mechanical and Structural Engineering
University of Trento
Via Mesiano 77, 38123 Trento, Italy
e-mail: oreste.bursi@ing.unitn.it, www.ing.unitn.it/~bursi/

[†] School of Civil Engineering, Chongqing University, Chongqing, China

Key words: Real-time, Rosenbrock integration methods, Interfield partitioned methods, Hardware-in-the-loop test.

Abstract. Hardware in the loop based on dynamic substructuring was conceived to be a hybrid numerical-experimental technique to simulate the non-linear behaviour of an emulated structure. Its challenge is to ensure that both numerical and physical substructures interact in real time by means of actuators –transfer systems-. With this objective in mind, the development and implementation of partitioned real-time compatible Rosenbrock algorithms are presented in this paper. In detail, we shortly introduce monolithic linearly implicit L-stable algorithms with two stages; and in view of the analysis of complex emulated structures, we present a novel interfield partitioned algorithm. Both the stability and accuracy properties of the proposed algorithm are examined through analytical and numerical studies carried out on Single-DoF model problems. Moreover, a novel test rig conceived to perform both linear and nonlinear substructure tests is introduced, and tests on a two-DoF split-mass system are illustrated. The drawbacks of this algorithm are underlined and improvements are introduced on a companion solution procedure.

1 INTRODUCTION

In recent years real-time hybrid testing techniques, as depicted in Figure 1, like the Hardware-in-the-Loop (HiL) technique with Dynamic Substructuring (DS), became more and more popular in order to study the performance of components and structures subject to dynamic loads [1,2]. With regard to relevant time-stepping methods, they can be broadly classified in monolithic and partitioned. In a monolithic approach, the method integrates: i) the Numerical Substructure (NS) only, whilst the Physical Substructure (PS) is considered a black box [2]; ii) both the NS and the PS by means of stiffness estimates [3], like in a typical pseudo-dynamic (PsD) test [2]. Conversely, a partitioned approach solves both NS and PS through different integrators and takes into account the interface problem, for instance by prediction, substitution and synchronization of Lagrange multipliers. In detail, partitioned

algorithms can be applied to the Euler-Lagrange form of the equations of motion -second-order in time- [4-6] or to the Hamilton form of the equations of motion -first-order in time- [7,8]. In this paper, we consider partitioned approaches based on L-stable two-stages real-time compatible Rosenbrock (LSRT) algorithms applied to equations of motion first-order in time.

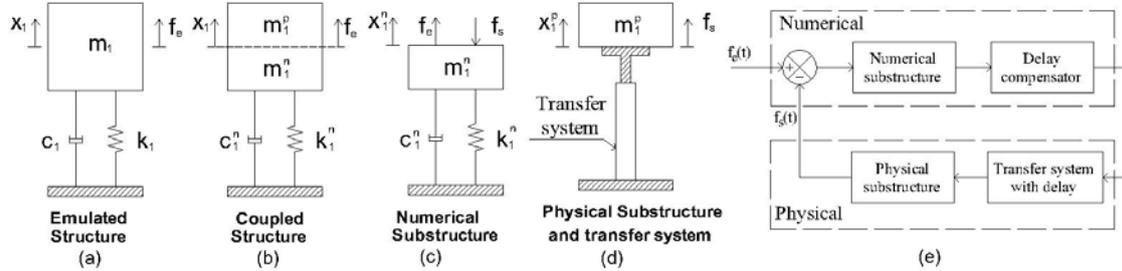


Figure 1: (a)-(d) Schematic representation of a SDOF split-mass system; (e) block diagram representation including delay.

Most of the research works carried out on substructure tests considered structural integrators applied to the equations of motion second-order in time. Nonetheless, it is well known that the motion of a PS in a substructure test, see Fig.1d, is driven by a transfer system –actuator- and sensors, governed by a control unit. Since the control system is typically described by first-order Differential Equations, the utilized integrators have to deal with mixed first- and second-order ODEs. In order to solve this problem, we suggest to employ first-order integrators like the LSRT Rosenbrock algorithms, both for structural and control systems, owing to the favourable properties of LSRT algorithms employed in control [9].

With regard to complex emulated structures, numerical and control requirements impose different time steps for a NS and a PS, respectively. As a result, two main techniques can be identified to tackle this problem: i) model reduction, that represents an effective way to lower computation burdens related to the integration of a complex NS, but becomes very inaccurate especially for non-linear systems; ii) multi-time methods that allow to employ different time integrators in distinct subdomains. Furthermore, subcycling permits to use different time steps in different subdomains. The last strategy is relatively simple to implement, but it can hinder stability and accuracy properties of the original schemes. Therefore, the paper proposes some novel multi-time method with subcycling strategies, investigates relevant stability and accuracy issues and presents hardware-in-the-loop tests. Also, limits and remedies of the treated multi-time method are suggested.

The remaining part of the paper is organized as follows. Firstly, LSRT algorithms with two stages (LSRT2) are introduced, which are nowadays used in real-time substructured tests [10], as an alternative to structural integrators applied to the equations of motion second-order in time [11,12]. Then, interfield parallel algorithms with subcycling strategies based on the progenitor LSRT algorithm that represent the main focus of the paper are considered. These linearly implicit algorithms first solve the interface problem by means of Lagrange multipliers and subsequently advance the solution in all subdomains. Thus, stability and accuracy properties of these algorithms are analysed through numerical experiments on a Single-DoF split-mass system, including subcycling too. Successively, real-time tests conducted by means of a novel test rig on a two-DoF systems are presented and commented. The limits of this algorithm are underlined and improvements are introduced on a companion interfield parallel

solution procedure. Lastly, main conclusions are drawn.

2 LINEARLY IMPLICIT ROSENBROCK-BASED ALGORITHMS

In this section, we introduce the LSRT compatible algorithms developed and suggested by Bursi et al. [10]. They are linearly implicit because eliminate the need to solve non-linear systems for nonlinear problems. To employ LSRT algorithms, the equations of motion

$$\mathbf{M}\ddot{\mathbf{u}} = \mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, t) \quad (1)$$

can be rewritten into a state-space form

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t), \text{ with } \mathbf{y} = \begin{Bmatrix} \mathbf{u} \\ \dot{\mathbf{u}} \end{Bmatrix}, \mathbf{f}(\mathbf{y}, t) = \begin{Bmatrix} \dot{\mathbf{u}} \\ \mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, t) \end{Bmatrix}, \quad (2)$$

where \mathbf{M} stands for the mass matrix which is assumed to be symmetric positive definite for simplicity $\mathbf{f}(\mathbf{y}, t)$ and $\mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, t)$ for the vectors of applied and internal forces, respectively. In a FE context, the force vector can be split as $\mathbf{r}(\mathbf{u}, \dot{\mathbf{u}}, t) = -\mathbf{K}\mathbf{u} - \mathbf{C}\dot{\mathbf{u}} + \mathbf{F}_e$ with a stiffness matrix \mathbf{K} , a damping matrix \mathbf{C} and a displacement vector \mathbf{u} . Differentiation with respect to time is expressed by a dot, and thus we set $\dot{\mathbf{u}}$ and $\ddot{\mathbf{u}}$ to define the corresponding velocity and acceleration vectors.

The two-stage L-stable real-time two-stage (LSRT2) method applied to $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, t)$ reads:

$$\mathbf{k}_1 = [\mathbf{I} - \gamma\Delta t\mathbf{J}]^{-1} \mathbf{f}(\mathbf{y}_k, t_k) \Delta t, \quad \mathbf{y}_{k+\alpha_{21}} = \mathbf{y}_k + \alpha_{21}\mathbf{k}_1, \quad (3)$$

$$\mathbf{k}_2 = [\mathbf{I} - \gamma\Delta t\mathbf{J}]^{-1} \left(\mathbf{f}(\mathbf{y}_{k+\alpha_{21}}, t_{k+\alpha_2}) + \mathbf{J}\gamma_{21}\mathbf{k}_1 \right) \Delta t, \quad \mathbf{y}_{k+1} = \mathbf{y}_k + b_1\mathbf{k}_1 + b_2\mathbf{k}_2, \quad (4)$$

where Δt is the step interval and $\mathbf{J} = (\partial\mathbf{f}/\partial\mathbf{y})$ is the Jacobian matrix evaluated at the first stage. Two sets of parameters that satisfy second-order accuracy, L-stability and real-time compatibility are introduced, namely $\gamma = 1 - \sqrt{2}/2$ and $\gamma = 1 + \sqrt{2}/2$, respectively, together with $\alpha_2 = \alpha_{21} = 1/2$, $\gamma_{21} = -\gamma$, $b_1 = 0$ and $b_2 = 1$. The favourable performance of the LSRT2 method with respect to low and high-frequency components of the response can be observed from Figure 2, where a comparison with the Generalized- α [13] method is illustrated.

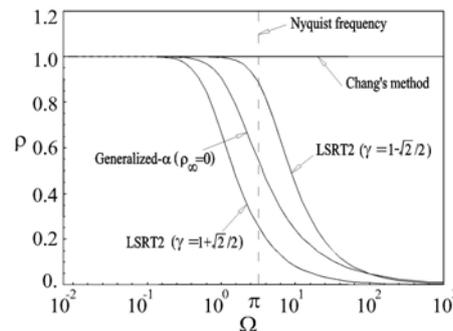


Figure 2: Spectral radii ρ of linearly implicit algorithms with respect to the Generalized- α method vs. the non-dimensional frequency Ω .

3 A PARTITIONED TIME INTEGRATION METHOD BASED ON ACCELERATION CONSTRAINT

In this section, we present a partitioned time integration method that adopts both element-based partitioning and the LSRT2 method presented in Section 2. Due to real-time compatibility, we consider the acceleration continuity at the interface of subdomains, and therefore, an explicit Lagrange multiplier formulation is obtained. Moreover, the novel partitioned method preserves favourable second-order accuracy, and very often, both unconditionally stability and high-frequency dissipation capabilities.

3.1 Derivation of an explicit Lagrange multiplier vector

In order to implement the LSRT2 method, we begin with a system of index one

$$\begin{cases} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M}^i \end{bmatrix} \begin{Bmatrix} \dot{\mathbf{u}}^i \\ \ddot{\mathbf{u}}^i \end{Bmatrix} = \begin{Bmatrix} \dot{\mathbf{u}}^i \\ \mathbf{f}(\mathbf{u}^i, \dot{\mathbf{u}}^i, t) \end{Bmatrix} + \begin{bmatrix} \mathbf{0} \\ (\mathbf{G}^i)^T \end{bmatrix} \boldsymbol{\Lambda} & i = 1, \dots, S \\ \sum_{i=1}^S \begin{bmatrix} \mathbf{0}, \mathbf{G}^i \end{bmatrix} \begin{Bmatrix} \dot{\mathbf{u}}^i \\ \ddot{\mathbf{u}}^i \end{Bmatrix} = \mathbf{0} \end{cases}, \quad (5)$$

which is modelled as a set of S non-overlapping subdomains constrained by acceleration continuity at the interface. For simplicity, we only consider the case with linear constraint equations [14]. With the assumption $\mathbf{y} = \{\mathbf{u}^T \quad \dot{\mathbf{u}}^T\}^T$, one obtains

$$\begin{cases} \mathbf{A}^i \dot{\mathbf{y}}^i = \mathbf{F}^i(\mathbf{y}^i, t) + (\mathbf{C}^i)^T \boldsymbol{\Lambda} \\ \sum_{i=1}^S \mathbf{C}^i \dot{\mathbf{y}}^i = \mathbf{0} \end{cases}, \quad (6)$$

For easiness of notation, both the matrices \mathbf{A}^i and \mathbf{C}^i refer to the i -th subdomain. Both the vector $\dot{\mathbf{y}}^i$ and the Lagrange multiplier vector $\boldsymbol{\Lambda}$ can be explicitly solved by means of (6), i.e.

$$\dot{\mathbf{y}}^i = (\mathbf{A}^i)^{-1} \mathbf{F}^i(\mathbf{y}^i, t) + (\mathbf{A}^i)^{-1} (\mathbf{C}^i)^T \boldsymbol{\Lambda} \quad (7)$$

with

$$\boldsymbol{\Lambda} = -\mathbf{H}^{-1} \sum_{i=1}^S \mathbf{C}^i (\mathbf{A}^i)^{-1} \mathbf{F}^i(\mathbf{y}^i, t) \quad (8)$$

and

$$\mathbf{H} = \sum_{i=1}^S \mathbf{C}^i (\mathbf{A}^i)^{-1} (\mathbf{C}^i)^T. \quad (9)$$

Equations (7) and (8) can be expressed in compact forms as

$$\dot{\mathbf{y}} = \mathbf{A}^{-1} \mathbf{F}(\mathbf{y}, t) + \mathbf{A}^{-1} \mathbf{C}^T \boldsymbol{\Lambda} \quad (10)$$

with

$$\Lambda = -[\mathbf{CA}^{-1}\mathbf{C}^T]^{-1}\mathbf{CA}^{-1}\mathbf{F}(\mathbf{y}, t), \quad (11)$$

where

$$\mathbf{y} = \{(\mathbf{y}^1)^T \quad \dots \quad (\mathbf{y}^S)^T\}^T, \quad \mathbf{A} = \text{Blockdiagonal}[\mathbf{A}^1 \quad \dots \quad \mathbf{A}^S] \quad (12)$$

$$\mathbf{F}(\mathbf{y}, t) = \left\{ \left[\mathbf{F}^1(\mathbf{y}^1, t) \right]^T \quad \dots \quad \left[\mathbf{F}^S(\mathbf{y}^S, t) \right]^T \right\}^T, \quad \mathbf{C} = [\mathbf{C}^1 \quad \dots \quad \mathbf{C}^S]$$

Hence, the proposed partitioned method is based on the explicit evaluation of Λ at the beginning of each time step/stage, and thus, the integration of each subdomain can independently advance.

3.2 Solution procedure of partitioned time integration methods

For simplicity, let us consider a system divided into two subdomains A and B. In the case that both subdomains be integrated by the LSRT2 algorithm with the same time step Δt , i.e. $ss=1$, the solution procedure of the partitioned method can be represented as in Figure 3, endowed with the task numbering. In detail, it can be characterized as follows:

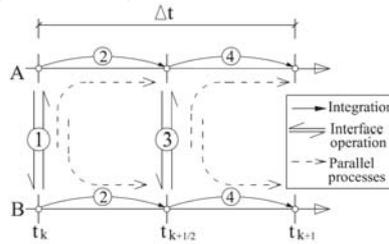


Figure 3 : The procedure of the LSRT2-based partitioned parallel method with $ss=1$

- (1) Evaluate \mathbf{F}_k^A and \mathbf{F}_k^B and calculate the Lagrange multiplier Λ_k at time t_k ,

$$\Lambda_k = -\mathbf{H}^{-1} [\mathbf{C}^A (\mathbf{A}^A)^{-1} \mathbf{F}_k^A + \mathbf{C}^B (\mathbf{A}^B)^{-1} \mathbf{F}_k^B]. \quad (13)$$

- (2) Compute \mathbf{k}_1^i where $i = A, B$ and evaluate the solutions $\mathbf{y}_{k+1/2}^i$ ($i = A, B$) -First stage-

$$\mathbf{k}_1^i = [\mathbf{I} - \gamma \Delta t \mathbf{J}^i]^{-1} (\mathbf{A}^i)^{-1} [\mathbf{F}_k^i + (\mathbf{C}^i)^{-1} \Lambda_k] \Delta t, \quad (14)$$

$$\mathbf{y}_{k+1/2}^i = \mathbf{y}_k^i + \frac{1}{2} \mathbf{k}_1^i. \quad (15)$$

- (3) Evaluate $\mathbf{F}_{k+1/2}^A$ and $\mathbf{F}_{k+1/2}^B$ and calculate $\Lambda_{k+1/2}$ at time $t_{k+1/2}$,

$$\Lambda_{k+1/2} = -\mathbf{H}^{-1} [\mathbf{C}^A (\mathbf{A}^A)^{-1} \mathbf{F}_{k+1/2}^A + \mathbf{C}^B (\mathbf{A}^B)^{-1} \mathbf{F}_{k+1/2}^B]. \quad (16)$$

- (4) Compute \mathbf{k}_2^i and advance the solution to \mathbf{y}_{k+1}^i in both subdomains, respectively, -Second

stage-

$$\mathbf{k}_2^i = [\mathbf{I} - \gamma \Delta t \mathbf{J}^i]^{-1} (\mathbf{A}^i)^{-1} [\mathbf{F}_{k+1/2}^i + (\mathbf{C}^i)^{-1} \boldsymbol{\Lambda}_{k+1/2} - \gamma \mathbf{J}^i \mathbf{k}_1^i] \Delta t, \quad (17)$$

$$\mathbf{y}_{k+1}^i = \mathbf{y}_k^i + \mathbf{k}_2^i. \quad (18)$$

Since the LSRT2 method contains two stages, inter-domain exchange of information is not only required at the beginning of each time step, but also at the beginning of the second stage. This algorithm can be defined as parallel, because the interconnection is only done at the beginning of each stage to compute the Lagrange multiplier; then each subdomain can independently advance in each stage.

4 AN INTERFIELD PARALLEL SOLUTION PROCEDURE

Along the line of [6], the Rosenbrock-based method developed above can be exploited to develop an interfield parallel procedure. To illustrate that, let us consider again two subdomains A and B . The solution procedure is highlighted in Figure 4 with the numbering of the two processes and the subscript i referred to the time step Δt . In detail, Subdomain A is integrated with the coarse time step $\Delta t_A = 4 \cdot \Delta t$, while Subdomain B with the fine time step $\Delta t_B = \Delta t/ss$, where $ss = 2$. The solution procedure for Subdomain A is as follows:

- (1) Evaluate \mathbf{F}_{i-2}^A and \mathbf{F}_{i-2}^B with the solutions \mathbf{y}_{i-2}^A and \mathbf{y}_{i-2}^B and then calculate the Lagrange multiplier $\boldsymbol{\Lambda}_{i-2}$,

$$\boldsymbol{\Lambda}_{i-2} = -\mathbf{H}^{-1} \left[\mathbf{C}^A (\mathbf{A}^A)^{-1} \mathbf{F}_{i-2}^A + \mathbf{C}^B (\mathbf{A}^B)^{-1} \mathbf{F}_{i-2}^B \right]. \quad (19)$$

- (2) Compute \mathbf{k}_1^A and advance the solution to \mathbf{y}_i^A ,

$$\mathbf{k}_1^A = [\mathbf{I} - 4\Delta t \gamma \mathbf{J}^A]^{-1} (\mathbf{A}^A)^{-1} \left[\mathbf{F}_{i-2}^A + (\mathbf{C}^A)^T \boldsymbol{\Lambda}_{i-2} \right] 4\Delta t, \quad (20)$$

$$\mathbf{y}_i^A = \mathbf{y}_{i-2}^A + \frac{1}{2} \mathbf{k}_1^A. \quad (21)$$

- (3) Evaluate \mathbf{F}_i^A and \mathbf{F}_i^B and then calculate $\boldsymbol{\Lambda}_i$,

$$\boldsymbol{\Lambda}_i = -\mathbf{H}^{-1} \left[\mathbf{C}^A (\mathbf{A}^A)^{-1} \mathbf{F}_i^A + \mathbf{C}^B (\mathbf{A}^B)^{-1} \mathbf{F}_i^B \right]. \quad (22)$$

- (4) Evaluate \mathbf{k}_2^A and advance the solution to \mathbf{y}_{i+2}^A ,

$$\mathbf{k}_2^A = [\mathbf{I} - 4\Delta t \gamma \mathbf{J}^A]^{-1} \left\{ (\mathbf{A}^A)^{-1} \left[\mathbf{F}_i^A + (\mathbf{C}^A)^T \boldsymbol{\Lambda}_i \right] - \gamma \mathbf{J}^A \mathbf{k}_1^A \right\} 4\Delta t, \quad (23)$$

$$\mathbf{y}_{i+2}^A = \mathbf{y}_i^A + \mathbf{k}_2^A. \quad (24)$$

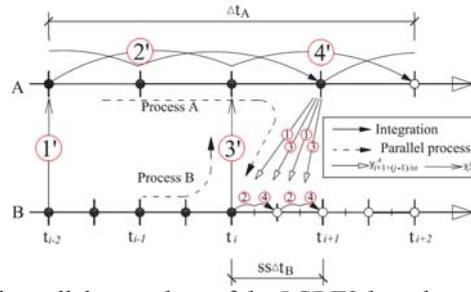


Figure 4 The interfield parallel procedure of the LSRT2-based partitioned method with $ss=2$.

(5) Calculate $\mathbf{y}_{i+1+\frac{in}{2ss}}^A$ by means of linear interpolation:

$$\mathbf{y}_{i+1+\frac{in}{2ss}}^A = \left(1 - \frac{in}{2ss}\right) \mathbf{y}_{i+1}^A + \frac{in}{2ss} \mathbf{y}_{i+2}^A, \quad (in = 1, 2, \dots, 2ss). \quad (25)$$

At the same time, the advancement procedure for $(j=1, \dots, s)$ substep in subdomain B , i.e. from \mathbf{y}_i^B to \mathbf{y}_{i+1}^B reads:

(1) Evaluate $\mathbf{F}_{i+\frac{j-1}{ss}}^A$ and $\mathbf{F}_{i+\frac{j-1}{ss}}^B$ and calculate $\Lambda_{i+\frac{j-1}{ss}}$,

$$\Lambda_{i+\frac{j-1}{ss}} = -\mathbf{H}^{-1} \left[\mathbf{C}^A (\mathbf{A}^A)^{-1} \mathbf{F}_{i+\frac{j-1}{ss}}^A + \mathbf{C}^B (\mathbf{A}^B)^{-1} \mathbf{F}_{i+\frac{j-1}{ss}}^B \right]. \quad (26)$$

(2) Calculate \mathbf{k}_1^B and advance the solution to $\mathbf{y}_{i+\frac{2j-1}{2ss}}^B$,

$$\mathbf{k}_1^B = \left[\mathbf{I} - \frac{\Delta t}{ss} \gamma \mathbf{J}^B \right]^{-1} (\mathbf{A}^B)^{-1} \left[\mathbf{F}_{i+\frac{j-1}{ss}}^B + (\mathbf{C}^B)^T \Lambda_{i+\frac{j-1}{ss}} \right] \frac{\Delta t}{ss}, \quad (27)$$

$$\mathbf{y}_{i+\frac{2j-1}{ss}}^B = \mathbf{y}_{i+\frac{j-1}{ss}}^B + \frac{1}{2} \mathbf{k}_1^B. \quad (28)$$

(3) Evaluate $\mathbf{F}_{i+\frac{2j-1}{2ss}}^A$ and $\mathbf{F}_{i+\frac{2j-1}{2ss}}^B$ and calculate $\Lambda_{i+\frac{2j-1}{2ss}}$,

$$\Lambda_{i+\frac{2j-1}{2ss}} = -\mathbf{H}^{-1} \left[\mathbf{C}^A (\mathbf{A}^A)^{-1} \mathbf{F}_{i+\frac{2j-1}{2ss}}^A + \mathbf{C}^B (\mathbf{A}^B)^{-1} \mathbf{F}_{i+\frac{2j-1}{2ss}}^B \right]. \quad (29)$$

(4) Calculate \mathbf{k}_2^B and advance the solution to $\mathbf{y}_{i+\frac{j+1}{ss}}^B$,

$$\mathbf{k}_2^B = \left[\mathbf{I} - \frac{\Delta t}{ss} \gamma \mathbf{J}^B \right]^{-1} \left\{ (\mathbf{A}^B)^{-1} \left[\mathbf{F}_{i+\frac{2j-1}{2ss}}^B + (\mathbf{C}^B)^T \Lambda_{i+\frac{2j-1}{2ss}} \right] - \gamma \mathbf{J}^B \mathbf{k}_1^B \right\} \frac{\Delta t}{ss}, \quad (30)$$

$$\mathbf{y}_{i+\frac{j}{ss}}^B = \mathbf{y}_{i+\frac{j-1}{ss}}^B + \mathbf{k}_2^B. \quad (31)$$

The method is not self-starting and to preserve second-order accuracy and parallel characteristics, we have chosen the LSRT2-based partitioned method with no subcycling to initiate the procedure.

5 NUMERICAL SIMULATIONS AND HARDWARE-IN-THE-LOOP TESTS

In order to examine numerical properties of the newly-developed method, spectral stability and convergence are analysed on a Single-Dof split mass system. Additionally, results provided with Hardware-in-the-loop tests on a two-Dof system are presented as well.

5.1 Stability and convergence analysis on a Single-Dof split mass system

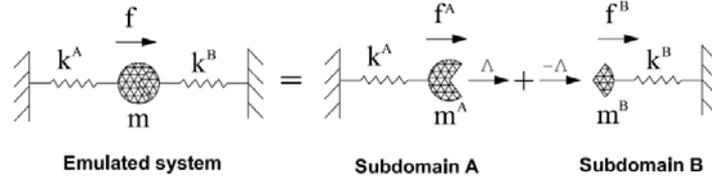


Figure 5: A Single-Dof emulated system with the relevant split-mass system.

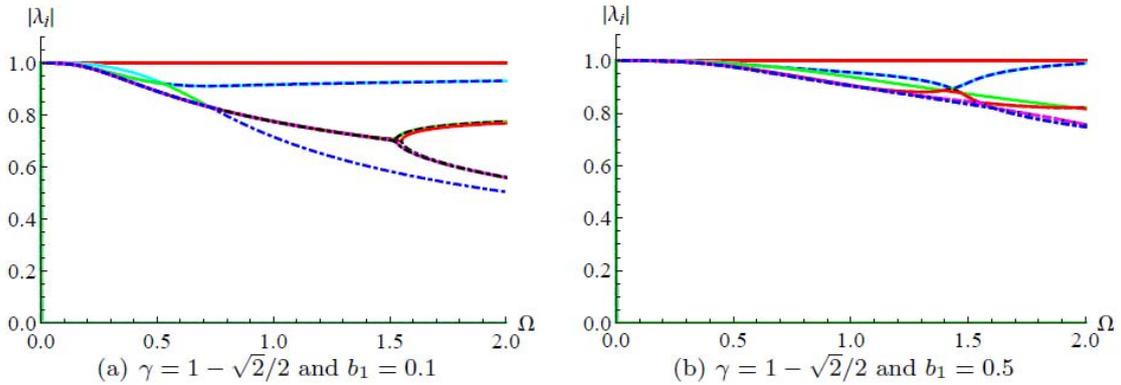


Figure 6: $|\lambda_j|$ of the SDoF problem for the interfield parallel procedure with $ss=10$ and (a) $b_1=0.1$, $\gamma = 1 - \sqrt{2}/2$; (b) $b_1=0.5$, $\gamma = 1 - \sqrt{2}/2$.

We consider the test problem depicted in Figure 5. For simplicity, we choose the following system variables $m = m_A + m_B = 1$ and $k = k_A + k_B = 1$; and b_1 defined as

$$b_1 = \frac{m_A}{m_B} = \frac{k_B}{k_A}. \quad (32)$$

In order to highlight the numerical dissipation properties of the method on the solution, we consider no physical damping. Moreover, no external force is involved as well. The application of the integration methods to this model problem leads to the recursive formula

$$\mathbf{y}_{k+1} = \mathbf{R}\mathbf{y}_k + \mathbf{L}_k, \quad (33)$$

where \mathbf{R} is the amplification matrix and \mathbf{L} is the load vector that depends on external forces, respectively. The spectral stability of the method is analysed through the spectral eigenvalues $|\lambda_j|$ of \mathbf{R} .

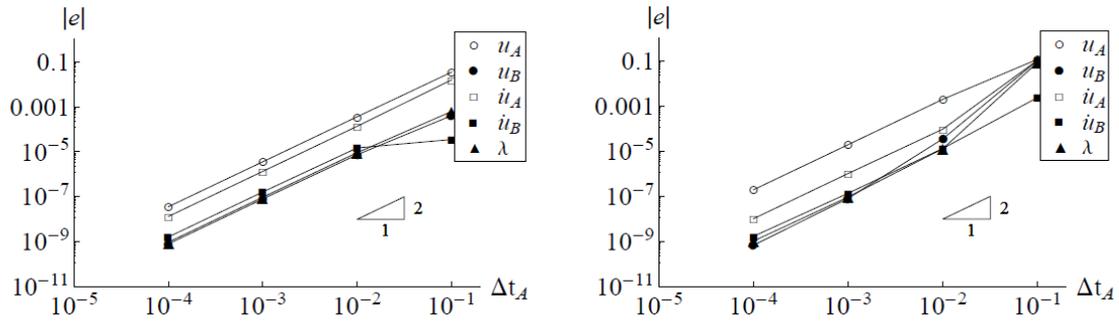


Figure 7: Global errors of the inter-field parallel method with $ss=10$ and (a) $\gamma = 1 - \sqrt{2}/2$;
(b) $\gamma = 1 + \sqrt{2}/2$

The parallel method is not self-starting and therefore the choice of proper state variables for the stability analysis represents the main difficulty. The initial solutions are composed of $\mathbf{y}_{k-2}^A, \mathbf{y}_{k-2}^B, \mathbf{y}_k^A, \mathbf{y}_k^B$ and \mathbf{y}_{k+1}^A , while the output contains only \mathbf{y}_{k+1}^B and \mathbf{y}_{k+2}^A . Moreover, another pair of intermediate solutions $\mathbf{y}_{k-1}^A, \mathbf{y}_{k-1}^B$ are needed. As a result, the state vector involved in the spectral stability analysis reads

$$\mathbf{X}_k = \left[\left(\mathbf{y}_{k-2}^A \right)^T \quad \left(\mathbf{y}_{k-2}^B \right)^T \quad \left(\mathbf{y}_{k-1}^A \right)^T \quad \left(\mathbf{y}_{k-1}^B \right)^T \quad \left(\mathbf{y}_k^A \right)^T \quad \left(\mathbf{y}_k^B \right)^T \quad \left(\mathbf{y}_{k+1}^A \right)^T \right]^T. \quad (34)$$

Consequently \mathbf{X}_k has a dimension $8n_A + 6n_B$, with n_A and n_B the DoFs of the two subdomains, respectively. For the single-DoF split-mass system, the dimension is 14. $|\lambda_i|$ relevant to the model problem integrated with the algorithm vs. the numerical frequency $\Omega = \sqrt{k/m} \Delta t_A$ are plotted in Figure 5. The number of nonzero eigenvalues is found to be 10: one is unitary, four pairs of them are complex conjugate and the other one is frequently less than 1. Besides, four zero eigenvalues are included. The method with $\gamma = 1 + \sqrt{2}/2$ exhibits unconditional stability, while the one characterized by $\gamma = 1 - \sqrt{2}/2$ sometimes is only conditionally stable (see Figure 6). Moreover, the stability of the parallel method depends on the parameter b_1 defined in (32). Larger values of b_1 introduce more damping and render the method more stable. For more information, readers are referred to [8].

Successively, the global error is analysed on the SDoF split mass system with the initial conditions $d(t_0)=1$ and $v(t_0)=1$. Figure 7 shows that the order of convergence of state variables is in agreement with the theoretical analysis and hence the method exhibits second-order accuracy [8].

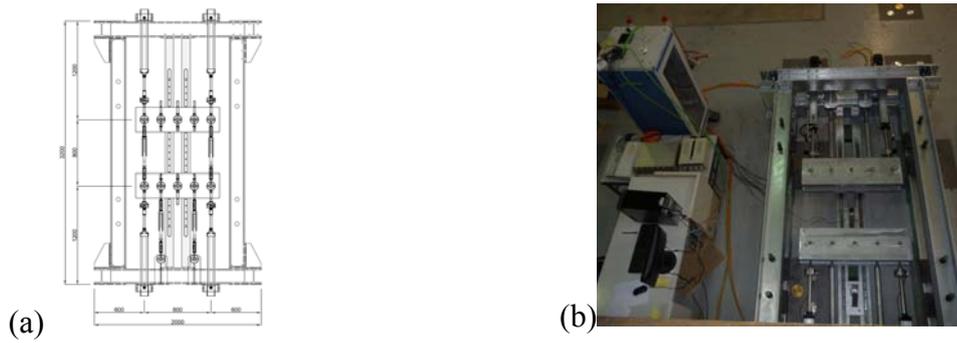


Figure 8: The test rig : (a) drawing; (b) photo.

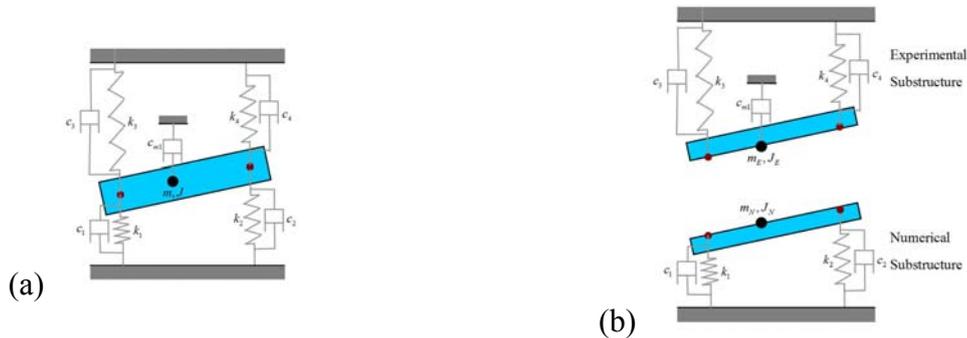


Figure 9: The simulation model: (a) 2Dof emulated structure; (b) the split system.

5.2 Tests on Two-Dof split-mass system

Table 1: Characteristics of both emulated and split subdomains in Hardware-in-the-loop tests

Struct. types Properties	Emulated system			Numerical substructure			Physical substructure		
	M	K	C	M _N	K _N	C _N	M _P	K _P	C _P
Translational	2210.9	346310	555.66	1658.2	306640	555.66	552.7	39670	0
Rotational	157.2	138524	22.226	117.9	12265	22.226	39.3	1711	0

Note: all valuables are in International Units.

In order to validate the effectiveness of the proposed methods in Hardware-in-the-loop tests, a versatile system was conceived and installed at the University of Trento, Italy. It consists of four actuators, one dSpace DS1103 control board and other high performance devices, shown in Figure 8. This section briefly describes the application of the new parallel method on the 2Dof split mass system, illustrated in Figure 9.

The system characteristics are collected in Table 1. In view of the different sampling times in different subdomains, multitasking techniques are exploited to make most of the advantages of the algorithm, such as multiple rates and parallel implementation. In the test, we selected $\Delta t_A = 4 \cdot \Delta t = 16\text{ms}$ and $\Delta t_B = \Delta t / ss = 2\text{ms}$. Additionally, the system delay of about 20ms was compensated for by means of a polynomial delay compensation scheme [14].

Test results compared with reference numerical simulations are presented in Figure 10. Both displacements fit well to the simulated ones considering the fact that friction forces existing in the system were not modelled. In addition, smaller limited drifts between displacements relevant to both the numerical and the physical substructure were observed.

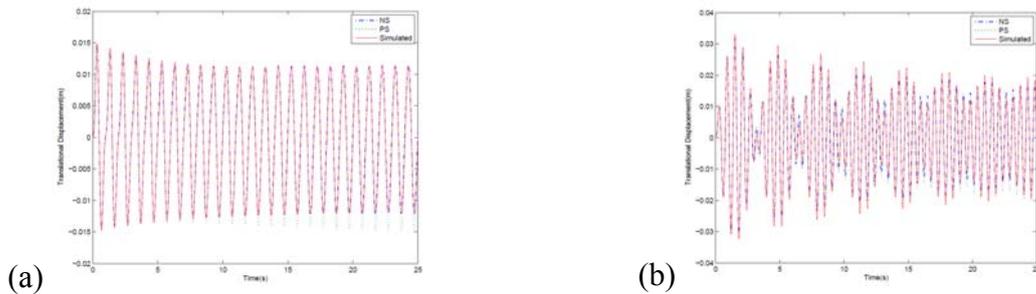


Figure 10: Test results: (a) translational displacement time histories; (b) rotational time histories.

6 AN IMPROVED PARALLEL SOLUTION PROCEDURE

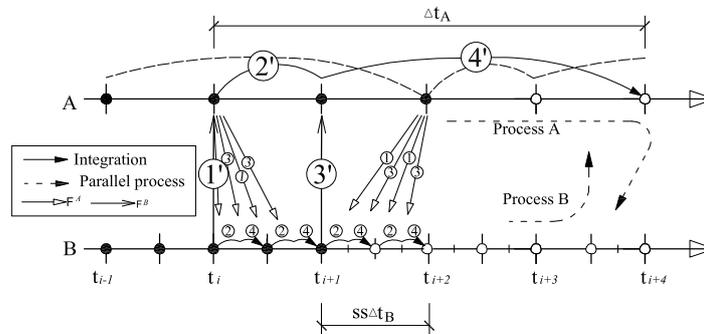


Figure 11: The solution procedure of the improved parallel algorithm

The interfield parallel method presented in Section 4 and tested in Section 5 is appealing, because of its flexibility of dealing with different substructure requirements. Unfortunately, the drift-off effects may limit its applications. Additionally, four parallel integration processes are required for Subdomain A. Based on this insight, the integration method was simplified by conducting the integration in Subdomain A with different stage sizes. Its characteristics can be observed in Figure 11. In addition, displacement drift observed in the progenitor algorithm was reduced via velocity projection at the end of each step [15]. The complete solution procedure of the improved method is also illustrated in Figure 11. Convergence analyses and applications to real-time Hardware-in-the-loop tests will be presented elsewhere.

7 CONCLUSIONS

Initially in this paper, we introduced and applied linearly implicit L-stable Rosenbrock methods with two-stages to real time hardware-in-the-loop substructure tests. The methods are endowed with several favourable characteristics, among which real-time compatibility, explicit evaluation of state variables and user-defined high-frequency dissipation capabilities. In detail and in view of hybrid testing of complex emulated structures, we developed and

illustrated a novel interfield parallel partitioned algorithm based on the progenitor Rosenbrock method, that can incorporate subcycling. Through spectral analysis and numerical simulations on an SDoF split-mass system, both stability and accuracy properties were shown. In a greater detail, the partitioned algorithm preserved second-order accuracy as the progenitor monolithic method and favourable stability properties. Moreover a novel test rig conceived to perform both linear and nonlinear substructure tests was introduced, and tests on a two-DoF split-mass system were illustrated. The drawbacks of this algorithm were commented and improvements were implemented on a companion solution procedure. Finally, these algorithms will allow an in-depth study of errors and control strategies of actuators.

REFERENCES

- [1] Saouma V. and Sivaselvan M.V. (Editors) 2008. *Hybrid Simulation - Theory, implementation and applications*, Taylor&Francis, London (2008).
- [2] Bursi, O. S. and Wagg, D. J. (Editors). *Modern testing techniques for structural systems –Dynamics and control*, Springer, Wien, New York (2008).
- [3] Jung, R. Y., Shing, P. B., Stauffer, E. and Thoen, B. Performance of a real-time pseudodynamic test system considering nonlinear structural response. *Earth. Engng. Struct. Dyn.* (2007), **36**:1785-1809.
- [4] Prakash, A. and Hjelmstad, K. D. 2004. A FETI-based multitime-step coupling method for Newmark schemes in structural dynamics. *Int. J. Num. Meth. Engng* (2004) **61**:2183-2204.
- [5] Bonelli, A., Bursi, O. S., He, L., Magonette, G. and Pegon, P. Convergence analysis of a parallel interfield method for heterogeneous simulations with substructuring. *Int. J. Num. Meth. Engng* (2008) **75**:800-825.
- [6] Bursi O.S., He L., Bonelli, A., Pegon, P., Novel generalized-alpha methods for interfield parallel integration of heterogeneous structural dynamic systems, *J. Comp. Appl. Math*, (2010), **234**:2250-2258.
- [7] Nakshatrala, K. B., Hjelmstad, K. D. and Tortorelli, D. A. 2008. A FETI-based domain decomposition technique for time-dependent first-order systems based on a DAE approach. *Int. J. Num. Meth. Engng* (2008) **75**:1385-1415.
- [8] Jia C., Bursi, O.S., Bonelli A. and Wang Z., Novel partitioned integration methods for DAE systems based on L-stable linearly implicit algorithms, *Int. J. Num. Meth. Engng* (2011), in print.
- [9] Bursi O.S., Stoten D.P., Tondini N., Vulcan L., Stability and accuracy analysis of a discrete model reference adaptive controller without and with time delay, *Int. J. Num. Meth. Engng* (2010) **82**:1158-1179.
- [10] Bursi, O.S., Jia C., Vulcan, L., Neild S.A. and Wagg, D. J., Rosenbrock-based algorithms and subcycling strategies for real-time nonlinear substructure testing, *Earth. Engng. Struct. Dyn.* (2011), **40**:1-19.
- [11] Chang S. Y. Explicit pseudodynamic algorithm with unconditional stability. *J. Engng. Mech.* (2002), **128**:935-947.
- [12] Bursi, O.S. He L., Lamarche C.P. and Bonelli A., Linearly implicit time integration methods for real-time dynamic substructure testing, *J. Engng. Mech.* (2010), **136**:1380-1389.
- [13] Chung, J. and Hulbert, G.M. 1993. A time integration algorithm for structural dynamics with improved numerical dissipation: the generalized-alpha method. *Journal of Applied Mechanics* **60**:371-375.
- [14] Lamarche CP. Development of real-time dynamic substructuring procedures for the seismic testing of steel structure. PhD thesis, Universite de Montreal and University of Trento, Italy, 2009.
- [15] Bauchau O. A. and Laulusa A. Review of Contemporary Approaches for Constraint Enforcement in Multibody Systems. *J. Comput. Nonlinear Dynam.* **3**: 01-10 (2008).

PERIODIC SOLUTIONS FOR ACOUSTO-ELASTIC SCATTERING PROBLEMS BY CONTROLLING THE TIME-DEPENDENT EQUATIONS

Sanna Mönkölä*, Sami Kähkönen

Department of Mathematical Information Technology, University of Jyväskylä,
P.O. Box 35 (Agora), FI-40014 University of Jyväskylä, Jyväskylä, Finland,
e-mail: {sanna.monkola, sami.h.kahkonen}@jyu.fi

Key words: Exact Controllability, Fluid-Structure Interaction, Spectral Element Method.

Abstract. The use of controllability methods for solving time-harmonic acousto-elastic wave equations is considered. In particular, the study concentrates on the coupling between two time-harmonic linear wave equations: the scalar valued Helmholtz equation concerning the propagation of acoustic waves and the vector valued Navier equation modeling the propagation of waves in an elastic medium. These fundamental equations occur in a number of physical applications such as acoustics, medicine, and geophysics. In this work, we concentrate on scattering problems but the same method can be used for other types of problems as well.

1 INTRODUCTION

Analytical solutions for fluid-structure interaction problems exist only for some simple geometries like sphere. Thus, a wide range of numerical methods have been used for solving time-harmonic wave equations. The traditional way of using complex-valued time-harmonic equations leads to large-scale sparse indefinite systems. Direct solvers, like LU or Cholesky factorization, require a significant amount of computing time and memory storage. Thus, despite the robustness and the fixed number of computing steps, they are not practical for solving very large problems. Developing iterative methods, such as GMRES or Bi-CGSTAB, to be more efficient, is a challenging task. Efficient preconditioners are needed to guarantee proper convergence speed for the iterative methods.

We focus on finding periodic solutions without solving the time-harmonic problems directly. Thus, we turn to time-dependent acousto-elastic equations (see, e.g. [1, 2, 3]). It is known that time-dependent equations can be simulated with respect to time until the time-harmonic solution is reached, but the approach suffers from poor convergence. We accelerate the convergence rate by employing the exact controllability method. The original time-harmonic problem is formulated as an equivalent least-squares control problem

for the time-dependent wave equation. The control problem is then solved by a conjugate gradient algorithm. The idea of the control approach was first suggested and developed for the acoustic and electromagnetic wave equations by Bristeau, Glowinski, and Périaux (see, e.g., [4, 5]). Recently, we have applied different discretization schemes to further improve the method (see, e.g., [6, 7, 8]).

Various formulations exist for the fluid-structure interaction between acoustic and elastic waves. Expressing the acoustic wave equation by the pressure in the fluid domain leads to a non-symmetric formulation (see, e.g., [9, 10, 11, 12]), while using the velocity potential results in a symmetric system of equations (see, e.g., [1, 3, 13, 2]). Finding a proper least-squares functional for the minimization problem of the non-symmetric formulation is not a straightforward task in the two-dimensional domain. That is why we use the velocity potential approach in the fluid domain and minimize the functional related to the natural energy norm of the symmetric formulation.

2 MATHEMATICAL MODEL

The computational domain $\Omega \subset \mathbb{R}^2$ is divided into the solid part Ω_s and the fluid part Ω_f by the interface Γ_i such that the structure is surrounded by the fluid (see Figure 1). Instead of solving directly the time-harmonic equation, we return to the corresponding time-dependent equation

$$\frac{1}{c(\mathbf{x})^2} \frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi = 0, \quad \text{in } \Omega_f \times [0, T], \quad (1)$$

$$\frac{1}{c(\mathbf{x})} \frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial \mathbf{n}_f} = y_{\phi \text{ext}}, \quad \text{on } \Gamma_{\text{ef}} \times [0, T], \quad (2)$$

$$\frac{\partial \mathbf{u}_s}{\partial t} \cdot \mathbf{n}_s + \frac{\partial \phi}{\partial \mathbf{n}_f} = 0, \quad \text{on } \Gamma_i \times [0, T], \quad (3)$$

$$\rho_s(\mathbf{x}) \frac{\partial^2 \mathbf{u}_s}{\partial t^2} - \nabla \cdot \sigma(\mathbf{u}_s) = \mathbf{0}, \quad \text{in } \Omega_s \times [0, T], \quad (4)$$

$$\sigma(\mathbf{u}_s) \mathbf{n}_s + \rho_f(\mathbf{x}) \frac{\partial \phi}{\partial t} \mathbf{n}_f = \mathbf{0}, \quad \text{on } \Gamma_i \times [0, T], \quad (5)$$

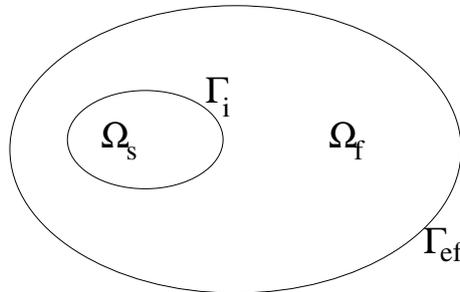


Figure 1: The domain Ω is divided into the solid part Ω_s and the fluid part Ω_f .

where length of the time interval is marked as T , ϕ denotes the velocity potential and \mathbf{u}_s is the displacement field. Coefficients $\rho_f(\mathbf{x})$ and $\rho_s(\mathbf{x})$ represent the densities of media in domains Ω_f and Ω_s , respectively. The speed of sound in the fluid domain is expressed as $c(\mathbf{x})$. The linearized stress tensor $\sigma(\mathbf{u}_s)$ depends on the Lamé parameters λ_s and μ_s . The outward normal vectors to domains Ω_f and Ω_s are marked as $\mathbf{n}_f = (n_{f1}, n_{f2})^T$ and $\mathbf{n}_s = (n_{s1}, n_{s2})^T$. The boundary Γ_{ef} truncates the fluid domain Ω_f (see, e.g., [14, 15]). In addition to the system (1)-(5), we take into account the initial conditions $\mathbf{e}_0 = (e_{\phi f0}, e_{s0})^T$ and $\mathbf{e}_1 = (e_{\phi f1}, e_{s1})^T$ such that

$$\phi(\mathbf{x}, 0) = e_{\phi f0}, \quad \frac{\partial \phi}{\partial t}(\mathbf{x}, 0) = e_{\phi f1}, \quad \text{in } \Omega_f, \quad (6)$$

$$\mathbf{u}_s(\mathbf{x}, 0) = e_{s0}, \quad \frac{\partial \mathbf{u}_s}{\partial t}(\mathbf{x}, 0) = e_{s1}, \quad \text{in } \Omega_s. \quad (7)$$

3 DISCRETIZATION

Spatial discretization is done with spectral element method [16, 10]. It allows convenient treatment of complex geometries and varying material properties. The basis functions are higher order Lagrange interpolation polynomials, and the nodes of these functions are placed at Gauss-Lobatto collocation points. The integrals in the weak form of the equation are evaluated with the corresponding Gauss-Lobatto quadrature formulas.

We denote by \mathbf{u} the global block vector containing the values of the displacement $\mathbf{u}_s(\mathbf{x}, t)$ and velocity potential $\phi(\mathbf{x}, t)$ at time t at the Gauss-Lobatto points of the quadrilateral mesh. Then, the semi-discretized coupled problem can be rewritten in the matrix form

$$\mathcal{M} \frac{\partial^2 \mathbf{u}}{\partial t^2} + \mathcal{S} \frac{\partial \mathbf{u}}{\partial t} + \mathcal{K} \mathbf{u} = \mathcal{F}, \quad (8)$$

where \mathcal{M} is the mass matrix, \mathcal{S} the matrix arising from absorbing boundary conditions and the coupling between acoustic and elastic wave equations, \mathcal{K} is the stiffness matrix, and \mathcal{F} is the vector due to the function $y_{\phi_{\text{ext}}}$. The time discretization is realized with central finite differences, and SuperLU is used for direct solution of the linear system at each timestep.

4 CONJUGATE GRADIENT ALGORITHM

After discretization, exact controllability problem is reformulated as a least-squares optimization problem,

$$\min \left(\frac{1}{2} (\mathbf{u}(T) - \mathbf{e}_0)^T \mathcal{K} (\mathbf{u}(T) - \mathbf{e}_0) + \frac{1}{2} \left(\frac{\partial \mathbf{u}(T)}{\partial t} - \mathbf{e}_1 \right)^T \mathcal{M} \left(\frac{\partial \mathbf{u}(T)}{\partial t} - \mathbf{e}_1 \right) \right), \quad (9)$$

which is solved with a preconditioned conjugate gradient algorithm. Each conjugate gradient iteration requires computation of the gradient of the discretized least-squares

functional, which involves the solution of the state equation and the corresponding adjoint equation, solution of a linear system with the block-diagonal preconditioner,

$$\begin{pmatrix} \mathcal{K} & 0 \\ 0 & \mathcal{M} \end{pmatrix}, \quad (10)$$

and some matrix-vector operations. Computation of the gradient of the functional is an essential point of the method, and we have done it with the adjoint state technique. To guarantee the smooth initial approximation for the algorithm we use a transition procedure suggested by Mur in [17]. The graph-based multigrid method [18], smoothed by successive over relaxation (SOR) with over-relaxation parameter 1.2, is used for preconditioning the conjugate gradient algorithm.

5 NUMERICAL TESTS

Firstly, we validate the solver. We consider scattering by an elastic circle Ω_s , having radius $a_s = 1$, embedded in the center of an acoustic domain Ω_f which is truncated by a circular absorbing boundary of radius $a_f = 2$ (see Figure 2). On the absorbing boundary we set the incident plane wave $\phi_{\text{inc}}(\mathbf{x}, t) = \cos(\kappa x_1 - \omega t)$ implying $y_{\phi_{\text{ext}}} = \kappa(1 - n_{f1}) \sin(\kappa x_1 - \omega t)$.

Thus, the solutions of the control algorithm are corresponding to the solutions of the time-harmonic problem,

$$-\kappa(\mathbf{x})^2 \Phi - \nabla^2 \Phi = 0, \quad \text{in } \Omega_f, \quad (11)$$

$$-i\kappa(\mathbf{x})\Phi + \frac{\partial \Phi}{\partial \mathbf{n}_f} = -i\kappa(1 - n_{f1})e^{i\kappa x_1}, \quad \text{on } \Gamma_{\text{ef}}, \quad (12)$$

$$i\omega \mathbf{U}_s \cdot \mathbf{n}_s + \frac{\partial \Phi}{\partial \mathbf{n}_f} = 0, \quad \text{on } \Gamma_i, \quad (13)$$

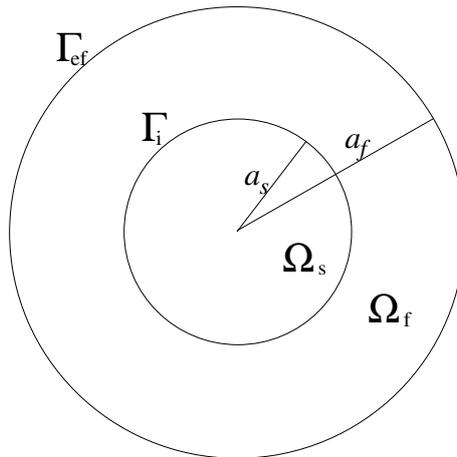


Figure 2: The elastic circle Ω_s surrounded by the acoustic media Ω_f .

$$-\omega^2 \rho_s(\mathbf{x}) \mathbf{U}_s - \nabla \cdot \sigma(\mathbf{U}_s) = \mathbf{0}, \quad \text{in } \Omega_s, \quad (14)$$

$$\sigma(\mathbf{U}_s) \mathbf{n}_s + i\omega \rho_f(\mathbf{x}) \Phi \mathbf{n}_f = \mathbf{0}, \quad \text{on } \Gamma_i, \quad (15)$$

where i is the imaginary unit, ω is the angular frequency, and $\kappa = \omega/c$ is the wavenumber describing how many waves there is in the fluid domain per a 2π unit. Respectively we can define wavenumbers in the solid domain $\kappa_p = \omega/c_p$ and $\kappa_s = \omega/c_s$, where c_p and c_s represent the speed of the pressure waves (P-waves) and the speed of the shear waves (S-waves). The relations between time-dependent and time-harmonic variables are $\phi(\mathbf{x}, t) = \text{Re}(\Phi(\mathbf{x})e^{-i\omega t})$ and $\mathbf{u}_s(\mathbf{x}, t) = \text{Re}(\mathbf{U}_s(\mathbf{x})e^{-i\omega t})$.

The solution of the time-harmonic problem can be presented in polar coordinates (r, α) as a Fourier series form (see, e.g. [19])

$$\Phi_r(r, \alpha) = e^{i\kappa r \cos \alpha} + \sum_{n=0}^{\infty} \left(A_n H_n^{(1)}(\kappa r) + B_n H_n^{(2)}(\kappa r) \right) \cos(n\alpha), \quad (16)$$

$$\mathbf{U}_{sr}(r, \alpha) = \begin{pmatrix} \sum_{n=0}^{\infty} \left(C_n \left(J_{n-1}(\kappa_p r) - \frac{n}{\kappa_p r} J_n(\kappa_p r) \right) \kappa_p + D_n \frac{n}{r} J_n(\kappa_s r) \right) \cos(n\alpha) \\ \sum_{n=0}^{\infty} \left(-C_n \frac{n}{r} J_n(\kappa_p r) - D_n \left(J_{n-1}(\kappa_s r) - \frac{n}{\kappa_s r} J_n(\kappa_s r) \right) \kappa_s \right) \sin(n\alpha) \end{pmatrix}, \quad (17)$$

where $e^{i\kappa r \cos \alpha} = J_0(\kappa r) + \sum_{n=1}^{\infty} 2i^n J_n(\kappa r) \cos(n\alpha)$ is the incident plane wave, J_n represents the Bessel functions and $H_n^{(1)}$ and $H_n^{(2)}$ are the Hankel functions. The coefficients A_n , B_n , C_n , and D_n can be solved from the linear system

$$\begin{aligned} H_n^{(1)}(\kappa a_s) A_n + H_n^{(2)}(\kappa a_s) B_n - \frac{i\omega}{a_s^2} (\kappa_p a_s J_{n-1}(\kappa_p a_s) - n J_n(\kappa_p a_s)) C_n - \frac{i\omega}{a_s^2} n J_n(\kappa_s a_s) D_n \\ = \begin{cases} -i^n J_n'(\kappa a_s), & n = 0, \\ -2i^n J_n'(\kappa a_s), & n = 1, \dots, \infty, \end{cases} \end{aligned}$$

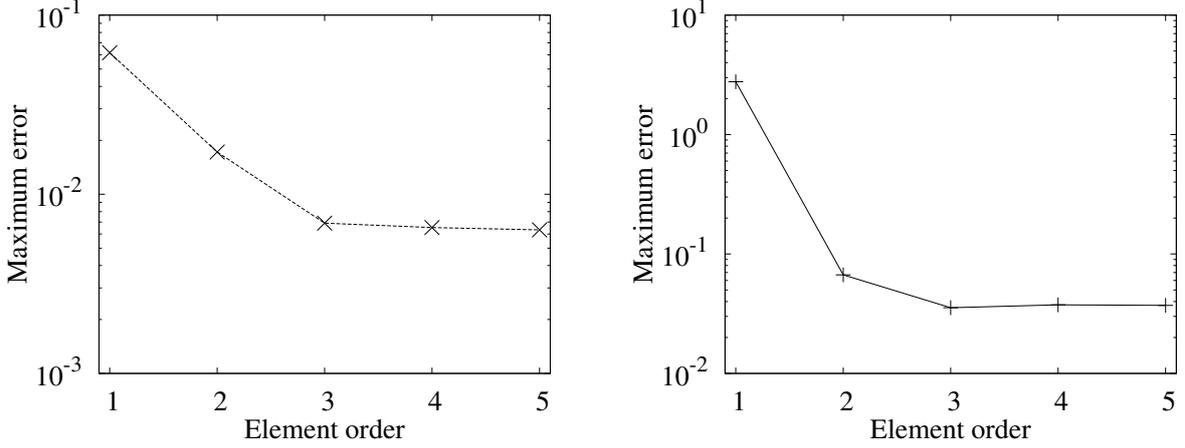
$$H_n^{(1)}(\kappa a_f) - i\kappa H_n^{(1)}(\kappa a_f) A_n + H_n^{(2)}(\kappa a_f) - i\kappa H_n^{(2)}(\kappa a_f) B_n = 0,$$

$$\begin{aligned} -i\omega \rho_f H_n^{(1)}(\kappa a_s) A_n - i\omega \rho_f H_n^{(2)}(\kappa a_s) B_n \\ + \frac{2\mu_s}{a_s^2} \left(\left(n^2 + n - \frac{\kappa_s^2 a_s^2}{2} \right) J_n(\kappa_p a_s) - \kappa_p a_s J_{n-1}(\kappa_p a_s) \right) C_n \\ + \frac{2\mu_s}{a_s^2} n \left((-n-1) J_n(\kappa_s a_s) + \kappa_s a_s J_{n-1}(\kappa_s a_s) \right) D_n = \begin{cases} \omega \rho_f i^{n+1} J_n(\kappa a_s), & n = 0, \\ 2\omega \rho_f i^{n+1} J_n(\kappa a_s), & n = 1, \dots, \infty, \end{cases} \\ -n \left((-n-1) J_n(\kappa_p a_s) + \kappa_p a_s J_{n-1}(\kappa_p a_s) \right) C_n \\ - \left(n^2 + n - \frac{\kappa_s^2 a_s^2}{2} \right) J_n(\kappa_s a_s) - \kappa_s a_s J_{n-1}(\kappa_s a_s) D_n = 0, \end{aligned}$$

constructed by substituting the formulas (16) and (17) into the boundary and coupling conditions (12), (13), and (15) in polar coordinates. Derivatives of the Bessel and Hankel functions are denoted as J_n' and H_n' , respectively. The test simulation has been run with angular frequency $\omega = 5\pi$, element orders $r = 1, \dots, 5$ ($r = 1$ corresponds to bilinear finite

Table 1: Mesh stepsizes and number of timesteps for different spectral orders with $\omega = 5\pi$.

r	1	2	3	4	5
Number of timesteps N	90	330	850	1370	1980

(a) Error $\|\text{Re}(\mathbf{U}_s) - \mathbf{u}_s\|_\infty$ in the solid domain.(b) Error $\|\text{Re}(\Phi) - \phi\|_\infty$ in the fluid domain.Figure 3: Errors between the control algorithm and the Fourier series solution measured in the L^∞ norm.

elements), and material parameters $\rho_f(\mathbf{x}) = 1$ and $\rho_s(\mathbf{x}) = 2.7$, $c(\mathbf{x}) = 1.5$, $c_p(\mathbf{x}) = 6.20$ and $c_s(\mathbf{x}) = 3.12$. The computational domains are presented by rectangular meshes. The minimum and maximum mesh stepsizes in the fluid domain are $h_{f,min} \approx 0.0833$ and $h_{f,max} \approx 0.1963$. Respectively, the minimum and maximum mesh stepsizes in the solid domain are $h_{s,min} \approx 0.0530$ and $h_{s,max} \approx 0.0982$. Stopping criterion of the control algorithm is achieved when the relative euclidean norm of the gradient of the least-squares functional is below 10^{-5} , whereas the Fourier modes n are computed until the relative difference attained by adding the next mode to the series is below 10^{-5} . The number of timesteps used for solving the time-dependent equations is shown in Table 1. The errors, measured using the L^∞ norm, are presented in Figure 3. In principle, the accuracy of the spatial discretization increases with the element order. The horizontal line in the error curves represent the level of dominating error source which is caused by some factor other than the spatial discretization. The possible sources interrupting the convergence of error level include the accuracy of time discretization, stopping criterion of the Fourier series solution, and approximation of the geometrical shapes. The control algorithm solution in the solid domain with $r = 2$ is illustrated in Figure 4, whereas the real part of the Fourier series solution in the solid domain is shown in Figure 5. Respectively, the control algorithm solution with $r = 2$ and the Fourier series solution of the values of velocity potential in the fluid domain are presented in Figures 6 and 7.

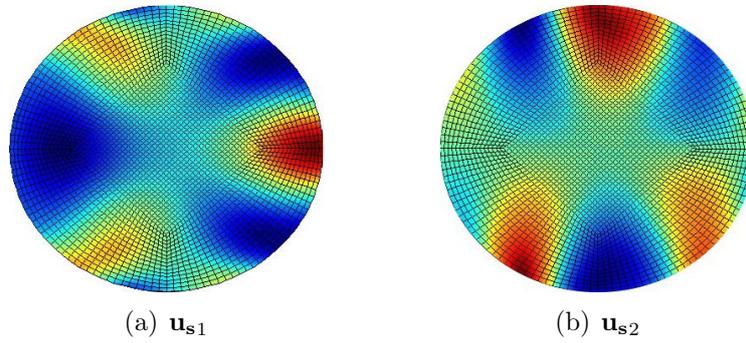


Figure 4: Control algorithm solution of the displacement components in the solid domain.

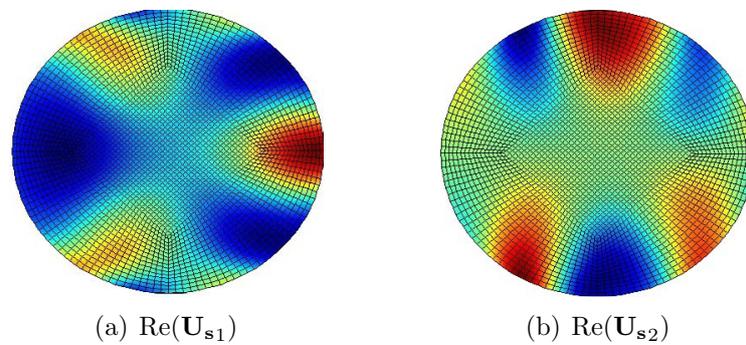


Figure 5: Fourier series solution of the displacement components in the solid domain.

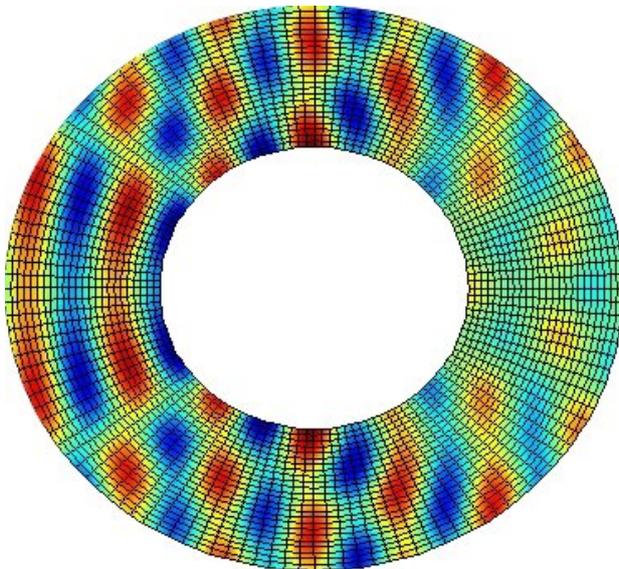


Figure 6: Control algorithm solution ϕ in the fluid domain.

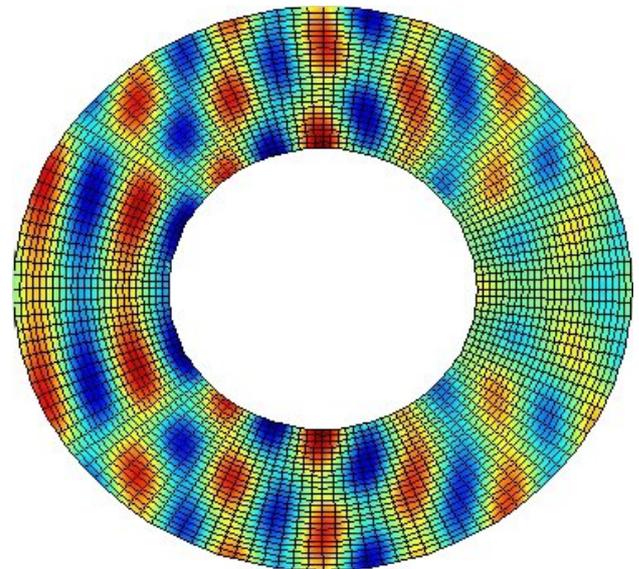
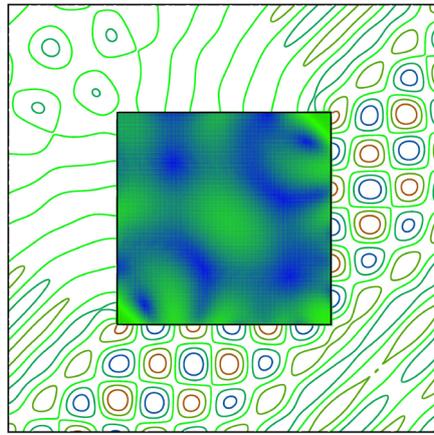
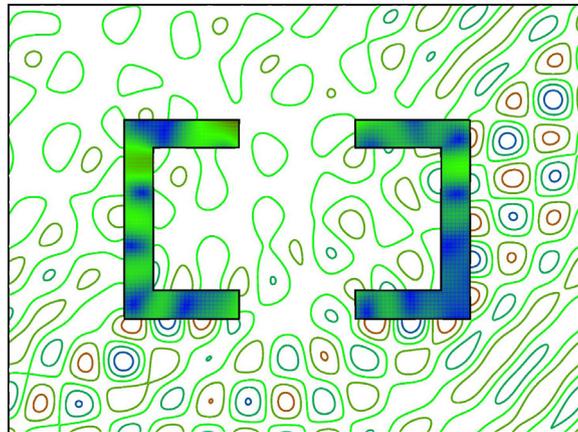


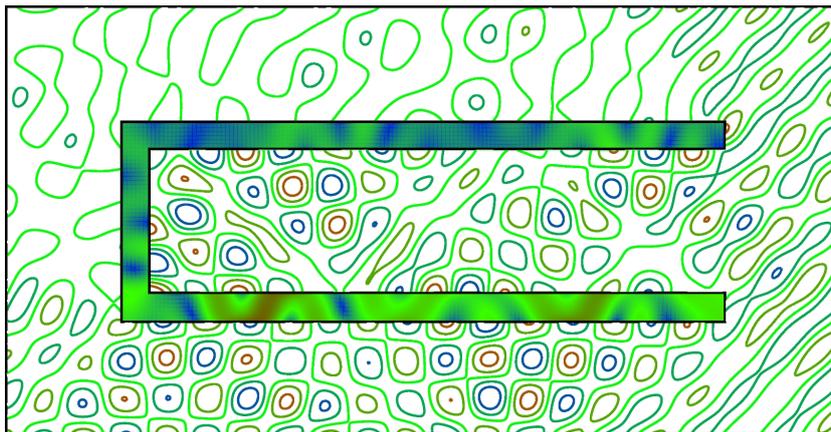
Figure 7: Fourier series solution $\text{Re}(\Phi)$ in the fluid domain.



(a) Convex square obstacle



(b) Two non-convex semi-open obstacles.



(c) Non-convex semi-open obstacle.

Figure 8: Scattering by elastic obstacles with polygonal boundaries with $r = 3$ and $h = 1/28$. The solutions are illustrated as velocity potential fields in the fluid domain and as displacement amplitudes in the structure domain.

Table 2: Mesh stepsizes and number of timesteps for different spectral orders with $\omega = 4\pi$.

r	1	2	3	4	5
Mesh stepsize h	1/80	1/40	1/28	1/20	1/16
Number of timesteps N	400	800	1500	1700	2000

Table 3: The number of iterations with different scatterers.

Type of the obstacle	1	2	3	4	5
convex obstacle (square)	107	105	101	102	104
two non-convex semi-open obstacles	177	166	165	165	165
non-convex semi-open obstacle	199	192	192	193	193

In the second set of experiments, we consider the acoustic wave equation with incident plane wave $\phi_{\text{inc}}(\mathbf{x}, t) = \cos(\vec{\omega} \cdot \mathbf{x} - \omega t)$ implying $y_{\text{fext}} = (\omega - \vec{\omega} \cdot \mathbf{n}_s) \sin(\vec{\omega} \cdot \mathbf{x} - \omega t)$. Three types of elastic obstacles with polygonal boundaries are considered (see Figure 8). In these experiments, we have used the angular frequency $\omega = 4\pi$, which implies that the artificial boundary is located at the distance of two wave lengths from the scatterer. Mesh stepsizes and number of timesteps used in this test are presented in Table 2. The propagation direction is chosen to be $\vec{\omega} = \omega \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$. We have set densities $\rho_f(\mathbf{x}) = 1$ and $\rho_s(\mathbf{x}) = 2.7$, and the propagation speeds $c(\mathbf{x}) = 1$, $c_p(\mathbf{x}) = 5.95$ and $c_s(\mathbf{x}) = 3.12$. The velocity potential fields and displacement amplitudes of the solutions are illustrated in Figure 8 with $r = 3$ and $h = 1/28$, and the number of iterations to reduce the relative euclidean norm of the gradient of the least-squares functional below 10^{-5} is reported in Table 3 for different scatterers. As we can see, the number of iterations is substantially less in the case of convex square scatterer than in the cases of non-convex scatterers. In all the experiments it appears that preconditioning keeps the number of CG iterations bounded with respect to r .

6 CONCLUSIONS

We considered solving time-harmonic acousto-elastic scattering problems by controlling the corresponding time-dependent equations. The validation of the accuracy of the control approach is done by comparing the results with the known time-harmonic solution computed by using the Hankel and Bessel functions. The accuracy of the spatial discretization is shown to increase with the element order until error factor such as time discretization, approximation of the geometrical shapes, or stopping criterion of the Fourier series solution, disturbs the approach. Furthermore, the simulation results show that, the number

of iterations required to attain the stopping criterion is independent of the element order and the number of iterations is substantially less in the case of convex square scatterer than in the cases of non-convex scatterers. The time discretization used in this article is only second order accurate, which restricts the accuracy of the method unless very fine time steps are used. Thus, it could be of interest to use higher order time schemes in future.

Acknowledgements

The authors thank Prof. Tuomo Rossi and Dr. Erkki Heikkola for useful discussions and Dr. Anssi Pennanen and Dr. Janne Martikainen for providing the multigrid solver for preconditioning. The research of the first author is partially carried out in the Digital Product Process technology program supported by the Finnish Funding Agency for Technology and Innovation (TEKES).

REFERENCES

- [1] E. Chaljub, Y. Capdeville, and J-P. Vilotte. Solving elastodynamics in a fluid-solid heterogeneous sphere: a parallel spectral element approximation on non-conforming grids. *Journal of Computational Physics*, 187(2):457–491, 2003.
- [2] L. L. Thompson and P. M. Pinsky. A space-time finite element method for structural acoustics in infinite domains, part I: Formulation, stability and convergence. *Computer Methods in Applied Mechanics and Engineering*, 132:195–227, 1996.
- [3] B. Flemisch, M. Kaltenbacher, and B. I. Wohlmuth. Elasto-acoustic and acoustic-acoustic coupling on non-matching grids. *International Journal for Numerical Methods in Engineering*, 67(13):1791–1810, 2006.
- [4] M. O. Bristeau, R. Glowinski, and J. Périaux. Using exact controllability to solve the Helmholtz equation at high wave numbers. In R. Kleinman, T. Angell, D. Colton, F. Santosa, and I. Stakgold, editors, *Mathematical and Numerical Aspects of Wave Propagation*, pages 113–127, Philadelphia, Pennsylvania, 1993. SIAM.
- [5] M. O. Bristeau, R. Glowinski, and J. Périaux. Controllability methods for the computation of time-periodic solutions; application to scattering. *Journal of Computational Physics*, 147(2):265–292, 1998.
- [6] S. Kähkönen, R. Glowinski, T. Rossi, and R. Mäkinen. Solution of time-periodic wave equation using mixed finite-elements and controllability techniques. *Journal of Computational Acoustics*. To appear.
- [7] E. Heikkola, S. Mönkölä, A. Pennanen, and T. Rossi. Controllability method for the Helmholtz equation with higher-order discretizations. *Journal of Computational Physics*, 225(2):1553–1576, 2007.

- [8] S. Mönkölä, E. Heikkola, A. Pennanen, and T. Rossi. Time-harmonic elasticity with controllability and higher order discretization methods. *Journal of Computational Physics*, 227(11):5513–5534, 2008.
- [9] A. Bermúdez, P. Gamallo, L. Hervella-Nieto, and E. Rodríguez. Finite element analysis of pressure formulation of the elastoacoustic problem. *Numerische Mathematik*, 95(1):29–51, 2003.
- [10] D. Komatitsch, C. Barnes, and J. Tromp. Wave propagation near a fluid-solid interface: A spectral-element approach. *Geophysics*, 65(2):623–631, 2000.
- [11] J. Mandel and M. O. Popa. Iterative solvers for coupled fluid-solid scattering. *Applied Numerical Mathematics*, 54(2):194–207, 2005.
- [12] X. Wang and K.-J. Bathe. Displacement/pressure based mixed finite element formulations for acoustic fluid-structure interaction problems. *International Journal for Numerical Methods in Engineering*, 40(11):2001–2017, 1997.
- [13] L. G. Olson and K.-J. Bathe. Analysis of fluid-structure interactions. A direct symmetric coupled formulation based on the fluid velocity potential. *Computers & Structures*, 21:21–32, 1985.
- [14] B. Engquist and A. Majda. Radiation boundary conditions for acoustic and elastic wave calculations. *Communications on Pure and Applied Mathematics*, 32:313–357, 1979.
- [15] A. Quarteroni, T. Tagliani, and E. Zampieri. Generalized Galerkin approximations of elastic waves with absorbing boundary conditions. *Computer Methods in Applied Mechanics and Engineering*, 163:323–341, 1998.
- [16] G. Cohen. *Higher-Order Numerical Methods for Transient Wave Equations*. Springer-Verlag, Berlin, 2001.
- [17] G. Mur. The finite-element modeling of three-dimensional electromagnetic fields using edge and nodal elements. *IEEE Transactions on Antennas and Propagation*, 41(7):948–953, 1993.
- [18] J.J. Heys, T.A. Manteuffel, S.F. McCormick, and L.N. Olson. Algebraic multigrid for higher-order finite elements. *Journal of Computational Physics*, 204(2):520–532, 2005.
- [19] T. Huttunen, J.P. Kaipio, and P. Monk. An ultra-weak method for acoustic fluid-solid interaction. *Journal of Computational and Applied Mathematics*, 213(1):166–185, 2008.

PROPER GENERALIZED DECOMPOSITION (PGD) TO SOLVE MIXED CONVECTION PROBLEM

A. DUMON*, C. ALLERY* AND A. AMMAR†

*LEPTIAB

Pôle Sciences et Technologie, Avenue Michel Crépeau
17042 La Rochelle Cedex 01, France
e-mail: adumon01@univ-lr.fr, cyrille.allery@univ-lr.fr

†Arts et Metiers ParisTech
2 Bvd du Ronceray
49035 Angers Cedex 01, France
e-mail: Amine.Ammar@ensam.eu

Key words: Reduced order model, Proper Generalized Decomposition, lid-driven cavity, mixed convection.

Abstract. In this communication, the Proper Generalized Decomposition method (PGD), which is an a priori reduction model method, consisting in searching a solution of EDP in separated form, will be applied to solve non-isothermal Navier-Stokes equations. The performances of the PGD method will be compared to the standard resolution technique in term of CPU time as well as in term of accuracy.

1 INTRODUCTION

The numerical simulation of complex fluids flows leads to very large system that cannot be easily solved numerically. This situation is not convenient for optimization problems for which multiple solutions are usually required or for feed-back control problems for which real-time solutions are needed. Consequently, model reduction methods have been developed, as the Proper Orthogonal Decomposition (POD), the Central Voronoi Tessellations (CVT) or the A Priori Reduction method (APR) . . .

In this communication we will focus on another model reduction method, the Proper Generalized Decomposition. It is an iterative method which consists in searching the solution u of an EDP in separating form:

$$u(x_1, \dots, x_N) = \sum_{i=1}^Q \prod_{k=1}^N F_{ki}(x_k) \quad (1)$$

where the N variables x_i can be any scalar or vectorial variable, involving space, time or any parameter of the problem. Thus, if M degrees of freedom are used to discretize each

variable, the total number of unknowns involved in the solution is $Q \times N \times M$ instead of the M^N degrees of freedom involved in traditional approaches. In most cases, where the field is sufficiently regular, the number of terms Q in the finite sum is generally quite reduced (a few tens) and in all cases the approximation converges towards the associated solution. It must be emphasized here that the functions are not 'a priori' known. At each iteration they are adaptatively computed by introducing the approximation separated representation into the model and then by solving the resulting non-linear problem.

The PGD method has been applied by using a separation on space and time variables in order to solve solids mechanics problem [3], or stochastic problems [2] (in this context, the PGD was initially called Generalized Spectral Decomposition). Method was extended to solve multidimensional problems by Ammar et al. [1]. Finally, the method was applied to solve isothermal 2D Navier-Stokes equations in the stationary and instationary cases with a separation only on the space variables [4, 5].

The aim of this communication is to show the ability of the PGD to treat non isothermal flows. After recall the general idea of the PGD, results for the case of the lid-driven cavity differentially heated will be presented.

2 “DESCRIPTION OF THE PGD

2.1 Preliminaries

For the sake of clarity and without losing its general scope, PGD will be examined in the case of a 2D space decomposition. The problem is expressed as follows :

$$\text{Find } U(x,y) \text{ as } \begin{cases} \mathcal{L}(U) = \mathcal{G} & \text{in } \Omega \\ \text{+Boundary Conditions} \end{cases} \quad (2)$$

where \mathcal{L} is a linear¹ differential operator and \mathcal{G} is the second member.

PGD consists in finding an approximation of the solution $U(x,y) \in \Omega = X \times Y \subset \mathbb{R}^2$ with $x \in X \subset \mathbb{R}$ and $y \in Y \subset \mathbb{R}$ as:

$$U(x,y) \approx U_m(x,y) = \sum_{i=1}^m \alpha^i F^i(x) G^i(y) \quad (3)$$

where $U_m(x,y)$ is the approximation of the solution of order m . At each iteration, the solution is enriched with an additional term $\alpha^{m+1} F^{m+1}(x) G^{m+1}(y)$. PGD should be decomposed in three steps. During the first step, “called the enrichment step”, the F^{m+1} and G^{m+1} functions are obtained by solving a small size non-linear problem. Then, for the second step, called the “projection step”, in order to improve the quality of the reconstruction, the $m+1$ α^i coefficients are determined by solving a linear system of size $(m+1)$. Finally, the “check convergence step” consists in the computing of the norm

¹If the operator is not linear, it is necessary to linearize it.

of the residual in order to decide if the solution need more enrichment or not. In the following these three steps will be described in details.

2.2 Enrichment step

At the $m + 1$ stage, the solution approximation of order m is supposed to be known. In this step we search to compute the functions $F^{m+1}(x)$ and $G^{m+1}(y)$. We search $U_m(x, y)$ as

$$U_m(x, y) = \sum_{i=1}^m \alpha^i F^i(x)G^i(y) + F^{m+1}(x)G^{m+1}(y) \quad (4)$$

By introducing equation (4) into problem (2) and by projecting onto each F^{m+1} and G^{m+1} , we obtain:

$$\langle \mathcal{L}\left(\sum_{i=1}^m \alpha^i F^i(x)G^i(y) + F^{m+1}(x)G^{m+1}(y)\right), F^{m+1} \rangle_{L^2(X)} = \langle \mathcal{G}, F^{m+1} \rangle_{L^2(X)} \quad (5)$$

$$\langle \mathcal{L}\left(\sum_{i=1}^m \alpha^i F^i(x)G^i(y) + F^{m+1}(x)G^{m+1}(y)\right), G^{m+1} \rangle_{L^2(Y)} = \langle \mathcal{G}, G^{m+1} \rangle_{L^2(Y)} \quad (6)$$

Solving this set of equations by a fixed point method, for example, gives the F^{m+1} and G^{m+1} .

2.3 Projection step

In order to increase the accuracy of the decomposition, the α^i coefficients are now searched in such a way that the residual is orthogonal to each of the $m + 1$ products of the $F^i G^i$ functions. At this step, we search $U_{m+1}(x, y)$ as,

$$U_{m+1}(x, y) = \sum_{i=1}^{m+1} \alpha^i F^i(x)G^i(y) \quad (7)$$

By introducing solution (7) into (2) and by projecting it according each $F^i G^i$, we obtained the following system of equations:

$$\langle \mathcal{L}\left(\sum_{i=1}^{m+1} \alpha^i F^i(x)G^i(y)\right), F^k G^k \rangle_{L^2(\Omega)} = \langle \mathcal{G}, F^k G^k \rangle_{L^2(\Omega)} \quad \text{for } 1 \leq k \leq m + 1 \quad (8)$$

The resolution of this linear system gives the α^i coefficients .

2.4 Check convergence step

At this step the residual is computed in the following way :

$$Res^{m+1} = \mathcal{L}\left(\sum_{i=1}^{m+1} \alpha^i F^i(x) G^i(y)\right) - \mathcal{G} \quad (9)$$

If the L^2 norm of this residual is lower than a coefficient ϵ set by the user, the PGD algorithm was converged.

3 GOVERNING EQUATIONS OF NON ISOTHERMAL FLOW

• Let us a domain Ω with boundary Γ . Flow in the domain is considered as Newtonian, incompressible and constant properties. Boussinesq approximation is performed. Equations of the problem could be write in dimensionless form :

$$\begin{cases} \nabla \cdot \mathbf{u} = 0 \\ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = \frac{1}{Re} \Delta \mathbf{u} - \nabla p + Ri \theta \vec{y} \\ \frac{\partial \theta}{\partial t} + (\mathbf{u} \cdot \nabla) \theta = \frac{1}{Re Pr} \Delta \theta \end{cases} \quad (10)$$

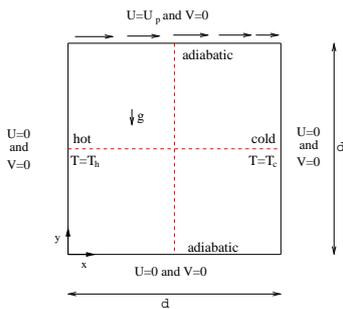
where $\mathbf{u} = (u_1, u_2)$ is the velocity field, θ the temperature and p the pressure. Re is the Reynolds number, Pr the Prandtl number and Ri is the Richardson number which characterizes the mixed convection flow.

These equations are solved by using a time splitting scheme, the so-called Van Kahn algorithm and a finite volume method.

• Solve this problem by PGD consists in searching the unknowns at each time step n in the following form :

$$u_i^n = \sum_{l=1}^{N_{U_i}} \alpha_{U_i}^l F_{U_i}^l(x) G_{U_i}^l(y) \quad p^n = \sum_{l=1}^{N_P} \alpha_P^l F_P^l(x) G_P^l(y) \quad \theta^n = \sum_{l=1}^{N_\Theta} \alpha_\Theta^l F_\Theta^l(x) G_\Theta^l(y) \quad (11)$$

4 APPLICATION OF THE PGD TO THE LID DRIVEN CAVITY



The definition sketch of the problem and the boundary conditions are shown in left figure. It is a square cavity with an incompressible fluid. Vertical walls have different temperature and horizontal walls are adiabatic. The top wall is moving right, and the the two velocity components vanish on the three others walls.

Simulations are made with $\delta t = 10^{-3}$ and for three Richardson number ($Ri = 0.1$, $Ri = 1$ and $Ri = 10$). Results of PGD will be compared to results of a standard solver (Bi-Conjugate gradient solver) in order to compare the accuarcy and CPU time between both methods.

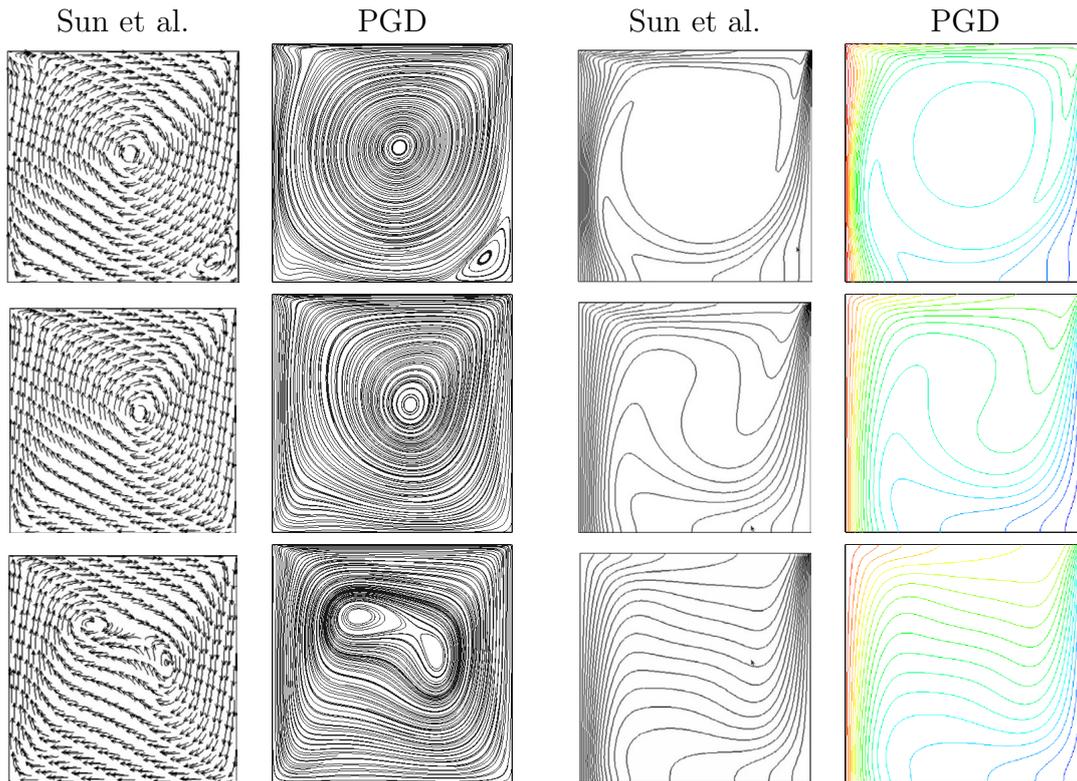


Figure 1: Comparison of the Streamline of velocity obtained by PGD with these from literature (left) and isovalue of temperature field obtained by PGD and from literature(right) for $Ri = 0.1$ (top), $Ri = 1$ (middle), $Ri = 10$ (bottom)

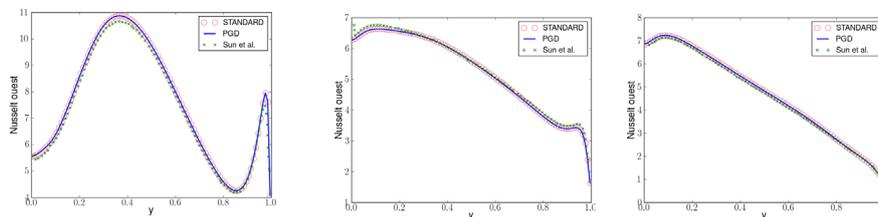


Figure 2: Local nusselt number at west for $Ri = 0.1$ (left), $Ri = 1$ (middle), $Ri = 10$ (right)

Figure 1 shows that the streamlines and isovalue of temperature obtained by PGD are similar to these obtained by Sun et al. [6]. Moreover, figure 2 shows that local Nusselt numbers which is defined by $Nu_w = -(\partial\theta/\partial X)_w/(\theta_h - \theta_c)$, computed from PGD and standard solver and from [6] are similar. About the computational duration, figure 3 shows that from approximately 150 nodes in each direction, PGD becomes faster than standards solver. For a mesh size of 500×500 , the CPU time was three times lower with the PGD solver for $Ri = 0.1$ and $Ri = 1$ and six times lower for $Ri = 10$.

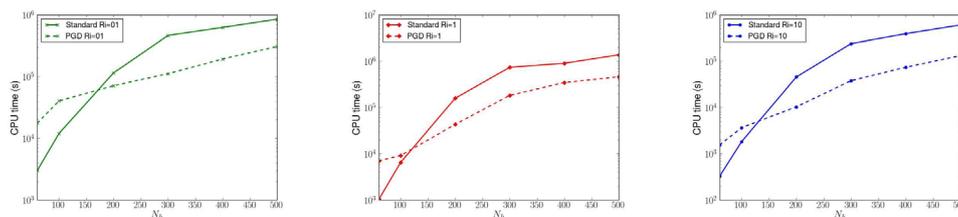


Figure 3: Computational duration in function of the number of nodes in each direction for $Ri = 0.1$ (left), $Ri = 1$ (middle), $Ri = 10$ (right)

4.1 Conclusion

We have observed that the PGD is able to solve this problem accurately with a CPU time saving in relation to the standard solver. This work is a first attempt to use PGD method to solve mixed convection problems. Some improvements are worth further developments. The extension to the 3D case is required to benefit more from the CPU time reduction. Furthermore, being able to consider time as a new variable of the tensorial decomposition could be an original alternative to the time integration scheme. However, the difficulty raised here will be related to the convergence rate of a 4D variables problem.

REFERENCES

- [1] Ammar, A. and Mokdad, B. and Chinesta, F. and Keunings, R. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids *Journal of Non-Newtonian Fluid Mechanics* (2006) **139**:153–176.
- [2] Nouy, A. and Le Maitre, O.P. Generalized spectral decomposition for stochastic nonlinear problems *J.Comput.Phys* (2007) **8**:283–288.
- [3] Ladeveze, P. New approaches and Non-Incremental Methods of Calculation. *Nonlinear Computational Structural Mechanics, Springer Verlag* (1999)
- [4] Dumon, A. and Allery, C. and Ammar, A. Proper general decomposition (PGD) for the resolution of Navier-Stokes equations *Journal of Computational Physics* (2011) **230**:1387–1407.
- [5] Dumon, A. and Allery, C. and Ammar, A. Proper Generalized Decomposition method for incompressible flows in stream-vorticity formulation *European Journal of Computational Mechanics* (2010) **19**:591–617.
- [6] Sun, C. and Yu, B. and Oztop, H.K. and Wang, Y and Wei, J. Control of mixed convection in lid-driven enclosures using conductive triangular fins *International Journal of Heat and Mass Transfer* (2011) **54**:894–909.

THE INFLUENCE OF DIFFERENT TIME INTEGRATION SCHEMES IN ALE DESCRIPTION APPLIED TO MOVING MESHES

FRANK FLITZ*[†], DÖRTE C. STERNEL* AND MICHAEL SCHÄFER*

*Institute of Numerical Methods in Mechanical Engineering
Technische Universität Darmstadt
Dolivostraße 15, 64293 Darmstadt, Germany
e-mail: {flitz, sternel, schaefer}@fmb.tu-darmstadt.de, <http://www.fmb.tu-darmstadt.de>

[†]Graduate School of Computational Engineering
Technische Universität Darmstadt
Dolivostraße 15, 64293 Darmstadt, Germany
e-mail: flitz@gsc.tu-darmstadt.de, <http://www.graduate-school-ce.de>

Key words: Time Discretization, Arbitrary Lagrangian Eulerian, Moving Mesh

Abstract. We analyze the error introduced by the arbitrary Lagrangian Eulerian description on a deliberately moved mesh with a flow field which is steady in Eulerian description. As governing equations, the Navier-Stokes equations for incompressible flow are considered. Two different discretization concepts are investigated for two time integration schemes. It turns out that the difference between the concepts is rather small, but that the difference in the time integration schemes and the dependency of the grid CFL number is much larger.

1 INTRODUCTION

Especially in the context of transient fluid-structure interaction the computational domain of the flow changes with time. When the computational grid is adapted in each time step, it is expensive to interpolate field variables in Eulerian description from one time instance to the next. The arbitrary Lagrangian Eulerian (ALE) description is an elegant extension to the Eulerian description that eliminates the need for this interpolation.

The main question we address is: What error is introduced to a solution which is steady in Eulerian description when it is calculated unsteadily in ALE description. For uniform flow similar tests were already performed in [1] to check for the fulfillment of the space conservation law.

In section 2 we introduce the basic equations for dealing with the ALE description followed by the problem specification in section 3 and the description of the methods

employed in section 4. The test cases and the corresponding results are shown in section 5.

2 BASIC EQUATIONS

The governing equations in ALE description are obtained via minor changes to the ones in Eulerian description. These changes manifest themselves in a time derivative over a time dependent control volume (CV) and a correction to the convective velocity in terms of a grid velocity.

The substantial derivative can be denoted in the Eulerian and in the ALE description as:

$$\frac{d\phi}{dt} = \frac{\partial\phi}{\partial t} + \mathbf{u} \cdot \nabla\phi = \left. \frac{\partial\phi}{\partial t} \right|_{\text{ALE}} + (\mathbf{u} - \mathbf{u}^g) \cdot \nabla\phi. \quad (1)$$

Here the term on the left hand side is the substantial derivative of some generic variable ϕ with respect to time t . $\frac{\partial\phi}{\partial t}$ is the Eulerian time derivative and \mathbf{u} is the material velocity

$$\mathbf{u} = \frac{d\mathbf{x}}{dt} \quad (2)$$

with the spatial coordinate \mathbf{x} . The grid velocity

$$\mathbf{u}^g = \left. \frac{\partial\mathbf{x}}{\partial t} \right|_{\text{ALE}}. \quad (3)$$

appears on the right hand side of (1). For incompressible flow, which is assumed here, the Navier-Stokes equations with the assumption of a Newtonian fluid and no external body force in Eulerian description read as follows:

$$\nabla \cdot \mathbf{u} = 0, \quad (4)$$

$$\rho \frac{\partial\mathbf{u}}{\partial t} + \rho\mathbf{u} \cdot \nabla\mathbf{u} + \mu\nabla^2\mathbf{u} - \nabla p = 0 \quad (5)$$

with the density ρ , the dynamic viscosity μ , and the pressure p . With the change to the ALE description the momentum equation changes correspondingly to (1), while the continuity equation remains the same due to the incompressibility assumption:

$$\nabla \cdot \mathbf{u} = 0, \quad (6)$$

$$\rho \left. \frac{\partial\mathbf{u}}{\partial t} \right|_{\text{ALE}} + \rho(\mathbf{u} - \mathbf{u}^g) \cdot \nabla\mathbf{u} + \mu\nabla^2\mathbf{u} - \nabla p = 0. \quad (7)$$

For the application of a finite-volume method the Navier-Stokes equations have to be rewritten in integral form. These change from the Eulerian description

$$\int_S \mathbf{u} \cdot \mathbf{n} dS = 0, \quad (8)$$

$$\frac{\partial}{\partial t} \int_V \rho\mathbf{u} dV + \int_S \rho\mathbf{u}\mathbf{u} \cdot \mathbf{n} dS + \mu \int_S \nabla\mathbf{u} \cdot \mathbf{n} dS - \int_V \nabla p dV = \mathbf{0} \quad (9)$$

with \mathbf{n} being the face normal vector, to the ALE description

$$\int_S \mathbf{u} \cdot \mathbf{n} \, dS = 0, \quad (10)$$

$$\frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V \rho \mathbf{u} \, dV + \int_S \rho \mathbf{u} (\mathbf{u} - \mathbf{u}^s) \cdot \mathbf{n} \, dS + \mu \int_S \nabla \mathbf{u} \cdot \mathbf{n} \, dS - \int_V \nabla p \, dV = \mathbf{0}. \quad (11)$$

For further details about the ALE description see for instance [2].

3 PROBLEM SPECIFICATION

We address the following question: What error is introduced to a solution which is steady in Eulerian description when it is calculated unsteadily in ALE description?

Similar tests were already performed for uniform flow ([1], [3]). Exactly resampling a uniform flow is a consequence of the fulfillment of the space conservation law ([1]) or geometric conservation law ([4]), as it is called more often. As it was shown in [3] this law is a necessary and sufficient condition for the time integration scheme in ALE description being as stable as the corresponding scheme in Eulerian description.

Since we calculate the grid velocity implicitly corresponding to the space conservation law our focus lies on the discretization of the time derivative term

$$\frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V u \, dV. \quad (12)$$

We examine its influence on a flow with a gradient in the direction of the grid movement, but we do not determine the accuracy of the time integration in ALE description itself, as it is done in [5].

Two concepts for discretizing the time derivative are introduced in the next section. For the sake of simplicity only 1d and quasi-1d configurations are considered.

4 METHODS

Before discretizing (12) in time it is discretized in the ALE domain according to the midpoint rule, yielding:

$$\frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V u \, dV \approx \frac{\partial}{\partial t} \Big|_{\text{ALE}} (u \delta V). \quad (13)$$

The two obvious ways to discretize this term are: *Concept A*: treat $u \delta V$ as one term and: *Concept B*: treat u and δV as two distinct terms.

For time integration we use the implicit Euler (EI) scheme and the second order backward differencing (BDF2) scheme. Applying concept A to the EI scheme yields

$$\frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V u \, dV \approx \frac{1}{\Delta t} (u^{n+1} \delta V^{n+1} - u^n \delta V^n) \quad (14)$$

with Δt being the constant time interval, $n + 1$ indicating the current time instance and n the old one. Applying concept B yields

$$\begin{aligned} \frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V u \, dV &\approx \frac{\partial}{\partial t} \Big|_{\text{ALE}} (u) \delta V + u \frac{\partial}{\partial t} \Big|_{\text{ALE}} (\delta V) \\ &\approx \frac{1}{\Delta t} \left((u^{n+1} - u^n) \delta V^{n+1} + u^{n+1} (\delta V^{n+1} - \delta V^n) \right) \\ &= \frac{1}{\Delta t} \left(u^{n+1} (2\delta V^{n+1} - \delta V^n) - u^n \delta V^{n+1} \right) \end{aligned} \quad (15)$$

with $n - 1$ the second last time instance. Applied to the BDF2 scheme concept A and concept B give

$$\frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V u \, dV \approx \frac{1}{\Delta t} \left(\frac{3}{2} u^{n+1} \delta V^{n+1} - 2u^n \delta V^n + \frac{1}{2} u^{n-1} \delta V^{n-1} \right) \quad (16)$$

and

$$\begin{aligned} \frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V u \, dV &\approx \frac{\partial}{\partial t} \Big|_{\text{ALE}} (u) \delta V + u \frac{\partial}{\partial t} \Big|_{\text{ALE}} (\delta V) \\ &\approx \frac{1}{\Delta t} \left(\left(\frac{3}{2} u^{n+1} - 2u^n + \frac{1}{2} u^{n-1} \right) \delta V^{n+1} \right. \\ &\quad \left. + u^{n+1} \left(\frac{3}{2} \delta V^{n+1} - 2\delta V^n + \frac{1}{2} \delta V^{n-1} \right) \right) \\ &= \frac{1}{\Delta t} \left(u^{n+1} \left(3\delta V^{n+1} - 2\delta V^n + \frac{1}{2} \delta V^{n-1} \right) \right. \\ &\quad \left. - 2u^n \delta V^{n+1} + \frac{1}{2} u^{n-1} \delta V^{n+1} \right), \end{aligned} \quad (17)$$

respectively.

5 NUMERICAL INVESTIGATIONS

The influence of the two concepts is analyzed for three test cases. The test cases concern flows varying in the direction of the grid movement. As a first test case a manual calculation on a concrete example flow is considered. The second test case has the aim to check the implementation and the validity of the boundary conditions. Finally, the third test case examines a general nonlinear flow.

The numerical simulations are carried out with the finite-volume code FASTEST ([6]), which is based on hexahedral block structured grids and uses a pressure-correction scheme with the SIMPLE algorithm. The discrete system is solved via a geometric multigrid method incorporating the incomplete LU decomposition method by Stone as a smoother. The fluxes are discretized via central differences. For the fine grid in the channel flow configuration (see section 5.3) 6 grid levels are used. The space conservation law is fulfilled implicitly as proposed in [1].

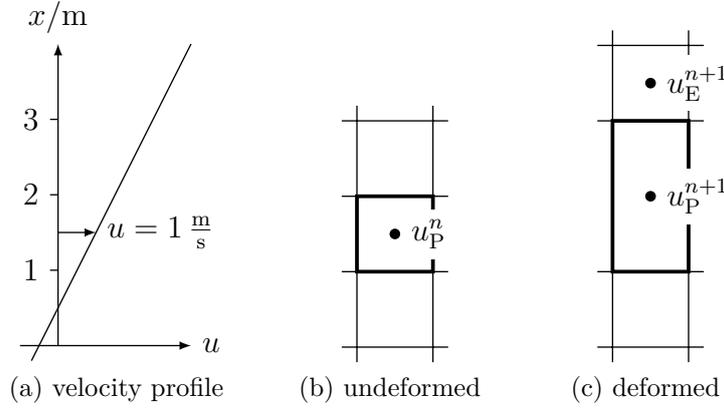


Figure 1: Velocity profile and configuration for the manual calculation

5.1 Manual test case

The manual calculations are conducted on a segment with three adjoining CVs immersed in a Couette flow (see Figure 1a). The extension of the CVs in the two directions orthogonal to the velocity gradient can be deliberately chosen, but is equal for all CVs and constant in time (see Figure 1b). Therefore this is a quasi 1d configuration. Constant pressure is assumed.

In order to judge whether a flow with a CV size change in ALE description will give the same result as the steady Eulerian solution, singly the time derivative and the grid velocity part of the convective term have to be considered. Therefore, the momentum equation in Eulerian description (9) in steady form is subtracted from the momentum equation in ALE description (11), yielding:

$$\frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V \rho \mathbf{u} \, dV - \int_S \rho \mathbf{u} \mathbf{u}^g \cdot \mathbf{n} \, dS = \mathbf{0}. \quad (18)$$

This equation has to be fulfilled for switching between the undeformed and the deformed configuration. We test this for the EI scheme in 1d on the configuration displayed in Figure 1. The time step size Δt is set to 1 s and the density ρ is $1 \frac{\text{kg}}{\text{m}^3}$. Concept A results in

$$\begin{aligned} & \frac{\partial}{\partial t} \Big|_{\text{ALE}} \int_V \rho u \, dV - \int_S \rho u u^g \, dS \\ & \approx \frac{\rho}{\Delta t} (u_P^{n+1} \delta V^{n+1} - u_P^n \delta V^n) - \frac{\rho}{\Delta t} (\delta V^{n+1} - \delta V^n) \left(\frac{1}{3} u_P^{n+1} + \frac{2}{3} u_E^{n+1} \right) \\ & = -0.5 \frac{\text{kg m}}{\text{s}^2} \neq 0 \end{aligned} \quad (19)$$

and concept B yields

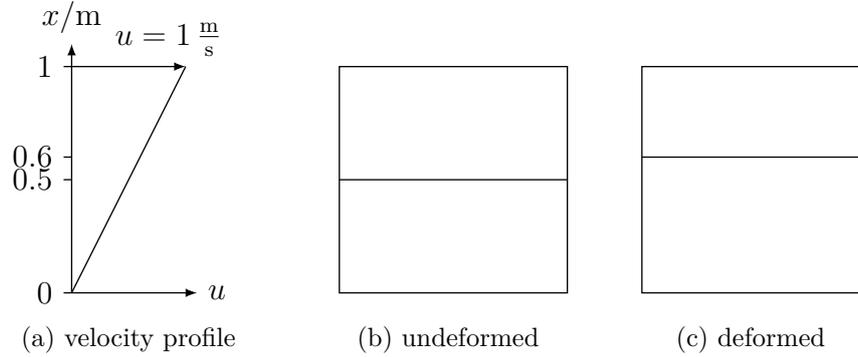


Figure 2: Velocity and CVs for the linear test case

$$\begin{aligned}
 & \left. \frac{\partial}{\partial t} \right|_{\text{ALE}} \int_V \rho u \, dV - \int_S \rho u u^g \, dS \\
 & \approx \frac{\rho}{\Delta t} \left(u_P^{n+1} (2\delta V^{n+1} - \delta V^n) - u_P^n \delta V^{n+1} \right) \\
 & \quad - \frac{\rho}{\Delta t} (\delta V^{n+1} - \delta V^n) \left(\frac{1}{3} u_P^{n+1} + \frac{2}{3} u_E^{n+1} \right) \\
 & = 0.
 \end{aligned} \tag{20}$$

As we can see, concept A fails, while concept B passes this simple test.

5.2 Linear test case

In the finite-volume solver FASTEST, both concepts are tested on a Couette flow with two CVs in order to test the implementation. The two CVs span an area of $1 \text{ m} \times 1 \text{ m} \times 1 \text{ m}$. At the beginning of the simulation both CVs have an extension of 0.5 m in the wall normal direction, which increases to 0.6 m for the lower one and decreases to 0.4 m for the upper one correspondingly (see Figures 2b and 2c).

The velocity of the lower wall is 0 and of the upper wall is $1 \frac{\text{m}}{\text{s}}$. In the directions orthogonal to the wall normal direction periodic boundary conditions are applied. For the viscosity a small but finite value ($\mu = 1 \times 10^{-8} \frac{\text{kg}}{\text{s m}}$) is used to avoid diffusion effects weakening the errors introduced by the schemes.

We obtain results confirming the observations of section 5.1. The converged solution with concept A is erroneous, while concept B gives the exact result.

5.3 Nonlinear test case

As a nonlinear test case a flow between two flat plates is considered. The calculation domain is a cube with edge length 1 m , which is splitted in wall normal direction into 4 blocks of equal size (Figure 3b) with 32 equally spaced CVs in each direction for the fine grid, giving a total number of $131\,072$ CVs. The boundary conditions are given by no slip

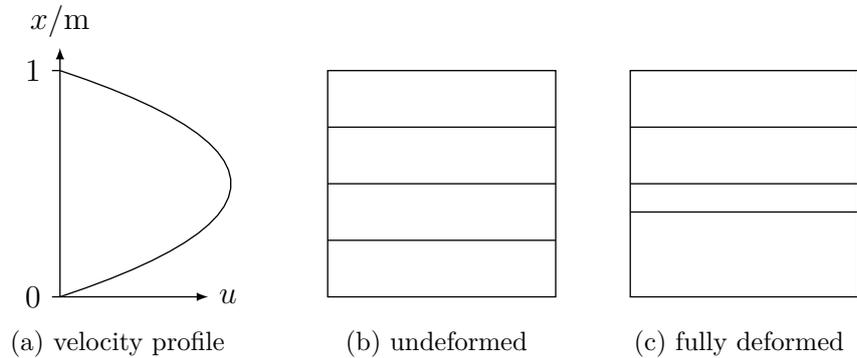


Figure 3: Velocity and blocks for nonlinear test case

walls and periodic conditions in the other two directions. The flow is driven by a pressure gradient of $1.2 \times 10^{-7} \frac{\text{Pa}}{\text{m}}$ applied as a body force. Together with a dynamic viscosity of $1 \times 10^{-8} \frac{\text{kg}}{\text{s m}}$ this yields a maximal velocity of $1.5 \frac{\text{m}}{\text{s}}$. The density is equal to $1 \frac{\text{kg}}{\text{m}^3}$.

During the computation the interface between the lowest block and its neighboring one is shifted by one eighth of the plate clearance towards the middle of the domain (see Figure 3c). The CVs in each of the two blocks are uniformly expanded and compressed, respectively. The other two blocks stay unmodified. As initial solution for the calculation the converged solution of the undeformed grid is used.

The following four parameters are varied:

- Time integration scheme,
- Grid resolution (see Table 1),
- Displacement velocity,
- Time step size.

Our calculations show that regarding the error a increase in the displacement velocity can be eliminated by a decrease in time step size and vice versa. Therefore these two parameters are replaced by a new one: Displacement steps. The full displacement is kept constant, as displayed in Figure 3c. Only the number of equally distributed displacement steps to achieve this full displacement is of interest.

We examine the error in the velocity component parallel to the pressure gradient after the full displacement. It is calculated as the CV size weighted L_2 -norm of the difference between the solution right after the full displacement and the steady solution on the fully deformed grid. The results are shown in Figure 4. The difference between the concepts

Table 1: Grid levels for the nonlinear test case

Grid	Control volumes
fine	131 072
medium	16 384
coarse	2048

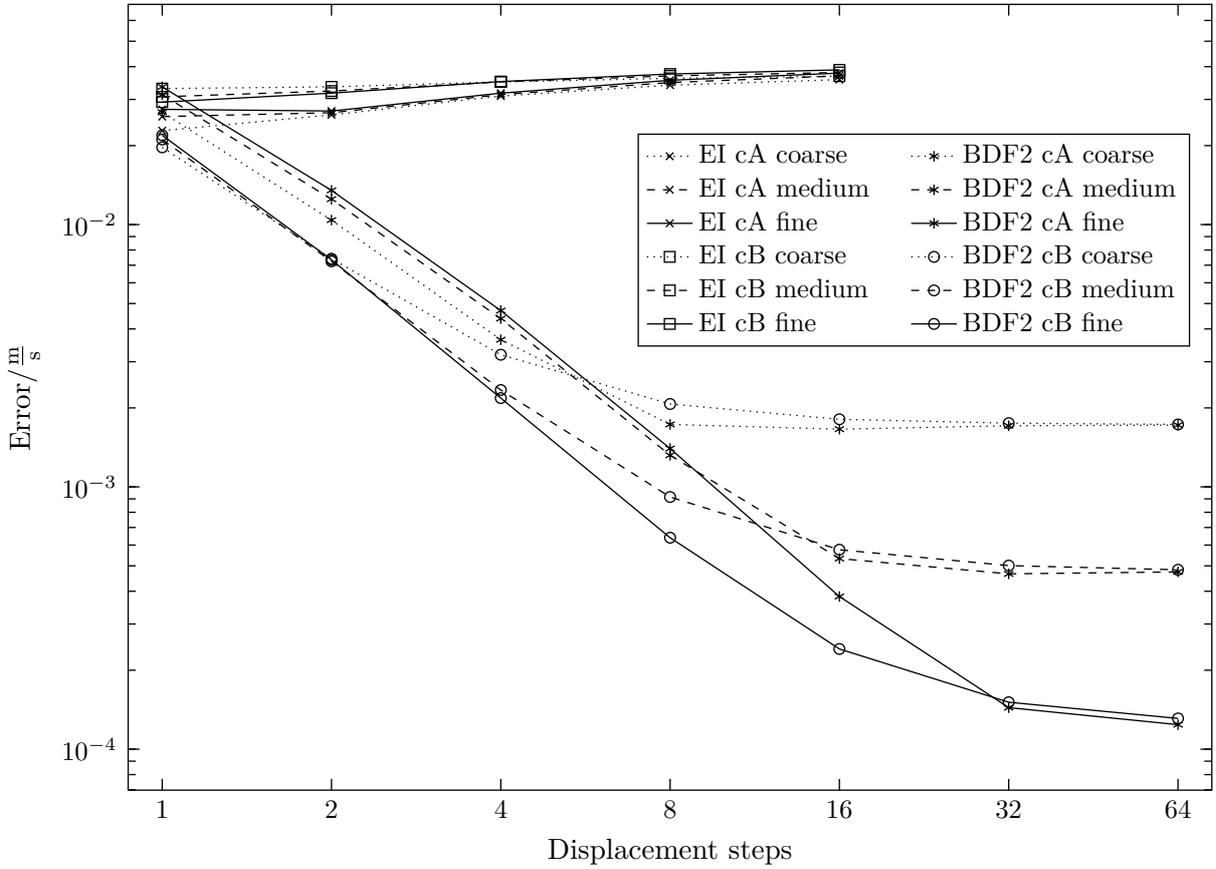


Figure 4: Errors for the nonlinear test case. cA denotes concept A and cB concept B.

is rather small in comparison to the difference between the schemes.

For the EI scheme there is no significant difference between the concepts. However, for the BDF2 scheme the error for concept A is about twice the error of concept B for a small number of displacement steps. For a large number of displacement steps there is also no significant difference.

While for the EI scheme the error is nearly independent of the number of displacement steps, the error of the BDF2 scheme shows a characteristic change between the behavior for a small and a large number of displacement steps. For the BDF2 scheme with a small number of displacement steps a decrease of the error with increasing number of displacement steps of the order of about 1.5 can be seen. For a large number of displacement steps the error stays nearly constant. The change between these two behaviors happens at a grid dependent number of displacement steps. In order to describe this behavior the grid CFL number ν^g is introduced:

$$\nu^g = \frac{u^g \Delta t}{\Delta x}. \quad (21)$$

For example, 16 displacement steps on the fine grid result in a grid CFL number of 1 at the beginning of the computation, while due to the decreasing CV size within the computation, the grid CFL number increases gradually to a value of 2 at the end of the computation. The intersection of the two asymptotes the error approaches for small and large numbers of displacement steps is in the vicinity of an average grid CFL number of 1. Generally, this is true for all grid levels.

6 CONCLUSIONS

Although the results from the linear test case induce the conclusion that the two considered discretization concepts have significant influence on the error, the results from the nonlinear test case show that this difference is negligible in contrast to the influence of the integration scheme or the grid CFL number. While the EI scheme shows no dependency of the grid CFL number in the analyzed range, the BDF2 scheme on the other hand does. The results imply, that for the BDF2 scheme attention should be paid to the grid CFL condition.

7 ACKNOWLEDGEMENTS

This work was supported by the 'Excellence Initiative' of the German Federal and State Governments and the Graduate School of Computational Engineering at Technische Universität Darmstadt. We would like to thank Stephen Sachs for his kind cooperation that led to this work.

REFERENCES

- [1] Demirdzic, I. and Peric, M. Space conservation law in finite volume calculations of fluid flow. *International Journal for Numerical Methods in Fluids* **8**, 1037–1050 (1988).
- [2] Donea, J., Huerta, A., Ponthot, J.-P. and Rodríguez-Ferran, A. *Arbitrary Lagrangian–Eulerian Methods* (John Wiley & Sons, Ltd, 2004).
- [3] Farhat, C., Geuzaine, P. and Grandmont, C. The Discrete Geometric Conservation Law and the Nonlinear Stability of ALE Schemes for the Solution of Flow Problems on Moving Grids. *Journal of Computational Physics* **174**, 669–694 (2001).
- [4] Thomas, P. D. and Lombard, C. K. Geometric Conservation Law and Its Application to Flow Computations on Moving Grids. *AIAA Journal* **17**, 1030–1037 (1979).
- [5] Geuzaine, P., Grandmont, C. and Farhat, C. Design and analysis of ALE schemes with provable second-order time-accuracy for inviscid and viscous flow simulations. *Journal of Computational Physics* **191**, 206–227 (2003).
- [6] Technische Universität Darmstadt, Institute of Numerical Methods in Mechanical Engineering, Darmstadt. *FASTEST Manual* (2005).

COMPUTATIONAL FLUID-STRUCTURE INTERACTION SIMULATIONS FOR WIND INDUCED VIBRATIONS IN SILO GROUPS

J. HILLEWAERE*, J. DEGROOTE[†], G. LOMBAERT*, J. VIERENDEELS[†]
AND G. DEGRANDE*

*Department of Civil Engineering
Katholieke Universiteit Leuven
Kasteelpark Arenberg 40, B-3001 Heverlee, Belgium
e-mail: jeroen.hillewaere@bwk.kuleuven.be, bwk.kuleuven.be/bwm

[†]Department of Flow, Heat and Combustion Mechanics
Ghent University
St. Pietersnieuwstraat 41, B-9000 Gent, Belgium
e-mail: joris.degroote@ugent.be, www.ugent.be/ir/floheacom

Key words: Silo, Cylinder Group, Owalling, Wind-Structure Interaction

Abstract. During a storm in October 2002, wind induced owalling vibrations were observed on several empty silos of a closely spaced group consisting of 8 by 5 silos in the port of Antwerp (Belgium). First, a thorough understanding of the fluid flow around the group is required to clarify the underlying mechanisms for the vibration. Since the configuration and orientation of the group drastically change the pressure distribution on the silos of the group, the flow regime around and within the silo group has been simulated for 7 angles of incidence between 0° and 90°, leaving other parameters unchanged (e.g. spacing ratio, Reynolds number,...). The flow regime shows similarities with the flow within tube arrays (e.g. heat exchangers) and the flow around rectangular cylinders. By a ‘one way coupling’ of static (time averaged) and dynamic (fluctuating) pressure loadings on the cylinder surfaces, two probable causes of wind induced silo vibrations in the group are observed. The first, as a result of large static wind pressures and fluctuating drag and lift coefficients, might lead to rigid body motions of the statically deformed silos. The second, due to higher dynamic pressure oscillations, can excite owalling oscillations in the third and fourth eigenmodes at the lee side of the group, corresponding with the lowest eigenfrequencies of the silos and the visually observed vibrations in 2002. Although it is shown by this ‘one way coupling’ that owalling vibrations can be excited in the group, more advanced ‘two way coupled’ fluid-structure interaction simulations are required to determine the underlying mechanism inducing these aeroelastic deformations.

1 INTRODUCTION AND MOTIVATION

During a storm in October 2002, ovalling was observed on several empty silos near the corners of a group of 40 silos in the port of Antwerp (Belgium). No explanation for these wind induced instabilities or appropriate design guidelines to avoid them can be found in standards, e.g. Eurocode 1 [1]. A more realistic estimation of wind pressures and forces on the silos is required.

Numerical simulations are used to study the observed wind induced ovalling vibrations in the closely spaced Antwerp silo group, organized in 5 rows of 8 silos (figure 1). First, the specific structural behaviour with natural frequencies and according ovalling mode shapes of the silos is presented in the next section. In the third section, numerical simulation results of the turbulent wind flow, modelled as incompressible (low Mach number) flow, around the Antwerp silo group are shown. The influence of the angle of incidence α of the wind flow is investigated while other parameters such as spacing ratio, Reynolds number, etc. are left unchanged. For the present 8 by 5 closely spaced group configuration in a storm regime at post-critical Reynolds number ($Re = 1.24 \times 10^7$), no experimental data are available. Therefore, validation of the numerical procedure is performed for the better documented case of 2D flow around a single cylinder in the post-critical regime. The more challenging simulations of the 2D flow around the entire silo group are furthermore validated qualitatively by assessing similarities of the present flow with the flow within tube arrays (e.g. heat exchangers) and the flow around rectangular cylinders. In the fourth section, the pressure distribution on the silos is investigated to verify whether and at which locations in the group pressure fluctuations can excite the ovalling eigenmodes of the silos. It is furthermore verified if this ‘one way coupling’ of structural and fluid dynamics model is sufficient to explain the existence of the ovalling vibrations in the silo group or whether more advanced ‘two way coupled’ fluid-structure interaction (FSI) simulations are required.

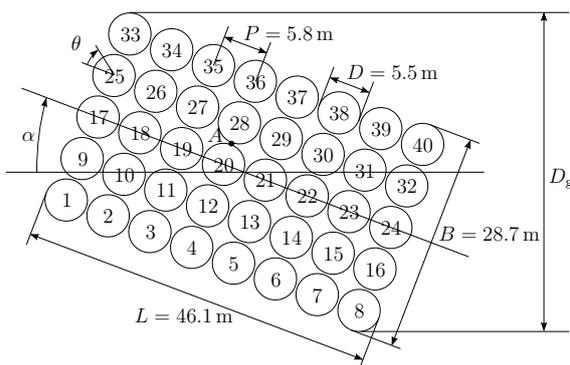


Figure 1: Plan view of the silo group with numbering of the individual silos. Normative dimensions are given as well as definitions for the angle of incidence α and the angle θ on the circumference of an individual cylinder.

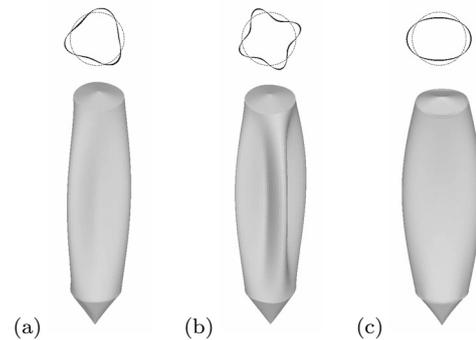


Figure 2: Selected ovalling eigenmodes of a single silo, (a) mode (1,3) at 3.93 Hz, (b) mode (1,4) also at 3.93 Hz and (c) mode (1,2) at 7.75 Hz [3].

2 STRUCTURAL BEHAVIOUR OF SILOS

Wind induced ovaling vibrations are an aeroelastic phenomenon where the cross section of the structure deforms as a shell without bending deformation with respect to the longitudinal axis of symmetry [2]. The ovaling mode shapes for the present thin walled empty silos have been studied by Dooms et al. [3] and are referred to by a couple (m, n) where m denotes the half wave number in the axial direction and n is the number of circumferential waves (figure 2). The lowest natural frequencies for these silos are found for ovaling mode shapes $(1, 3)$ and $(1, 4)$ at $f_n = 3.93$ Hz.

3 AIR FLOW AROUND THE 8 BY 5 CYLINDER GROUP

The turbulent air flow around the 8 by 5 silo group is simulated numerically for 7 angles of incidence ($0^\circ \leq \alpha \leq 90^\circ$). Other influence parameters are left unchanged (e.g. spacing ratio, Reynolds number,...). After the introduction of the computational procedure, the case of a single cylinder in cross flow is calculated for validation. Similarities of the present flow around the silo group with the flow within tube arrays (e.g. heat exchangers) and the flow around rectangular cylinders are examined for qualitative validation of the simulation results. Hence, a distinction is implicitly made between vibrations related to the periodicity of the interstitial flow and vibrations caused by the large vortex structures behind the entire cylinder bundle.

3.1 Computational procedure

The 2D unsteady Reynolds averaged Navier-Stokes (URANS) discretized set of equations is solved in the Ansys FLUENT software package, using the hybrid shear-stress transport (SST) turbulence model. While 3D flow simulations over complex bodies have become possible in recent years, they remain very expensive and are therefore limited to moderate Reynolds numbers. On the contrary, 2D simulations are quite feasible, even for complex geometries and relatively high Reynolds numbers [4]. A coupled pressure-based calculation with a second order interpolation of the pressure, a second order upwind interpolation of momentum, turbulent kinetic energy k and specific dissipation rate ω is performed, while a second order implicit, unconditionally stable, time stepping method is used.

In the computations, the air density is $\rho = 1.25 \text{ kg/m}^3$ and its dynamic viscosity is $\mu = 1.76 \times 10^{-5} \text{ Pa}\cdot\text{s}$. The boundaries of the rectangular computational domain are placed at distances of $9D$ to the central cylinder for the inlet and the lateral boundaries and $30D$ for the outlet of the domain, with D the diameter of the cylinder. Equivalently, $9D_g$ and $30D_g$ are used for the group configuration, with D_g the projected width of the silo group (figure 1). At the velocity inlet, an imposed free stream velocity $v_f = 31.8 \text{ m/s}$ is applied, based on Eurocode 1 for the present storm conditions in the vast and flat suburban surroundings of the silo group [1]. The outlet boundary is modelled as a pressure outlet with static pressure equal to the reference pressure. At the lateral boundaries symmetry

is imposed. The cylinder walls are considered smooth and no-slip boundary conditions are applied.

In these transient calculations, both grid and time step independency have been checked. The optimal mesh refinement is chosen and a time step of $\Delta t = 0.005$ s is applied in the simulations.

3.2 Validation of single cylinder simulations

The Building Block Approach, introduced by the AIAA [5], allows for the validation of a proposed computational procedure with a simpler sub-system for which experimental data are available. The lack of experimental data for the 8 by 5 silo group makes this approach particularly appealing. The computational procedure is hence validated for the flow around a single cylinder.

For validation, the present numerical results are compared with experimental data and results of other (2D and 3D) numerical simulations. Several parameters are compared: the Strouhal number $St = f_{vs}L/v_f$, the separation angle θ_s (figure 3) and the pressure coefficient, with f_{vs} the vortex shedding frequency, v_f the free stream velocity of the fluid and L the characteristic length, equal to the diameter D of the cylinder in the present case. The pressure coefficient along the circumference of a cylinder at a certain point in time is defined as

$$C_p(\theta, t) = \frac{p(\theta, t) - p_f}{\rho v_f^2 / 2} \quad (1)$$

with p_f the free stream pressure. The time averaged pressure coefficient $\overline{C}_p(\theta)$ is calculated as the average over multiple vortex shedding periods in time. The time averaged pressure coefficient for the present, single cylinder simulation is shown in figure 3 with $\theta_s = 116^\circ$ and $St = 0.32$.

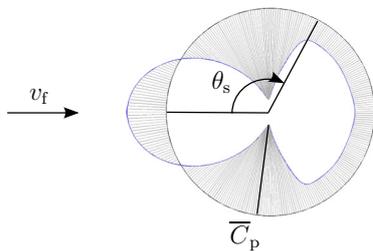


Figure 3: Time averaged pressure coefficient $\overline{C}_p(\theta)$ on the circumference of the cylinder with indication of the free stream velocity v_f and the separation angle θ_s .

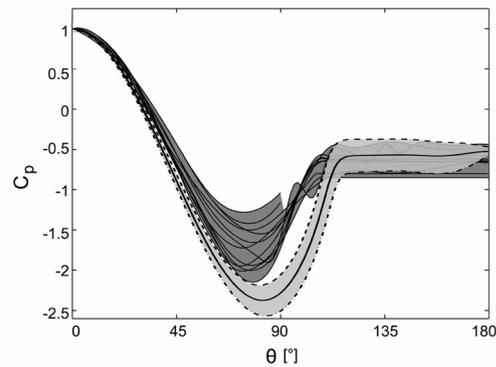


Figure 4: Measured pressure coefficients at Reynolds numbers from 0.73×10^7 to 3.65×10^7 [6] (dark grey zone) vs. present calculated maximal (dashed line), minimal (dash-dotted line) and time averaged pressure coefficients $\overline{C}_p(\theta)$ (solid line) at $Re = 1.24 \times 10^7$.

Zdravkovich [6] gives an elaborate overview of experimental pressure coefficients at Reynolds numbers from 0.73×10^7 to 3.65×10^7 (figure 4) where separation occurs between $\theta_s = 100^\circ$ and 110° . For Reynolds numbers larger than 0.5×10^7 , experimental smooth flow data of Zan [7] indicate that the Strouhal number remains at 0.2, whereas Schewe [8] found that it rises to about 0.3 as the Reynolds number approaches a value of 10^7 ; consistent with the tendency of the Strouhal number to rise from 0.2 to 0.3 in the range of Reynolds numbers between 10^6 and 10^7 [6].

Several numerical simulations have been reported in the literature for highly turbulent cross flows around circular cylinders. Younis et al. [9] performed 2D URANS simulations at $Re = 0.35 \times 10^7$ with different turbulence models and report a Strouhal number of 0.28 and separation at $\theta_s = 120^\circ$. Travin et al. [10] applied 3D DES for Reynolds numbers up to 3×10^6 and found Strouhal numbers 0.35 with separation at $\theta_s = 111^\circ$.

The experimental and numerical data from literature show considerable scatter due to differences in Reynolds number, applied turbulence model, etc. However, generally good agreement is found between the present simulations ($St = 0.32$, $\theta_s = 116^\circ$) and the data from literature.

3.3 Discussion of the flow around the cylinder group

At the transverse corner cylinders of the group (e.g. cylinders 8 and 33 for $\alpha = 30^\circ$, see figure 5c), shear layers in the outer flow are separated while approximately 10% of the flow is forced through the interstitial spaces in the group. These interstitial flows emerge at the lee side, join up and form several local recirculation zones in the wake that coalesce as they are carried downstream. One large scale vortex street is formed in the wake of the entire group, with a flow periodicity depicted by the Strouhal number St (table 1) with characteristic length $L = D_g$. For the smallest angles of incidence ($\alpha = 0^\circ$ and 15° , figures 5a and 5b), it is clear that the emerging interstitial flows on the upper downstream side of the group (cylinders 33 to 40) are joined up and dragged downstream without forming local recirculation zones, due to the proximity of the separated shear layer. The same applies for the highest angles of incidence ($\alpha = 75^\circ$ and 90°), where no such recirculation zones can be formed on the lower side of the group (cylinders 8 to 40).

The flow around the group as a whole resembles the behaviour of a single bluff body in cross flow, similarly to what Kareem et al. [11] observed for two closely spaced cylinders in tandem arrangement. Comparison with experimental data of the flow around a bluff rectangular cylinder in cross flow might hence be useful to assess the influence of porosity and rounded corners of the present 8 by 5 silo group. However, no experimental data can be found in literature for the present high Reynolds number. Knisely [12] performed experiments for a rectangular cylinder ($L/B = 1.67$) in cross flow, but at much lower Reynolds number ($1.2 \times 10^4 \leq Re \leq 2.4 \times 10^4$). He found significantly lower Strouhal numbers (due to differences in Reynolds number and the rounded corners of the silo group), but also a sudden fall in Strouhal numbers for very small and very high angles of incidence ($\alpha \rightarrow 0^\circ$ and $\alpha \rightarrow 90^\circ$) when the separated shear layer reattaches to the

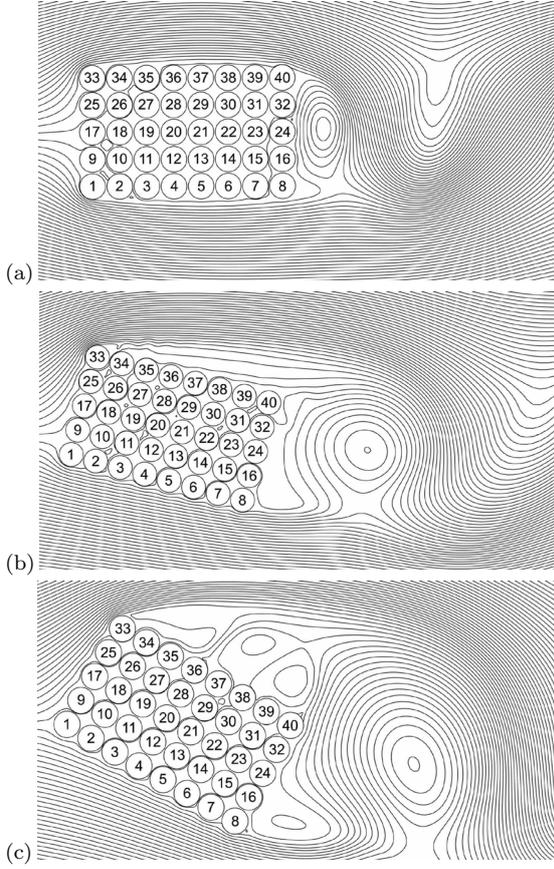


Figure 5: Velocity streamlines of the flow around the 8 by 5 cylinder group for an angle of incidence (a) $\alpha = 0^\circ$ at $t = 80.0$ s, (b) $\alpha = 15^\circ$ at $t = 82.5$ s, and (c) $\alpha = 30^\circ$ at $t = 77.0$ s.

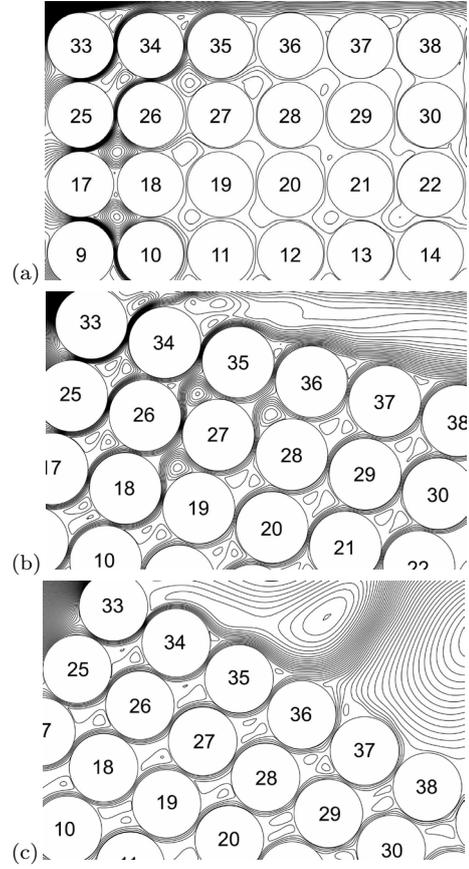


Figure 6: Detail of velocity streamlines for the interstitial space in the 8 by 5 cylinder group for an angle of incidence (a) $\alpha = 0^\circ$ at $t = 78.5$ s, (b) $\alpha = 15^\circ$ at $t = 77.0$ s, and (c) $\alpha = 30^\circ$ at $t = 79.0$ s.

cylinder surface. This sudden fall in Strouhal numbers cannot be observed in the present simulations: the emerging interstitial flows at the downstream side of the group prevent the shear layer from reattaching.

Table 1: Strouhal frequencies (f_{vs}) and Strouhal numbers (St) as a function of the angle of incidence (α) with respective projected width of the silo group (D_g).

α	[$^\circ$]	0	15	30	45	60	75	90
D_g	[m]	28.7	38.4	45.9	50.6	52.3	50.7	46.1
f_{vs}	[Hz]	0.31	0.24	0.17	0.16	0.14	0.18	0.20
St		0.28	0.29	0.25	0.25	0.23	0.29	0.29

3.4 Discussion of the interstitial flow in the cylinder group

Despite important differences (e.g. close spacing between cylinders and limitation to 40 cylinders), the interstitial flow pattern in the present simulations can be related to the flow through tube bundles that has been widely studied for the design of heat exchangers. Tube bundles are typically divided in two categories [13]: the in-line category where cylinders are arranged in square or rectangle arrays and the interstitial flow is mostly straight through the arrays, and the staggered category where cylinders are arranged in rotated square or triangle arrays and the flow is forced along wavy paths. The in-line, square configuration could apply to the cases with $\alpha = 0^\circ$ and $\alpha = 90^\circ$ for the cylinder group while for all other angles of incidence the staggered, rotated square arrangement would be applicable. As shown in figure 6, this is not always the case.

No straight flow pattern for $\alpha = 0^\circ$ (figure 6a) and $\alpha = 90^\circ$ is found. For in-line tube bundles, the presence of the subsequent row prevents the transitional eddies to form and roll-up and the eddies are carried away between the tubes by the jet-like interstitial flow [14]. The present cylinders, however, are too closely packed and these eddies are believed to partially or completely disappear in the distorted flow. Instead, the interstitial flow is not separated from the cylinder wall and follows a wavy path through the array, deflecting the flow up- and downward to the sides of the group, following the shortest path from the high pressures at the leading side of the group to the lower pressures at the lee side of the group. For other angles of incidence, interstitial flows resemble the wavy interstitial flow pattern of staggered tube bundles [13], e.g. for $\alpha = 30^\circ$ (figure 6c). However, for $\alpha = 15^\circ$ (figure 6b) and $\alpha = 60^\circ$ (not shown), the regular wavy pattern is interrupted at arbitrary points in the array, where the interstitial flow separates from the cylinder surface and forms small recirculation zones or even results in local vortex shedding. These irregularities are probably related to the 2D character of the simulations and would not exist in 3D simulations where spanwise velocities are allowed [4].

4 Wind induced ovaling vibrations

Pressure distributions on the walls of the cylinders indicate whether wind induced vibrations of the silos can be excited. Distinction should be made between time averaged pressures on the one hand, which provide an indication of the static deflection of the silos and fluctuating pressures on the other hand, which represent the dynamic excitation of the silos. The silos on the transverse upstream corners of the group where the shear layer is separated (e.g. cylinders 1 and 33 for $\alpha = 0^\circ$, figure 5a, or cylinders 8 and 33 for $\alpha = 30^\circ$, figure 5c) are subject to the largest static pressures for all angles of incident flow. Combined with larger fluctuating drag and lift at these corners, this may result in observable rigid body motions of the statically deformed silos. This vibration phenomenon is however fundamentally different from the observed ovaling of the silos.

Ovaling vibrations can only be triggered by the fluctuating pressures on the cylinder

wall. Therefore, fluctuating pressure coefficients are determined as follows:

$$C'_p(\theta, t) = C_p(\theta, t) - \overline{C}_p(\theta) \quad (2)$$

To investigate the contribution of these fluctuating pressures in the excitation of the eigenmodes of the silos, the pressure coefficients are harmonically decomposed into a series of cosine functions with circumferential wavenumber n , corresponding to the ovaling mode shapes of the axisymmetric structure (figure 2):

$$C'_p(\theta, t) = \sum_{n=0}^{\infty} C_p'^n(t) \cos(n\theta + \overline{\phi}_n) \quad (3)$$

Afterwards, the time history of the fluctuating pressure amplitudes $C_p'^n(t)$ is transformed to the frequency domain by means of a FFT algorithm. Results for angle of incidence $\alpha = 30^\circ$ and for circumferential wave numbers $n = 3$ and $n = 4$ are shown in figure 7 for cylinders 1, 8, 21, 33 and 40.

The frequency spectra for cylinders 1 and 33 (figures 7c and 7a) show no periodicities other than the low frequency contributions related to the large vortex shedding in the wake of the group. However, moving towards the lee side of the group, irregularities appear: higher frequencies also prevail in the frequency spectra for cylinders 8 and 40 (figures 7d and 7b). This frequency content at around 3 Hz to 4 Hz indicates that the third and fourth circumferential eigenmodes of the silos (both at eigenfrequencies of 3.93 Hz) will probably be excited. Moving downstream within the group, contributions at even higher frequencies are also encountered (e.g. cylinder 21, figure 7e). For other angles of incidence, these peaks in the frequency range between 3 Hz and 4 Hz are also found, confirming that the eigenmodes with the lowest eigenfrequencies, i.e. modes (1,3) and (1,4) (figure 2), will most likely be excited at the lee side of the silo group.

Hence, from the present ‘one way coupling’ of pressure fluctuations to the structural eigenmodes, it is found that ovaling vibrations may very well be excited at the lee side silos of the group. However, the underlying mechanism inducing these vibrations has not yet been determined.

It is generally accepted that there are three distinct mechanisms leading to vibrations in tube arrays [15, 16]. Firstly, forces can arise due to coincidence of a structural natural frequency with the vortex shedding frequency in the tube wake. Secondly, fluid-elastic instability (FEI) is based on self-excited forces which are caused by the interaction between tube motion and fluid flow [13]. Finally, turbulent buffeting forces arise due to turbulent fluctuations of the flow pressure. These forces arise as a response to flow turbulence, either initiated upstream or induced within the array itself [15].

Considering the large difference between the natural ovaling frequencies (figure 2) and the vortex shedding frequencies (f_{vs} , table 1), resonance effects can be excluded as a mechanism inducing ovaling vibrations in the silo group. Although periodicities in the interstitial flow may be very different from classical vortex shedding, these do not seem

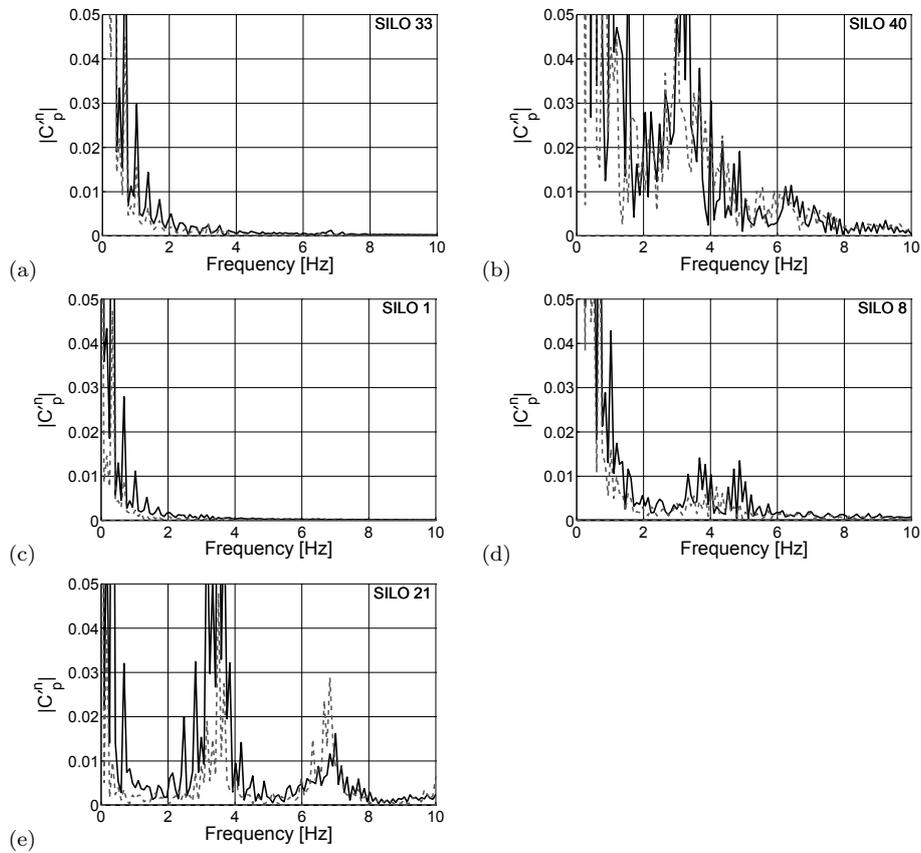


Figure 7: Frequency spectra for the amplitude of $C_p^{(3)}$ (solid line) and $C_p^{(4)}$ (dashed line) for angle of incidence $\alpha = 30^\circ$ for (a) cylinder 33, (b) cylinder 40, (c) cylinder 1, (d) cylinder 8 and (e) cylinder 21.

to be related to ovalling, since they occur throughout the entire group whereas ovalling is only observed on the corner silos. Hence, FEI and/or turbulent buffeting are believed to be the primary causes of wind induced ovalling vibrations on the corners of the silo group.

Although it is confirmed by the present ‘one way’ coupling that ovalling vibrations may exist in the group arrangement, more advanced ‘two way coupling’ FSI calculations are required to verify the influence of FEI and turbulent buffeting. Since problems of aeroelasticity are typically considered weakly coupled problems, a partitioned coupling of the numerical structural model (e.g. finite element model of the silos) and the numerical fluid model (e.g. finite volume approach, considered here), with explicit coupling at the interface, could be applied. However, when modelling the incompressible flow around light and flexible structures, numerical instabilities as the artificial added mass effect may occur [17, 18]. The study of coupling schemes for wind-structure interaction problems is the topic of ongoing research.

5 CONCLUSIONS

In order to elucidate the occurrence of ovaling oscillations on the empty corner silos of a 8 by 5 silo group in the port of Antwerp, the post-critical flow around this closely spaced cylinder group was simulated numerically. 2D URANS simulations for the entire group were performed for 7 angles of incidence α between 0° and 90° .

The group configuration and orientation of the group drastically change the flow regime, showing similarities with the fluid flow around bluff rectangular cylinders. The rounded corners and the porosity of the group have an important influence on the flow regime around the group. Approximately 10% of the incident flow penetrates the group and emerges at the lee side, preventing the shear layer from reattaching at very low and high angles of incidence ($\alpha \rightarrow 0^\circ$ and $\alpha \rightarrow 90^\circ$).

The flow in the interstitial spaces of the group is somewhat similar to the flow in tube bundles. Although the flow pattern is clearly different when the group is oriented parallel to the incident flow ($\alpha = 0^\circ$ or $\alpha = 90^\circ$), for an inclined orientation, the interstitial flow is very similar, following wavy paths through the array. For $\alpha = 15^\circ$ and $\alpha = 60^\circ$, at arbitrary locations in the array, irregularities are observed which are attributed to the 2D character of the numerical simulations.

To verify whether ovaling vibrations can be excited, the pressure distributions on the silos in the group are ‘one way coupled’ to the dynamic structural properties of the silos. Both static deflection (time averaged pressures) and dynamic excitation (fluctuating pressures) of the silos in the group configuration are considered. The silos near the transverse corners of the silo group, where the shear layer is separated, are subject to the largest static pressures for all angles of incident flow. However, to explain ovaling vibrations, dynamic fluctuating pressures have to be considered.

For all angles of incidence α , fluctuating pressures on the silos at the lee side of the group are seen to most likely excite the third and fourth structural ovaling eigenmodes, corresponding with the lowest natural frequencies of the silos. This observation corresponds with the visually detected ovaling eigenmodes with three and four circumferential wavelengths at the corner silos of the group during the 2002 storm in Antwerp.

The ‘one way coupling’ technique, presented here, is sufficient to explain the existence of the ovaling vibrations on the corner silos at the lee side of the silo group. Based on these simulations, the underlying physical mechanisms producing the flow periodicities and eventually inducing the ovaling vibrations are believed to be turbulent buffeting and/or FEI while resonance with some periodic vortex shedding frequency can be excluded. However, more advanced ‘two way coupling’ (FSI) simulations are required to verify the influence of FEI and turbulent buffeting.

ACKNOWLEDGEMENTS

The research in this paper has been performed within the frame of the FWO project G.0275.08 ”Efficient analysis of fluid-structure interaction problems in structural dynam-

ics”, funded by the Research Foundation Flanders (FWO Vlaanderen). The support of FWO is gratefully acknowledged.

REFERENCES

- [1] BIN. *NBN EN 1991-1-4:2005 Eurocode 1: Actions on structures - Part 1-4: General actions - Wind actions*. Belgisch Instituut voor Normalisatie, (2005).
- [2] Païdoussis, M.P., Price, S.J. and Suen, H.C. Ovalling oscillations of cantilevered and clamped-clamped cylindrical-shells in cross flow: An experimental-study. *Journal of Sound and Vibration*, (1982) **83**(4):533-553.
- [3] Dooms, D., Degrande, G., De Roeck, G. and Reynders, E. Finite element modelling of a silo based on experimental modal analysis. *Engineering Structures*, **28**(4):532-542.
- [4] Mittal, R. and Balachandar, S. Effect of three-dimensionality on the lift and drag of nominally two-dimensional cylinders. *Physics of Fluids*, (1995) **7**(8):1841-1865.
- [5] Versteeg, H.K. and Malalasekara, W. *An Introduction to Computational Fluid Mechanics: The Finite Volume Method*. Pearson Education Limited, Essex, England, second edition (2007).
- [6] Zdravkovich, M.M. *Flow Around Circular Cylinders, Volume 1: Fundamentals*. Oxford University Press, Oxford, England, (1997).
- [7] Zan, S.J. Experiments on circular cylinders in crossflow at Reynolds numbers up to 7 million. *Journal of Wind Engineering and Industrial Aerodynamics*, (2008) **96**(6-7):880-886.
- [8] Schewe, G. On the force-fluctuations acting on a circular-cylinder in cross-flow from subcritical up to transcritical Reynolds-numbers. *Journal of Fluid Mechanics*, (1983) **133**(AUG):265-285.
- [9] Younis, B.A. and Przulj, V.P. Computation of turbulent vortex shedding. *Computational Mechanics*, (2006) **37**(5):408-425.
- [10] Travin, A., Shur, M., Strelets, M. and Spalart, P. Detached-eddy simulations past a circular cylinder. *Flow, Turbulence and Combustion*, (2000) **63**(1-4):293-313.
- [11] Kareem, A., Kijewski, T. and Lu, P.C. Investigation of interference effects for a group of finite cylinders. *Journal of Wind Engineering and Industrial Aerodynamics*, (1998) **77-78**:503-520.
- [12] Knisely, C.W. Strouhal numbers of rectangular cylinders at incidence: a review and new data. *Journal of Fluids and Structures*, (1990) **4**(4):371-393.

- [13] Zdravkovich, M.M. *Flow Around Circular Cylinders, Volume 2: Applications*. Oxford University Press, Oxford, England, (2003).
- [14] Hunt, J.C.R. and Eames, I. The disappearance of laminar and turbulent wakes in complex flows. *Journal of Fluid Mechanics*, (2002) **457**:111-132.
- [15] Price, S.J., Païdoussis, M.P., Macdonald, R. and Mark, B. The flow-induced vibration of a single flexible cylinder in a rotated square array of rigid cylinders with a pitch-to-diameter ratio of 2.12. *Journal of Fluids and Structures*, (1987) **1**(3):359-378.
- [16] Weaver, D.S., Lian, H.Y., and Huang, X.Y. Vortex shedding in rotated square arrays. *Journal of Fluids and Structures*, (1993) **7**(2):107-121.
- [17] Causin, P., Gerbeau, J.F. and Nobile, F. Added-mass effect in the design of partitioned algorithms for fluid-structure problems. *Computer Methods in Applied Mechanics and Engineering*, (2005) **194**(42-44):4506-4527.
- [18] Förster, C., Wall, W.A. and Ramm, E. Artificial added mass instabilities in sequential staggered coupling of nonlinear structures and incompressible viscous flows. *Computer Methods in Applied Mechanics and Engineering*, (2007) **196**(7):1278-1293.

FSI SIMULATIONS FOR EXPLOSIONS VERY NEAR REINFORCED CONCRETE STRUCTURES

MATTHEW A. PRICE^{*}, ALEX K. H. LEE, ORLANDO SOTO[†] AND
OI YIN KAREN CHONG[‡]

^{*} Institute of High Performance Computing (IHPC)
1 Fusionopolis Way, #16-16, Singapore 138632
e-mail: pricama@ihpc.a-star.edu.sg, www.ihpc.a-star.edu.sg

[†] Science Applications International Corporation (SAIC)
1710 Saic Drive MS 2-6-9, Mclean, VA 22102, USA
e-mail: orlando.a.soto@saic.com, www.saic.com

[‡] Defence Science & Technology Agency (DSTA)
167 Jalan Bukit Merah, Tower 5 #10-10, Singapore 150167
e-mail: COIYIN@dsta.gov.sg, www.dsta.gov.sg

Key words: Fluid Structure Interaction, Air Blast, Explosion, Reinforced Concrete, K&C Concrete Damage Plasticity Model.

Abstract. The analysis of explosives in contact or very near to reinforced concrete (RC) structures is an important aspect in the design of protective structures and vulnerability assessments. Although this remains a topic of high importance for defence, a more widespread interest has developed as civilian structures become the targets of terrorism. This type of assessment requires a robust simulation method for coupled fluid-structural interactions (FSI) which can handle the explosive detonation, air blast propagation, structural deformation, and damage evolution. This paper describes the application of a loose-coupling method which combines the FEFLO CFD code and SAIC's CSD code for 3D numerical simulations of unconfined and semi-confined explosions near RC structures. This approach takes advantage of the unstructured tetrahedral mesh for the CFD and an embedded method for CSD structures inside the fluid domain. Comparisons of simulations with experiment provide validation, but also reveal some weaknesses of the method. A good agreement between simulation and experiment is found with moderate explosive loading. However, a severe explosive loading with confinement results in extensive damage to the structure which is difficult to reproduce in simulations.

1 INTRODUCTION

In this work, we focus on the prediction of damage to reinforced concrete (RC) structures from the detonation of high explosives. In particular, we are interested in scenarios where the explosive is very near or in contact with the concrete. This presents a challenging problem for simulation as there is significant deformation and damage of the structure which will alter the

blast flow field and it becomes crucial to capture the fluid-structure interaction (FSI) phenomena. To accomplish this, we use a loose-coupled method which combines computational fluid dynamics (CFD) and computational structural dynamics (CSD) techniques.

Our current work focuses on the assessment of the loose-coupled FSI method through comparison with experiments. We follow the simulation methodology established by SAIC for high strain-rate and large deformation response of concrete structures during blast and impact scenarios [1]. First, we investigate a moderate blast loading on a single concrete slab to identify an appropriate set of model parameters. Then we test how well the same set of parameters and simulation method works for moderate and severe semi-confined explosive loading of a more complex RC structure.

2 NUMERICAL APPROACH

2.1 Fluid-Structure Interaction (FSI) methodology

The simulations use a loose-coupled approach which combines previously validated and established CFD and CSD codes through a controller code [2]. Alternatively, a strong coupled method would require a completely re-written single coupled code. The disparity in stiffness and timestep between the fluid and solid domains also creates difficulties for a strong coupled method.

The position and motion of the solid inside the fluid domain is handled with an embedded approach. This has several advantages over a glued-mesh approach. In particular, the CFD and CSD surface meshes do not have to be matching at the interface, which prevents problems created when tracking fragments or the breakup of the solid structures [3]. Figure 1 shows the basic FSI approach with exchange of information between the CFD and CSD codes via the controller code [1, 4].

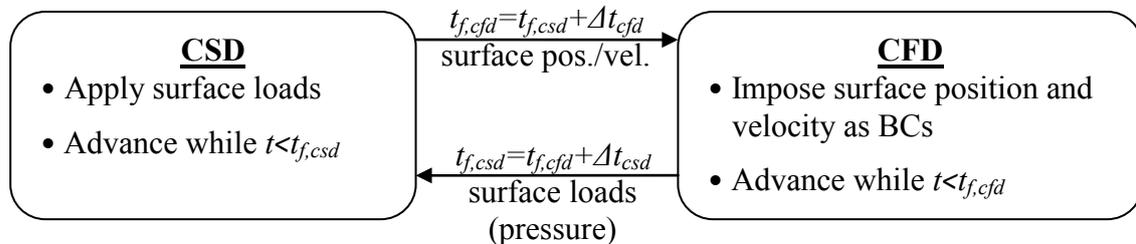


Figure 1: General CFD/CSD coupling procedure during simulation. Each code calculates its own timestep (Δt) and continues from the final time (t_f) of the other code in a staggered manner.

2.2 Computational Fluid Dynamics (CFD) solver

The FEFLO98 code is a 3D, unstructured, edge-based fluid dynamics solver using an Arbitrary Lagrangian-Eulerian (ALE) formulation of the Navier-Stokes and Euler equations. The code uses the FEM-FCT method for shock capturing which has been well-established for applications involving explosions and shock propagation around complex geometries [3, 4]. The explosive is modeled using the Jones-Wilkins-Lee (JWL) equation of state with a programmed burn detonation model and the air is modeled with a “Real Gas” EOS (non-

constant gamma behavior).

2.3 Computational Structural Dynamics (CSD) solver

The structural solver employed is the SAIC-CSD code which uses an explicit second order finite element method. The primary features applicable for RC simulations are ^[1]:

- Fully-integrated Q1/P0 solid elements
- A large strain rate (FE) convective formulation
- The K&C phenomenological plasticity model for damage in the concrete
- Embedded method for reinforcement bars inside solid concrete elements
- Numerical damping to avoid spurious velocities near failed elements
- A general contact algorithm using bin technology for fast node-face searching operations
- All procedures are fully parallelized with a quasi-optimum speed-up on shared memory architectures

The K&C concrete plasticity model has been highly validated for concrete subjected to explosions, impact, and high strain rate events. This material model has three independent strength surfaces (yield, failure, and residual) with consideration of all three stress invariants (I_1 , J_2 , and J_3) ^[5]. Additionally, there is an extension of the plasticity model in tension and a radial path strain rate enhancement which are critical for producing accurate simulations of concrete under rapid loading conditions. The modified plastic strain λ is calculated separately for compressive or tensile loading by:

$$\lambda = \begin{cases} \int_0^{\bar{\epsilon}_p} \frac{d\bar{\epsilon}_p}{r_f(1 + p/r_f f_t)^{b_1}} & \text{for } p \geq 0 \\ \int_0^{\bar{\epsilon}_p} \frac{d\bar{\epsilon}_p}{r_f(1 + p/r_f f_t)^{b_2}} & \text{for } p < 0 \end{cases} \quad (1)$$

where p is pressure, and the effective plastic strain increment is defined as $d\bar{\epsilon}_p = \sqrt{\left(\frac{2}{3}\right) \epsilon_{ij}^p \epsilon_{ij}^p}$. This also accounts for the strain rate enhancement r_f (dynamic increase factor, DIF) of the concrete which is based on the modified CEB formulation ^[6]:

$$DIF_{conc} = \begin{cases} \left(\frac{\dot{\epsilon}}{\dot{\epsilon}_s}\right)^\delta & \text{for } \dot{\epsilon} \leq 1 \text{ s}^{-1} \\ \beta \left(\frac{\dot{\epsilon}}{\dot{\epsilon}_s}\right)^{1/3} & \text{for } \dot{\epsilon} > 1 \text{ s}^{-1} \end{cases} \quad (2)$$

where $\dot{\epsilon}_s$ is a constant, and parameters δ and β are functions of the concrete yield strength. As λ increases, the K&C model damage parameter η will increase from 0 to 1 at λ_m and then decrease back to 0. A scaled damage variable is then calculated by ^[7]:

$$\delta = 2\lambda/(\lambda + \lambda_m) \quad (3)$$

The steel is modeled with a bilinear elasto-plastic material and employes a similar form of strain rate enhancement ^[8]:

$$DIF_{steel} = \left(\frac{\dot{\epsilon}}{10^{-4}} \right)^\alpha \quad (4)$$

where α is a function of the steel yield strength. The reinforced bars are modeled by embedding beam elements into solid hexahedral elements during the CSD initialization. During each simulation time step, the kinematic variables (displacement, velocity, and acceleration) of the “slave” embedded re-bars are interpolated from the “master” solid element’s degrees of freedom, and their internal forces are applied back to their respective master elements by a standard finite element extrapolation.

3 NUMERICAL SIMULATIONS

3.1 Explosion near a concrete slab

A validation of the loose-coupled method was performed by comparing simulations with experimental work described in Zhou *et al.* [9]. The test consisted of a 0.5kg Comp-B charge located 10cm above a 1-way reinforced concrete slab which had dimensions of 100x130x10cm and was clamped along the short edges. The concrete strength was 50MPa and steel rebar assumed properties of ASTM Grade 60. The simulations utilized one plane of symmetry and were run on 4 to 16 CPU of an SGI-Altix system. A similar element size between the CFD and CSD domains was maintained as shown in Figure 2a.

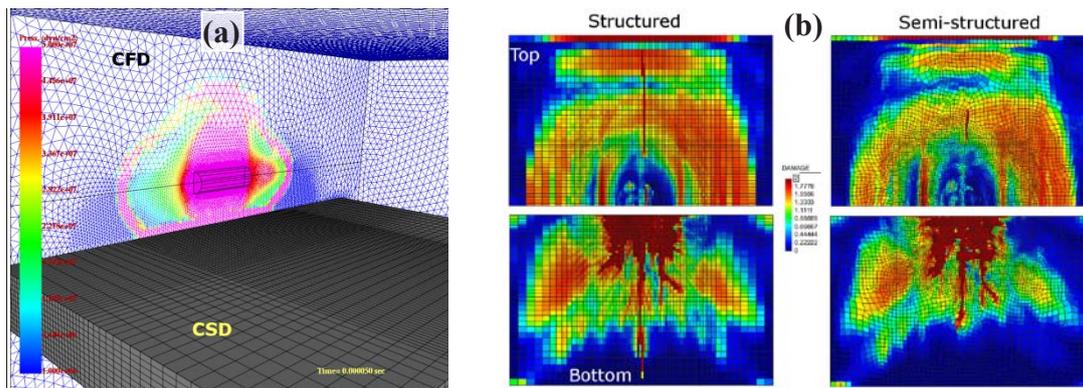


Figure 2: Blast loading on the concrete slab showing CFD and CSD discretizations (a). Comparison of concrete damage variable on the top and bottom surfaces using structured and semi-structured 8-node elements (b).

The validation models were used to investigate the influence of mesh, material, and element parameters on the simulations results. A comparison of simulations using structured and semi-structured 8-node solid elements for the concrete shows some influence on the predicted damage (Fig. 2b). Defining areas of mesh refinement was easier with a semi-structured mesh. The structured mesh produced element failure along straight paths which was questionable. However, a structured mesh with uniform element size was useful when the damage locations are not known *a priori*. Decreasing the element size resulted in a finer resolution of damage paths in the concrete, and also increased the localized damage and element failure. A small change in CSD/CFD element size very rapidly increased the total number of elements and computational resources required which can become restrictive. Table 1 demonstrates this, and also reveals that number of elements for the CFD was at least an order of magnitude

larger than the CSD.

Table 1: Number of elements for the CSD, CFD, and ratio of CFD/CSD for three levels of mesh refinement. “Size” refers to the dimension of the smallest CSD elements (at the slab center).

Mesh	Size	CSD	CFD	Ratio
Coarse	8mm	5.10E+04	3.46E+06	67
Medium	6mm	1.18E+05	4.65E+06	39
Fine	4mm	3.88E+05	8.20E+06	21

Other important input parameters were the damping fraction α and the concrete maximum damage value δ_{max} which triggers element erosion (when $\delta_{max} > \delta$)^[1]. It was found that a few combinations of δ_{max} and α in simulations produced a reasonable agreement with experiment. However, keeping α as low as possible was desired to avoid unrealistic over-damping of the structural response. Values of $\alpha = 0.05-0.1$ and $\delta_{max} = 1.90-1.95$ gave the best results. The same set of parameters determined from this validation study were applied to the more complex semi-confined structures.

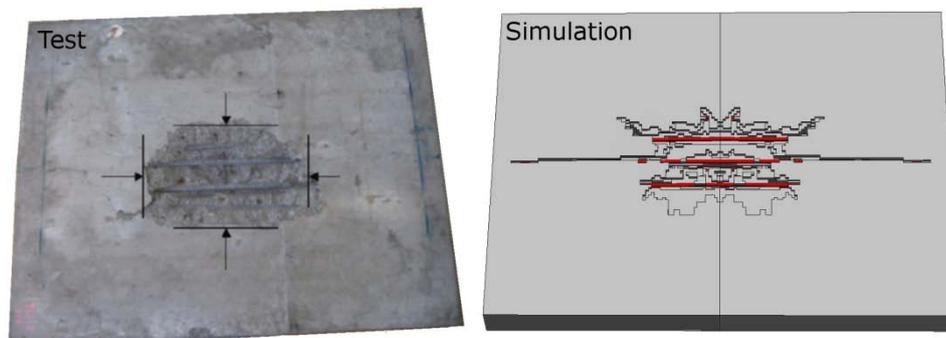


Figure 3: Comparison of damage to the bottom of the concrete slab observed in the experiment (image reproduced with permission)^[9] and simulation ($\alpha=0.1$, $\delta_{max}=1.90$).

3.2 Semi-confined explosions inside RC structures

The semi-confined explosion tests consisted of two slabs with different thickness separated by columns at the corners. An explosive charge of TNT was placed in contact with the bottom slab. A complex arrangement of three rebar types (different steel strengths and bar diameters) provided reinforcement in the structure. Using the embedded rebar approach, the complete rebar configuration was constructed and modified independently from the concrete structure in the pre-processing. This makes it considerably faster to develop the RC model compared to the more traditional approach of constructing a single part containing both the steel and concrete materials and adjusting element sizes to match the rebar size. The simulations utilized two planes of symmetry and were run on 4 to 16 CPU of an SGI-Altix system.

The test structure was subjected to a moderate explosive loading which produced spalling, a small through-hole in the top slab, and a larger hole in the bottom slab. For this case, the simulations were able to predict the extent of damage and relative size of the damaged areas. The simulation results are shown at 5 ms, (Fig. 4c), after the blast loading and damage to the structure was complete.

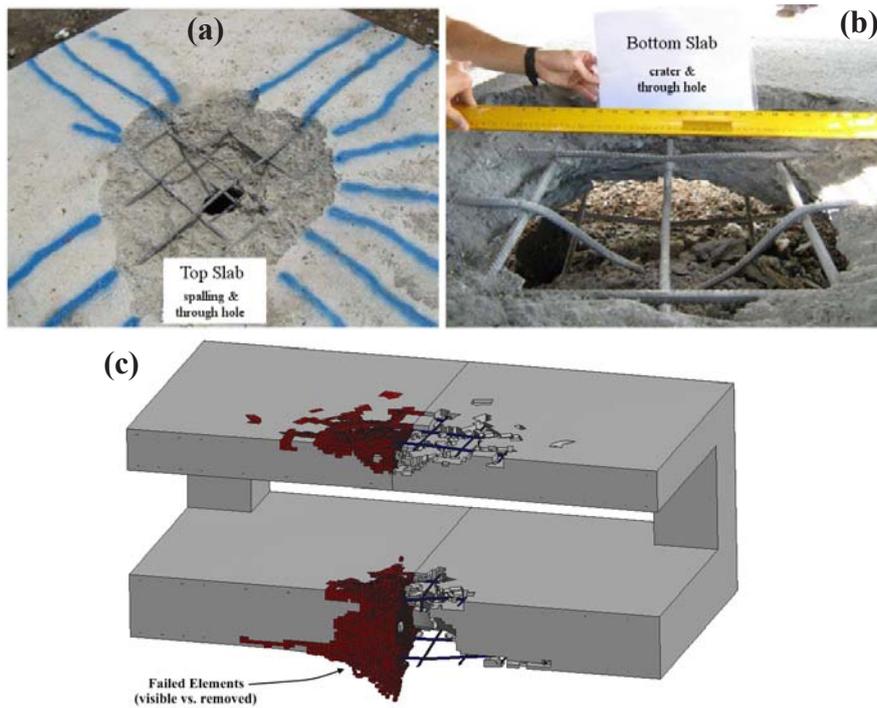


Figure 4: Damage comparison for a moderate explosive loading scenario. Test results (a,b) and simulation with visualization for half of the structure (c).

Next, a more severe explosive loading scenario was investigated which consisted of a larger charge in a similar structure having an additional edge beam around the top slab. The edge beam provided more confinement for the blast which increased the structural loading. In this case, there was extensive damage to the structure with the entire top slab removed and columns destroyed. There was also a through-hole produced in the bottom slab and a large amount of deformation and damage to the rebar. Initial results of the simulations (5ms) show a reasonable prediction of the bottom slab through-hole size while the top slab and columns remain intact. However, simulations run beyond this time have a large amount of failed elements without the removal of the top slab or rebar damage observed in the test (Fig. 5a).

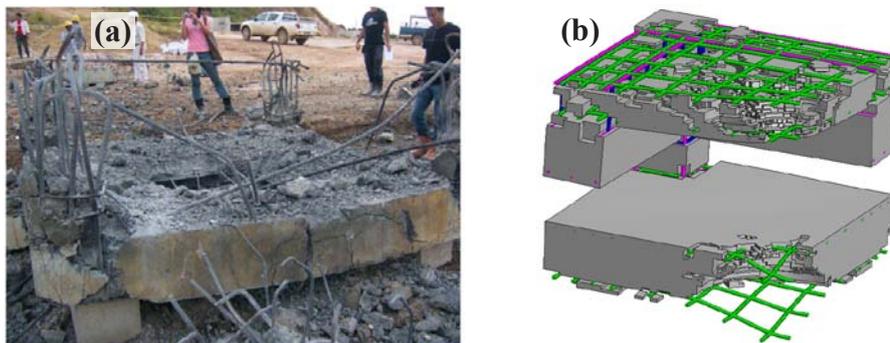


Figure 5: Damage comparison for a severe explosive loading scenario. Test results (a) and simulation at 5ms (b).

3.3 Treatment of failed concrete elements

Once a concrete element has failed, the internal nodes are duplicated and the element is allowed to fly free from the structure. The failed CSD elements can either be invisible to the CFD or have their surfaces sent to the CFD to interact with the flow field. Passing the failed concrete CSD surfaces to the CFD solver can become very computationally expensive as the number of failed elements increases during simulation, and is usually not necessary.

In the case of the severe loading or a confined blast, the gas can escape through the holes created in the concrete and prematurely reduce the structural loading. Two simulations of the severe loading scenario were used to compare the difference when failed CSD elements were invisible to the flow and when they interact with the flow. The differences in the blast pressures and inclusion of the failed CSD surfaces can be observed in Figure 6. As the number of failed concrete elements increases, the difference in computational time becomes apparent. At 5ms, the speed of the simulation with failed CSD elements in the CFD (Fig. 6b) was reduced to approximately 50% of the original (Fig. 6a) on the same number of CPU. Interestingly, the final damage observed in the two simulations was essentially identical. This implies that the pressure reduction due to the hole formed in the concrete is not significant to the overall structural loading.

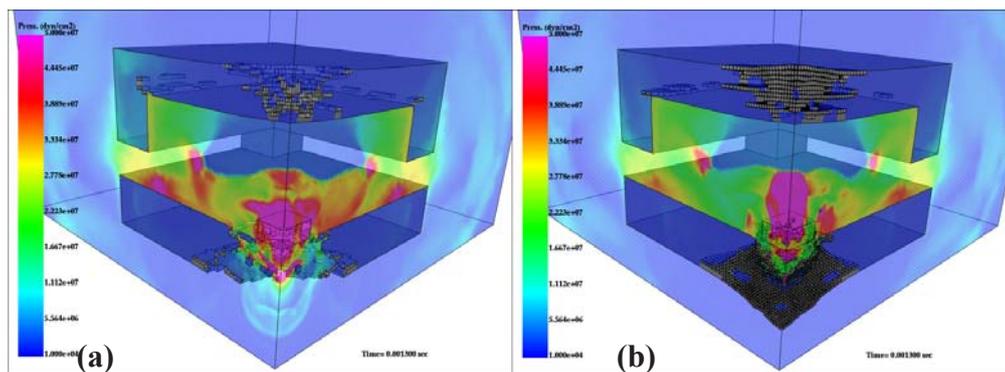


Figure 6: Pressure contours on the symmetry planes and CSD surfaces passed to the CFD solver for simulations without (a) and with (b) failed concrete elements.

In actual RC structures, the rebar provides a large amount of strength for the structure and acts as a cage keeping the damaged concrete fragments from flying apart. Using the embedded rebar approach, once the concrete element containing a rebar beam is failed then the exposed rebar does not interact with the CSD elements or CFD flow. That is, the failed concrete elements will “float” through the exposed reinforcement bars. This lack of interaction between failing structural components could contribute to the under-prediction of damage observed in the simulations for the severe loading scenario. A possible solution to this limitation is to use solid elements for the rebar inside the concrete, but at increased model complexity and computational expense.

5 CONCLUSIONS

This work has demonstrated the ability of the loose-coupled CFD/CSD method with a solid-embedding technique for simulating explosions near RC structures. The primary

observations made through this work are:

- Preliminary simulations for validation and parameters studies are essential in determining the best set of solver, material, and mesh parameters.
- Simulations for the moderate explosive loading of a concrete slab and a semi-confined RC structure have good agreement with damage observed in tests.
- Simulations for the severe explosive loading of a semi-confined RC structure are initially able to predict the extent of damage, but under-predict the amount of damage to the structure and rebar at later times.
- Including failed CSD concrete elements as individual flying bodies in the CFD flow increases computational cost significantly, and does not noticeably alter the predicted damage to the structure.
- Treatment of rebar by embedding beam elements in solid concrete elements method does not account for the interaction of exposed rebar with other CSD elements or CFD pressures.

REFERENCES

- [1] Soto, O.A., J.D. Baum, C. Charman, and R.L. Löhner. Coupled fluid/solid scheme to deal with blast and impact loads over reinforced concrete structures. *Comp. Meth. for Coupled Problems, COUPLED PROBLEMS 2009*. Ischia, Italy. June (2009).
- [2] Löhner, R., C. Yang, J. Cebral, J.D. Baum, H. Luo, D. Pelessone and C. Charman. Fluid-Structure Interaction Using a Loose Coupling Algorithm and Adaptive Unstructured Grids. *AIAA-95-2259* (1995).
- [3] Baum, J.D., E. Mestreau, H. Luo, R. Löhner, D. Pelessone and C. Charman. Recent Development of a Coupled CFD/CSD Methodology Using an Embedded Approach. *Proc. 24th Inter. Shock wave Symposium*. Beijing, China. July (2004).
- [4] Baum, J.D., H. Luo, R. Löhner, C. Yang, D. Pelessone and C. Charman. A Coupled Fluid/Structure Modeling of Shock Interaction with a Truck. *AIAA-96-0795* (1996).
- [5] Malvar, L.J., J.E. Crawford, J.W. Wesevich, and D. Simmons. A plasticity concrete material model for Dyna3D. *Int. J. Impact Eng.* (1997) 19(9–10):847–73.
- [6] Malvar, L.J. and Crawford, J.E. Dynamic Increase Factors for Concrete. *28th DDESB Seminar*. Orlando, FL. Aug. (1998).
- [7] LS-DYNA keyword user's manual, version 970. Livermore Software Technology Corporation. April (2003).
- [8] Malvar, L. J. and Crawford, J. E. Dynamic Increase Factors for Steel Reinforcement Bars. *28th DDESB Seminar*. Orlando, FL. Aug. (1998).
- [9] Zhou, X.Q., V.A. Kuznetsov, H. Hao, and J. Waschl. Numerical prediction of concrete slab response to blast loading. *Int. J. Impact Eng.* (2008) 35:1186-1200.

INVESTIGATING THE EFFECT OF ROTATIONAL DEGREE OF FREEDOM ON A CIRCULAR CYLINDER AT LOW REYNOLDS NUMBER IN CROSS FLOW

SEYED HOSSEIN MADANI*, JAN WISSINK, HAMID BAHAI

School of Engineering and Design - Brunel University West London

*Islamic Azad University – South Tehran Branch
email: Hossein.Madani@brunel.ac.uk

Key words: circular cylinder; Vortex-Induced Vibrations, rotational d.o.f.

Summary. Numerical simulations of Vortex-Induced Vibrations (VIV) of a circular cylinder in cross flow with a rotational degree of freedom about its axis have been carried out by means of a finite-volume method. The study is performed in two dimensions at a Reynolds number of $Re_D = 100$, based on the free stream velocity and the diameter, D , of the cylinder. The effect of the rotational degree of freedom on the cylinder's lift and drag forces are compared with the baseline simulation results of flow around a stationary cylinder. The introduction of a rotational degree of freedom (d.o.f) is observed to cause the lift and drag forces to change. Also, the pattern of vortex shedding behind the cylinder is found to drastically change when the cylinder is allowed to rotate.

1. INTRODUCTION

The study of flows around cylinders has a long history [1, 2]. The early studies were focussed on the flow around a stationary cylinder at various Reynolds numbers. Subsequently, investigations have been carried out of flow around a cylinder with a prescribed rotational velocity [2]. Although the study of flows around rotating circular cylinders is not new, most of the previous works consider the rotational speed as a parameter that can be used to decrease the effect of vortex shedding on the cylinder. In other words, angular velocity is viewed as a way to reduce the root mean square of the lift force. This is the reason why rotation of the cylinder is used in feedback control of wakes [3]. Forced oscillatory rotation of a circular cylinder has also been investigated numerically as well as experimentally [4]. All of these studies focused on imposed oscillatory angular velocities. Etienne and Fontaine [5] studied the effect of vortex shedding on a two dimensional cylinder with two spatial d.o.f. They observed that the cylinder was mainly oscillating transversely and slightly in line with the flow. When they added a rotational degree of freedom, for an arbitrary rotational moment of inertia, the transverse amplitude of oscillation was found to be reduced by a factor of two, while the mean in-line deflection was also found to decrease by a factor between 1.5 to 2. In their case, the Magnus effect was found to be negligible as the maximum angular velocity was only on the order of 5% of the free-stream velocity U [5].

In this study, we evaluate the effect of introducing a rotational degree of freedom, on the flow around a circular cylinder. To achieve this we introduce the rotational angle of the cylinder as an unknown that is affected by friction-induced torque.

Below, we present some of the results obtained by performing a parametric study in which the moment of inertia (I) and the rotational spring rigidity (k) are varied. The spring rigidity is used to control the rotational degree of freedom (d.o.f) and both k and I determine the natural frequency (f) of the system. An important parameter in this context is the so-called reduced velocity, $U_r = U/fD$, where U is the free-stream velocity and D is diameter of the cylinder. It should be noted that both k and I are defined for a unit-length cylinder.

2. GENERAL SPECIFICATION

The cylinder was allowed to rotate about its axis. The rotation was controlled by adding a torsional spring with stiffness K . Without the presence of the spring the cylinder was observed to rotate rigidly in one direction.

Because of the simplicity of the problem and the low Reynolds number, we were able to model the set up as a two-dimensional flow problem. The computational domain is shown in figure 2. At the inlet, the flow is assumed to be uniform with $u=U_0$ and $v=0$, where u and v are the velocities of the flow in x -direction and y -direction respectively. A free-slip boundary condition is applied along the upper and lower boundaries while a convective outflow boundary condition is applied at the outlet. At the surface of the cylinder, finally, a no-slip boundary condition is prescribed.

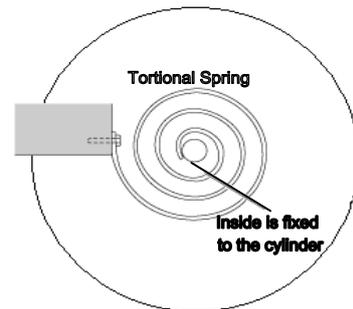


Figure 1 : Cylinder with Rotational degree of freedom, a Torsional spring- stiffness of K and the moment of Inersial of I

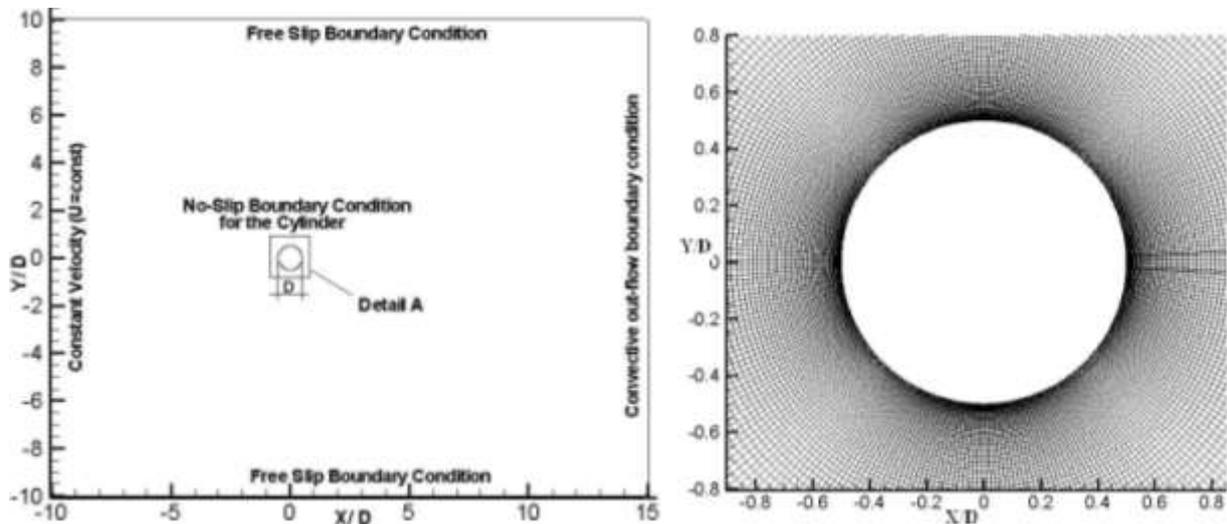


Figure 2: Left: The computational domain showing the boundary conditions. Right: Zoomed view of the O-mesh close to the cylinder corresponding to “Detail A” at the left

This problem is an example of a rotational harmonic oscillator that can oscillate about the axis of the cylinder. This behaviour is analogous to linear spring-mass oscillators. The general equation of motion is given by:

$$I \frac{d^2\theta}{dt^2} + C \frac{d\theta}{dt} + \kappa\theta = \tau(t) \quad (1)$$

If the damping is small, $C \ll \sqrt{\kappa I}$, as is the case in this study, the frequency of vibration is very close to the natural resonance frequency of the system:

$$f_n = \frac{\omega_n}{2\pi} = \frac{1}{2\pi} \sqrt{\kappa/I} \quad (2)$$

In the absence of a driving force ($\tau = 0$), the general solution of the resulting homogeneous problem is given by: (3)

$$\theta = Ae^{-\alpha t} \cos(\omega t + \phi)$$

Where:

$$\omega = \sqrt{\omega_n^2 - \alpha^2} = \sqrt{\kappa/I - (C/2I)^2} \quad (4)$$

Table 1 : Definition of terms in the equations

Definition of terms		
Term	Unit	Definition
θ	Radians	Angle of deflection from rest position
I	kg m ²	Moment of inertia
C	kg m ² s ⁻¹ rad ⁻¹	Rotational friction (damping)
κ	N m rad ⁻¹	Coefficient of torsion spring
\mathcal{T}	N m	Drive torque
f_n	Hz	Undamped (or natural) resonance frequency
ω_n	rads ⁻¹	Undamped resonance frequency in radians
f	Hz	Damped resonance frequency
ω	rads ⁻¹	Damped resonance frequency in radians
α	s ⁻¹	Reciprocal of damping time constant
ϕ	Rad	Phase angle of oscillation
L	M	Distance from axis to where force is applied

The angular velocity of the cylinder is determined by a numerical approximation of eq. (1) with C=0, using an Euler scheme for the integration of time. The torque $\tau(t)$ is calculated every time step by integrating the tangential frictional forces of the flow on the cylinder.

3. NUMERICAL RESULTS

For this simulation the LESOCC flow solver has been used. LESOCC has been developed at the Institute of Hydromechanics at Karlsruhe, Institute of Technology, Germany. In Wissink and Rodi [6] it has been extensively tested for the simulation of flow around a cylinder at $Re=3200$. LESOCC uses a second-order accurate discretization of the convection and diffusion, combined with a three-stage Runge-Kutta method for the time-integration. It uses a collocated variable arrangement combined with momentum interpolation to avoid a decoupling of the pressure and velocity fields.

For the present study, a mesh independency test was carried out and, as a result, a mesh with (360×126) points in the circumferential and radial direction respectively was chosen. Numerous runs have been carried out on the computing cluster at Brunel University. To simulate each case using 8 processors it takes nearly 2000 hours for the results to converge. Figure 3 shows how the results converged for one specific case. The Reynolds number was kept constant at $Re=100$ for all cases. To initiate the vortex shedding, we applied a random perturbation to the flow. **Figure 4** shows that the results are not dependent on the initial perturbation.

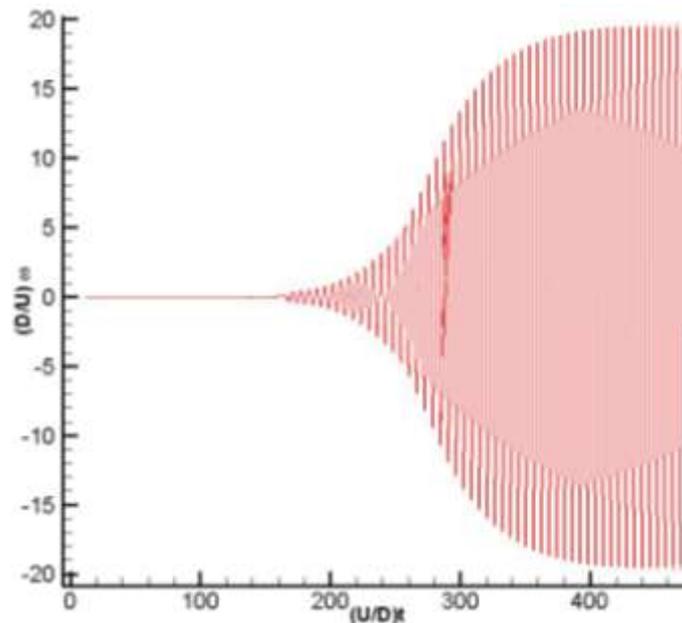


Figure 3: The results of the simulation Converges- $Re=100$, $I=0.333$, $K=0.3648$ $Ur=6$

In this study, the effect of a rotational degree of freedom on vortex shedding and lock-in phenomena was investigated and the results were compared with flow around a stationary cylinder. Initially an attempt was made to perform a simulation with a rotational d.o.f without any restoring force ($K=0$). As a result, the cylinder was observed to rotate in only one direction. It was therefore decided to add a restoring force by the introduction of a rotational spring ($k>0$). To establish which moment of inertia, I , would be relevant to our problem, we assumed a solid cylinder with the same density of water (1000kg/m^3). The diameter of the cylinder was chosen to be 20 cm; as a result $I = (1/8)mD^2 = (1/32)\pi\rho D^4 = 0.157\text{ kg/m}^2$. (For the

case with $D=0.2\text{m}$ and a density of water equal to 1000 kg/m^3 , the equivalent moment of inertia for the cylinder becomes $I=0.157\text{ kg.m}^2$) and the corresponding non-dimensional moment of inertia becomes $I=0.5$. Figure 5 on the left shows the effect of inertia of the cylinder on the frequency of the vortex shedding of the cylinder (with a constant rotational stiffness $k=0.05$). For high (low) amounts of inertia the frequency decreases (increases) dramatically. Figure 5, right, depicts the relation between K/I and the natural frequency of the system and the frequency of the rotational velocity (ω). This graph clearly proves equation (2). The power of the K/I is 0.5045, which is almost the same as predicted by theory (0.5) and the coefficient is 0.1598 which is very close to the coefficient $1/2\pi=0.1591$ in eq. (2).

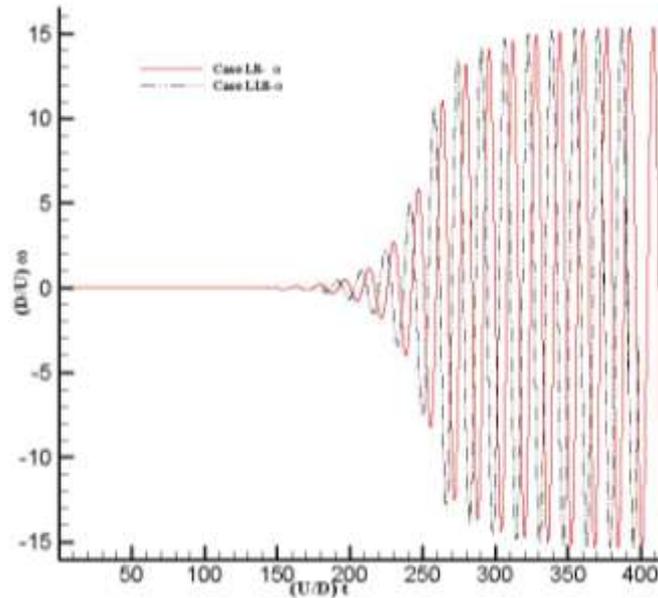


Figure 4: The effect of changes in random perturbation on the convergence of the results for two similar cases- $Re=100, I=0.333, K=0.05, Ur=16$

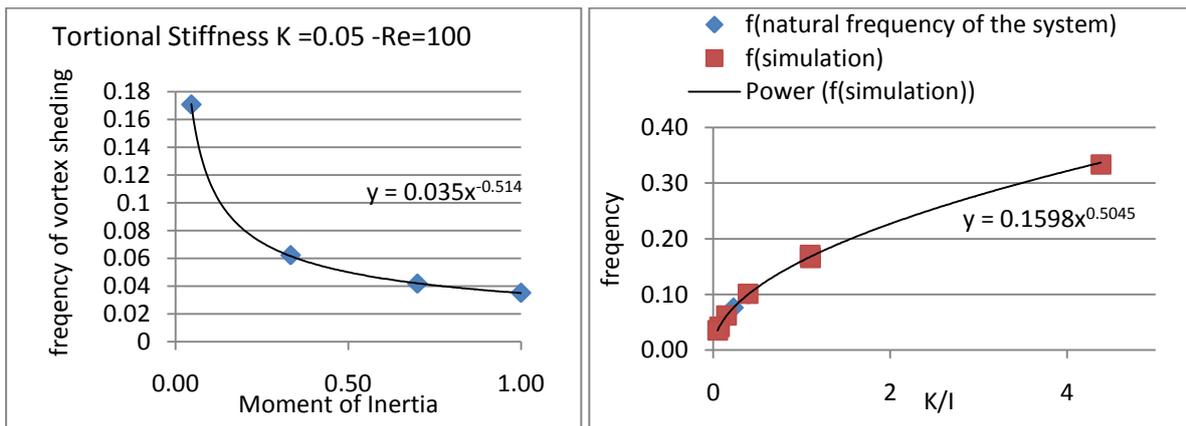


Figure 5 –left: Natural frequency verses inertia when $K=0.05$ constant. Right: natural frequency and vortex shedding frequency verses (K/I) , the parameters are non-dimensional.

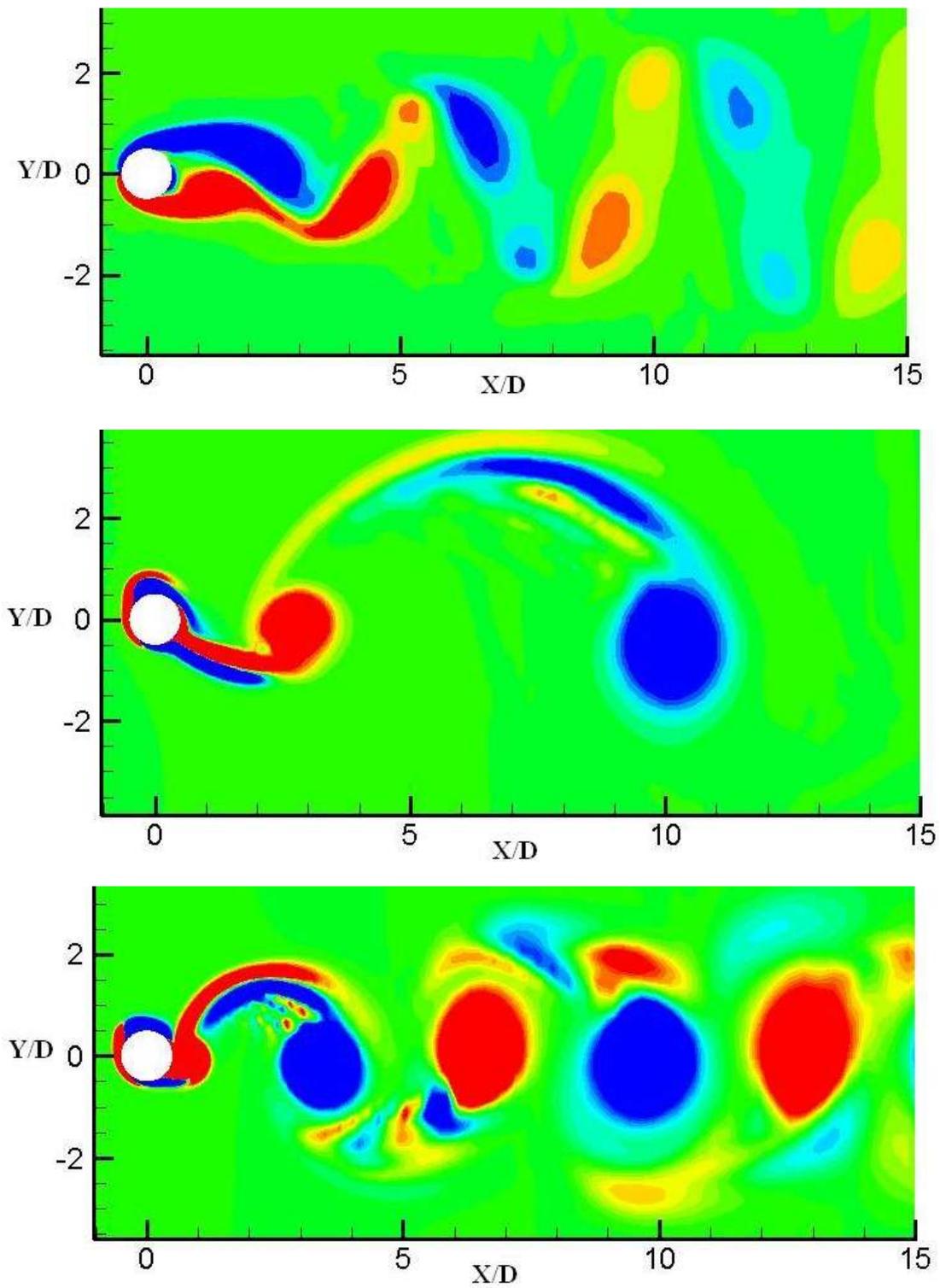


Figure 6 : Vortex Shedding $Re=100$ - first case: stationary cylinder, $St=0.167$, Second case $I=0.333$ and $K=0.05$, $Ur=16$, $St=0.0625$, third case: $I=0.333$, $K=0.365$, $Ur=6$, $St=0.167$

Figure 6 shows vortex shedding for 3 different cases in a two dimensional flow at $Re=100$. The first one is the stationary cylinder; the frequency of the vortex shedding for this case is 0.17. These result exactly match the results of Roshko [7], who measured the frequencies using a hot-wire velocity probe. For the low Reynolds number laminar region Roshko condensed his results to an equation of the form $St = 0.212 (1 - 21.2 / Re)$ where St is the Strouhal number $S=FD/U$ [7]. The second and third pictures in figure 6 show the vortex shedding from a cylinder with a rotational degree of freedom. For the cases $I=0.333$ combined with $k=0.05$ and $k=0.365$, the frequencies of vortex shedding become $f=0.0625$ and $f=0.167$ respectively. In figure 7, the effect of rotational d.o.f was compared with the stationary cylinder. Etienne and Fontaine [5] observed that the introduction of a rotational degree of freedom causes a reduction in the vortex-induced vibration in the transverse direction with the flow [5]. It implies that we should expect a lower lift when we have a rotational d.o.f. in combination with spatial degrees of freedom. In the absence of a spatial degree of freedom, our results show a completely different behaviour and predict a significant increase in unsteady lift forces acting on the cylinder due to the Magnus effect.

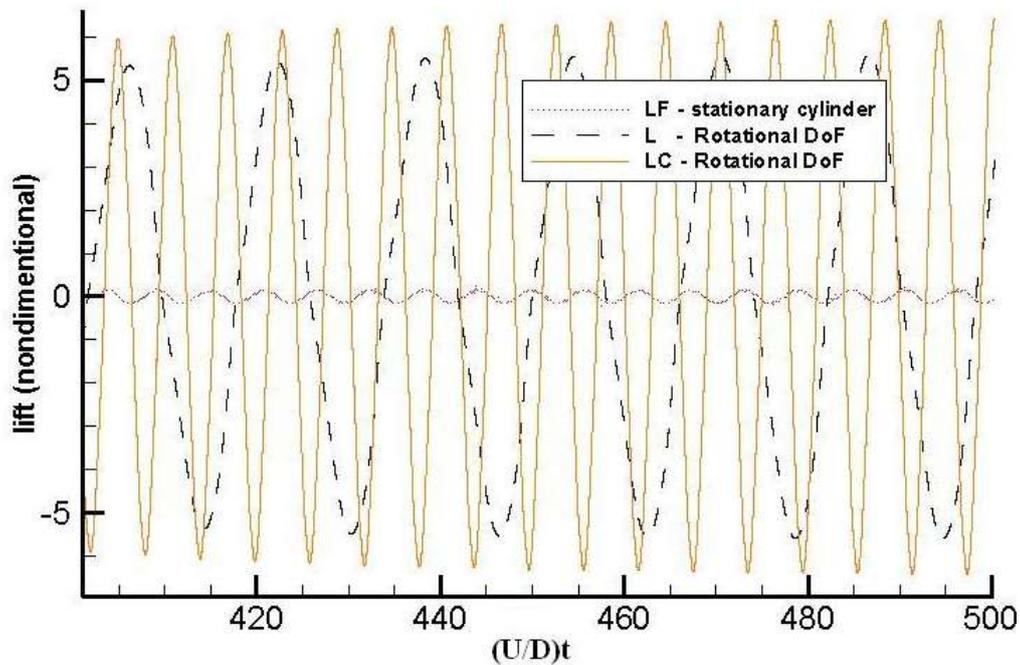


Figure 7- lift for three cases. first case : stationary cylinder, second case: Rotational d.o.f $I=0.333$ and $K=0.05$, third case: Rotational degree of freedom $I=0.333$, $K=0.365$

4. Conclusion

The introduction of a rotational degree of freedom which allows the cylinder to rotate about its axis, has a significant effect on the pattern of vortex shedding at low Reynolds numbers. In all cases considered, the vortex shedding locks-in to the natural frequency of the inertial/spring system. Compared to the baseline simulation of flow around a stationary cylinder, the addition of a rotational degree of freedom to the cylinder was observed to

significantly increase unsteady lift forces because of the Magnus effect, while also the drag forces were not diminished.. In the near future, we aim to complete the present parametric study of the effects of inertia/spring stiffness on the flow pattern and the lift and drag forces.

References

- [1] C.H.K. Williamson and R. Govardhan, “ Vortex- Induced Vibration”, *Annu. Rev. Fluid Mech.*(2004). 36:413-55.
- [2] T.Sarpkaya, “A critical review of the intrinsic nature of vortex-induced vibrations”, *Journal of fluids and structure* 19 (2004) 389-447.
- [3] Lee S B, Baek S J, Sung H J (2009). Feedback control of a circular cylinder wake with rotation oscillation, *Fluid Dyn. Res.*, **41**, 1-24.
- [4] Lam K M (2009). Vortex shedding flow behind a slowly rotating circular cylinder. *Journal of Fluids and Structures*, **25**, 245–262,
- [5] S. Etienne, E. Fontaine, “ Effect of Rotational degree of freedom on Vortex-Induced Vibrations of a circular cylinder in Cross-Flow”, *Proceedings of the international offshore and polar engineering conference*, 3, pp. 1089-1093, 2010.
- [6] J.G. Wissink, W. Rodi, “Numerical study of the near wake of a circular cylinder”, *International journal of heat and fluid flow*, 29(2008) 1060-1070.
- [7] Roshko, Anatol (1954). On the drag and shedding Frequency of two dimensional bluff bodies, NACA technical note 3169.

MODELING OF AIRBLAST PROPAGATION THROUGH AN ENCLOSED STRUCTURE

JESSE A. SHERBURN^{*,†}, DONALD H. NELSON[†], CAREY D. PRICE[†]
AND THOMAS R. SLAWSON[†]

[†] U.S. Army Corps of Engineers

Engineer Research and Development Center

3909 Halls Ferry Rd., Vicksburg, MS 39180, USA

e-mail: {jesse.a.sherburn, donald.h.nelson, carey.d.price, thomas.r.slawson}@usace.army.mil

Key words: Fluid-Structure Interaction, CTH, DYSMAS, Confined Airblast

Abstract. *The ability to model explosively formed shock waves propagating through a structure is of particular interest to engineers concerned with structural responses to blasts. Accurate peak pressure and impulse values are critical to understanding blast loads on structures and predicting the resulting structural behavior, but are sometimes difficult to determine analytically. Experiments are necessary to determine the true structural response, but the experiments alone may not identify all the details involved in an explosive event that may be important for design purposes. When tied to experiments, computational modeling of explosive events can be an invaluable tool for an engineer. The most difficult part of modeling structural response to a close-in explosive event is capturing the fluid-structure interaction of the resulting flow of the detonation products. In this paper, we compare the results of numerical simulations of an explosive experimental event in an enclosed structure, or “attic space”, using two different computational codes, CTH and DYSMAS. Both adequately model the explosive event in attic space when compared to the experiment. We also compare the two codes’ ability to produce explosive-induced pressure-time histories in the free field. The advantage of using a coupled code like DYSMAS is that structural response can be more accurately captured than by using a hydrocode like CTH alone. The differences between the two codes’ ability to model the event are analyzed and described as well as a general description of the shock wave propagation in the attic space.*

1 INTRODUCTION

Designing enclosed structures to withstand explosively-generated shock waves presents many difficulties for the military structural engineer. The original studies on confined blast started in World War II, but the first useful study published was that of Weibull [1] in 1968 that correlated peak quasi-static pressure versus the charge weight for a series of TNT charges detonated within a vented enclosure. Since then numerous studies were conducted to understand confined blast on structures [2-6]. Some recent studies [7-8] have numerically modeled confined explosive events for understanding the structural loading and mitigation techniques.

An engineer needs to know the blast pressure-time history, or as a minimum the pressures and impulses a structure will need to withstand, in order to produce an adequate design. Analytical methods work for simple structures; but in complicated structural configurations, it can be difficult or impossible to predict localized loading throughout the structure. With experimental validation, computational models have the ability to aid the engineer in understanding explosively-generated loads on a complicated structure. Many different tools exist such as LS-DYNA, AUTODYN, ABAQUS, and others, but for this study CTH and DYSMAS were used to perform simulations on a structure undergoing close-in airblast. CTH, an advanced Eulerian hydrocode developed by Sandia National Laboratories for the purpose of modeling materials under large deformations, was used for analysis of shock propagation through the structure [9]. In contrast to CTH, DYSMAS is a coupled Eulerian-Lagrangian code that allows the fluid and the structure to be modeled separately. DYSMAS was originally developed by the United States Navy for modeling under water explosions on structures [10-11]. This study performed simulations on an enclosed structure using CTH and DYSMAS and compared the results to experimental data. Simple free field calculations in open air were performed to compare CTH's and DYSMAS's ability to propagate airblast. This paper shows the merits of using a coupled fluid-structure code like DYSMAS for this type of scenario.

2 MODEL SETUP

2.1 Experimental setup

Figure 1 shows the experimental scale model of the structure investigated in this study. The structure consisted of a roof deck suspended 0.188 m above a ceiling deck, forming an attic space between. The structural members (frames and purlins) were simple rectangular cross sections. A mild steel was used in order to produce a near non-responding reaction structure in which to measure the pressure wave as it propagated and reflected through the attic space. The final dimensions of the structure measured 2.3 m in width, 0.91 m high at the ridgeline, and 2.5 m in length. A 0.17 m circular hole was cut (shown in Figure 1) in the upper roof to simulate the breached area in the roof and to remove the complexity of having to model the effects of breaching on the blast propagation into the attic space, and a spherical RDX-based explosive was placed just above the center of the hole and flush with the roof line. The amount of explosive used in this study was 36.45 g.



Figure 1: Elevation and plan view of experimental structure.

The pressure-time histories were measured by pressure gages on the top of the ceiling deck as shown in Figure 2. The only pressure gage that will be discussed in this study is gage 12 which is located on top of the ceiling deck in the center of the structure, as shown in Figure 2. Figure 2 also shows the charge location relative to gage 12.

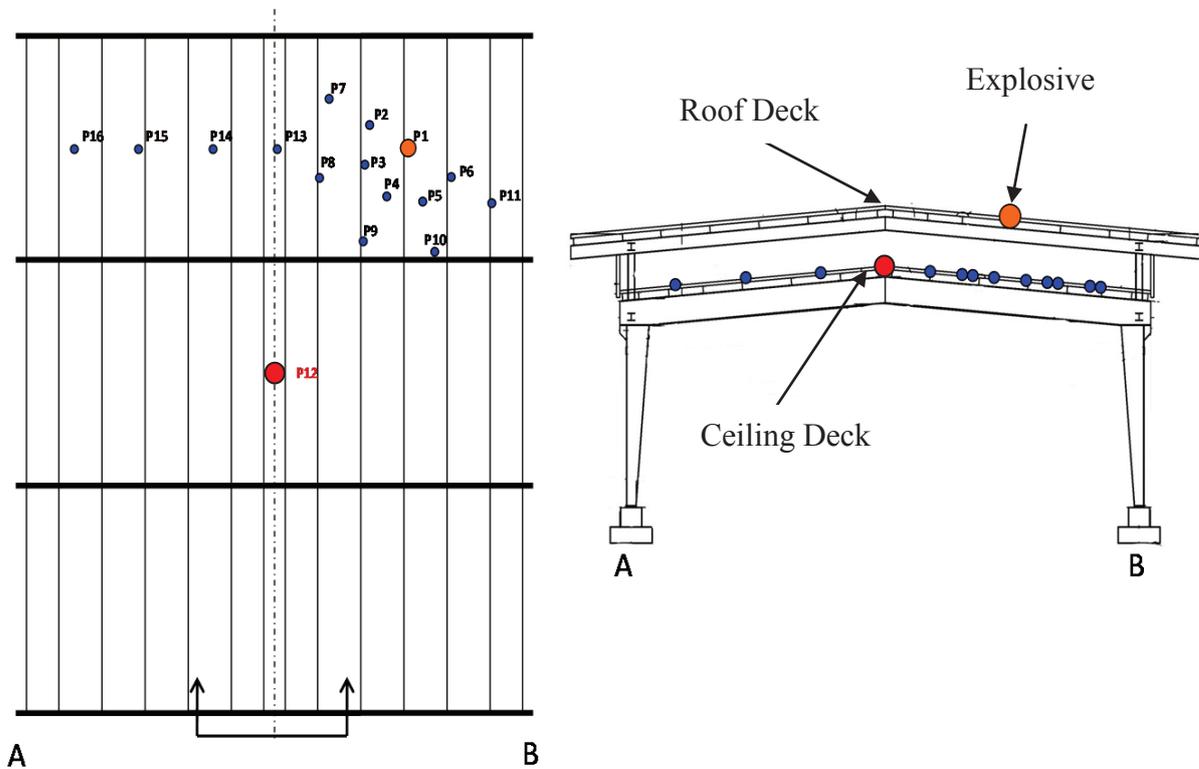


Figure 2: Plan and elevation views of gage layout. The large orange circle represents explosive location. The gage under consideration is in the center of the structure on top of the ceiling deck (large red dot). Note the simulation setup is mirror image of this gage layout.

2.2 CTH model

The CTH model incorporated a simplified upper portion of the experimental structure discussed in section 2.1. Figure 3 displays a three-dimensional view of the CTH model setup with the hole and the spherical charge clearly shown. The Jones-Wilkins-Lee (JWL) equation-of-state (EOS) was used to model the charge, and the structure's material used a SESAME tabulated EOS with a simple elastic-plastic von mises yield model for mild steel. The JWL EOS parameters are listed in Table 1. Air was also included using the SESAME tabulated EOS in CTH in order to properly propagate the pressure wave through the structure. The total size of the CTH model was identical to that of the experimental structure discussed in section 2.1. The domain size was limited to 750 x 750 x 200 cells, or 3.75 m x 3.75 m x 1.0 m. The sides and top boundary conditions were transmitting boundaries while the bottom boundary was a wall condition. A mesh sensitivity study was conducted to determine an adequate mesh size for accuracy and computational efficiency. The study determined that a uniform mesh with a cell size of 0.005 m for a total of 112.5 million cells over the domain fulfilled the balance between accuracy and efficiency. Stationary tracer particles were placed in the same

gage locations as those in the experimental model shown in Figure 2. The simulation was conducted in two phases. First, a one-dimensional simulation propagated the initial pressure wave up until the time just before it would impinge on the structure. Secondly, the one-dimensional results were rezoned and placed into the three-dimensional domain and the simulated propagation was allowed to continue. Stationary tracer particles at corresponding gage locations recorded the pressure-time histories throughout the simulation.

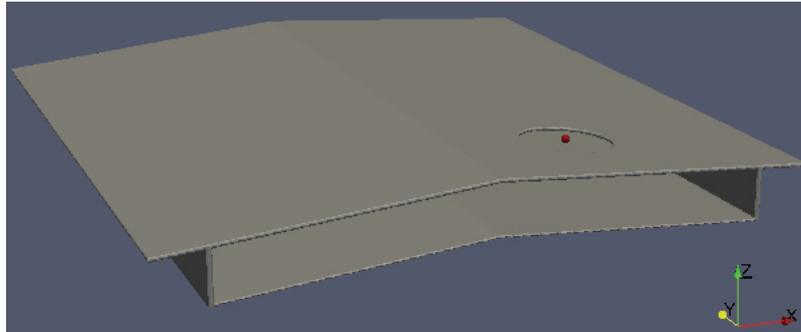


Figure 3: Simplified simulation of the attic space structure. Note the location of the charge above the hole.

Table 1: JWL parameters used for modeling the explosive.

Constant	Value
Density (g/cm ³)	1.601
AG (MPa)	6.0977×10^5
BG (MPa)	1.295×10^4
Omega	0.25
R1	4.5
R2	1.4

2.3 DYSMAS model

The DYSMAS model included the same simplified attic space as the CTH model. The main difference between the CTH model and the DYSMAS model was the DYSMAS model had the solid and fluid parts separated. The fluid part, or Eulerian domain, was where the air and the explosive detonation products were modeled. The size of the Eulerian domain was identical in size to the CTH model described in section 2.2. The domain boundaries were identical to the CTH simulation. Air was modeled using a gamma law EOS, and the explosive was modeled with the JWL parameters in Table 1. The solid part, or Lagrangian structure, was modeled using a near uniform finite element mesh. Figure 4 displays the mesh which consisted of about 30,000 shell elements with each element having an average side length of 0.02 m and a shell thickness of 0.0127 m. The same material model for mild steel was used for the structure. The Lagrangian structure was then placed in the Eulerian domain in the same configuration as the CTH model. DYSMAS requires the elements that were in contact with the fluid domain must be specified as a wetted element. This specification of wetted elements allows DYSMAS to handle the coupled fluid-structure interaction when a pressure wave and/or detonation products impinges on a structure. Just like the CTH calculation a one-dimensional Eulerian simulation was performed for the initial expansion of the pressure wave and detonation products and then was rezoned into the three-dimensional Eulerian domain.

The simulation was then allowed to continue to completion. Stationary tracer particles at analogous CTH locations were placed in the Eulerian DYSMAS domain. An added benefit to DYSMAS was an element's history near a pressure gage location could also be recorded to measure exactly what pressure loads the structure endured.

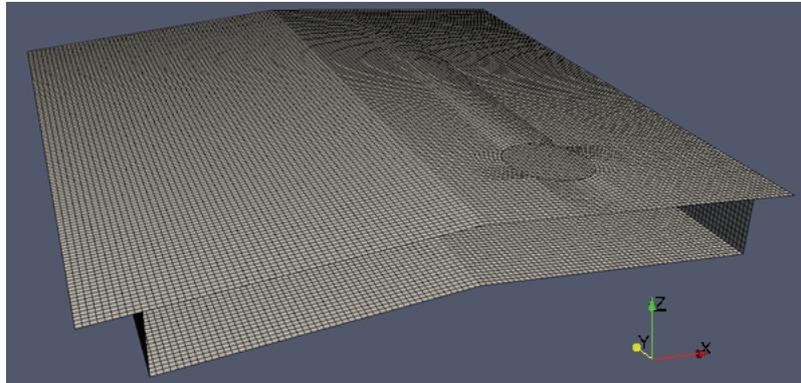


Figure 4: Lagrangian mesh used for the DYSMAS calculation.

2.4 Free field model

Simple free field simulations were conducted in both CTH and DYSMAS in order to compare the propagation of a pressure wave through air. This was to ensure they could be compared directly. The two-dimensional axisymmetric domain consisted of a 1.0 m x 2.0 m size with the explosive at the center. The same explosive weight was used (36.45 g) in the free field simulations as the previously described CTH and DYSMAS calculations. The uniform cell size was identical to the CTH and DYSMAS calculations (0.005 m) and used the same rezoning technique. The rezoning in the free field case was done from a one-dimensional domain to a two-dimensional domain. A single stationary tracer particle was placed 0.5 m away from the explosive's center and recorded the pressure-time history.

3 RESULTS

3.1 Free field results

The free field simulations compared adequately between CTH and DYSMAS for this scenario. Figure 5 shows the pressure-time history at the tracer particle for both CTH and DYSMAS. The arrival times were comparable with CTH arriving slightly earlier and the CTH peak pressure was also slightly smaller in magnitude. These differences were expected because CTH and DYSMAS used a different EOS for air. The SESAME tabulated EOS used by CTH is more realistic (i.e., more physics included) than the gamma law, but for the purposes of this study the similarity between the results was satisfactory.

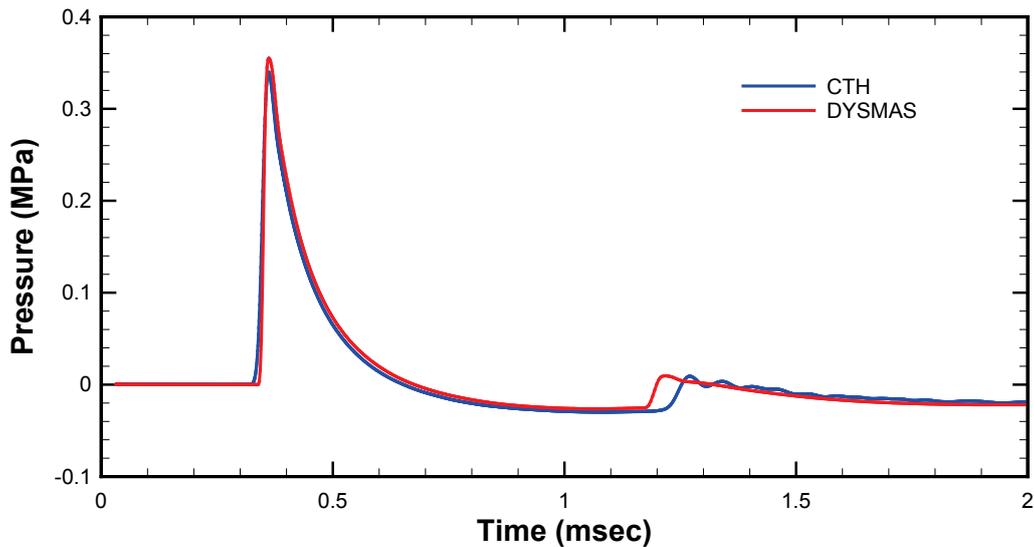


Figure 5: Free field pressure-time histories for CTH and DYSMAS located 0.5 m away from charge center.

Figure 6 shows the pressure field contours for the CTH and DYSMAS free field simulations at time 1.0 msec. The pressure contours were near identical in size and shape. CTH included more fluctuations in pressure due to the presence of more Richtmyer-Meshkov instabilities between the detonation products and the air. The DYSMAS simulation did have some of the instabilities, but not as many as CTH. The second pressure wave was produced by the creation of a low pressure void in the expanding detonation products that eventually collapsed the detonation products inward. The collapse then rebounds to produce the small pressure peak shown in Figure 5 around 1.2 msec. When the second pressure wave passes through the instabilities, the resultant pressure field has a non uniform look as shown in Figure 6a. The DYSMAS detonation products did not have as many instabilities, so the second pressure wave had a near circular shape. The differences between the two simulations were minimal which gave confidence that comparing the loading on the structure would be valid.

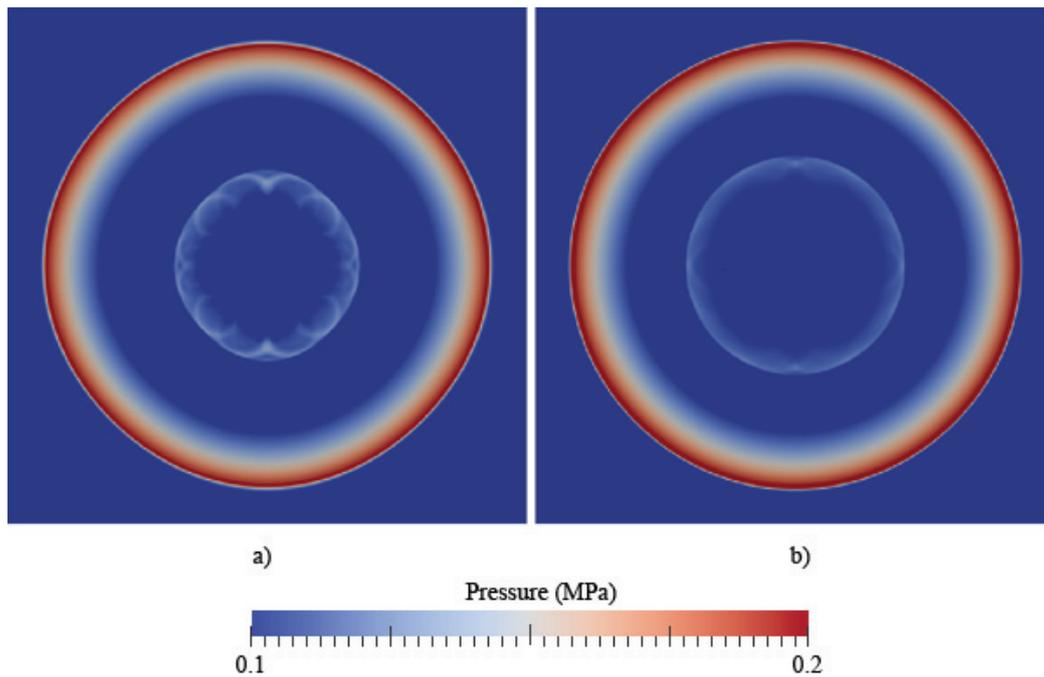


Figure 6: Free field pressure contours at time 1.0 msec for a) CTH and b) DYSMAS.

3.2 CTH and DYSMAS results

The overall response of the simplified CTH and DYSMAS simulations was adequate when compared to the experimental data. There were some differences noted in the initial peak pressures and subsequently the impulse on the structure. Figure 7 shows the detonation products at time equals 2.0 msec. The simulations have similar shapes but have some distinct differences. The overall expanse in the detonation products were the same size but interacted differently on the structure's surface. The detonation products in the CTH calculation tended to adhere to the structure's surface more than the DYSMAS calculation. The DYSMAS detonation products slid fairly well compared to CTH along the roof's surface, as shown in Figure 7b. Inside of the attic space the same adhering happened in CTH, and the detonation products did not get pushed by the sidewall reflected pressure wave as much as the DYSMAS calculation.

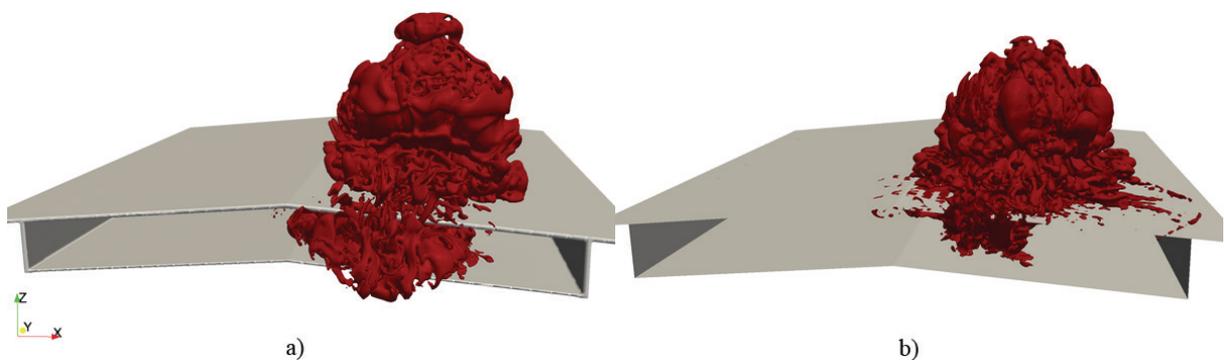


Figure 7: Detonation products at time 2.0 msec for a) CTH and b) DYSMAS.

The pressure field contours at time equals 1.15 msec for the CTH and DYSMAS calculations are shown in Figure 8 for a cross section in the center of the structure. The pressure wave on the top of the roof surface looked identical in both CTH and DYSMAS cases. Inside the attic space the pressure waves had similar magnitudes, but the arrival times of the inner reflected wave were quite different. Some of this was due to the fact that DYSMAS uses a shell element with an assumed thickness for response purposes which made the attic space cross section area slightly larger than the CTH case as shown in Figure 8. This would make the wave reflect quicker in the CTH case.

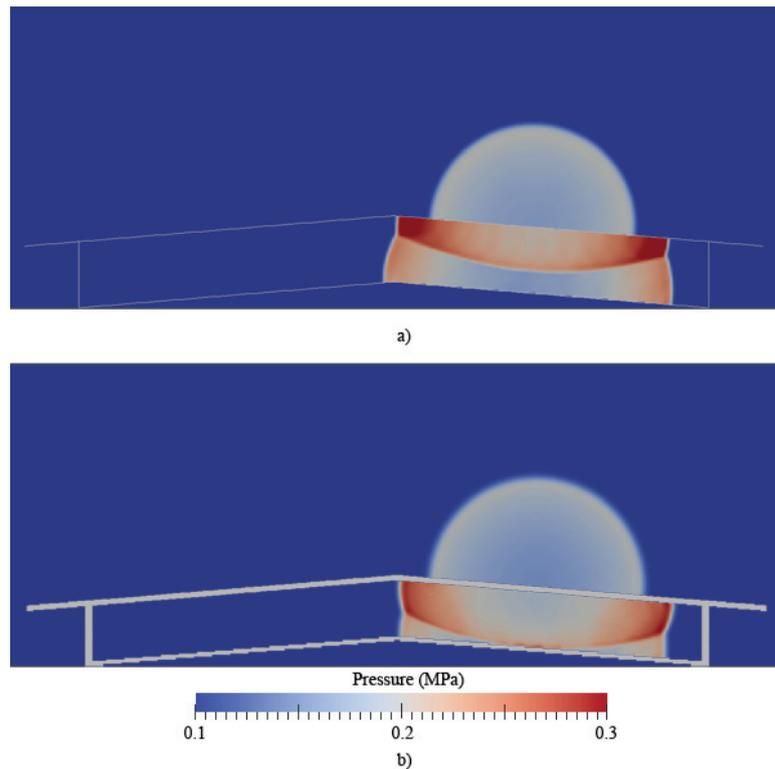


Figure 8: Pressure contours at time 1.15 msec for a) DYSMAS and b) CTH.

Figure 9 displays the pressure-time histories for CTH, DYSMAS, and the experiment at gage 12. Both of the CTH and DYSMAS simulations adequately predicted the character of the pressure-time history of the experiment data. Both the CTH and DYSMAS calculations look smoother because there were no interior frame or purlin members to cause the complex reflections observed in the experimental data, as shown in Figure 9. CTH nearly matched the experimental initial peak pressure at 6.7% higher than the experimental data, while DYSMAS undershot the peak pressure data by 20%. A large difference between the experimental data and the models is the presence of a high secondary peak which shows up in the experimental data at time 2.3 msec and between 2.5 msec and 3.0 msec in CTH and DYSMAS, as shown in Figure 9. The difference between the models and the experiment is clearly shown when impulse-time history is plotted as shown in Figure 10. This secondary peak causes the overall impulse to be much higher in the experimental data, which is the reason the simulations deviate from the experimental impulse shown in Figure 10. The explanation for this

difference could be threefold. First the simplified model in this study may not capture a large reflection off of an interior member that is present in the experimental data. A second explanation is that afterburning effects were not included in the hydrodynamic simulations which would account for the extra pressure in the secondary pressure wave shown in Figure 9. Research has shown afterburning effects in RDX-based explosives can be quite substantial in any enclosed or semi enclosed environment [8]. A third explanation is the JWL EOS did not adequately capture shock propagation through preshocked air which implies the second pressure wave in the simulations did not propagate as quickly as the experimental test. More research needs to be done in order to quantify this difference for this particular structure.

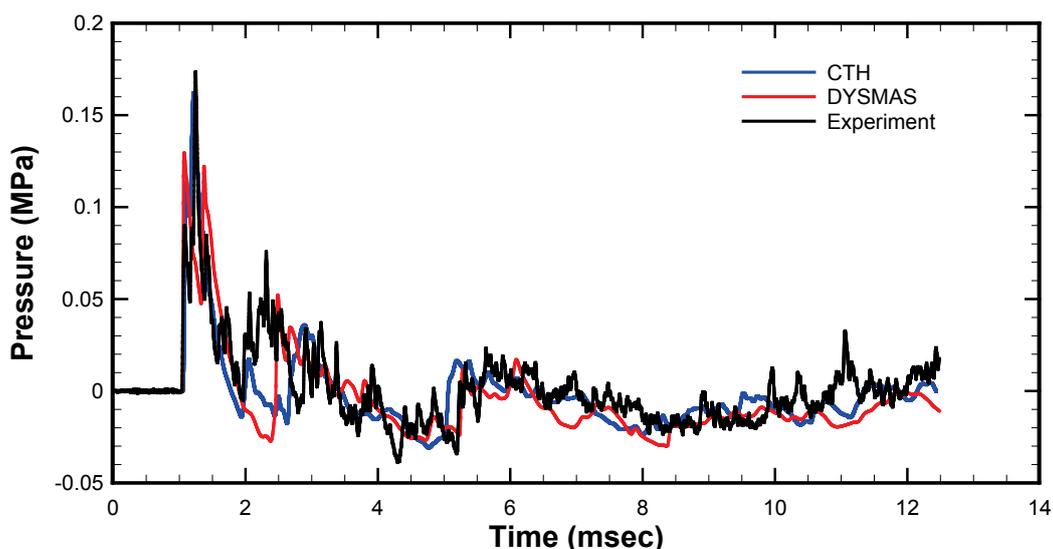


Figure 9: Comparison of pressure-time histories at gage 12 for CTH, DYSMAS, and experiment.

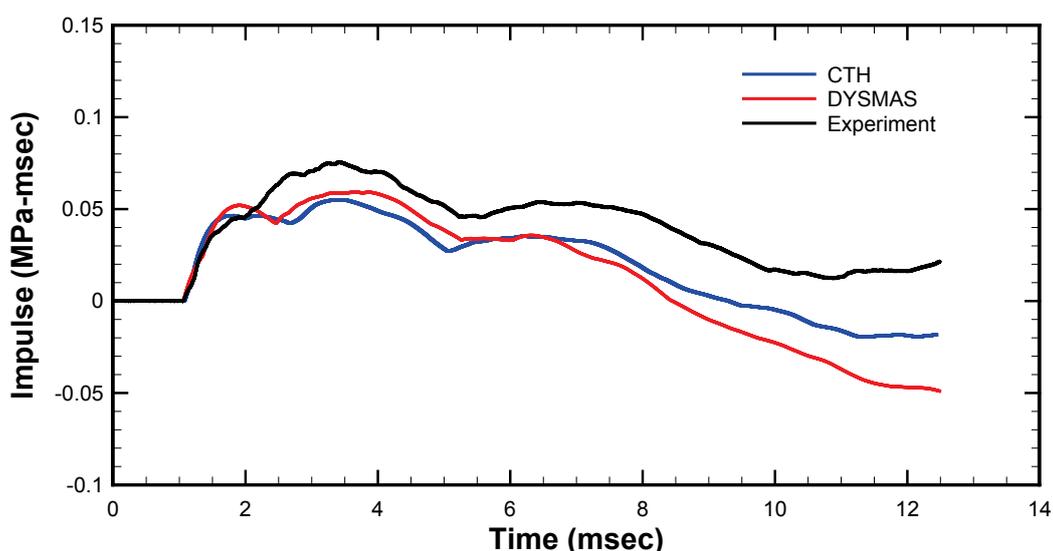


Figure 10: Comparison of impulse-time histories at gage 12 for CTH, DYSMAS, and experiment.

One of the advantages of using a coupled code like DYSMAS is the ability to extract relevant data on the structure being loaded from the finite element mesh. Extracting this data from CTH is impossible due to the fact that the mesh is fixed in space. Figure 11 compares the experimental data, the data extracted from the Eulerian domain tracer, and the data from the element closest to the gage location. The structure's element has about a 12% lower peak pressure than the Eulerian domain tracer which indicates the structure actually feels less of the pressure wave than the Eulerian domain tracer reports. The two data points track similar trends until late time when the structure's element tracks the increasing pressure trend past 8 msec better than the Eulerian domain tracer. Figure 12 displays the matching trend by the impulse-time history. The experimental impulse in earlier times, 2-4 msec, is much higher due to the presence of the larger secondary pressure wave, but the structure's data tracks the experimental impulse trend far better than the Eulerian domain tracer. The Eulerian domain tracer data never turns upward in late times (past 8 msec) like the structure's element and the experimental data. This highlights the fact that the structure responds differently than a tracer in the Eulerian domain reports, which shows the importance of modeling the fluid and the structure separately.

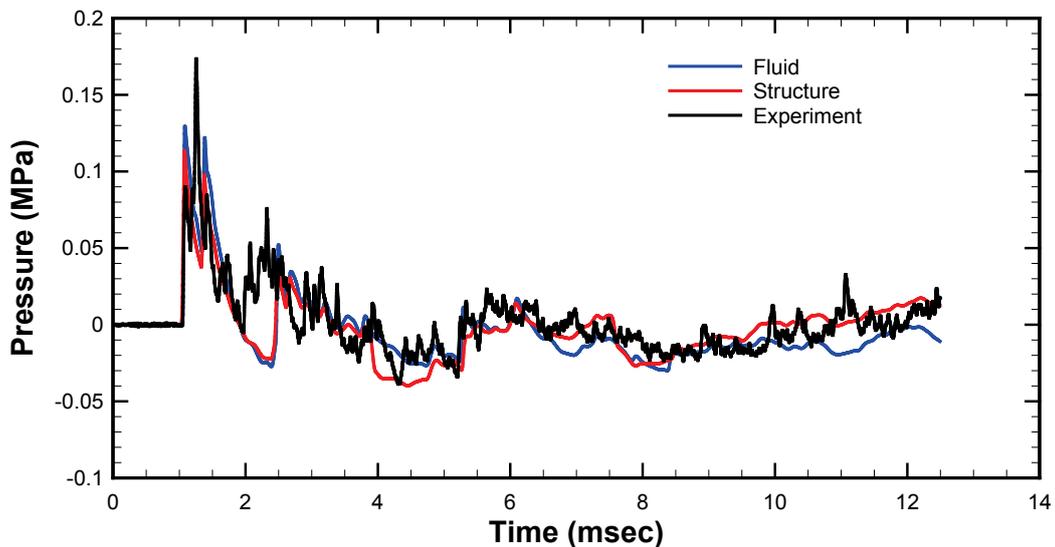


Figure 11: Comparison between fluid and structure pressure-time histories at gage 12 for the DYSMAS calculation.

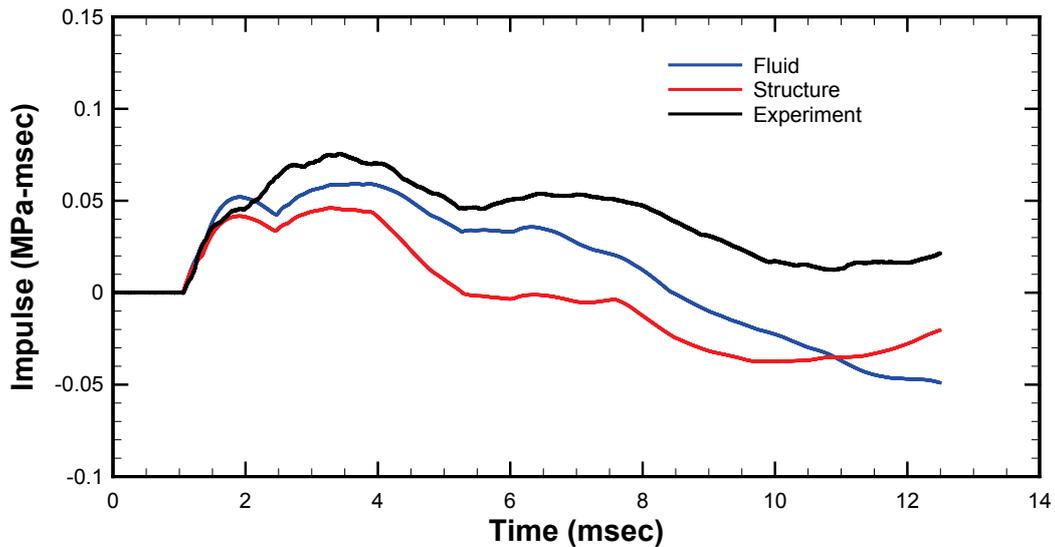


Figure 12: Comparison between fluid and structure impulse-time histories at gage 12 for the DYSMAS calculation.

4 CONCLUSIONS

For the scenario presented in this paper, CTH and DYSMAS both propagated the pressure wave through the attic space adequately compared to the experimental data. DYSMAS had the advantage of more accurately modeling the structure's response to the explosively-generated pressure wave. The comparison of the experimental pressure-time history and impulse-time history results to the simulations allows the design engineer to confidently design a real responding structure with the simulation data. Now that the simulations had success matching the experimental data, different loadings could be tested in DYSMAS on the structure that may be helpful for the design engineer.

This was an initial effort in modeling the attic space that can be extended in complexity in a few different ways. First, the internal members need to be modeled in order to capture the complex reflections present in the attic space. Second, the effect of afterburning needs to be investigated and modeled in order to quantify its effects on the pressure wave and impulse. A third area of future study is the possible EOS issues in the preshocked attic space. A final area of study is to use DYSMAS to study the roof breaching effects on the attic space in a full scale event.

5 ACKNOWLEDGEMENTS

Permission to publish was granted by the Director, Geotechnical & Structures Laboratory. All simulations were performed on Department of Defense Supercomputer Resource Center high performance computers.

REFERENCES

- [1] Weibull, H.R.W. Pressures recorded in partially closed chambers at explosion of TNT charges. *Ann. N. Y. Acad. Sci.* (1968) **152**:256-261.
- [2] Baker W.E., Anderson Jr., C.E., Morris, B.L and Wauters D.K. Quasi-static pressure, duration and impulse for explosions (e.g. HE) in structures. *Int. J. Mech. Sci.* (1983) **25**:455-464.
- [3] Kingery, C.N., Schumacher, R. and Ewing Jr., W.O. Internal pressure from explosions in suppressive structures. Memorandum Report ARBLRL-MR-02848, U.S. Army Ballistic Research Laboratory, Aberdeen Proving Ground, MD (1978).
- [4] Esparza, E.D., Baker, W.E. and Oldham G.A., Blast pressures inside and outside suppressive structures. Report EM-CR-76042, Edgewood Arsenal, Aberdeen Proving Ground, MD (1975).
- [5] TM5-1300, Structures to resist the effects of accidental explosions. U.S. Departments of the Army, Navy and Air Force (1969).
- [6] Beshara, F.B.A. Modelling of blast loading on aboveground structures-II. Internal blast and ground shock. *Comp. Struct.* (1994) **51**(5):597-606.
- [7] Gelfand, B.E., Silnikov, M.V. and Chernyshov, M.V. On the efficiency of semi-closed blast inhibitors. *Shock Waves* (2010) **20**:317-321.
- [8] Togashi, F., Baum, J.D., Mestreau, E., Löhner, R. and Sunshine, D. Numerical simulation of long-duration blast wave evolution in confined facilities. *Shock Waves* (2010) **20**:409-424.
- [9] McGlaun, J.M., Thompson, S.L. and Elrick, M.G. CTH: A three-dimensional shock wave physics code. *Int. J. Impact Eng.* (1990) **10**(1-4):351-360.
- [10] Naval Surface Warfare Center (NSWC) Indian Head. *Gemini User's Manual, Release 5.00.00*. 10 May 2010.
- [11] Naval Surface Warfare Center (NSWC) Indian Head. *DYNA3D: A Nonlinear, Explicit, Three-Dimensional Finite Element Code for Solid and Structural Mechanics, User's Manual*. July 2009.

MULTIDISCIPLINARY INVESTIGATION BY FLUID-STRUCTURE- MOTION INTEGRATED SIMULATION

S. TAKAHASHI* AND N. ARAI*

* Tokyo University of Agriculture and Technology (TUAT)
2-24-16, Naka-cho, Koganei, Tokyo, Japan
e-mail: takahass@cc.tuat.ac.jp, www.tuat.ac.jp/~arailab

Key words: Fluid-Structure Interaction, Biomimetic System, Cartesian Grid method, Overset Grid Method.

Abstract. This study is devoted to realize the flight system based on biomimetic mechanisms by using multidisciplinary numerical simulation that consists of flow simulation, structural simulation and motion analysis. In this paper, the effective motion is investigated to generate sufficient force to make a flight. The flow simulation is carried out by finite difference and finite volume method in overset grid while the structural analysis is conducted by linear elastic simulation with finite element method. The coupled algorithm is constructed from the weak coupling with inner iteration. As a result, it was revealed that the fluid force is increased by the deformation of the object.

1 INTRODUCTION

The motion of the wings is essential to generate lift and thrust together for the flight of living beings. From the viewpoint of the physics, the phenomenon consists of physics of fluid, structure and motion so that one-sided study is not sufficient to investigate it precisely. Therefore, multidisciplinary analysis should be performed to reveal the coupled mechanism and characteristics. The flight mechanism of living beings has been investigated by experimental and numerical studies [1,2,3]. Especially, one of insects acting in low Reynolds number flow may be closer to the swimming in sticky fluid rather than flying in the air. Many researches are conducted to develop a micro aerial vehicle (MAV) based on the insects because of the high agility flight. These studies are attractive to realize the absolute flight envelope including complete vertical takeoff and landing (VTOL). By the way, what is occurred as the scale becomes larger and the flight speed becomes faster? Obviously, the flow acting on the wings becomes turbulent as Reynolds number becomes high. In the case of birds, the frequency of the flapping motion becomes lower and the shape of the wing section becomes closer to the airfoil with increase of Reynolds number [4,5,6]. It is thought that the trend is caused from the limitation of the muscle power and the efficiency of the wings for the forward flight. It is correct that only small birds can achieve VTOL due to the flapping motion with high frequency.

The objective of this study is to develop a manned aerial vehicle that can realize VTOL characteristics based on biomimetic flight systems. In this paper, coupled simulation is

discussed to perform parametric study for investigation of the aerodynamic characteristics. The coupled simulation consists of incompressible Navier-Stokes simulation and linear elastic structural analysis. The coupling system is constructed by weak coupling algorithm with inner iteration. In this paper, low Reynolds number flows are adopted to investigate the validity of the present solver as a primary stage of this work.

2 NUMERICAL METHODOLOGY

2.1 Numerical method for flow simulation

The flow simulation is carried out by two-dimensional incompressible Navier-Stokes equations with overset grid composed of boundary-fitted coordinate (BFC) grid and multi-block Cartesian grid. The present Cartesian grid method is called as Building-Cube Method (BCM) that is used for efficient spatial discretization based on equally-spaced Cartesian grid [7,8,9,10]. Although one of the most important points in using a Cartesian grid method is surface approximation of the object, the object is represented by BFC grid to estimate the fluid force correctly in the present study. Furthermore, the numerical dissipation can be suppressed owing to the background Cartesian grid. It can be significant to capture the hysteresis of the present flowfield to estimate the fluid force exactly. Arbitrary Lagrangian Eulerian (ALE) method is utilized in the BFC grid [11], while Eulerian method is adopted in the background BCM grid [12]. Overset grid system in the present simulation is shown in Fig. 1. Blue thick lines are boundaries of cubes including same number of meshes.

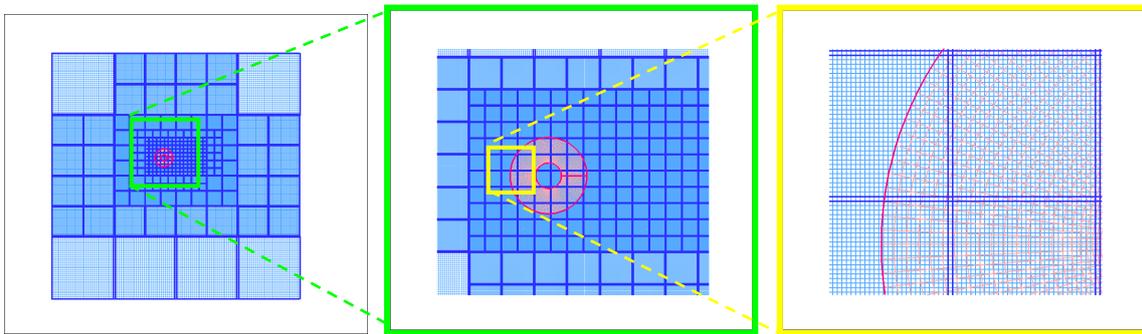


Fig. 1 Overset grid system of BFC grid and BCM grid

The governing equation in BFC grid is expressed in Eq. (1). The grid velocity term is appeared in Eq. (1) because of the ALE method.

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0 \\ \frac{\partial \mathbf{u}}{\partial t} + \left\{ (\mathbf{u} - \mathbf{u}_g) \cdot \nabla \right\} \mathbf{u} &= -\frac{\nabla p}{\rho} + \frac{\mu}{\rho} \Delta \mathbf{u} \end{aligned} \quad (1)$$

The governing equation is divided into three stages for fractional-step method [13] and discretized by finite difference method in regular grid arrangement. The convective flux and

the diffusive flux are approximated by UTOPIA scheme and second-order central difference scheme, respectively. Temporal velocity field of fractional-step method is solved by Crank-Nicolson scheme. Poisson equation of the pressure is discretized by second-order central difference scheme. Both the temporal velocity field and the pressure field are solved by SOR method. Finally, real velocity field is solved by the correction of the pressure gradient to the temporal velocity field. These procedures are implemented based on Eq. (2).

$$\begin{aligned} \frac{\mathbf{u}^* - \mathbf{u}^n}{\Delta t} + \left\{ (\mathbf{u}^n - \mathbf{u}_g^n) \cdot \nabla \right\} \mathbf{u}^* &= \frac{\mu}{\rho} \Delta \mathbf{u}^* \\ \frac{\nabla^2 p^{n+1}}{\rho} &= \frac{\nabla \cdot \mathbf{u}^*}{\Delta t} \\ \frac{\mathbf{u}^{n+1} - \mathbf{u}^*}{\Delta t} &= -\frac{\nabla p^{n+1}}{\rho} \end{aligned} \quad (2)$$

The governing equation in BCM grid is expressed in Eq. (3). In BCM grid, the equation is also divided into three stages as same as BFC grid for fractional-step method. The spatial discretization is implemented by finite volume method in collocated grid arrangement.

$$\begin{aligned} \nabla \cdot \mathbf{u} &= 0 \\ \frac{\partial \mathbf{u}}{\partial t} + \nabla \cdot (\mathbf{u}\mathbf{u}) &= -\frac{\nabla p}{\rho} + \frac{\mu}{\rho} \Delta \mathbf{u} \end{aligned} \quad (3)$$

The convective flux and diffusive flux in solving the temporal velocity field are approximated by fully-conservative second-order finite difference scheme and second-order central difference scheme, respectively. In adaptive refined Cartesian grid method, the numerical fluxes through the boundary between the different-sized grids may cause critical numerical instability. In the present method, the convective flux is calculated by using first-order upwind scheme based on the face flux estimated from Rhie-Chow momentum interpolation [14] and the average of the volume flux. The procedure is shown in the schematic as Fig. 2. The cell-averaged volume flux in large grid f_L is adopted as fluxes in fringe cells of the small grid. At the same time, the flux of a large fringe cell is calculated as $(f_{S1} + f_{S2})/2$. The face fluxes for small and large grids are evaluated as $(f_L + f_{S1})/2$, $(f_L + f_{S1})/2$ and $\{(f_{S1} + f_{S2})/2 + f_L\}/2$ to calculate the convective fluxes. In the case of the face between same-sized grids, the convective flux is calculated by fully-conservative second-order central difference scheme.

In the second stage of fractional-step method, Poisson equation is solved for the pressure that should be taken care of the treatment since the gradient of the pressure is also used to evaluate the real velocity field. The right hand side of the Poisson equation is calculated by the face fluxes based on the arithmetic average of the volume fluxes in usual collocated grid method. In the present AMR grid system, however, the face fluxes should be continuous through the different-sized grid face. The process is as same as the one for the estimation of the convective fluxes. Though the calculation process is so simple, it has much importance to

perform stable computation with fully-conservative scheme in AMR grid system.

The real velocity field is calculated by the correction using the pressure gradient to the temporal velocity field in the third stage of fractional-step method. Though the contravariant flux is also calculated by Rhie-Chow momentum interpolation at the same time, it should be treated properly at the boundary of different-sized grid. Therefore, the face flux of the large grid is estimated by the same manner as the RHS of the Poisson equation.

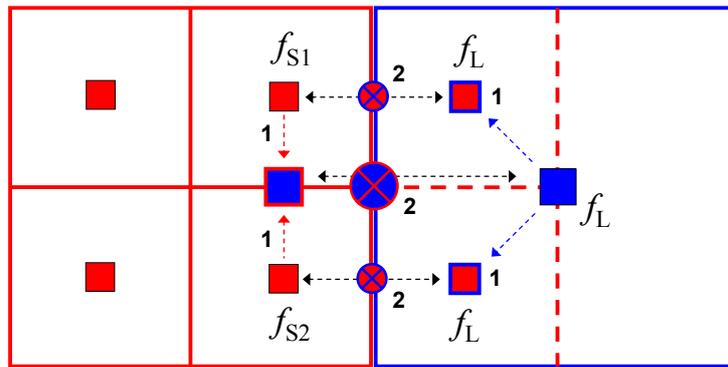


Fig. 2 Schematic of interpolation of fluxes at different-sized grid boundary

The data exchange process is necessary in overset grid system. In the present study, the relationship between a donor cell and a child node is constructed with the motion of the BFC grid repeatedly. Here, the donor cell means data supplier to the child node. The process is performed with weighted averaged interpolation based on the area-coordinate rule as shown in Fig. 3 and Eq. (4). Blue lines and nodes belong to BCM grid, while red lines and nodes belong to BFC grid. A variable of child node V_C is interpolated from weighed values by using the areas of triangles of S_{12} , S_{23} , S_{34} and S_{41} of the donor cell. Child nodes of the BFC grid are the vertices belonging in outermost cells as shown in the left schematic of Fig. 3, while one of the BCM grid are cell center points surrounded by BFC cells except the outermost nodes as shown in right schematic of Fig. 3.

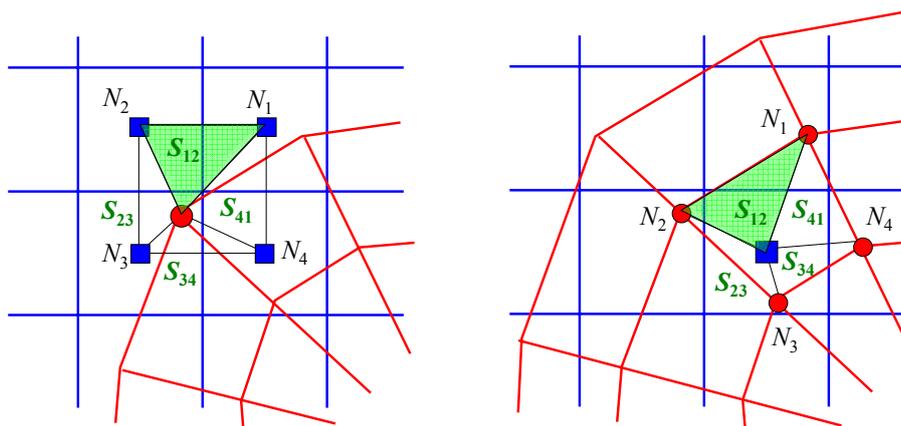


Fig. 3 Relationship between BCM donor cell and BFC child node (left) and BFC donor cell and BCM child node (right)

$$V_C = \frac{S_{23}S_{34}V_1 + S_{34}S_{41}V_2 + S_{41}S_{12}V_3 + S_{12}S_{23}V_4}{S_{12} + S_{23} + S_{34} + S_{41}} \quad (4)$$

2.2 Numerical method for structural analysis

The deformation of the object is solved by two-dimensional finite element method as linear elastic deformation with triangle elements. Shape function is first-order and the relationship between stress and strain assumes plane-stress state. The deformation at each timestep is calculated from the stiffness matrix at initial state without any stress by BiCGstab method. In this procedure, the phase error can be occurred since the fluid force is estimated around deformed object. This problem is solved by implementation of inner iteration for the coupling of the flow analysis and structural analysis as shown in Fig. 4. After the object deformation, the spatial grid for the flow simulation is deformed as the surface deformation of the object by using simple algebraic method.

2.3 Numerical method for motion analysis

The motion analysis is carried out to investigate the trajectory and the stability of the flight. However, fluid-structure coupled simulation is discussed in this paper to find out appropriate motion to achieve a vertical flight.

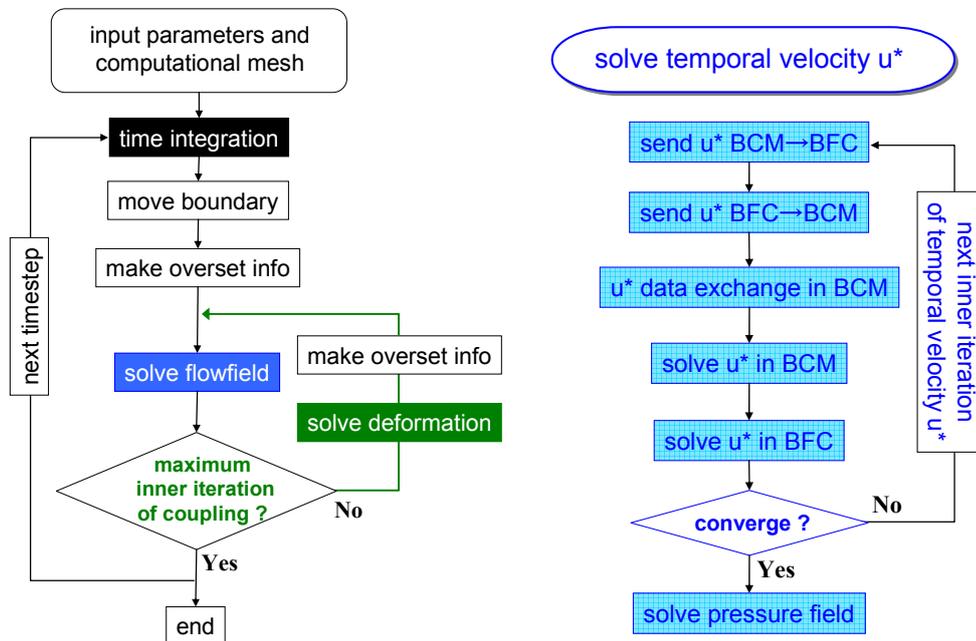


Fig. 4 Flowchart of the present coupled simulation (left) and algorithm for solving temporal velocity field (right)

2.4 Algorithm for coupled simulation

As mentioned before, the present coupled simulation is based on fluid-structure simulation to explore the optimal motion to make a flight. The flowcharts in Fig. 4 show the schematics of the present coupled algorithm and flow simulation with overset grid method. The coupled simulation is implemented by weak coupling algorithm with inner iteration. The flow simulation with overset grid method also includes inner iteration in each stage of fractional-step method to exchange data through the overlapping boundary. Furthermore, the data exchange process is basically needed for BCM grid because of the multi-block structure.

3 RESULTS AND DISCUSSION

In this paper, three kinds of flows are discussed. Firstly, validation of the present method is demonstrated by flow around a circular cylinder. Secondly, flow simulation around translating circular cylinder is discussed as an example of a flow around moving object. Finally, flow simulation around translating elastic circular cylinder is investigated by the present coupled simulation. The diameter of the circular cylinder is 0.10[m]. Freestream velocity of the validation and the maximum translational velocity are 5.0×10^{-2} [m/s]. In all cases, viscosity of fluid is assumed as 1.82×10^{-5} [Pa·s] so that Reynolds number based on freestream velocity or maximum translational velocity is around 330.

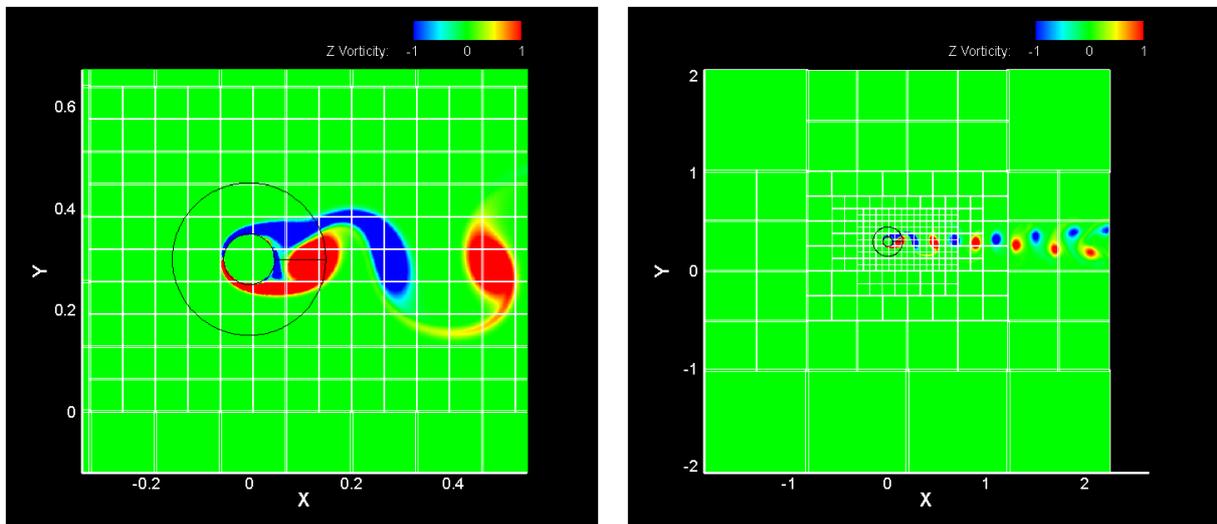


Fig. 5 Vorticity distribution at Re# 330 close to cylinder (left) and far from cylinder (right)

3.1 Flow simulation around fixed circular cylinder

The validation of the present overset flow solver is conducted by using fixed BFC grid

with a circular cylinder at Reynolds number 330. Total numbers of grid points of BFC grid and BCM grid are 4,530 (=151×30) and 274,432 (=268[cube]×32²[cell]), respectively. Uniform flow is assumed at the inflow boundary while Neumann boundary condition is used for the outflow velocity. The pressure at inflow and outflow boundaries is also determined by Neumann boundary condition. At upper and lower boundaries, slip boundary condition is used for the velocity and Dirichlet boundary condition is used for the pressure.

Figure 5 shows fully developed vorticity fields close to the cylinder and far from the cylinder. The connectivity of the vorticity through the overlapping boundary between BFC and BCM grid is sufficiently smooth. The variation of the aerodynamic coefficient is also comparable to general value. The validity of the present flow solver was confirmed from these results.

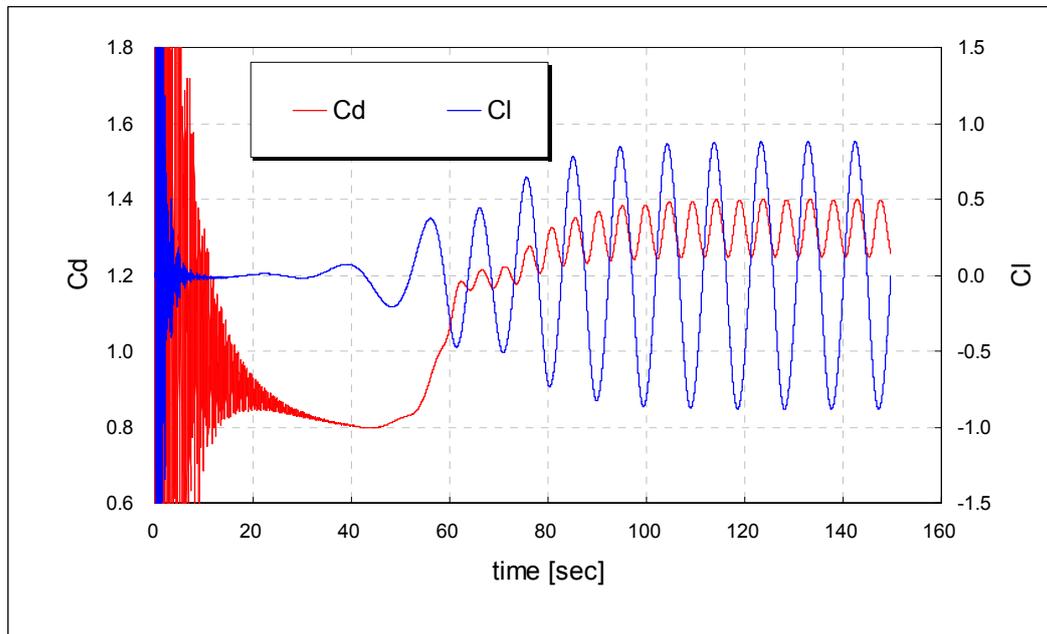


Fig. 6 Aerodynamic coefficient around circular cylinder at Re# 330

3.2 Flow simulation around translating circular cylinder

Next simulation is carried out to investigate the aerodynamic characteristics around translating circular cylinder in quiet fluid. Only BFC grid is moved backward and forward along with x -axis by force in background BCM grid. The translational motion is expressed in trigonometric function as Eq. (5).

$$x(t) = A \cos(\omega t) \quad (5)$$

The value x in the equation (5) represents the center of the circular cylinder. The variable A and ω mean the amplitude of the translation and the angular velocity. The amplitude and the angular velocity are defined as 0.10[m] and 0.50[rad/s] in the present study. The maximum

translational velocity $A\omega$ is 5.0×10^{-2} [m/s] so that Reynolds number based on the maximum velocity is around 330.

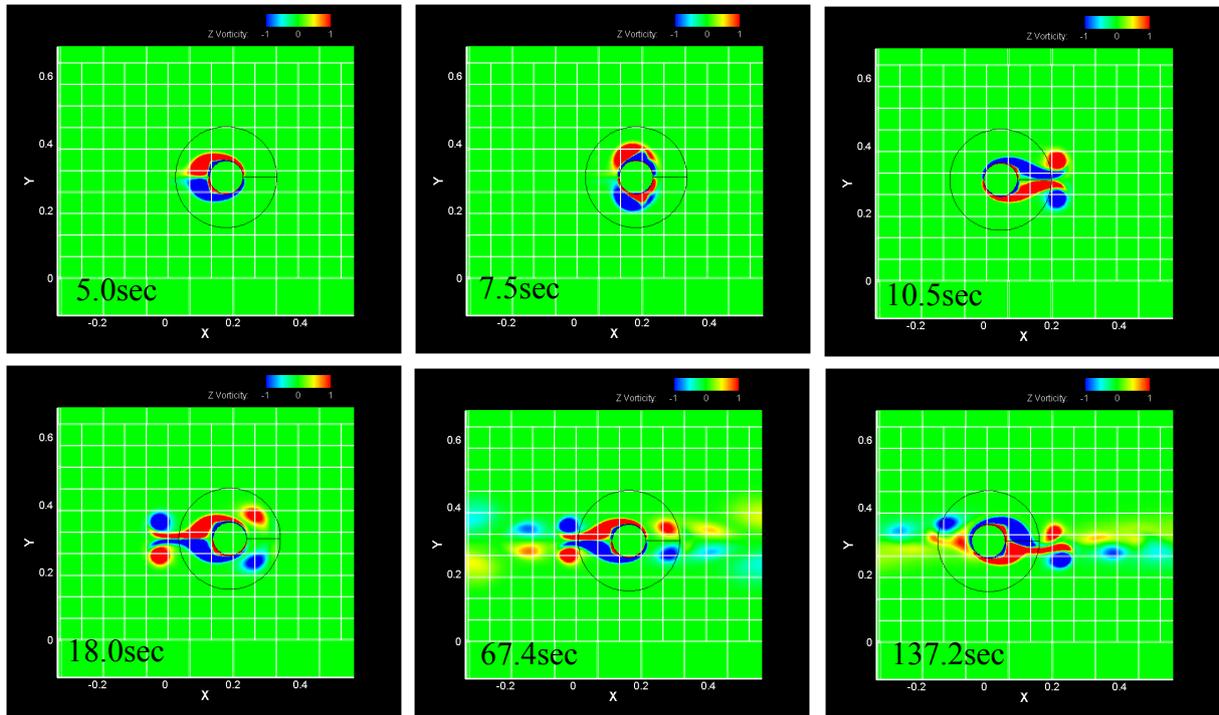


Fig. 7 Vorticity field around translating circular cylinder

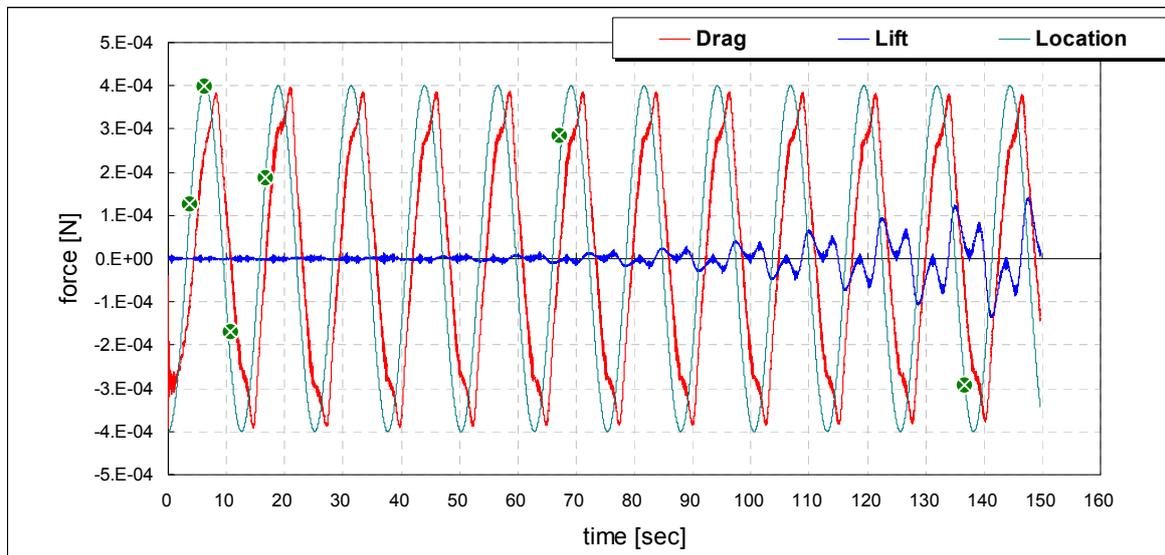


Fig. 8 Aerodynamic force around translating circular cylinder

Figure 8 shows aerodynamic forces history in that the instantaneous vorticity fields are

captured as shown in Fig. 7. The green symbols in the graph show the timings of capturing the flowfield. When the cylinder is accelerated to the right direction at 5.0 sec twin vortices are grown in the wake of the cylinder. Then, the motion is reversed at 7.5 sec to the left direction. The maximum drag force is generated just after the timing since the cylinder moves into the wake made by the previous motion. The twin vortices are left in the wake and it flows to the right direction due to the self-induced velocity at 10.5 sec. After that, vortex streets are occurred in both sides at 67.4 sec. After further, the flowfield shows asymmetrical vortex streets in both sides, gradually. Finally, complete asymmetric vortex streets are produced at 137.2 sec. The phenomenon is corresponded to the lift generation in Fig. 8. Although Karman vortex street is occurred in this Reynolds number in uniform flow as well known, the present asymmetric flow induced by the motion is similar phenomenon with Karman vortex street in self-induced flow.

3.3 Coupled simulation around translating elastic circular cylinder

Finally, the fluid force is investigated around an elastic circular cylinder by coupled simulation. Flow conditions are same as previous simulation around rigid circular cylinder. The deformation of the cylinder is solved by 3,000 grid points with 6,000 degrees of freedom with triangular elements. The computational mesh is shown in Fig. 9. Inner circle drawn in red line is treated as fixed points by the assumption that hard material core is included in the structural mesh. Though the shape of element is quadrangle in Fig. 9, it is treated as triangular element in the analysis. Young's modulus and Poisson ratio are defined as 4×10^{-2} [Pa] and 0.30 to investigate the affect of the deformation, apparently. The width of the material is 1.0[m]. The deformation is solved as linear elastic problem from the initial state at all times. Inner iteration is performed twice at each timestep to suppress the phase error from the weak coupling algorithm.

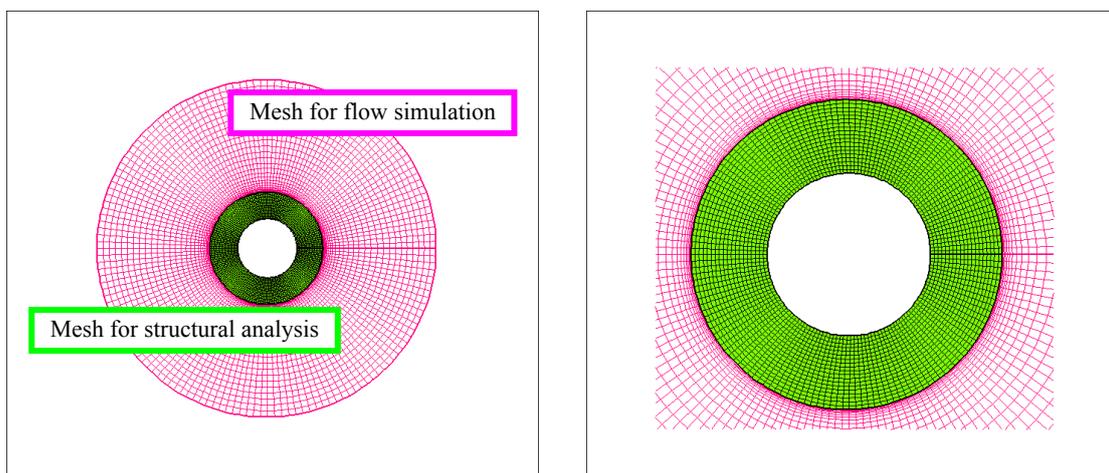


Fig. 9 Computational mesh for structural analysis

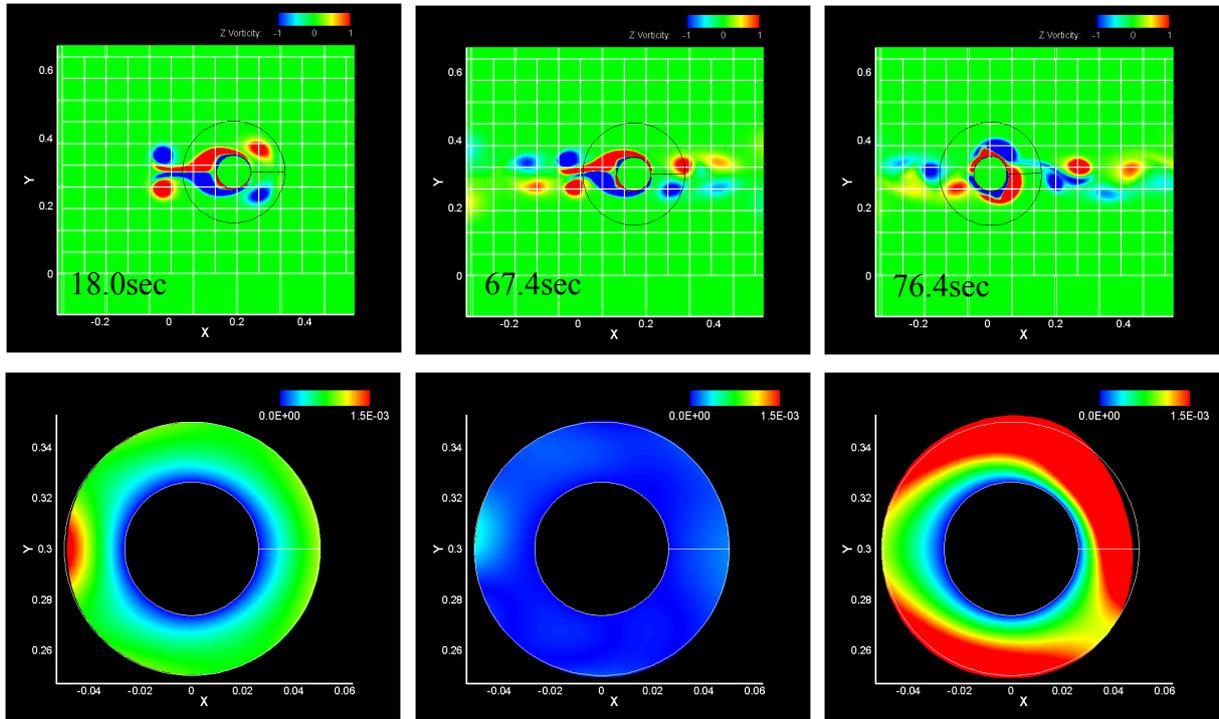


Fig. 10 Vorticity field around translating elastic circular cylinder (up) and the deformation of the elastic circular cylinder (down)

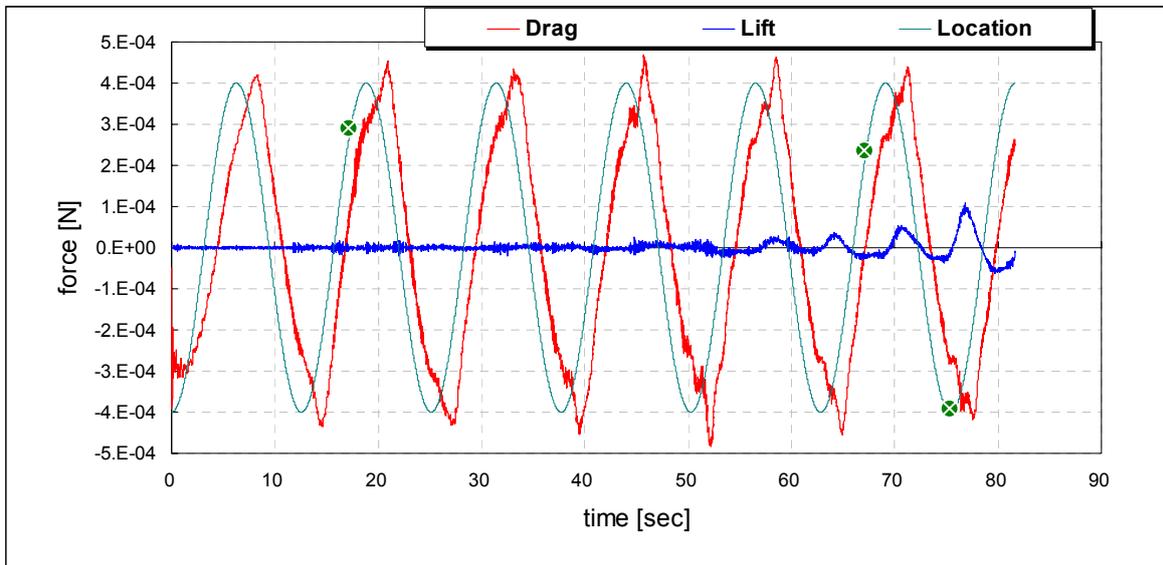


Fig. 11 Aerodynamic force around translating elastic circular cylinder

The visualizations show vorticity fields and the deformation of the cylinder at that time. The extending deformation is occurred at 18.0 sec due to the catching up of the vortices from the wake in the deceleration phase. In the significant point, the maximum fluid force is increased at the peak of drag as shown in Fig. 11. Moreover, the flowfield at 67.4 sec becomes more asymmetric than the previous result shown in Fig. 7, obviously. The visualization at 76.4 sec shows large deformation at the timing of the reversing of the motion. It is considered that the deformation of the cylinder is occurred from the wake-motion interaction. Furthermore, the passive deformation can be effective to receive (generate) the fluid force since the maximum drag is observed just after the timing in Fig. 11.

12 CONCLUSIONS

- Fluid-structure coupled solver was developed by using overset grid method based on BFC grid and BCM grid. The validity of the numerical method was confirmed by the flow simulation at Reynolds number 330.
- The flow around translating circular cylinder was solved at Reynolds number 330 based on maximum translational velocity. As a result of the simulation, twin vortices in the wake of the motion and vortex streets were observed. Moreover, it was also observed that the flowfield became asymmetric gradually.
- The coupled simulation in same flow condition with the previous one around elastic circular cylinder was carried out to investigate the effect of the deformation. Consequently, the fluid force was increased by the effect of the passive deformation by the wake-motion interaction.

REFERENCES

- [1] Liu, H. Integrated Modeling of Insect Flight: From Morphology, Kinematics to Aerodynamics. *J. Comp. Phys.*, (2009) **228**:439-459.
- [2] Wang, Z.J., Birch, J.M. and Dickinson, M.H. Unsteady Forces and Flows in Low Reynolds Number Hovering Flight: Two-dimensional Computations vs Robotic Wing Experiments. *J. Exp. Bio.*, (2004) **207**:449-460.
- [3] Bergou, A.J., Ristroph, L., Guckenheimer, J., Cohen, I. and Wang, Z.J. Fruit Flies Modulate Passive Wing Pitching to Generate In-Flight Turns. *Phys. Rev. Lett.*, (2010) **104**.148101:1-4
- [4] Greenewalt, C.H. Dimensional Relationships for Flying Animals. *Smithsonian Miscellaneous Collections*, (1962) **144**:1-46.
- [5] Takahashi, S., Yamazaki, W. and Nakahashi, K. Aerodynamic Design Exploration of Flapping Wing, Viewpoint of Shape and Kinematics. AIAA Paper, (2007) 2007-0481.
- [6] Liu, T., Kuykendoll, K., Rhew, R., and Jones, S. Avian Wings. AIAA Paper, (2004) 2004-2186.
- [7] Nakahashi, K. High-Density Mesh Flow Computations with Pre-/Post-Data Compressions. AIAA Paper, (2005) 2005-4876.

- [8] Takahashi, S., Ishida, T., Nakahashi, Building-Cube Method for Incompressible Flow Simulations of Complex Geometries. *Comp. Flu. Dynamics 2008*, (2009) **25**:473-478.
- [9] Sasaki, D., Takahashi, S., Ishida, T., Nakahashi, K., Kobayashi, H., Okabe, K., Shimomura, Y., Soga, T. and Musa, A. Large-Scale Flow Computation of Complex Geometries by Building-Cube Method. *HPC on Vect. Sys. 2009*, (2009) **4**:167-178.
- [10] Ishida, T., Takahashi, S. and Nakahashi, K. Efficient and Robust Cartesian Mesh Generation for Building-Cube Method. *J. Comp. Sci. Tech.*, (2008) **2.4**:435-446.
- [11] Arai, N., Houzu, H. and Takakura, Y. Oscillation Pattern of Parachute and Concave Body. *AIAA Paper*, (2007) 2007-2531.
- [12] Takahashi, S., Monjugawa, I. and Nakahashi, K. Unsteady Flow Computation around Moving Airfoils by Overset Unstructured Grid Method. *Trans. Japan Soc. Aero. Space Sci.*, (2008) **51**.172:78-85.
- [13] Kim, J. and Moin, P. Application of a Fractional-step Method to Incompressible Navier-Stokes Equations. *J. Comp. Phys.*, (1985) **59**:308-323.
- [14] Rhie, C.M. and Chow, W.L. Numerical Study of the Turbulent Flow Past an Airfoil with Trailing Edge Separation. *AIAA J.*, (1983) **21**.11:1525-1532.

NUMERICAL STUDY OF THE FLUID – STRUCTURE INTERACTION IN THE DIFFUSER PASSAGE OF A CENTRIFUGAL PUMP

ALFRED FONTANALS*, ALFREDO GUARDO^{†*}, MIGUEL COUSSIRAT^{††} AND
EDUARD EGUSQUIZA[†]

*Escola Universitaria d'Enginyeria Tècnica Industrial de Barcelona.
C\ Compte d'Urgell, 187. 08036. Barcelona, Spain.

[†] Centre de Diagnòstic Industrial i Fluidodinàmica (CDIF), Universitat Politècnica de Catalunya
Av. Diagonal, 647. ETSEIB. – 08028 Barcelona, Spain
e-mail: alfredo.guardo-zabaleta@upc.edu – Web page: <http://www.upc.edu/cdif/>

^{††} Electromechanic Department. FRM, Universidad Tecnológica Nacional
Coronel Rodríguez 273 – 5500. Mendoza, Argentina

Key words: Fluid-structure interaction.

Abstract. Reliable design of turbomachinery is a complex task. In order to gain overall efficiency in the machine performance, designers may have to reduce the gap between the impeller and the diffuser, forcing them to be as closely spaced as possible. In these situations, there may be a strong interaction between them that influences both the aerodynamics and the structural performance of blades and vanes. This phenomenon is called rotor-stator interaction (RSI), and it has a strong influence on the machine behavior. These interactions can have a significant impact on the vibrational and acoustical characteristics of the machine [1-2]. Sometimes, this interaction has led to blade or vane failure [3].

Unsteadiness and turbulence play a fundamental role in RSI [4-5], and the use of computational fluid dynamics (CFD) is becoming a usual requirement in turbomachinery design due to the difficulties and elevated cost of the experimentation required to identify RSI phenomena. Nowadays, a CFD analysis based on Reynolds-averaged Navier-Stokes equations (RANS) and a coupled eddy viscosity turbulence model (EVM) is commonly applied in turbomachinery design. Therefore, the choice of an appropriate turbulence model and the boundary layer treatment is far from trivial, and a suitable turbulence modeling plays an important role for successful CFD results.

In this work, an entire stage of a diffuser pump was modeled by means of a commercial CFD code in order to study the pressure fluctuations due to the interaction between the impeller and the diffuser of the pump. The obtained numerical results were compared against the experimental results of Tsukamoto et al., [6]. Full RANS equations coupled with several EVM were solved for a diffuser pump stage in order to establish the most accurate modeling strategy for a diffuser pump. Boundary layer sensitivity tests were performed, and numerical discretization influence on results was also tested and established. Frequencies of the pressure fluctuations in the diffuser passage are also obtained with several EVM and compared against experimental results.

1 INTRODUCTION

In turbomachinery design, the Rotor-Stator Interaction (RSI) is an important phenomenon, with a strong influence on the machine behavior. These interactions can have a significant impact on the vibrational and acoustical characteristics of the machine [2, 3]. Unsteadiness and turbulence play a fundamental role in complex flow structure [7, 8], and the use of Computational Fluid Dynamics (CFD) is becoming a usual requirement in design in turbomachinery due to the difficulties and high cost of the necessary experiments needed to identify RSI phenomena.

The RSI can be divided into two different mechanisms: potential flow interaction and wake interaction [1]. The nature of the flow due to the wake interaction is unsteady and turbulent and, in the case of flow in turbomachines, there also appear three-dimensional boundary-layer, curvature and system rotation effects.

Nowadays, a CFD analysis based on Reynolds-Averaged Navier-Stokes equations (RANS) and a coupled eddy viscosity turbulence model (EVM) is commonly applied in turbomachinery design. Therefore, the choice of an appropriate turbulence model and the boundary layer treatment is far from trivial, and a suitable turbulence modeling plays an important role for successful CFD results [4].

In this work, an entire stage of a diffuser pump was modeled, by means of a commercial CFD code, in order to study the pressure fluctuations due to the interaction between the impeller and the diffuser of the pump, and the obtained numerical results were compared against the experimental results of Tsukamoto et al. [6]. Full RANS equations coupled with several EVM were solved for a diffuser pump stage in order to establish the most accurate modeling strategy for a diffuser pump. Boundary layer sensitivity tests were performed, and numerical discretization influence on results was also tested and established.

2 NUMERICAL MODELING

2.1 Geometry and Grid Generation

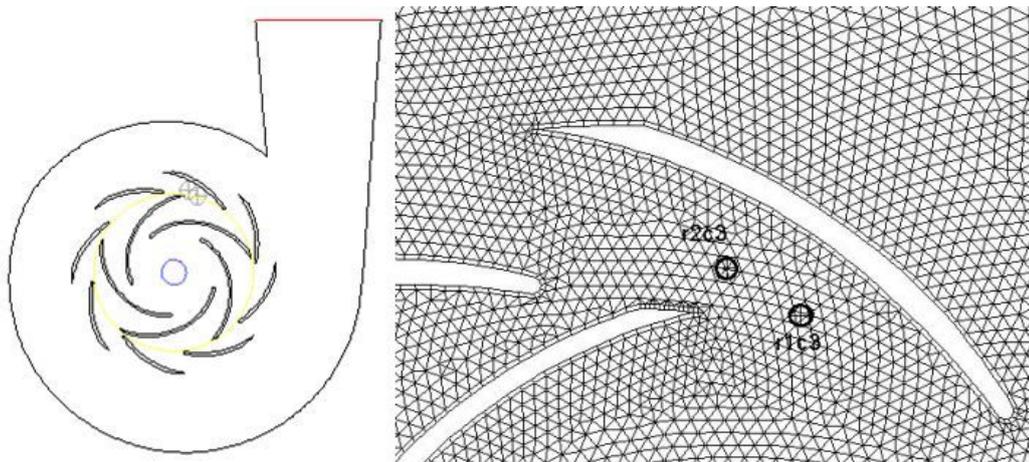
The modeled test pump was a single-stage diffuser pump with five impeller blades, $N_i=5$, eight diffuser vanes, $N_v=8$, and volute casing (Figure 1). A detailed description of the pump is given by Tsukamoto et al. [6].

Table 1: Mesh sensitivity test

Mesh	Number of cells	y^+
1	$4,8 \times 10^4$	20-200
2	$9,7 \times 10^4$	20-200
3	$3,2 \times 10^5$	20-200
4	$8,1 \times 10^4$	1-10
5	$1,4 \times 10^5$	1-10
6	$3,0 \times 10^5$	1-10

A mesh sensitivity test was performed in order to evaluate the independency of the numerical results from the mesh density. Also, the effect of the boundary layer treatment on the results was performed. For accomplishing the aforementioned, for 3 different mesh densities the boundary layer around the blades, vanes and walls was modeled using a wall function (WF), with a y^+ ranging from 20 to 200, and another three mesh densities were tested using a two layer model (TLM), with a y^+ ranging from 1 to 10. For further details on the meshes used see Table 1.

Figure 1: Geometry and studied point position in the vaned diffuser stage.



2.2 Unsteady Calculations Setup

Two dimensional, unsteady incompressible Reynolds-averaged Navier-Stokes equations were solved by means of the commercial CFD code Ansys Fluent 12.1 [9]. A constant pressure value was imposed at the fluid inlet and a constant pressure value was imposed at the pump outlet. A non slip boundary condition was imposed at the runner blades, diffuser vanes and volute casing wall. A rotational speed of 2066 rpm was imposed to the blade impeller.

One of the goals of this work was to benchmark the ability of different turbulence models in capturing the pressure fluctuations and the characteristic frequencies in the diffuser according to the experimental results reported by Tsukamoto et al. [6]. Turbulence models employed in its work are listed in Table 2.

In order to accomplish the proposed turbulence model performance study, pressure fluctuations in two points of the vaned diffuser passage, which correspond to the reported measurement points for the experimental data set, were recorded for every numerical simulation developed. Location of the monitor points r1c3 and r2c3 can be seen in Figure 1.

The unsteady formulation used was a second-order implicit velocity formulation and a pressure-based solver was chosen. The SIMPLE pressure-velocity coupling algorithm was used, and second order scheme discretization was selected for the numerical experiments.

Table 2: Turbulence models used

#	Turbulence model	Reference
1	Spalart – Allmaras	Spalart and Allmaras [10]
2	Standard k- ϵ	Launder and Spalding [11]
3	Realizable k- ϵ	Shi et al. [12]
4	RNG k- ϵ	Yakhot and Orszag [13]
5	Standard k- ω	Wilcox [14]
6	Shear Stress Transport (SST) k- ω	Menter[15]
7	Reynolds Stress Model (RSM)	Launder et al. [16]

The maximum number of iterations for each time step was set to 40 in order to reduce all computed normalized numerical residuals to 1×10^{-5} . The interface between the rotor blade and the diffuser vane was set to a sliding mesh, in which the relative position between the rotor and the stator was updated every time step. The adopted computational time step was about 1/360 of the rotor revolution time.

3 RESULTS AND DISCUSSION

Due to the unsteady nature of the flow, it is required that the whole flow domain is affected by the unsteady fluctuations. In order to check the aforementioned situation, a flow rate monitoring was made at the domain outlet. Uniform unsteady flow behavior was reached after 10 revolutions.

3.1 Pressure Fluctuations

For the two monitored points, r1c3 and r2c3, in the diffuser vane the unsteady static pressure was obtained, in order to determine the effect of the RSI in the diffuser vane. The results obtained for several turbulence models of the non-dimensional unsteady pressure, $\Delta\Psi$, are shown in Figures 2 and 3.

Results obtained show that the highest pressure value in the diffuser is due to the potential effect generated by the movement of the rotor vane in front of the diffuser vane. The wakes at the rotor outlet enter the diffuser and interact with the vanes producing high pressure peaks.

When comparing the numerical results obtained with each one of the turbulence models tested, it can be observed that there's no significant difference in the results of pressure fluctuation. Only the Standard k- ω turbulence model diverge from others turbulence models results and experimental results.

When analyzing the effect of the boundary layer treatment on the results, it can be observed that WF are able to accurately reproduce the pressure fluctuations due to RSI, while a TLM does not lead to an improve of the results, even when using a turbulence model that takes into account the boundary layer transition from laminar to turbulent flow, as the Standard k- ω or the SST k- ω .

A detailed analysis of the obtained results allow to notice that using a WF, the Spalart-Allmaras, Realizable $k-\epsilon$ and SST $k-\omega$ turbulence models show the best fitting for the experimental pressure fluctuations (Figure 4).

When a TLM is used as near-wall treatment, the best fitting obtained for pressure fluctuation correspond also to Spalart-Allmaras, Realizable $k-\epsilon$ and SST $k-\omega$ turbulence models.

Previous studies carried out by our research group [4] showed that the SST $k-\omega$ turbulence model presented a good performance when determining the velocities in the boundary layer and the wake vortex shedding when studying flow around an isolated foil, and also a low computational cost if compared with more accurate (but more resource-consuming) turbulence models as RSM or Large Eddy Simulation (LES).

Figure 2: Computed non dimensional unsteady pressure, $\Delta\Psi$, at point r1c3. Boundary layer treatment: left, wall functions; right, two layer model.

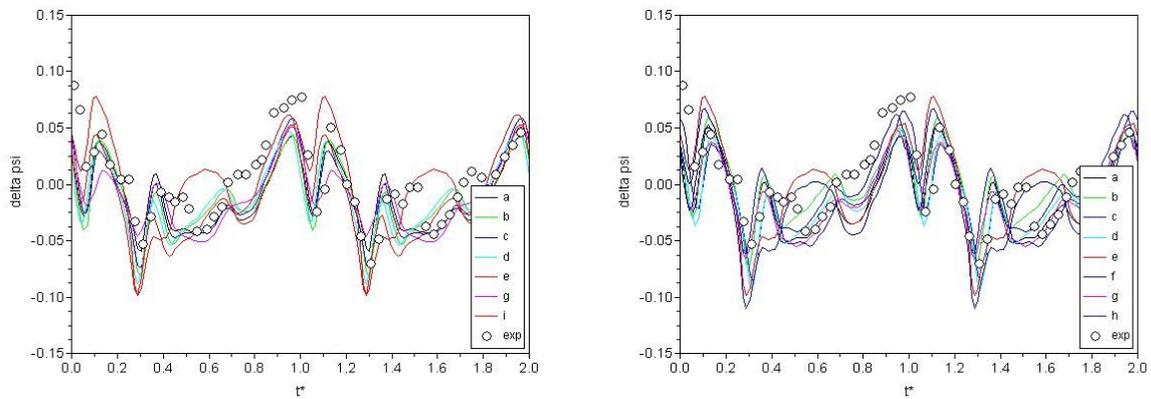


Figure 3: Computed non dimensional unsteady pressure, $\Delta\Psi$, at point r2c3. Boundary layer treatment: left, wall functions; right, two layer model.

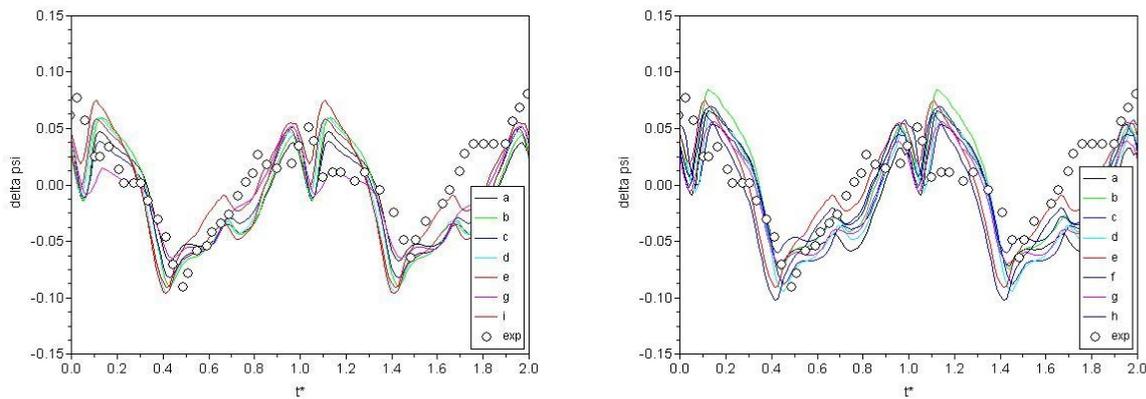


Figure 4: Computed non dimensional unsteady pressure, $\Delta\Psi$. Wall function boundary layer treatment: left, point r1c3; right, point r2c3.

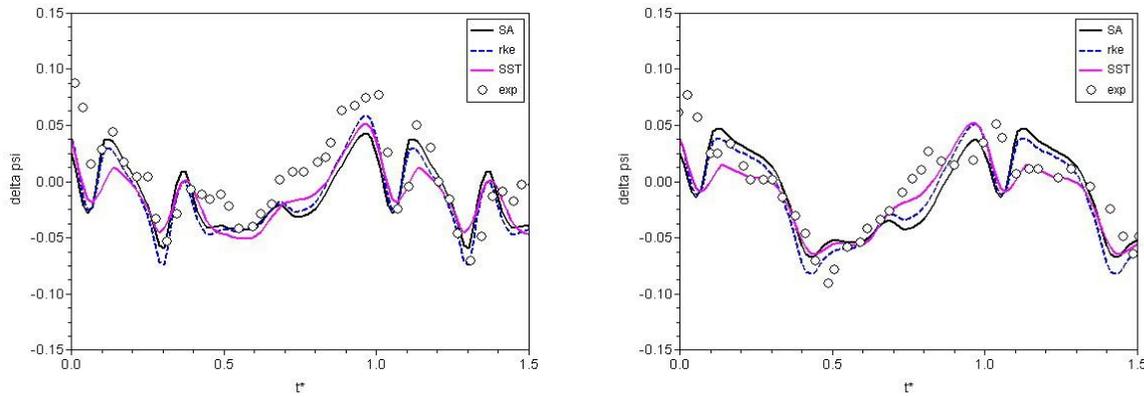


Table 3: Pressure fluctuations in frequency domain

		Frequencies	1Z _i N	2Z _i N	3Z _i N	4Z _i N	5Z _i N
		Experimental	172	344	517	689	861
	Turbulence model	Pressure Point					
a	Spalart-Allmaras	r1c3	171	343	518	689	861
		r2c3	171	343	518	689	861
b	Standard k-ε	r1c3	171	346	517	689	860
		r2c3	171	346	517	689	860
c	Realizable k-ε	r1c3	171	343	518	689	861
		r2c3	172	343	518	689	861
d	RNG k-ε	r1c3	171	343	517	689	861
		r2c3	171	343	517	689	861
e-f	Standard k-ω	r1c3	-	345	516	688	862
		r2c3	171	-	516	-	861
g-h	SST k-ω	r1c3	171	343	-	689	862
		r2c3	171	-	515	689	862
i	RSM	r1c3	172	343	517	689	861
		r2c3	172	343	517	689	861

3.1 Frequencies

In order to capture the RSI effects, the relationship between the pressure fluctuations and the movement of the rotor vanes in front of the diffuser vanes was determined. Using a Fourier transform, the characteristic frequencies of the pressure fluctuations have been obtained at points r1c3 and r2c3, resulting that the pressure fluctuates with the impeller blade passing frequency $Z_i N$ and its higher harmonics. Table 3 shows the obtained frequencies with different turbulence models and the experimental results. It can be noticed that the k-ω family turbulence models are not able to capture the representative frequencies while all the other models benchmarked accurately capture the characteristic frequencies of the phenomena.

Figure 5: Spalart-Almaras model. Left: calculated unsteady pressure at point r1c3; right: frequency domain.

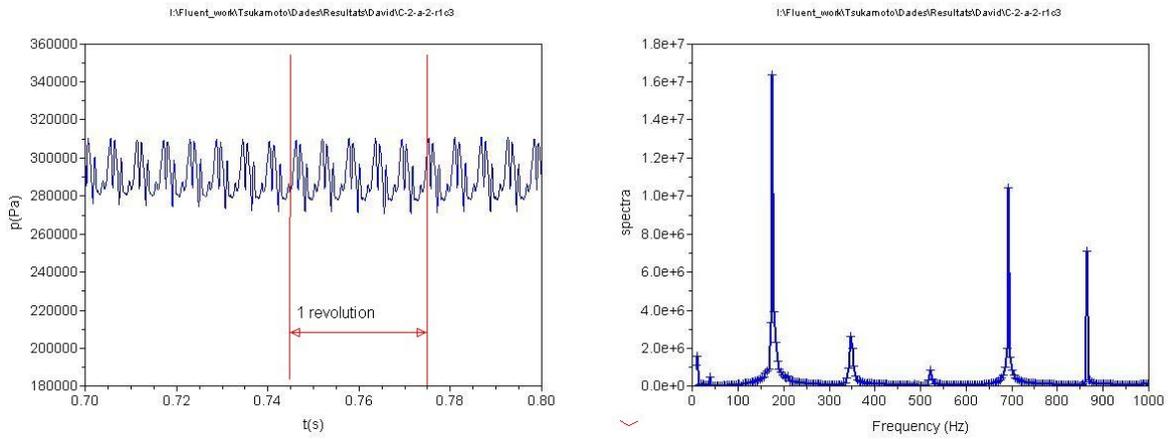


Figure 6: Realizable $k-\epsilon$ model. Left: calculated unsteady pressure at point r1c3; right: frequency domain.

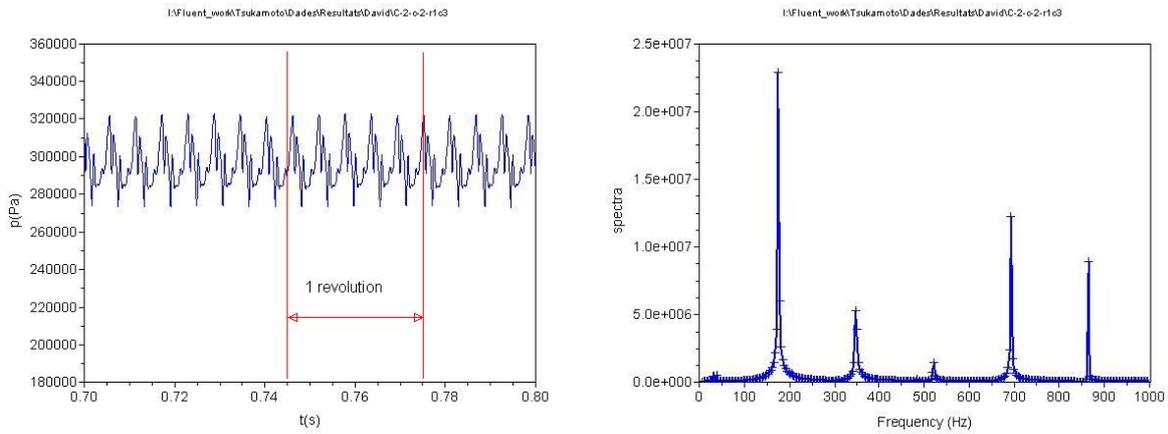
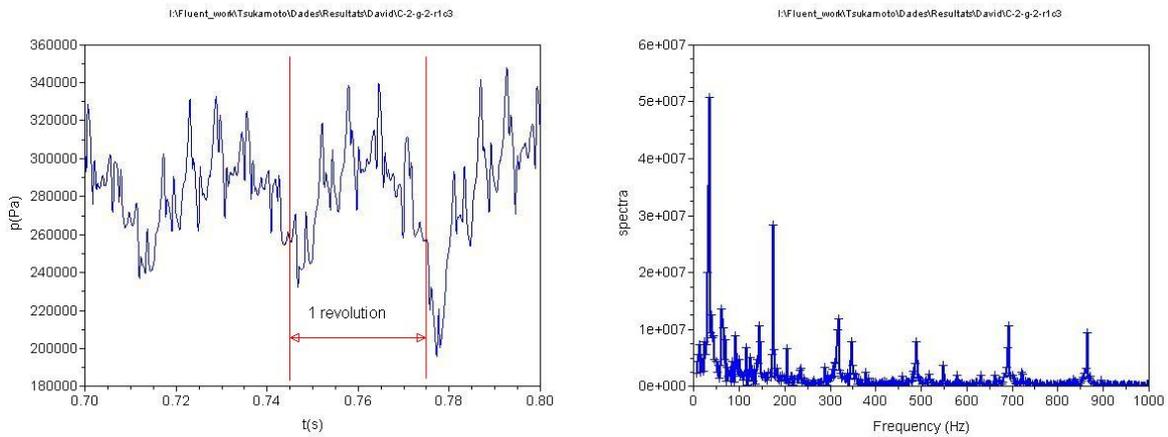


Figure 7: SST $k-\omega$ model. Left: calculated unsteady pressure at point r1c3; right: frequency domain.



Figures 5, 6 and 7 show in detail the pressure fluctuations at point r1c3 and its correspondent Z_iN frequencies (after treating the obtained pressure data with a Fourier transform) for 3 different turbulence models: Spalart-Allmaras, Realizable $k-\varepsilon$ and SST $k-\omega$.

For the Spalart-Allmaras and Realizable $k-\varepsilon$ turbulence model (Figures 5 and 6), it can be noticed that for 1 revolution of the rotor, the pressure fluctuation generated by the movement of the rotor vanes follows a periodic pattern, and the frequency domain accurately registers the data of the impeller blade passing frequency Z_iN .

For the SST $k-\omega$ turbulence model (Figure 7), pressure fluctuation show different values along 1 revolution of the rotor for each passing of the impeller blade in front of the diffuser vanes, showing also a cyclic pattern after each revolution. The presence of all this ‘noise’ in the numerical results obtained makes it difficult to determine the impeller blade passing frequency Z_iN , as it translates into non-physical peaks in the frequency domain. This behavior pattern was also noticed for the Standard $k-\omega$ turbulence model.

4 CONCLUSIONS

A 2-D CFD has been applied to the study of RSI in the vaned diffuser of pump. Reynolds-averaged Navier-Stokes equations with several turbulence models were tested in order to check their capability to capture RSI phenomena in the pump. Models tested showed good results of pressure fluctuation in the vaned diffuser when comparing against experimental results, except the Standard $k-\omega$ model.

A detailed analysis of the obtained results allow to notice that the Spalart-Allmaras, Realizable $k-\varepsilon$ and SST $k-\omega$ turbulence models show the best fitting for the experimental pressure fluctuations

When analyzing the effect of the boundary layer treatment on the results, it can be observed that WF are able to accurately reproduce the pressure fluctuations due to RSI, while a TLM does not lead to an improve of the results, even when using a turbulence model that takes into account the boundary layer transition from laminar to turbulent flow, as the Standard $k-\omega$ or the SST $k-\omega$.

The relationship between the pressure fluctuations and the movement of the rotor vanes in front of the diffuser vanes was determined. The characteristic frequencies of the pressure fluctuations were obtained, resulting that the pressure fluctuates with the impeller blade passing frequency Z_iN and its higher harmonics. It was noticed that the $k-\omega$ family turbulence models were not able to capture the representative frequencies while all the other turbulence models benchmarked accurately captured the characteristic frequencies of the phenomena.

All the turbulence models tested shoed a periodic pattern in the pressure fluctuation in which each cycle is produced by the movement of a rotor blade in front of a diffuser vane, except for the $k-\omega$ turbulence models family. For these models, pressure fluctuations value was different for each pass of an impeller blade in front of a diffuser blade, but the overall behavior for 1 revolution of the impeller followed a regular pattern.

Spalart-Allmaras and Realizable $k-\varepsilon$ turbulence models show the most accurate results of the benchmarked models, making them a suitable option for CFD modeling of RSI in turbomachinery. It is important to remark that Spalart-Allmaras turbulence model has a smaller computational cost, as it is a 1-equation model.

Previous work of our research group showed that SST k- ω turbulence models was a good choice for estimating boundary layer velocities and wake vortex shedding around isolated foils, but results obtained in present work show that when RSI potential effect is present, the numerical response of the model generates background noise in the results, making it difficult to clearly differentiate the characteristic frequencies of the phenomena..

In order to obtain a better understanding of the RSI phenomena it would be necessary to determine the influence of the discretization schemes, the time step value. These numerical studies are to be performed in the near future.

5 NOMENCLATURE

ψ = instantaneous pressure coefficient = $(p - P_s) / (\rho U_2^2 / 2)$

$\Delta\psi$ = non-dimensional unsteady pressure = $\tilde{p}^* / (\rho U_2^2 / 2)$

p = instantaneous static pressure = $\bar{p} + \tilde{p}$

\bar{p} = time average p

\tilde{p} = unsteady component of p

p^* = relative pressure = $p - P_s$

P_s = total pressure at pump suction port

U_2 = peripheral speed of impeller = $\pi D_2 N$

N = rotational speed

ρ = density

t^* = non-dimensional time = t/T_i

t = time

T_i = time required to traverse one pitch of impeller blade = $1/(N Z_i)$

r = radius

c = symbol of pressure traverse line

Z = number of blades

ACKNOWLEDGEMENTS

Funding from the Spanish Ministry of Science and Innovation (Grant No. DPI 2009 – 12827) is appreciated. The authors also appreciate the contribution from EUETIB - UPC in the development of this work.

REFERENCES

- [1] Ardnt, Acosta, Brennen and Caughey. Rotor-Stator Interaction in a Diffuser Pump. *J. of Turbomachinery* (1989) **111**:213-221.
- [2] Chow, Y., Uzol, O. and Katz, J. Flow nonuniformities and turbulent “hot spots” due to wake-blade and wake-wake interaction in a multi-stage turbomachine. *J. of Turbomachinery* (2002) **124**:553-563.

- [3] Uzol, O., Chow, Y., Katz, J. and Meneveau, C. Experimental investigation of unsteady flow field within a two-stage axial turbomachine using particle image velocimetry. *J. of Turbomachinery* (2002) **124**:542-552.
- [4] Coussirat, M., Fontanals, A., Grau, J., Guardo, A. and Egusquiza, E. CFD study of the boundary layer influence on the wake for turbulent unsteady flow in rotor-stator interaction. *IAHR 4th. Symposium on Hydraulic Machinery and Systems*. Foz do Iguassu, Brazil (2008).
- [5] Dring, Joslyn, Hardin And Wagner H. Turbine Rotor-Stator Interaction. *J. Eng. for Power* (1982) **104**:729-742.
- [6] Tsukamoto, H., Uno, M., Hamafuku, N., And Okamura, T. Pressure fluctuations downstream of a diffuser pump impeller. *The 2nd Joint ASME/JSME Fluids Engineering Conference, Forum of unsteady flow, FED* (1995) **216**:133-138.
- [7] Henderson, A., Walker, G. and Hughes, J. The influence of turbulence on wake dispersion and blade row interaction in an axial compressor. *J. of Turbomachinery* (2006) **128**:150-165.
- [8] Soranna, F., Chow, Y., Uzol, O. and Katz, J. The effect of inlet guide vanes wake impingement on the flow structure and turbulence around a rotor blade. *J. of Turbomachinery* (2000) **128**:82-95.
- [9] Fluent Inc. *Fluent 6.3. User's guide*, (2006).
- [10] Spalart, P.R. and Allmaras, S.R. A one equation turbulence model for aerodynamic flow. *La Recherche Aéronautique* (1994) **1**:5-21.
- [11] Launder, B.E. and Spalding, D.B. *Lectures in mathematical models of turbulence*, Academic Press, London (1972).
- [12] Yakhot, V. and Orszag, S. Renormalization group analysis of turbulence: I Basic theory, *J Sci Comput* (1986) **1**:3-51.
- [13] Shih, T. Liou, W. Shabbir, A. Yang, Z. and Zhu, J. A new k- ϵ eddy viscosity model for high Re turbulent flow – Model development and validation, *Comput Fluids* (1995) **24**:227-238.
- [14] Wilcox, D.C. *Turbulence model for CFD*, DCW Industries Inc., California (1998).
- [15] Menter, F.R., Two equations eddy-viscosity turbulence models for engineering applications, *AIAA J* (1994) **32**:1598-1605.
- [16] Launder, B.E. Reece, G.J. and Rodi, W. Progress in the development of a Reynolds stress turbulence closure, *J Fluid Mech* (1975) **68**:537-566.

ONSET OF TWO-DIMENSIONAL TURBULENCE WITH HIGH REYNOLDS NUMBERS IN THE NAVIER-STOKES EQUATIONS.

A. Nicolás-Carrizosa * and B. Bermúdez-Juárez[†]

*Depto. de Matemáticas, 3er. Piso Ed. AT-Diego Bricio,
Uam-Iztapalapa, 09340, México DF, México.
e-mail: anc@xanum.uam.mx

[†]Facultad de Ciencias de la Computación
Benemérita Universidad Autónoma de Puebla
14 sur y San Claudio, Puebla, Pue. México.
e-mail: bbj@cs.buap.mx

Key words: High Reynolds numbers, time-dependent flow, asymptotic steady state

Abstract. Even though turbulence is a tri-dimensional phenomenon, two-dimensional flows at high Reynolds numbers Re give some clues of transition to real turbulence, mainly through the vorticity. This transition is shown here with flows that are obtained computationally from the unsteady Navier-Stokes equations in stream function and vorticity variables on the well known un-regularized driven cavity problem. The work covers the range of $5000 \leq Re \leq 31000$; the results are obtained with a numerical scheme based on a fixed point iterative process applied to the elliptic nonlinear system that results after time discretization, it was reported in [1] for lower Re ; it started since WCCM V, 2002, [2]-[3], and we are still working on it. The scheme has the ability to start from rest, initially, regardless of Re and has shown to be robust enough to handle high Reynolds numbers, which is not an easy task to deal with.

1 INTRODUCTION

The main goal of this paper is to present numerical results for high Reynolds numbers in the range of $5000 \leq Re \leq 31000$. Actually, in [2] the range $10000 \leq Re \leq 20000$ is covered, in [3], with primitive variables, the range $25000 \leq Re \leq 40000$, and in [1] the range $400 \leq Re \leq 5000$ to capture the steady state flow and the range $10000 \leq Re \leq 20000$ for time-dependent flows; all the cases of time-dependent flows at high Reynolds numbers are displayed at bigger times than the one considered here as well as on different meshes. The results are obtained using a simple numerical scheme for the unsteady Navier-Stokes equations in stream function and vorticity variables. They give us some clues of transition

to real tri-dimensional turbulence, based on the fact that the Navier-Stokes equations is one of the three classical approaches to turbulence; the other two being: the dynamical systems approach and the conventional statical theory of turbulence, [4]. The numerical scheme is based mainly on a fixed point iterative process applied to the steady subproblem that results after a convenient time discretization is applied, [1].

At moderate Reynolds numbers, say for instance $Re \leq 7500$, the flow approaches to an asymptotic steady state as t tends to ∞ . For higher Reynolds numbers, like the ones reported here, as time elapses the flow does not seem to be "stationed" somewhere, indicating that the solution is time-dependent.

The flows are obtained from the well known un-regularized driven (or lid-driven) cavity problem which originates recirculation phenomena due to the nonzero velocity boundary condition on the top wall: the recirculation is originated by the fluid flow coming from the upstream top corner, and then hitting the downstream top corner.

To get the results, unlike in [3] where very coarse meshes are used since an up-winding effect is considered, here no stabilization process is used; then the meshes in this work follow the size dictated by the thickness of the boundary layer (of order of $Re^{-\frac{1}{2}}$) and no refining on the mesh is required near the boundary. The results clearly show that as the Reynolds number increases the mesh has to be refined and this in turn leads to decrease the time step: numerically, by stability matters and physically, to capture the fast dynamics of the flow. We have already pointed out in earlier works that to get the right vorticity contours, say the ones given by the values in [5], which are supposed to be correct is more difficult than to get the right streamlines of the stream function; this the reason that some published works do no report the iso-vorticity contours, at best due to oscillations on the top right corner of the cavity for insufficient mesh refining, [6]; for instance, the result in [7] for $Re = 10000$, with a mesh $\frac{1}{128}$, shows a reasonable streamlines but awful iso-vorticity contours, due to high oscillations, (not reported by them but computed by us using the same mesh and agreeing with them with the streamlines flow). Actually, concerning the mesh size for our results we have chosen, at this stage, the one for which such oscillations are reduced to a minimum, smaller than the ones in [6].

Last but not least, as it will be seen in the Numerical Results section, from some Reynolds number on the iso-vorticity contours resemble the bi-dimensional view we see in the real 3D hurricane pictures, close to the hurricane's eye; it may be due to the recirculation that is originated by the fast fluid water flow hitting the big waves, thermal effects; and high wind velocities interacting with the ocean nonlinearly, and viceversa, [8].

2 THE CONTINUOUS PROBLEM AND THE NUMERICAL METHOD

Let $\Omega \subset R^N$ ($N=2,3$) be the region of the flow of a viscous incompressible fluid, and Γ its boundary. It is well known that this kind of unsteady flow is governed by the

non-dimensional Navier-Stokes equations given by

$$\begin{cases} \mathbf{u}_t - \frac{1}{Re}\Delta\mathbf{u} + \nabla p + (\mathbf{u} \cdot \nabla)\mathbf{u} = \mathbf{f} & \text{in } \Omega, t > 0, & (a) \\ \nabla \cdot \mathbf{u} = 0 & \text{in } \Omega, t > 0, & (b) \end{cases} \quad (1)$$

where \mathbf{u} , and p are the velocity and pressure of the flow, respectively. The parameter Re is the Reynolds number. The momentum equation (1a) must be supplemented with appropriate initial condition and boundary condition, for instance $\mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x})$ in Ω ($\nabla \cdot \mathbf{u}_0 = 0$) and $\mathbf{u} = \mathbf{f}_1$ on Γ , $t \geq 0$ ($\int \mathbf{f}_1 \cdot \mathbf{n}d\Gamma = 0$) respectively.

Restricting the domain for equations (1a) – (1b) to the two dimensional case, taking the curl on both sides of (1a) and taking into account the relations

$$u_1 = \frac{\partial\psi}{\partial y}, \quad u_2 = -\frac{\partial\psi}{\partial x}, \quad (2)$$

which follow from (1b), with ψ the stream function, and $(u_1, u_2) = \mathbf{u}$; the component in the direction of \mathbf{k} gives the scalar system

$$\begin{cases} \nabla^2\psi & = -\omega & (a) \\ \omega_t - \nu\nabla^2\omega + \mathbf{u} \cdot \nabla\omega & = 0 & (b) \end{cases} \quad (3)$$

where $\frac{1}{Re}$ has been replaced by the viscosity parameter ν , and ω is the vorticity, which from $\omega\mathbf{k} = \nabla \times \mathbf{u} = -\nabla^2\psi\mathbf{k}$, is given by

$$\omega = \frac{\partial u_2}{\partial x} - \frac{\partial u_1}{\partial y} \quad (4)$$

System (3) is the Navier-Stokes equations in stream function-vorticity variables associated with the one in primitive variables (1), considering the external force given by $\mathbf{f} = \mathbf{0}$. It should be noted that because of the relations (2) the incompressibility condition (1b) is automatically satisfied in Ω , an advantage against the disadvantage that no boundary condition is given for the vorticity. Actually in [9], a procedure is given to get the boundary condition for ω in general domains.

For two dimensional rectangular domains, equations (3) are set in the domain $\Omega = (0, a) \times (0, b)$; with $a, b > 0$. The motion boundary condition in terms of the primitive variable \mathbf{u} is defined by $\mathbf{u} = (1, 0)$ at the moving boundary (the top one $y = b$) and $\mathbf{u} = (0, 0)$ elsewhere.

A translation of the boundary condition in terms of the velocity primitive variable \mathbf{u} to the $\psi - \omega$ variables has to be performed. Following [1], ψ is a constant function on solid and fixed walls; at the moving wall $y = b$, a constant function for ψ is also obtained, then $\psi = 0$ is chosen in Γ . By Taylor expansion of (3a) on the boundary, with h_x and h_y the space steps, one obtains

$$\begin{cases} \omega(0, y, t) = -\frac{1}{2h_x^2}[8\psi(h_x, y, t) - \psi(2h_x, y, t)] + O(h_x^2) \\ \omega(a, y, t) = -\frac{1}{2h_x^2}[8\psi(a - h_x, y, t) - \psi(a - 2h_x, y, t)] + O(h_x^2) \\ \omega(x, 0, t) = -\frac{1}{2h_y^2}[8\psi(x, h_y, t) - \psi(x, 2h_y, t)] + O(h_y^2) \\ \omega(x, b, t) = -\frac{1}{2h_y^2}[8\psi(x, b - h_y, t) - \psi(x, b - 2h_y, t)] - \frac{3}{h_y} + O(h_y^2). \end{cases} \quad (5)$$

About time discretization, the time derivative ω_t is approximated by the second-order scheme

$$f_t(\mathbf{x}, (n + 1)\Delta t) \approx \frac{3f^{n+1} - 4f^n + f^{n-1}}{2\Delta t}, \quad (6)$$

where $n \geq 1$, $\mathbf{x} \in \Omega$, $\Delta t > 0$ the time step.

Then, at each time level the following nonlinear elliptic system is obtained

$$\begin{cases} \nabla^2\psi = -\omega, & \psi|_\Gamma = 0; & (a) \\ \alpha\omega - \nu\nabla^2\omega + \mathbf{u} \cdot \nabla\omega = f_\omega, & \omega|_\Gamma = \omega_{bc}, & (b) \end{cases} \quad (7)$$

where $\alpha = \frac{3}{2\Delta t}$ and $f_\omega = \frac{4\omega^n - \omega^{n-1}}{2\Delta t}$. To obtain (ψ^1, ω^1) , any second order strategy using a combination of one step can be applied and systems of the form (7) are also obtained.

Taking into account that the elliptic system (7) in addition to be nonlinear is of non-potential (or transport) type, a fixed point iterative process is used to solve it. A distinctive aspect here is that the iterative process is extended until the boundary to handle the ω boundary conditions given implicitly by unknown interior values of ψ .

Denoting

$$R_\omega(\omega, \psi) \equiv \alpha\omega - \nu\nabla^2\omega + \mathbf{u} \cdot \nabla\omega - f_\omega \quad \text{in } \Omega,$$

system (7) is equivalent to

$$\begin{cases} \nabla^2\psi & = -\omega \quad \text{in } \Omega, & \psi = 0 \quad \text{on } \Gamma \\ R_\omega(\omega, \psi) & = 0 \quad \text{in } \Omega & \omega|_\Gamma = \omega_{bc} \end{cases} \quad (8)$$

Then, (8) is solved at time level $(n+1)$, by the fixed point iterative process: Given ω^0 and θ^0 solve until convergence in ω and θ

$$\begin{cases} \nabla^2\psi^{m+1} = -\omega^m \quad \text{in } \Omega, \\ \psi^{m+1} = 0 \quad \text{on } \Gamma \\ \omega^{m+1} = \omega^m - \rho_\omega(\alpha I - \nu\nabla^2)^{-1}R_\omega(\omega^m, \psi^{m+1}) \quad \text{in } \Omega, \\ \omega^{m+1} = \omega_{bc}^{m+1} \quad \text{on } \Gamma, \quad \rho_\omega > 0; \end{cases} \quad (9)$$

and then, take $(\omega^{n+1}, \psi^{n+1}, \theta^{n+1}) = (\omega^{m+1}, \psi^{m+1}, \theta^{m+1})$.

Finally, system (9), with the corresponding ω^0 , is equivalent to

$$\begin{cases} \nabla^2 \psi^{m+1} = -\omega^m & \text{in } \Omega, \\ \psi^{m+1} = 0 & \text{on } \Gamma \\ (\alpha I - \nu \nabla^2) \omega^{m+1} = (\alpha I - \nu \nabla^2) \omega^m - \rho_\omega R_\omega(\omega^m, \psi^{m+1}) & \text{in } \Omega, \\ \omega^{m+1} = \omega_{bc}^{m+1} & \text{on } \Gamma, \rho_\omega > 0. \end{cases} \quad (10)$$

Two uncoupled elliptic linear problems associated with the operators $-\Delta$, $\alpha I - \nu \Delta$ have to be solved. Therefore, the solution of the original system, at each iteration of each time level, leads to the solution of standard symmetric linear elliptic operators.

It is well known that for the space discretization of problems like (10), either finite differences or finite elements may be used, as far as rectangular domains are concerned; it is also known that in either case very efficient solvers exist. For the specific results in this work, the second order approximation of the Fishpack solver [10] has been used, where the linear systems are solved through an efficient cyclic reduction iterative method; then, such second order approximation in space, combined with the second order one for vorticity boundary conditions and the second order approximation in time (6) imply that the whole approximate problem is second order.

3 Numerical Experiments.

The experiments take place on the well known un-regularized driven cavity problem. Then, the problem is set in the region $\Omega = (0, 1) \times (0, 1)$, with boundary given by the four walls of the cavity; the top one is moving with a nonzero velocity given by $(1, 0)$ and the other are solid and fixed, the velocity (by viscosity) is given by $(0, 0)$. The range that is considered for the Reynolds numbers is $5000 \leq Re \leq 31000$. The results are reported through the iso-vorticity contours; the size mesh and the time step are denoted respectively by h and Δt , and they will be specified by each case under study. All the flows that are reported are flows at time $t = 5$.

A) Figures 1, 2, and 3 picture the iso-vorticity contours for $Re = 5000$, 10000, and 20000 with mesh size given by $h = 1/256$, $3/384$, and $1/640$ respectively, and time step given by $\Delta t = 0.0025$ for all of them. B) Figures 4, 5 and 7 display the iso-vorticity contours for $Re = 25000$, 30000 and 31000, with mesh size $h = 1/768$ for the first two, $1/512$ for third one, and time step given by $\Delta t = 0.00025$ for all of them. C) In connection with the flow that is showed in Figure 5, Figure 6 displays the profile for the vorticity along the line $y = x$ which clearly shows that the great variation, positive and negative, occurs close to the top right corner which is in concordance with what Figure 5 shows. Actually, for $Re = 30000$ in Figure 5 we are talking about the Max/min values for the vorticity are given by $\text{Max}/\text{min} = 2.845 \times 10^3 / -2.1812 \times 10^3$ and they occur at $(x, y) = (0.999, 1) / (x, y) = (0.999, 0.00195)$.

Some discussion follows: 1) It is known that the flow for $Re = 5000$, Figure 1 in A), arrives at a steady state, we have reported this in earlier works as a validation matter. It is displayed here just to compare it, at the same time $t = 5$, with the others which are

supposed to be time-dependent flows (they do not arrive at some steady state); this flow at this short time shows that the vorticity goes to accumulate around the walls in some uniform manner. As Re increases, Figures 2 and 3 in A), the vorticity tends to spread in all the cavity and the uniform manner tends to disappear, Figure 3. **2)** Surprisingly, from $Re = 25000$, Figure 4, the structure changes drastically, a "turbulent" part appears; then, the end of the "turbulent" part tends to go down as long as Re increases; Figures 5 and 6; which is a consequence that the fluid motion is faster. In this connection, the flow for $Re = 31000$ in Figure 7, with mesh size coarser than the one in Figure 5, shows a bigger oscillation in the top right corner; nevertheless, it is not a significant one since it shows that the "turbulent" end is below the previous one for $Re = 30000$ in Figure 5. **3)** Outside the "turbulent" part in pictures 4, 5, and 7, the vorticity is zero which could be in disagreement with the position Max/min given just above but it is not: it can be verified, for instance with the corresponding 3D vorticity picture, that close to the bottom left corner there appear the lowest values but they are so few that the plotter does not take care of them in Figure 6.

4 Conclusions

We have presented fluid flows at high Reynolds numbers using the Navier-Stokes equations approach to two-dimensional turbulence through the vorticity which is caused by recirculation in driven cavity problem, solving the stream function-vorticity formulation through a simple numerical scheme. For the range of Reynolds numbers considered $5000 \leq Re \leq 31000$, our results, at the fixed time $t = 5$, distinguish three stages of Reynolds number sizes: a low one that is supposed to reach its steady state; two which are supposed to be time-dependent flows; and four for which a "turbulent" part appears close to the downstream top corner. Preliminary calculations for higher Reynolds numbers, $Re \geq 40000$, with finer mesh than $1/768$ and Δt smaller than 0.00025 , show that the "turbulent" part fills more the cavity downwards, and more at a bigger time than $t = 5$. These results will be reported elsewhere. It must be taken into consideration that these kind of two-dimensional fluid flow views, mainly for $Re \geq 25000$, are nowhere realised in nature or the laboratory but only in computer simulations, [11]; however, as we have been already pointed out in the Introduction, for this specific recirculation problem these views have to do, at least qualitatively, with the hurricanes phenomenon.

REFERENCES

- [1] A. Nicolás and B. Bermúdez, 2D Incompressible Viscous Flows at moderate and High Reynolds Numbers, CMES Vol.6, No. 5, (2004),441-451.
- [2] B. Bermúdez, A. Nicolás and C. Ortiz, On Numerical Solutions of the Navier-Stokes equations in Stream function-Vorticity variables at High Reynolds Numbers, Fifth World Congress on Computational Mechanics; Vienna Austria, July 7-12, 2002. On line publication ISBN 3-9501554-0-6.

- [3] A. Nicolás and B. Bermúdez, On some Numerical Solutions of the time-dependent Navier-Stokes equations at High Reynolds Numbers. Fifth World Congress on Computational Mechanics; Vienna Austria, July 7-12, 2002. On line publication ISBN 3-9501554-0-6.
- [4] C. Foias, O. Manley, R. Rosa, R. Temam, Navier-Stokes Equations and Turbulence, Cambridge University Press, (2001).
- [5] U. Ghia, K.N.Ghia, C. T.Shin, High-Re solutions for incompressible flow using the Navier-Stokes equations and a multigrid method, *J. Comput. Phys.*, 48, (1982), 387-411.
- [6] R. Schreiber, H.B. Keller, Driven Cavity flow by efficient numerical Techniques, *J. Comput. Phys.* 40, (1983), 310-333.
- [7] E. Weinan and J.G. Liu, Vorticity boundary condition and related issues for finite difference schemes, *J. Comput. Phys.*, 124, (1996), 368-382.
- [8] J.L. Liond, R. Temam, S. Wang, Models for the coupled atmosphere and ocean, *Computational Mechanics Advances*, Vol. 1, No.1, North-Holland, 1993.
- [9] E. J. Dean, R. Glowinski and O. Pironneau, Iterative solution of the stream function-vorticity formulation of the Stokes problem, application to the numerical simulation of incompressible viscous flow, *Comput Methods Appl. Mech. Engrg.*,87,(1991), 117-155.
- [10] J. Adams, P. Swarztrauber and R. Sweet R, FISHPACK: A Package of Fortran Subprograms for the Solution of Separable Elliptic PDE's', *The National Center for Atmospheric Research*, (1980), Boulder, CO, USA.
- [11] R.H. Kraichnan and D. Montgomery, Two-dimensional turbulence, *Rep. Prog. Phys.*, Vol. 43 (The Institute of Physics), (1980), 547-619.

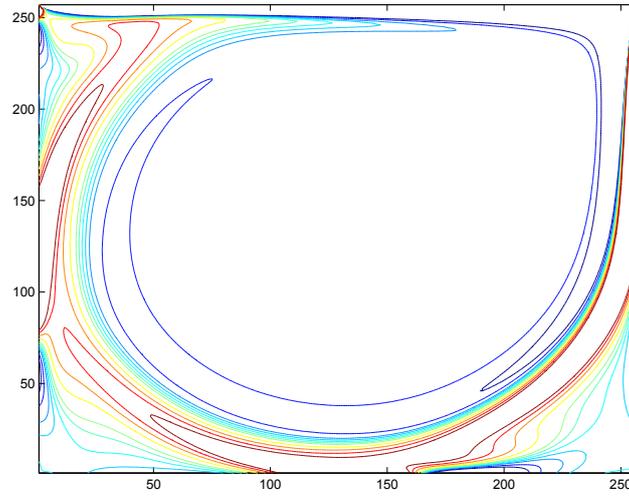


Figure 1: Vorticity: $Re = 5000$, $h = 1/256$, $dt = 0.0025$; $t = 5$

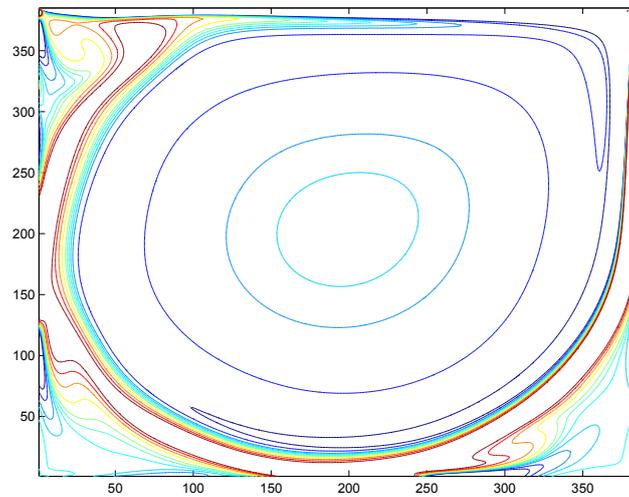


Figure 2: Vorticity: $Re = 10000$, $h = 1/384$, $dt = 0.0025$; $t = 5$

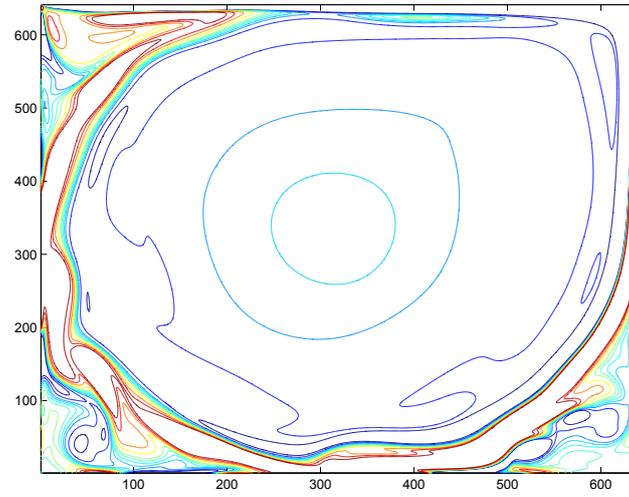


Figure 3: Vorticity: $Re = 20000$, $h = 1/640$, $dt = 0.0025$; $t = 5$

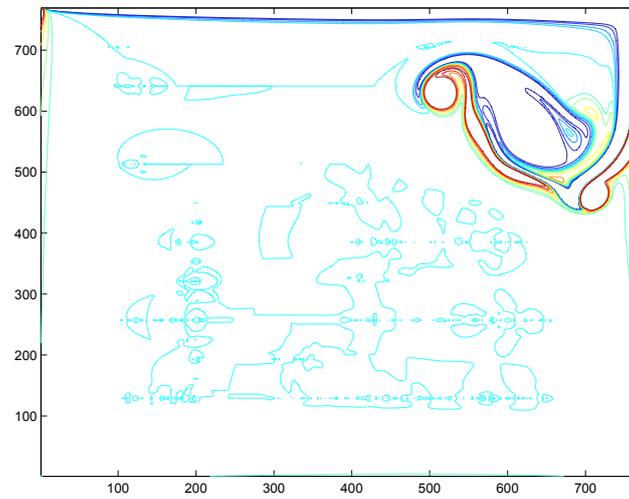


Figure 4: Vorticity: $Re = 25000$, $h = 1/768$, $dt = 0.00025$; $t = 5$

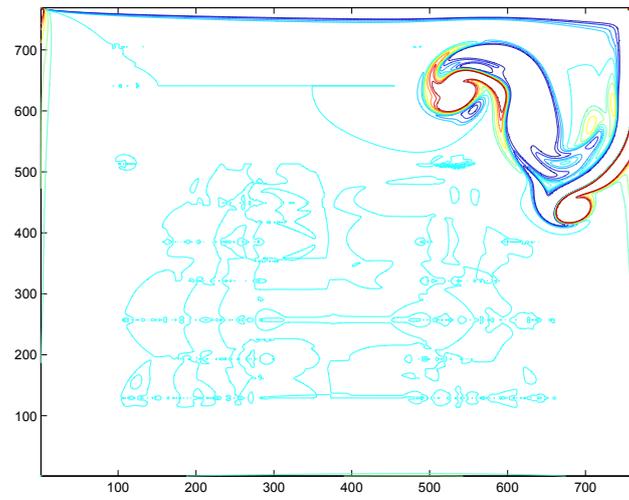


Figure 5: Vorticity: $Re = 30000$, $h = 1/768$, $dt = 0.00025$; $t = 5$

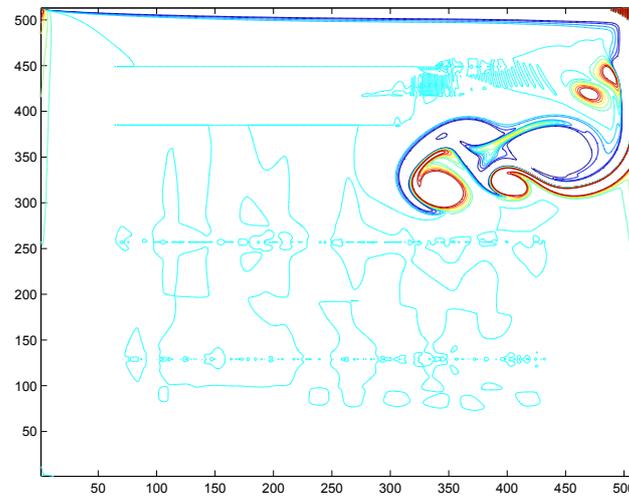


Figure 6: Vorticity: $Re = 31000$, $h = 1/512$, $dt = 0.00025$; $t = 5$

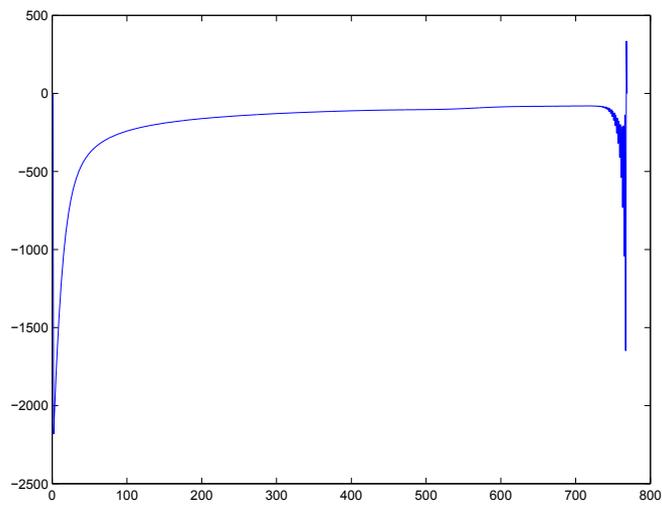


Figure 7: Vorticity profile along $y = x$ for $Re = 30000$: $h = 1/768$, $dt = 0.00025$; $t = 5$

VORTEX-INDUCED VIBRATIONS OF AN ELASTICALLY MOUNTED CYLINDER WITH LOW MASS-RATIO AT $RE=3900$

YOANN JUS*, ELISABETH LONGATTE*, JEAN-CAMILLE CHASSAING[†] and
PIERRE SGAUT[†]

*LAMSID-UMR EDF/CEA/CNRS n° 2832
Clamart, France
e-mail: yoann.jus@gmail.com, elisabeth.longatte@edf.fr

[†]UPMC Univ Paris 06, UMR 7190, Institut Jean le Rond d'Alembert
F-75005 Paris, France

Key words: Fluid-structure Interaction, Vortex-Induced Vibrations

Abstract. The present paper deals with the numerical study of vortex-induced vibrations (VIV) of an elastically mounted cylinder in a cross flow with uniform inflow. It is well known that, in the case of low mass-damping, three distinct types of transverse amplitude response can be observed depending on the range of the reduced velocity. However, the accurate numerical simulation of the VIV amplitudes at the lock-in upper branch remains a great challenge. Moreover, few studies deal with the investigation of the hysteretic loop due to the jump between the initial excitation branch and the upper branch.

Here, we propose to compute the transverse motion of the structure at a Reynolds number equal to 3900 Large Eddy Simulation. The Navier-Stokes equations are solved on a moving and deforming grid by means of an Arbitrary Lagrangian Eulerian (ALE) co-located finite volume method for unstructured meshes. An iterative algorithm is used at each time step based on Newton's or fixed point method. This algorithm uses convergent explicit predictions of the coupled fluid structure system and sub-cycling is involved to get convergence towards the implicit solution of the fully coupled system. A criterion based on the structure velocity is used to stop the numerical sub-cycling process.

First, we compare the transverse amplitude response of the cylinder for various reduced velocities with the DNS results obtained by Lucor et al. (2005) for zero structural damping and a mass ration equal to 2. Next, the experimental data of Hover et al. (1998) are considered to demonstrate the ability of the present solver to predict the VIV response for low mass-damping.

1 INTRODUCTION

Solving multi-physics problems is still a challenge requiring advanced computational methods. This work aims to investigate flow-induced vibrations of dynamical systems subjected to

complex flows. In very confined areas such as those encountered in steam exchanger and heat generator tube arrays we are proposing a decomposition of physical problems by considering, as a first step, mechanical systems made of single cylinder subjected to cross-flows.

Typical response of tube array under cross-flow are described below. When there is no mean flow, this is the so-called region of Fluid Structure Interaction (FSI). Cylinders are subjected to two kinds of excitations : inertial effect on one hand, damping effect on the other hand. These effects are combined and generated by the presence of fluid inertia and viscosity impacting shear stress exerted on solid walls. They may largely depend on confinement and they act on vibratory response frequency and damping. In the presence of mean flow, flow patterns across cylinder array are very sensitive to cylinder arrangement as well as other parameters. Cylinder spacing is one of the most important geometric parameters. Conventional results propose a classification in flow regimes in standard in-line and staggered tube arrays. In several configurations, all cylinders shed Karman vortices but a jet swing associated with vortex shedding is also possible. Sometimes Karman vortices can not develop because the free shear layer of a front cylinder may become attached to the downstream cylinder. On the contrary, when the confinement decreases, vortex streets can be the same as those shed by isolated cylinders. There are several possible flow regimes and many studies have been performed to identify the borderlines between these flow regimes as flow regimes depend on all of the geometric and hydraulics parameters like Reynolds and Stokes numbers. There are three typical kinds of action exerted by fluid and flow on solid walls responsible for three different dynamical behaviours of tube array under cross flow : (1) Vortex-Induced Vibration (VIV) : in the lock-in frequency range, a synchronization occurs between vortex shedding frequency and solid response frequency, which implies an increase of displacement magnitude as well as a change of cylinder frequency. (2) Turbulence-Induced Vibration (TIV) : whenever the reduced velocity, cylinders are submitted to an external action exerted by fluid whose level is directly governed by the Reynolds number. The higher the Reynolds number value, the higher the level of turbulence and the associated loading which induces a moderate increase of vibration magnitude - without any change however in the dynamical stability regime of the cylinders. (3) Finally Motion-Induced Vibration (MIV) : when the reduced velocity reaches the critical threshold called the fluid-elastic instability threshold, vibration magnitude dramatically increases linearly and only non-linear effects like collision or breakdown can stop this motion.

The paper is organised as follows. A brief review of simulation of VIV are recalled in a first part. In the second part, numerical methods are described. The last section focuses on VIV and simulation of lock-in for a dynamic single cylinder under cross-flow.

2 STATE OF ART

Before dealing with multi-tube configurations (in-line or staggered tube arrays), single-cylinder configurations are investigated in this article. In the case of single cylinder, only the VIV phenomenon occurs. This phenomenon has been widely studied. So, it is well known that, in the case of an elastically mounted cylinder with low mass-damping, three distinct types of transverse amplitudes response can be observed. The distinction between the three response

branches arising in different ranges of the reduced velocity U^* has been described in the literature [13][14][15]. The reduced velocity U^* is defined by $U_0/f_n D$ with U_0 the flow velocity, f_n the system natural frequency and D the cylinder diameter.

For low reduced velocities, there exists an *initial* branch associated with a 2S vortex shedding mode (two single vortices shed per cycle) and the mean forces and cylinder response are in phase. For intermediate and larger reduced velocities there is an *upper* and a *lower* branch associated with a 2P vortex shedding mode emission [13] [14] (two pairs of vortices shed per cycle). The 'lock-in' is characterized by a synchronization of the vortex shedding frequency (f_s) and the system natural frequency (f_n). The reduced frequency f^* is defined by the ratio between frequency of oscillation f_0 and f_n . At the lock-in, f^* tends to be equal to 1. At the 'lock-in', the reduced velocity U^* is directly related to the Strouhal number St . For a given reduced velocity, the amplitude of the response and the mode emission of the vortex shedding depend on the mass-damping parameter $m^*\xi$. m^* designates the mass ratio m/m_f with m_f the mass of displaced fluid for a single cylinder ($m_f = \rho\pi L_z D^2/4$) and ξ the reduced damping. However, very few numerical results have been able to accurately reproduce the three-branch response model obtained from experiments.

Reaching the *upper* branch response still remains a challenge. Hysteresis, three dimensional as well as low mass-damping effects have been highlighted in previous work (2D-FEM simulations of Singh and Mittal (2005) [16], 3D-DVM simulations of Yamamoto et al. (2004) [17] and 2D-FVM simulations of Placzek et al. (2007) [18]) but without the capture of the *upper* branch response of the cylinder. As the 'upper branch' regime is observed for moderate and high Reynolds numbers and therefore appears to be Reynolds-dependent. Al Jamal and Dalton (2004) [19] introduce turbulence modelling and perform 2D-LES simulations at $Re = 8000$, without reaching the *upper* branch. However Guilmineau and Queutey (2004) [20] and Pan et al. (2007) [21] perform 2D-RANS simulations in a wide range of Reynolds numbers ($2000 < Re < 10000$), which yield promising results, since the *upper* branch regime is observed by using an ' U^* -increasing' initial condition for simulations (i.e. by performing continuous simulations and gradually increasing the fluid upstream velocity - and the Reynolds number - while keeping all other parameters constant). However, the *upper* branch is not observed when using the 'from rest' conditions (i.e. setting the fluid upstream velocity at a given value and letting the cylinder freely oscillate under VIV). Recently, Sigrist et al. (2008) [22] yield the *upper* branch with a 2D RANS simulation but the extent of this higher amplitude zone is rather narrow when compared to the experiments. Vortex-Induced Vibrations involve complex three dimensional phenomena that can not be accurately simulated by using 2D modeling. Only Lucor et al. (2005) [23] obtain cylinder response results closely matching the 'three-branch response' mode proposed in experimental works. In particular, the existence of an *upper* branch with large amplitude response is confirmed by performing Direct Numerical Simulations (DNS) with low mass-damping (zero structural damping and mass ratio $m^* = 2$) for Reynolds number range ($Re = 1000$ to 3000).

Numerical simulation of the turbulent wake of static and dynamic cylinders in cross flows is considered. Large Eddy Simulation (LES) is involved in the subcritical Reynolds number

range enabling large temporal spectra evaluations. A moving mesh formulation is used through an Arbitrary Lagrange Euler (ALE) method in order to deal with solid boundary motion and an iterative solver is used to account for fluid structure interfacial coupling. As far as fully-coupled fluid solid system computation is concerned, in the presence of strong non linearity of fluid, solid or interface dynamics (like turbulence or large magnitude motion), small perturbation development procedures relying on linearization of interfacial boundary conditions are prohibited. Therefore alternative methods involving iterative algorithms are prescribed. The present work proposes the resolution of a fully-coupled fluid solid system by using a staggered procedure suitable for non-matching interfaces. Thanks to a partitioned approach both fluid and solid subsystems are formulated, discretized and solved separately and an iterative algorithm is involved to ensure the convergence towards the solution ensuring energy transfer consistency through the interface.

3 NUMERICAL ASPECTS

Simulations presented in the framework of the present article rely on Navier-Stokes equation computations¹ by using a collocated finite volume method for unstructured meshes devoted to incompressible flows and turbulence with a fractional time step procedure for fluid pressure velocity coupled computation through a projection method [1].

The flow is assumed Newtonian with a constant density ρ . If \bar{u} and \bar{p} stand for spatially filtered velocity and pressure, the filtered Navier-Stokes equations can be written as :

$$\frac{\partial \bar{u}_i}{\partial x_i} = 0 \quad (1a)$$

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial \bar{u}_i \bar{u}_j}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + \nu \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j} - \frac{\partial \tau_{ij}}{\partial x_j} \quad (1b)$$

with local coordinates (x, t) in the space time domain.

The subgrid scale tensor $\underline{\tau}$ has to be modeled. The dynamic Smagorinsky model based on Germano identity and Lilly minimization is used [2] [3] to take into account the small structures that are mainly dissipative. The deviatoric part of the subgrid-scale tensor is given by :

$$\tau_{ij} - \frac{1}{3} \tau_{kk} \delta_{ij} = -2\nu_t \bar{S}_{ij} = -2(C_s \bar{\Delta})^2 \|S\| \bar{S}_{ij} \quad (2)$$

where \bar{S}_{ij} represents the filtered strain rate tensor, $\|S\| = \sqrt{2\bar{S}_{ij}\bar{S}_{ij}}$, ν_t the subgrid-scale viscosity, $\bar{\Delta}$ the filter width and C_s the dynamic Smagorinsky constant. In this work all computational cells are hexahedral, therefore $\bar{\Delta} = 2\Omega^{\frac{1}{3}}$ can be used, where Ω is the volume of a computational cell. The term $\frac{1}{3}\tau_{kk}\delta_{ij}$ is taken into account in the pressure gradient. The filter which appears in equation (1) is implicit as it is introduced by the discretization errors in space

¹With *Code_Saturne* (www.code-saturne.org) involving a SIMPLEC algorithm with a Rhie and Chow interpolation to avoid odd-even decoupling on structured meshes [1].

and time and by the model itself. The explicit filter which is applied to compute the dynamic constant C_s uses the neighbors sharing a node with the computational cell. No averaging in the homogeneous directions (the spanwise direction in these cases) is performed. No negative value is allowed to the subgrid-scale viscosity and the maximum value of the dynamic constant is set to 0.065 (the standard Smagorinsky model value).

In the framework of collocated finite volume approach, all variables are located at the center of gravity of the cells. The momentum equations are solved by considering an explicit mass flux (the three components of velocity are thus uncoupled). A second order centred scheme in space and time is used. A Crank-Nicholson time scheme with a linearized convection and a second order Adams-Bashforth method for the part of the diffusion involving the transposed gradient operator, coupling velocity components are involved. A centred scheme is used for the convection operator. The non-orthogonalities are taken into account with an implicit reconstruction technique explained in [1]. When a non-orthogonal grid is used, the matrix contains only the orthogonal contributions of the different operators. The non-orthogonal part is added to the right hand side of the transport equation (thus, inner iterations are needed for the velocity and pressure equations to make the gradient reconstruction implicit). This approach is suitable for several academic cases (decaying isotropic turbulence, channel flow) and also for industrial ones (T-junctions, combustor chamber, tube arrays) [4][5].

Solving fluid solid coupled systems involves the resolution of fluid and solid dynamics simultaneously and in a coupled manner particularly at the interface where energy transfer takes place. Therefore a formulation of boundary conditions compatible with both fluid and solid systems is required. For a given interface, energy exchanges between fluid and solid domain occur through the interfaces of solid and fluid domains ($\Gamma_{s/f}$ and $\Gamma_{f/s}$). The energy transferred per unit of time and surface is defined by the product of mechanical stress acting on the interface and the interface displacement velocity. Therefore, the energy flux evaluated on each cell boundary located on fluid and solid interface models is consistent if at least two conditions are fulfilled in the continuum space : the continuity of the velocity and the stress at the interface. After space and time discretizations these conditions must be preserved with the required accuracy whatever the spacetime evolution of the boundaries. This implies an additional condition on the geometry continuity. At the interface :

$$\begin{cases} u_i = \frac{Du_i^s}{Dt} & \text{on } \Gamma_{f/s} \\ \sigma_{ij}n_j = T_{ij}n_j & \text{on } \Gamma_{s/f} \end{cases} \quad (3)$$

where u_s represents the displacement of the interface, $\frac{D}{Dt}$ the material derivative and T_{ij} the solid stress tensor. In this work, cylinder motion under crossflow is investigated. Considering small magnitude motion in the cross direction only if there is no degree-of-freedom of motion in the drag direction and assuming rigid body motion, the solid dynamic equation is expressed as follows :

$$m \ddot{y} + c \dot{y} + k y = F_y \quad (4)$$

where y is the transverse cylinder displacement, m the oscillating structural mass, c the structural damping, k the structural stiffness and F_y the action exercised by the fluid in the lift direction (small magnitude fluctuations around the mean equilibrium position). In the framework of rigid motion theory (similarly for linear elasticity), a lagrangian formulation is used to describe the time evolution of the solid kinematics. System remains linear and time integration relies on an implicit Newmark second order unconditionally stable algorithm.

In the fluid domain however standard Eulerian formulation is not suitable for describing solid boundary motion. Therefore an Arbitrary Lagrangian Eulerian (ALE) approach is involved. This moving grid method consists in introducing an arbitrary referential domain for Navier-Stokes system computation and using associated time-depending reference space mapping to derive the system to be solved in the computational coordinate system.

Thus, the governing equations in the ALE framework for an incompressible flow are written as follows :

$$\begin{aligned} \rho \frac{\partial u_i}{\partial x_i} &= 0 \\ \frac{\partial u_i}{\partial t} + (u_i - v_i) \frac{\partial u_i}{\partial x_j} &= -\frac{1}{\rho} \frac{\partial p}{\partial x_i} + \nu \frac{\partial^2 u_i}{\partial x_j \partial x_j} \end{aligned} \quad (5)$$

where v represents the cell velocity evaluated at the cell's center of gravity in a collocated finite volume approach. The introduction of an arbitrary computational reference system, from a mathematical point of view, means the introduction of a grid mesh impacting convective terms in the momentum equation. For incompressible flow, the Geometry Conservation Law (GCL) [6] is ensured for uniform flows with a first order approximation. It ensures numerical conservation of physical fields : the variation of an elementary volume during a time step Δt is balanced by the flux through the volume faces during the same period. The integral form of the GCL is :

$$\frac{\partial}{\partial t} \int_V dV = \int_{\delta V} w \cdot n dS \quad (6)$$

Several formulations are possible to impose the grid velocity. In this work, a Poisson elliptic equation is introduced where variable $\bar{\lambda}$ allows cell deformation [7] to be controlled :

$$\begin{cases} \nabla \cdot (\bar{\lambda} \nabla(v)) &= 0 \\ v &= \frac{Du_s}{Dt} \quad \text{on } \Gamma_{f/s} \\ v &= 0 \quad \text{on } \partial\Omega_f \setminus \Gamma_{f/s} \end{cases} \quad (7)$$

The assumption that LES filtering commutes with partial derivatives is generally considered valid on fixed grids with uniform cell width. On deforming unstructured grid this is not the case anymore and temporal commutation errors (TCE) may have to be considered. In this work, the TCE are neglected in a first approximation. Finally to deal with computation of the fully-coupled fluid solid system, an iterative method is involved in order to look for a solution ensuring continuity conditions through the interface at each step of the calculation [7][8][9][10][11].

A fixed point method is used with consistent predictor and corrector terms for kinematics and stress field transfer associated with a projection and a condensation method in order to enable one-degree-of-freedom systems for modeling solid dynamics [12]. Under-relaxation may be introduced to improve the convergence properties of the iterative scheme.

4 NUMERICAL RESULTS AND DISCUSSION

A single cylinder motion under crossflow is studied. This is a first step before investigating cylinder arrangements in the future. The main objective of this part is to investigate numerically the behavior of a rigid cylinder at low mass-damping parameter, constrained to oscillate transversely to a free stream. The *upper* branch is particularly considered.

Numerical calculations are performed for two configurations : with and without damping. The experimental work of Hover et al. (1998) [24] and numerical results of Lucor et al. [23] are taken respectively as experimental and numerical references. Simulations are performed in a range of reduced velocities which extends from $U^* = 2$ to $U^* = 10$ for the case with damping and for $U^* = 5, 6, 7$ for the case without damping. Response of the cylinder undergoing VIV is analyzed in terms of amplitude and frequency of oscillations.

Before presenting results, some numerical parameters need to be specified. The size of the computational domain is $20D \times 20D \times 4D$. The length of the computational domain upstream the cylinder is equal to $10D$. This is necessary to allow the pressure field to reach an upstream asymptotic behaviour. A mesh with 32 points in the transversal direction is used and the total number of cells is equal to $2 \cdot 10^6$. Constant boundary conditions are introduced at the inlet (constant velocity in space and time) and no synthetic method is involved to generate turbulence. The flow is laminar upstream of the cylinder. No wall functions have been necessary during the unsteady simulations. Periodic boundary conditions are used in the spanwise direction. The outlet conditions are standard, Dirichlet condition for the pressure and homogeneous Neumann condition for the velocity. Symmetry boundary conditions are used for the upper and lower boundary faces. At the moving boundary, a Dirichlet condition is imposed for velocity with respect to the kinematic consistency condition. The Reynolds number $Re = 3900$ has been used. The maximum CFL number is equal to 0.8. The corresponding time step is $\Delta t = 0.001D/U_0$ where U_0 is the inlet velocity. Initial conditions for the flow field are set by a static computation.

4.1 Case without damping

In a first case, we set the structural damping to be zero as we are interested in the maximum response of the system. The choice of mass ratio ($m^* = 2$) and structural damping corresponds to a low mass-damping parameter range. The parameters are the same as the numerical references [23]. We investigate three different reduced velocities $U^* = 5, 6, 7$ for the oscillator, see table 1. All the cases corresponds to the upper branch region.

A_{mean} corresponds to the mean amplitude averaged on 10 oscillation periods, A_{max} to the maximum amplitude and A_{min} to the minimum amplitude computed. The deviation corresponds to the standard deviation from the mean amplitude A_{mean} .

m^*	$m^*\xi$	V_r	A_{mean}/D	A_{max}/D	A_{min}/D	Deviation
2	0	5	0.795	0.909	0.579	0.096
2	0	6	0.694	0.836	0.632	0.062
2	0	7	0.532	0.593	0.441	0.043

Table 1: Average, maximum, minimum amplitude of response and deviation versus reduced velocity without damping

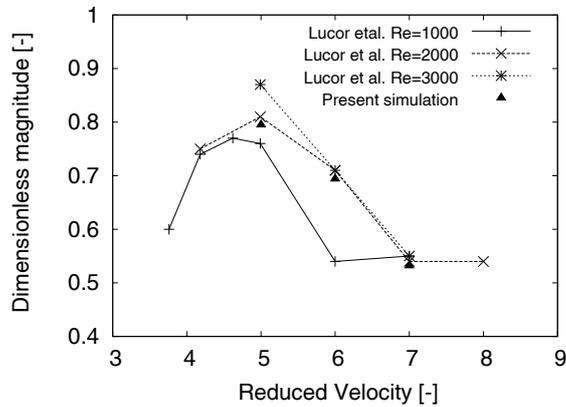


Figure 1: Dimensionless displacement at Reynolds 3900 at mass ratio $m^* = 2$ without damping versus reduced velocity with comparison to numerical [23] references

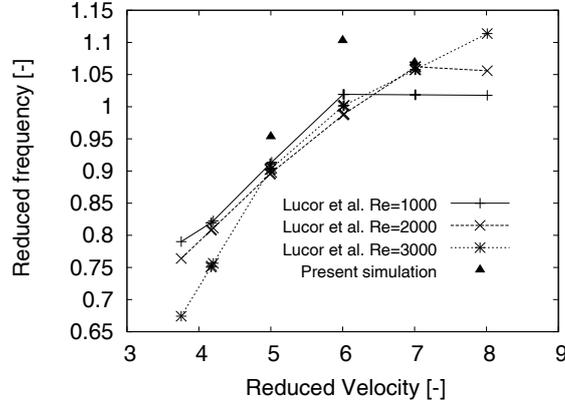


Figure 2: Evolution of reduced frequency at Reynolds 3900 at mass ratio $m^* = 2$ without damping (right) versus reduced velocity with comparison to numerical [23] references

As illustrated by the figure 1, the amplitude response of *upper* branch is Reynolds-dependent. For $U^* = 5$, amplitude response is lower than DNS results at Reynolds number 3000 and similar for other reduced velocity values ($U^* = 6, 7$). The first reason is clearly due to the difference of turbulence modelling and a second explanation can be the number of points along the spanwise direction (LES simulations were performed using 32 points whereas DNS simulations used 64 points).

Figure 4.1 gives the evolution of the frequency ratio f^* of the cylinder oscillations throughout the reduced velocity. The numerical results correlate well with the 'lock-in' zone except the reduced frequency for $U^* = 6$ is much larger than the numerical reference [23] for lower Reynolds number.

The simulation results are satisfactorily when compared to DNS results at least from a CPU cost point of view. This approach will be compared to an experimental reference : a cylinder with structural damping.

4.2 Case with damping

Numerical calculations are performed at low mass ratio ($m^* = 1$) and low mass-damping ($m^*\xi = 0.04$), which corresponds to the conditions of experimental references [24].

We present amplitude response results for $Re = 3900$ and investigate 8 cases for the oscillator with damping (table 2). A_{mean} gives the mean amplitude averaged on 10 oscillation periods, A_{max} the maximum amplitude and A_{min} the minimum amplitude computed throughout the reduced velocity range $2 < U^* < 10$. The deviation corresponds to the standard deviation from the mean amplitude A_{mean} .

For the case with damping, good correlations are observed as far as the *upper* branch is concerned. The highest average of the mean magnitude response of the cylinder around $A_{mean} \sim 0.795$ is obtained for a reduced velocity about $U^* = 5$. The extent of this higher amplitude zone correlates with the experimental data.

m^*	$m^*\xi$	V_r	A_{mean}/D	A_{max}/D	A_{min}/D	Deviation
1	0.04	2	0.081	0.121	0.052	0.019
1	0.04	3	0.380	0.416	0.352	0.33
1	0.04	4	0.794	0.892	0.663	0.052
1	0.04	5	0.795	1.01	0.639	0.056
1	0.04	5.5	0.721	0.898	0.562	0.107
1	0.04	6	0.715	1.08	0.547	0.149
1	0.04	7	0.661	0.821	0.550	0.072
1	0.04	10	0.488	0.546	0.413	0.045

Table 2: Average, maximum, minimum amplitude of response and deviation versus reduced velocity with damping

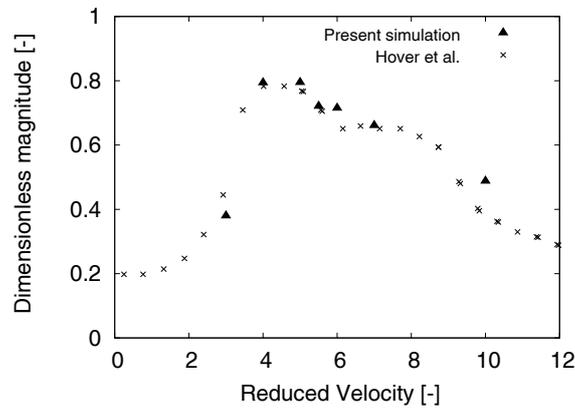


Figure 3: Dimensionless displacement at Reynolds 3900 at mass ratio $m^* = 1$ and damping ($m^*\xi = 0.04$) versus reduced velocity with comparison to experimental [24]

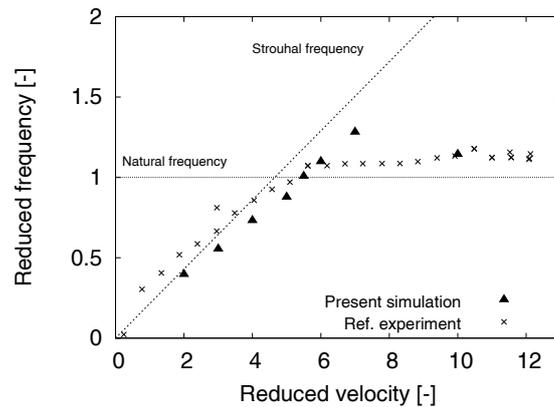


Figure 4: Evolution of reduced frequency at Reynolds 3900 at mass ratio $m^* = 1$ and damping ($m^*\xi = 0.04$) versus reduced velocity with comparison to experimental [24] references

Figure 4.2 gives the evolution of the frequency ratio f^* of the cylinder oscillations throughout the reduced velocity. The numerical results correlate well with the 'lock-in' zone.

The different branches of amplitude response are very well characterized by plotting not only the amplitude, but also the phase ϕ between the lift force and displacement, and the Lissajou figures or phase plane portraits of lift versus displacements (figure 5).

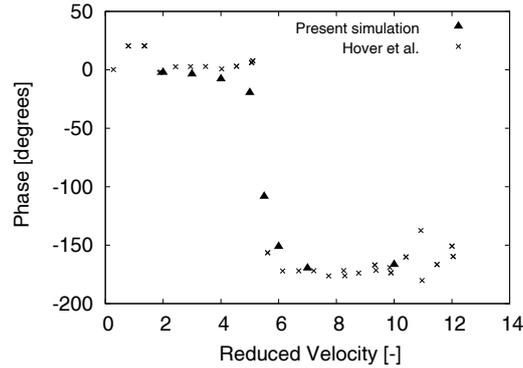


Figure 5: Phase between the lift force fluctuations and the displacement at Reynolds 3900 at mass ratio $m^* = 1$ and mass-damping ($m^*\xi = 0.04$) as a function of reduced velocity with comparison to experimental work [24]

The phase delay is evaluated by computation of cross power spectral density. The jump in phase angle ϕ in this figure is associated with the transition from the *upper* to the *lower* branch. If the phase ϕ is defined as the lead of the lift force fluctuations over the displacement, then the *upper* is near 0° , whereas the *lower* branch is nearer 180° , such as that found for a linear forced system going through resonance. When the dynamics shifts from the *initial* to the *upper* branch, the phase angle ϕ remains at just above 0° . The Lissajou figures indicate the very periodic nature of the oscillations in the *lower* branch (figure 6 left) and the relatively less steady dynamics of the *upper* branch (figure 6 right). In the latter case, shown by the Lissajou on figure 7 left, the phase changes of around 180° as the oscillations seem to wander to the smaller amplitude of the *lower* branch, and then back by -180° as the oscillations switch back to the larger amplitude of the *upper* branch. There is an intermittent switching between the *upper* and *lower* branch amplitudes and phases, whereas we see a hysteretic change from the *initial* to the *upper* branch. The phase portraits (figures 6, 7) can be compared to those plotted by [13] (see figure 13 in this reference).

The maximum lift (r.m.s.) occurs (figure 8) at the transition between the initial excitation and the *upper* branch. It is indeed very low in *lower* branch. As Khalak and Williamson (1999) [13], the lift increases rapidly before the transition and decreases after the transition.

The presented simulations yield the *upper* branch response which has never been reported so far with a 3D LES approach. The simulation results are therefore satisfactorily when compared to DNS results at least from a CPU cost point of view, since this approach can be considered

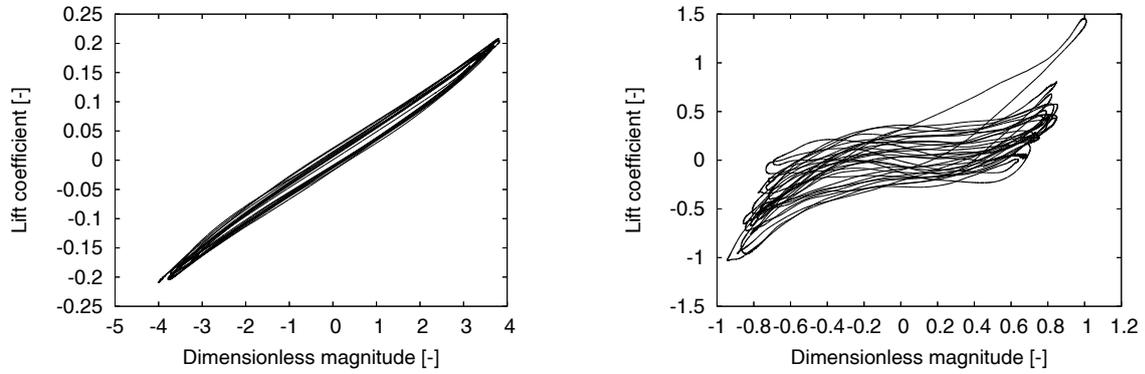


Figure 6: Phase plane for $U^* = 3$ (left) and $U^* = 5$ (right) at mass ratio $m^* = 1$ and damping ($m^*\xi = 0.04$)

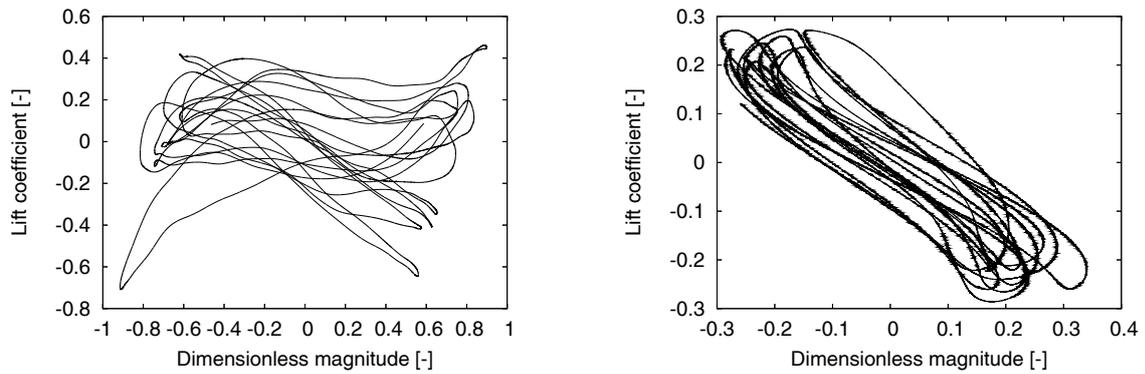


Figure 7: Phase plane for $U^* = 5.5$ (left) and $U^* = 10$ (right) at mass ratio $m^* = 1$ and damping ($m^*\xi = 0.04$)

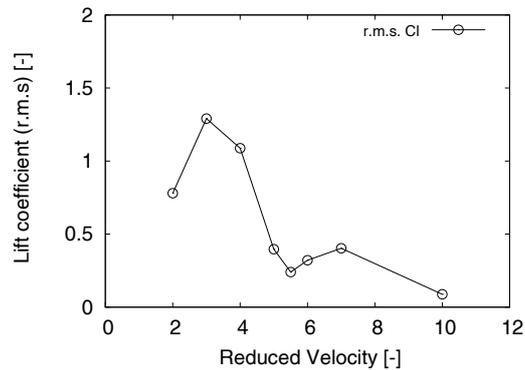


Figure 8: Variation of lift force coefficient versus reduced velocity for $m^* = 1$ and $m^*\xi = 0.04$

for practical applications. Further investigations will be devoted to the transitions between the branches.

5 CONCLUSIONS

Numerical simulations of the VIV response of a cylinder freely vibrating transversely to a fluid flow at moderate Reynolds number has been performed in the present study under low mass-damping conditions. Simulations are performed by using a coupled CFD/CSM procedure, with a single degree-of-freedom solid system and LES for modeling turbulent flow. Numerical results have been exposed and compared with experiments of Hover et al. (1998) on the one hand and simulations of Lucor et al. (2005) on the other hand. The presented simulations yield the *upper* branch response which has never been reported so far with a 3D LES approach. The simulation results are therefore satisfactorily when compared to DNS results at least from a CPU cost point of view, since this approach can be considered for practical applications. Qualitative agreements are highlighted as far as the cylinder response oscillation amplitude and frequency is concerned. Furthermore, the presented numerical simulations yield the *upper* branch response which has never been reported so far in CFD computations. Further investigations will be devoted to the transitions between the branches.

REFERENCES

- [1] F. Archambeau, N. Méchitoua, M. Sakiz, *A finite volume code for the computation of turbulent incompressible flows - industrial applications*, International Journal of Finite Volumes, 1, 2004.
- [2] M. Germano, U. Piomelli, P. Moin, W. Cabot, *A dynamic subgrid-scale eddy viscosity model*, Physics of Fluids, 3(7), 1760-1765, 1991.
- [3] D. Lilly, *A proposed modification of the germano subgrid-scale closure method*, Physics of Fluids, 4, 633-635, 1992.
- [4] S. Benhamadouche, *Large eddy simulation with the unstructured colocated arrangement*, Ph.D. thesis, Manchester University, 2006.
- [5] S. Benhamadouche, D. Laurence, *LES, coarse LES, and transient RANS comparisons on the flow across tube bundle*, International Journal Heat and Fluid Flow, 4, 470-479, 2003.
- [6] M. Lesoinne, C. Farhat, *A geometric conservation for flow problems with moving boundaries and deformable meshes, and their impact on aeroelastic computations*, Computer Methods in Applied Mechanics and Engineering, 134, 71-90, 1996.
- [7] F. Huvelin, *Couplage de codes en interaction fluide-structure et applications aux instabilités fluide-élastiques*, PhD thesis, 2008.

- [8] E. Longatte, V. Verreman, M.Souli, *Time marching for simulation of fluid structure interaction problems*, Journal of Fluids and Structures, 25, 95-111, 2009.
- [9] S. Piperno, *Explicit/implicit fluid/structure staggered procedure with a structural predictor and fluid subcycling for 2D inviscid aeroelastic simulations*, International Journal of Numerical Methods In Fluids, 25, 1207-1226, 1997.
- [10] S. Piperno, C. Farhat, *Partitionned procedures for the transient solution of coupled aeroelastic problems*, Computer Methods in Applied Mechanical and Engineering, 190, 3147-3170, 2001.
- [11] M. Schaefer, M. Heck, S. Yigit *An Implicit Partitioned Method for the Numerical Method of Fluid-Structure Interaction*, p. 171-194, H.-J. Bungartz and M. Schäfer, editors, Fluid-Structure Interaction, LNCSE 53, Springer, 2007.
- [12] N. Maman, C. Farhat, *Matching fluid and structure meshes for aero-elastic computations : a parallel approach*, Computers & Structures, 54 (4), 779-785, 1995.
- [13] A. Khalak, C.H.K. Williamson, *Motions, Forces and Mode Transitions in Vortex-Induced Vibrations at Low Mass Damping*, Journal of Fluids and Structures, 13, 813-851, 1999.
- [14] R. Govardhan, C.H.K Williamson, *Modes of Vortex Formation and Frequency Response of a Freely Vibrating cylinder*, Journal of Fluid Mechanics, 420, 85-130, 2000.
- [15] C.H.K. Williamson, R. Govardhan, *Vortex-induced vibrations*, Annual Review of Fluid Mechanics, 36, 413-455, 2004.
- [16] S.P. Singh, S. Mittal, *Vortex-Induced Vibration at Low Reynolds Numbers : Hysteresis and Vortex-Shedding modes*, Journal of Fluids and Structures, 20, 1085-1104, 2005.
- [17] C.T. Yamamoto, J.R. Meneghini, F. Saltara, R.A. Fregonesi, J.A. Ferrari, *Numerical Simulation of Vortex-Induced Vibration on Flexible Cylinders*, Journal of Fluids and Structures, 19, 467-489, 2004.
- [18] A. Placzek, J.F. Sigrist, A. Hamdouni, *Numerical Simulation of Vortex Shedding Past a Circular Cylinder at Low Reynolds Number with Finite Volume Technique. Part II : Flow Induced Vibrations.*, Pressure Vessel and Piping, San Antonio, 22-26 July 2007.
- [19] H. Al-Jamal, C. Dalton, *Vortex Induced Vibrations Using Large Eddy Simulation at a Moderate Reynolds number* Journal of Fluids and Structures, 19, 73-92, 2004.
- [20] E. Guilmineau, P. Queutey, *Numerical Simulation of Vortex-Induced Vibration of a circular cylinder with Low Mass-Damping in a Turbulent Flow*, Journal of Fluids and Structures, 19, 449-466, 2004.

- [21] Z.Y. Pan, W.C. Cui, Q.M. Miao, *Numerical Simulation of Vortex-Induced Vibration of a Circular Cylinder at Low mass-damping Using RANS Code*, Journal of Fluids and Structures, 23, 23-37, 2007.
- [22] JF. Sigrist, C. Allery, C. Beghein, *A Numerical Simulation of Vortex Induced Vibrations on a Elastically Mounted Circular rigid Cylinder at Moderate Reynolds Numbers*, Pressure Vessel and Piping, Chicago, 27-31 July 2008.
- [23] D. Lucor, J. Foo, G.E. Karniadakis, *Vortex mode selection of a rigid cylinder subject to VIV at low mass damping*, Journal of Fluids and Structures, 20, 483-503, 2005.
- [24] F.S Hover, A.H. Techet, M.S. Triantafyllou, *Forces on oscillating uniform and tapered cylinders in crossflow*, Journal of Fluid Mechanics, 363, 97-114, 1998.

MATHEMATICAL FORMULATION OF A COUPLED DYNAMIC MODEL FOR ENVIRONMENTAL BUSINESS ACTION PLANNING

SERGIO A. DAVID*, CLIVALDO de OLIVEIRA[†] AND DERICK D. QUINTINO[‡]

* Departamento de Ciências Básicas (ZAB)
Universidade de São Paulo -USP
Av. Duque de Caxias Norte 225, 13635-900 Pirassununga (SP), Brazil
e-mail: sergiodavid@usp.br

[†] Departamento de Engenharia de Produção (FACET)
Universidade Federal da Grande Dourados -UFGD
Rodovia Dourados/Itahum, Km 12, 79804-970 Dourados (MS), Brazil
email: ClivaldoOliveira@ufgd.edu.br

[‡] Departamento de Economia, Administração e Sociologia (LES)
Universidade de São Paulo- ESALQ -USP
Av. Pádua Dias 11, 13418-900 Piracicaba (SP), Brazil
email: derickdq@usp.br

Key words: dynamic optimization, rate of return, financial dynamics.

Abstract. The growing internationalization of markets, backed by the free movement of capital flows, has redefined the past quarter century's business strategies and tends to continue driving economic and financial integration throughout this century.

In this context, firms that aim to stand out in such markets should use the essence of the theoretical apparatus to allocate scarce resources efficiently. This means seeking the best possible benefits to offset the constraints that are inherent to the nature of the business environment. This new environment is also characterized by more far-reaching advances in technologies and their subsequent obsolescence.

In a scenario of coupled and complex interrelationships among variables and of strong uncertainties, corporate strategies must be relentlessly dynamic, i.e., constant readjustment is fundamental in dealing with the cyclical dynamics of the financial characteristics of self-regulated markets.

In this turbulent and competitive world, there is an increasing need to devise planning models to address the multiple issues (coupled) that affect competitiveness, such as: planned rate of return, price adjustment, technological obsolescence, optimal investment path, among others. In an effort to contribute to solutions for this need, we propose a coupled dynamic model that addresses the aforementioned issues. The model is based on a mathematical formulation using the Hamiltonian approach considering the Cobb Douglas function, Pontryagin conditions and is intended for application in environmental business action planning.

1 INTRODUCTION

In today's tumultuous world with its scenario of complex interrelated variables and its intense uncertainties, corporate strategies must be unfailingly dynamic. Hence, constant readjustments are fundamental in dealing with the cyclical dynamics of self-regulated markets. Today, strategy and modeling cannot be taken as absolutes, for there are countless uncertainty factors that may be very difficult to resolve. The importance of business strategies and modeling of action planning increases in proportion to increasing uncertainty. Strategies must therefore be dynamic and change constantly to deal with unpredictable external factors. In view of the above, we believe that the best way to face today's business uncertainties is through strategizing and modeling, followed by continuous revisions of the strategy and the model.

Established companies should brace themselves for a future of hypercompetition and be prepared to respond to rapid changes in the business environment by adopting a new approach – one that combines modeling and strategic thinking.

One of the basic notions in economics is that a company's capital accumulation is a future-oriented activity, and as such, it should depend on what the company believes is likely to occur. The fact that the future cannot be known with certainty has led to several investigations of the effects of increased uncertainty on the company's decision-making ability. Nickell [6], for instance, examined these effects based on several assumptions.

Moreover, in the globally competitive business world there is an increasing need to create planning models to deal with a wide variety of problems [2]. We believe there is an increasing need to devise planning models that address multiple issues (coupled) which affect competitiveness, such as planned rate of return, price adjustment, technological obsolescence, and optimal investment path, among others.

In an effort to contribute to solutions to satisfy this need, we propose a dynamic coupled model that addresses the aforementioned issues. The model is based on a mathematical formulation using the Hamiltonian approach [5] considering the Cobb Douglas function and Pontryagin's minimum principle, which is intended for application in environmental business action planning. Although this work is still incipient and therefore many aspects remain to be investigated, we believe this paper offers a contribution to this field.

This paper is divided as follows. Section 2 outlines the dynamic optimization model considering the Cobb Douglas function and Pontryagin's conditions. Section 3 describes the conditions for optimality. Section 4 presents the results of simulations. Lastly, section 5 offers a discussion, our conclusions and a future outlook.

2 THE DYNAMIC OPTIMIZATION MODEL

The development of a business must include a planned rate of return. Once the rate of return is determined, it can be written as a product of the profit margin vs. circulation of capital, as expressed by the following equation:

$$s' = M.G \quad (1)$$

where,

$$M = \frac{\text{Pr ofit}}{\text{Re venue}} \quad (2)$$

and

$$G = \frac{\text{Re venue}}{\text{Capital}} \quad (3)$$

With this fact in mind, we examine the behavior of the rate of return (ROR) more closely. This close scrutiny reveals why the ROR does not follow its predicted course, and indicates how to change its course, or planning, depending on the type of business in question.

In addition, the model allows one to observe the optimal investment path, on both the finite and infinite horizons (steady-state values). In the case of the finite horizon, in particular, the model is more qualitative than quantitative. However, in both cases, it allows for a variety of sensitivity analyses.

The study involved two types of input market structures: an input market representing imperfect competition (convex adjustment costs) and a market of scale economies (concave adjustment costs) [3,7].

The cumulative profit to be maximized is then [4,5]:

$$\bar{\pi} = \max_{K,I} \int_0^T e^{-rt} [pF(K) - wL - C(I)] dt + e^{-rT} C_s K(T) + \lambda (-\dot{K} + I - \alpha K) + \mu \left(-\dot{p} + S \left[\frac{sK + wL}{F(K)} - p \right] \right) \quad (4)$$

Equation (4) describes the company's instantaneous cash flow constraints:

$$\dot{K}(t) = I(t) - \alpha K(t) \quad (5)$$

$$\dot{p}(t) = S \left[\frac{sK(t) + wL}{F(K(t))} - p(t) \right] \quad (6)$$

Factors K and L are the two input factors, i.e., capital stock and labor stock, respectively. The $\dot{K}(t)$ function is the time derivative of $K(t)$.

The $p(t)$ function is the unit price. The $\dot{p}(t)$ function is the time derivative of $p(t)$ and states the price adjustment mechanism [8].

As an initial assumption, due to the minor variation in the value of L (number of employees), L is considered constant and can therefore be written as $\Delta = wL$, while F depends only on K .

The F function can be seen as the Cobb Douglas function. Thus,

$$F(K, L) = a_0 K^{a_1} L^{a_2} \quad (7)$$

is the firm's production function.

Since L is a constant value, one has $F_K(K) = \frac{\partial F}{\partial K} > 0$, with $F(0) = 0$ and $F_{KK}(K) = \frac{\partial^2 F}{\partial K^2} < 0$.

We also considered $I(t)$ as the gross investment and $C(I)$ as the total cost investment function, with $C(I) > 0$, $C'(I) = \frac{dC(I)}{dI} > 0$ and $C(0) = 0$.

The remaining parameters are as follows:

- $\bar{\pi}$ is the cumulative profit;
- w is the wage rate;
- C_s is the scrap value of the capital goods;
- S is a speed for price adjustment;
- r is the discount rate;
- α is the depreciation rate;

- s' is the capital return rate, $s' \leq \frac{\pi}{C_a K}$, C_a is the unit value of the capital goods, and π is the annual profit.

We know that

$$\pi = pF - \Delta - \alpha C_a K \quad (8)$$

Thus, s' can be written as:

$$s' \leq \frac{pF - \Delta - \alpha C_a K}{C_a K} \quad (9)$$

Therefore,

$$pF - \Delta - sK \leq 0 \quad (10)$$

Where

$$s = C_a(\alpha + s'), \quad K(0) = K_0 \text{ and } p(0) = p_0 \quad (11)$$

We would like to point out that, in this context, the meanings of equations (6) and (10) are not regulated [1], but there is a planned objective for profit. Thus, it can be demonstrated that, under particular conditions, equations (6) and (10) are equivalent.

After some manipulation, Pontryagin's necessary conditions yield:

$$\dot{q} = (r + \alpha)q - pF_K(K) - \theta S \frac{[F(K)s - (sK + \Delta)F_K(K)]}{[F(K)]^2} \quad (12)$$

$$\dot{\theta} = \theta(S + r) - F(K) \quad (13)$$

$$q = C'(I) \quad (14)$$

where $q = \lambda e^{rt}$ and $\theta = \mu e^{rt}$ are the shadow prices related to $K(t)$ and $p(t)$ respectively.

With regard to the input market (suppliers), two conditions can be stated [7]:

Convex Total Investment Costs

The supplier market is an imperfect capital market, where $C''(I) > 0$ and $C(I)$ is assumed to be: $C(I) = A.I^2 + B.I$, A and B positive constants.

Concave Total Investment Costs

In this case, the costs decrease due to scale economies. Therefore, $C''(I) < 0$ and $C(I) = \sqrt{AI}$, where A is a positive constant. The problem here is a non-convex one. According to [7], it is possible to solve the artificially "convexified" problem by replacing the $C(I)$ function with a function constructed by setting $\frac{\bar{C}(I)}{I} = \bar{q}$ for all I , $0 < I \leq I_0$, where $\bar{q} = \frac{C(I_0)}{I_0}$. Thus, for this "convexified" problem, (\bar{q}, \bar{K}) is a feasible steady-state with a level of investment $\bar{I} = \alpha \bar{K}$ and an average cost of investment \bar{q} .

3 CONDITIONS FOR OPTIMALITY

In addition to the conditions stated in the previous section, several important conditions for the model's application and for investigation are considered in this section, namely: planned ROR, price adjustment and optimal investment path, technological obsolescence involving a depreciation model and various general conditions.

3.1 Planned rate of return

This model works with a planned ROR. Therefore, s' is a planned objective of the company. For instance, if the objective is to achieve a ROR of 12%, the model maximizes the cumulative profit for this kind of premise. In addition, one can write: $s' = M.G$, where $M = \frac{\pi}{pF}$ (profit margin) and $G = \frac{pF}{K}$ (circulation of capital). If s' changes, s also changes because $s = C_a(\alpha + s')$. Thus, it is possible to evaluate the two effects separately to achieve the same s' value. In some situations it is better to vary M and in others it is more advantageous to vary G .

3.2 Price adjustment and optimal investment path

The dynamics that governs the price is that of equation (10), based on the planned ROR (the company's objective). Combining equation (10) with equation (1) and using a numerical technique to converge to a fixed point (which will be explained in greater detail in Section 4), the model leads the price and capital values to converge at an equilibrium point, indicating the optimal investment path.

3.3 Technological obsolescence and depreciation model

The historical depreciation model is not the best one to explain technological obsolescence. Instead, technological obsolescence is best explained by the sum-of-digits model. However, because the time between the several operation ages (vintages) [9,10] is extremely discrepant, it is more convenient to model depreciation as a general model based on the Fisher-Pry model (S shape model). Instead of a constant value of α , the depreciation rate has a time variation. The Fisher-Pry mathematical model can be written as:

$$y(\tau) = \frac{1}{1 + e^{-b(\tau-a)}} \quad (15)$$

where $y(\tau)$ is the fraction of the new technology in time τ . Parameter a describes the time at which the new technology reaches 50% of the entire time between the new and the old technologies. Parameter b measures the speed of substitution. Based on the Fisher-Pry model, the mean depreciation rate can be written as:

$$\alpha(\tau) = \alpha_0 + \frac{\alpha_1}{1 + e^{-b(\tau-a)}} \quad (16)$$

where α_0 is the initial depreciation rate. The final depreciation rate is $\alpha = \alpha_0 + \alpha_1$. Thus, equation (16) can be incorporated into the dynamic equations, keeping in mind that $s = C_a(\alpha + s')$.

Figure 1 shows the evolution of the depreciation rate.

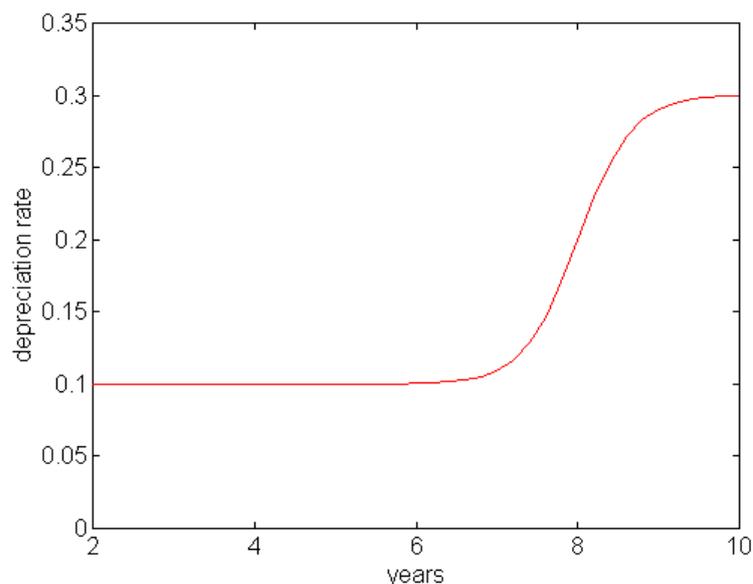


Fig. 1: Depreciation rate evolution

3.4 General conditions

The company is always constrained to satisfy demand and presumably knows all the relevant information about the supply and demand functions (output market, customers). Assuming a demand function $p = a - bQ$, the maximum revenue value occurs at point $a/2b$. Figure 2 indicates that a slight variation around the equilibrium point, assuming an admissible variation in p , causes a slight variation in $Q = F(K)$.

This slight variation in Q causes a slight admissible variation in the revenue. The model exploits this range of variations because the solution depends on a type of contraction to a fixed point before converging to stationary values due to a numerical solution.

Figure 2 illustrates the range of applications of the model. The upper portion of this figure contains the supply and demand functions. The red line corresponds to the supply function and the blue line to the demand function. In the lower portion of this figure, the green line corresponds to the revenue curve as a function of $Q = F(K)$.

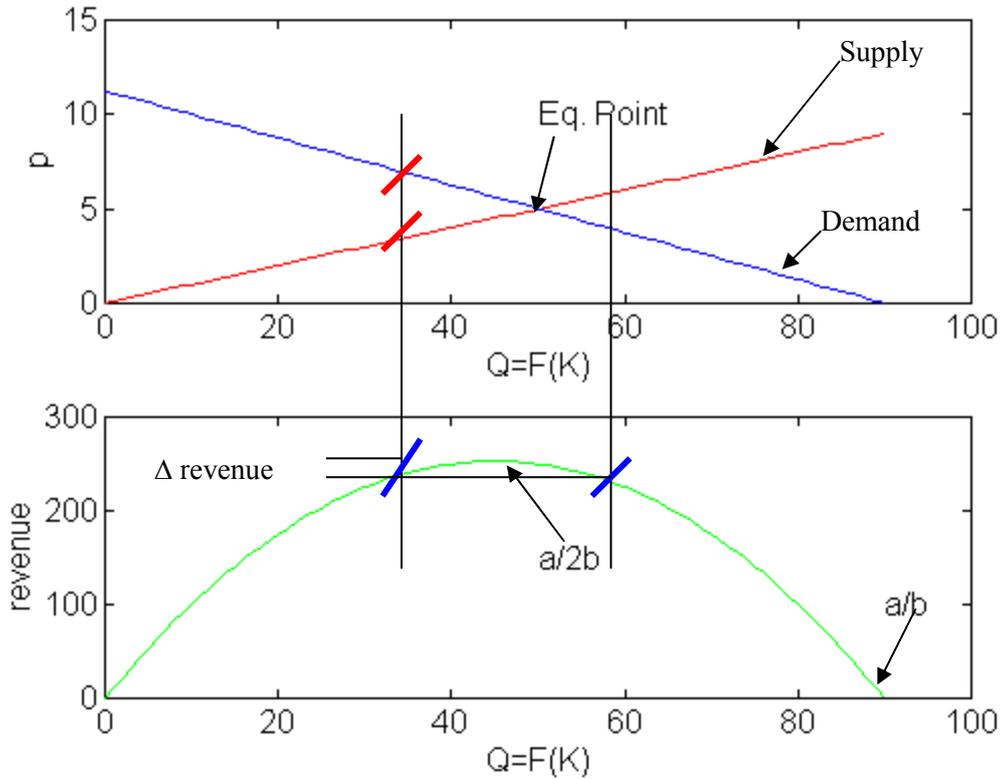


Fig. 2: Range for Model Application

4 SIMULATIONS RESULTS

Assuming a fixed value of $I(t)$, or equivalently, its infinite value, equation (5) can be written as follows:

$$K(t) = \frac{I_\infty}{\alpha} (1 - e^{-\alpha t}) + e^{-\alpha t} K_0 \quad (17)$$

It is interesting to note that $K_\infty = \frac{I_\infty}{\alpha}$, as expected.

The model assumes the maintenance of $I(t)$ as a constant and calculates $K = K(t)$ as a time function for each different time. Proceeding in the same way, equation (6) will become:

$$p(t) = \frac{sK_\infty + \Delta}{F(K)} (1 - e^{-st}) + e^{-st} p_0 \quad (18)$$

Equations (19) and (20) can be treated similarly, but their integration constants must disappear over time. The time it takes for them to reach their stationary values can be investigated (this is a sort of “warm-up” time). Their stationary values will thus be:

$$q_\infty = \frac{1}{r + \alpha} \left[p_\infty F_K(K_\infty) + \theta_\infty S \frac{[F(K_\infty)s - (sK_\infty + \Delta)F_K(K_\infty)]}{[F(K_\infty)]^2} \right] \quad (19)$$

$$\theta_{\infty} = \frac{F(K_{\infty})}{S+r} \tag{20}$$

The four differential equations (17), (18), (19) and (20) are then coupled and $I(t)$ is calculated from equation (14) and with $I(q) = (C'(\dot{I}(q)))^{-1}$. See the additional observation for concave adjustment costs in [7].

The model is actually more qualitative than quantitative, but it gives a good idea about the optimal path and, in addition, it allows for a thorough sensitivity analysis.

The equations resulting from the Hamiltonian approach (12), (13) and (14) and the model's equations (3), (6) and (10) were solved using the Matlab SimulinkTM software package.

Next, we describe some of the results of these simulations and briefly discuss each one.

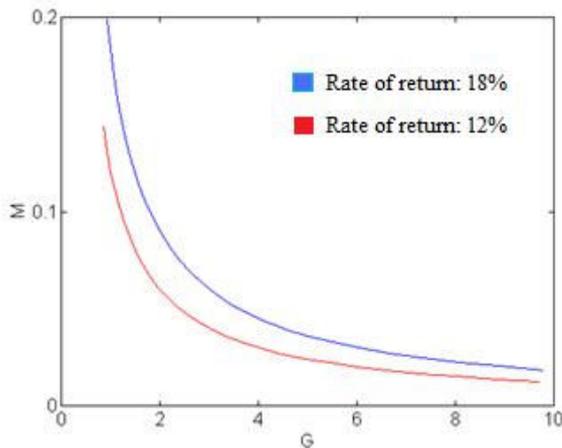


Fig.3: C(I) convex shape

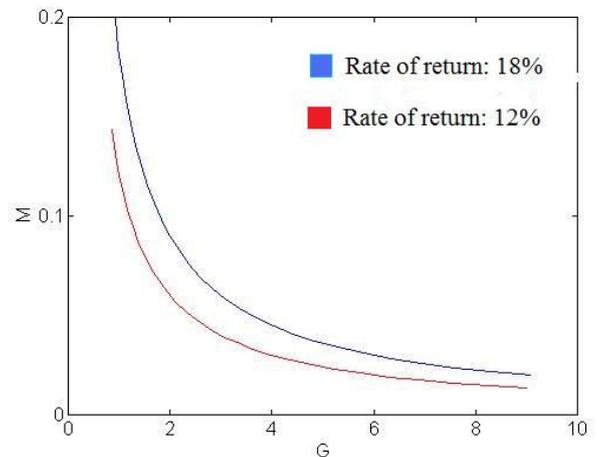


Fig.4: C(I) concave shape

Figures 3 and 4 show that when the ROR increases, the relationship between M and G also changes.

The company can choose how to achieve a ROR of 18%. For instance, it can increase G or, alternatively, it can increase M between the admissible limits.

Price variations achieved by changing the ROR are presented in Figures 5 and 6.

When the ROR s' changes from 12% to 18%, for instance, the price level goes to a higher steady-state value.

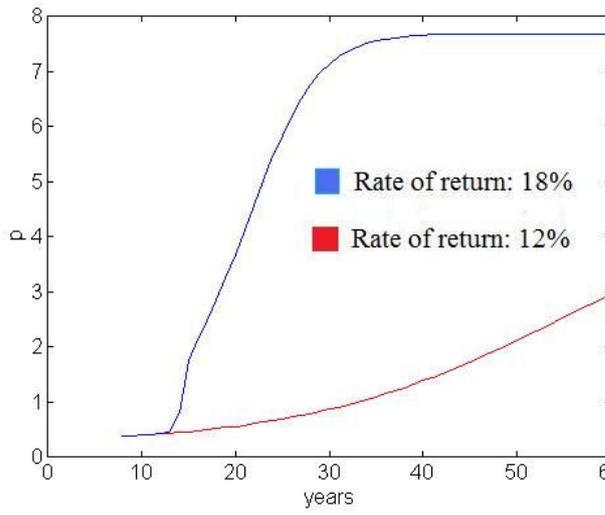


Fig. 5: C(I) concave shape

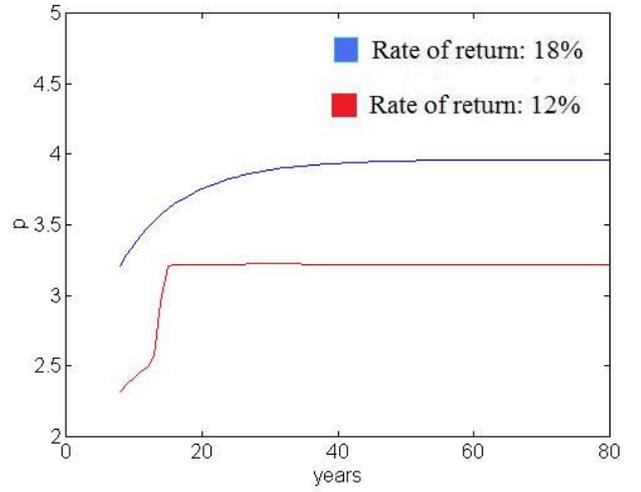


Fig.6: C(I) convex shape

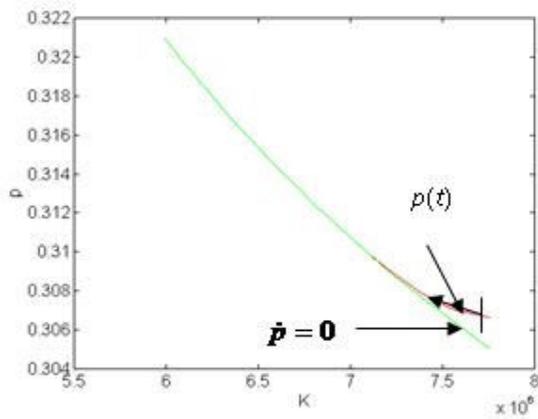


Fig.7:Phase diagram for C(I) convex shape

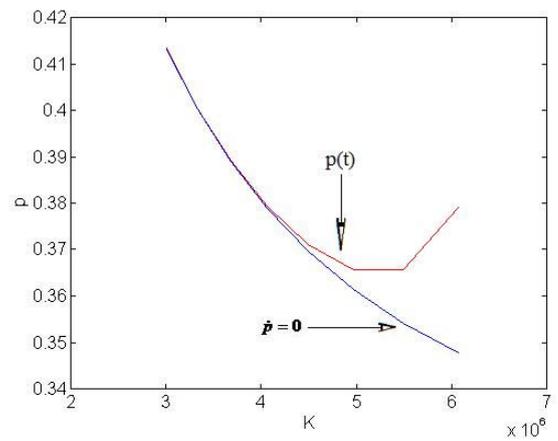


Fig.8:Phase diagram for C(I)concave shape

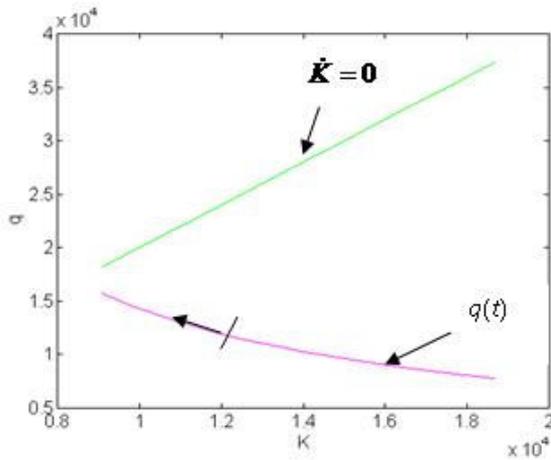


Fig9: Phase diagram for C(I) convex shape and depreciation 0.1

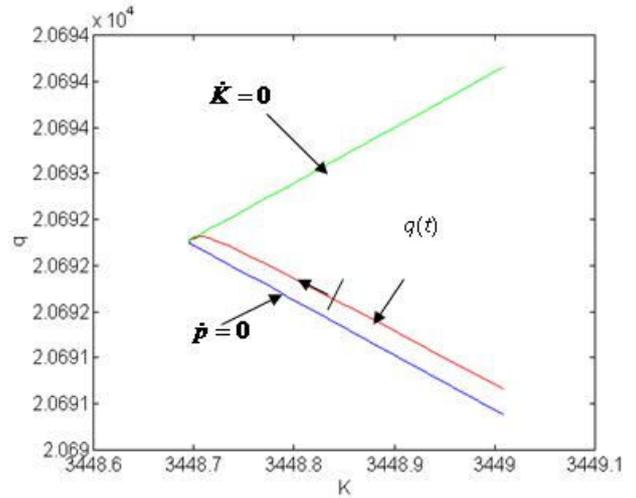


Fig10: Phase diagram for C(I) convex shape and depreciation 0.3

Figure 7 shows the phase diagram of Price (p) vs. Capital (K) for the cost convex shape, with $s'=12\%$ and a depreciation rate of 0.1. In this figure, note that the Price reaches the steady-state.

On the other hand, Figure 8 shows the phase diagram of Price (p) vs. Capital (K) for a cost concave shape with $s'=18\%$ and a depreciation rate of 0.3.

The phase diagram in Figure 9 shows the cost convex shape and a depreciation of 0.1, as well as the investment path to reach the steady-state for a given K_0 value

Figure 10 shows the phase diagram for a cost convex shape and depreciation equal to 0.3. A comparison of this figure with Figure 9 with the same value of K_0 indicates that the steady-state was reached at a lower value. This implies a lower ROR at the equilibrium point.

5 CONCLUSIONS

An evaluation of the model's variations of the steady-state price indicates that the concave cost model is more sensitive and therefore has a higher probability of leaving a given operating point stipulated by the company's objective as the depreciation rate increases.

However, as the depreciation rate increases, two conditions can be observed in the model, i.e., an increase in the steady-state price and an increase in the *circulation of capital* (G), although the *profit margin* (M) may sometimes decrease. In fact, the best situation would be increase in G and M , but some companies have large amounts of capital and which they cannot easily put into circulation.

Another fact about our model is that the stationary point has a lower profit value, since the profit is given by the expression $s'C_a K(t)$, where s' is the planned ROR and C_a is the value of capital goods.

With the increase in ROR, s' , the main condition observed is a price increase. In certain cases, regulatory agencies may establish a maximum limit for prices (or tariffs), causing the planned ROR to become unattainable. Moreover, taking into account possible changes in the

price (or tariff), another relevant issue is undoubtedly the competition (output market). However, a possible shift from the supply and demand operating point can be foreseen.

A possible way to achieve greater accuracy and confidence in the model's application, even for qualitative purposes, would be to make a data survey using regression and data fitting models, employing the two main functions involved: the Cobb-Douglas type $F(K,L)$ function, and the production function (or $F(K)$), if it depends only on K and $C(I)$. The same applies to total investment costs or adjustment costs.

This newly developed model is very comprehensive and allows for the evaluation of other types of parameters, such as the sensitivity of discount or interest rates, r , or variations in the cost of a capital good, C_a .

We hope that the use of this optimization model may make action planning clearer and lead to easier, more accurate and efficient decisions.

REFERENCES

- [1] S. Katayama, *Optimal Investment Policy of the Regulated Firm*, Journal of Economic Dynamics and Control, 13, 1989, pp. 532-552, North Holland.
- [2] J. C. Eckalbar, "Closed-form solutions to bundling problems", *Journal of Economics & Management Strategy*, Vol. 19, Issue 2, pp. 513-544, (2010).
- [3] L.E.Ohanian, E. C. Prescott, N.L. Stokey, "Introduction to dynamic general equilibrium", *Journal of Economic Theory*, Vol. 144, Issue 6, pp. 2235-2246, (2009).
- [4] E. Silberberg, "The structure of economics: A mathematical analysis", 2th Edition, McGraw Hill, (1990).
- [5] H. Goldstein, C.P. Poole, J. Safko, "Classical Mechanics", 3th Edition, Addison Wesley, (2001).
- [6] S. Nickell, *Uncertainty and lags in the investment decisions of firms*, The review of economic studies, vol. 44, pp. 249-263, (1977).
- [7] R. Davidson and R. Harris, *Non-convexities in Continuous-time Investment Theory*, Review of Economic Studies 48, pp. 235-253, (1981).
- [8] E. G. Davis, *A Dynamic Model of the Regulated Firm with a Price Adjustment Mechanism*, The Bell Journal of Economics and Management Science, pp. 270-282, (1972).
- [9] L. K. Vanston, R. L. Hodges, *Depreciation Lives for Telecommunications Equipment*, Technology Futures, Inc. (1996)
- [10] S. L. Barreca, *Comparison of Economic Life Techniques*, Technology Futures, Inc. (1999).

MODERNIZING SCIENCE&ENGINEERING SOFTWARE SYSTEMS

LILIANA FAVRE^{*}, LILIANA MARTINEZ[†] AND CLAUDIA PEREIRA[†]

^{*} Universidad Nacional del Centro de la Provincia de Buenos Aires
Comisión de Investigaciones Científicas de la Provincia de Buenos Aires
Tandil, Argentina
e-mail: lfavre@exa.unicen.edu.ar

[†] Universidad Nacional del Centro de la Provincia de Buenos Aires
Tandil, Argentina
email: {cpereira, lmartine}@exa.unicen.edu.ar}

Key words: Reverse Engineering, Model Driven Architecture, Software Modernization, Modeling, Reliability.

Abstract. As the demands for modernized legacy systems rise, so does the need for frameworks for information integration and tool interoperability. The Object Management Group (OMG) has adopted the Model Driven Architecture (MDA), which is an evolving conceptual architecture that aligns with this demand. MDA could help solve coupling problems of multidisciplinary character in science and engineering that consist of one or more applications, supported by one or more platforms. The objective of this paper is to describe rigorous techniques to control the evolution from science & engineering software legacy systems to MDA technologies. We propose a rigorous framework to reverse engineering code in the context of MDA. Considering that validation, verification and consistency are crucial activities in the modernization of systems that are critical to safety, security and economic profits, our approach emphasizes the integration of MDA with formal methods.

1 INTRODUCTION

Nowadays, software industry is evolving and tackling new approaches aligned with Internet, object orientation, distributed components, new modeling languages and new platforms. However, the majority of the large engineering applications running today in many research organizations were developed many years ago with technology that is now obsolete. These old systems, known as legacy systems, are still business-critical. New platforms and applications must interoperate with legacy software systems. Their complete replacement is dangerous and their maintenance is increasingly expensive. That is the reason for the demand of modernization of legacy system to extend their useful lifetime [18].

As the demands for modernized legacy systems rise, so does the need for frameworks for information integration and tool interoperability. The Object Management Group (OMG) has adopted the Model Driven Architecture (MDA), which is an evolving conceptual architecture that addresses the challenges of networked and changing system environments. MDA could

help solve coupling problems of multidisciplinary character in science and engineering that consist of one or more applications, supported by one or more platforms [14].

The outstanding ideas behind MDA are separating the specification of the system functionality from its implementation on specific platforms, managing the software evolution from abstract models to implementations. MDA shifted the center of software development from code to models. In the MDA context, code is considered as an artifact derived from model transformations. The concepts of model, metamodel and model transformation are central in MDA. Models play a major role in MDA which distinguishes at least the following models:

- Platform Independent Model (PIM): a model with a high level of abstraction that is independent of any implementation technology.
- Platform Specific Model (PSM): a tailored model to specify the system in terms of the implementation constructs available in one specific platform.
- Implementation Specific Model (ISM): a description (specification) of the system in source code.

The initial diffusion of MDA was focused on its relation with UML as modeling language [20,21]. However, there are UML users who do not use MDA, and MDA users who use other modeling languages such as Domain Specific Languages (DSL). The essence of MDA is MOF (Meta Object Facility) metamodel that allows different kinds of software artifacts to be used together in a single project [15]. The basic idea is to create a common specification for communication between applications. The MOF 2.0 Query, View, Transformation (QVT) metamodel is the standard for expressing transformations [17].

Validation, verification and consistency are crucial activities in the modernization of legacy systems that are critical to safety, security and economic profits. Reasoning about models of systems is well supported by automated theorem provers and model checkers, however these tools are not integrated into CASE tools environments.

The objective of this paper is to describe rigorous techniques to control the evolution from science & engineering software legacy systems to MDA technologies. We propose a rigorous framework to reverse engineering code in the context of MDA that emphasizes the integration of techniques related to MDA, such as metamodeling, with formal methods.

The paper is organized as follows. In the next section we provide background material on modernization of systems and Case tools. The next four sections consist of the main ideas of our work. Section 3 describes a framework for reverse engineering. Section 4 describes how to transform code into models through static and dynamic analysis. Section 5 and Section 6 describe how to formalize reverse engineering processes in terms of MOF metamodels and formal specifications respectively. Finally, conclusions are drawn in Section 7.

2 RELATED WORK

The article [4] compares existing work, discusses success and provides a road map for possible future developments in the area. The reengineering of a deteriorated object oriented industrial program written in C++ is described in [9]. In order to deal with this problem, the authors designed and implemented several restructuring tools and used them in specific reengineering scenarios. A variety of techniques for object oriented reengineering based on patterns are distinguished in [7].

Many works are linked to MDD-based reverse engineering. A rigorous framework for automatic legacy system migration in MDA called MOMENT is described in [3]. A tool-assisted way of introducing models in the migration towards MDA is presented in [12]. The article [13] describes model-based dynamic analysis techniques that relate system execution traces and its models such as testing whether a system run satisfies a property that a certain model specifies and measuring how various model features materialize in a system run.

The material presented in [19] is based on techniques developed during a collaboration with CERN (Conseil Européen pour la Recherche Nucléari) in the introduction of tools for software quality assurance, among which a reverse engineering tool called RevEng was presented. This tool extracts UML diagrams (class diagrams, object diagrams, state diagrams, sequence and collaboration diagrams and package diagrams) from C++ code.

Modernizing large industrial software systems is impossible without appropriate tool support. All of the MDA tools are partially compliant to MDA features. They provide good support for modeling and limited support for automated transformation in reverse engineering.

Techniques that currently exist in MDA CASE tools provide little support for validating models in the design stages. Reasoning about models of systems is well supported by automated theorem provers and model checkers, however these tools are not integrated into CASE tools environments. Another problem is that as soon as the requirements specifications are handed down, the system architecture begins to deviate from specifications. Only research tools provide support for formal specification and deductive verification. Many CASE tools support reverse engineering, however, they only use more basic notational features with a direct code representation and produce very large diagrams.

To be able to reason about software systems, MDA CASE tools need a common information base aligned to MOF. However few MDA-based CASE tools support MOF and QVT or at least, any of the QVT languages. As an example, IBM Rational Software Architect and Spark System Enterprise Architect do not implement QVT. Other tools partially support QVT, for instance Together allows defining and modifying transformations model-to-model (M2M) and model-to-text (M2T) that are QVT-Operational compliant [5].

The Eclipse Modeling Framework (EMF) [8] was created for facilitating system modeling and the automatic generation of Java code. EMF started as an implementation of MOF resulting Ecore, the EMF metamodel comparable to EMOF. EMF has evolved starting from the experience of the Eclipse community to implement a variety of tools and to date is highly related to Model Driven engineering (MDE). Commercial tools such as IBM Rational Software Architect, Spark System Enterprise Architect or Together are integrated with Eclipse-EMF [5].

3 A RIGOROUS FRAMEWORK FOR REVERSE ENGINEERING

We propose to reverse engineering MDA models from object oriented code starting from the integration of compiler techniques, metamodeling and formal specification. Figure 1 shows a framework for reverse engineering that integrates static and dynamic analysis, metamodeling and formal specification. It distinguishes three different abstraction levels linked to models, metamodels and formal specifications.

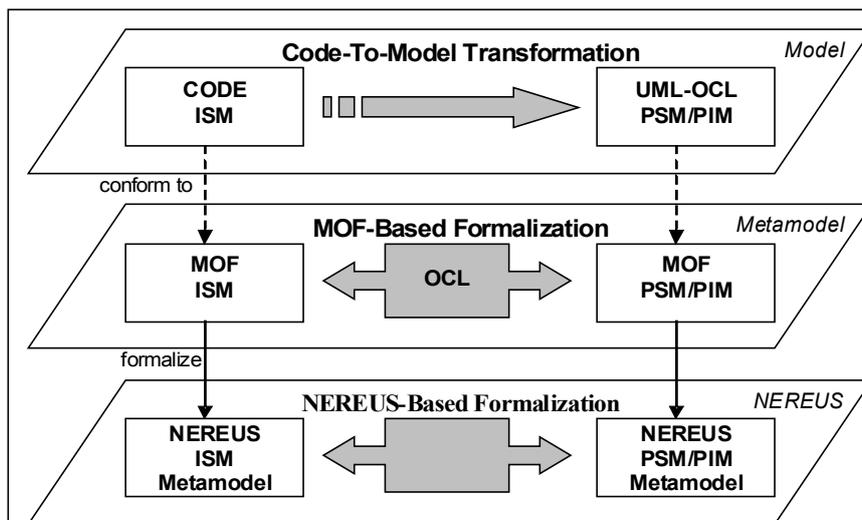


Figure 1. MDA-based reverse engineering

The model level includes code, PIM and PSM. A PIM is a model with a high level of abstraction that is independent of an implementation technology [14]. A PSM is a tailored model to specify a system in terms of specific platform such J2EE or .NET. PIM and PSM are expressed in UML and OCL [20,21,16]. The subset of UML diagrams that are useful for PSM includes class diagrams, object diagrams, state diagrams, interaction diagrams and package diagrams. On the other hand, a PIM can be expressed by means of use case diagrams, activity diagrams, interactions diagrams to model system processes and, state diagrams to model lifecycle of the system entities. An ISM is a specification of the system in source code.

At model level, transformations are based on classical compiler construction techniques. They involve processes with different degrees of automation, which can go from totally automatic static analysis to human intervention requiring processes to dynamically analyze the resultant models. All the algorithms that deal with the reverse engineering share an analysis framework. The basic idea is to describe source code or models by an abstract language and perform a propagation analysis in a data-flow graph called in this context object-data flow. This static analysis is complemented with dynamic analysis supported by tracer tools.

The metamodel level includes MOF metamodels that describe the transformations at model level [15]. A metamodel is an explicit model of the constructs and rules needed to construct specific models. MOF metamodels use an object modeling framework that is essentially a subset of UML 2.3 core [20]. The modeling concepts are classes which model MOF metaobjects, associations, which model binary relations between metaobjects, data types which model other data, and packages which modularize the models. At this level MOF metamodels describe families of ISM, PSM and PIM. Every ISM, PSM and PIM conforms to a MOF metamodel. Metamodel transformations are specified as OCL contracts between a source metamodel and a target metamodel. MOF metamodels “control” the consistency of these transformations.

The formal specification level includes specifications of MOF metamodels and metamodel transformations in the metamodeling language NEREUS that can be used to connect them with different formal and programming languages [10,11].

To sum up, at model level, instances of ISM, PSM and PIM are generated by applying static and dynamic analysis. Static analysis builds an abstract model of the state and determines how the program executes to this state. Dynamic analysis operates by executing a program and evaluating the execution traces of the program. Contracts based on MOF-metamodels “control” the consistency of these transformations and NEREUS facilitates the connection of the metamodels and transformations with different formal languages.

Our work could be considered as an MDA-based formalization of the process described by Tonella and Potrich in [19]. Additionally, we propose a different algorithm for extracting UML State diagrams and new processes to recover PIM including use case diagrams and activity diagrams. We also propose to include OCL specifications (preconditions, postconditions and invariants) in UML diagrams. Other contributions are linked to the automation of the formalization process and interoperability of formal languages [10,11].

The following sections describe reverse engineering at three different abstraction levels corresponding to code-to-model transformation, MOF-metamodel formalization and algebraic formalization.

4 RECOVERING MODELS FROM CODE

At model level, transformations are based on static and dynamic analysis. Static analysis extracts static information that describes the software structure reflected in the software documentation (e.g., the text of the source code) whereas dynamic analysis information describes the structure of the run-behavioral. Static information can be extracted by using techniques and tools based on compiler techniques such as parsing and data flow algorithms. Dynamic information can be extracted by using debuggers, event recorders and general tracer tools.

We suppose that the reverse engineering process starts from an ISM that could reflect, for instance, the migration of legacy code to object oriented code. The first step in the migration towards MDA is the introduction of PSMs. Then, a PIM is abstracted from the PSMs omitting platform specific details.

Next, we describe the process for recovery PSMs from code. Figure 2 shows the different phases. The source code is parsed to obtain an abstract syntax tree (AST) associated with the source programming language grammar. Then, a metamodel extractor extracts a simplified, abstract version of a language that ignores all instructions that do not affect the data flows, for instance all control flows such as conditional and loops.

The information represented according to this metamodel allows building the data-flow graph for a given source code, as well as conducting all other analysis that do not depend on the graph. The idea is to derive statically information by performing a propagation of data. Different kinds of analysis propagate different kinds of information in the data-flow graph, extracting the different kinds of diagrams that are included in a PSM.

The static analysis is based on classical compiler techniques and abstract interpretation [1]. On the one hand, data-flow graph and the generic flow propagation algorithms are specializations of classical flow analysis techniques [1]. On the other hand, abstract interpretation allows obtaining automatically as much information as possible about program executions without having to run the program on all input data and then ensuring computability or tractability. These ideas were applied to optimizing compilers.

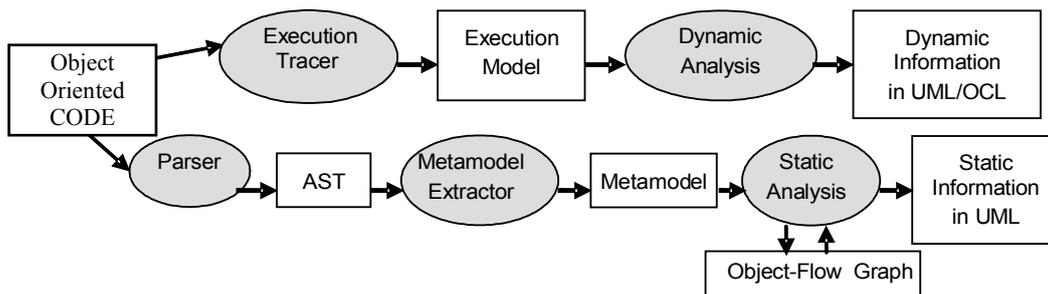


Figure 2. From code to models: Static and dynamic analysis

The static and dynamic information could be shown as separated views or merged in a single view. In general, the dynamic behavior could be visualized as execution sceneries which describe interaction between objects. To extract specific information, it is necessary to define particular views of these sceneries. Although, the construction of these views can be automated, their analysis requires some manual processing in most cases.

In the MDA context, we can distinguish code-based dynamic analysis and model-based dynamic analysis. The first is based on an execution model including the following components: a set of objects, a set of attributes for each object, a location and value of an object type for each object, and a set of messages. On the other hand, model-based dynamic analysis is based on model-level debugging and evolution.

4.1 Recovering Class Diagrams

A class diagram is a representation of the static view that shows a collection of static model elements, such as classes, interfaces, methods, attributes, types as well as their properties (e.g., type and visibility). Besides, the class diagram shows the interrelationships holding among the classes [21].

Reverse engineering of class diagrams from code is a difficult task that cannot be automated due to certain elements in the class diagram carry behavioral information that cannot be inferred just from the analysis of the code. By analyzing the syntax of the source code, internal class features such as attributes and methods and their properties (e.g. the parameters of the methods and visibility) can be recovered. From the source code, associations, generalization, realizations and dependencies may be inferred too. However, to distinguish between aggregation and composition we need to capture system states through dynamic analysis. Figure 3 shows relationships that can be detected statically between a C++ program and a UML class diagram.

Dynamic analysis allows generating execution snapshot to collect life cycle traces of object instances and reason from tests and proofs. Execution tracer tools generate execution model snapshots that allow us to deduce complementary information, for instance information to detect compositions. A composition is a particular aggregation in which the lifetime of the part is controlled by the whole (directly or transitively) and we can identify it by generating tests and scanning dependency configurations between the birth and the death of a part object according to those of the whole. In the same way, the execution traces of different instances of the same class or method could guide the construction of invariants or preconditions and postconditions respectively.

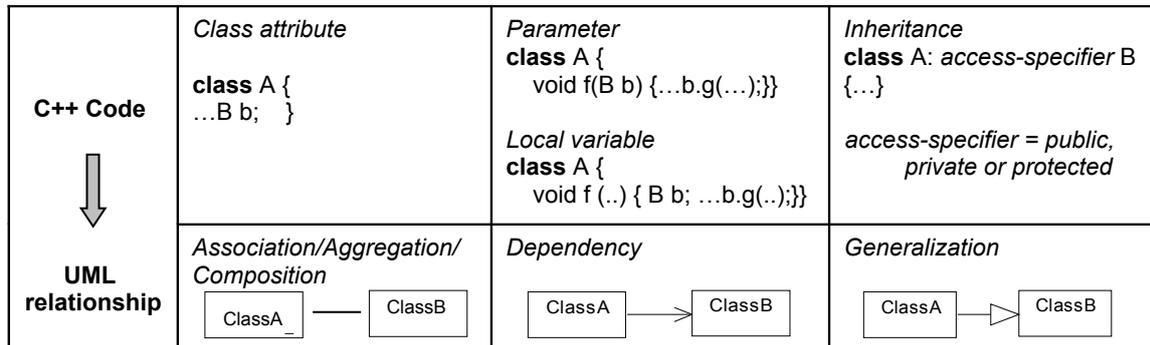


Figure 3. From C++ code to UML class diagram

4.2 Recovering State Diagrams

A state transition diagram describes the life cycle of objects that are instances of a class from the time they are created until they are destroyed. Object state is determined by the value of its attributes and possibly by the variables involved in attribute computations. The basic elements of a state diagram are states, identified as equivalence classes of attribute values, and transitions triggered by method invocation.

Our approach to recover state diagrams has similar goals to abstract interpretation that allows obtaining automatically as much information as possible about program executions without having to run it on all input data and then ensuring computability or tractability. These ideas were applied to optimizing compilers, often under the name data-flow analysis [1]. In our context, an abstract interpretation performs method invocation using abstract domains instead of concrete attribute values to deduce information about the object computation on its actual state from the resulting abstract descriptions of its attributes. This implies to abstract equivalence classes that groups attribute values corresponding to the different states in which the class can be and the transitions among state equivalence classes.

Then, the first step is to define an appropriate abstract interpretation for attributes (which give the state of the object) and modifier class method (which give the transitions from state to state to be represented in the state diagram).

The recovery algorithm iterates over the following activities: the construction of a finite automata by executing abstract interpretations of class methods and the minimization of the automata for recovering approximate state equivalence classes. To ensure tractability, our algorithm proposes an incremental minimization every time a state is candidate to be added to the automaton. When it is detected that two states are equivalents, they are merged in an only state. This could lead to modification of the parts of the automaton that had been previously minimized. To optimize the comparison of pairs of states, these are classified according to their emerging transitions. Let m be a bound of the number of transformer methods of a class, the idea is to generate subsets of the set of transformer methods. The subset of emerging transitions of a new state belongs, in a particular snapshot, to one of them. Two states are candidates to be equivalent if they belong to the same subset. Then, it is sufficient to compare all the pairs composed by the state and one element of the subset. Considerable human interaction to select which abstract interpretations should be executed is required [6].

As an example, Figure 4 (a) shows a diagram including states (s_1, s_2, \dots, s_8) and transitions (m_1, m_2, \dots, m_6). Fig. 4 (b) shows a simplified snapshot of the automaton when a transition to s_5 is added. Then, the shaded states could belong to the same equivalence state class. s_8 belongs to the same subset of s_4 and an equivalence analysis is carried out concluding that s_8 and s_4 can be merged. Figure 4 (c) (d) (e) shows the successive transformations.

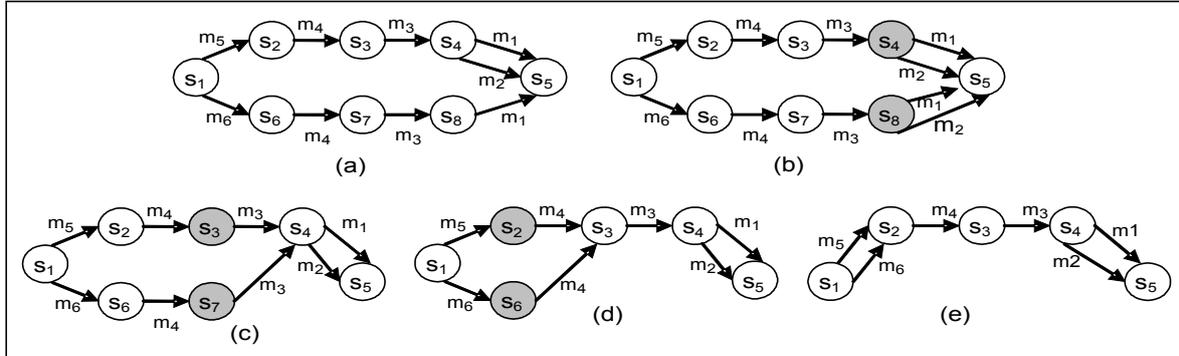


Figure 4. Recovering minimum State Diagram

A C++ implementation was developed to test the feasibility of our recovering algorithm. In the following, the algorithm for the identification of the states is described:

```

/* initialization of different sets*/
set-of-states initialStates = {}; /* states defined by class constructors*/
set-of-states pendingStates = {}; /* set of states pending of analysis*/
set-of-states allStates = {}; /* set of all states*/
/*definition of initial states for the objects of the class*/
for each class constructor c
    /*executing an abstract interpretation of each class constructor*/
    state s = abstractInterpretationState (c, {});
    initialStates = initialStates  $\cup$  {s};
    pendingStatesPending = pendingStates  $\cup$  {s};
    allStates = allStates  $\cup$  {s}
endfor
set-of-transitions transitionSet = {}; /* initialization of transition set*/
set-of-bins b = classifiedStates (allStates); /*generation of subsets of transformer methods*/
while |pendingStates| > 0
    state r = extract (pendingStates);
    pendingState = pendingStates - {r};
    for each transformer class method m
        /*generating transitions of the state r*/
        s = abstractInterpretationState (m, r);
        if s  $\notin$  allStates
            pendingStates = pendingStates  $\cup$  {s};
            allStates = allStates  $\cup$  {s}
        endif
    transitionSet = transitionSet  $\cup$  abstractInterpretationTransition (m,r,s)
endfor
/* updating subsets of transformer methods*/
b = modifyBins (r, transitionSet, allStates);
for each e  $\in$  b
    if s  $\in$  b /*defining equivalence of states and merging equivalent states*/
        for each q  $\in$  bin and s < q
            if equivalent(s, q) mergeStates (transitionSet, allStates, s, q) endif
        endfor
    endif endfor endwhile

```

5 REVERSE ENGINEERING AT METAMODEL LEVEL

We specify reverse engineering processes as MOF-defined transformations. MOF allows capturing all the diversity of modeling standards and interchange constructs that are used in MDA. We call anti-refinement the process of extracting from a more detailed specification (or code) another one, more abstract, that is conformed by the more detailed one.

Figure 5 shows partially an ISM-C++ metamodel that includes constructs for representing classes, variables and functions. It also shows different kind of relationships between classes, for instance, a C++ class can have super classes or nested classes. Figure 6.a shows partially a PSM-C++ metamodel that includes constructs for representing classes, attributes, association ends and functions. The main difference between an ISM-C++ and a PSM-C++ is that the latter includes constructs for associations. The state diagram metamodel (Figure 6.b) defines a set of concepts that can be used for modeling discrete behavior through finite state transition systems such as state machines, states and transitions. We specify metamodel-based model transformations as OCL contracts that are described by means of a transformation name, parameters, preconditions, postconditions and additional operations. Transformation semantics is aligned with QVT, with the QVT Core in particular. In Figure 7 we partially exemplify a transformation from an ISM-C++ to a PSM-C++. This transformation uses both the specialized UML metamodel of C++ code and the UML metamodel of a C++ platform as source and target parameters respectively. The postconditions state relations at metamodel level between the elements of the source and target model. The transformation specification guarantees that for each class in C++ code there is a class in the PSM-C++, both of them with the same name, the same parent class, equivalent operations and so on. Besides, the PSM-C++ has a 'stateMachine' for each class having a significant dynamic behavior.

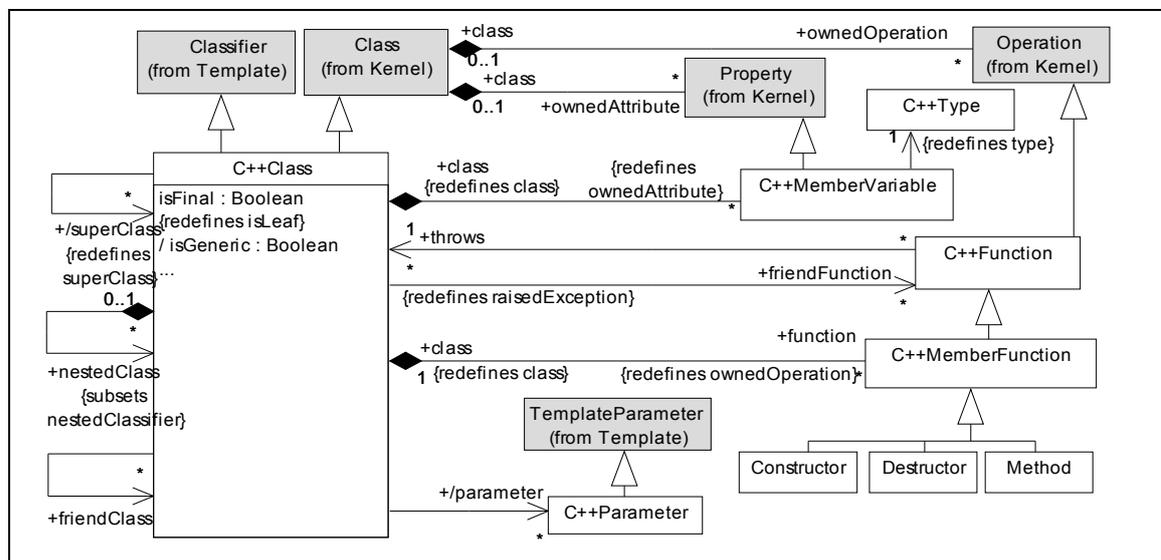


Figure 5. ISM C++ Metamodel

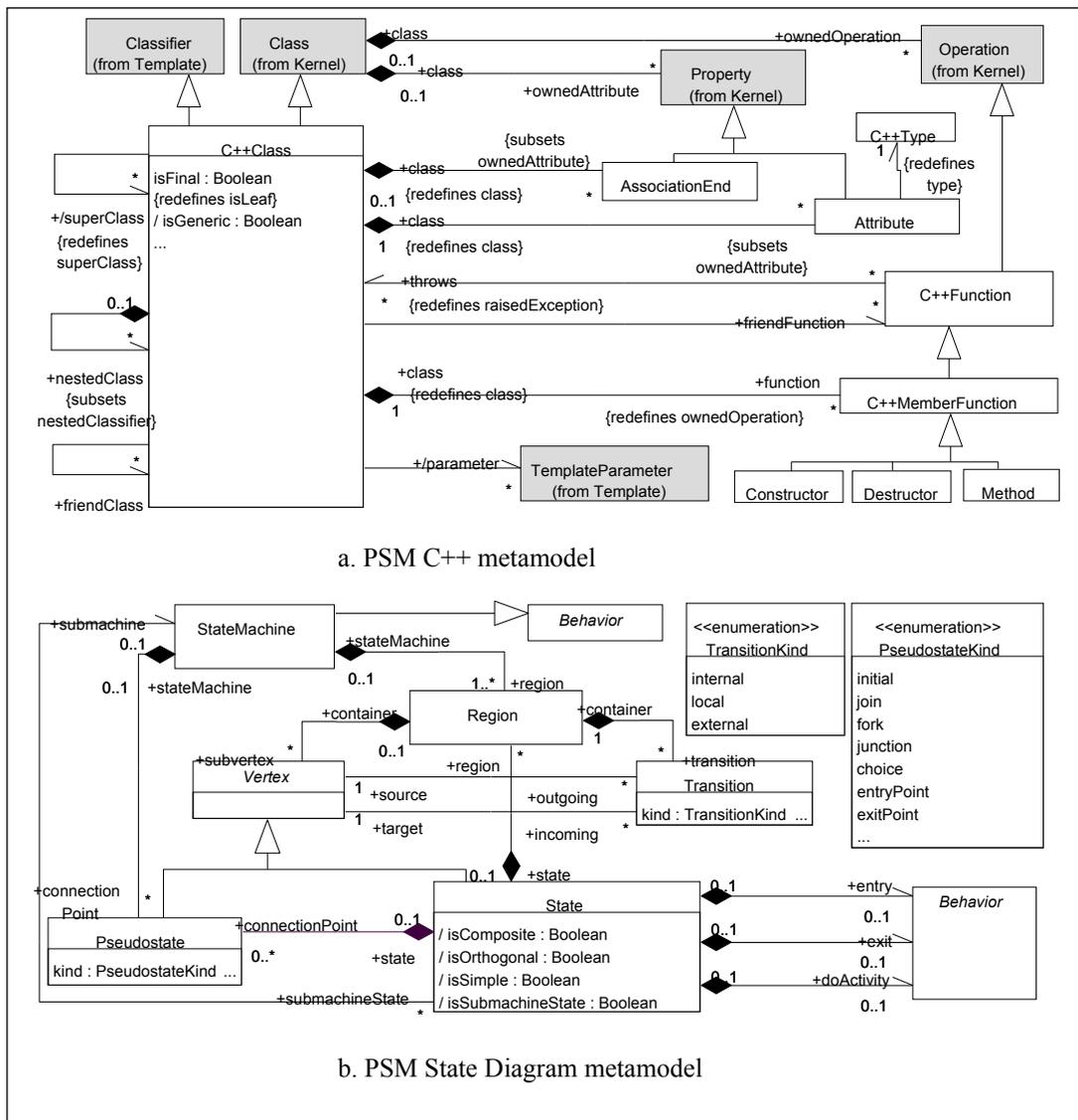


Figure 6. PSM Metamodel

6 TRANSFORMATION AT FORMAL LEVEL

This work is strongly integrated with previous ones that show how to formalize metamodels and metamodel-based transformations in terms of a formal metamodeling language called NEREUS. This language takes advantage of existing theoretical background on formal specifications and can be integrated with different algebraic languages.

NEREUS focused on interoperability of formal languages in MDD. A detailed description may be found at [10] that defines a bridge between MOF metamodels and NEREUS consisting of a system of transformation rules to convert automatically MOF into NEREUS. Also, the articles [10,11] shows how to integrate NEREUS with the Common Algebraic Specification Language (CASL) [2]. On the other hand, NEREUS allows specifying metamodels such as the Ecore metamodel, the specific metamodel for defining models in EMF (Eclipse Modeling Framework) [8]. Today, we are integrating NEREUS in EMF.

```

Transformation ISM-C++ to PSM-C++ {
parameter   sourceModel: ISM-C++Metamodel:: C++Project
              targetModel: PSM-C++Metamodel:: Package
postconditions
/* For each class 'sourceClass' in the sourceModel*/
sourceModel.ownedMember->select(oclIsTypeOf(C++Class))->forall(sourceClass |
/*there is a class 'targetClass' in the targetModel so that both classes have the same name,*/
  targetModel.ownedMember-> select(oclIsTypeOf(C++Class))-> exists (
    targetClass | targetClass.name = sourceClass.name and
/*For each superClass of 'sourceClass' there is a superClass in 'targetClass' so that both super classes are
equivalent*/
  sourceClass.superClass-> forall ( superClass | targetClass.superClass-> exists (
    superCtarget | superCtarget.classMatch(superCsource) ) and
/*For each member function of 'sourceClass' there is an operation in targetClass so that both operations are
equivalent*/
  sourceClass.C++MemberFunction->forall(sourceF | targetClass.C++MemberFunction ->
    exists(targetF | targetF.operationMatch(sourceF) ) ) and
/*For each variable in 'sourceClass' whose type is a primitive or library type there is an attribute in
'targetClass' so that the attribute conform to the variable*/
  sourceClass.C++MemberVariable-> select ( v |
    v.C++Type.oclIsTypeOf(Primitive) or v.C++Type.oclIsTypeOf(Library) ) -> forall
    (sourceVar | targetClass.attribute -> exists ( targetAtt |
      targetAtt.conformTo (sourceVar) ) ) and
/*For each variable in 'sourceClass' whose type is a user defined type there is an association end in
'targetClass' so that the association end correspond to the variable:*/
  sourceClass.C++MemberVariable->select(v | v.C++Type.oclIsTypeOf(UserC++Class) )
    ->forall (sourceVar | targetClass.associationEnd -> exists ( targetAssocEnd |
      targetAssocEnd.correspondTo (sourceVar) ) ) and
/*If 'sourceClass' has some significant dynamic behavior, targetModel has a 'stateMachine' so that:*/
  sourceClass.hasSignificantDynamicBehavior() implies
    targetModel.ownedMember->select(oclIsTypeOf(StateMachine))-> exists (targetMachine |
/*'targetMachine' and 'sourceClass' have the same name and*/
    targetMachine.name = sourceClass.name and
/*For each modifier function in the 'sourceClass' there is a transition in 'targetMachine'*/
    sourceClass.C++MemberFunction-> select (f| f.isModifier() )-> forall( f |
      targetMachine.region.transition-> exists( t | t.isCreatedFrom(f)) and ... )
and ...
}

```

Figure 7. ISM-C++ to PSM-C++ Transformation

7 CONCLUSIONS

MDA addresses the challenges of networked, changing system environment, providing a conceptual architecture that promotes portability, cross-platform interoperability, platform independent, better code quality and easier maintainability. Therefore, it is possible to define Domain-Specific languages adapted to new, industry-specific applications over diverse platforms.

This paper describes rigorous techniques to control the evolution of legacy systems in the context of science and engineering applications. We propose an MDA-based framework to reverse engineering code. Considering that testing, verification and consistency analysis are crucial activities in the modernization of systems that are critical to safety, security and economic profits, our approach stresses the integration of historic reverse engineering

techniques with MDA and formal specifications.

We describe how to transform code into models in a higher level of abstraction that allows moving from these abstractions to new implementations.

REFERENCES

- [1] Aho, A., Sethi, R. and Ullman, J. *Compilers: Principles, Techniques, and Tools* (Edition 2). Addison-Wesley (1985).
- [2] Bidoit, M. and Mosses, P. *CASL User Manual- Introduction to Using the Common Algebraic Specification Language* (LNCS 2900). Heidelberg: Springer-Verlag (2004).
- [3] Boronat, A., Carsi, J. and Ramos, I. Automatic reengineering in MDA using rewriting logic a transformation engine. *Proceedings of the Ninth European Conference on Software Maintenance and Reengineering (CSMR'05)* Los Alamitos: IEEE Computer Society (2005) 228-231.
- [4] Canfora, G. and Di Penta, M. New Frontiers of reverse Engineering. Future of Software Engineering. *Proceedings of Future of Software Engineering, FOSE 2007*. Los Alamitos: IEEE Press (2007) 326-341.
- [5] CASE TOOLS (2010). www.modelbed.net/mda_tools.html
- [6] Daciuk, J. Incremental Construction of Finite-State Automata and Transducers, and their use in the Natural Language Processing. Ph. D. Thesis. Technical University of Gdansk (1998).
- [7] Demeyer, S., Ducasse, S. and Nierstrasz, O. *Object oriented Reengineering Patterns*. Amsterdam: Morgan Kaufmann (2002).
- [8] Eclipse: *The eclipse modeling framework* (2010) <http://www.eclipse.org/emf/>
- [9] Fanta, R. and Rajlich, V. Reengineering of object oriented code. *Proceedings of the International Conference on Software Maintenance*. IEEE Press (1998) 238-246.
- [10] Favre, L. A Formal Foundation for Metamodeling. Reliable Software Technologies- ADA Europe 2009. LNCS 5570. Springer-Verlag (2009) 177-191.
- [11] Favre, L. *Model Driven Architecture for Reverse Engineering Technologies: Strategic Directions and System Evolution*. Engineering Science Reference, USA (2010).
- [12] MacDonald, A., Russell, D. and Atchison, B. Model driven Development within a Legacy System: An industry experience report. *Proceedings of the 2005 Australian Software Engineering Conference*. ASWEC 05. Los Alamitos: IEEE Press (2005) 14-22.
- [13] Maoz, S. and Harel, D. On Tracing Reactive Systems. *Software and Systems Modeling*, Vol 9, Springer-Verlag (2010)
- [14] MDA. *The Model Driven Architecture* (2005) www.omg.org/mda
- [15] MOF: *Meta Object facility 2.0*. OMG Specification formal/2006-01-01 (2006).
- [16] OCL: *Object Constraint Language. Version 2.2*. OMG: formal/2010-02-01 (2010).
- [17] QVT: *MOF 2.0 Query, View, Transformation*. Formal/2008-04-03 (2008)
- [18] Sommerville, I. *Software Engineering* (7th ed.). Addison Wesley (2004).
- [19] Tonella, P. and Potrich, A. *Reverse Engineering of Object Oriented Code*. Monographs in Computer Science. Heidelberg: Springer-Verlag (2005).
- [20] UML *Infrastructure*. OMG Specification formal/ 2010-05-03 (2010).
- [21] UML *Superstructure*. OMG Specification: formal/2010-05-05 (2010).

MULTI-PHYSICS MODELLING AND SIMULATION OF SAND TRANSFER AROUND CUBE

KAZUTO MATSUI^{*}, MASAYA SUZUKI[†] AND MAKOTO YAMAMOTO[‡]

Department of Mechanical Engineering
Tokyo University of Science
1-14-6 Kudankita, Chiyoda-ku, Tokyo, 102-0073, Japan

^{*} e-mail: j4510648@ed.kagu.tus.ac.jp

[†] e-mail: masaya@rs.kagu.tus.ac.jp

[‡] e-mail: yamamoto@rs.kagu.tus.ac.jp

Key words: Multi-Physics Simulation, Sand Transfer, Gas-Solid Two-Phase Flow, Turbulent Flow, Computational Fluid Dynamics, Desertification.

Abstract. Yearly, concerns on environmental problem of the earth are rapidly growing on. One of the typical issues is desertification. To stop or suppress the harmful effects of desertification, the prediction method which can clarify the mechanism of desertification is required. For the purpose it is expected that numerical simulations are very useful. A number of research papers on sand transfer have been published in the literatures. For example, Yuasa et al. reasonably simulated the sand-transfer experiment conducted by Tominaga. However, they assumed that the sand particle trajectory is parabolic. This assumption would oversimplify the physics. Our final goals are to construct the multi-physics simulation technique which can correctly reproduce sand transfer, and, by applying it, to develop an effective prevention method for desertification. The purposes in the present study are to improve the particle trajectory computation, and to simulate the sand transfer around a cube more correctly. Comparing the numerical results with the experiments, we confirmed that the present method can quantitatively simulate the sand transfer phenomena.

1 INTRODUCTION

In late years, the various environmental problems of the earth are rapidly growing on. One of the typical issues is desertification. Desertification causes decline of agricultural productivity, disappearance of road and town, and so on. To stop or suppress the progress of desertification, a lot of prevention activities have been applied such as tree plantation, wind-breaking fence and so on. However, desertification does not stop and is expanding on 10-20 percent of arid region every year, according to Millennium Ecosystem Assessment (2006). The main process of desertification is sand transfer. Therefore, the researches on the method to predict and suppress sand transfer blown by wind are needed against desertification. In the past, this kind of researches has mainly been performed with experimental methods. On the other hand, theoretical studies which satisfy experimental observations are not adequate (e.g. effect of tree plantation). Therefore, it is thought that more holistic knowledge or

understandings about sand transfer might help us to create an effective prevention method of desertification. For this purpose, a wide variety of experiments and established prediction methods are required. One of the useful prediction methods for sand transfer is CFD (Computational Fluid Dynamics). Advantages of CFD are that it is easy to apply in various conditions and reasonable cost to perform simulations as compared with experiments. In our final purpose, employing an engineering approach with CFD, we would like to look for an effective prevention method of desertification.

The pioneering researcher for sand transfer by wind is Bagnold [1]. He provided mechanism and mathematical models for sand mass transfer blown away by wind. Kawamura [2] modified the model, introducing the idea of threshold friction velocity. This modified model is useful to simulate the sand surface geometry. Yuasa et al. [3] employed their modelling associated with existing models, and computed sand transfer phenomena around a cube. The numerical results were well compared with the experimental results conducted by Tominaga [4]. However, their numerical procedure and the physical models for predictions assumed that the particle trajectory is parabolic. The assumption would over-simplify the physics.

Our final goals are to establish the multi-physics simulation technique which can correctly reproduce sand transfer phenomena, and, by applying it, to develop an effective prevention method for desertification. The purposes in the present study are to improve the particle trajectory computation, and to simulate the sand transfer around a cube more correctly. The turbulent flow, the sand particle trajectory and the temporal change of sand surface are iteratively computed with a RANS technique. Comparing the numerical results with the experiments measured by Tominaga [2], it is confirmed that the present method can quantitatively predict the sand transfer phenomena, especially the temporal change of sand surface geometry.

2 NUMERICAL PROCEDURES

2.1 Assumption

The number of sand particles in a desert is infinite. If all sand particles are traced, we must spend a huge amount of time for simulations. Thus, in the present study sand particles are treated as a lump. In each grid point, mass and direction of sand transfer are estimated from the flow field. Since the sand transfer needs a long period and the time scale is much longer than that of the flow field, steady flow field is assumed for sand surface geometry. That is, sand transfer depends on the flow field in a quasi-steady state. The flow field is turbulent. Although we can use Large Eddy Simulation (LES) or Reynolds-Averaged Navier-Stokes Simulation (RANS), considering the short computing time, RANS is employed.

In this study, we employ the computational procedures for sand transfer as follows:

- (1) Simulate turbulent air flow field
- (2) Simulate sand transfer
- (3) Simulate sand surface geometry
- (4) Change computational grid
- (5) Return to (1)

These procedures are iteratively repeated, until the computational time reaches the prescribed terminal time.

2.2 Flow field

We assume that the flow field is incompressible and fully turbulent and sand particles have no influence on the flow field because of the low concentration of sand. The governing equations are Reynolds-averaged continuity equation and Navier-Stokes equation (i.e. RANS), as described above. The governing equations are expressed by

$$\frac{\partial U_i}{\partial x_i} = 0 \quad (1)$$

$$\frac{DU_i}{Dt} = -\frac{1}{\rho} \frac{\partial P}{\partial x_i} + \frac{\partial U_i}{\partial x_i} \frac{\partial}{\partial x_i} \left\{ v \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} \right) - \overline{u_i u_j} \right\} \quad (2)$$

where U_i is the mean velocity component in x_i -direction, t is time, ρ is air density, P is pressure, v is kinetic viscosity coefficient and $\overline{u_i u_j}$ is Reynolds stresses. Reynolds stresses are expressed as follows,

$$\overline{u_i u_j} = \frac{2}{3} k \delta_{ij} - v_t \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} \right) \quad (3)$$

where δ_{ij} is Kronecker delta and v_t is eddy kinetic viscosity coefficient.

We use RNG k - ε turbulence model proposed by Yakhot-Orszag [5] to predict Reynolds stresses. Then the turbulence model equations are as follows:

$$\frac{Dk}{Dt} = D_k + P_k - \varepsilon \quad (4)$$

$$\frac{D\varepsilon}{Dt} = D_\varepsilon + \frac{\varepsilon}{k} (C_{\varepsilon 1} P_k - C_{\varepsilon 2} \varepsilon) \quad (5)$$

$$v_t = C_\mu \frac{k^2}{\varepsilon} \quad (6)$$

where k is turbulent kinetic energy, ε is its dissipation rate, C_μ is model constant and has the value of 0.085, P_k is production term, D_k and D_ε are diffusion terms, as follows,

$$P_k = \frac{1}{2} v_t \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} \right)^2 \quad (7)$$

$$D_k = \frac{\partial}{\partial x_j} \left\{ \left(\frac{v_t}{\sigma_k} + v \right) \frac{\partial k}{\partial x_j} \right\} \quad (8)$$

$$D_\varepsilon = \frac{\partial}{\partial x_j} \left\{ \left(\frac{v_t}{\sigma_\varepsilon} + v \right) \frac{\partial \varepsilon}{\partial x_j} \right\} \quad (9)$$

$$C_{\varepsilon 1} = 1.42 - C_{1R} \quad (10)$$

$$C_{1R} = \frac{\eta(1 - 0.228\eta)}{1 + 0.015\eta^3} \quad (11)$$

$$\eta = \sqrt{2S_{ij}S_{ij}} \frac{k}{\varepsilon} \quad (12)$$

$$S_{ij} = \frac{1}{2} \left(\frac{\partial U_i}{\partial x_j} + \frac{\partial U_j}{\partial x_i} \right) \quad (13)$$

where σ_k , σ_ε and $C_{\varepsilon 2}$ are 0.7179, 0.7179, and 1.92, respectively.

These governing equations are solved by using the MAC method.

2.3 Sand transfer

According to the research by Bagnold [1], there are three types of sand transfer. They are suspension, saltation and surface creep. This classification depends on sand particle diameter. Suspension means that sand particles are suspended by wind in long stretches, which has less than about 20 μm in diameter. It is noted that suspension brings on yellow sand phenomena in Japan. Saltation means that sand particles jump on sand surface by wind, which has about 70-500 μm in diameter. Surface creep means that sand particles are rolled on sand surface by wind, which has more than about 500 μm . Generally, surface creep is generated by saltation. Bagnold [1] found that saltation is dominant in sand transfer. Therefore, we assume that sand is transferred only by saltation.

Bagnold [1] proposed the equation which expresses the relation between the sand transfer and the friction velocity as follows:

$$q = C \sqrt{\frac{d}{D}} \frac{\rho}{g} u_*^3 \quad (14)$$

where q is the mass transfer of sand, d is the mean particle diameter, D is the reference particle diameter which is equal to 0.25 mm, ρ is air density, g is gravitational acceleration, u_* is friction velocity, and C is a constant that ranges from 1.5 to 2.8 depending upon the distribution of sand particles being transported. C has the value of 1.5 for uniform sand particles. Friction velocity is calculated by the following log law,

$$\frac{u_z}{u_*} = \frac{1}{\kappa} \ln \left(\frac{z}{z_0} \right) \quad (15)$$

where u_z is mean wind velocity over vertical height z from sand surface, κ is a constant which has 0.4 and z_0 is roughness length.

Equation (14) has been modified by a lot of experiments. Kawamura [4] proposed a model with an idea to describe the threshold friction velocity,

$$q = C \frac{\rho}{g} (u_* - u_{*t})(u_* + u_{*t})^2 \quad (16)$$

where u_{*t} is threshold friction velocity which means the minimum velocity that is able to move sand particles and C is a constant. He suggested that C is 2.78. In this study, we use Kawamura's equation to estimate sand mass transfer.

Sand particles dash out from a point of the sand surface, draw the arc, and collide with another point of the sand surface. Yuasa et al. [3] assumed the particle trajectory as a parabola.

Under this assumption, saltation length and height are as follows:

$$L_s = \alpha \frac{u_*^2}{g}, H_s = \beta \frac{u_*^2}{g} \quad (17)$$

where α and β are constants of 7.2 and 0.5, respectively. However, this assumption would over-simplify the physics.

In this study, the particle-phase is treated by the Lagrangian approach, in which each particle is tracked over time along their trajectories in the flow field. We assume that the particles are spherical and non-rotating. As mentioned above, particle-particle collisions and the interaction with the flow field from the particle-phase are negligible because the solid loading is sufficiently small. In addition, the pressure gradient force, the added mass force, and the Basset history force are ignored because the difference between the density of solid particles and the fluid is large and the free-stream velocity is high (approximately 10 m/s). Under these assumptions, the equations of particle motion are as follows:

$$\frac{dx_{p,i}}{dt} = u_{p,i} \quad (18)$$

$$\frac{du_{p,i}}{dt} = F_{D,i} + F_{B,i} \quad (19)$$

where subscript p denotes the particle; $F_{D,i}$, $F_{B,i}$ are the drag force and the buoyancy force, respectively. The fluid velocity used for calculating the drag force is the time-averaged velocity obtained by RANS. The changes of sand surface geometry are simulated by

$$\rho_s \frac{dh}{dt} = -\frac{dQ}{dz} \quad (20)$$

where ρ_s is density of sand, h is the normal distance from the base plane, Q is source of sand mass transfer and z is specified length scale. In the present study, z is treated as the first grid point length from the sand surface.

Additionally, it is well known that there is sand mass transfer called avalanche. When the slope of sand surface becomes locally larger than the repose angle, avalanche arises, and mass of sand propagates downward so as to keep the repose angle. Generally it occurs when the slope angle is more than 32-35 degrees [6]. In this study, if the slope of sand exceeds the repose angle of 35 degrees, the height of sand at the grid point is changed artificially so as to keep the repose angle and also satisfy the mass conservation of the sand.

3 COMPUTATIONAL CONDITIONS

Tominaga [4] carried out the wind tunnel experiment for the sand transfer around a cube. We follow the experimental conditions and simulate the sand transfer with our modeling. The computational domain is three dimensional, and it is shown in Figure 1. The conditions are listed in Table 1. Figure 2 exhibits the computational grid used in all computations. The grid number of the flow field is 161×77×56. Sand layer whose initial thickness is 0.03 m evenly covers on the ground surface. If sand height is less than or equal to zero, the sand mass transfer does not occur and the ground surface is not eroded. This limitation is same as that in

Table 1 Computational conditions

Length	3.0	[m]
Width	0.9	[m]
Height	0.9	[m]
Cube	0.2	[m]
Initial sand height	0.03	[m]
Free stream velocity	12.0	[m/s]

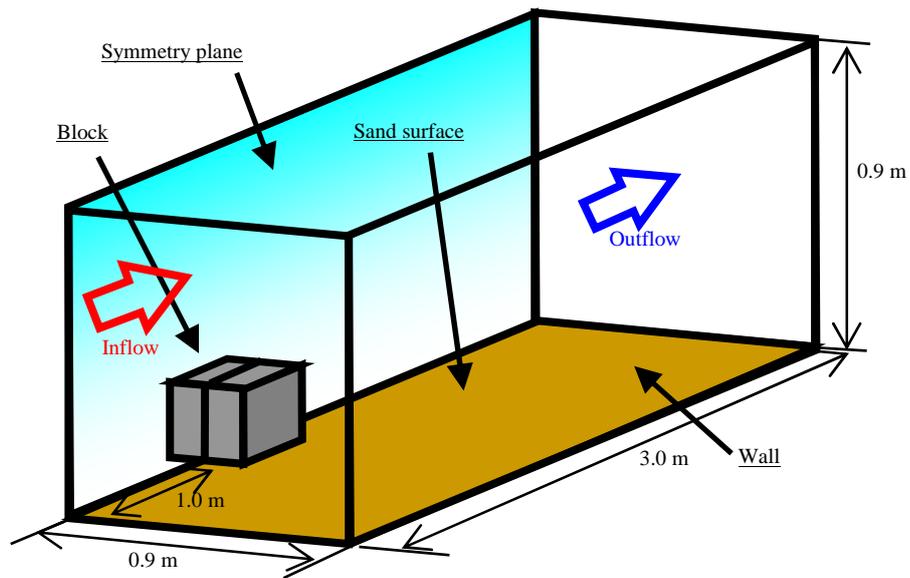


Figure 1 Schematic of computational domain

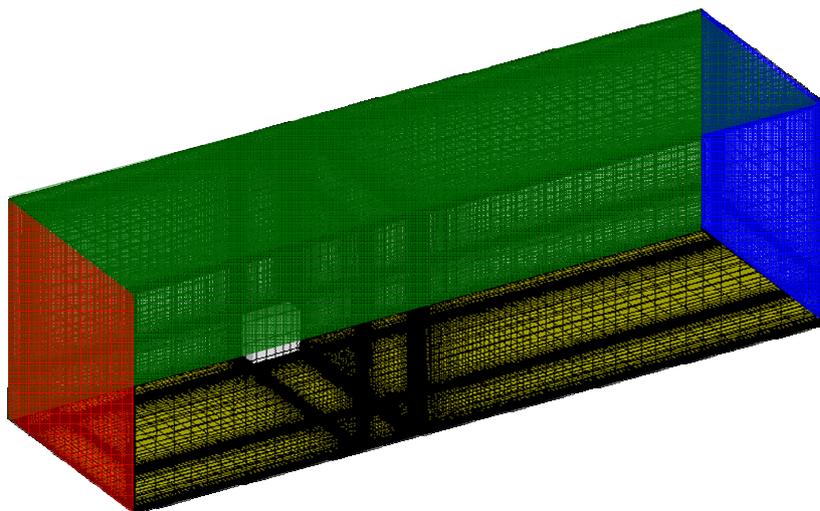


Figure 2 Computational grid around a cube

the experiment. The cube is mounted on the ground surface. At the beginning, a part of the cube (i.e. 0.03 m from the ground) is buried in the sand layer.

The following boundary conditions are applied. At the inflow boundary, inflow velocity, turbulent kinetic energy and its dissipation rate are fixed. In addition, pressure is extrapolated from the computational region and the velocity profile is assumed to be the one-fourth power law according to the experiment. At the outflow boundary, the velocity in the free stream direction and other variables are extrapolated from the computational region. On the sand surface, the cube surface and the wall boundary, non-slip condition is applied for the velocity, and the pressure is extrapolated from the computational domain. The turbulent kinetic energy and its dissipation rate are estimated by using the wall function. At the upper and symmetry boundaries, slip condition is imposed, and thus there is no outflow from these boundaries. For other variables same conditions for the outflow boundary are used.

There is no sand supply from the inflow boundary, and sand transfer whose saltation length goes over the outflow boundary is not taken into account. This means that the total mass of sand particles always decreases through the computation. In addition, if a saltation particle hits against the cube surface, we assume that elastic collision occurs.

4 NUMERICAL RESULTS AND DISCUSSION

Figure 3 shows the initial streamlines around a cube, colored by wind speed (see color bar). It is apparent that a lot of vortices are generated because of the influence of the cube. We can find a strong horseshoe vortex in the windward and lateral sides of the cube. And two vortices exist on the side surface. In the back of the cube, a large recirculation can be observed. It is not clear, but we can confirm a separation vortex on the roof of cube. Obviously, these vortices influence the velocity profile around the cube, and thus the friction velocity changes drastically. We can easily expect that, since high velocity generally leads to high friction velocity, severe aeolian erosion occurs around high speed region. On the windward region of the cube, the horseshoe vortex makes high velocity near the sand surface, and enhances active

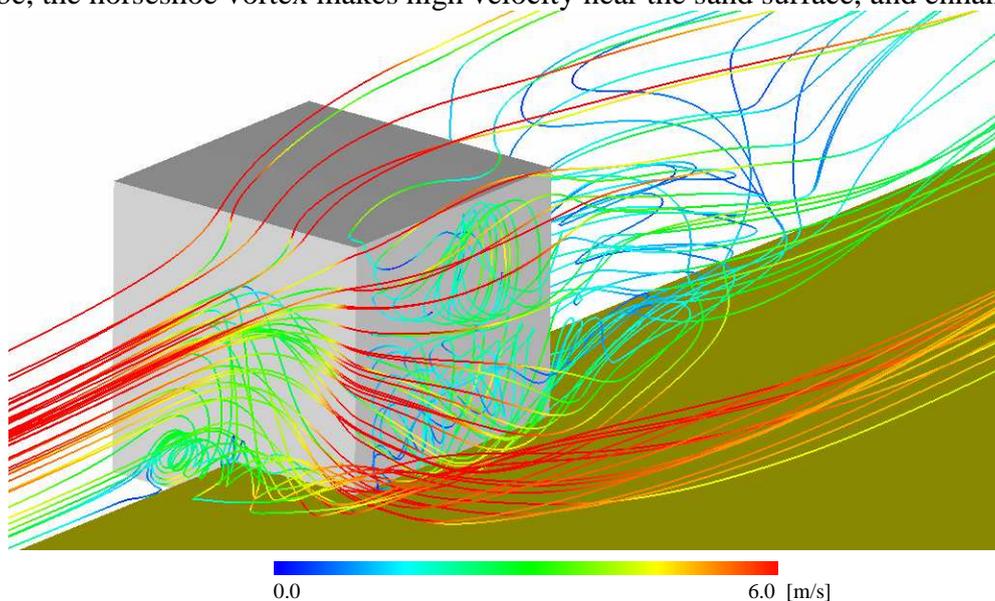


Figure 3 Streamlines colored by wind speed ($t=0$)

active sand transfer in the opposite direction of the free stream. Therefore, aeolian erosion can be found in the windward side of the cube. Moreover, the flow velocity rapidly increases in the vicinity of the windward corner. As a result, sand transport is promoted in the vicinity of the windward corner, and aeolian erosion is intensely generated there.

Figure 4 shows the temporal change of sand surface around the cube after five and ten minutes. Color denotes the height of the sand surface from the ground level. It is clear that the strongest sand transfer takes place around the windward corner of the cube, and the second strongest sand transfer does behind the cube. The former comes from the horseshoe vortex, and the latter caused from the separation vortex on the back of the cube. On the other hand, the sand transfer by saltation makes some crests around the cube, especially behind the cube and the windward region of the horseshoe vortex. Roughly speaking, the erosion pattern does not depend on time. We confirmed that these results are in agreement with the experimental data conducted by Tominaga [4].

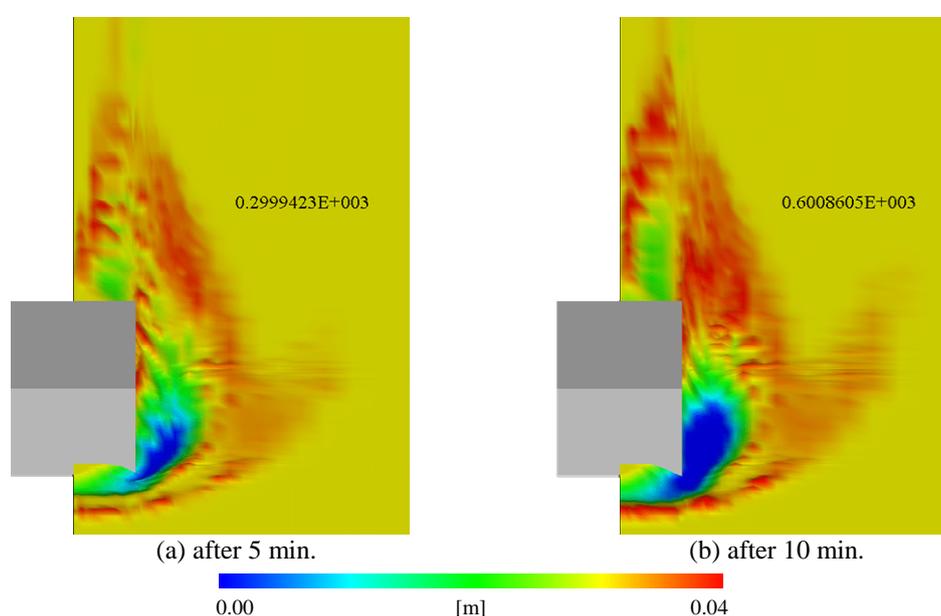


Figure 4 Height of sand surface at different time

Figures 4(a) and 4(b) exhibit the heights of the sand surface in the symmetry plane and in the central cross section on the lateral side of the cube, respectively. The present results are compared with the experimental data by Tominaga [4] and also the computational results by Yuasa et al. [3]. In the figures, blue symbols and lines correspond to the results at five minutes, and red ones correspond to those at ten minutes. Globally, the present results reasonably agree with the experiments. Especially, in the symmetry plane, our results are clearly better than those by Yuasa et al. [3]. However, in the lateral cross section, our method overestimates sand transfer and thus the erosion depth. Probably, this is because the threshold value of friction velocity in our model is lower than the experiment. Therefore, in order to obtain more realistic results, we have to improve the physical models, by properly estimating the mass of sand transfer (i.e. Eq.(16)) and the threshold friction velocity. In addition, we should point out that turbulence model may affect the results.

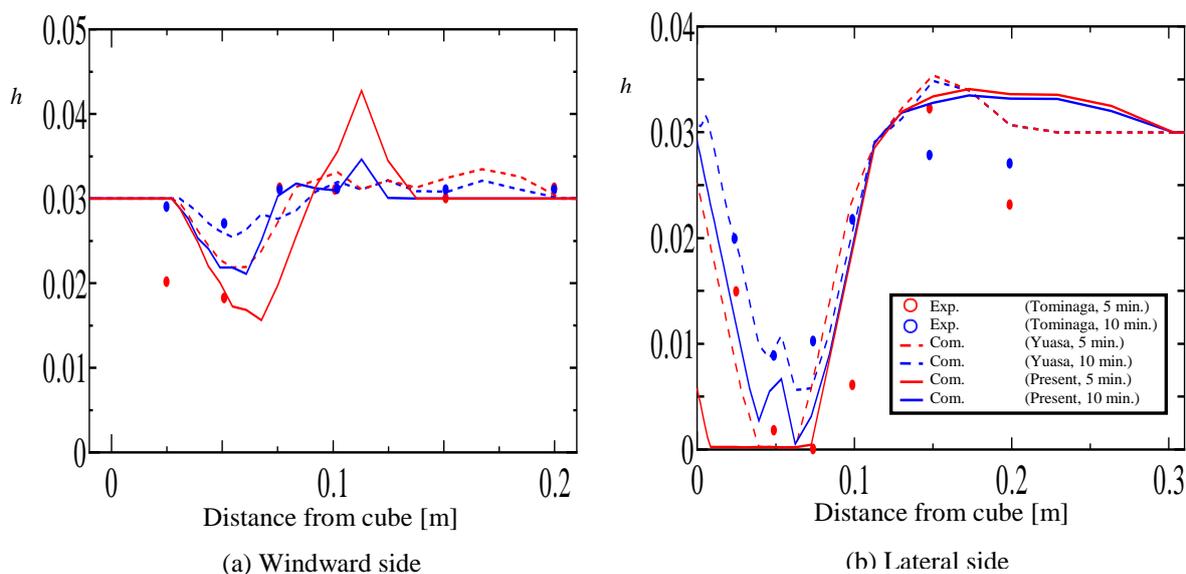


Figure 5 Comparison of height of sand surface

5 CONCLUSIONS

We carried out the three-dimensional and multi-physics computations for the sand transfer around a cube. Through the present study, we obtained following remarks:

- Aeolian erosion around the cube is reasonably predicted by the present method.
- Aeolian erosion is severe around the windward corner of the cube and behind the cube. These are caused from the vortex motions of the flow field.
- Sand accumulation is observed in the windward of the horseshoe vortex and behind the cube. Saltation particles make remarkable crests.
- The erosion depth in the symmetry plane matches with the experiment.
- It is necessary to modify our models for sand transfer and the method to estimate the threshold friction velocity, in order to obtain more accurate sand transfer.

REFERENCES

- [1] Bagnold, R. A., *The physics of blown sand and desert dune*. Methuen, London. (1941).
- [2] Kawamura, R., *Study on sand movement by wind*. The institute of science and technology, university of Tokyo (1951) **5**:95-112.
- [3] Yuasa, T., Yamamoto, M. and Kawaguchi, Y., *Numerical Simulation of Sand Transfer in Desert*. Proceedings of the 23rd Computational Fluid Dynamics Symposium (2009), 1-6
- [4] Tominaga, S. *Chanel experiment and CFD analysis on erosion and bank of sand around object*. Niigata Institute of Technology (2007) **11**:43-50.
- [5] Yakhot, V. and Orszag, S. A. *Renormalization group analysis of turbulence. I. Basic theory*. Journal of Scientific Computing (1986)
- [6] Kroy, K., Sauermaun, G. and Herrmann, H. J. *Minimal model for sand dunes*. Review letter (2002)

RECURRENCE PLOTS ANALYSIS OF PRESSURE FLUCTUATIONS IN FLUIDIZED BEDS

M. Tahmasebpour, R. Sotudeh-Gharebagh, R. Zarghami, N. Mostoufi

Multiple Systems Research Laboratory, School of Chemical Engineering,
University of Tehran, Tehran, Iran
e-mail: tahmasebpour@ut.ac.ir

Key words: Fluidized bed, Recurrence plot, Recurrence quantification analysis, Pressure fluctuation.

Summary. *Recurrence plot (RP) and recurrence quantification analysis (RQA), as powerful statistical techniques, have been used for studying the dynamic behavior of gas-solids fluidized beds. The method of delays was used to reconstruct the state space attractor to carry out analysis in the reconstructed state space. In this work, variance of recurrence rate, which indicates density of recurrence points in RP, against different epoch lengths (time windows) for time series of pressure fluctuation of fluidized bed was calculated. It was concluded that the characteristic parameters of RPs could reflect the extent of chaos in fluidization behavior. The average cycle frequency and entropy as nonlinear dynamical invariants were evaluated with RQA at different aspect ratios. The estimated entropy showed a similar trend of average cycle frequency for different aspect ratios. The results also indicated that the entropy and average cycle frequency are higher in smaller aspect ratios showing that the importance of the finer structures. In addition, a minimum in average cycle frequency and entropy of the pressure fluctuations indicated a minimum deviation from periodicity or, in other words, a minimum deviation from the larger structures, of the bed. The results of this study allow the deep understanding the fluidized bed hydrodynamics which can then be used for scale up.*

1 INTRODUCTION

Fluidized bed reactors have a numerous advantages over other reactor types that make them suitable for industrial applications. They have good particle mixing, high heat and mass transfer rates in addition to low pressure drop. However, due to complexity of the hydrodynamics, design and scaling of this type of chemical reactor are still not straightforward [1-3]. The governing equations of fluidized bed system are rather complex. Since the performance of a fluidized bed is dependent on their hydrodynamic states of fluidization, many investigations reported the hydrodynamic properties of fluidized bed properly such as transition velocities, bubble and cluster characteristics. There are many techniques to determine the hydrodynamic properties of fluidized bed such as optical fiber probes, pressure fluctuations measurements and etc. However, a great advantage of the pressure signals is that they are easy to measure consisting different dynamic phenomena taking place in the bed, such as bubble formation, bubble coalescence and splitting, bubble passage as well as particles behaviors [4].

Traditionally, time series of pressure signals are analyzed using spectral (e.g. Fourier transform, power spectrum) or statistical (e.g. standard deviation, averages) analysis. These analysis techniques assume that the irregular time dependant behavior is due to the linear summing up of random and periodic fluctuations. These techniques do not include the complex hydrodynamic of fluidized beds [1, 5]. Most researchers who investigated fluidized bed based on pressure fluctuations have accepted them as a nonlinear system [5-9]. The new technique that takes account of the nonlinearity of the dynamics is called chaos analysis, in comparison to statistical and spectral analysis [1].

All methods of nonlinear time series analysis are based on the attractor reconstruction of the underlying system in the state space. However, different reconstruction methods may lead to different embedding parameters. In other words, these methods are accompanied by some limitations such as uncertainty on attractor reconstruction methods [9]. Many researchers believe that the two-phase structure of the fluidized bed has a low-dimensional chaotic behavior (typically more than 3 and less than 5) in the state space [9-13]. Thus, attractors with dimensions more than three can be figured only by projection into the two or three-dimensional spaces. On the other hand, long-term data sampling, which is required for typical nonlinear evaluation of the pressure fluctuations in bubbling fluidized bed is usually involved with some difficulties (e.g. steady state sampling with practical fluctuation feed flow, data saving, data acquisition, etc.) during experimental measurement [12, 14-15].

Recurrence is a basic property of dynamical systems, which can be exploited to describe the system's behavior in phase space [16]. While in the state space, attractors with dimensions more than three cannot be visualized due to constrains in figuring the high dimensional attractors, any phase space trajectory can be represented in a 2-dimensional plot using recurrence plot (RP). In Addition, while embedding is required for reconstruction of attractor in state space, RP may be constructed without embedding. All information contained in the embedded RP can be attained in the non-embedded one [17]. Moreover, the remarkable properties of RP are its ability to evaluate non-stationary and short-term data [18, 19]. These features make RP a very potent tool to study fluidized bed hydrodynamics and eliminates needs for time consuming and difficult long-term data sampling required in typical methods of nonlinear analysis. The aim of this work is to apply the RP and recurrence quantification analysis (RQA) to study of scaling aspect ratio (L/D) effect on the dynamic features of the gas-solids fluidized bed using the local pressure fluctuation signals.

2 RECURRENCE QUANTIFICATION ANALYSIS (RQA)

2.1 Recurrence plots

RP technique, derived from nonlinear properties, is based on a graphical explanation of system's dynamics. Ekmann et al. [16] introduced the conception of recurrence plot, as a graphical tool that can determine recurrent behavior in a phase-space of a dynamical system. Briefly, a RP provides a qualitative picture of the correlations between the states of a time series over all available time-scales. A phase-space is generally a high dimensional space and can only be visualized by the projection onto smaller two- or three-dimensional sub-spaces. RPs enable investigation of a m-dimensional phase-space trajectory through a two-dimensional representation of its recurrences to be possible.

Recurrence plot is a 2-dimensional plot expressed by the matrix:

$$R_{i,j} = \Theta(\varepsilon - \|x_i - x_j\|) \quad i, j = 1, 2, 3, \dots, N \quad (1)$$

where N is the number of measured points, $x_i, x_j \in R^d$ represent the i -th and j -th points of the d -dimensional state space trajectory, $\| \cdot \|$ represent the norm, ε is a threshold distance and Θ is the Heaviside function. The RP is obtained by plotting the recurrence matrix, Eq. (1). If $R_{i,j} \equiv 1$, it is considered as a recurrence point and appears as a black dot, if $R_{i,j} \equiv 0$, it forms a white dot [18]. March et al. [17] showed that RP can be constructed without embedding. Thus, it was thought desirable to choose the delay time of 1 based on the Takens theorem [20]. Therefore, in the present work, the RP of time series of pressure fluctuations was constructed without embedding.

2.2. Determining Parameters for RQA

The graphical representation of RPs may be complicated to evaluate, since they are considered as qualitative tools to detect hidden rhythms graphically. The quantification analysis of the recurrence plots involves estimation of some parameters (recurrence parameters) that describe the structures in the plots such as single dots and diagonal, vertical and horizontal lines. The structures within a recurrence plot are related to the different dynamics of the system [18]. Recurrence rate and entropy are two of RQA variables that were used in this work.

Recurrence rate (RR) expresses the density of repeated states throughout the trajectory and is mathematically defined as:

$$RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j} \quad (2)$$

where $\sum R_{i,j}$ is the total number of repeated points. Usually, RR is used to determine the value of radius threshold. The radius threshold should not be selected so large that makes the value of RR be greater than 20 % [18].

Entropy (ENT) refers to the Shannon information entropy of all diagonal line lengths distributed over integer bins in a histogram. Individual histogram bin probabilities ($p(l) = P(l) / N_l$) are computed for each non-zero bin and then summed according to Shannon's equation.

$$Entropy = - \sum_{l_{min}}^N p(l) \log_2 p(l) \quad (3)$$

Entropy is related to complexity of the system. For example, entropy would be expected to be 0.0 bits/bin for a periodic systems in which all diagonal lines are of equal length, but relatively high within chaotic systems [18, 21].

3 EXPERIMENTS

Experiments were carried out in a gas-solid fluidized bed made of a Plexiglas-pipe of 15 cm inner diameter and 2 m height. Air at ambient temperature entered the column through perforated plate distributor with 435 holes of 7 mm triangle pitch. A cyclone was used to separate air from particles at high superficial gas velocities. Sand particles (Geldart B) with

mean size of 150 and a particle density of 2640 kg/m^3 were used in the experiments. The bed was operated with different loaded sand heights (L/D of 1, 1.5 and 2) and at gas velocities ranging from 0.1 to 1.1 m/s.

Pressure probe (model SEN-3248 (B075), Kobold Company) was screwed onto the gluing studs located 10 cm above the distributor. Pressure fluctuations were recorded in approximately 164 s corresponding to 65535 data. The measured signals were band-pass (hardware) filtered at lower cut-off frequency of 0.1 Hz to remove the bias value of the pressure fluctuations and upper cut-off Nyquist frequency (200 Hz). The sampling frequency was 400 Hz. This sampling frequency is also in according with criterion of 50 to 100 times of the average cycle frequency (typically between 100 to 600 Hz) which is required for nonlinear evaluation of the pressure fluctuations in bubbling fluidized bed [13, 14].

4 RESULTS AND DISCUSSION

In fluidized beds, the main frequency of pressure fluctuations is normally below 10 Hz with a maximum at about 2.5-3 Hz. A difference between the average cycle frequency, f_c (the number of times per time unit the signal crosses its average) and the dominant frequency of the spectrum, f_d , indicates deviations from a perfect periodicity of the macro scale, since f_d is related to macro-scale structures [14]. Figure 1 indicates the average cycle frequency of the pressure fluctuations at aspect ratios of 1, 1.5 and 2. As can be seen, at lower aspect ratio, f_c has the higher deviation from the dominant frequency (2.5-3 Hz) which indicates that the finer structures have significance. In addition, as it can be observed, average cycle frequency of all three different aspect ratios initially decreases and approaches to the peak dominant frequency of the bed and then increases with an increase in velocity.

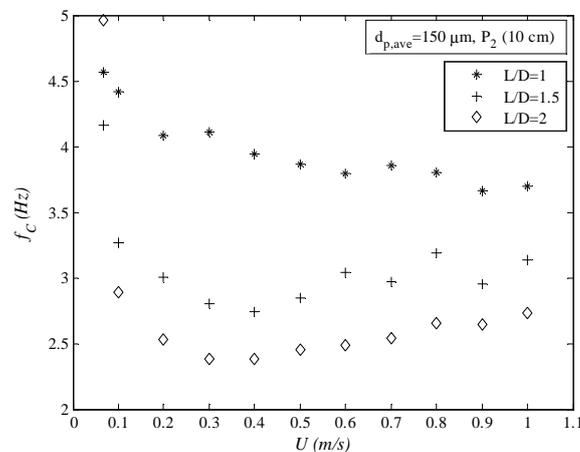


Figure 1: Average cycle frequency (f_c) of the pressure fluctuations at different aspect ratios

Figures 2a-c show the RP constructed from the pressure time series at aspect ratios of 1, 1.5, and 2, respectively. The repeated structures shown in these figures can be categorized in four groups of short diagonal lines, small bold areas, white bands (strips) repeated approximately regular and stretched vertically or horizontally, and quasi-square shapes made of horizontal and vertical lines with white area within them. Each of these typical patterns is

linked to a specific behavior of the system [18]. In this paper, only characteristics related to diagonal lines are considered.

The presence of single spots and the large diagonal lines with invariable distance between them is the obvious property of the stochastic and periodic recurrence plot, respectively. Short diagonal lines with irregular distance between them are one of the indications of chaotic dynamics [18, 22]. However, single points are rarely found in the RP of the fluidized bed. In other hands, short diagonal lines in Figures 2a-c show complexity of fluidized bed and indicate that the bed's hydrodynamic behavior is predictable only for short times. This qualitative pattern of RP is further quantified by RQA in terms of % RR and ENT .

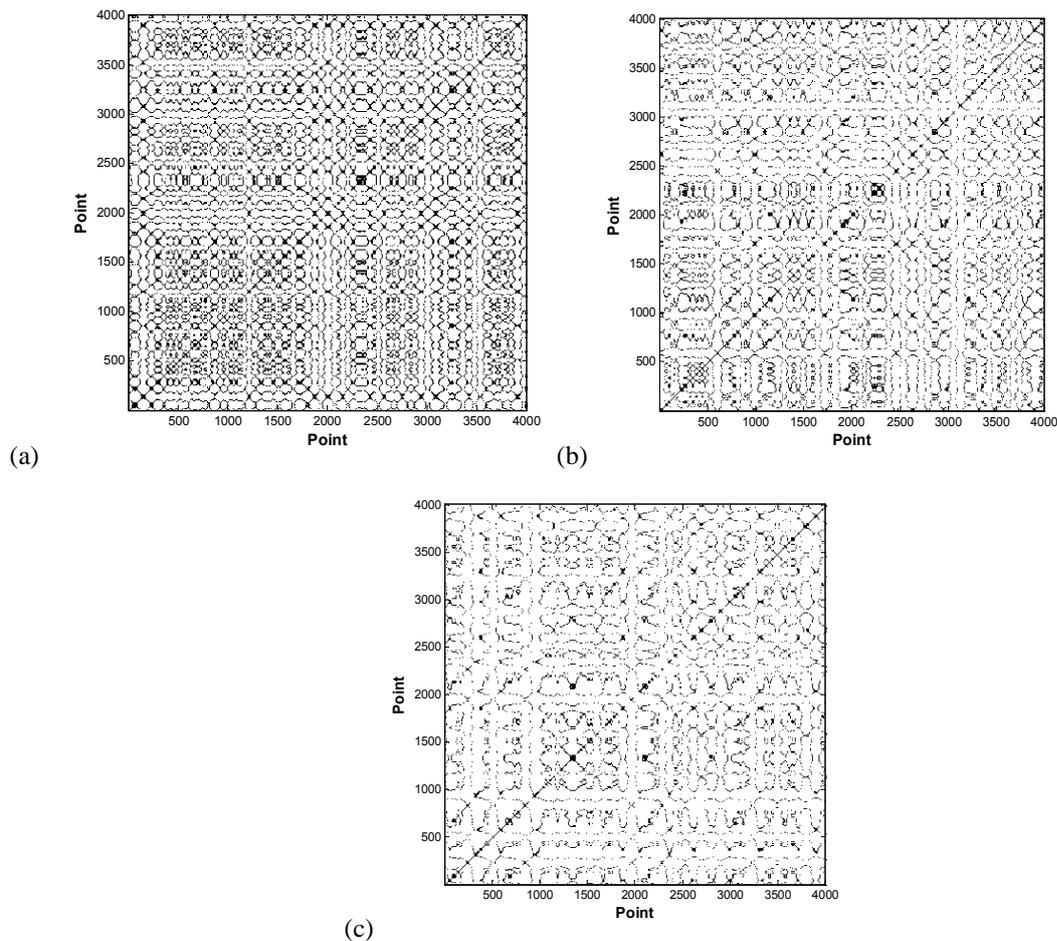


Figure 2: Recurrence plot of the fluidized bed at aspect ratios of (a) 1 (b) 1.5, and (c) 2; $U=0.5$ m/s; particles size $150\ \mu\text{m}$; $N=4000$; $\varepsilon=0.1$

RQA variables are usually calculated in consecutive epochs and dynamics of the system is inspected through them. Figure 3 shows the plots of the values for each of these quantified parameters (% RR and ENT) in consecutive epochs. As shown in this figure, the values of RR are smaller than 20 % for all aspect ratios. This shows that the value of radius threshold has been chosen reasonably. By comparing of RR values of the fluidized bed with the values of the stochastic and periodic systems, it was found that the dynamic behavior of fluidized bed is

between long-term predictable and complete unpredictable systems. In addition, RR values increased with an increase in aspect ratio. Therefore recurrence rate verifies that the bed with higher aspect ratio shows a more periodic behavior which had been shown by the average cycle frequency results. At this condition, effect of macro phenomena on the pressure fluctuations of the bed are dominant against meso and micro phenomena and the pressure signal approaches periodic behavior. It is expected that a more periodic system has the lower entropy which can be confirmed by the plot of entropy values. As can be seen in this plot, the entropy is smaller in higher aspect ratios. This shows that contribution of the larger structures becomes more important in higher aspect ratios and cause to lower complexity.

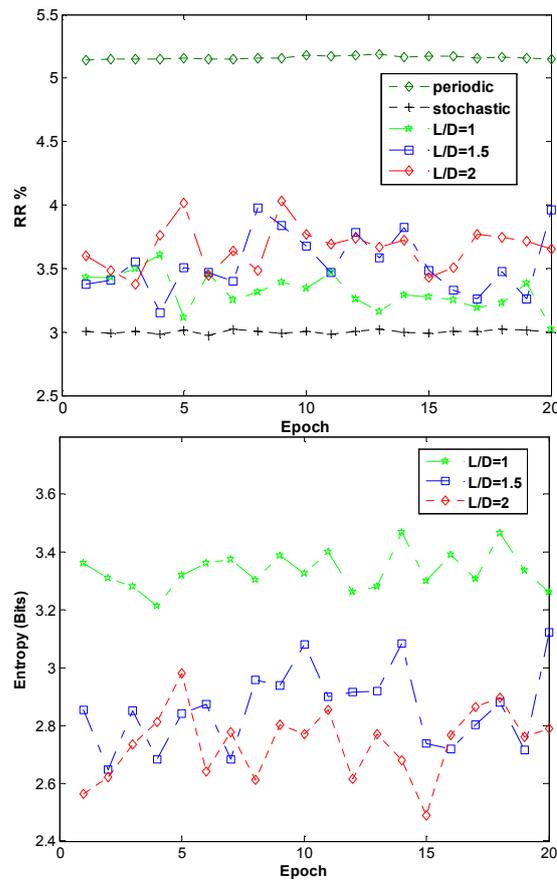


Figure 3: RQA variables (Recurrence Rate and Entropy) of the fluidized bed pressure signal at $U=0.5$ m/s and different aspect ratios; Epoch length=2000; $N=40000$; $\varepsilon=0.05$; $l_{min}=2$.

Figure 4 shows the entropy of the pressure fluctuations measured 10 cm above distributor as a function of gas velocity for different aspect ratios and particles size $150 \mu\text{m}$. As shown in this figure, the entropy for all three aspect ratios initially decreases and then increases with an increase in gas velocity. Comparing Figures 1 and 4 reveals that the trends of average cycle frequency and the entropy against gas velocity are approximately similar. It can be concluded that when there is a minimum deviation from periodicity of the bed, entropy are minimum.

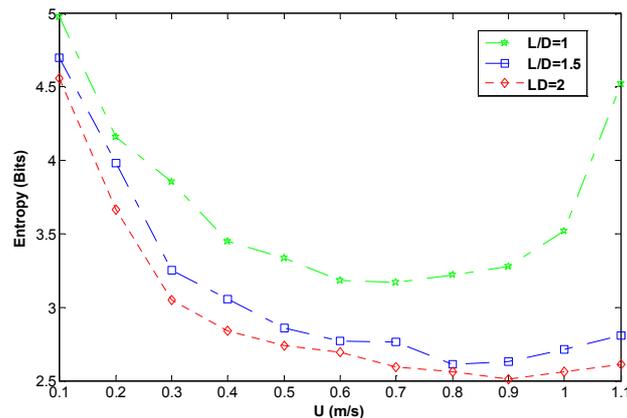


Figure 4: Entropy of the pressure fluctuations at different L/Ds of 1, 1.5 and 2, Epoch length=2000; $N=40000$; $\varepsilon=0.05$; $l_{min}=2$.

5 CONCLUSIONS

Recurrence plot (RP) and recurrence quantification analysis (RQA) were used to study of scaling aspect ratio effect on the dynamic features of the gas-solid fluidized bed. The presence of short diagonal lines in the RP showed that the fluidized bed is predictable only for short times. The higher aspect ratio provides greater amplitude due to larger bubbles. This trend is confirmed by the average cycle frequency results too, since at higher aspect ratio, f_c has the lower deviation from the dominant frequency which indicates that the larger structures have importance. At higher aspect ratio, effect of macro phenomena on the pressure of the bed are dominant against meso and micro phenomena and the pressure signal approaches periodic behavior. Recurrence rate and entropy verify this result and show that the bed with higher aspect ratio has a more periodic behavior. The results of the present work showed that the RQA is a powerful and easy method that its variables can be used for scaling, monitoring, study of hydrodynamic behavior within the fluidized bed system.

6 REFERENCES

- [1] Schouten J.C., Vander Stappen M.L.M., Van Den Bleek C.M. Scale-up of chaotic fluidized bed hydrodynamics. *Chemical Engineering Science* (1996) 51:1991.
- [2] Vander Stappen M.L.M. Chaotic hydrodynamics of fluidized beds. *PhD Thesis*, Delft University of Technology, (1996) ISBN 90-407-1375-8.
- [3] Knowlton T.M., Karri S.B.R., Issangya A. Scale-up of fluidized-bed hydrodynamics. *Powder Technology* (2005) 150:72.
- [4] Sasic S., Leckner B., Johnsson F. Characterization of fluid dynamics of fluidized beds by analysis of pressure fluctuations. *Progress in Energy and Combustion Science*. (2007) 33:453.
- [5] Glicksman L.R., Hyre M.R., Farrell P.A. Dynamic similarity in fluidization. *Int. J. Multiphase Flow* (1994) 20:331.
- [6] Daw, C.S., Halow, J.S. Characterization of voidage and pressure signals from fluidized bed using deterministic chaos theory. In: Anthony, E.J. (Ed.), *Proc. Eleventh Int. Conf. Fluidized bed combustion*. (1991) 778.

- [7] Schouten, J.C., Van den Bleek, C.M. Monitoring the quality of the fluidization using the short-term predictability of pressure fluctuations. *AIChE J.* (1998) 44:48.
- [8] Van der Stappen, M.L.M, Schouten, J.C., Van den Bleek, C.M. Application of deterministic chaos theory in understanding the fluid dynamic behavior of gas-solid fluidization. *AIChE Symp. Series* (1993) 89:91.
- [9] Zarghami R., Mostoufi N., Sotudeh-Gharebagh R. Nonlinear characterization of pressure fluctuations in fluidized beds. *Industrial and Engineering Chemistry Research* (2008) 47: 9497.
- [10] Fan L.T., Kang Y., Neogi D., Yashima M. Fractal analysis of fluidized particle behavior in liquid-solid fluidized beds. *AIChE J.* (1993) 39:513.
- [11] Skrzyzce D.P, Nguyen K., Daw C.S. Characterization of the fluidization behavior of different solid types based on chaotic time-series analysis of pressure signals. In: Rubow L, Commonwealth G (Eds.), *Proc. Twelfth Int. Conf. Fluidized Bed Combustion, ASME Book No. I0344B, New York.* (1993) 155.
- [12] Hay J.M., Nelson B.H., Briens C.L., Bergougnou M.A. The calculation of the characteristics of a chaotic attractor in a gas-solid fluidized bed. *Chem Eng Sci.* (1995) 50:373.
- [13] Franca F., Acikgoz M., Lahey R.T., Clause A. The use of fractal techniques for flow regime identification. *Int J Multiphase Flow.* (1991) 17:545.
- [14] Johnsson F., Zijerveld R.C., Schouten J.C., Van den Bleek C.M., Leckner B. Characterization of fluidization regimes by time-series analysis of pressure fluctuations. *Int J Multiphase Flow* (2000) 26:663.
- [15] Schouten J.C., Takens F., Van den Bleek C.M. Estimation of the dimension of a noisy attractor. *Phys Rev E.* (1994) 50:1851.
- [16] Eckmann J.P., Kamphorst S.O., Ruelle D. Recurrence plots of dynamical systems. *Europhys Lett.* (1987) 4:973.
- [17] March T.K., Chapman S.C., Dendy R.O. Recurrence plot statistics and the effect of embedding. *Physica D.* (2005) 200:171.
- [18] Marwan N., Carmen Romano M., Thiel M., Kurths J. Recurrence plots for the analysis of complex systems. *Physics Reports.* (2007) 438:237.
- [19] Thiel M., Carmen M.R., Kurths J. How much information is contained in a recurrence plot?. *Physics Letters A.* (2004) 330:343.
- [20] Takens F. Detecting strange attractors in turbulence. In *Proceedings of Dynamical Systems and Turbulence, Lecture Notes in Mathematics* 898; Rand DA, Yong LS. Eds.; Springer Verlag, Berlin. (1981) 366.
- [21] Gandhi A.B., Joshi J.B., Kulkarni A.A., Jayaraman V.K., Kulkarni B.D., SVR-based prediction of point gas hold-up for bubble column reactor through recurrence quantification analysis of LDA time-series, *International Journal of Multiphase Flow* (2008) 34:1099.
- [22] Webber C.L., Zbilut J.P. Recurrence quantification analysis of nonlinear dynamical systems. In: *Tutorials in contemporary nonlinear methods for the behavioral sciences.* (Chapter 2), Riley MA, Van Orden G, Retrieved December 1, (2004) 26.

FABRICATION AND MULTISCALE MECHANICAL CHARACTERIZATION OF Ti ALLOY/AL₂O₃ FUNCTIONALLY GRADED MATERIALS FOR ORTHOPAEDIC APPLICATIONS

DAVIDE CARNELLI^{*,†}, EMANUELE BERTARELLI^{*,†}, DARIO GASTALDI^{*},
TOMASO VILLA^{*,†}, FRANCESCO CASARI[§], ALBERTO MOLINARI[§] AND
PASQUALE VENA^{*,†}

^{*} Laboratory of Biological Structure Mechanics
Department of Structural Engineering, Politecnico di Milano
Piazza Leonardo da Vinci 32, 20133 Milano, Italy
e-mail: vena@stru.polimi.it

[†] IRCCS Istituto Ortopedico Galeazzi
Via Riccardo Galeazzi 4, 20161 Milano, Italy

[§] Department of Materials Engineering and Industrial Technologies
University of Trento, Via Mesiano 77, 38123 Trento, Italy

Key words: Functionally graded materials (FGM), Multiscale characterization, Grid indentation, Nanoindentation.

Abstract. In this study, the mechanical and tribological properties of Ti6Al4V-alloy/Al₂O₃ Functionally Graded Materials (FGM) were evaluated through multiscale indentation tests and metallographic analysis. The FGM, consisting of 11 uniform layers graded in 10% composition steps of 1 mm thickness each, was generated by Spark Plasma Sintering (SPS) technique with a non-uniform sintering temperature. Metallographic analyses evidenced the presence of a Ti₃Al intermetallic phase at the metal/ceramic interface of the intermediate FGM layers. Nano and micro-scale indentation tests were performed with a NanoTest Indenter (Micro Materials Ltd., Wrexham, U.K.), while an ad hoc set-up was designed and realized to characterize the FGM at the macroscopic scale. Elastic and inelastic material properties have been obtained on both the cross-section and the surfaces of the external pure metal and pure ceramic layers. The mechanical characterization exhibited the capability to quantify the problems related to the sintering techniques already evidenced in the metallographic analyses. The lack of an optimal sintering temperature led to sintering defects that most likely resulted in imperfect interfaces. This work proved that the combination of advanced micro and nano mechanical characterization techniques in a multiscale investigation framework represents a reliable feedback for the development and optimization of the SPS as a process for the fabrication of FGM.

1 INTRODUCTION

Functionally Graded Materials (FGM), i.e. materials with spatial gradients in composition, undoubtedly represent a promising class of functional materials with the potential of application in diverse contexts such as tribology, fracture mechanics and nanotechnology. A special interest can be identified in the development of structural materials for orthopaedic applications, where different properties such as reliability, biocompatibility and wear resistance are required. This can be achieved exploiting an FGM architecture.

As an example, in joint replacement the commonly used Co–Cr against UHMWPE bearings show a very good performance in terms of structural reliability, while third body wear phenomenon may lead to periprosthetic inflammation and implant loosening [1]. Fully ceramic implants (mainly Alumina) have been used too. The advantage of ceramic surfaces is the drastic reduction of the bearing wear rate. On the other hand, the low fracture toughness of ceramic materials and the susceptibility to failure by slow crack growth still remain a crucial issues [2]. In this direction, efforts have been made for exploiting the FGM principle for the enhancement of the structural reliability [3,4]. Besides, efforts are being made for exploiting the FGM principle for the enhancement of the tribological behaviour, with the realization of graded ceramic coatings on a metal substrate [5].

In this work, a multiscale approach based on the instrumented indentation technique has been adopted to characterize the mechanical and tribological behaviour of different material systems all obtained with the Spark Plasma Sintering (SPS) process: homogeneous Ti6Al4V, homogeneous AL₂O₃, composite Ti6Al4V-alloy/Al₂O₃ with different phase concentrations, and Ti6Al4V-alloy/Al₂O₃ Functionally Graded Materials (FGM). The study of the mechanical behaviour of the materials has been coupled with Light Optical Microscopy (LOM) imaging, Scanning Electron Microscopy (SEM), chemical (Energy Dispersive Spectrometry, EDS) and metallographic (X-Ray Diffraction) characterizations. The aim of the work is to show how the use of advanced multiscale indentation techniques represents a reliable feedback for the development and optimization of the SPS as process for the fabrication of Ti6Al4V-alloy/Al₂O₃ Functionally Graded Materials.

2 MATERIALS AND METHODS

2.1 Samples deposition

Disk samples were realized by means of the SPS technique [6]. In case of homogeneous Ti6Al4V and AL₂O₃ samples, sintering temperature was set to 1200°C and 1450°C, respectively; 60% Ti6Al4V-40% Al₂O₃ and 80% Ti6Al4V-20% Al₂O₃ composites were sinterized at 1350°C; finally, a non-uniform sintering temperature varying between 1450°C (Al₂O₃ side) and 1200°C (Ti side) was adopted for the FGM, consisting in 11 uniform layers graded in 10% composition steps of 1 mm thickness each.

2.2 Imaging, metallographic and chemical analyses

LOM and SEM imaging techniques were employed to preliminary evaluate the material microstructure. In particular, the presence of a third phase at the metal/ceramic interface of the intermediate FGM layers and composites was evidenced in the 60% Ti6Al4V-40% Al₂O₃

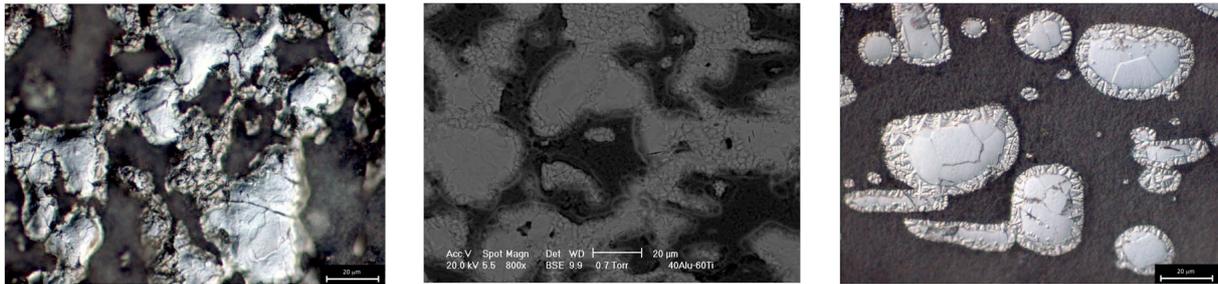


Figure 1: LOM (left) and SEM (middle) of the 60% Ti6Al4V-40% Al₂O₃ composites sinterized with SPS at 1350°C/10min/30MPa; bright is the Ti6Al4V, dark is Al₂O₃; LOM (right) of the 80% Ti6Al4V-20% Al₂O₃ composites. The third phase at the metal/ceramic interface is clearly visible.

and 80% Ti6Al4V-20% Al₂O₃ composites (Figure 1).

EDS chemical analyses carried out at the interface region between metal and ceramic enabled to identify Ti₃Al as the intermetallic phase. As an example, Figure 2 reports the EDS spectrum of the interface region of Ti/Al₂O₃ composite. This result is further evidenced with X-Ray Diffraction measurements of the composite phases. The intermetallic phase Ti₃Al is formed in spite of the pure Ti6Al4V and Al₂O₃ phases (Figure 3).

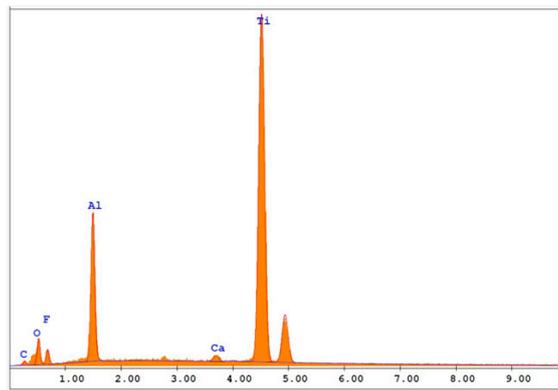


Figure 2: EDS spectrum of the interface region of Ti/Al₂O₃ composite.

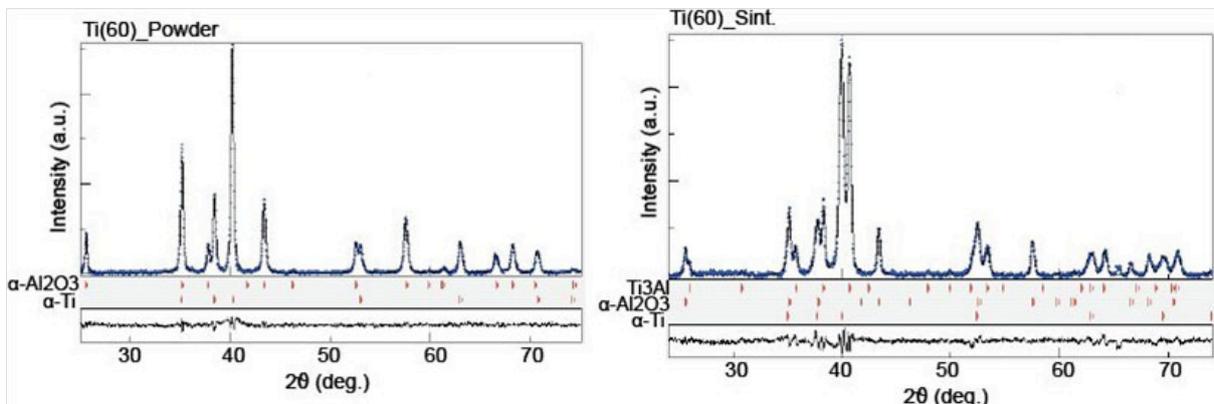


Figure 3: XRD spectrum of the interface region of Ti/Al₂O₃ composite.

2.3 Instrumented indentation

Instrumented indentation tests have been carried out at three different scales. Nanoscale (load range 10-70 mN) and microscale (load range 3-15 N) indentation tests were performed with a NanoTest (Micro Materials Ltd., Wrexham, U.K.) indenter (Figure 4a). An ad-hoc experimental set-up was designed and realized to characterize the samples at the macroscopic scale adopting a load range from 200 to 300 N (Figure 4b).

Homogeneous Ti6Al4V, homogeneous AL₂O₃, and Ti6Al4V-alloy/Al₂O₃ composites, and Ti6Al4V-alloy/Al₂O₃ FGM have been tested at the nano and micro scales. In particular, the composite samples were subjected to tests adopting the grid indentation technique [7] in order to identify the different phases of the composite as well as their mechanical properties (Figure 4a). Further, in case of FGM samples, elastic and inelastic material properties have been obtained on both the cross-section and the surfaces of the external pure metal and pure ceramic layers at the macro scale (Figure 5).

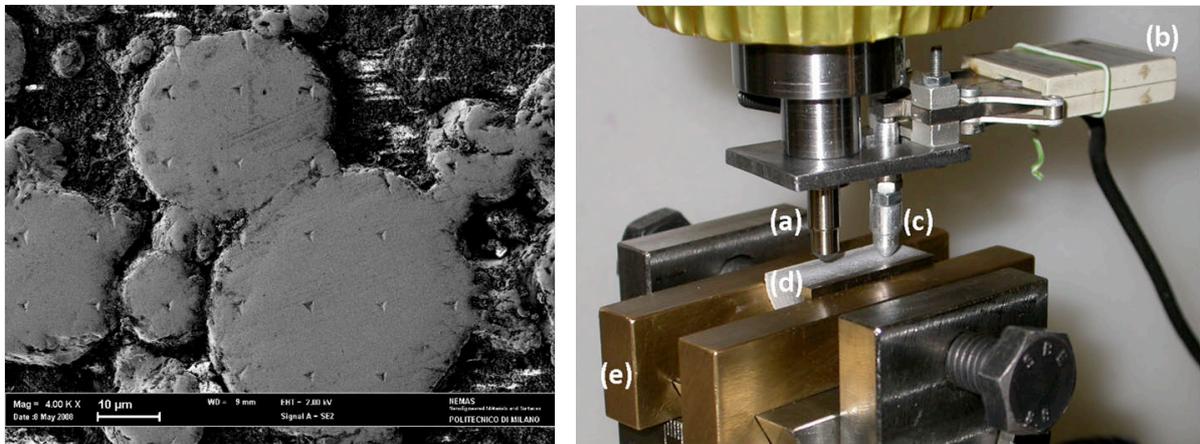


Figure 4: SEM image showing nanoindentation residual imprints of the grid indentation technique applied on the 60% Ti6Al4V-40% Al₂O₃ composite (left). Set-up of the instrumented macroscale indentation tests (right): (a) indenter tip; (b) extensometer; (c) extensometer comparator used to detect the tip penetration into the sample; (d) FGM sample; (e) sample holding system.

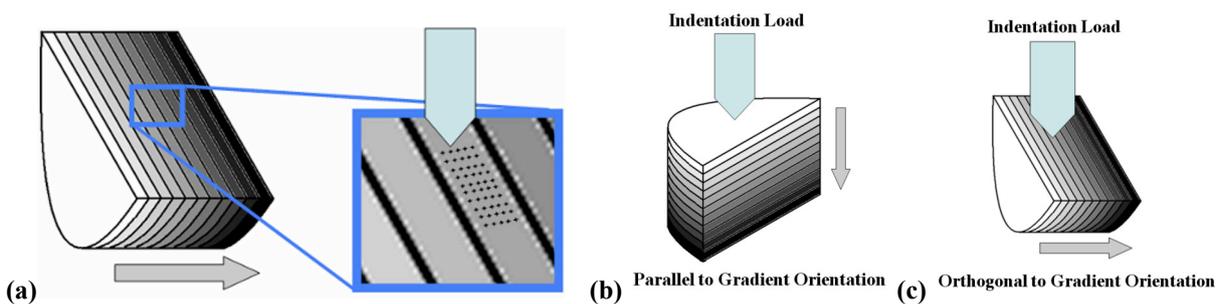


Figure 5: (a) Sketch of the nanoindentation tests performed on a single layer of the FGM structure. Testing configurations for macro scale indentation tests performed on the external layers surface (along the compositional gradient orientation direction) (b) and on the sample cross-section (orthogonal to the compositional gradient orientation direction) (c).

2.4 Coating performances assessment

In this work, the indentation loading-unloading curves are analyzed by means of the method proposed by Oliver and Pharr [8]. Following this procedure, the reduced modulus E_r and hardness H are evaluated as:

$$E_r = \frac{S}{2\beta} \sqrt{\frac{\pi}{A}} \quad (1)$$

$$H = \frac{P}{A} \quad (2)$$

where S is the initial unloading contact stiffness, A is the projected contact area, P is the maximum load applied during the test, and β a correction factor dependent on the indenter geometry ($\beta = 1.034$ for the Berkovich tip here adopted).

The ratio between the hardness and the reduced modulus (H/E_r) is expected to be a good reference parameter to evaluate the tribological behaviour of materials [9]: indeed, the H/E_r ratio has been proved to provide a closer agreement to the coatings performances in terms of wear than the hardness value only.

Finally, the elastic recovery parameter ERP , determined as a ratio of the maximal depth to the plastic depth of penetration during indentation, has been also calculated.

3 RESULTS

The comparison between the homogeneous materials (Ti6Al4V and AL2O3) and the Ti6Al4V-alloy/Al2O3 composites showed a higher dispersion of the mechanical properties of the composite with respect to the homogeneous materials. The reduced modulus and hardness values obtained from the nanoindentation tests on the homogeneous Ti6Al4V, homogeneous AL2O3, and 60% Ti6Al4V-40% Al2O3 composite samples are reported in Figure 6.

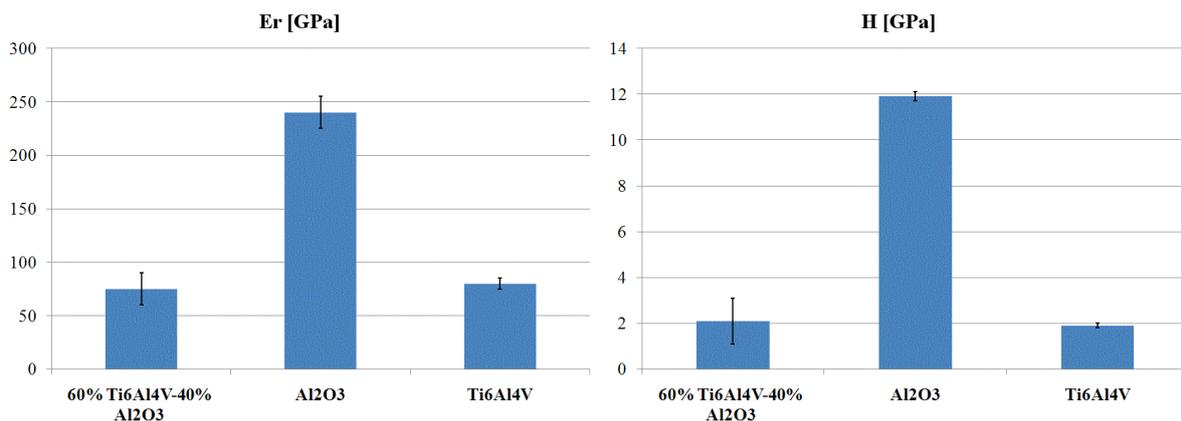


Figure 6: Reduced modulus and hardness values obtained from the nanoindentation tests on homogeneous Ti6Al4V, homogeneous AL2O3 and 60% Ti6Al4V-40% Al2O3 composite.

The mechanical properties of the composite samples are substantially close to the one of the pure Ti samples, the contribution of the Al phase being negligible. In fact, reduced modulus E_r and hardness H are both lower than the expected values estimated with homogenization techniques. These two results are in agreement with the presence of a third phase created during the SPS process. The grid indentation technique performed at the nanoscale confirms this evidence too (data not shown).

Macroindentation tests on the outermost layers, i.e. Al₂O₃ and Ti₆Al₄V, of the FGM sample evidenced an influence of the indentation direction with respect to the compositional gradient. Figure 7 shows the hardness and reduced modulus values for the outermost Ti₆Al₄V and AL₂O₃ layers of the FGM structure. The reduced modulus and hardness obtained along the layer cross-section are remarkably higher than normal to the surface in case of the AL₂O₃ phase, while lower differences were found for the Ti₆Al₄V phase.

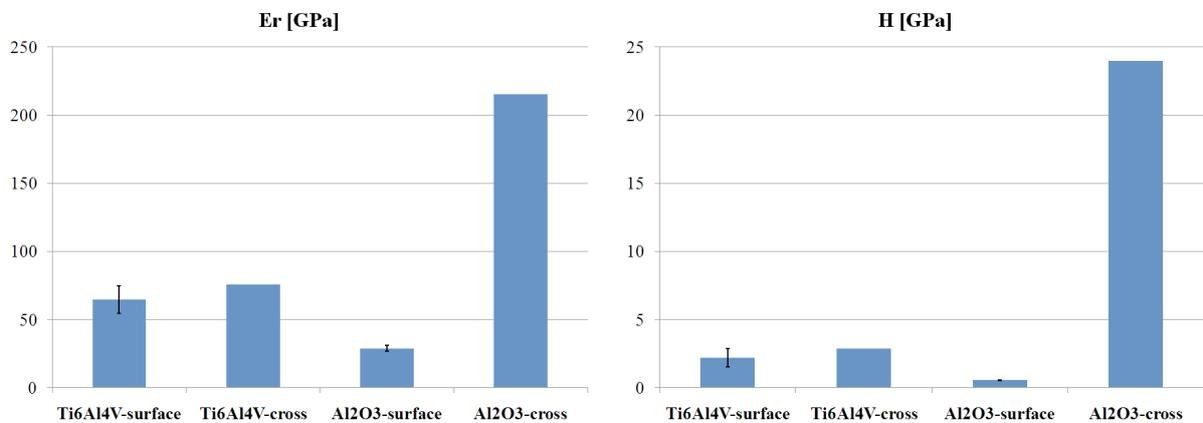


Figure 7: Reduced modulus and hardness values for the outermost Ti₆Al₄V and AL₂O₃ layers of the FGM structure tested along and normal to the compositional gradient.

Regarding the indentations on the cross section of the FGM sample, although the compositional gradient is linear between the pure metal and ceramic phases, a linear behaviour of the mechanical properties was not found (Figure 8). Figure 8 shows the comparison between reduced modulus, hardness, H/Er ratio, and ERP between the nano and macro indentations on each layer of the FGM cross section.

A reduction of the mechanical properties in the central layers, where lower reduced moduli and hardness than the expected values estimated with homogenization techniques are found, is evident from the measures both at the nano and macro scale.

Further, even if the overall trends at the nano and macro scales are comparable, absolute values are affected by the different mechanical properties evidenced at the two scales of measure. This is especially true for the Al-rich layers of the FGM system where the H , H/E_r and ERP values at the macroscale are in some cases more than twice the ones at the nanoscale, thus suggesting an evident size effect in the coating mechanical properties. Whereas, while lower differences have been found in terms of elastic behaviour for the Al-rich layers, the Ti-rich layers show size-dependent mechanical properties especially in the elastic range.

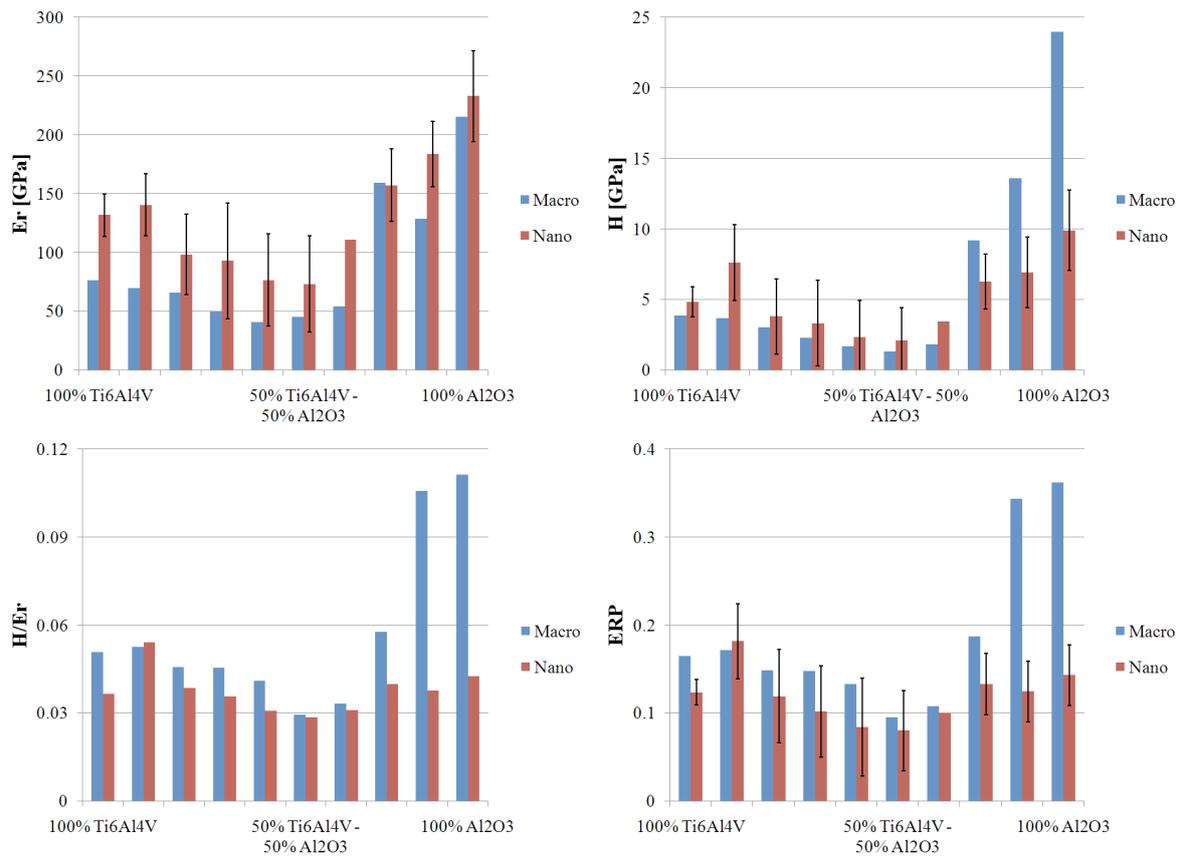


Figure 8: Comparison between reduced modulus (upper-left), hardness (upper-right), H/E_r ratio (lower-left), and ERP (lower-right) between the nano and macro indentations performed along the FGM cross section.

Finally, delamination and cracking of some FGM samples occurred during the mechanical testing. As an example, Figure 9 shows two different cases: on the left side, complete delamination is seen both at the 30% Ti6Al4V layer level and then between the 70% Ti6Al4V and 60% Ti6Al4V layers; instead, on the right side, a trans-layer crack crosses the sample between the 70% Ti6Al4V and 90% Ti6Al4V layers.

4 DISCUSSIONS AND CONCLUSION

The mechanical characterization had the capability to quantify the problems related to the sintering techniques already evidenced by the metallographic analyses.

In general, the mechanical testing results showed reliable mechanical properties only at the outermost layers, where the SPS process allowed obtaining materials with the mechanical properties expected for purely Al and Ti alloy; whereas, a significant reduction of the mechanical properties in the central layers was evidenced. The reason for this decay of mechanical properties in the innermost layers is attributable to the sintering temperature field applied, which is the result of a compromise between the limit temperatures of each material.

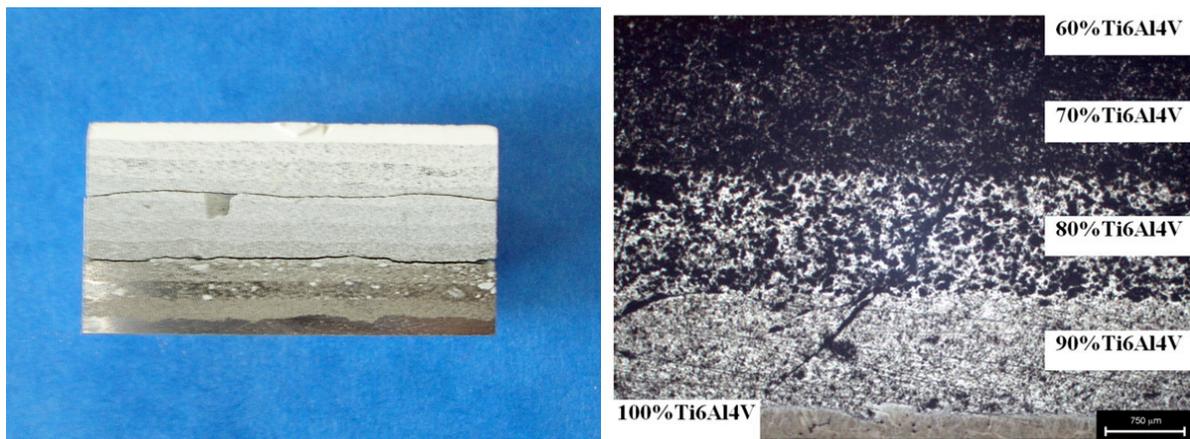


Figure 9: Delamination of an FGM sample (left) along the layer boundaries and a crack propagating perpendicular to the graded layers (right).

The lack of an optimal sintering temperature led to sintering defects that most likely resulted in imperfect interfaces. This hypothesis is confirmed both by the metallographic characterization and by the hardness and reduced moduli found for the intermediate FGM layers, which are lower than the expected values, estimated with homogenization techniques. Comparison between nano and macro scale tests evidenced the presence of a size effect in the FGM mechanical properties which affects the mechanical performances of both the Al-rich and Ti-rich layers. However, further investigations are required to shed light on this point.

This work proved that the use of advanced multiscale indentation techniques represent a reliable feedback for the development and optimization of the SPS as process for the fabrication of metal/ceramic FGM. These material solutions represent a forthcoming development for component of orthopaedic devices.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the “Cassa di Risparmio di Trento e Rovereto” for financially supporting the research project.

REFERENCES

- [1] Raimondi, M., Vena, P. and Pietrabissa, R., Qualitative evaluation of the prosthetic head damage induced by microscopic third-body particles in total hip replacement. *J. Biomed. Mater. Res.* (2001) **58**:436-448.
- [2] Rahaman, M.N., Yao, A., Bal, B.S., Garino, J.P. and Ries, M.D., Ceramics for prosthetic hip and knee joint replacement. *J. Am. Ceram. Soc.* (2007) **90**:1965-1988.
- [3] Sglavo, V.M. and Bertoldi, M., Design and production of ceramic laminates with high mechanical resistance and reliability. *Acta Mater.* (2006) **54**: 4929-4937.
- [4] Vena, P., Bertarelli, E., Gastaldi, D. and Contro, R., Energy-based and local approaches to the strength analysis of ceramic laminates with thermal residual stresses through the finite element method. *Mech. Res. Comm.* (2008) **35**:576-582.
- [5] Bertarelli, E., Carnelli, D., Gastaldi, D., Tonini, D., Di Fonzo, F., Beghi, M., Contro, R.

- and Vena, P., Nanomechanical testing of Alumina-Titanium functionally graded thin coatings for orthopedic applications. *Surf. Coat. Tech.* (2011) **205**:2838-2845.
- [6] Tokita, M., Development of Square-shaped Large-size WC/Co/Ni system FGM fabricated by Spark Plasma Sintering (SPS) method and its industrial applications. *Mat. Sci. Forum* (2005) **492-493**:711-718.
- [7] Constantinides, G., Ravi Chandran, K.S., Ulm, F.-J. and Van Vliet, K.J., Grid indentation analysis of composite microstructure and mechanics: Principles and validation. *Mat. Sci. Eng. A* (2006) **430**:189-202.
- [8] Oliver, W.C. and Pharr G.M., An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. *J. Mater. Res.* (1992) **7**:1564-1583.
- [9] Leyland, A. And Matthews A., On the significance of the H/E ratio in wear control: a nanocomposite coating approach to optimized tribological behaviour. *Wear* (2000) **246**:1-11.

MAGNETIZATION MODELS FOR PARTICLE-BASED SIMULATIONS OF MAGNETORHEOLOGICAL FLUIDS

HANNA G. LAGGER*, JOËL PEGUIRON*, CLAAS BIERWISCH* AND
MICHAEL MOSELER*

* Fraunhofer Institute for Mechanics of Materials IWM
Wöhlerstrasse 11, 79108 Freiburg, Germany
e-mail: hanna.lagger@iwm.fraunhofer.de, www.iwm.fraunhofer.de

Key words: Magnetorheological Fluid, Magnetization Model, Discrete Element Method, Numerical Simulation

Abstract. In this study, three-dimensional particle-based simulations are used to model magnetorheological fluids. The numerical model of the MRF is implemented in the framework of the Discrete Element Method (DEM) and takes into account the coupling of the magnetic dipoles, the hydrodynamic drag forces and steric forces between particles. To accurately treat the magnetic interaction between particles, the magnetic field at the particles' position is computed and an appropriate magnetization model is implemented. DEM simulations with different volume fractions of the MRF are carried out and the resulting magnetization curves are put in comparison with experimental data.

1 INTRODUCTION

Typical magnetorheological fluids (MRF) consist of magnetically permeable particles (e.g. iron) in carrier oil. Upon activation of an external magnetic field, the apparent viscosity of the MRF changes within a few milliseconds by orders of magnitude, inducing a change of the MRF from liquid to solid. By controlling the strength of the magnetic field, the viscosity of the MRF can be adjusted very accurately. For this reason, MRF are highly interesting for several industrial applications, such as controllable dampers or automotive clutches.

In this work we use numerical simulations to investigate the mechanisms which govern the behaviour of the MRF when subjected to a magnetic field. Inside the MRF, the contribution of the magnetized particles to the local magnetic field cannot be neglected [1]. Therefore an accurate description of the particles' magnetization is required. We can experimentally assess magnetization curves of magnetorheological fluids, but not of single particles. Different magnetization functions to model the response of an iron particle to a magnetic field can be found in the literature [2, 3, 4], however there is no well-established

magnetization function which is suitable for every MRF. A specific magnetization function is an arbitrary choice and can also be different for different materials [5]. Here we choose to perform simulations based on different single-particle magnetization functions, compare the resulting MRF magnetization curves to experiments [4, 6] and adapt the curve parameters.

In Sec. 2 the simulation method is presented. The calculation of the magnetization of a particle is explained in Sec. 3. Different magnetization models from the literature are discussed. In Sec. 4 we derive an expression to evaluate the magnetization of the MRF in the simulation. In Sec. 5 we show first simulation results and bring them in comparison with experimental results from the literature.

2 DISCRETE ELEMENT MODEL

The numerical description of the MRF is based on the Discrete Element Method (DEM), originally proposed by Cundall and Strack to compute the motion of a large number of particles [7]. After the computation of the forces \mathbf{F}_i on each particle, Newton's equation of motion

$$m_i^G \ddot{\mathbf{r}}_i = \mathbf{F}_i \quad (1)$$

is solved numerically for an ensemble of particles $i = 1, \dots, N$ with masses m_i^G and center of mass positions \mathbf{r}_i . We model the MRF clutch with a three-dimensional simulation box containing spherical particles, terminated by solid walls in one direction and with periodic boundary conditions in the other two directions.

The forces included in the model are magnetic interaction forces between the particles, elastic repulsion, and the Stokes' drag of the fluid on the particles. The wall is modeled as a dense and flat ensemble of non-magnetic particles. Gravity and Brownian forces are neglected.

2.1 Magnetic forces

The dipole-dipole interaction energy $E_{i,j}$ of two magnetic particles i and j with magnetic moments \mathbf{m}_i and \mathbf{m}_j is [8]

$$E_{i,j} = \frac{\mu_0}{4\pi} \left[\frac{\mathbf{m}_i \cdot \mathbf{m}_j}{r_{i,j}^3} - \frac{3}{r_{i,j}^5} (\mathbf{m}_i \cdot \mathbf{r}_{i,j}) (\mathbf{m}_j \cdot \mathbf{r}_{i,j}) \right], \quad (2)$$

where $\mathbf{r}_{i,j} = \mathbf{r}_i - \mathbf{r}_j$ is the vector connecting the centers of particle i and j , $r_{i,j} = |\mathbf{r}_{i,j}|$ its norm and $\mu_0 = 4\pi \cdot 10^{-7} \frac{\text{Tm}}{\text{A}}$ the vacuum permeability.

The force on dipole i caused by dipole j is given by

$$\begin{aligned} \mathbf{F}_{i,j} &= -\nabla_i E_{i,j} \\ &= \frac{3\mu_0}{4\pi} \left[\frac{(\mathbf{m}_i \cdot \mathbf{m}_j) \mathbf{r}_{i,j} + (\mathbf{m}_j \cdot \mathbf{r}_{i,j}) \mathbf{m}_i + (\mathbf{m}_i \cdot \mathbf{r}_{i,j}) \mathbf{m}_j}{r_{i,j}^5} - 5 \frac{(\mathbf{m}_i \cdot \mathbf{r}_{i,j}) (\mathbf{m}_j \cdot \mathbf{r}_{i,j}) \mathbf{r}_{i,j}}{r_{i,j}^7} \right]. \end{aligned} \quad (3)$$

The total magnetic force on particle i is then the sum of the contributions of all other particles j

$$\mathbf{F}_i = \sum_{\substack{j \\ j \neq i}} \mathbf{F}_{i,j}. \quad (4)$$

The magnetic moments of the particles depend on the local magnetic field. The dependence is specified by a magnetization function which is described in detail in Sec. 3.2.

2.2 Hydrodynamic forces

The fluid acts on the particles via Stokes' drag,

$$\mathbf{F}_i^{\text{Stokes}} = 6\pi\eta R_i (\mathbf{v}_{\text{fluid}} - \mathbf{v}_i), \quad (5)$$

where \mathbf{v}_i is the velocity of particle i , R_i is the particle radius, $\mathbf{v}_{\text{fluid}}$ is the velocity of the carrier oil and η its viscosity.

2.3 Contact forces

A repulsive contact force which prevents the particles from overlapping is included. The normal force of particle j on particle i is given by the Hertzian repulsion

$$\mathbf{F}_i^n = \left(\frac{\frac{2}{3}Y}{1 - \nu^2} \sqrt{\frac{R_i R_j}{R_i + R_j} \xi_{i,j}^{3/2}} \right) \frac{\mathbf{r}_{i,j}}{r_{i,j}}, \quad (6)$$

where Y is Young's modulus, ν is Poissons's ratio, and $\xi_{i,j} = \max\{R_i + R_j - |\mathbf{r}_{i,j}|, 0\}$ the overlap [9].

As a first approach, no tangential forces are assumed.

3 TREATMENT OF THE MAGNETIC INTERACTIONS

In this section we introduce a selfconsistent algorithm for the calculation of the local magnetic field as well as general features of the single-particle magnetization function (Sec. 3.1). Three specific magnetization models are then discussed in Sec. 3.2.

3.1 Selfconsistent calculation of local magnetic induction and particle magnetization

The local magnetic induction at the position of particle i ,

$$\mathbf{B}_{i,\text{ext}} = \mathbf{B}_{\text{applied}} + \sum_{\substack{j \\ j \neq i}} \mathbf{B}_{i,j}, \quad (7)$$

is given by the sum of the externally applied magnetic field $\mathbf{B}_{\text{applied}} = \mu_0 \mathbf{H}_{\text{applied}}$ and the sum of the contributions from other particles

$$\mathbf{B}_{i,j} = \frac{\mu_0}{4\pi} \left[3 \frac{(\mathbf{m}_j \cdot \mathbf{r}_{i,j}) \mathbf{r}_{i,j}}{r_{i,j}^5} - \frac{\mathbf{m}_j}{r_{i,j}^3} \right]. \quad (8)$$

The magnetization \mathbf{M}_i of particle i is $\mathbf{M}_i = \frac{\mathbf{m}_i}{V_i}$ where V_i is the volume of particle i . The magnetic field $\mathbf{H}_{i,\text{in}}$ at the interior of particle i is given by [3]

$$\mathbf{H}_{i,\text{in}} = \frac{1}{\mu_0} \mathbf{B}_{i,\text{ext}} + \alpha \mathbf{M}_i. \quad (9)$$

The parameter α thus describes the contribution of the particle's own magnetization to its inner field $H_{i,\text{in}}$.

For the magnetization of the particle, we consider a model of the form

$$\mathbf{M}_i = |\mathbf{M}_S| f(b|\mathbf{H}_{i,\text{in}}|) \frac{\mathbf{B}_{i,\text{ext}}}{B_{i,\text{ext}}}, \quad (10)$$

where \mathbf{M}_S is the saturation magnetization of the particle and f is a scaled magnetization function which is linear for small $|\mathbf{H}_{i,\text{in}}|$ and tends to 1 for large $|\mathbf{H}_{i,\text{in}}|$. The parameter b is chosen such that the slope for small fields matches the low-field susceptibility χ of the material.

As \mathbf{H}_{in} itself depends on \mathbf{M} , Eq. (10) is an implicit equation. We solve this equation numerically by a root-finding algorithm that combines Newton-Raphson with the bisection method for a fail-safe routine [10]. We approach the solution of the coupled equations (7), (8) and (10) by the following iterative algorithm.

1. Initial step: $\mathbf{M}_i = 0$, $\mathbf{B}_{i,\text{ext}} = \mathbf{B}_{\text{applied}}$
2. Solve Eq. (10) with root-finding algorithm, update \mathbf{M}_i for all i
3. Update $\mathbf{B}_{i,\text{ext}}$ for all i with equations (7) and (8)
4. Repeat steps 2. and 3. until the convergence criterion

$$\max_i |\mathbf{M}_i(n) - \mathbf{M}_i(n-1)| < \epsilon \quad (11)$$

is reached, where $\mathbf{M}_i(n)$ is the magnetization of particle i in the n -th iteration.

The positions of the particles are not changed during the selfconsistency loop.

3.2 Different magnetization models proposed in the literature

Magnetorheological fluids are mostly suspensions of carbonyl iron powder, a highly pure iron, where magnetic hysteresis is negligible [4]. Therefore we consider only models of the form of Eq. (10), describing anhysteretic magnetization behaviour. In this section three different choices for the magnetization function f as introduced in Eq. (10) are presented, the role of the model parameters is discussed, and the different models are compared.

3.2.1 Three magnetization functions

The three models we want to compare are the following:

1. A simple magnetization function for axially anisotropic materials with applied field along the easy axis, found in [2]

$$\mu_0 M = \mu_0 M_S \tanh\left(\frac{\chi}{M_S} H_{\text{in}}\right) \quad (12)$$

2. A phenomenological model for isotropic materials, based on the modified Langevin function [3]

$$\mu_0 M = \mu_0 M_S \left(\coth\left(\frac{3\chi}{M_S} H_{\text{in}}\right) - \frac{1}{\frac{3\chi}{M_S} H_{\text{in}}} \right) \quad (13)$$

3. The Fröhlich-Kennely law for the magnetization of a particle located inside an infinite chain [4]

$$\mu_0 M = \mu_0 M_S \frac{\frac{\chi}{M_S} H_{\text{in}}}{1 + \frac{\chi}{M_S} H_{\text{in}}} \quad (14)$$

3.2.2 How to choose the parameters

Saturation magnetization M_S is the saturation magnetization of the material of the spheres. This value can be obtained experimentally. Following [11], the saturation magnetization M_S of the suspension of particles and carrier oil can be related to the saturation magnetization of the bulk magnetic solid $M_{S,\text{bulk}}$ by the volume fraction ϕ of the suspension, $M_S = \phi M_{S,\text{bulk}}$.

Susceptibility and α All chosen models are constructed in a way that

$$M \sim \chi H_{\text{in}} \text{ for small } H_{\text{in}}, \quad (15)$$

where $\chi = \chi_{\text{mat}}$ is the susceptibility of the material. The susceptibility χ_{part} of a particle is given by

$$M \sim \chi_{\text{part}} H_{\text{ext}} \text{ for small } H_{\text{ext}}. \quad (16)$$

Combining equations (9), (15) and (16), we get

$$\chi_{\text{mat}} = \frac{\chi_{\text{part}}}{1 + \alpha \chi_{\text{part}}}. \quad (17)$$

Unfortunately, the single particle parameters are usually not known. In experiments, the magnetization curve of the whole MRF is measured instead [12]. Thus we have to infer

the susceptibility of the particle from the experimentally measured magnetization curve of the MRF.

Regarding the parameter α , one obtains for uniformly magnetized spheres $\alpha = -\frac{1}{3}$ and thus $\chi_{\text{sphere}} = \frac{3\chi_{\text{mat}}}{3+\chi_{\text{mat}}}$ [8]. In other cases, α is a coupling parameter depending on the coupling of the magnetic domains inside the particle and on the particle geometry and has to be fitted to experiments [3, 8].

3.2.3 Comparison of the magnetization models

Figure 1 shows the different magnetization models for parameter values $\mu_0 M_S = 1.709$ T (taken from [4]), $\alpha = -1/3$ (uniformly magnetized spheres) and a susceptibility $\chi_{\text{mat}} \approx 5.7$.

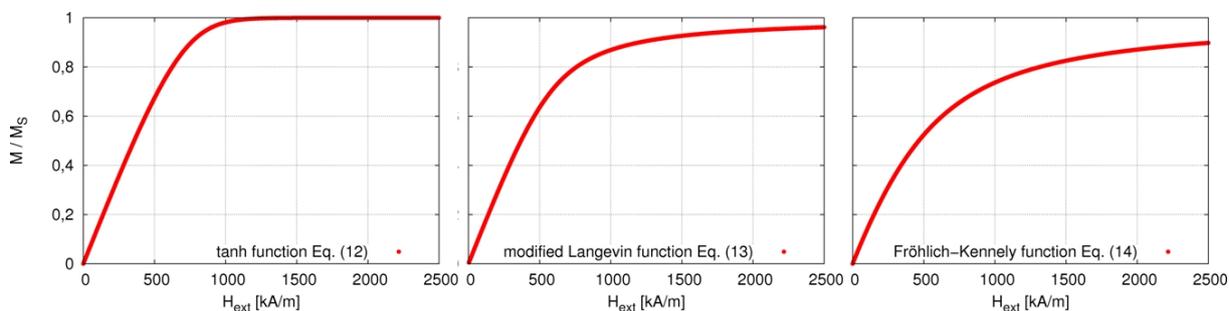


Figure 1: Comparison of the three magnetization functions Eq. (12), Eq. (13), and Eq. (14).

We see that the three models differ strongly in their behaviour for intermediate field strength. The tanh function is very steep and approaches saturation early. The modified Langevin function deviates from the linear regime for small fields earlier and needs stronger fields to reach saturation. The same holds for the Fröhlich-Kennely function, only that the effects are still more pronounced.

4 DEFINITION OF THE MRF MAGNETIZATION

Experimentally the magnetization of the MRF is obtained by

$$\mu_0 M = B_{\text{measured}} - \mu_0 H_{\text{applied}}, \quad (18)$$

where H_{applied} is the externally applied magnetic field, controlled by a electromagnet, and B_{measured} is the magnetic induction measured at a location close to the MRF.

If we want to mimic the experimental procedure in the simulation, we need to choose a point where to measure the magnetization. We can—similarly to the experiment—apply a certain external field H and measure the magnetic induction B at some point close to the MRF. Figure 2 shows the value of $B_{\text{measured}} - B_{\text{applied}}$ for a setup where the MRF is between two discs at $z = \pm 37.5 \mu\text{m}$. B_{measured} is the measured magnetic induction in the simulation (at $y = 0$, averaged in x-direction) as a function of the measuring position,

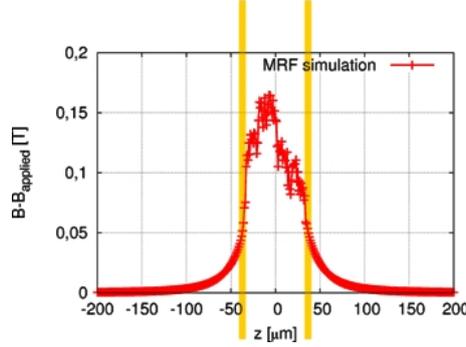


Figure 2: Locally measured magnetic induction (only the contribution of the MRF, without externally applied field). The MRF is located between the yellow bars.

called z_0 , in z -direction. We see that the measured value depends significantly on the distance of the measuring point from the MRF. However, we want to have a value for the magnetization of the MRF, which is independent of the measuring point. To this purpose, we analyze the contribution of the MRF to the magnetic field B at distances far from the MRF. The behaviour of the absolute value of the magnetic field for large distances from

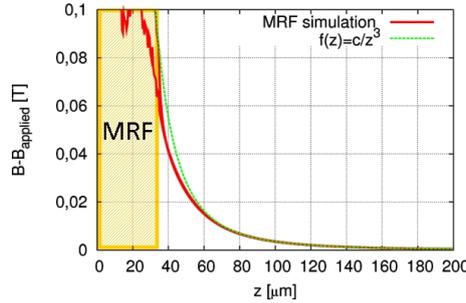


Figure 3: Magnetic field B induced by the MRF, far from the MRF.

the MRF can be described by

$$|B_{\text{measured}}(z) - B_{\text{applied}}| = \frac{c}{z^3} \quad (19)$$

with a constant c , as shown in Fig. 3. Thus the field of the MRF shows the same decay as a magnetic dipole field. We therefore define the magnetic moment of the MRF \mathbf{m}_{MRF} as the magnetic dipole moment inducing the same magnetic field B as the MRF at larger distances from the MRF.

A magnetic dipole located at the origin with dipole moment $\mathbf{m} = m\mathbf{e}_z$ creates a magnetic field

$$\mathbf{B}_{\text{dipole}}(z) = \frac{\mu_0}{2\pi} \frac{\mathbf{m}}{|z|^3} \quad (20)$$

at location $(0, 0, z)$. The value of $m_{\text{MRF}} = |\mathbf{m}_{\text{MRF}}|$ can be found by fitting the large-distance magnetic field B (19), yielding

$$m_{\text{MRF}} = c \frac{2\pi}{\mu_0}. \quad (21)$$

Another expression for the MRF magnetization can be obtained by considering that, seen from a larger distance from the MRF, the magnetic dipoles carried by the particles inside the MRF are approximately at the same spot. With the superposition principle we get for the field created by the MRF:

$$\mathbf{B}_{\text{MRF}}(z) = \sum_i \mathbf{B}_i(z) = \sum_i \frac{\mu_0}{2\pi} \frac{\mathbf{m}_i}{|z|^3} = \frac{\mu_0}{2\pi} \frac{1}{|z|^3} \sum_i \mathbf{m}_i, \quad (22)$$

where \mathbf{B}_i is the magnetic field created by particle i and \mathbf{m}_i is the magnetic moment of particle i . On the other hand

$$\mathbf{B}_{\text{MRF}}(z) = \frac{\mu_0}{2\pi} \frac{\mathbf{m}_{\text{MRF}}}{|z|^3}. \quad (23)$$

Combining the equations (22) and (23) we get

$$m_{\text{MRF}} = \left| \sum_i m_i \right|. \quad (24)$$

Thus, we expect the two definitions (21) and (24) for m_{MRF} to be equivalent. To test this hypothesis, we performed simulations with an externally applied field of $\mu_0 H = 0.14$ T, 0.42 T, and 1.4 T. The results are shown in Table 1.

Table 1: Comparison of the two definitions for the magnetic moment of the MRF.

$\mu_0 H_{\text{applied}}$ [T]	0.14	0.42	1.4
$m = c \frac{2\pi}{\mu_0}$ [Am^2]	2.40×10^{-8}	5.94×10^{-8}	6.92×10^{-8}
$m = \sum_i m_i$ [Am^2]	2.38×10^{-8}	5.90×10^{-8}	6.86×10^{-8}

The differences between the definitions amount to less than 1 % of the magnitude of the magnetic moment. Therefore the two definitions can be regarded as equivalent. For our simulations, we take $\mathbf{m}_{\text{MRF}} = \sum_i \mathbf{m}_i$ as the working definition. \mathbf{M}_{MRF} is then defined as

$$\mathbf{M}_{\text{MRF}} = \frac{\mathbf{m}_{\text{MRF}}}{V_{\text{MRF}}} = \frac{N \langle \mathbf{m}_{\text{part}} \rangle}{V_{\text{MRF}}} = \phi \langle \mathbf{M}_{\text{part}} \rangle. \quad (25)$$

The MRF magnetization for monodisperse suspensions can thus be expressed as the average particle magnetization $\langle \mathbf{M}_{\text{part}} \rangle = \frac{\langle \mathbf{m}_{\text{part}} \rangle}{V_{\text{part}}}$ times the volume fraction $\phi = \frac{N V_{\text{part}}}{V_{\text{MRF}}}$, where V_{part} is the volume of one particle, V_{MRF} the volume of the MRF, and N the number of particles in the MRF.

Equation (25) provides a convenient definition for the magnetization of the MRF in the simulation. The experimentally measured magnetization however still depends on the unknown measuring point. We can get independence from the measuring point by dividing the measured magnetization by the measured high-field saturation value at the same measuring point. This gives transformed experimental values

$$\frac{|B(z_0) - B_{\text{applied}}|}{|B_S(z_0) - B_{\text{applied}}|} = \frac{\frac{\mu_0}{2\pi} \frac{m_{\text{MRF}}^{\text{exp}}}{|z_0|^3}}{\frac{\mu_0}{2\pi} \frac{m_{\text{S,MRF}}^{\text{exp}}}{|z_0|^3}} = \frac{m_{\text{MRF}}^{\text{exp}}}{m_{\text{S,MRF}}^{\text{exp}}} \quad (26)$$

The same can be done for the computed magnetization

$$\frac{M_{\text{MRF}}}{M_{\text{S,MRF}}} = \frac{m_{\text{MRF}}}{m_{\text{S,MRF}}} \quad (27)$$

With this scaling procedure, the experimental [Eq. (26)] and computed [Eq. (27)] magnetization values can be put in comparison.

5 RESULTS

Here we present DEM-simulation results at different volume fractions with spherical particles of $1 \mu\text{m}$ in diameter, comparable to the experiments reported in [4]. We used the tanh model with values from Sec. 3.2.3. In the simulations we see that the magnetization curve of the MRF deviates significantly from the single sphere curve (Fig. 4). Figure 5

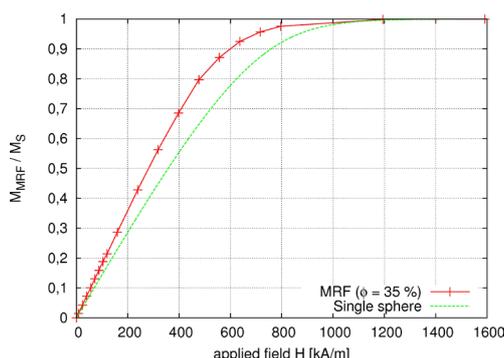


Figure 4: Scaled magnetization curve of the MRF ($\phi = 35 \%$) compared with the single sphere curve.

shows the unscaled MRF magnetization curves simulated at different volume fractions ϕ . In Ref. [4], the relative differential permeability curves of MRF are shown instead of the

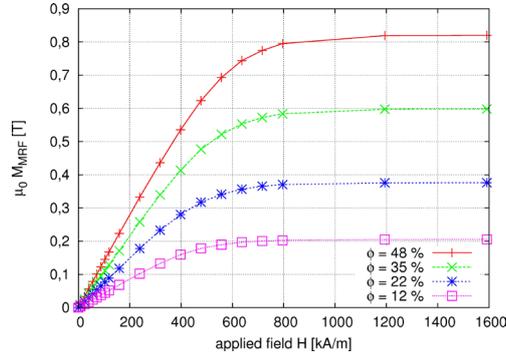


Figure 5: Magnetization curves of the MRF for different volume fractions ϕ .

magnetization. From simulations, we get the relative differential permeability $\mu_{r,dif}$ of the MRF as

$$\mu_{r,dif}(B) = 1 + \mu_0 \frac{dM_{MRF}}{dB} \quad (28)$$

The simulation results are shown in Fig. 6. As stated in Sec. 3.2.2, we can get experimental

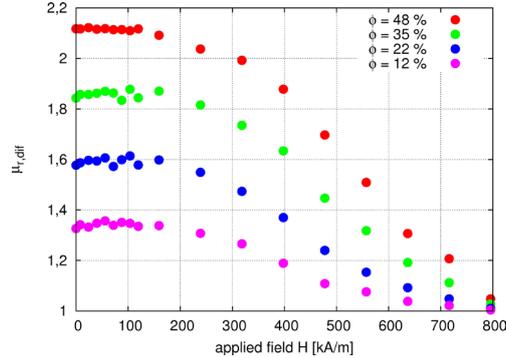


Figure 6: Relative differential permeability $\mu_{r,dif}$ for different volume fractions ϕ of the MRF from DEM-simulations.

values for the low-field susceptibility χ_{MRF} of the MRF, whereas the susceptibility of the material χ_{mat} is unknown. The susceptibility χ_{MRF} is related to the relative differential permeability by

$$\chi_{MRF} = \mu_{r,dif}(0) - 1. \quad (29)$$

The computed values are shown in Fig. 7. The linear dependence of χ_{MRF} on volume fraction is also observed in experiments [4]. The susceptibility there is much higher though. By increasing the susceptibility of the material χ_{mat} we also get higher susceptibilities for the MRF (see figure 8). However, the exact dependence still has to be investigated.

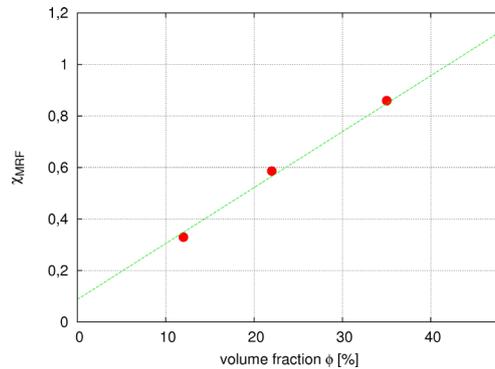


Figure 7: Low-field susceptibility of the MRF from DEM-simulations vs. volume fraction ϕ .

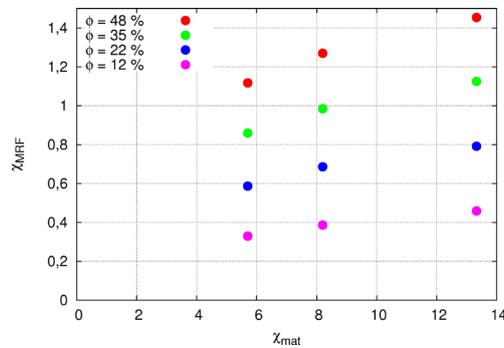


Figure 8: Low-field susceptibility of the MRF vs. material susceptibility χ_{mat} for different volume fractions ϕ .

6 CONCLUSIONS

A numerical model of a magnetorheological fluid based on the Discrete Element Method with emphasis on the magnetization model was presented. The magnetic interactions were treated using a selfconsistent algorithm. For the description of the anhysteretic magnetization curve of the particles, three magnetization models from the literature were compared. An expression for the magnetization of the MRF in the simulation was derived, thus enabling the comparison of experimental and computed magnetization curves.

From simulations we see that the magnetization curve of the MRF deviates significantly from the single sphere curve. The low-field susceptibility of the MRF depends linearly on volume fraction. The material's susceptibility, which is needed as a parameter for the magnetization model, can be obtained by comparing the resulting MRF susceptibility to experiments.

REFERENCES

- [1] Furst E M and Gast A P, Phys. Rev. E **61**, 6732 (2000).
- [2] Raghunathan A, Melikhov Y, Snyder J E, and Jiles D C, Appl. Phys. Lett. **95**, 172510 (2009).
- [3] Jiles D C and Atherton D L, J. Magn. Magn. Mater. **61**, 48 (1986).
- [4] de Vicente J, Bossis G, Lacia S, and Guyot M, J. Magn. Magn. Mater. **251**, 100 (2001).
- [5] Jiles D C, Thoelke J B, and Devine M K, IEEE Trans. Magn. **28**, 27 (1992).
- [6] Ehrlich J, Böse H, Private communication.
- [7] Cundall P A and Strack O D L, Geotechnique **29**, 47 (1979)
- [8] Jackson J D, *Classical Electrodynamics*. Wiley (1998).
- [9] Hertz H, J. reine angew. Math. **92**, 156 (1881).
- [10] Press W H, Teukolsky S A, Vetterling W T, and Flannery B P, *NUMERICAL RECIPES in Fortran 77*. Cambridge Univ. Press (1992).
- [11] Rosensweig R E, *Ferrohydrodynamics*. Dover Publications (1997).
- [12] Jolly M R, Carlson J D, and Muñoz B C, Smart Mater. Struct. **5**, 607 (1996).

A THREE-PHASE FINITE ELEMENT MODEL OF WATER-INFILTRATED POROUS MATERIALS SUBJECTED TO FREEZING

MENGMENG ZHOU* AND GÜNTHER MESCHKE†

*Institute for Structural Mechanics
Ruhr-University Bochum
Universitätsstr. 150, 44780 Bochum, Germany
e-mail: mengmeng.zhou@rub.de, www.sd.rub.de

†Institute for Structural Mechanics
Ruhr-University Bochum
Universitätsstr. 150, 44780 Bochum, Germany
e-mail: guenther.meschke@rub.de, www.sd.rub.de

Key words: Thermo-hydro-mechanical Coupling, Finite Element Method, Porous Materials, Soil Freezing

Abstract. In areas of geotechnical engineering, artificial ground freezing is commonly used as an effective way to deal with various ground construction challenges such as groundwater control and temporary excavation support. For the description of the coupled thermo-hydro-mechanical behavior of soil exposed to frost action, this paper presents a three-phase Finite Element model of porous materials, consisting of solid skeleton, liquid water and crystal ice, where the liquid phase contains both weakly-bound pore water and strongly-bound water film. Within the theory of thermo-poroelasticity proposed by Coussy [1, 2], poroelastic constitutive relations are provided from an energy approach of poromechanics. In addition, the phase transition between water and ice is characterized by a purely temperature-dependent thermodynamic state function named liquid saturation degree considering the pore size distribution. The cryo-suction mechanism that induces migration of water towards the frozen sites is impelled by the chemical potential difference existing between the pre-melted water film and the adjacent pore water. By choosing solid displacement, liquid pressure and mixture temperature as principal unknowns, the model is implemented in a geometrically-linear Finite Element context base upon the governing balance equations for the soil constituents and their mixture. The validation procedure is shown by selected examples with analyses of different aspects of the model behavior.

1 INTRODUCTION

Freezing of porous materials is of great interest to a wide range of fields such as geotechnical engineering, food industry, environmental engineering and biomechanics. In areas of geotechnical engineering, artificial ground freezing, which can be achieved by using either a large portable refrigeration plant or liquid nitrogen, is often used as an effective ground improvement technique to deal with various geotechnical construction challenges such as groundwater control and temporary excavation support. Yet, the potential problem of this technique is that it may produce frost heave and thaw settlement at the ground surface, which have deleterious consequences among surface infrastructure. Moreover, the closure of the frozen-soil body is significantly influenced by the groundwater flow especially when the seepage velocities are high. Consequently, for the sake of a safe design and execution, such applications require a correct prediction of the coupled thermo-hydro-mechanical behavior of porous materials upon freezing.

However, the mechanics of porous materials subjected to frost action is way more complex than simply that of a sealed water-filled bottle, since porous solids with pores of various sizes behave intriguing and even counterintuitive. A well-known interpretation is the unexpected expansion observed during the freezing of a sealed cement paste sample filled up with benzene, which contracts when solidifying [3]. During freezing, fully coupled thermo-hydro-mechanical processes are generally involved such as phase transition between constituents as well as the cryo-suction mechanism driving the liquid towards the frozen sites. Recent works have been devoted to the modeling of the behavior of porous materials exposed to freezing. A mathematical freezing model for saturated soil has been established by Mikkola & Hartikainen [4, 5, 6] based on the mixture theory and basic principles of thermomechanics. Similarly, a three-phase freezing model of saturated porous media (solid, pore fluid and pore ice) has been developed by Bluhm et al. [7, 8] within the framework of Theory of Porous Media (Boer [9] and Ehlers [10]). Nevertheless, both models provide little physical explanation for the observed phenomena such as the cryo-suction process that sucks liquid water towards the already frozen sites. Coussy has proposed a comprehensive thermo-poroelastic theory for the description of mechanical behavior of water-infiltrated materials upon freezing [1, 2, 11, 12], which specifies the multi-scale physics of confined crystallization of ice and provides more physics-based understanding by means of exploring how the macroscopic properties can be upscaled from the knowledge of microscopic properties. Thereby, as a result of careful evaluation, the theory of thermo-poroelasticity proposed by Coussy is adopted as the basis for present research.

Based upon the adopted theory, a three-phase Finite Element model of porous materials, consisting of solid skeleton, liquid water and crystal ice, is present in this paper to describe the behavior of water-infiltrated materials upon freezing. The setting up of the liquid-crystal thermodynamic equilibrium provides a relation between the crystal pressure and liquid pressure, which in the sequel plays an important role on the cryo-suction mech-

anism. The phase transition between water and ice is incorporated via a temperature-dependent liquid saturation function characterized by pore size distribution. Poroelastic constitutive relations are obtained from an energy approach of poromechanics. Stemming from the governing balance equations, the coupled set of differential equations for solid displacement, liquid pressure and mixture temperature is formulated in a geometrically-linear Finite Element context and implemented in an object-oriented Finite Element code (KRATOS). The validity of the model is demonstrated by three selected examples with analyses of the main characteristics of the behavior of freezing porous materials.

2 MODELING OF FREEZING POROUS MATERIALS

According to the adopted poromechanics theory [1, 2], the behavior of freezing porous materials can be essentially described at the macroscopic scale by the following set of equations: (1) liquid-crystal equilibrium relation, (2) liquid saturation curve, and (3) constitutive equations.

2.1 The liquid-crystal equilibrium relation

Thermodynamic equilibrium between the liquid pore water (L) and the adjacent crystal ice (C) requires the equality of the chemical potential of both phases

$$d\mu_L = d\mu_C, \quad \text{with} \quad d\mu_J = \frac{dp_J}{\rho_J} - s_J dT \quad (J=L, C), \quad (1)$$

where p_J , ρ_J and s_J are the pressure, the specific mass density and the entropy per unit mass of phase J, respectively. Integrating the above equation between the reference state ($p = 0 \text{ Pa}$, $T = T_f = 273 \text{ K}$) and the current state provides the liquid-crystal equilibrium condition

$$p_C - p_L = \Sigma_f (T_f - T), \quad \text{with} \quad \Sigma_f = \frac{\rho_C L_f}{T_f}, \quad (2)$$

where T_f , Σ_f and L_f are the freezing temperature, the freezing entropy and the latent heat of freezing, respectively.

2.2 The liquid saturation curve

The mechanical equilibrium of the liquid-crystal interface is governed by the YOUNG-LAPLACE law

$$p_C - p_L = \frac{2\gamma_{CL}}{R}, \quad (3)$$

where γ_{CL} is the liquid-crystal interface energy and R is the mean curvature radius of the interface. If a zero contact angle between the liquid and the solid skeleton is assumed for simplicity, R denotes also the pore radius where the current liquid-crystal interface locates. Combining Eqs. (2) and (3) yields the GIBBS-THOMSON relation

$$R = \frac{2\gamma_{\text{CL}}}{\Sigma_{\text{f}}(T_{\text{f}} - T)}, \quad (4)$$

which implies that all the pores having an entry radius greater than R will freeze at the given temperature T , and the remaining ones being filled with liquid water. This allows us to retrieve the existence of a state relation linking the liquid saturation and the temperature.

The determination of this relation can be inferred from the knowledge of the capillary curve, which relates the liquid saturation to the capillary pressure. For a porous material partially saturated by liquid water, and whose remaining porous space is filled by air (G), the capillary curve can be described by the function of VAN GENUCHTEN [13]

$$p_{\text{G}} - p_{\text{L}} = \mathcal{N}(S_{\text{L}}^{-\frac{1}{m}} - 1)^{1-m}, \quad 0 < m < 1, \quad (5)$$

where p_{G} , \mathcal{N} and m are the gas pressure, the capillary modulus and a constant representing the shape of the capillary curve, respectively. Similarly to Eq. (3), the YOUNG-LAPLACE law at the gas-liquid interface reads

$$p_{\text{G}} - p_{\text{L}} = \frac{2\gamma_{\text{GL}}}{R}, \quad (6)$$

where γ_{GL} is the liquid-air interface energy. Combination of Eqs. (5) and (6) provides a relationship between the liquid saturation and the pore size

$$S_{\text{L}} = \left(1 + \left(\frac{2\gamma_{\text{GL}}}{\mathcal{N}R} \right)^{\frac{1}{1-m}} \right)^{-m} = S(R), \quad (7)$$

which denotes that the remaining liquid saturation S_{L} equals to the cumulative fraction $S(R)$ of pore volume occupied by pores having a pore entry smaller than R , according to the standard mercury porosimetry [14, 12].

Replacing the pore size R in (7) by the GIBBS-THOMSON law (4) gives a relationship between the liquid saturation and the temperature

$$S_{\text{L}} = \left(1 + \left(\frac{\Delta T}{\Delta T_{\text{ch}}} \right)^{\frac{1}{1-m}} \right)^{-m}, \quad (8)$$

where $\Delta T = T_{\text{f}} - T$ is defined as the current cooling and $\Delta T_{\text{ch}} = \frac{\mathcal{N}\gamma_{\text{CL}}}{\Sigma_{\text{f}}\gamma_{\text{GL}}}$ the characteristic cooling. The parameters ΔT_{ch} and m are closely related to the pore size and pore size distribution (cf. Boukpeti [15]). Their influences on the shape of the saturation curve are illustrated in Fig. 1.

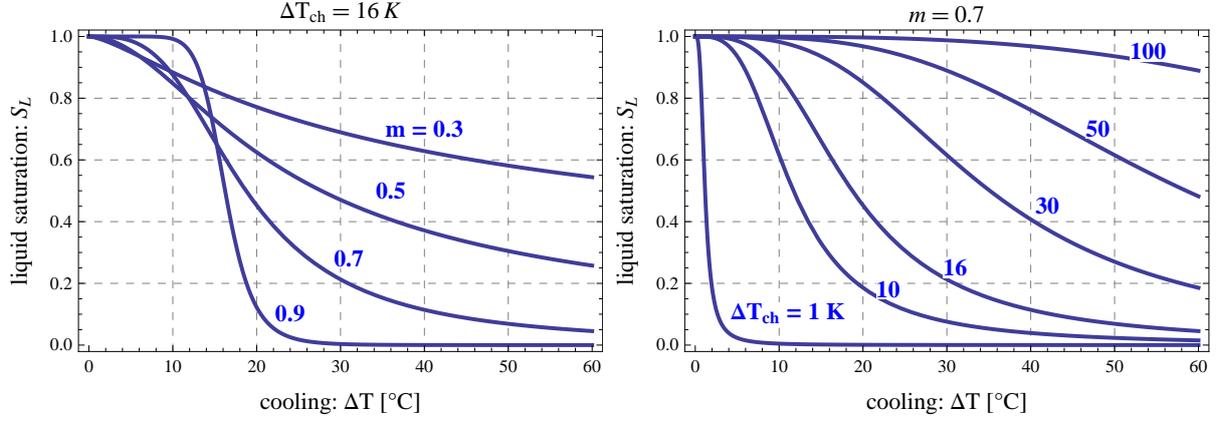


Figure 1: Liquid saturation curve during freezing by influences of m (left) and ΔT_{ch} (right).

2.3 Constitutive equations

To describe the mechanical behavior of freezing porous materials, the thermo-poroelastic constitutive equations, developed by Coussy [2] using an energy approach, are adopted. Consider a representative volume $d\Omega_0$ extracted from a partially frozen porous material with an initial porosity ϕ_0 , an initial temperature T_0 and a (zero) reference pore pressure of atmospheric pressure. The current (LAGRANGIAN) porosity ϕ can be written as

$$\phi = \phi_L + \phi_C, \quad \phi_J = \phi_0 S_J + \varphi_J, \quad (9)$$

where ϕ_J is current partial porosity occupied by phase J, φ_J the respective change in partial porosity, and S_J the respective current saturation with $S_L + S_C = 1$. Recalling the standard micro-macro relations [1] and assuming that all pores deform the same when subjected to the same pore pressure, the total stress tensor $\boldsymbol{\sigma}$ and the partial porosity change φ_J are related to the strain tensor $\boldsymbol{\varepsilon}$, the pore pressures p_J and the temperature T as following

$$\begin{aligned} \sigma_{ij} &= (K - \frac{2}{3}G)\epsilon \delta_{ij} + 2G\varepsilon_{ij} - b(S_L p_L + S_C p_C)\delta_{ij} - 3\alpha_S K(T - T_0)\delta_{ij}; \quad (10) \\ \varphi_L &= b\epsilon S_L + \frac{p_L}{N_{LL}} + \frac{p_C}{N_{CL}} - \alpha_S(b - \phi_0)(T - T_0)S_L; \\ \varphi_C &= b\epsilon S_C + \frac{p_L}{N_{CL}} + \frac{p_C}{N_{CC}} - \alpha_S(b - \phi_0)(T - T_0)S_C, \end{aligned}$$

where $\epsilon = \varepsilon_{kk}$ is the volumetric dilation; K and G are respectively the bulk modulus and the shear modulus of the solid skeleton subjected to condition $p_L = p_C = 0$; k_S and α_S are respectively the bulk modulus and the linear thermal dilation coefficient of the solid matrix, that is the solid part of the porous solid; while b and N_{JK} are respectively the BIOT coefficient related to the solid skeleton and the generalized BIOT coupling moduli satisfying the MAXWELL symmetry relation $N_{CL} = N_{LC}$ and $\frac{1}{N_{JJ}} + \frac{1}{N_{LC}} = \frac{b - \phi_0}{k_S} S_J$. Hence, the current porosity yields

$$\phi = \phi_0 + b \epsilon + \frac{b - \phi_0}{k_S} (S_L p_L + S_C p_C) - \alpha_S (b - \phi_0) (T - T_0). \quad (11)$$

It is worthwhile to note that if the approximation $\varphi_J \approx (\varphi_L + \varphi_C) S_J$ is made, the current partial porosities can be obtained via $\phi_J = \phi S_J$.

3 GOVERNING BALANCE EQUATIONS

With the above formulation of the constitutive relations, adopting linearized geometrical relations, the problem of a freezing water-infiltrated porous material exposed to freezing temperatures is described by the governing balance equations, which form the set of differential equations to be solved by the Finite Element Method (FEM). In the present work, the model development was restricted to isotropic thermo-poroelasticity under the assumption of small deformations. For the formulation of the balance equations it is assumed that the pore volume is at all time fully-saturated – partly occupied by ice crystals and partly occupied by the water remaining in liquid form. By choosing the displacement \mathbf{u} , the liquid water pressure p_L and the mixture temperature T as primary variables, the set of field equations was derived as summarized below.

3.1 Mass balance of liquid water and crystal ice

Taking into account the possible phase transition between liquid water and crystal ice, the mass balance equation relative to each phase can be written as

$$\frac{d m_L}{d t} + \nabla \cdot \mathbf{w}_L = -\overset{\circ}{m}_{L \rightarrow C}, \quad \frac{d m_C}{d t} + \nabla \cdot \mathbf{w}_C = \overset{\circ}{m}_{L \rightarrow C}, \quad (12)$$

where $m_J = \rho_J \phi_J$ stands for the current mass content related to phase J per unit of initial volume $d\Omega_0$, while $\overset{\circ}{m}_{L \rightarrow C}$ is the rate of liquid water mass changing into crystal ice per $d\Omega_0$. Normally, the ice flow is slow, so that $\mathbf{w}_C = \mathbf{0}$ can be assumed. Summation of these two equations yields the mass balance for the liquid water and the crystal ice

$$\frac{d m_L}{d t} + \frac{d m_C}{d t} + \nabla \cdot \mathbf{w}_L = 0. \quad (13)$$

In addition the liquid water flow \mathbf{w}_L is assumed to be governed by DARCY's law.

3.2 Overall momentum balance

Disregarding dynamic effects, the momentum balance equation for the mixture is given as:

$$\nabla \cdot \boldsymbol{\sigma} + \rho \mathbf{g} = 0, \quad (14)$$

where ρ stands for the overall mass density of the porous material and \mathbf{g} for the gravity forces per unit volume.

3.3 Overall entropy balance

Under the assumption of small perturbations, small variations of the temperature, the entropy balance for the mixture can be derived as:

$$T \left(\frac{d \Sigma_S}{dt} + m_L \frac{d s_L}{dt} + m_C \frac{d s_C}{dt} + (s_C - s_L) \frac{d m_C}{dt} \right) + \nabla \cdot \mathbf{q} = 0, \quad (15)$$

where $\Sigma = \Sigma_S + m_L s_L + m_C s_C$ is the overall density of entropy per unit of volume, while Σ_S is the entropy of the solid matrix, s_J the specific entropy related to liquid water or crystal ice, and \mathbf{q} the outgoing heat flow vector. Considering isotropy, FOURIER's law is given as $\mathbf{q} = -k \nabla T$, where k is the thermal conductivity.

4 FINITE ELEMENT FORMULATION

In the context of the Finite Element formulation of the model in a geometrically linear setting, the balance equations (13)-(15) are transformed to their variational forms. For the spatial discretization of the initial boundary value problem, quadratic Lagrangian shape functions are used for the approximation of the displacement field and linear shape functions are used for the approximations of the liquid pressure and the temperature. With such approximations the BABUSKA-BREZZI stability constraint is fulfilled [16]. For the temporal discretization, a modified midpoint rule, denoted as the generalized- α method is used, which ensures unconditional stability and second order accuracy for an appropriate choice of its parameters [17]. The discretized weak form, evaluated at the generalized midpoint, yields a highly nonlinear system of equations that is solved iteratively using NEWTON's method. For the sake of simplicity, the tangent stiffness matrices required to solve the linearized system of algebraic equations are generated numerically according to the methodology presented by Lee & Park [18]. Finally, the three-phase freezing model for saturated soils has been implemented into the object-oriented FE-code KRATOS (Dadvand et al.[19]).

5 MODEL VALIDATION

Once the numerical model has been established, a validation procedure has been initiated in order to demonstrate, that the proposed formulation can be used to reproduce the main characteristics of the behavior of freezing porous materials. The adopted validation strategy is characterized by three steps, in which different aspects of the model behavior are analyzed.

5.1 Phase transition with latent heat effect

The model performance with regards to the phase change behavior and the latent heat effect of freezing soils is first investigated by comparing the numerical results with the phase-change model presented by Lackner et al. [20], where only the thermal problem is

considered. A cuboid fully saturated sand specimen with a height of 0.09 m and a cross-section of 0.41 m by 0.41 m and an initial temperature $T_i = 1^\circ\text{C}$ is subjected to freezing initialized at the bottom surface with a heat flux $q^* = -100 \frac{\text{W}}{\text{m}^2}$, whereas all other surfaces are assumed as thermally isolated.

The obtained simulation results of the developed model is compared with the numerical results generated from the Lackner's model [20]. Both models indicate that, as soon as phase transition starts, the release of latent heat prevents the temperature from dropping. As long as the total released energy is consumed, a rapid temperature decrease is observed (Fig. 2 left). Moreover, during the freezing process, the freezing front propagates through the specimen from the bottom to the top until the entire specimen is frozen (Fig. 2 right). The comparison shows a good correlation of the numerical results for the evolution of both the temperature and the crystal saturation. Only a slight shape difference appears in the latter curve, due to use of different liquid saturation functions.

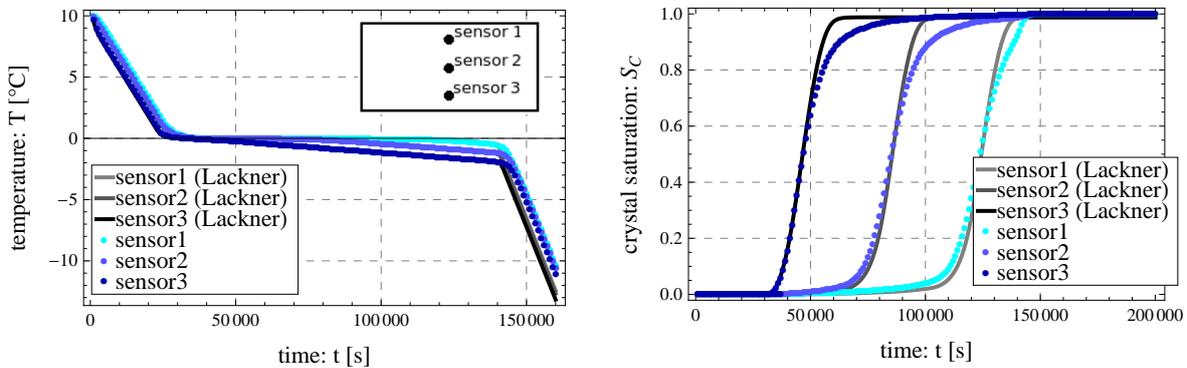


Figure 2: Model validation – phase transition with latent heat effect: Temperature (left) and ice saturation evolutions for three sensor positions

5.2 Terzaghi's consolidation problem

As a second validation example, a one dimensional consolidation test is reanalyzed numerically using the developed model. To this end, a soil layer resting on a rigid impervious base at depth $z = 1$ m, while its upper surface at $z = 0$ m remains drained with a reference pressure $p_L = 0$ Pa, is considered. For sake of simplification, gravity loading is ignored. A vertical constant load $t^* = -1000 \frac{\text{N}}{\text{m}^2}$ is instantaneously applied at the upper surface.

Initially, the applied load is fully carried by the liquid water leading to an overpressurization owing to its low compressibility. Subsequently this overpressure progressively vanishes as a result of the diffusion process of the fluid towards the boundary of the soil layer which remains drained. The obtained numerical results conform perfectly with the analytical solution [1], as shown in Fig. 3. The temporal evolution of the vertical displacement of the upper surface follows an exponential characteristics, with a value of

$u = -0.09258$ mm at $t = 1000$ s. This (nearly asymptotic) value is close to the analytical solution of the soil layer's drained settlement at $t = \infty$, $s_\infty = 0.09286$ mm, where the fluid overpressure has dissipated completely.

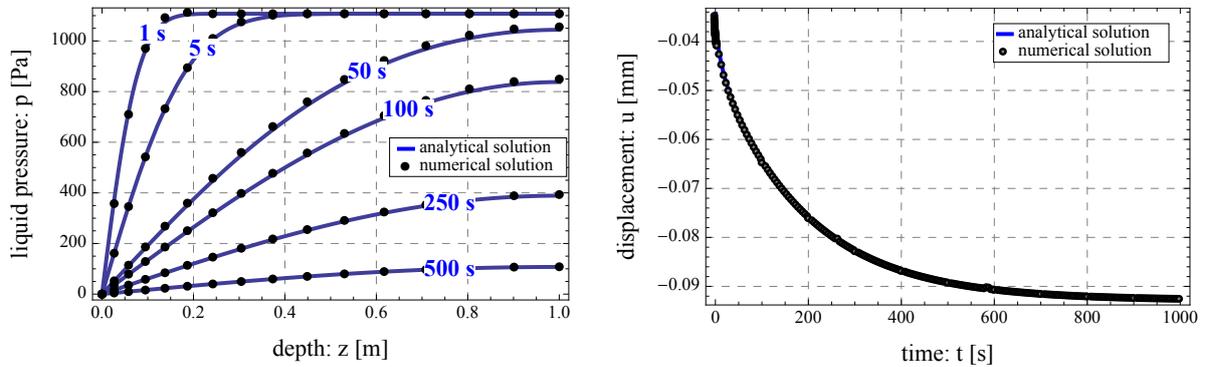


Figure 3: Model validation – consolidation test: Water pressure profiles over depth for different time instants (left); and vertical displacement evolution of upper surface (right)

5.3 Strain analysis during freezing

For the final validation of the model behavior related to the thermo-hydro-mechanical couplings, an experiment performed by Beaudoin & MacInnis [21] has been re-analyzed numerically. A cubic cement paste specimen (side length: 0.1 m), initially fully saturated by liquid water, is exposed to uniform cooling. The specimen is assumed as stress-free and undrained, i.e. no external load is applied and the total mass of water existing either in solid or liquid form remains constant.

In the absence of air voids, the porous material undergoes in general a cryo-swelling during freezing (Fig. 4). The strain analysis (Coussy [14]) shows that there are three contributions to the volumetric dilatation during freezing (Fig. 4 left). The main contribution $\varepsilon_{\Delta\rho}$ accounts for the hydraulic effect because the excess of liquid water expelled from the freezing sites, as a result of the liquid-crystal density difference, cannot escape from the sample. The other two contributions ε_{th} and ε_{Σ_f} are consequences of the thermal contraction of the mixture and the micro-cryo-suction process caused by the difference between the liquid pressure and crystal pressure, respectively. Actually, it has been concluded that ε_{Σ_f} is always positive even for liquids that usually contract when they solidify, and would vanish only for a zero freezing entropy [14]. This explains the dilation still observed in the experiment [3] using benzene as saturating liquid. In Fig. 4 (right) the numerical result for the overall volumetric strain during cooling is compared to analytical solution. Only small difference caused by the discretization error is observed.

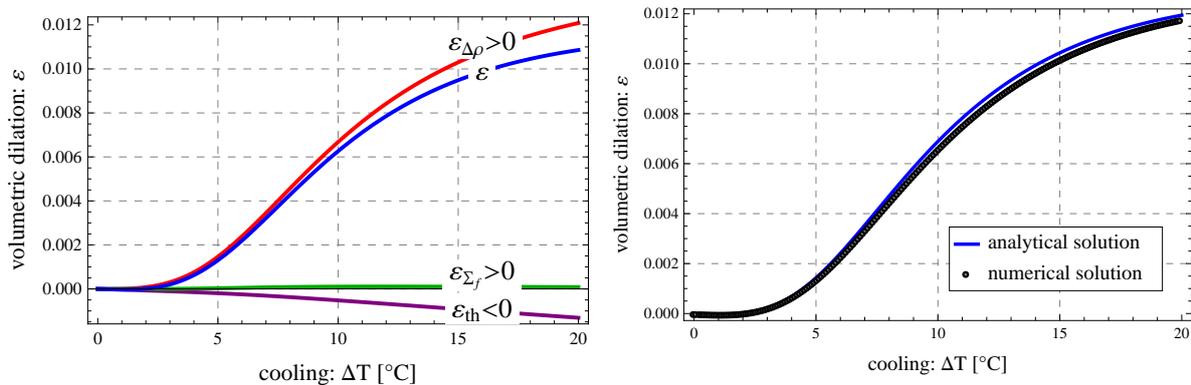


Figure 4: Model validation – thermo-hydro-mechanical couplings: Three contributions to the volumetric dilatation (left) and comparison of the volumetric dilatation as obtained from the model and from analytical solution (right)

6 CONCLUSIONS AND OUTLOOK

A Finite Element model formulated within the theory of thermo-poroelasticity proposed by Coussy [1, 2] has been presented in this paper to describe the fully-coupled thermo-hydro-mechanical behavior of water-saturated soil subjected to freezing. Within the developed model, soil is treated as a three-phase porous material composed of solid skeleton, liquid water and crystal ice, where the liquid phase contains both unconfined pore water and confined water film. Instead of being specified, the latent heat associated phase transition has been incorporated via the temperature-dependent liquid saturation curve characterized by the pore size distribution. The cryo-suction mechanism, identified as the driving force for frost heave phenomenon, is impelled by the chemical potential difference existing between the supercooled pore water some distance from the freezing site and the pre-melted liquid film directly in contact with the formed crystals (cf. Eq. (2)), and has been demonstrated in the third validation example.

In the subsequent stage, a supplementary step, taking into account the formation and growth of ice lenses within the frozen fringe, is to be carried on for an integrated description of the frost heave phenomenon. Besides, a more realistic inelastic material description of the freezing soil is to be considered.

ACKNOWLEDGEMENTS

This work has been financed by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation). This support is gratefully acknowledged. In addition, the first author thanks the Ruhr-University Research School funded by Germany's Excellence Initiative for further support [DFG GSC 98/1].

REFERENCES

- [1] O. Coussy. *Poromechanics*. J. Wiley & Sons, Ltd., 2004.

- [2] O. Coussy. Poromechanics of freezing materials. *J. Mech. Phys. Solids*, 53:1389–1718, 2005.
- [3] A. Fabbri, T. Fen-Chong, and O. Coussy. Dielectric capacity liquid water content, and pore structure of thawing-freezing materials. *Could Regions Science and Technology*, 44:52–66, 2006.
- [4] M. Mikkola and J. Hartikainen. Mathematical model of soil freezing and its numerical implementation. European Conference on Computational Mechanics, pages 159–174, München, Germany, August 1999.
- [5] M. Mikkola and J. Hartikainen. Mathematical model of soil freezing and its numerical implementation. *International Journal for Numerical Methods in Engineering*, 52:543–557, 2001.
- [6] M. Mikkola and J. Hartikainen. Computational aspects of soil freezing problem. In H.A. Mang, F.G. Rammerstorfer, and J. Eberhardsteiner, editors, *Fifth World Congress on Computational Mechanics*, volume 5 of *WCCM*, Vienna, July 7-12 2002.
- [7] J. Kruschwitz and J. Bluhm. Modeling of ice formation in porous solids with regard to the description of frost damage. *Computational Materials Science*, 32:407–417, 2005.
- [8] Joachim Bluhm and Tim Ricken. Modeling of freezing and thawing processes in liquid filled thermo-elastic porous solids. In M. J. Setzer, editor, *Proceeding of the 5th International Conference on Transport in Concrete*, pages 41 – 57, 2007.
- [9] Reint de Boer. *Theory of Porous Media - highlights in the historical development and current state.*, volume 2. Springer-Verlag, Berlin, Heidelberg, New York, 2000.
- [10] W. Ehlers. Poröse medien, ein kontinuumsmechanisches modell auf der basis der mischungstheorie. Technical report, Universität Gesamthochschule Essen, 1989.
- [11] O. Coussy. Poroelastic model for concrete exposed to freezing temperatures. *Cement and Concrete Research*, 2007.
- [12] Olivier Coussy. *Mechanics and Physics of Porous Solids*. John Wiley and Sons, 2010.
- [13] M.T. van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America*, 44:892–898, 1980.
- [14] Oliver Coussy and Paulo J.M. Monteiro. Poroelastic model for concrete. *Cement and Concrete Research*, 38:40 – 48, 2008.
- [15] N. Boukpeti. One-dimensional analysis of a poroelastic medium during freezing. *Int. J. Numer. Anal. Meth. Geomech.*, 32:1661–1691, 2008.

- [16] F. Brezzi and M. Fortin. Mixed and hybrid finite element methods. *Springer*, 1991.
- [17] D. Kuhl and M.A. Crisfield. Energy-conserving and decaying algorithms in non-linear structural dynamics. *Int. J. Num. Meth. in Eng.*, 45(5):569–599, 1999.
- [18] Yongjoo Lee and K. C. Park. Numerically generated tangent stiffness matrices for nonlinear structural analysis. *Computer Methods in Applied Mechanics and Engineering*, 191(51-52):5833–5846, 2002.
- [19] P. Dadvand, J. Mora, C. González, A. Arraez, P.A. Ubach, and E. Oñate. KRATOS: An Object-Oriented Environment for Development of Multi-Physics Analysis Software. In H.A. Mang, F.G. Rammerstorfer, and J. Eberhardsteiner, editors, *WCCM V, Fifth World Congress on Computational Mechanics*, 2002.
- [20] R. Lackner, A. Amon, and H. Lagger. Artificial ground freezing of fully saturated soil: Thermal problem. *Journal of Engineering Mechanics*, 131(2):211–220, February 2005.
- [21] J.J. Beaudoin and C. MacInnis. The mechanism of frost damage in hardened cement paste. *Cement and Concrete Research*, 4:139–147, 1974.

ADAPTIVE CALIBRATION OF A NONLOCAL COUPLED DAMAGE PLASTICITY MODEL FOR ALUMINIUM ALLOY AA6082 T0

J.P. BELNOUE* AND A.M. KORSUNSKY*

* Department of Engineering Science, Parks Road
University of Oxford
Oxford, UK OX1 3PJ

Email: alexander.korsunsky@eng.ox.ac.uk, Web page: <http://www.eng.ox.ac.uk/solidmech/>

Key words: Continuum Damage Mechanics, Ductile materials, Nonlocal damage, Parameter identification, Damage function formulation

Abstract. Continuum Damage Mechanics (CDM) accounts for material degradation (softening and ultimately failure) by modifying the load-bearing properties of the material (stiffness and strength) through a special state variable referred to as damage. Damage is typically represented by a scalar or a higher dimension object (such as vector or tensor) with values between zero for virgin material and unity for the material that lost all its bearing capacity. Considered in this way, damage becomes an additional field quantity that needs to be considered along with strain and stress, and can be computed either incrementally, or as a certain function of a suitable physical parameter such as inelastic strain. The advantage of enriching the formulation of a continuum deformation problem with a damage parameter is that it allows considering the material post-critical behaviour, i.e. its response under deformations exceeding those when the maximum load-bearing capacity is reached. Typically, this post-critical behaviour is associated with strain localisation, initiation, growth and interaction of discontinuities, and final fracture. Within the CDM framework, cracks are represented by diffuse regions of material damaged so that it lost all its strength in at least one direction. Computationally, modelling the post-critical (softening) behaviour of material represents a challenge in terms of the numerical stability of algorithms. Nonlocal description of damage appears to offer a rational route towards stable modelling. Nonlocal averaging of the plastic strain for the evaluation of damage also renders CDM models independent of the mesh size and orientation, and helps overcome numerical instabilities. The formulation that emerges can be referred to as coupled nonlocal damage-plasticity modelling [1, 2].

An important challenge remains, however, in developing this general approach into a flexible and material-specific modelling tool. This concerns the need to calibrate a large number of material parameters that emerge in this formulation. In order to address this challenge, recently we developed an approach for the calibration of CDM models of ductile materials that we propose to refer to as *adaptive calibration*. The calibration of the damage function is accomplished by matching the model prediction to the experimental data obtained from a single tensile test with multiple gauge length extensometry [3] used to capture strain localisation and size effects. We describe the application and validation of this approach to the damage function parameter calibration for the aluminium alloy AA 6082 T0. Excellent agreement with experimental measurements is obtained.

1 INTRODUCTION

The approaches to the issue of structural analysis in the aeronautical industry have evolved over the decades. Original methodologies relied on the principle of avoiding crack initiation in order to ensure fatigue resistance and durability. Later, damage tolerant design principles were introduced, in which the presence of small defects (cracks) was accepted as unavoidable, and emphasis was placed on ensuring that they do not grow to critical lengths within the exploitation intervals between inspections. In the current view, not only crack initiation, but also crack propagation and trajectory analysis are very important for ensuring safe design and operation. Aeroengine components are subjected to complex loading induced by the combination of mechanical loading, changing temperatures and thermal gradients, inducing plastic deformation and creep that ultimately may lead to crack initiation and propagation. Along with an increasing proportion of composite materials, many aerospace components continue to be made from ductile metallic materials, such as nickel-base superalloys, titanium and aluminium alloys. In these materials rupture (material separation) is preceded by damage that finds its physical representation in void nucleation, growth and coalescence. The development of populations of defects in ductile materials is responsible for material softening that is played out in competition with the usual strain hardening behaviour often observed when ductile metal alloys are subjected to (tensile) loading. Modelling the behaviour of these materials in the post-peak softening regime and the correct and reliable prediction of crack propagation even today represent a serious challenge.

Predicting efficiently crack propagation rates and directions has been the aim of many researches since the emergence and the development, in the 1970's, of numerical methods such as Finite Elements (FE). The traditional approach consists in representing the crack as a discontinuity of the modelled structure, for example, introducing cohesive zone elements [4] on the assumed crack path or using the Extended Finite Element Method (XFEM) [5]. This so-called discontinuous approach to crack problems is well-suited to describing the failure process of brittle materials in which the crack path is known in advance (e.g. laminate composites). However, if one wants to model a failure process (e.g. ductile failure) happening in a less sudden manner, in which the deformation process is more homogeneously distributed and for which the crack path is not known *a-priori*, then discontinuous crack approach becomes problematic. Therefore, simultaneously with the development of discontinuous crack modelling methods, a second approach, called Continuous Damage Mechanics (CDM) [6, 7, 8] has been proposed and explored. Material degradation is taken into account through the use of a state variable whose value varies between zero (for virgin material) and unity (for fully damaged material) and that modifies the material stiffness and strength. Coupling between damage and plasticity is introduced by making the damage variable dependent on the equivalent plastic strain. The method is particularly attractive in the sense that it allows to take into account some of the most fundamental aspects of the failure process that are particularly important in the context of ductile rupture (e.g. as material softening, strain localization and size effects) and also to grow a crack without knowing the path in advance. However, in the course of advancing this approach to modelling both crack initiation and propagation in an internally consistent fashion, two major shortcomings of the CDM method have been identified:

1. Material softening and strain localization that occur at the late stage of the deformation process (i.e. when the damage becomes high enough to overcome the strain hardening) are responsible for numerical instabilities and loss of ellipticity of the governing equation [9].
2. In terms of the effect of the finite element mesh size, it is observed that the energy dissipated during the failure process is reduced with mesh refinement. In the extreme case of an infinitely fine mesh, the classical CDM models predict no energy dissipation during the failure process; a conclusion that is clearly physically unrealistic. This is a manifestation of the phenomenon known as mesh-dependency.

A further characteristic feature of CDM-type models is that they do not capture the discontinuous nature of cracks precisely in the form of interfaces that open up between adjacent volumes of material. Instead, cracks are represented as diffuse regions (bands) of damaged material. The damage variable reaches the ultimate value of unity at the mid-line of these bands, so that the mid-line can be associated with the crack line.

Over the years, various modifications have been elaborated aiming at making CDM models more reliable. A classical solution consists in replacing the “conventional” equivalent plastic strain by its weighted average over a sphere of radius R (the nonlocal equivalent plastic strain) in the damage function formulation [10]. Nonlocal models have been first applied to describe the failure behaviour of brittle materials [10, 11]. Recently, the authors of the present paper proposed a model aiming at describing the failure behaviour of ductile materials [1, 2, 12]. Until now, most efforts have been directed to resolving the numerical issues in the 1D [1] and 2D (plane stress) [2, 12] formulations. In the present contribution, a closer look is taken at the identification of the material parameters used by the model.

In the approach described here, the model parameter calibration is realized using experimental data obtained from a single tensile test in which multiple gauge length extensometry is used to capture the strain localisation and size effects. The material considered in the present study is a ductile aluminium alloy AA6082 T0 [3]. The material plastic hardening behaviour is described by a combination of linear and Chaboche isotropic hardening law. The damage behaviour is captured using a novel approach called *adaptive calibration* that allows to extract the material behaviour in the post-peak softening regime *without* recourse to complex or iterative parameter identification procedures (e.g. Markovian optimisation [13], least squares fitting [14], Levenberg-Marquardt algorithm [15]).

Firstly, the idea of adaptive damage calibration method [16] is briefly introduced and reviewed. Secondly, the damage law (the damage evolution as a function of the nonlocal equivalent plastic strain) that is obtained by the adaptive calibration procedure is approximated as a simple piece-wise function in order to ease its use within the finite element implementation. Finally, advantages and drawbacks of the method are discussed.

2 EXPERIMENTAL PROCEDURE

Flat dogbone specimens of the shape shown in Figure 1 were prepared from an aluminium alloy AA 6082 T0 by mechanical milling of 3 mm-thick sheets. The T0 heat treatment was chosen so as to bestow low strength and high ductility on the material. The specimens were solution treated at 570°C for various amounts of time depending on their thickness, and air cooled afterwards.

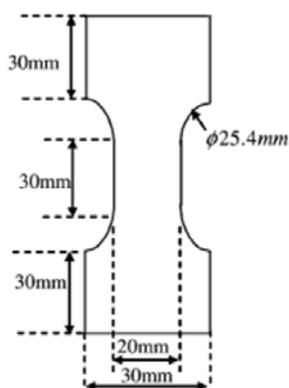


Figure 1: Dog-bone specimen design (from [3]).

Tensile testing until complete failure by rupture was performed using a screw-driven Hounsfield universal testing machine with Servocon control and acquisition system (see Figure 2). The crosshead speed was set to 2 mm/min. As illustrated in Figure 2(b), white bands of flexible paint were printed on the specimen surface so as to obtain a screen of contrasting markers. The positions of band boundaries were continuously measured with a Fiedler Optoelektronik laser scanning extensometer (with a spatial resolution of about 0.5 μm). Continuous recording of band boundaries positions allows the determination of extension (and average strain) between any two band boundaries at any time during the deformation process. Laser extensometry thus provides the records of load-extension curves for multiple gauge length specimens within one experiment.

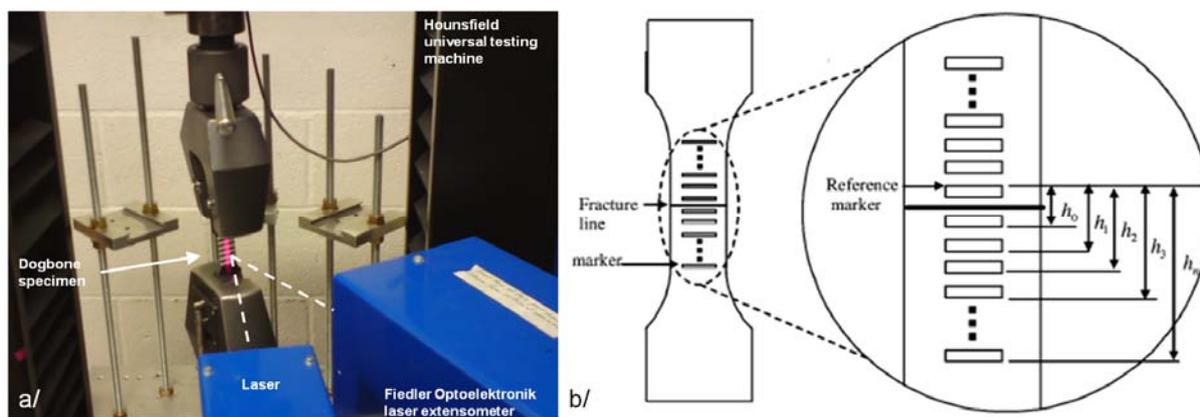


Figure 2: Laser extensometry a/ Experimental setup - b/ Dog-bone specimen: markers and the multiple gauge length definition (from [3]).

The experimental technique chosen for this study thus allows plotting multiple Load-Extension curves based on the data collected from a single dog-bone specimen experiment. As illustrated in Figure 2(b), a marker located at some small stand-off distance above the fracture line can be chosen as reference. All other markers located below the fracture line are used to define “sub-specimens” of gauge lengths h_i . Knowing the time evolution of marker positions allows to calculate the extension of every "sub-specimen" (of initial length h) at any time during the deformation process. The interpretation of laser extensometry results gives rise to experimental curves such as those presented in Figure 3. It also makes it possible to calculate the strain anywhere in the bar and at any time during the deformation process. This particular capability is of crucial importance for the determination of the characteristic material length (the nonlocal radius R).

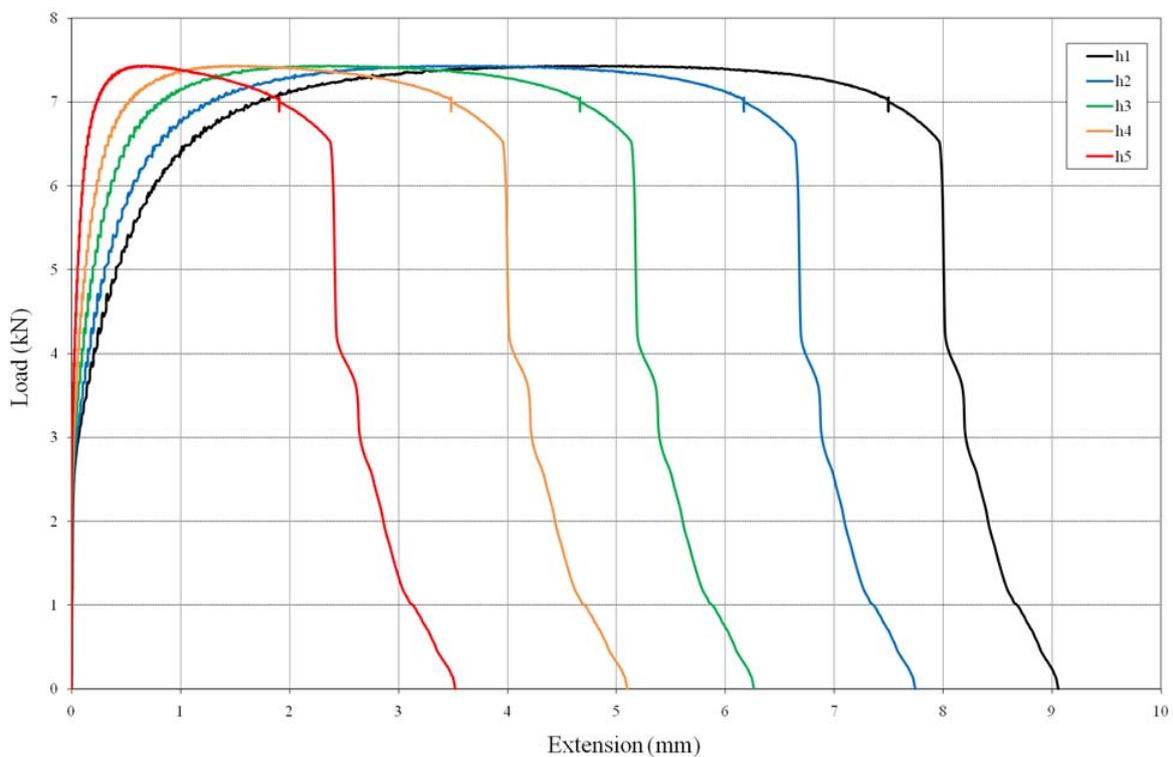


Figure 3: Load-extension curves of a 3 mm-thick dog-bone specimen obtained for sub-specimen of initial lengths: $h_1=23$ mm, $h_2=17$ mm, $h_3=11$ mm, $h_4=7$ mm, and $h_5=3$ mm.

As stated in previous publications on the subject [17, 18], the information needed to calibrate CDM models (i.e. the characteristic material length, the damage function shape and the associated parameters values) must be contained in the Load-Extension curves presented in Figure 3. Providing that a suitable calibration procedure for the model parameters is used, a successful finite element simulation of the tensile test on a dog-bone specimen must provide simultaneous satisfactory agreement with all of the individual stress-strain curves presented in Figure 3.

3 DUCTILE FAILURE MODELLING

3.1 The model

The multiple gauge length tensile test described in Section 2 above was first represented in the form of an 1D FE model. The model consisted of a bar with one extremity kept fixed whilst the other one was subjected to incremental outward displacement (see [1]). The deformation of the material elements along the bar that ensued is described using the 1D version of the coupled nonlocal damage-plasticity model. The simulation involves sequential calculation of increments of stress, strains, damage within each element, and the tangential stiffness that must show agreement with the experimental results. This provides the basis for determining the damage function evolution with deformation by the means of the adaptive calibration procedure described below.

3.2 The adaptive calibration method

For the purpose of parameter identification, the assumption is made that at early stage of the deformation process, the damage has very little effect on the material stress-strain response. The Young's modulus E of 73 GPa was found by fitting the initial linear behaviour by a straight line, the material yield stress at 0.2% of plastic strain, σ_{Y0} , was fixed at 50 MPa. An Excel spreadsheet was setup in order to vary the description of the plastic behaviour by a combined linear and Chaboche type isotropic hardening function [19, 16] (see (1) below). The best fit material parameters of the model were found to be $h=50$ (to account for the linear contribution to strain hardening), and $b=37.5$ and $Q=67.5$ (Chaboche function parameters), leading to the overall expression:

$$\sigma_Y = \sigma_{Y_0} + h\varepsilon_p + Q(1 - e^{-b\varepsilon_p}) \quad (1)$$

In equation (1), σ_Y and ε_p stand respectively for the current yield stress (i.e. modified by the deformation history) and the plastic strain at instant t . An optimal fit of the experimental curve can be achieved using automatic methods (e.g. least-squares) for the determination of the parameters h , Q and b .

During testing, the continuous recording of paint band boundary positions allowed following the evolution of the strain distribution along the specimen. At the early stages of the deformation process, the strain in the bar remained low and homogeneously distributed. During the elastic stage of the deformation process, and all the way to the critical point of maximum overall load borne by the specimen, the strain distribution remained uniform. This part of the process corresponds to the conditions when strain “levelling” effects due to hardening dominate over localization (promoted by damage). The localization of plastic strain, into an area of approximately 4 mm, happened at a relatively late stage of the deformation process, once the critical elongation was exceeded, and necking (damage and strain localization) began. Previous analysis [16] of the effect of the nonlocal radius on the strain distribution just before failure suggest that R should therefore be set to 4 mm.

The concept of *adaptive damage calibration* was described extensively in [16]. The idea of the adaptive damage function simulation procedure is as follows. Firstly, let us note the similarity between the incremental description of the processes of damage and plastic deformation. For a given material volume at any stage in its deformation history, the prior behaviour is already known and stored in the computer memory. The subsequent behaviour is found in a stepwise fashion, at each stage only for a short time period corresponding to a small increment of deformation. Wishing to apply this principle to the description of the damage process, we note that at any particular stage in the damage process it is not necessary to know the entire damage function. Rather, it suffices to know the instantaneous slope of the damage function. This allows the calculation of the effective tangent stiffness, and hence the determination of the subsequent deformation increment.

Therefore, the following procedure for adaptive damage function calibration was put forward. In the first instance, the threshold value of plastic strain was chosen beyond which the damage process was likely to begin. As will be seen further, guessing a value for this parameter that is too low is not detrimental to the accuracy, and may only result in some slowing down of the algorithm. Other aspects of the calculation also need to be set at this stage, namely, the strain increment to be used, and the initial guesses for the slope of the damage function (these can be conveniently chosen to be zero and a non-negative number, e.g. 0.1). The calculation of the next segment of the stress-strain curve is then carried out twice: without damage (or, more precisely, with the damage function slope of zero), and with the damage switched on. This results in two predictions in the form of linear segments describing the stress-strain curve over the next increment. The data available from the experimental stress-strain curve is then used to guide the correct selection of the damage function to deliver the best match. This process can be accelerated by using linear interpolation, and repeated incrementally, until certain tolerance is attained. The procedure is then repeated for the next and subsequent increments. In this way, a piecewise linear description of the damage function dependence on the plastic strain is built up. One remarkable aspect of this approach is that this piecewise linear description readily affords itself to be used as a lookup table for further calculation. Therefore only the damage in the most deformed element need to be adjusted. Damage in the other elements just follows the evolution that has already happened and been recorded in the most deformed element.

The procedure described here is adaptive in the sense that it determines the damage function profile in a way that ensures incremental adjustment so as to preserve the agreement with the experimental data throughout the deformation history. The results of the application of this procedure are illustrated in Figure 4 for the rising (box 1) and falling (box 2) parts of the stress-strain curve. It is apparent that the step size and initial guesses for the damage function slope could be chosen to deliver satisfactory agreement for both parts of the experimental curve.

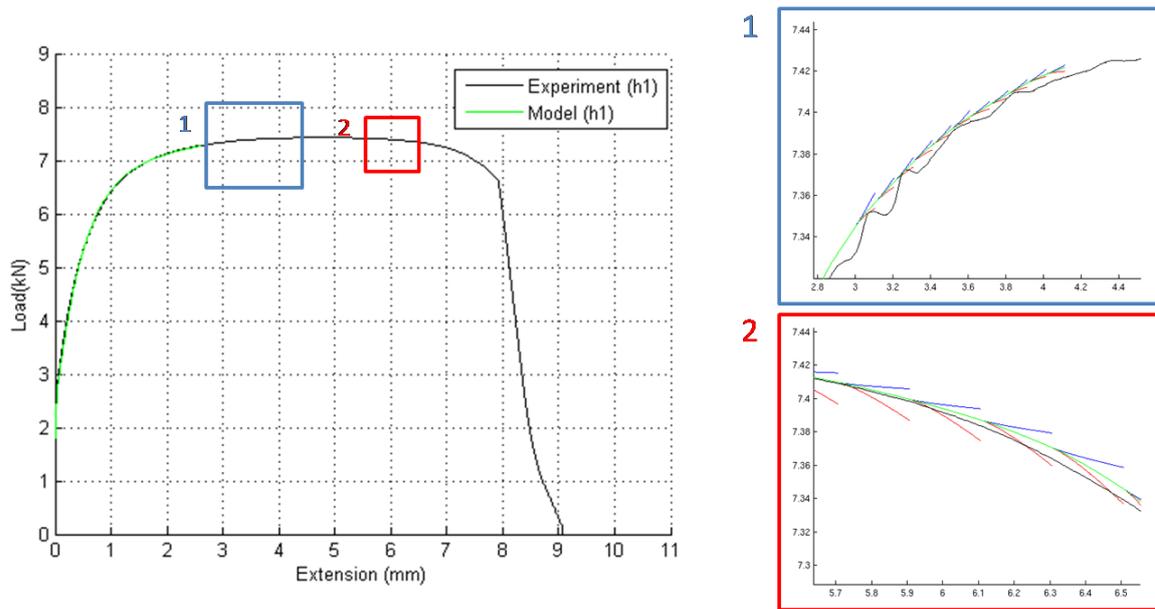


Figure 4: The adaptive damage calibration method makes the assumption that the damage function is piecewise linear. Two initial guesses are made for the initial slope of the damage function, and computation is carried out based on both guessed values. The intermediate slope of the damage function that allows good match to the experimental data is then determined by linear interpolation.

Figure 5(a) illustrates the damage function curve that was obtained after the adaptive simulation was carried out until complete sample failure. It is worth noting here that an artificial modification was applied to the damage curve for the range of damage parameter values $0.4 < \alpha_d < 1$. When the damage approaches 0.4, unstable and rapid failure takes place. This process is too fast for the quasistatic loading machine to capture correctly, so that the experimental points obtained beyond this stage are not fully meaningful any more. Subsequent damage evolution to the value of unity is modelled by prescribing a very high slope to the damage function. This smooth approach to unity is dictated by numerical stability reasons.

Figure 5(b) illustrates the quality of the resulting simulation. The stress-strain curves for all the sub-specimens considered are captured equally well, indicating that the size effect and the energetic characteristics of the plastic rupture process are adequately represented by the model. It is worth noting here that the adaptive damage process has been applied only to the analysis of the load-displacement response of the longest sub-specimen. The observed high quality agreement between the model and the experimental data for all other sub-specimens thus furnishes a good evidence of the accuracy of the nonlocal characteristic length determination, and of the overall coupled nonlocal damage model.

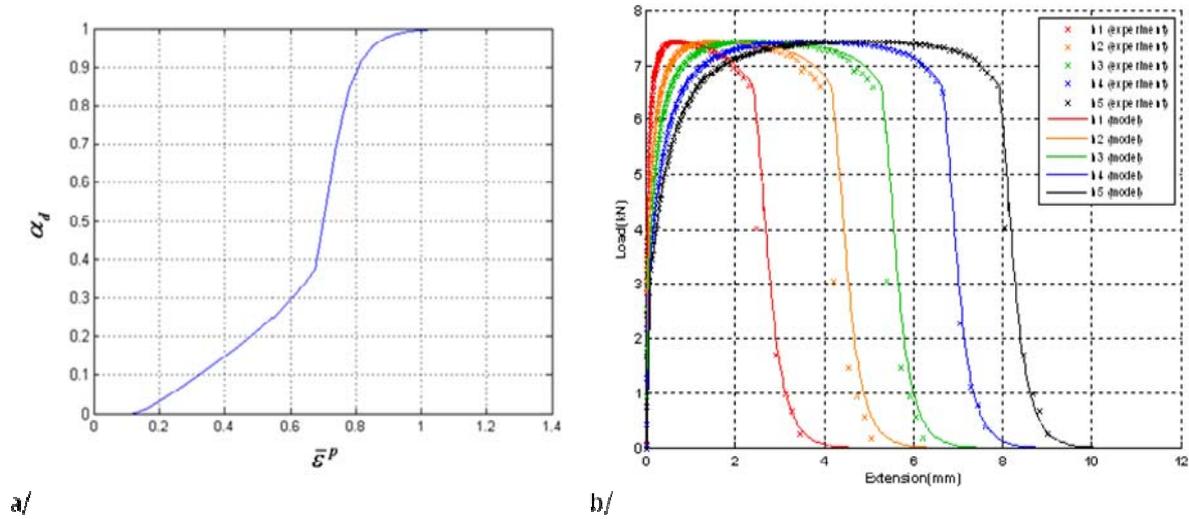


Figure 5: a) Damage function obtained from the adaptive damage calibration procedure – b) Load-extension curves for a 3 mm-thick dog-bone specimen obtained for sub-specimens of initial lengths: $h_1=23$ mm, $h_2=17$ mm, $h_3=11$ mm, $h_4=7$ mm, and $h_5=3$ mm. Model predictions are continuous curves superimposed on the experimental data (markers).

3.3 Damage function: a simple analytical formulation

Although the match between the model and the experimental data looks satisfactory, the adaptive damage calibration method suffers from one major drawback: the resulting damage function is stored in the form of a table corresponding to a certain number of data points. This makes the use of the damage function in an FE code (e.g. in 2D plane stress) somewhat tedious. We therefore propose to fit the obtained damage function by a piece-wise combination of two functions, one for the range $0 < \alpha_d < 0.4$, and the other for the range $0.4 < \alpha_d < 1$. Figure 6 illustrates that satisfactory description is obtained by using the following equation:

$$f = \begin{cases} 0.5785 \times (\bar{\epsilon}_p - 0.1293) & \text{if } 0 < \alpha_d < 0.4 \\ 1 - 8 \times (\bar{\epsilon}_p - 1.3)^4 & \text{if } 0.4 \leq \alpha_d < 1 \end{cases} \quad (2)$$

Figure 7 illustrates the load-extension curves obtained from the simulation based on the approximation of the damage function described above. The load-extension curves for all sub-specimens considered appear to have been captured relatively well. It is clear that equation (2) provides only an approximation of the correct description of the damage function evolution. However, the accuracy obtained appears to be good enough to justify using this formulation to streamline and speed up FE simulations of the damage process. It can be concluded, furthermore, that abrupt change in the damage function slope introduces a similar discontinuous slope feature in the simulation. Consequently, it appears logical to adopt a rule that relative changes in the damage function slope must be restricted to ensure smoothness of the simulated load-displacement curves.

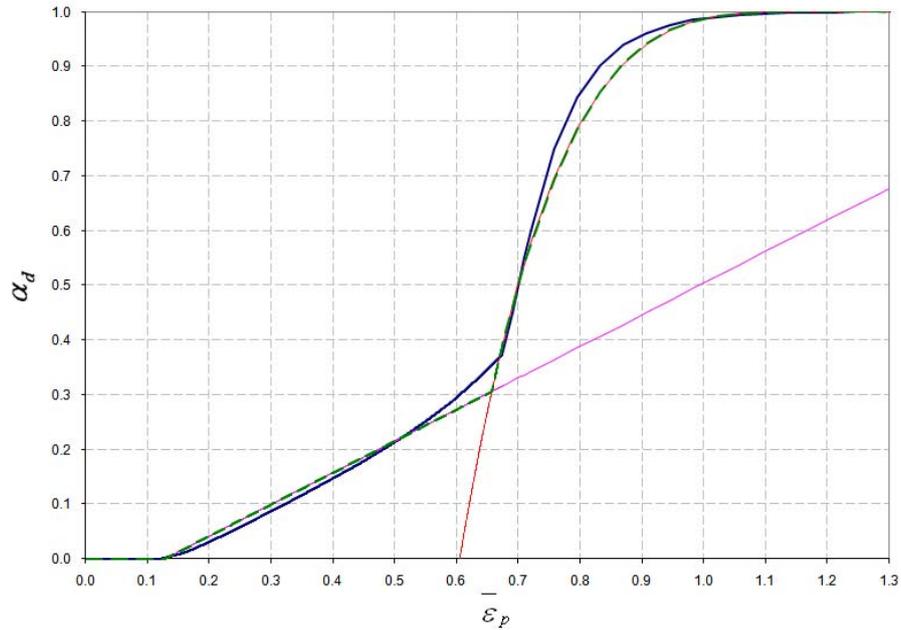


Figure 6: The damage function obtained using the adaptive calibration method (in blue) is fitted by a linear function (in pink). Unstable crack propagation is modelled by a quadratic function (in red).

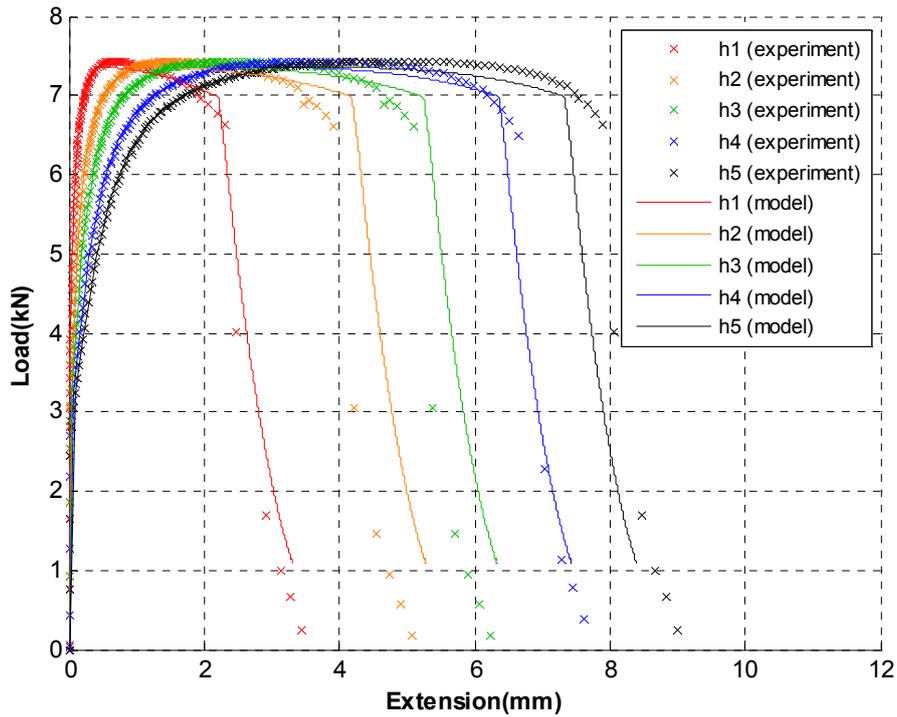


Figure 7: Load-extension curves of a 3 mm-thick dog-bone specimen obtained for sub-specimen of initial lengths: $h_1=23$ mm, $h_2=17$ mm, $h_3=11$ mm, $h_4=7$ mm, and $h_5=3$ mm. Model predictions are superimposed on experimental data. The damage function proposed in Figure 7 (green line) is used.

4 DISCUSSION AND CONCLUSION

The adaptive damage calibration procedure combined with the approximate analytical description of the damage function provides a convenient framework for simulating post-critical softening. Once the calibration is obtained, the damageable material model can be run efficiently within the finite element framework.

It is apparent from Figure 3 and Figure 7 that the post-critical part of the load-displacement curve is only captured well until the damage parameter value reaches the value of 0.4 or so. In order to capture the subsequent evolution of the load-displacement curves beyond this stage, high rate acquisition of displacement and load is required.

Accomplishing both of these tasks may be possible by employing a suitably fast imaging camera, such as the Vision Research Phantom v7.3 high speed camera capable of over 6000 frames per second at full resolution of 800×600 pixels (and up to 0.5M frames per second at reduced resolution). However, it is important to note that dynamic deformation effects would need to be taken into account, so that an explicit Finite Element formulation must be used.

REFERENCES

- [1] Belnoue JP, Nguyen GD, Korsunsky AM, A one-dimensional nonlocal damage-plasticity model for ductile materials, *Intl. J. Fracture*, **144**, pp. 53-60 (2007).
- [2] Belnoue JP, Garnham B, Bache M, et al. The use of coupled nonlocal damage-plasticity to predict crack growth in ductile metal plates, *Eng. Frac. Mech.*, **77**, pp.1721-1729 (2010).
- [3] Korsunsky AM, Kim K, Determination of essential work of necking and tearing from a single tensile test, *Intl. J. Fracture*, **132**, pp. L37-44 (2005).
- [4] Needleman A, A Continuum Model for Void Nucleation by Inclusion Debonding, *J. Appl. Mech.*, **54**(3), pp.525-531 (1987).
- [5] Moes N, Dolbow J, Belytschko T, A finite element method for crack growth without remeshing, *Int. J. Numer. Meth. Engng.*, **46**, pp. 131-150 (1999).
- [6] Kachanov LM, Rupture time in creep conditions [in Russian], *Izd. AN SSSR*, **8**, (1958).
- [7] Rabotnov YN, Creep of Structural Elements [in Russian], *Nauka*, Moscow (1966).
- [8] Lemaitre J, A course on damage mechanics, Springer (1992)
- [9] Benallal A, Billardon R, Geymonat G, Some mathematical aspects of the damage softening rate problem. In: Cracking and damage (Mazars, J., Bazant, Z. P., eds.), pp. 247-258. Amsterdam London: Elsevier (1988).
- [10] Pijaudier-Cabot G, Bazant ZP, Nonlocal damage theory, *J. Engrg. Mech. ASCE*, **113**, pp. 1512–1533 (1987).
- [11] Nguyen GD, A Thermodynamic Approach to Constitutive Modelling of Concrete using Damage Mechanics and Plasticity Theory, D.Phil. Thesis, Oxford University (2005).
- [12] Belnoue JP, Nguyen GD, Korsunsky AM, Consistent tangent stiffness for local-nonlocal damage modelling of metals, *Procedia Engineering*, **1**(1), pp. 177-180 (2009).

- [13] Carmeliet J, Optimal estimation of gradient damage parameters from localisation phenomena in quasi-brittle materials. *Int. J. Mech. Cohesive Frict. Mats.*, **4**, pp. 1–16 (1999).
- [14] Mahnken R, Kuhl E, Parameter identification of gradient enhanced damage models with the finite element method, *Eur. J. Mech. A Solids*, **18**, pp. 819–835 (1999).
- [15] Le Bellégo C, Dube JF, Pijaudier-Cabot G, Gerard B, Calibration of nonlocal damage model from size effect tests, *European Journal of Mechanics A/Solids*, **22**, pp. 33–46 (2003).
- [16] Belnoue JP, Korsunsky AM, A damage function formulation for nonlocal coupled damage-plasticity model of Aluminium Alloy AA 6082 T0, *Intl. J. Fatigue*, under review
- [17] Korsunsky AM, Nguyen GD, Houlsby GT, Analysis of essential work of rupture using non-local damage-plasticity modeling, *Int. J. Fract.*, **135**, pp. 19-26 (2005)
- [18] Korsunsky AM, Nguyen GD, Kim K, The analysis of deformation size effects using multiple gauge length extensometry and the essential work of rupture concept. *Materials Science and Engineering A*, 423(1-2), pp. 192-198 (2006).
- [19] Chaboche JL, Constitutive equations for cyclic plasticity and cyclic viscoplasticity, *Int. J. Plast.*, **5**(3), pp. 247-302 (1989).

ADVANCED STOCHASTIC FEM-BASED ARTIFICIAL NEURAL NETWORK FOR CRACK DAMAGE DETECTION

C. SBARUFATTI^{*}, A. MANES[§] AND M. GIGLIO[†]

^{*} Politecnico di Milano (Dipartimento di Meccanica)
Via la Masa 1, 20156 Milano, Italy

E-mail: claudio.sbarufatti@mail.polimi.it - Web page: <http://www.mecc.polimi.it>

[§] Politecnico di Milano (Dipartimento di Meccanica)
Via la Masa 1, 20156 Milano, Italy

E-mail: andrea.manes@polimi.it - Web page: <http://www.mecc.polimi.it>

[†] Politecnico di Milano (Dipartimento di Meccanica)
Via la Masa 1, 20156 Milano, Italy

E-mail: marco.giglio@polimi.it - Web page: <http://www.mecc.polimi.it>

Key words: Structural Health Monitoring, Artificial Neural Network, structural diagnosis, crack, helicopter.

Abstract. Structural Health Monitoring (SHM) is nowadays one of the most challenging research fields. As a matter of fact, if from one hand the aerospace industry is trying to extend the duration of life-limited components, from the other hand a deep control is necessary over the structures to guarantee both the machine availability and reliability. In effect, thanks to the advance in the evaluation of the actual structural health by means of a SHM system, it could be possible to set a Condition Based Maintenance (CBM). This approach means substituting a component according to its real structural conditions instead of relying just on the design assumptions. The final aim is to update the scheduled maintenance intervals according to the actual condition of the structures. However this is not an easy task, as it is governed and influenced by many variables, each one characterized by a stochastic distribution. In particular, the key factor is the disposal of detection and monitoring systems as reliable as possible in order to conjugate safety with economics objective. On the basis of this all the machine stops can be optimized in order to exploit the machine availability with the minimum loss of reliability. Thus, the first step for developing such advanced technology would be the disposal of a robust damage detection system, able to recognise, locate and quantify the damage in a certain component. The aim of the present work is to define a methodology that combines the use of Finite Element Models (FEM) with Artificial Neural Networks (ANN) [1] for crack detection over a typical aerospace structure consisting of a riveted aluminium skin stiffened with some reinforcing elements [2]. Numerical models, in fact, could be used to train ANN. A basic system knowledge would result, upon which to introduce the variability by means of real sensor network data [3], in order to consider the problem from a statistical point of view. Finally, a proposal for the sensor network characterization in terms of Probability of Detection (PoD) and False Alarm (PFA) is also reported.

1 INTRODUCTION

Helicopters are very critical machines for a fatigue standpoint. In fact they are subjected to a very wide range of load cycles; high frequency load cycles, from one hand, due to the interaction between the rotors and the air and characterized by lower amplitude; low frequency loads, due to the wide variety of manoeuvres and characterized by higher amplitude. In addition, due to the harsh environment where such machines operate, low velocity impacts and corrosion issues are sources for further flaw nucleation. Maintenance of a helicopter is thus one of the most critical costs, being about 25% of the operating cost. In order to assess the damage criticality of aging aircraft, as well as to minimize the scheduled periodic inspection and eliminate the costly unscheduled maintenance in new aircraft, application of an 'on-board' Structural Health Monitoring (SHM) system would be required. On-board health monitoring systems of real structures require installation of networks of embedded sensors. Such sensory networks should be clustered around structural "hot spots" and be connected to a data repository and/or processor, enabling to monitor the structure and identify the presence of defects and incipient damage.

However, damage identification is a very difficult challenge, especially when dealing with a structure which is subjected to an extremely wide range of loads as well as noise. The first thing to be addressed is the identification of the best parameter for crack characterization (detection, localization and extent) as well as the most suitable sensor technology for parameter monitoring. Many technologies are available to date for Non-Destructive Evaluation (NDE) of structural health, among which strain measurements, mechanical wave based sensor, piezoelectric sensors, potential drop measurements, eddy current based detection systems, MEMS technologies, crack gauges, etc. In particular, the choice undergone in the current paper is to monitor cracks by investigating the induced changes in the strain field [2], which can be measured by means of classical resistance-based strain gauges [3] or through optical Fiber Bragg Gratings (FBG) [4]. The consecutive step consists in the choice of the algorithms for data processing, the aim of which is to infer over the actual state of the structure, according to the information extracted from the measurements provided by the available sensor network. It is possible to distinguish two typical methodologies to do inference in complex SHM systems, respectively based upon the Maximum Likelihood and the Bayesian Approach, that is to say Artificial Neural Networks (ANN) [5,6] and Bayesian Networks [7,8]. The main difference between the two methods is that, while ANN has the final aim to obtain one deterministic network, the best one, or the one related with the smallest error, the Bayesian approach doesn't refuse any possible parameter configuration a priori, storing all the possible network evaluated during the training thus keeping as test output the expected value coming from all the outputs of all the possible networks. Both the methodologies have been widely developed in the recent past and they appear to be applicable for structural diagnosis. However, ANNs look very promising, because of their ability to "reason" on the basis of the knowledge created during the training phase. Even better, this basic knowledge might be provided through finite element simulations, thus reducing the cost of the design phase. Nevertheless, it is important to appreciate the on-field noise as well as the deviation of the reality with respect to the FE model, by means of experimental tests, thus allowing for a more optimized network training. Coupling of FE models with advanced diagnostic algorithms is the most suitable direction for a feasible design of Structural Health

Monitoring (SHM) systems, as emerges from the recent studies applied to various sensor technologies. In [9], FE models are used to investigate the transient scattering of Rayleigh wave induced by a surface crack in a plate. ANNs are thus applied to establish the mapping relationship between the characteristic of the reflected waveform and the crack depth in aluminum plates, proving that the method can be applied to online structural damage detection and health monitoring for various industrial structures. Moreover, in [10] parameterized modeling for finite element analysis and an information mapping approach are applied to cost-effectively produce the damage parameter database for ANN training. A good accuracy was obtained in damage parameters (central position, size, and orientation) by monitoring the Lamb Wave signal scattering. However, not many attempts have been carried out to build a complete SHM system based upon strain measures over the monitored region inside the helicopter.

The performance evaluation of an Artificial Neural Network for damage characterization over a helicopter fuselage stiffened skin is presented herein, based upon strain measures. The disposal of a detailed Finite Element Model (FEM) allows for a low-cost training of the network. However it is necessary to introduce some noise both at training and testing level, in order to appreciate the effective performances of the simulated sensor network, thus designing a system as reliable as possible.

2 THE ARTIFICIAL NEURAL NETWORK PROBLEM

Artificial Neural Networks are algorithms able to learn from experience, thus elaborating new answers even in the case of new inputs. Practically, they can model reality by learning the underlying rules, even for those events for which a mathematical formulation is not available; the network is thus able to automatically reproduce the link between causes and consequences, or the inputs and outputs of the ANN. An Artificial Neural Network can be defined as a dynamic model represented through a flow chart, where nodes and links respectively symbolize artificial neurons and synaptic weights. Nodes are then organized into layers (*figure 1a*): the first one is the input layer, followed by a certain number of hidden layers, finally ending with the output layer. Each node in a layer is connected by means of some weights to each node of the consecutive layer. Moreover, each node (except those belonging to the input layer) is also connected to a bias node by means of a special threshold weight (w_0). Thus, in equation (1) the *activation function* calculated in each j -th node is reported, where x_i and y_j are respectively the inputs and the output of the considered j -th node, while w_i are the weights associated to the i -th synapsis in input.

$$y_j = f\left(w_0 + \sum_i w_i x_i\right) \quad (1)$$

$$y_j = \frac{1}{1 + e^{-\left(\sum_{i=1}^n x_i w_{ji} + w_{j0}\right)}} \quad (2)$$

The ability of an ANN to learn depends on the choice of the activation function. Generally, the most used is the *sigmoid* one, reported in *figure 1b* and expression (2), where n is the number of synapsis entering the j -th node.

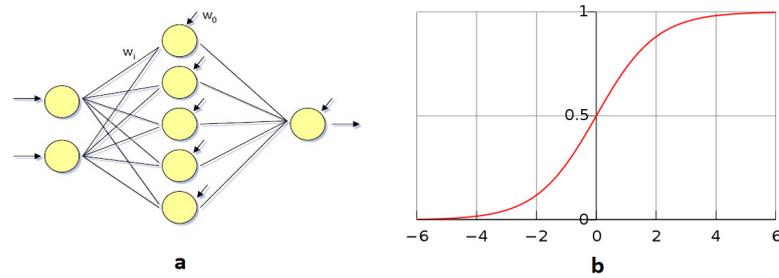


Figure 1: (a) ANN with single 5-node hidden layer, 2 input nodes, 1 output node (b) sigmoid function

During the training phase, all the synapsis weights have to be optimized in order to best fit the training data set, but without overfitting the data, thus maintaining the capability of the algorithm to generalize and to give the correct estimation even with input values different from those used in the training phase. The network optimization is carried out with the Back Propagation algorithm, which is based upon the Gradient Descent method, or on minimizing the error committed by the network, defined as in formula (3), where T_i and O_i are respectively the i -th target and i -th evaluated output.

$$E = \frac{1}{2} \sum_i (T_i - O_i)^2 \quad (3)$$

$$\Delta w_{ij}(n) = \eta \cdot D + \alpha \cdot \Delta w_{ij}(n-1) \quad (4)$$

In equation (4) one can notice that the variation of the weight parameters for each synapsis at the n -th iteration (Δw_{ij}), obtained from the derivation of the error E with respect to each weight, is mainly composed by two terms. The first one depends on the gradient (D) of the error E with respect to the considered weight (layer l and node j) and it is multiplied by a learning coefficient η . The second term is called momentum and serves as an inertia term, dependent on the amount of changes in weight parameters in the previous iteration, multiplied by an inertia coefficient α . The ANN algorithm optimization should thus be carried out by modifying α and η parameters as well as the number of hidden layers and neurons, thus finding the most performing solution.

In conclusion, the advantages of ANN algorithms are that no mathematical or physical modeling effort is needed and the system is able to perform an automatic parameter adjustment in the training phase based on available input/output data, maybe coming from FE models or joint FE/experimental simulations. The drawbacks are that it behaves as a “black box”, difficult to be interpreted in a physical way, and sometimes requiring a huge amount of input/output data for a proper training, increasing with the number of involved parameter, with the extent of non-linearity and with the complexity of the approximated functions.

3 DAMAGE PARAMETER DEFINITION

In an helicopter structure, as in the majority of aeronautical structures, the most susceptible areas for crack damage evolution are located either in joints or in particularly stressed zones. The former can be easily identified while the latter must be recognized by means of a detailed model of the entire fuselage of the machine, in order to highlight the most critical areas. An

on-board real-time monitoring system, which could accurately and reliably detect and quantify the crack, would be of great importance from a SHM point of view. Most of the proposed approaches for solving this problem are based on a wide variety of physical characteristics, such as, damping, vibration frequency, Lamb waves [9,10], etc. Although the above approaches have certain advantages, they also exhibit a series of serious disadvantages, most important of which are sensitivity to boundary and/or ambient conditions (e.g. applied load, temperature), utilization of intrusive sensors and in many cases inability to quantify the defects. To date, only few of the developed SHM methodologies are based on strain measurements [1,4,11]. The interested reader might refer to [1] for a detailed reference list.

Concerning the work presented in the current paper, the aim is to develop an SHM system based upon the strain field modifications, due to the presence of a damage. In [2] was demonstrated that the strain field sensitivity to crack damage can be exploited for structure monitoring; it was also confirmed the possibility to accurately model with Finite Elements the static and fatigue dynamic behavior of a stiffened structure. However, the main issue was the strain sensitivity also to the applied load. In fact, the helicopter fuselage is experiencing a load that is absolutely varying in time and, even worse, not easily predictable, due to the environmental influences. The first thing to do is to try to filter out the effect of the boundary load. Two approaches are suggested. The former is based upon the normalization of the actual strain measured by each sensor with respect to a reference value, estimated either from the tri-axial acceleration of the machine or from a sensor located far away from the damage. The latter, used hereafter, consists in normalizing each sensor of a confined region with respect to the average value measured among all the sensors in the same region. In fact, this normalizing factor should be robust to the presence of the damage but only sensible to the applied load, thus allowing to filter the load effect, without significantly altering the sensitivity to the crack. For that reason, the method remains valid only for linear material behavior as well as under the assumption that the damage effect remains a localized effect affecting the minority of the sensors, otherwise the damage information would enter inside the normalizing factor. In addition, it is important that the sensors participating to the average value are measuring in the same direction as the normalized sensor, in order to refer to the same load, which will be thus filtered out. Finally, assuming N strain sensors measuring exactly in the same direction, the expression of the normalized output for the k -th sensor becomes the following:

$$\varepsilon_k^{norm} = \frac{\varepsilon_k}{\sum_{i=1}^N \varepsilon_i / N} \quad (5)$$

4 FEASIBILITY STUDY ON CRACK CHARACTERIZATION ON SIMPLE SKIN

A simple FE model (figure 2) has been created by using ABAQUS STD V6.9 in order to represent an aluminum fuselage skin and to evaluate the performances of an ANN for crack detection and evaluation. The thickness of the skin has been set to 0.81mm, while the panel dimensions are 500mm x 600mm, as happened also in [2]. A quadratic mesh was set in order to better represent the non-linearity occurring near the cracks, thus allowing for a more precise description of the crack effect over the strain field. Finally, the panel was constrained in its lower edge, while the upper edge was loaded with a shell edge load, corresponding to a

totality of 2 tons, as [3]. The crack damage was modeled by means of mesh node release in randomly generated locations over the structure. In fact the damage has been parameterized in its center position, angle and length in order to automatically generate a database to be used for ANN training. In particular, damage characteristics are reported in table 1.

Table 1: Damage characteristics

x-position	Randomly generated
y-position	Randomly generated
Length	80mm (Fixed)
Angle (w.r.t. horiz. dir.)	0° (Fixed)

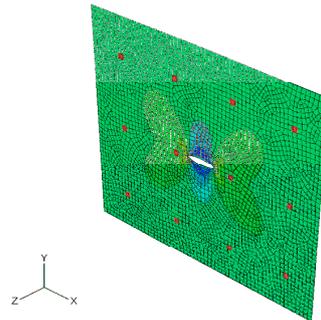


Figure 2: FE model of a fuselage skin; stringers are omitted in this work; nevertheless it is possible to clearly appreciate the effect of strain field modification due to the damage.

The sensor network has been designed with 12 sensors, arranged in a matrix with 3 rows and 4 columns, being the dimensions typical of a 4 bays panel [3]. The model has been solved many times by varying the crack center coordinates and always collecting strain information in correspondence of the sensor elements, shown in figure 2 in red. The sensor array has been chosen under the hypothesis that the optimum locations are with equal-distributed sensors over the skin surface, reasonable assumption for the regular structure and loading under study.

4.1 Crack detection and localization performances

A database of 350 cracks (80mm) was generated (figure 3) and the MATLAB “Pattern Recognition ANN” was used inside this framework. In particular, the neural network consisted of 12 inputs (strain simulated at sensor elements) and 1 binary target (to indicate whether the damage is present or not). The 70% of the samples were used to train the network, while the 15% for validation (check for generalization) and the remaining 15% for testing (the latter reported in figure 4). Being 350 damaged panels samples available, additional 350 undamaged samples were assigned as inputs, but with the addition of a random Gaussian noise (under the hypothesis that the variability introduced before is Gaussian). Thus, the test is about the algorithm capability to distinguish between the strain pattern (actually without additional noise) generated by a damaged panel and the one induced by an undamaged and noised panel. The performances of the ANN are evaluated for two different noise extents (5% - 10%), superposed to the undamaged samples and calculated with respect to the expected strain field.

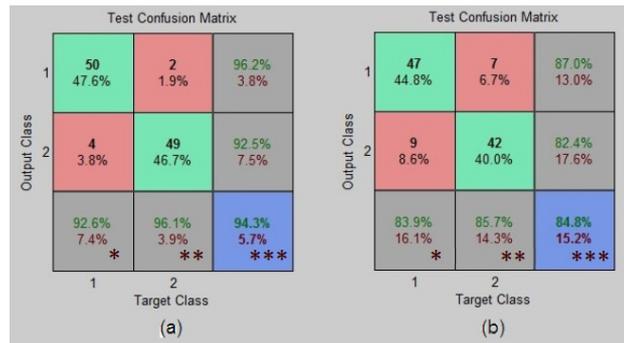


Figure 3: Confusion matrix (False Alarms and Missed Detection) for the case with 5% (a), 10% (b). Label 1 means “presence of a crack” while label 2 stands for “healthy panel”. *Probability of Missing Event (PME), **Probability of False Alarm (PFA), ***Probability of Wrong Assumptions (PWA).

In figure 3, the green squares represent the correct decisions while the red squares correspond to the wrong decisions; finally, the blue square reports the global performance of the system. Performances are evaluated by means of 3 parameters, PME (missed cracks divided by the totality of cracked cases), PFA (false alarm number divided by the number of alarms) and PWA (wrong assumptions divided by the totality of analysed cases). The ANN algorithm showed relatively good performances for the case with 5% (figure 3a) Gaussian random noise. However, when the noise increased beyond this limit, the performances got worse (10% in figure 3b). An attempt to improve the detection capability was carried out by increasing the number of available samples. It was clear that by increasing the number of samples the ANN was able to better fit the phenomenon, with a significant improvement in the global performance of the ANN. The results reported in figure 3 are obtained by consequently training the Neural Network with different hidden layer numbers and selecting the case with better performance

Once the damage has been detected, the step further consists in its localization. Inside this work frame, information on strain field damage induced modifications are extracted again from the FE model, in particular by sampling 350 cases for crack position in order to train the MATLAB “Fitting ANN”. In figure 5 the regression plot is reported relatively to the samples used for testing phase. One can notice that the localization capability appears to be acceptable. However this performance was obtained without the introduction of noise.

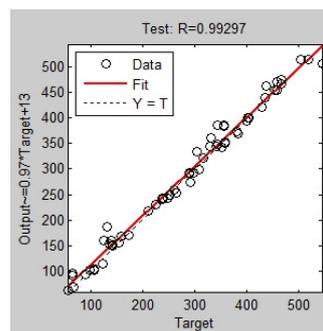


Figure 4: Localization performances for the ANN

5 CRACK CHARACTERIZATION ON STIFFENED RIVETED SKIN

Given the relatively good performances obtained for the simple skin model, the prosecution of the work is related to a more complex problem, that is to say the crack detection over the stiffened panel. The introduction of the stringers should reduce the sensitivity because of a decrease in the strain amount. In addition, also the stress and strain distribution would result more complex with respect to the simple skin model, as clearly noticeable by comparing figures 1 and 5.

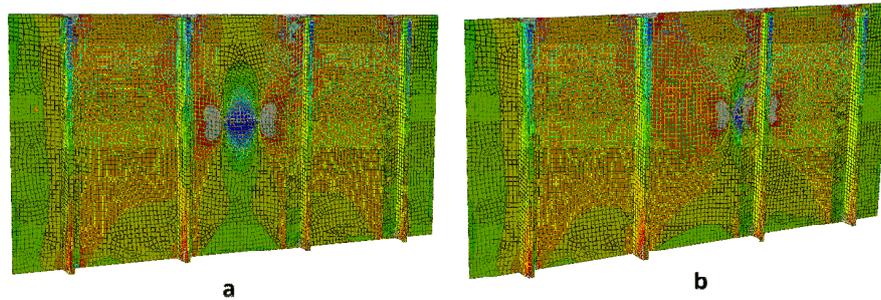


Figure 5: stress distribution in the direction parallel to stringer for two damaged panels with crack in the centre of a bay (a) and crack starting from a rivet (b)

Apart from the geometry, which was changed to a 600x300mm panel, the model description presented in Ch. 4 is valid for thicknesses and mesh. Rivet connections are modelled with ties in order to speed up the database generation for ANN training.

5.1 The problem definition

As reported in [12,13], five levels of damage identification can be distinguished:

- Level 1: (DETECTION) The method gives a qualitative indication that damage might be present in the structure.
- Level 2: (CHARACTERIZATION) The method estimates the type of damage occurring inside the structure.
- Level 3: (LOCALIZATION) The method also gives information about the probable position of the damage.
- Level 4: (QUANTIFICATION) The method gives an estimate of the extent of the damage.
- Level 5: (PROGNOSIS) The method offers information about the safety of the structure, e.g. estimates a residual life.

Inside this paper, the aim is to develop a methodology for SHM up to Level 4, thus allowing an extension to Level 5 in further works, on the basis of the results presented hereafter and, more in detail, in [13].

It is however necessary to make some assumptions and choices in order to proceed toward the best SHM system design, as reported in [13].

Firstly, the input and output of the ANN have to be decided. This is a critical phase, as many different design paths are allowed, but related with different performances. An

important choice is whether to use a single ANN that contemporaneously gives an alert with an estimation of the location and magnitude of the damage or to split the algorithm in more ANNs, each one related to a Level of damage identification.

Another critical point to be defined is the target crack length that the detection system has to identify. In fact, it could be worth to focus the ANN training on a certain crack length, critical according to aerospace structural design specifications. Actually it must be considered that one cannot expect the same performances of classical Non-Destructive-Evaluation (NDE) technologies mainly because of the completely different purpose. If, from one hand, NDE technologies are used by maintenance operators to find cracks or damages at their first evolution step over a structure that is not monitored on-board, from the other hand SHM aims at guaranteeing that the crack will be detected before it reaches a critical extent with respect to flight conditions, with a continuous on-board monitoring. Even more, the specification for damage tolerance of helicopter stiffened-skin structures suggest a design able to tolerate a 2-bay crack (about 250mm length), which means that detecting a 60mm target crack length could be a good result from a SHM point of view.

Last but not least comes the sensor number and positioning. For a detailed reporting on sensitivity analysis with respect to sensor number and performance comparison for sensor locations over the skin or on the stringers, the interested reader could refer to [13]. A fixed network configuration is used hereafter, defined according to the results presented in [2,13].

5.2 Noise quantification

Noise characterization is also a key factor for the improvement of ANN performances. Two ways to proceed can be highlighted. The first one exploits the classical formulation of ANN algorithm and consists in training the network with FE simulation data alone, optimizing the synapsis weights until the ANN starts to over-fit data, as described in Ch. 2. The limit in fitting is practically due to the fact that further optimization would bring the algorithm to learn also the noise inside the data, thus being too fitted for the training data and losing in terms of generality and performances when new samples are processed. The second approach consists in training the ANN with noised FE data, by adding a certain percentage noise to the FE output, which however will be considered as a mean value. By this way, the ANN recognises that the same output can be obtained with slightly different (noised) strain inputs. However, even in this case, the synapsis weight optimization will be carried out taking into account the problem of over-fitting.

Being some strain measurement repetitions for crack propagation in similar structures available from [2,3], according to some verification reported in [13], the noise percentage suggested for training and test of ANN is reported in table 2.

Table 2: Noise parameters for ANN training and testing

Purpose	Sensor Position	Noise Percentage
ANN training	Skin	10%
ANN training	Stringer	8%
ANN testing	Skin	10-15%
ANN testing	Stringer	5-10%

6 RESULTS AND CONSIDERATIONS

Some results are presented hereafter about the various information levels investigated with ANN algorithms. In particular, some neural networks have been designed considering results coming from [13], thus receiving as input the strain measurements collected by 6 equally distributed sensors for each stringer (24 sensors for the whole panel).

6.1 Level 1: Pattern Recognition for Anomaly Detection

Level 1 information has been investigated by means of a *Pattern Recognition ANN*, receiving 24 strain measures as input and giving a binary number as output, indicating whether the damage is present or not. The ANN has been trained with FE calculation (with additional noise as reported in table 2) for cracks in each rivet and cracks randomly positioned inside the bays, the latter without crossing the stringer. In particular the training set was executed with 60mm cracks, then evaluating the detection performances variation when a different crack length is given as input. In figure 6, Level 1 performances are evaluated through Probability of Missing Event, Probability of False Alarms and Probability of Wrong Assumption (wrong decisions with respect to totality of cases analysed in testing phase). Though the ANN was trained with just data relative to 60mm cracks and given the important noise defined in table 2, it is noticeable how the algorithm maintain a PoD of about 90% with 5% of PFA for 60mm cracks. Obviously, performances got worse for smaller cracks and better when larger cracks are tested.

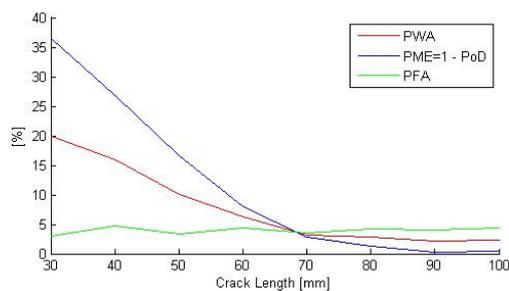


Figure 6: PWA, PME and PFA as a function of crack length (training was executed with 60mm cracks), acquiring from 6 sensors for each stringer

6.2 Level 2: Pattern Recognition for Damage Characterization

The next step consists in identifying the damage type occurring over the structure. It could be useful in case of dealing with many different damages to identify the type as soon as possible in the hierarchical level structure in order to best suit the next level ANN. For instance it could be applied to distinguish between a broken stringer or a skin crack. However, no FE data are available in the present work for the broken stringer effect. For that reason a *Pattern Recognition ANN* has been used trying to distinguish two very similar damages, that is to say skin bay cracks or skin rivet crack. Again, the ANN was trained for 60mm cracks and tested with different length cracks. In figure 7 it is clear that the influence of crack length in performances (PWA) is less influent. Moreover, though performances in damage

recognition seems to be low, a better behaviour is expected when really different damages are compared, as explained above.

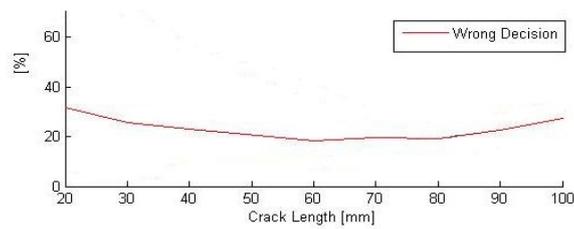


Figure 7: PWA as a function of crack length for damage characterization (training was executed with 60mm cracks), acquiring from 6 sensors for each stringer

6.3 Level 3: Damage localization and quantification

Once the anomaly has been identified (Level 1) and characterized (Level 2), the damage has to be localized. Two different *Function Fitting ANN* have been trained to localize a crack developed in a certain bay or starting from a rivet. Results are respectively reported in figures 8 and 9.

Concerning with skin bay cracks, ANN was trained with FE noised data relative to 60mm cracks, giving as input the strain measures and as output the $[x,y]$ estimation of the position inside the panel. It can be noticed how the distribution of ANN test output is centred on the target, unless the requested target is too near the boundaries of the controlled region. It could be due to a lack of knowledge about that zone, maybe requiring some additional training sets.

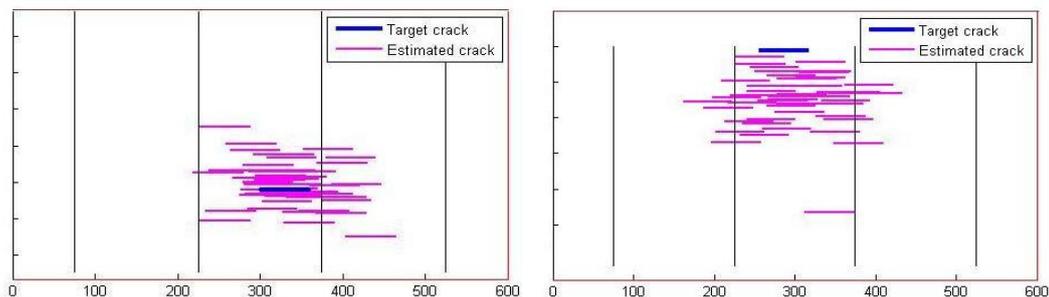


Figure 8: Crack localization for a bay damage, (training and testing were executed with 60mm cracks), acquiring from 6 sensors for each stringer

On the other hand, regarding skin rivet cracks, ANN was trained again with FE noised data relative to 60mm cracks, but giving as input the strain measures and as output the estimation of the rivet and stringer identification number. This method for output setting could seem more practically usable, however its relative performances appeared to be lower with respect to a $[x,y]$ output, shown before. The suggestion is thus to identify the rivet with respect to an $[x,y]$ coordinate system. In figure 9 the contour plots for crack position estimation have been reported. The problem undergone for the localization of the target in figure 9c was solved by increasing the number of sensors to 12. In fact, in figure 9d it is clear that the contour lines are focused on the right position, indicating a correct damage estimation.

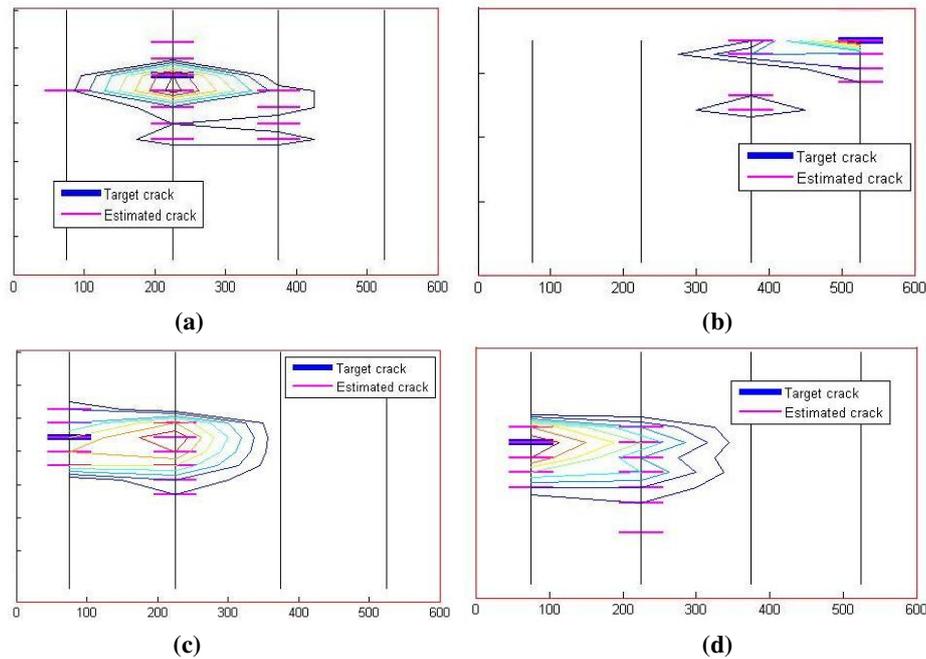


Figure 9: Localization for a rivet crack, (training and testing were executed with 60mm cracks), acquiring from 6 sensors for each stringer in (a,b,c) and 12 sensors for each stringer (d)

6.4 Level 4: Crack length quantification

The last step of structural diagnosis consists in damage quantification. In figure 10 an attempt is reported to acquire information regarding the length of the developing crack contemporaneously to localization. A bigger *Function Fitting ANN* as to be trained with cracks of different length and different position, demanding as output the estimation of crack length and position during the test phase. It can be noticed that the ANN was able to correctly estimate the crack length, while the accuracy of damage location got a bit worse than that presented in Ch. 6.4.

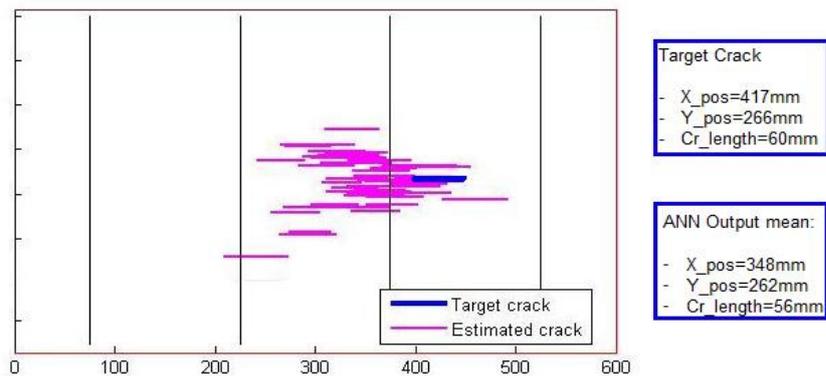


Figure 10: ANN for crack location and quantification (training was executed with crack ranging from 30mm to 100mm, testing with random crack length), acquiring from 6 sensors for each stringer

7 CONCLUSIONS

The feasibility of a structural health monitoring (SHM) system has been practically demonstrated. Good performances were obtained with a sensor network constituted by 6 sensors on each stringer of a skin stiffened aluminium panel. 90% of PoD associated with 5% of PFA is the limit obtained with such amount of sensors, however it could be by far improved allowing a more dense network, but always taking into consideration the practical installation and maintenance of such a complicate system. A 5 layer knowledge is suggested in order to organise the network. In practice, 4 layers are needed to infer about the actual state of the structure (diagnosis) while the last layer is necessary in order to do prognosis by exploiting the outputs of the diagnostic system.

REFERENCES

- [1] C.E. Katsikeros and G.N. Labeas, *Development and validation of a strain-based Structural Health Monitoring system*. Mechanical Systems and Signal Processing, Vol. 23/ 2, (2009).
- [2] C. Sbarufatti, A. Manes and M. Giglio, *Probability of detection and false alarms for metallic aerospace panel health monitoring*. Proc. 7th Int. Conf. on CM & MFPT, BINDT, (2010).
- [3] M. Giglio and A. Manes, *Crack propagation on helicopter panel: experimental test and analysis*. Engineering fracture mechanics, Vol. 75, pp. 866-879, (2008).
- [4] S. Hideki, F. Shin-Etsu, O. Tomonaga, T. Nobuo and Y. Jr. Toshimitsu, *Structural health monitoring of cracked aircraft panels repaired with bonded patches using fiber bragg grating sensors*. Source: Applied Composite Materials, Vol. 13/2, (2006).
- [5] T. Mitchell, *Machine Learning*, McGraw Hill, (1997)
- [6] C. Bishop, *Neural Networks and Pattern Recognition*, Oxford University Press, (1995)
- [7] I.Y. Hoballah and P.K. Varshney, *Distributed Bayesian Signal Detection*, *IEEE transaction on information theory*, Vol. 35, Nr. 5, (1989)
- [8] B. Chen and P.K. Varshney, *A Bayesian Sampling Approach to Decision Fusion Using Hierarchical Models*, *IEEE transactions on signal processing*, Vol. 50, Nr. 8, (2002)
- [9] J. Guan, Z. Shen, J. Lu and X. Ni, *Surface crack detection for Al plate using the surface acoustic waves and neural network identification*. Proc. SPIE 5629, 466 (2005)
- [10] Y. Lu, L. Ye, Z. Su, L. Zhou and L. Cheng, *Artificial Neural Network (ANN)-based Crack Identification in Aluminum Plates with Lamb Wave Signals*. *Journal of Intelligent Material Systems and Structures* January vol. 20 no. 1 39-49 (2009)
- [11] *Health Monitoring for Airframe Structural Characterization*, NASA/CR-2002-211428.
- [12] K. Worden and A.P. Burrows, *Optimal sensor placement for fault detection*, *Engineering Structures* 23 885–901(2001).
- [13] C. Sbarufatti, A. Manes and M. Giglio, *Sensor network optimization for damage detection on aluminum stiffened helicopter panels*, Proc. Coupled Problems in Science and Engineering (Coupled 2011)

COMPUTATIONAL STUDY OF THE INTERACTION BETWEEN A NEWTONIAN FLUID AND A CELLULAR BIOLOGICAL MEDIUM IN A STRAIGHT VESSEL

TERPSICHORI S. ALEXIOU^{*†}, GEORGE E. KAPELLOS^{*}, STAVROS PAVLOU^{*†}

^{*}Department of Chemical Engineering
University of Patras
Karatheodori Str. 1, GR-26504 Patras, Greece
e-mail: xalexiou@chemeng.upatras.gr, gek222@chemeng.upatras.gr, sp@chemeng.upatras.gr
www.chemeng.upatras.gr

[†]Institute of Chemical Engineering and High Temperature Chemical Processes (ICEHT)
Stadiou Str., Platani, GR-26504, Patras, Greece
www.iceht.forth.gr

Key words: Fluid-Biomaterial Interactions, Navier-Stokes equations, Mixture Theory, Finite Element Method.

Abstract. In this work, we solve numerically the governing equations for quasi-steady Newtonian flow past and through a cellular biological medium, which is attached to the surface of a straight vessel. The flow past the cellular biological medium is described by the Navier-Stokes equations. For the modeling of momentum transfer within the cellular biological medium, we consider that the cellular biological medium constitutes a biphasic fluid-solid mixture with poroelastic behaviour. The system of governing equations is solved numerically with the mixed finite element method. The computational domain is discretized using an unstructured, variable density triangular element mesh. From the numerical solution we obtain the spatial distributions of: (i) the fluid velocity and pressure, and (ii) the displacement and stresses of the solid matrix within the cellular biological medium. Also, the components of the overall hydrodynamic force exerted by the flowing fluid on the cellular biological medium are calculated. A parametric analysis is performed with regard to the Reynolds and Darcy numbers that characterize the flow past and through the cellular biological medium.

1 INTRODUCTION

The interaction between a flowing fluid and a cellular biological medium (e.g. biofilm, or tissue) attached to the surface of a vessel is of key importance in several natural phenomena, and processes of physiological and technological significance. The biodegradation of organic contaminants by microbial biofilms in soil and aquifers, the in vitro construction of artificial tissues from human stem cells in synthetic porous scaffolds, the flow of blood in vessels containing thrombus and atheromatous plaque formations, are some of the processes in which fluid-cellular biological medium interactions are important.

In all the aforementioned processes, the interaction between the flowing extracellular fluid and the deformable solid matrix (cells plus extracellular matrix) plays multiple important roles: (a) affects the internal architecture (spatial arrangement of cells and extracellular matrix) and the external morphology (overall size and shape) of the cellular biological medium [1], (b) enhances the mass transfer rate of chemical species (nutrients, wastes, chemical signaling molecules, etc) within the cellular biological medium [2], and (c) regulates the function of cells through the action of mechanical stresses, which are either applied directly on the outer surface of the biological cell, or transmitted indirectly through the extracellular polymeric matrix [3].

Significant research effort has been directed toward the elucidation of the exact mechanisms through which the complex interaction between a flowing fluid and a cellular biological medium occurs. Along this direction, it is standard practice to employ experimental techniques that allow the study of the effect of a well-controlled flow field on the function and morphology of individual biological cells, layers of cells, or cellular biological media samples. In particular, flow chambers with parallel plate configuration have been used in several studies to observe the response of cellular biological media to fluid shear.

Mathematical modelling is an indispensable tool, which is complementary to experimental investigation and provides qualitative interpretation and quantitative correlation of the fluid-structure interactions in these systems in terms of velocity, stress fields, etc. A comprehensive survey on the available approaches for theoretical modelling of momentum transport in cellular biological media, with focus on the formulation of governing equations and the calculation of material properties, is given in [4].

In this work, we present preliminary results from computer simulations of flow past and through a poroelastic biomaterial, which is attached to the surface of a straight vessel. Fluid flow in the clear fluid regions is described by the Navier-Stokes equations, and momentum transfer in the cellular biological medium is described in the context of biphasic mixture theory. The effect of the Reynolds and Darcy numbers that characterize the flow past and through the biological medium is investigated.

2 MATHEMATICAL MODEL

We consider the quasi-steady flow of an incompressible Newtonian fluid in a straight vessel with rigid solid walls, and a cellular biological medium attached to the wall of the vessel as shown in Figure 1. In the clear fluid region, which is denoted by Ω_f , fluid flow is described by the Navier-Stokes and continuity equations

$$\nabla \cdot \mathbf{v}_f = 0 \quad \text{in } \Omega_f \quad (1)$$

$$\rho_f \mathbf{v}_f \cdot \nabla \mathbf{v}_f = -\nabla P_f + \mu_f \nabla^2 \mathbf{v}_f \quad \text{in } \Omega_f \quad (2)$$

where \mathbf{v}_f is the velocity, P_f is the pressure, ρ_f is the density, and μ_f is the viscosity of the fluid. The flow in the vessel is driven by a prescribed, constant pressure drop ΔP_{ref} between the inflow and outflow boundaries of the vessel. Further, the gradient of the velocity components along the flow direction (x-axis) is nil at the inflow and outflow boundaries because the flow is considered to be fully developed there. At the fluid-solid interface the fluid velocity is nil based on the no-slip and no-penetration assumptions.

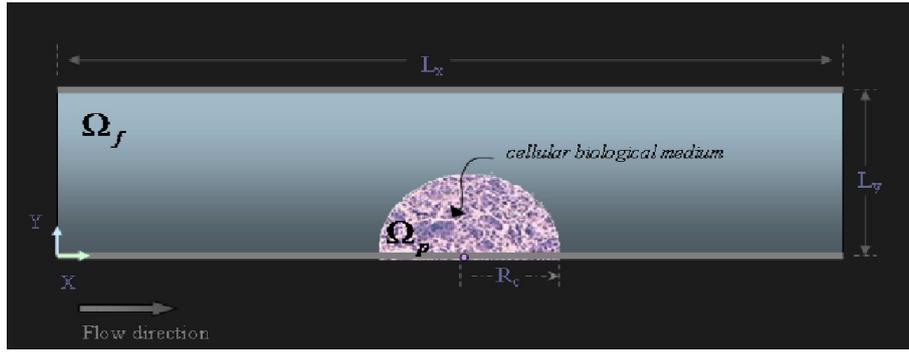


Figure 1: Schematic representation of the flow geometry

The cellular biological medium is treated as a biphasic mixture consisting of: (a) a viscous fluid phase denoted by f (extracellular aqueous solution), and (b) an elastic solid phase denoted by s (cells and extracellular matrix). The mathematical formulation of the mass and momentum balances for the cellular biological medium is based on the theory of mixtures, which was developed for the mathematical description of momentum, mass, and energy transport in multicomponent, multiphase systems at a scale of observation with resolution much larger than the characteristic length of the individual constituents and phases (the structure of the system at finer spatial scales is disregarded completely). The mathematical foundation of the theory relies upon the concept that the constituents of the mixture can be modeled as superimposed, interacting continua. In this way, a material point is assigned for each constituent at every point in space, which is occupied by the mixture. Exhaustive reviews on the theory of mixtures are given in [5,6]. The governing conservation laws for each constituent of the mixture contain the usual terms (that appear in formulations for a single phase), and an additional term which accounts for the interaction between the reference constituent and the other constituents of the mixture.

We consider that the fluid (f -constituent) behaves as a Newtonian fluid, and the solid (s -constituent) behaves as a linearly elastic isotropic solid. For the systems under consideration, it is reasonable to neglect the convective momentum transfer and mass production–consumption (for time scales of observation much smaller than the characteristic time of cell division). Furthermore, we assume that the flow is in quasi-steady state. Under these assumptions, in the context of biphasic mixture theory, the equations that govern momentum transfer in the cellular biological medium are

$$\nabla \cdot \mathbf{v}_f = 0 \quad \text{in } \Omega_p \quad (3)$$

$$\mathbf{0} = \nabla \cdot \boldsymbol{\sigma}_f + \mathbf{F}_{s \rightarrow f} \quad \text{in } \Omega_p \quad (4)$$

$$\mathbf{0} = \nabla \cdot \boldsymbol{\sigma}_s + \mathbf{F}_{f \rightarrow s} \quad \text{in } \Omega_p \quad (5)$$

The stress tensors and the interaction force are given by the following constitutive expressions

$$\boldsymbol{\sigma}_f = -\phi_f P_f \mathbf{I} + \mu_f \left[\nabla \mathbf{v}_f + (\nabla \mathbf{v}_f)^T \right] \quad (6)$$

$$\boldsymbol{\sigma}_s = -\phi_s P_f \mathbf{I} + \lambda_s (\nabla \cdot \mathbf{u}_s) \mathbf{I} + \mu_s \left[\nabla \mathbf{u}_s + (\nabla \mathbf{u}_s)^T \right] \quad (7)$$

$$\mathbf{F}_{f \rightarrow s} = -\mathbf{F}_{s \rightarrow f} = \frac{\mu_f \phi_f}{k_{eff}} \mathbf{v}_f + P_f \nabla \phi_s \quad (8)$$

Here, ϕ_a is the volume fraction of the a th constituent in the cellular biological medium (with $\phi_f + \phi_s = 1$), \mathbf{u}_s is the solid displacement, μ_s and λ_s are the Lamé parameters for the solid, and k_{eff} is the hydraulic permeability of the material.

The mathematical formulation is completed with the boundary conditions at the interface between the cellular biological medium and the free fluid. We consider continuity of the fluid velocity and, further, that the normal stress exerted on the interface by each constituent of the cellular biological medium equals the total normal stress weighted by the corresponding volume fraction [7].

3 NUMERICAL SOLUTION

For the numerical solution of the governing equations of the problem we used the Galerkin finite element method. The spatial discretization of the computational domain is performed with triangular elements. An unstructured, variable density triangular element mesh is generated by combining a Delaunay triangulation algorithm with a force-based mesh smoothing methodology, so as to optimize the mesh quality. A typical mesh used for the numerical calculations is shown in Figure 2.

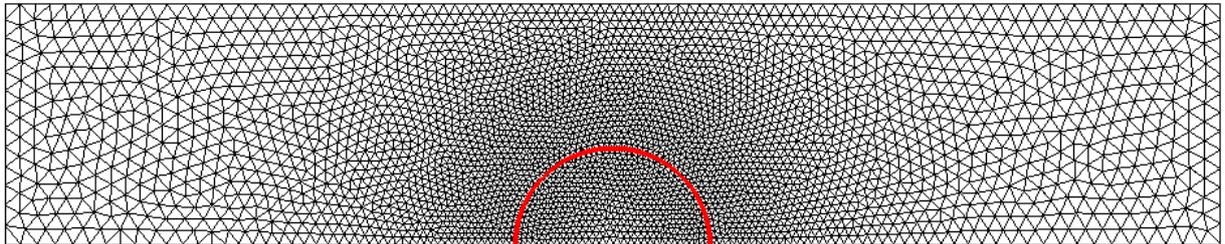


Figure 2: Typical unstructured mesh of triangular elements.

For the approximation of the fluid pressure we use 3-node linear basis functions, while the fluid velocities and solid displacements are approximated by 6-node quadratic basis functions. For the solution of the non-linear systems of equations that result from the discretization of the fluid momentum balances we employ a combination of Picard iterations and the Newton-Raphson method. Due to the assumptions postulated in section 2, the governing equations and boundary conditions of the problem constitute a one-way coupled system, so that the fluid and solid constituents momentum balances can be sequentially solved.

4 RESULTS AND DISCUSSION

In this section, we present preliminary results regarding the effects of the *Reynolds* and *Darcy* numbers that characterize the flow *past* and *through* the cellular biological medium, respectively. These dimensionless numbers are defined as follows

$$Re = \frac{\rho_f \langle U \rangle R_c}{\mu_f} \quad (9)$$

$$Da = \frac{k_{eff}}{R_c^2} \quad (10)$$

where $\langle U \rangle$ is the magnitude of the average velocity for flow in a clear vessel, under the same pressure drop. All other parameters of the system are held constant for the simulations. The size ratio of the radius of the semicircular obstacle to the width of the vessel is equal to 0.4, the length to width ratio of the vessel is equal to 5.0, and the Young's modulus and Poisson's ratio for the solid are set equal to 1MPa, and 0.45. Figure 3 shows the fluid streamlines in the vessel for three different values of the Reynolds number (0, 1, and 10) and two different values of the Darcy number (10^{-2} and 10^{-4}). We observe that as the Re number increases, the flow pattern past and through the obstacle changes qualitatively and, gradually, obtains an asymmetrical structure. This effect becomes more pronounced for the lower Da number, where an extended recirculation region is formed behind the obstacle. Interestingly, the recirculation zone extends into the permeable biological medium (Figure 3F). Furthermore, we observe that as the Da number decreases, the fluid streamlines begin to divert significantly from the permeable obstacle, meaning that the amount of fluid passing through the cellular biological medium undergoes noticeable decrease.

Figure 4 shows the corresponding maps of longitudinal and lateral displacement of the solid in the cellular biological medium. An increase in the Re number affects the solid displacement field both quantitatively and qualitatively. The values of the horizontal and vertical displacement increase in response to the elevated values of the forces experienced by the solid due to fluid flow (namely, term $\mathbf{F}_{f \rightarrow s}$ in the momentum balance for the solid). In addition, the gradual loss of symmetry of the flow affects strongly the solid displacements (see for example Figure 4F). The Da number also has an important effect on the displacements. A decrease in the value of Da leads to an increase of the solid displacements. This is attributed to the increase of the fluid-solid interaction force, $\mathbf{F}_{f \rightarrow s}$, caused by the decrease of the permeability of the medium.

In Figure 5, the effect of Re on the dimensionless drag and lift forces exerted by the fluid on the cellular biological medium is depicted for two different Darcy numbers, namely $Da=10^{-2}$ and $Da=10^{-4}$. The dimensionless drag and lift forces (per unit length) are defined by the following expressions:

$$F_{drag} = \frac{1}{L_y \Delta P_{ref}} \oint_S \mathbf{e}_x \cdot \boldsymbol{\sigma}_f \cdot \mathbf{n} dS \quad (11)$$

$$F_{lift} = \frac{1}{L_y \Delta P_{ref}} \oint_S \mathbf{e}_y \cdot \boldsymbol{\sigma}_f \cdot \mathbf{n} dS \quad (12)$$

Here, L_y is the width of the vessel, and ΔP_{ref} is a reference pressure drop.

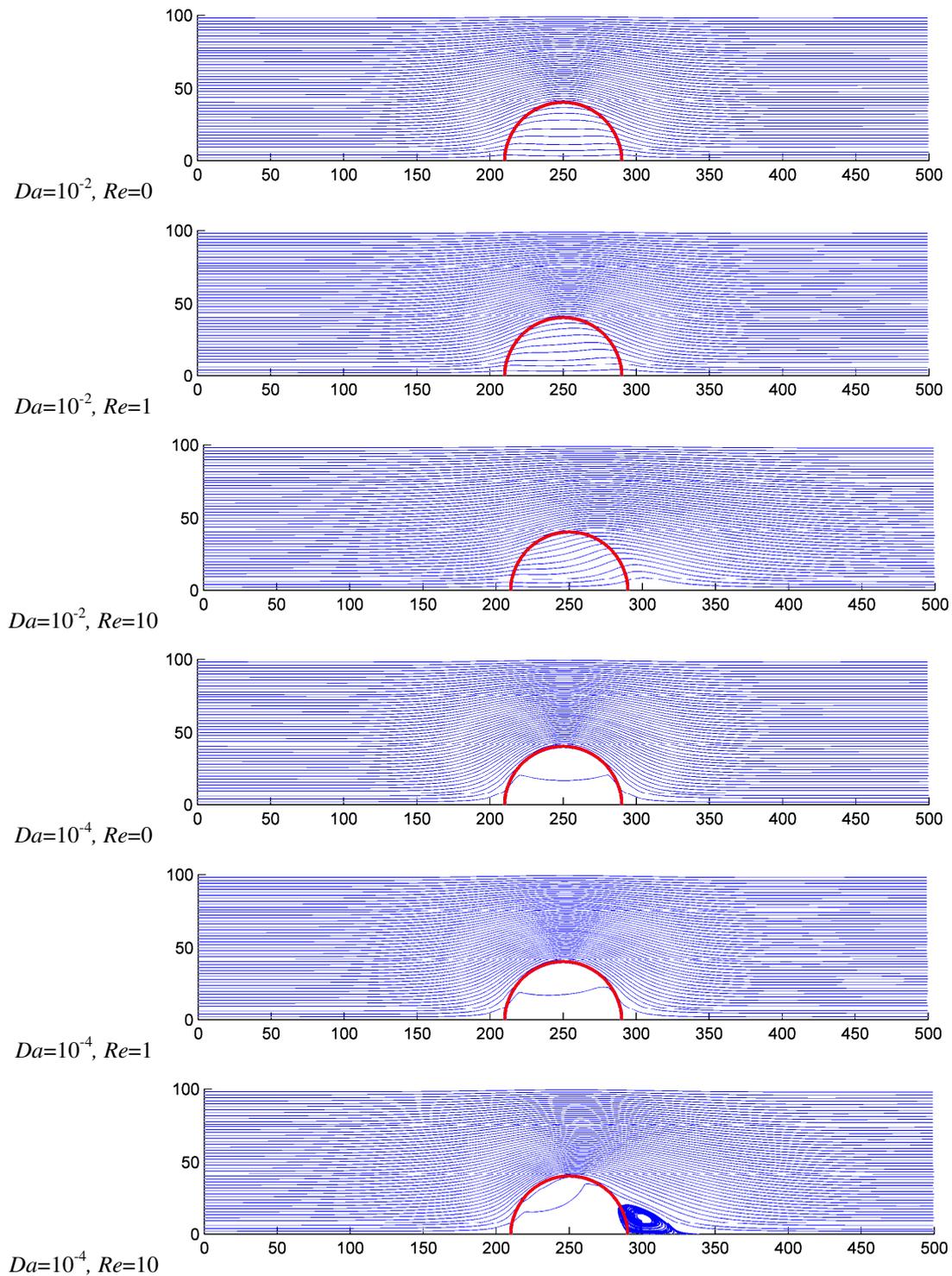


Figure 3: Fluid streamlines for representative values of the Reynolds (Re) and Darcy (Da) numbers.

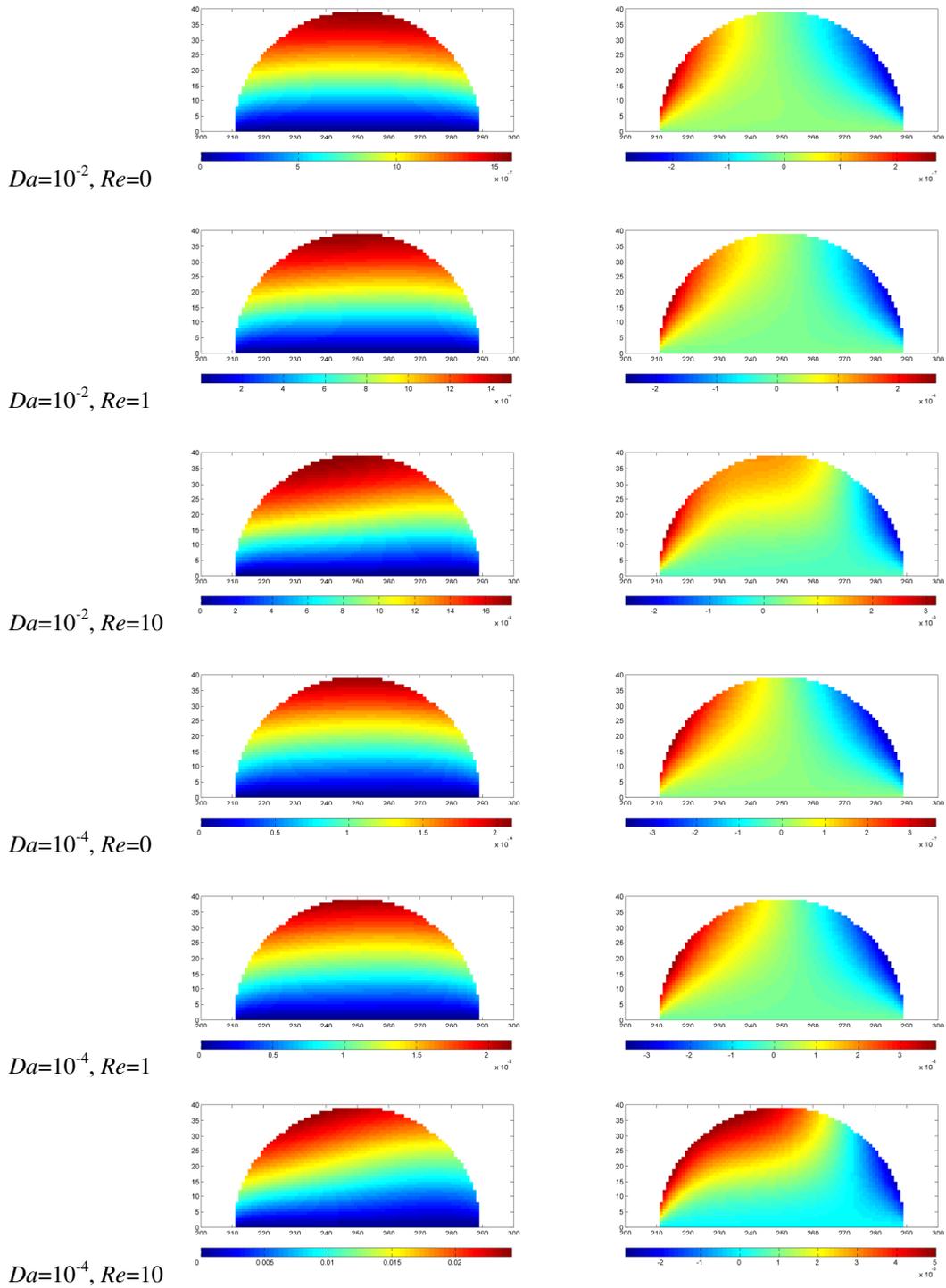


Figure 4: Maps of the longitudinal (left column) and lateral (right column) displacements of the solid in the cellular biological medium for representative values of the Reynolds and Darcy numbers.

In Figure 5, we observe that an increase in the Re number is followed by an increase in the values of both the drag and lift forces, as expected. Interestingly, an increase in the Da number leads to a decrease in the value of drag and an increase in the value of lift.

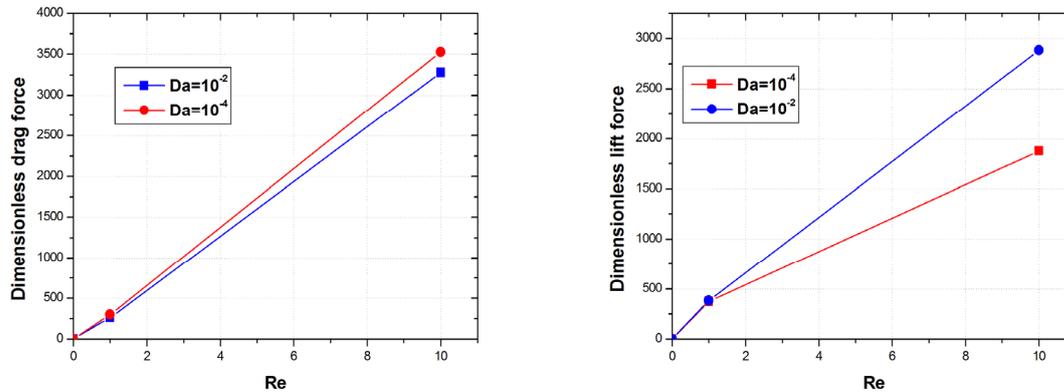


Figure 5: Effect of Re on the dimensionless drag and lift forces for two different Da numbers.

5 CONCLUSIONS

- As the Reynolds number increases and the Darcy number decreases, an extended downstream recirculation region forms which reaches into the cellular biological medium.
- The drag force exerted by the flow on the cellular biological medium increases for increasing Reynolds number and decreasing Darcy number, while the lift force increases for increasing Reynolds number and increasing Darcy number.

REFERENCES

- [1] Ng, C.P. and Swartz, M.A. Mechanisms of interstitial flow-induced remodeling of fibroblast-collagen cultures. *Ann. Biomed. Eng.* (2006) 34(3): 446-454.
- [2] Evans, R.C. and Quinn, T.M. Solute convection in dynamically compressed cartilage. *J. Biomech.* (2006) 39:1048-1055.
- [3] Rutkowski, J.M. and Swartz, M.A. A driving force for change: interstitial flow as a morphoregulator. *Trends Cell Biol.* (2007) 17(1):44-50.
- [4] Kapellos, G.E., Alexiou, T.S. and Payatakes A.C. Theoretical modeling of fluid flow through cellular biological media: An overview. *Math. Biosci.* (2010) 225(2):83-93.
- [5] Bedford, A. and Drumheller D.S. Theories of immiscible and structured mixtures. *Int. J. Eng. Sci.* (1983) 21(8): 863-960.
- [6] Rajagopal, K.R. and Tao L. *Mechanics of Mixtures*. World Scientific (1995).
- [7] Hou J.S., Holmes M.H., LAi W.M., and Mow VC. Boundary conditions at the cartilage-synovial fluid interface for joint lubrication and theoretical verifications. *J Biomech Eng.* (1989) 111(1): 78-87.

CRACK PROPAGATION IN CIVIL ENGINEERING BRIDGE CABLES: COUPLED PHENOMENA OF FATIGUE AND CORROSION

V. PÉRIER^{*}, L. DIENG[†] AND L. GAILLET[†]

^{*} IFSTTAR (*French Institute of Science and Technology for Transport, Development and Networks*)

Bridges and Structures Department

Route de Bouaye – CS4, 44341 Bouguenais, France

e-mail: virginie.perier@ifsttar.fr, web page: <http://www.ifsttar.fr>

Key words: Bridge cables, corrosion, fatigue, fretting-fatigue, steel wires.

Abstract. This paper deals with the propagation of a crack initiated by fretting-fatigue in bridge cables. They are submitted to traffic loads and exposed to environmental conditions (rain, wind, marine environment, de-icing salts ...). These loads can lead to two main damage causes: fretting-fatigue and corrosion. These both phenomena can be coupled and lead to premature failures of drawn steel wires. In fact, at specific contact points, cracks are often initiated by fretting-fatigue. Fretting-fatigue is due to the cable bending. Near anchorages, the bending deformations are the highest and can generate damage (wire cracks). Cracks may propagate under traffic or wind fatigue load. They are also often submitted to a severe environment. Indeed, even if cables are initially protected, these protections can deteriorate and then cables are subject to corrosion. Experimental results show that wire cracks are often initiated by fretting-fatigue and crack propagation is due to fatigue or fatigue corrosion. In this study wire crack is initiated by fretting-fatigue first and fatigue tests are set up after with several environments in order to evaluate their influence on the crack propagation. The studied corrosive environments are: distilled water, sodium chloride solution and thiocyanate ammonium solutions with different concentrations. Lifetimes of specimens are compared. Metallographic and fractographic observations show the influence of the different solutions on the crack propagation. Indeed, since the microstructure of the steel wires is fully oriented in the longitudinal direction, corrosion changes the crack propagation modes and favours mixed mode propagation or longitudinal propagation.

1 INTRODUCTION

Wires of civil engineering cables are submitted to two main damage mechanisms: fretting-fatigue and corrosion. Fretting-fatigue is generally observed near the anchorage where the cable is submitted to the most important free bending deformations. Usually, wires are protected against corrosion in several ways such as lubrication, zinc coating, etc. However, if these means of protection disappear during the cable life, and if corrosive solution is in contact with the cable, then the wires would be submitted simultaneously to fretting-fatigue and corrosion. The wires of cables are generally elaborated by cold drawing. Those wires are cylindrical with a diameter ranging between 3 to 7 millimetres. They have a high-grade of

carbon (about 0.8%) and are obtained by cold drawing. These cable wires have a high tensile strength and a low ductility.

The different types of cables are described in [1] and [2]. Due to cable geometry, inter-wires contacts are subject to stress concentration which can lead to crack initiation. Indeed, stay cables are submitted to vibrations induced by climatic loads (wind, rain, etc.) and traffic loads. These vibrations lead to bending deformations which are maximal near the anchorages and then there are small relative displacements between wires (fretting-fatigue).

Studies of fretting-fatigue phenomena were already conducted by several authors (Fouvry [5], Siegert [4], Zhou [6], etc.). They concluded that the fatigue limit is highly reduced for specimens undergoing fretting-fatigue. In the case of cables [4], this reduction is mostly due to contact friction forces between wires. In order to study this interwire contact mechanism, Siegert [6] determined normal contact force in a multilayer strand and the relative displacement amplitude between wires. Then, a fretting-fatigue device was developed; it aims at reproducing the contact fatigue conditions in spiral strands undergoing free bending deformation.

Aging cables are also subjected to severe environmental conditions and can undergo corrosion damage. There are three main types of corrosion for cables: uniform corrosion, localized corrosion and mechanically assisted corrosion as stress corrosion cracking or fatigue-corrosion. In stay cables with low tension, corrosion appears generally only after a certain age. However, for aging bridge cables where protective coatings are damaged, corrosion is an important source of cables degradation [2].

A previous study aimed at investigating the influence of a sodium chloride (NaCl) solution on the behaviour of wires undergoing fretting-fatigue [1]. In this paper, coupled phenomena between fatigue and corrosion are studied in wires of bridge cables. After initiating cracks in a wire by fretting-fatigue, their propagations are realized by fatigue-corrosion in several environments.

2 MATERIAL

The material used for the tests is mechanically a high strength steel and chemically a low alloy steel.

Table 1 : Mechanical properties and characteristics of test wires and strands

	E (MPa)	UTS (MPa)	σ_{Rg} (MPa)	$\sigma_{RP0,02}$ (MPa)	ϵ (%)
Wire	202 000	2 020	1 860	1 750	5,5
Strand	193 000	1 940	1 860	1 480	5,2

Those wires are manufactured by cold drawing process, which consists in reducing the wire section by passing through decreasing section dies. This process improves the wire resistance by hardening. The mechanical properties and characteristics of a wire and of a strand are reported in Table 1. They were obtained by carrying out several tension tests. Figure 1 shows the stress-strain curves obtained from these tests.

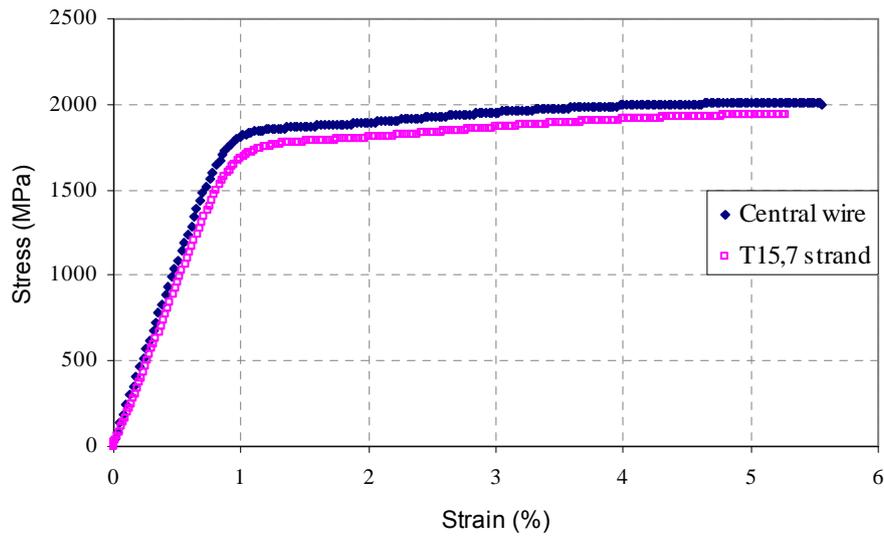


Figure 1: Tensile test on a wire and on a strand

Tested specimens are low alloy steel wires whose chemical composition is given in Table 2. They are cylindrical with a diameter of about 5.4 millimetres. They exhibit a pearlitic structure with fully oriented grains in the longitudinal direction.

Table 2: Chemical composition of cold drawn wires (main elements)

Fe	C (carbon)	Si (silicon)	Mn (manganese)	S (sulphur)	P (phosphor)
Balance	0,8 %	0,23 %	0,52 %	0,018 %	0,017 %

3 EXPERIMENTAL TEST DEVICE

For the tests, a fatigue machine is used: it is associated with a fretting grip for fretting-fatigue tests and with corrosive cells for fatigue-corrosion tests.

The fretting-fatigue testing device is shown in Figure 2. This experimental device is used to reproduce the loading conditions of wires in a spiral strand like the free bending of stay-cables. The tested specimens consist of steel wires which are subjected to a tensile fatigue loading with a mean stress value of 600 MPa, corresponding approximately to 30% of the wire UTS (1860 MPa), and a stress variation $\Delta\sigma$. The constant normal contact force F_c is added perpendicularly to the cable length, using two pairs of pads made up with the same material that crossed perpendicularly the tested specimen. The two pairs of pads were fixed in a rigid way at a distance L on a pair of supports. δ is the amplitude of the relative displacement between the pads and the tested wire.

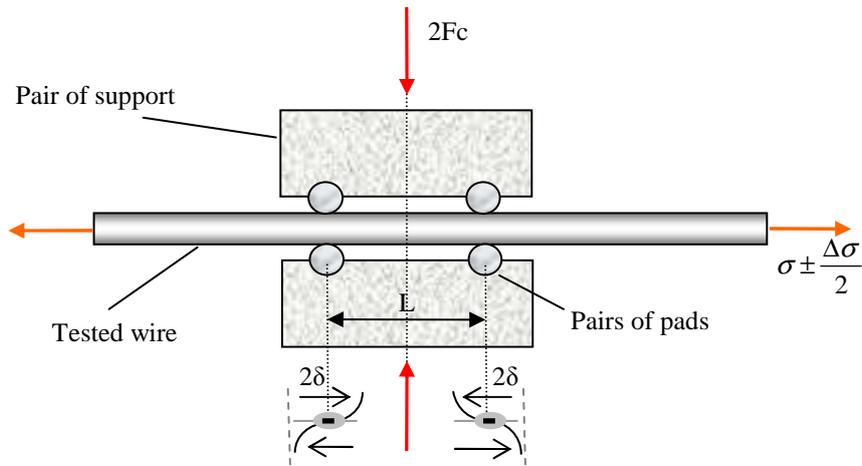


Figure 2: Schematic of the fretting-fatigue test set-up (side view)

The modelling of the mechanical behaviour of a strand subjected to bending deformations makes it possible to calculate the local contact between wires under simplified conditions (Siegert [7]). Moreover one first experimental study made possible to identify value ranges of the test parameters representative of the inter wire contacts in the strand which lead to an important reduction of the stress limit for bare wires (not galvanized and not lubricated). The tests were carried out with a contact force of 200 N, a distance between pairs of pads of 20 mm and a variation of tensile stress of 200 MPa at a frequency of 2 Hz.

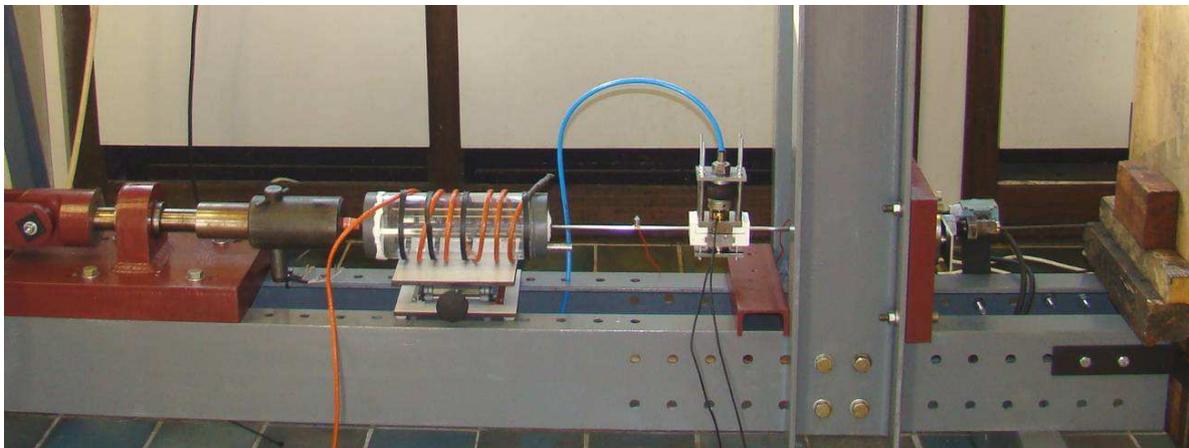


Figure 3 : Experimental device: fretting-fatigue test under realization and a corrosive cell which is translated after the fretting-fatigue

For the fatigue-corrosion test, the fretting grip is removed and a cell, which contains a corrosive solution at 50 °C is translated on the contact scars. The fatigue parameters are the same as for the fretting-fatigue test (mean stress of 600 MPa, stress variation of 200 MPa), excepted for the frequency which is reduced to 1 Hz (Figure 3).

So, the tests consisted in two steps. The first step aims at initiating a crack by fretting-

fatigue. The second step consists in propagating the crack by fatigue-corrosion.

For the first step, 350 000 cycles of fretting-fatigue are sufficient to initiate cracks. The presence of these cracks is controlled with an ultrasonic device.

For the second step, the pre-cracked tested wire is submitted to fatigue-corrosion, and corrosion cells are added. For tests in distilled water or sodium chloride (NaCl) solutions, two cells are added, one cell is placed around the tested wire, and there is another cell in which the solution is aerated. For tests in thiocyanate ammonium (NH_4SCN) solutions, only one cell is added around the wire.

4 RESULTS AND DISCUSSION

The main results of the tests are the lifetimes of the tested specimens and their metallographic observations. The results are compared for the different studied environments: air, distilled water, NaCl solution and NH_4SCN solutions.

Lifetimes are given on the histogram of the figure 4. The pink part of the bars represents the fretting-fatigue phase (350 000 cycles). The blue part corresponds to the fatigue phase in each environment. The dispersion of the results is also reported.

For tests in air, the average lifetime is about 600 000 cycles, and it can be seen a large dispersion of the results, this dispersion is inherent of fretting-fatigue phenomena.

In distilled water, the average lifetime is about 750 000 cycles and in NaCl solution, the mean lifetime is about 800 000 cycles.

In NH_4SCN solutions, the average lifetime is between 370 000 and 500 000 cycles depending on the concentration.

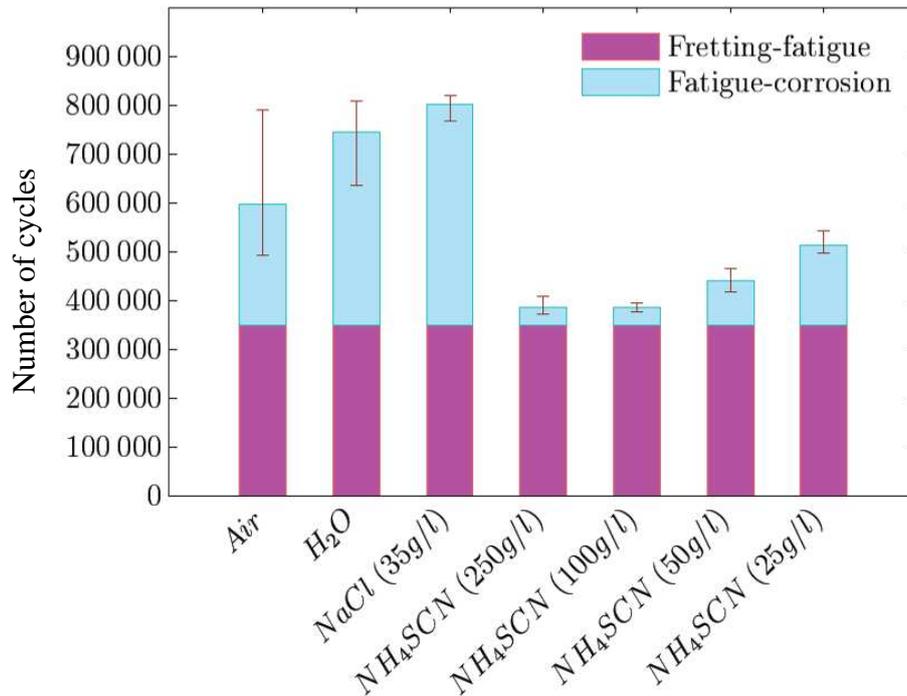
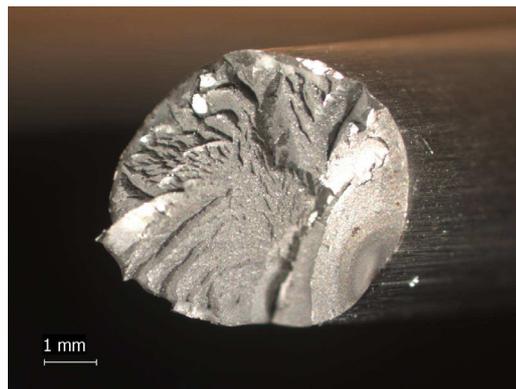


Figure 4 : Lifetimes of specimens submitted to fretting-fatigue first and then to fatigue in several environments

In NH_4SCN solutions, the lifetimes are reduced compared to air, more significantly for the concentrations of 100 and 250 g/l. This result means that hydrogen accelerate the crack propagation by a Hydrogen-Induced Stress Corrosions cracking (HI-SCC) mechanism.

On the contrary, in distilled water or in NaCl solution, lifetimes increase. These solutions seem to reduce the crack propagation. These results are obtained in the retained conditions and may not be representative of what happens on bridges.

However, some authors made similar observations. For example, Pao & *al.* [8] noticed that for low values of the variation of the stress intensity factor ΔK , the crack growth rate in an aluminum alloy can be slower in a NaCl solution than in air.



(a) Air



(b) Distilled Water



(c) NaCl solution



(d) NH_4SCN solution – 50g/l



(e) NH_4SCN solution – 100g/l

Figure 5: Fractures surfaces of specimens submitted to 350 000 cycles of fretting-fatigue and to fatigue until rupture in different environments

For each type of tests, the specimen rupture surfaces were observed with binocular, optical and/or scanning electron microscopes.

The fracture surfaces for tests in air are quite smooth in the crack area (Figure 5 (a)). The crack initiation and the beginning of its propagation are influenced by the fretting loading and make an angle of about 45°. The crack propagation is then due to the fatigue loading and is perpendicular to the wire until the final shear rupture.

For the tests in distilled water or in NaCl solution, the crack propagates transversally to the wire first and then longitudinally (Figure 5 (b) and (c)).

For tests in NH₄SCN solutions (Figure 5 (d) and (e)), after the transversal part of the crack, the crack propagates in a mixed mode (modes I and II).

The increase in lifetime of specimens immersed in distilled water or NaCl solution may be explained by the changes in the propagation of the crack. Indeed, there is a very long longitudinal propagation which is not observed for specimens undergoing fatigue in air. Moreover, some authors [8],[9] also made similar observations for aluminum alloys and they made hypothesis that may explain this effect. Three hypotheses are given by Menan [9]. The first one consists in the blunting of the crack tip which results from a competition between mechanical kinetic and anodic dissolution kinetic under the action of chloride ions. The second one supposes crack closure effects which might be due to corrosion products as said by Pao *et al.* [8]. The third one is the passivation of the crack tip which would protect the steel and then slow down the crack growth rate.

5 CONCLUSIONS

Cables of bridges have two main damage causes: fatigue and corrosion. In this paper, coupled phenomena are studied. Indeed cracks are initiated by fretting-fatigue and then propagate under fatigue-corrosion. Several environments are studied and compared. Air is chosen as a reference environment. An increase of lifetime is observed in distilled water or in NaCl solutions. This increase may be due to changes in the propagation of cracks in these environments (changes of modes, of crack growth rate, etc.). Indeed, there is a very long longitudinal propagation (mode II) which is not observed in air, and chloride ions may induce crack tip blunting. A reduction of lifetime is observed in the more aggressive corrosion solutions, it is probably due to hydrogen embrittlement.

The study of interaction between these phenomena may contribute to improve the management of bridges which are getting older. This study might be completed by pure fatigue-corrosion studies with different level of stress.

REFERENCES

- [1] V. Périer, L. Dieng, L. Gaillet, C. Tessier and S. Fouvry. Fretting-fatigue behaviour of bridge engineering cables in a solution of sodium chloride, *Wear* (2009) **267** (1-4): 308-314.
- [2] L. Dieng, V. Périer, L. Gaillet, and C. Tessier. Degradation mechanisms and protective methods of civil engineering cables, *Mécanique & Industries* (2009) **10**:33-42.
- [3] J.M. Walton. Developments in steel cables, *Journal of constructional steel research*

- (1996) **39**(1): 3-29.
- [4] D. Siegert, J. Royer and P. Brevet. Fretting fatigue in steel stay cables, *Int. Symposium on Fretting*, China (1997).
- [5] S. Fouvry, Ph. Kapsa, L. Vincent and K. Dang Van. Theoretical analysis of fatigue cracking under dry friction for fretting loading conditions, *Wear* (1996) **195**:21-34.
- [6] Z.R. Zhou, K. Nakazawa, M.H. Zhu, N. Maruyama, Ph. Kapsa and L. Vincent. Progress in fretting maps, *Tribology International* (2006) **39**:1068-1073
- [7] D. Siegert. Initiation of fretting fatigue cracks in spiral multilayer strands, *OIPEEC Bulletin* (1999) **78**.
- [8] P. S. Pao, S. J. Gill, C. R. Feng and K. K. Sankaran. Corrosion-fatigue crack growth in friction stir welded Al 7050. *Scripta Materialia* (2001) **45**(5):605-612.
- [9] F. Menan. *Influence de la corrosion saline sur la tolérance aux dommages d'un alliage d'aluminium aéronautique 2XXX*. PhD. Thesis, University of Poitiers (2008).

DEVELOPMENT AND IMPLEMENTATION OF AN EULERIAN APPROACH FOR EFFICIENT SIMULATION OF FRICTIONAL HEATING IN SLIDING CONTACTS

N. Strömberg*

*Department of Mechanical Engineering
Jönköping University
SE-551 11 Jönköping, Sweden
e-mail: stni@jth.hj.se

Key words: Eulerian, Thermomechanical, Frictional heating

Abstract. Thermal stresses as a result from frictional heating must be considered when designing disc brakes. The rotational symmetry of a disc brake makes it possible to model this system using an Eulerian approach instead of a Lagrangian framework. In this paper such an approach is developed. The sliding object is formulated in an Eulerian frame where the convective terms are defined by the sliding velocity. A node-to-node formulation of the contact interface is utilized. The energy balance of the interface is stated by introducing an interfacial temperature. Both frictional power and contact conductance are included in this energy balance. The contact problem is solved by a non-smooth Newton method. By adopting the augmented Lagrangian approach, this is done by rewriting Signorini's contact conditions to a system of semi-smooth equations. The heat transfer in the sliding body is discretized by a Petrov-Galerkin approach, i.e. the numerical difficulties due to the non-symmetric convective matrix appearing in a pure Galerkin discretization is treated by following the streamline-upwind approach. In such manner a stabilization is obtained by adding artificial conduction along the streamlines. For each time step the thermoelastic contact problem is first solved for the temperature field from the previous time step. Then, the heat transfer problem is solved for the corresponding frictional power. In such manner a temperature history is obtained via the trapezoidal rule. In particular the parameter is set such that both the Crank-Nicolson and the Galerkin methods are utilized. The method seems very promising. The method is demonstrated for two-dimensional benchmarks as well as a real disc brake system in three dimension.

1 INTRODUCTION

In this paper an Eulerian approach for sliding contacts is developed. In the design of machine components like brakes and clutches it is of importance to consider effects

from frictional heating. Today, this is mostly done by experiments. The Lagrangian approaches in our commercial softwares usually fail due convergence difficulties in the contact algorithms and too long computational times. An idea to improve these drawbacks is to formulate the problem in an Eulerian frame instead. This is the topic of this paper. The approach is presented for a two-dimensional translating problem. In a forthcoming paper the approach will also be developed for rotating systems as well as for the three-dimensional case.

Previously, we have studied thermo-mechanical contact problems in the setting of small displacements. In Strömberg [1] thermo-mechanical wear problems were studied for a thermo-elastic body in unilateral contact with a rigid foundation. The development of hot spots was studied by solving the fully coupled equation system using Newton's method. The influence of wear on these hot spots was also investigated numerically. That work was later extended to the case of two thermo-elastic bodies in unilateral contact in Ireman, Klarbring and Strömberg [2]. An earlier work on this topic in the same research group was done by Johansson and Klarbring [3]. In this paper, the thermo-mechanical framework developed in our previous works is now extended to also include large rotations with superimposed small displacements and this is done in an Eulerian framework.

Examples of early works on frictional heating in large displacements are e.g. the papers by Oancea and Laursen [4], and by Agelet de Saracibar [5]. A more recent paper is the one by Rieger and Wriggers [6] where the accuracy of the contact solution, which is most important in order to represent the frictional power sufficiently well, was controlled by adaptive techniques. Another way to improve the contact solution in large displacements is to use the mortar technique. Recently, this was investigated by Hieber and Wohlmuth [7] for thermo-mechanical friction problems. A nice feature with the presented Eulerian approach in this paper is that the contact region is always well defined and a node-to-node based approach can be adopted, producing very accurate contact solutions. The contact equations are then treated with the celebrated augmented Lagrangian approach where the corresponding equation system is solved by a non-smooth Newton algorithm. The details can be found in Strömberg [8].

One can find several other works where a Lagrangian formulation has been utilized for treating frictional heating, e.g. [9], but it is not easy to find any paper where an Eulerian framework is used. One example is the paper by Pauk and Yevtushenko [10] where a cylinder sliding over a half-space was considered. In this work we present a finite element approach using an Eulerian framework for solving frictional heating in sliding contacts. The fully coupled problem is decoupled in one mechanical part and another thermal problem. These two equation systems are then solved sequentially by using Crank-Nicolson's and Galerkin's settings of the trapezoidal rule in the time discretization. Other possibilities of performing the time discretization are of course also available. For instance, Laursen [11] proposed thermodynamically consistent algorithms for this class of problems. The convective term is stabilized by the streamline-upwind approach. For this task the excellent text-book by Donea and Huerta [12] has been consulted.

The proposed method is implemented on 64-bits Windows using Intel Fortran and the sparse Cholesky and LU solvers of Matlab. The pre- and postprocessing are performed on Abaqus/CAE by Python scripts. The implementation seems to be very robust and produce accurate solutions at low computational times. This is demonstrated by presenting numerical examples.

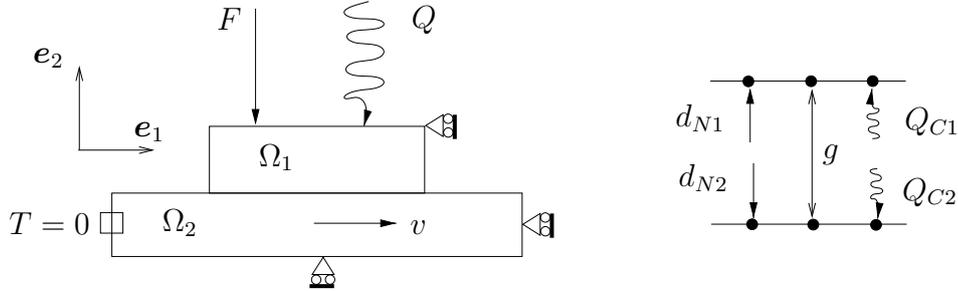


Figure 1: Two linear thermoelastic bodies in unilateral contact.

2 GOVERNING EQUATIONS

Let us consider two linear thermo-elastic bodies Ω_m in unilateral contact, see Figure 1. The first body is subjected to external forces $\mathbf{F} = \{F_i^A\}$ and heat powers $\mathbf{Q} = \{Q^A\}$ at finite element nodes \mathbf{x}_A on the top of the first body. The second body is translating with a constant velocity $\mathbf{v} = v\mathbf{e}_1$ and has superimposed small displacements onto the current rigid body configuration at time t . For each body Ω_m , the nodal displacements are collected in $\mathbf{d}_m = \{d_j^A\}$ and the nodal temperature vectors is represented by $\mathbf{T}_m = \{T^A\}$, respectively.

At the contact surface of each body, the normal displacements are given by

$$\mathbf{d}_{Nm} = \mathbf{C}_{Nm}\mathbf{d}_m, \quad (1)$$

where the rows of the transformation matrices \mathbf{C}_{Nm} contain surface normals in proper positions, i.e.

$$\begin{aligned} \mathbf{C}_{N1}^{\text{row}} &= [\mathbf{0} \ [0 \ -1] \ \mathbf{0}], \\ \mathbf{C}_{N2}^{\text{row}} &= [\mathbf{0} \ [0 \ 1] \ \mathbf{0}]. \end{aligned} \quad (2)$$

The corresponding normal contact forces \mathbf{F}_{Nm} are obtained by

$$\mathbf{F}_{Nm} = -\mathbf{C}_{Nm}^T \mathbf{P}, \quad (3)$$

where \mathbf{P} is a vector of Lagrange multipliers which are governed by Signorini's contact conditions:

$$\mathbf{P} \geq \mathbf{0}, \quad \mathbf{d}_{N1} + \mathbf{d}_{N2} \leq \mathbf{g}, \quad \mathbf{P} \circ (\mathbf{d}_{N1} + \mathbf{d}_{N2} - \mathbf{g}) = \mathbf{0}. \quad (4)$$

Here, \mathbf{g} represents a vector of initial gaps g^A between contact nodes in node-to-node contact, see Figure 1, and \circ is the Hadamard product.

It is assumed that sliding is always developed and that the corresponding frictional forces are given by

$$\mathbf{F}_{T_m} = \mu \mathbf{C}_{T_m}^T \mathbf{P}, \quad (5)$$

where

$$\begin{aligned} \mathbf{C}_{T_1}^{\text{row}} &= [\mathbf{0} \ [1 \ 0] \ \mathbf{0}], \\ \mathbf{C}_{T_2}^{\text{row}} &= [\mathbf{0} \ [-1 \ 0] \ \mathbf{0}]. \end{aligned} \quad (6)$$

These assumptions are in agreement with Coulomb's law of friction when sliding is developed.

By introducing finite element shape functions $N^A = N^A(\mathbf{x})$ and performing a finite element discretization, the following equilibrium equations in forces can be derived:

$$\begin{aligned} \mathbf{K}_1 \mathbf{d}_1 - \hat{\mathbf{K}}_1 \mathbf{T}_1 &= \mathbf{F} + \mathbf{F}_{N_1} + \mathbf{F}_{T_1}, \\ \mathbf{K}_2 \mathbf{d}_2 - \hat{\mathbf{K}}_2 \mathbf{T}_2 &= \mathbf{F}_{N_2} + \mathbf{F}_{T_2}, \end{aligned} \quad (7)$$

where

$$\begin{aligned} \mathbf{K}_m &= [K_{ik}^{BA}], \quad K_{ik}^{BA} = \int_{\Omega_m} E_{ijkl} \frac{\partial N^A}{\partial x_l} \frac{\partial N^B}{\partial x_j} dV, \\ \hat{\mathbf{K}}_m &= [\hat{K}_i^{BA}], \quad \hat{K}_i^{BA} = \int_{\Omega_m} \alpha(3\lambda + 2G) N^A \frac{\partial N^B}{\partial x_i} dV, \end{aligned} \quad (8)$$

$E_{ijkl} = \lambda \delta_{ij} \delta_{kl} + G(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})$, λ and G are Lamé's coefficients, respectively, and α represents the thermal expansion coefficient.

The energy balance for the first body reads

$$\mathbf{M}_1 \dot{\mathbf{T}}_1 + \mathbf{O}_1 \mathbf{T}_1 = \mathbf{Q} + \mathbf{Q}_{C_1}, \quad (9)$$

where

$$\begin{aligned} \mathbf{M}_m &= [M^{BA}], \quad M^{BA} = \int_{\Omega_m} \rho c N^A N^B dV, \\ \mathbf{O}_m &= [O^{BA}], \quad O^{BA} = \int_{\Omega_m} k \frac{\partial N^A}{\partial x_i} \frac{\partial N^B}{\partial x_i} dV, \end{aligned} \quad (10)$$

ρ is the mass density, c is the heat capacity and k is the thermal conductivity. Furthermore, by introducing the contact conductance ϑ and the frictional dissipation at each contact pair as

$$\mu \mathbf{P} v, \quad (11)$$

we can define the heat power transferred at the first contact surface as

$$\mathbf{Q}_{C_1} = \frac{\vartheta}{2} \mathbf{P} \circ (\mathbf{S}_2 \mathbf{T}_2 - \mathbf{S}_1 \mathbf{T}_1) + \frac{1}{2} \mu \mathbf{P} v. \quad (12)$$

Here, we have also introduced \mathbf{S}_i , where $\mathbf{S}_i^{\text{row}} = [\mathbf{0} [1] \mathbf{0}]$, in order to obtain the nodal temperatures at the contact surfaces. In a similar way, we can define the heat power transferred at the second contact surface as

$$\mathbf{Q}_{C_2} = \frac{\vartheta}{2} \mathbf{P} \circ (\mathbf{S}_1 \mathbf{T}_1 - \mathbf{S}_2 \mathbf{T}_2) + \frac{1}{2} \mu \mathbf{P} v. \quad (13)$$

In the energy balance for the second body convective terms appear due to the speed v . These are represented by $\mathbf{N} \mathbf{T}_2$, where

$$\mathbf{N} = [N^{BA}], \quad N^{BA} = \int_{\Omega_2} \rho c v \frac{\partial N^A}{\partial x_1} N^B \, dV, \quad (14)$$

Thus, the convection matrix \mathbf{N} is non-symmetric. When this matrix dominates over the symmetric conduction matrix \mathbf{O}_2 , then the thermal solution might be unstable. This might be stabilized by adding artificial conduction along the streamlines by $\mathbf{R} \mathbf{T}_2$, where

$$\mathbf{R} = [R^{BA}], \quad R^{BA} = \bar{k} v \int_{\Omega_2} \frac{\partial N^A}{\partial x_1} \frac{\partial N^B}{\partial x_1} \, dV, \quad (15)$$

and \bar{k} is an artificial conduction coefficient. By using (14) and (15), we obtain the following energy balance for the second body:

$$\mathbf{M}_2 \dot{\mathbf{T}}_2 + (\mathbf{N} + \mathbf{R} + \mathbf{O}_2) \mathbf{T}_2 = \mathbf{Q}_{C_2}. \quad (16)$$

3 NUMERICAL TREATMENT

The equations presented in the previous section are treated sequentially for each time step by decoupling the mechanical and thermal equations. That is, for a given temperature distribution the thermo-mechanical contact problem is first solved, then for the obtained contact force distribution the energy balance is solved. Details are presented in this section.

The contact problem is treated by the augmented Lagrangian approach. The key idea is to rewrite (4) as

$$\mathbf{P} = (\mathbf{P} + r(\mathbf{d}_{N_1} + \mathbf{d}_{N_2} - \mathbf{g}))_+, \quad (17)$$

where $r > 0$ is a penalty coefficient and $(x)_+ = (x + |x|)/2$. (7) and (17) are then put together to form an equation system as

$$\mathbf{h} = \mathbf{h}(\mathbf{d}_1, \mathbf{d}_2, \mathbf{P}, \mathbf{T}_1, \mathbf{T}_2) = \mathbf{0}. \quad (18)$$

This is a semi-smooth equation system which is efficiently solved by a Newton algorithm with an inexact line-search procedure for given temperatures \mathbf{T}_m .

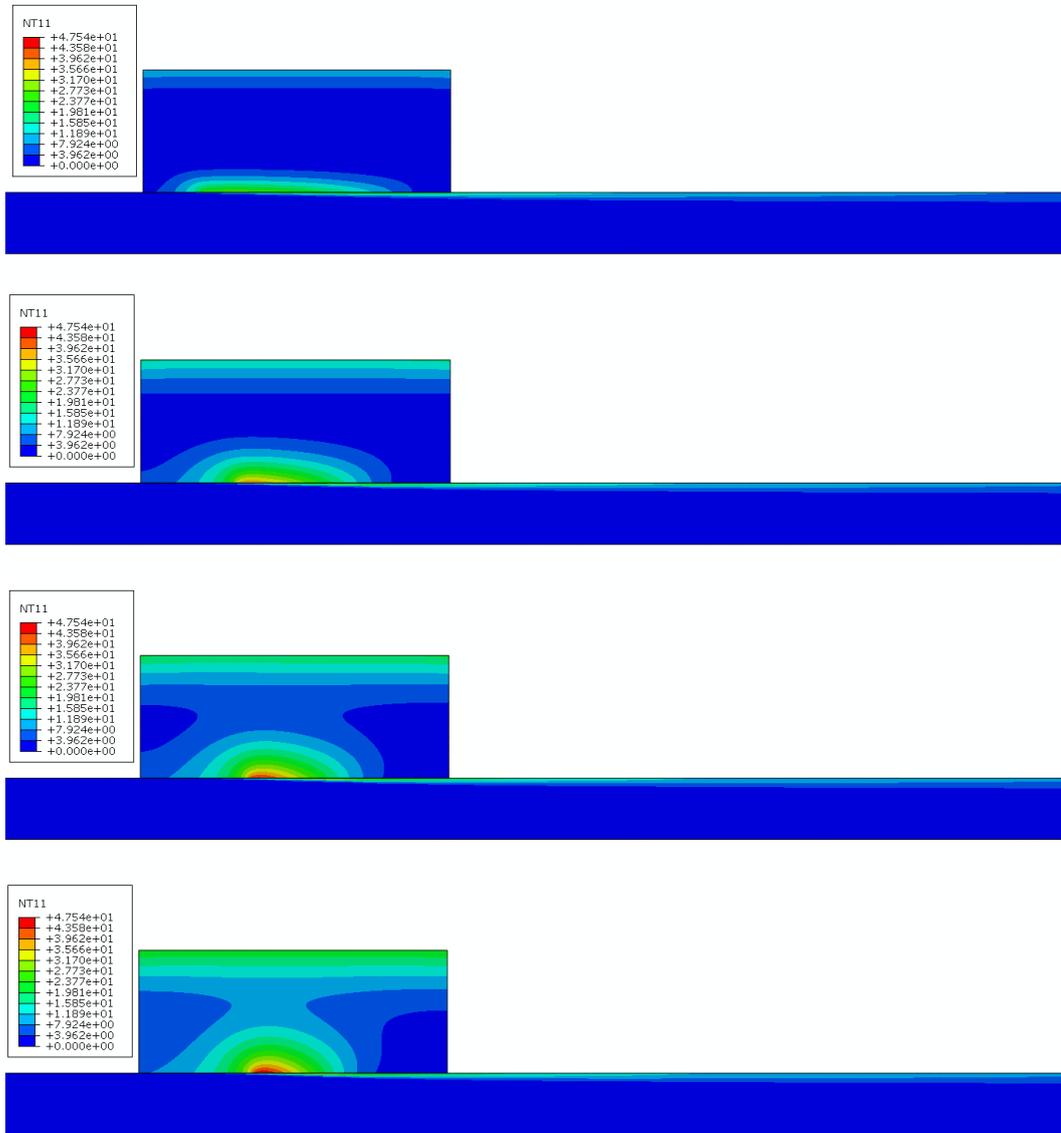


Figure 2: The nodal temperatures plotted at different times: after 20, 40, 60 and 80 increments, respectively.

The time rates appearing in the energy balances are discretized by the trapezoidal rule. Let $\mathbf{T}_m^n = \mathbf{T}_m(t_n)$ at time t_n , then the temperatures at the next time step t_{n+1} are updated according to

$$\mathbf{T}_m^{n+1} = \mathbf{T}_m^n + \Delta t \left((1 - \xi) \dot{\mathbf{T}}_m^n + \xi \dot{\mathbf{T}}_m^{n+1} \right), \quad (19)$$

where $\Delta t = t_{n+1} - t_n$ and $\xi = 1/2$ (Crank-Nicolson) or $\xi = 2/3$ (Galerkin). (19) inserted in (9) and (16) yields

$$\begin{aligned} \left(\mathbf{O}_1 + \frac{1}{\xi \Delta t} \mathbf{M}_1 \right) \mathbf{T}_1^{n+1} &= \mathbf{Q}_1^{\text{eff}} + \mathbf{Q}_{C1}, \\ \left(\mathbf{N} + \mathbf{R} + \mathbf{O}_2 + \frac{1}{\xi \Delta t} \mathbf{M}_2 \right) \mathbf{T}_2^{n+1} &= \mathbf{Q}_2^{\text{eff}} + \mathbf{Q}_{C2}, \end{aligned} \quad (20)$$

where

$$\begin{aligned} \mathbf{Q}_1^{\text{eff}} &= \mathbf{Q} + \frac{(1 - \xi)}{\xi} \mathbf{M}_1 \dot{\mathbf{T}}_1^n + \frac{1}{\xi \Delta t} \mathbf{M}_1 \mathbf{T}_1^n, \\ \mathbf{Q}_2^{\text{eff}} &= \frac{(1 - \xi)}{\xi} \mathbf{M}_2 \dot{\mathbf{T}}_2^n + \frac{1}{\xi \Delta t} \mathbf{M}_2 \mathbf{T}_2^n. \end{aligned} \quad (21)$$

Furthermore, (12) and (13) can also be written as

$$\begin{aligned} \mathbf{Q}_{C1} &= \mathbf{S}_{P2} \mathbf{T}_2 - \mathbf{S}_{P1} \mathbf{T}_1 + \mathbf{Q}_\mu, \\ \mathbf{Q}_{C2} &= \mathbf{S}_{P1} \mathbf{T}_1 - \mathbf{S}_{P2} \mathbf{T}_2 + \mathbf{Q}_\mu, \end{aligned} \quad (22)$$

where \mathbf{S}_{P_i} and \mathbf{Q}_μ all depend on \mathbf{P} . By putting together (20) and (22), one obtains an equation system on the following form:

$$\mathbf{A}(\mathbf{P}) \begin{Bmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \end{Bmatrix} = \mathbf{Q}(\mathbf{P}), \quad (23)$$

which of course becomes a linear system for given multipliers \mathbf{P} .

In conclusion, let $\mathbf{d}_m^n, \mathbf{T}_m^n, \mathbf{P}^n$ be given at time t_n , then $\mathbf{d}_m^{n+1}, \mathbf{T}_m^{n+1}, \mathbf{P}^{n+1}$ are obtained by the following steps:

Step 1:

$$\mathbf{h}(\mathbf{d}_1^{n+1}, \mathbf{d}_2^{n+1}, \mathbf{P}^{n+1}, \mathbf{T}_1^n, \mathbf{T}_2^n) = \mathbf{0}$$

is solved by Newton's method, details can be found in [8].

Step 2:

$$\begin{Bmatrix} \mathbf{T}_1^{n+1} \\ \mathbf{T}_2^{n+1} \end{Bmatrix} = \mathbf{A}(\mathbf{P}^{n+1})^{-1} \mathbf{Q}(\mathbf{P}^{n+1}).$$

Step 3:

$$\dot{\mathbf{T}}_m^{n+1} = -\frac{1 - \xi}{\xi} \dot{\mathbf{T}}_m^n + \frac{\mathbf{T}_m^{n+1} - \mathbf{T}_m^n}{\xi \Delta t}.$$

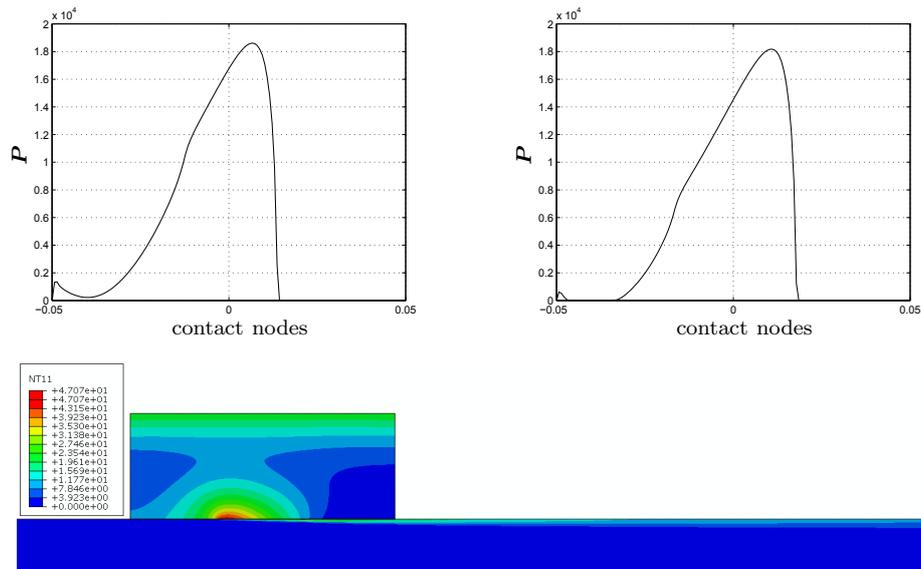


Figure 3: A comparison when the friction force is neglected in the equilibrium equations.

4 NUMERICAL EXAMPLES

The problem in Figure 1 is here considered as a numerical benchmark. The dimensions of the two bodies are taken to be 0.1×0.04 [m²] and 0.5×0.02 [m²], respectively. The plain strain assumption is adopted with a thickness of 1 [m]. The first body is meshed using 8151 elements and for the second body 20735 elements are used. Young’s modulus is 2.1×10^{11} [Pa], Poisson’s ratio is 0.3, the expansion coefficient is 1.2×10^{-5} [1/K], the density is 7800 [kg/m³], the heat capacity is 460 [J/kgK], the conductivity is 46 [W/mK] and the conduct conductance is taken to be $\varphi = 1$ [W/NK].

A total heat power of $Q = 5760$ [W] is applied on the top of body Ω_1 as well as a total force of $F = 72 \times 10^4$ [N] (corresponding to a pressure of 7.2 [MPa]). Both the total heat power and the total force are equally distributed over all contact nodes on the top surface. The heat power is applied at time zero and the force is ramped up using a log-sigmoid function for 20 time increments. The problem is solved for 80 time increments with a constant time step $\Delta t = 0.125$ [s]. The speed of the second body is $v = 1$ [m/s]. Thus, the total sliding distance is 10 [m]. The evolution in temperatures for this problem when $\mu = 0.1$ are plotted in Figure 2.

The bottle-neck of the algorithm is to solve the linear system appearing in the Newton algorithm. Typically 4-6 such linear systems have to be solved for getting convergence in each time step. The system is also non-symmetric due to the friction force. One approach to speed up these calculation is to assume that the friction force has a little influence on the thermal solution. If the friction force is neglected in the mechanical problem, then only

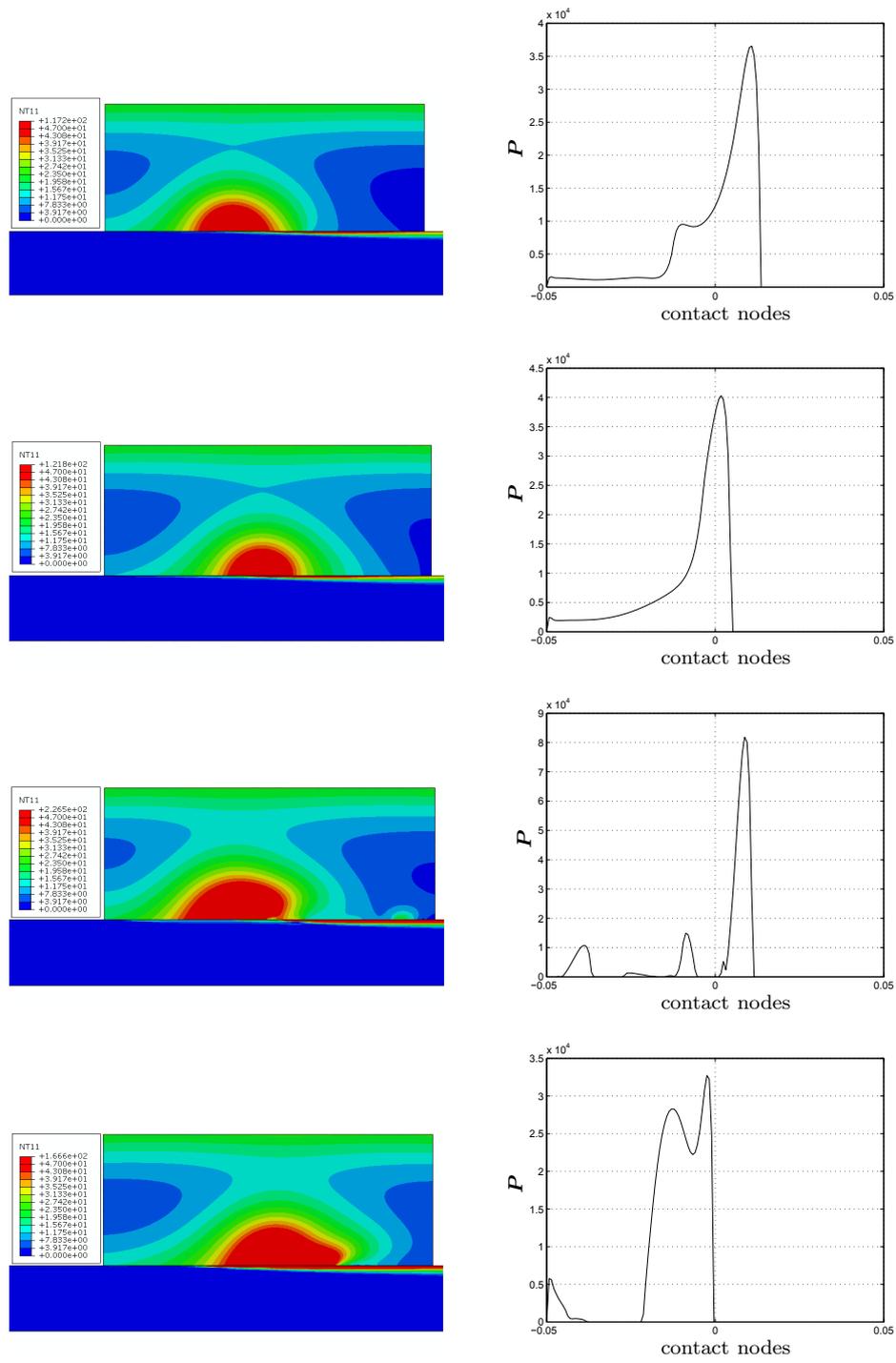


Figure 4: Comparison of temperatures and contact forces for $\mu = 0.2$ and 0.3 when friction forces are included as well as neglected.

2-3 Newton steps are needed and the linear system also becomes symmetric. In Figure 3 we have utilized this approach and compare the final solution to the original one presented in Figure 2. The resemblance of the two solutions are very close. A similar comparison is performed when the friction coefficient is taken to be 0.2 and 0.3, respectively. The results are presented in Figure 4. Also here a close resemblance between the results is obtained for the two different approaches, with and without frictional forces. The difference shown in the plots for $\mu = 0.3$ depends mostly on a time shift. This will be explained in more detail at the conference.

Another approach for speeding up the calculations but still consider the frictional force is to first solve the frictionless problem and then letting the friction force be defined by the obtained frictionless contact pressure, and solving the friction problem for this constant friction force. That is, at each iteration, (18) is first solved for $\mathbf{F}_{T_m} = \mathbf{0}$. Let $\hat{\mathbf{P}}$ denote the solution and then solve (18) again but now with $\mathbf{F}_{T_m} = \mu \mathbf{C}_{T_m}^T \hat{\mathbf{P}}$. In general, the number of iterations will be twice the number of iterations for the frictionless case. Of course, for this case, we will also have a symmetric Jacobian which is most beneficial for large size problems.

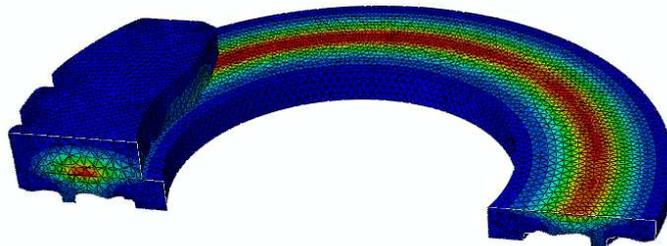


Figure 5: A heat band developed in a disc brake.

5 CONCLUSIONS

In this work a method for simulating frictional heating in sliding contacts is developed and implemented. A key idea of the approach is to use an Eulerian frame for the sliding object. The convective term appearing in this approach is stabilized by the streamline-upwind technique. The method seems promising. This is shown by solving a two-dimensional benchmark with a translating object for different coefficients of friction. The next step in the development will be to consider rotating objects. The ultimate goal is to solve frictional heating in disc brakes efficiently. A preliminary result is presented in Figure 5.

Acknowledgement *The mesh of the three-dimensional disc-pad problem was prepared by Asim Rashid. This project was financed by Vinnova (FFI-Strategic Vehicle Research and Innovation)*

and Volvo 3P.

REFERENCES

- [1] Strömberg, N., Finite element treatment of two-dimensional thermoelastic wear problems, *Computer Methods in Applied Mechanics and Engineering* (1999) **177**:441–455.
- [2] Ireman, P., Klarbring, A. and Strömberg, N., Finite element algorithms for thermoelastic wear problems, *European Journal of Mechanics, A/Solids* (2002) **21**:423–440.
- [3] Johansson, L. and Klarbring, A., Thermoelastic frictional contact problems: modelling, FE-approximation and numerical realization, *Computer Methods in Applied Mechanics and Engineering* (1993) **105**: 181–210.
- [4] Oancea, V.G. and Laursen, T.A., A finite element formulation of thermomechanical rate-dependent frictional sliding, *International Journal for Numerical Methods in Engineering* (1997) **40**: 4275–4311.
- [5] Agelet de Saracibar, C., Numerical analysis of coupled thermomechanical frictional contact problems. Computational model and applications, *Archives of Computational Methods in Engineering* (1998) **5**:243–301.
- [6] Rieger, A. and Wriggers, P., Adaptive methods for thermomechanical coupled contact problems, *International Journal for Numerical Methods in Engineering* (2004) **59**:871–894.
- [7] Hieber, S. and Wohlmuth, B.I., Thermo-mechanical contact problems of non-matching meshes, *Computer Methods in Applied Mechanics and Engineering* (2009) **198**:1338–1350.
- [8] Strömberg, N., An augmented Lagrangian method for fretting problems, *European Journal of Mechanics, A/Solids* (1997) **16**:573–593.
- [9] Pantuso, D., Bathe, K.J. and Bouzinov, P.A., A finite element procedure for the analysis of thermo-mechanical solids in contact, *Computers & Structures* (2000) **75**:551–573.
- [10] Pauk, V.J. and Yevtushenko, A.A., Frictional heating in sliding contact of two thermoelastic bodies, *International Journal of Solids and Structures* (1997) **34**:1505–1516.
- [11] Laursen, T.A., On the development of thermodynamically consistent algorithms for thermomechanical frictional contact, *Computer Methods in Applied Mechanics and Engineering* (1999) **177**: 274–287.
- [12] Donea, J., and Huerta, A., *Finite element methods for flow problems*, Wiley, (2003).

DYNAMICALLY COUPLED MODELS OF THE SLIDING AND SPINNING FRICTION BASED ON PADÉ EXPANSIONS

ALEXEY A. KIREENKOV*

* Institute for Problems in Mechanics of the Russian Academy of Sciences
Prospekt Vernadskogo 101, korp.1, 119526 Moscow, Russia
e-mail: kireenk@ipmnet.ru, kireenk@mail.ru

Key words: Models of the sliding and spinning friction, Pade expansions.

Abstract. It is presented a new approach for dry friction modeling under conditions of combined kinematics. The main distinguish feature of this approach is building of friction models which are suitable for using in differential equations of motion. Under the proposed models of friction are understudied the interrelations between friction force components, torques and velocities which are represented by the analytical functions. The procedure of the models constructing consists of the two parts. In the first part, the exact integral expressions for the net vector and torque are formed with the assumption that Coulomb's friction law in classical forms or generalized differential forms is valid at each point of the contact area. In addition, in process of the exact integral models construction there are used well known results from the theory of elasticity that tangent stresses lead to shift in the symmetric diagram of the normal contact stresses in the direction of the instantaneous sliding velocity. To use the theory of elasticity results in the dynamics problems, it is proposed the simple asymptotic representations for the contact stresses distributions based on their general properties known from the theoretical results of the theory of elasticity. In the second part the exact integral models are replaced by appropriate Pade expansions. The approximate models preserve all properties of the models based on the exact integral expressions and correctly describe the behaviour of the net vector and torque of the friction forces and their first derivatives at zero and infinity. Moreover, one does not have even to calculate the integrals to determine the coefficients of the Pade approximations. The corresponded coefficients can be identified from experiments. Consequently, the models based on Pade expansions may be considered as phenomenological models of combined dry friction.

1 INTRODUCTION

One of the first models describing the relation between the sliding friction and the whirling friction in the case of nonpoint contact between the moving bodies was proposed by in [1]. A principally new development of the theory was given by in [2], where exact analytic expressions for the resultant vector and the frictional moment for circular contact sites were obtained under the assumption that the distribution of contact stresses in the contact spot obeys the Hertz law. In [2], to apply the obtained dependencies to problems of dynamics, the linear-fractional Pade approximations of these dependencies were

constructed. The developed in [2] theory was used in [3] to study the dynamics of a homogeneous circular disk sliding with rotation on a plane. Under the assumption that the distribution of contact stresses obeys the Galin law, exact analytic expressions for the resultant vector and the frictional moment were obtained and their linear-fractional Pade' approximations were constructed.

The convenience in the use of the Pade approximations, which permit describing the effects of combined dry frictions for the entire range of angular and linear velocities, allowed one to construct principally new the two-dimensional coupled models of the sliding and spinning friction on the basis of these approximations [4].

The two-dimensional friction model was constructed under supposition that the classical Coulomb law in differential form is validated for an infinitesimal area inside of contact spot. Its generalizing for the case of more realistic dry friction characteristic (validity of Coulomb law in generalized differential form) was given in [5]. It was shown that in the case of combined kinematics using of the Coulomb law in generalized differential form leads to new qualitative properties of the friction force dependence on the sliding and spinning velocities, but does not change the model dimension. All these models of the sliding and spinning friction were constructed in the assumption that, in the case of circular contact sites, the distributions of normal contact stresses depend only on the position vector with origin at the contact spot center. But, it is known [6] that in the case of the rigid solids sliding it is appears tangent stresses that leads to shifting in the symmetric diagram of the normal contact stresses in the direction of the instantaneous sliding velocity. Investigations carried out in [7] shown this shifting even for uniform distribution of the normal contact stresses cause, in the case of combined kinematics, the dynamics coupling between components defining the force state of rubbed solids.

Proposed below the dry friction models generalizing permits to take into account, simultaneously, both the dynamics coupling of the components defining force state and the more realistic representations about dry friction characteristics and the normal contact stresses distributions in the case of combined kinematics.

2 COUPLED MODELS OF THE SLIDING AND SPINNING FRICTION

2.1 Basic relationships

The combined model of sliding and rolling friction is constructed for circular contact sites under the assumption that the Coulomb law in differential form holds for the small surface element dS in the interior of the contact spot, according to which the differentials of the resultant vector $d\mathbf{F}$ and the moment of friction dM_c with respect to the disk center are determined by the formulas:

$$d\mathbf{F} = -f\sigma \frac{\mathbf{V}}{|\mathbf{V}|} (1 + \mu_1 |\mathbf{V}|^3 - \mu_2 |\mathbf{V}|) dS, \quad dM_c = -f\sigma \frac{\mathbf{r} \times \mathbf{V}}{|\mathbf{V}|} (1 + \mu_1 |\mathbf{V}|^3 - \mu_2 |\mathbf{V}|) dS, \quad (1)$$

$$\mathbf{V} = (v - \omega y, \omega x), \quad \mathbf{r} = (x, y)$$

where f is the coefficient of friction, $\mathbf{r} = (x, y)$ is the position vector of an elemental area in the interior of the contact spot with respect to its center (Fig. 1), ω is the angular

velocity of rotation of the contact spot center, but μ_1 and μ_2 are the coefficients which can be defined in practice from experiments.

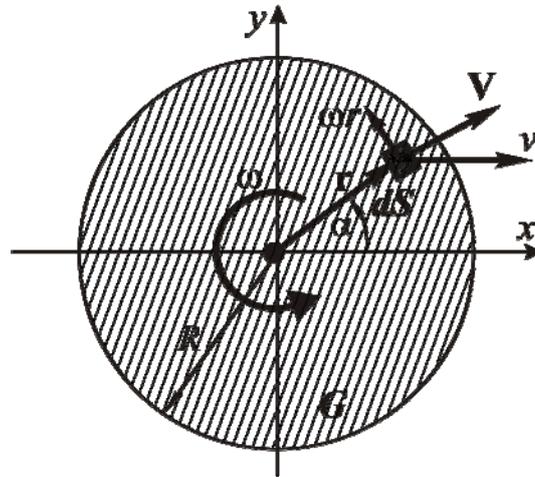


Figure 1. Kinematics inside the contact spot letters

To use the theory of elasticity results in the dynamics problems, a simple linear approximation of the normal contact stresses distribution is proposed:

$$\sigma(x, y) = \sigma_0 (1 + kx/R) \quad (2)$$

where $\sigma_0 = \sigma_0(r)$ - distribution of normal contact stresses at absence of motion having the properties of central symmetry, R - radius of contact spot, x - α e of the rectangular coordination systems with origin in the center of contact circle (Fig. 1) which is directed parallelly to vector of the instantenious sliding.

To calculate coefficient k in the formula (1) it is used the condition of equality of the external force F torque to the normal reaction force N torque which is appears from the shifting of the center of gravity of the contact spot in the direction of sliding on the value s :

$$Fh = Ns \quad (3)$$

where h - distance from the moving solid center mass to the plane of sliding. On the other hand the shifting s of the gravity center relatively of the contact spot center can be defined by the following formula:

$$s = \frac{\iint_G x\sigma(x, y)dx dy}{\iint_G \sigma(x, y)dx dy}, \quad G = \{(x, y) : x^2 + y^2 \leq R^2\} \quad (4)$$

Substitution of the representation (2) to the (4) yields:

$$s = \frac{\pi k}{R} \int_0^R \sigma_0(r) r^3 dr \quad (5)$$

Equalization values s calculated from the formulas (3) and (5) allows to calculate coefficient k which is characterized the dynamical coupling of the components defining the force state inside of contact spot.

If the distribution of normal contact stresses $\sigma_0(r)$ at the moving absence are obeyed by the Hertz $\sigma_0 = 3N\sqrt{1-r^2/R^2} / (2\pi R^2)$ or Galin $\sigma_0 = N / (2\pi R^2 \sqrt{1-r^2/R^2})$ laws then $k = 5Fh/(NR)$ or $k = 3Fh/(NR)$, correspondently.

2.1 Integral model

To obtain the resultant vector and the moment of friction, it is necessary to integrate the expressions (1) over the contact spot. The obtained dependencies, where F_{\parallel} and F_{\perp} denote the respective components of the resultant vector directed along the tangent and the normal to the trajectory of motion, present an exact combined integral model of sliding and spinning friction

$$\begin{aligned} F_{\parallel}(\omega, v) &= -f \iint_G \left(\frac{(v - \omega y)}{\sqrt{\omega^2(x^2 + y^2) + v^2 - 2\omega v y}} + \mu_1 v^3 - \mu_2 v + 2\mu_1 v \omega^2 (x^2 + y^2) \right) \sigma_0 dx dy \\ F_{\perp}(\omega, v) &= -\frac{kf}{R} \iint_G \frac{\omega x^2 \sigma_0}{\sqrt{\omega^2(x^2 + y^2) + v^2 - 2\omega v y}} dx dy, \quad G = \{(x, y) : x^2 + y^2 \leq R^2\} \\ M_C(\omega, v) &= -f \iint_G \left(\frac{(\omega(x^2 + y^2) - vy)}{\sqrt{\omega^2(x^2 + y^2) + v^2 - 2\omega v y}} + (2\mu_1 v^2 - \mu_2)\omega(x^2 + y^2) + \mu_1 \omega^3 (x^2 + y^2)^2 \right) \sigma_0 dx dy \end{aligned} \quad (6)$$

After introducing dimensionless variables: $x = \hat{x}R$, $y = \hat{y}R$ and $\sigma(\hat{x}, \hat{y}) = \hat{\sigma}(\hat{x}, \hat{y})N/R^2$ it is convenient to calculate the modulus of integrals (6) in the polar coordinates: $x = r \cos \alpha$, $y = r \sin \alpha$, $r \in [0, 1]$, $\alpha \in [0, 2\pi]$ (Fig. 1) in which the functions (6) take the form

$$\begin{aligned} F_{\parallel} &= fN \int_0^{2\pi} \int_0^1 \frac{(v - ur \sin \alpha) r \sigma_0(r)}{\sqrt{u^2 r^2 + v^2 - 2uvr \sin \alpha}} dr d\alpha + 2\pi f \left((\mu_1 v^3 - \mu_2 v) \int_0^1 r \sigma_0(r) dr + 2\mu_1 v u^2 \int_0^1 r^3 \sigma_0(r) dr \right) \\ F_{\perp} &= kfN \int_0^{2\pi} \int_0^1 \frac{ur^3 \sigma_0(r) \cos^2 \alpha}{\sqrt{u^2 r^2 + v^2 - 2uvr \sin \alpha}} dr d\alpha, \quad u = \omega R \\ M_C &= fRN \int_0^{2\pi} \int_0^1 \frac{(ur^2 - vr \sin \alpha) r \sigma_0(r)}{\sqrt{u^2 r^2 + v^2 - 2uvr \sin \alpha}} dr d\alpha + 2\pi f \left((2\mu_1 v^2 - \mu_2) u \int_0^1 r^3 \sigma_0(r) dr + \mu_1 u^3 \int_0^1 r^5 \sigma_0(r) dr \right) \end{aligned} \quad (7)$$

where the ‘‘hat’’ symbol is omitted for brevity

If $k = 0$, then model (7) is fully agree to the model, investigated in [3] and can be considered as the first approximation, but presented in this investigation as the second approximation. Thus, we have substantial approximation to the real situation in dependence on the general properties of the normal contact stresses distribution. At the supposition that external forces are absence, the coefficient k in formula (1), (5), (6), (7) is defined by the friction force component F_{\parallel} from the first expressions in the relations (6-8) and, consequently, the dynamically coupled integral friction model is

$$\begin{aligned}
 F_{\parallel} &= fN \int_0^{2\pi} \int_0^1 \frac{(v - ur \sin \alpha) r \sigma_0(r)}{\sqrt{u^2 r^2 + v^2 - 2uvr \sin \alpha}} dr d\alpha + 2\pi f \left((\mu_1 v^3 - \mu_2 v) I_1 + 2\mu_1 v u^2 I_3 \right) \\
 F_{\perp} &= kfN \int_0^{2\pi} \int_0^1 \frac{ur^3 \sigma_0(r) \cos^2 \alpha}{\sqrt{u^2 r^2 + v^2 - 2uvr \sin \alpha}} dr d\alpha, \quad k = \frac{fhR}{\pi I_3} \int_0^{2\pi} \int_0^1 \frac{(v - ur \sin \alpha) r \sigma_0(r)}{\sqrt{u^2 r^2 + v^2 - 2uvr \sin \alpha}} dr d\alpha \\
 M_C &= fRN \int_0^{2\pi} \int_0^1 \frac{(ur^2 - vr \sin \alpha) r \sigma_0(r)}{\sqrt{u^2 r^2 + v^2 - 2uvr \sin \alpha}} dr d\alpha + 2\pi f \left((2\mu_1 v^2 - \mu_2) u I_3 + \mu_1 u^3 I_5 \right)
 \end{aligned}
 \tag{8}$$

where coefficients of polynomials terms in formulas (3) are the first moments of the normal contact stresses distribution: $I_1 = \int_0^1 r \sigma_0(r) dr$ - moment of the first order,

$I_3 = \int_0^1 r^3 \sigma_0(r) dr$ - moment of the third order and $I_5 = \int_0^1 r^5 \sigma_0(r) dr$ - moment of the fifth order.

They can be calculated in elementary functions for the most used functions of the normal contact stresses distributions [5].

If the distribution of normal contact stresses is obeyed to the Hertz law: $\sigma(r) = 3\sqrt{1-r^2}/(2\pi)$ then: $I_1 = 1/2\pi, I_3 = 1/5\pi, I_5 = 4/35\pi$.

If the distribution of normal contact stresses is obeyed to the Galin law: $\sigma(r) = (2\pi\sqrt{1-r^2})^{-1}$ then: $I_1 = 1/2\pi, I_3 = 1/3\pi, I_5 = 4/15\pi$.

In the case of thin circle, the distribution of normal contact stresses can be described by the following function: $\sigma(r) = \delta(r-1)/(2\pi)$, where $\delta(r-1)$ - Dirac delta function in the point $r = 1$ and $I_1 = I_3 = I_5 = 1/4\pi$.

Plots of the tangent F_{\parallel} (left figure) and normal F_{\perp} (right figure) friction force components normalized on the their maximum values as function of velocity of sliding v at the constant velocity of whirling $u = 1$ are presented on the Fig. 2. As concerned friction torque then, qualitatively, its behavior is the same as case of using classical form Coulomb law: there are only small quantitative distinctions.

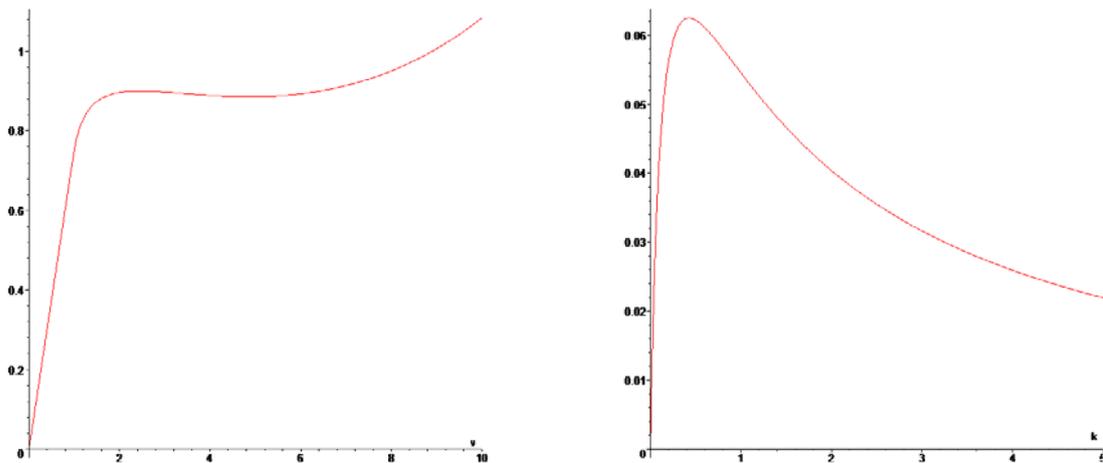


Figure 2. Tangent and normal friction force components

The expressions for the components of the resultant vector and the moment of friction in relations (8) have several important properties as functions of u and v .

Property 1. The distortion in symmetric diagram of the normal contact stresses distribution results in the appearance of the resultant vector component F_{\perp} directed along the normal to the trajectory of motion. The resultant vector is not directed opposite to the velocity of sliding.

Property 2. The distortion in the symmetric diagram of distribution of normal stresses does not affect to the moment M_C and the resultant vector component F_{\parallel} directed along the tangent to the trajectory.

Property 3. The first terms of the tangent F_{\parallel} force component and torque M_C , just as normal F_{\perp} force component, are homogeneous functions of the variables u and v of zero order of homogeneity and hence are invariant under the similarity group:

Property 4. The expressions (9), for the moment and both components of the friction force as functions of u and v have a singularity at the point $(u, v) = (0, 0)$, because they do not have any limit at this point with respect to both of the variables u and v .

Property 5. In the case of pure sliding $u = 0$ or spinning $v = 0$, the moment M_C and the tangential component F_{\parallel} are homogeneous models corresponding to the usual Coulomb law:

$$F_{\parallel}(0, v) = F_0 \equiv fN, \quad M_C(u, 0) = M_0, \quad M_0 = 2\pi fNR I_2, \quad I_2 = \int_0^1 \sigma_0(r)r^2 dr$$

Property 6. In the case of pure sliding, the normal component vanishes: $F_{\perp}(0, v) = 0$, and hence the friction force is directed opposite to the velocity vector; in the case of pure spinning, it is equal to $F_{\perp}(u, 0) = \mu F_0$, $\mu = fhR/(\pi I_3)$.

Property 7. The moment M_C and both components of the friction force F_{\parallel} and F_{\perp} have only one nonzero first partial derivative (the others are zero):

$$\left. \frac{\partial M_C}{\partial u} \right|_{u=0} \neq 0, \quad \left. \frac{\partial F_{\parallel}}{\partial v} \right|_{v=0} \neq 0, \quad \left. \frac{\partial F_{\perp}}{\partial u} \right|_{u=0} \neq 0$$

2.2 Models based on Pade expansions

The integral models (8) give a good description of the combined sliding and spinning friction, but are inconvenient to be used in problems of dynamics, because it is required to calculate multiple integrals in the right-hand sides of the equations of motion. This difficult procedure can be eliminated by replacing the exact integral expressions by the corresponding Pade approximations. The simplest of them is the linear-fractional approximation preserving the value at zero and at infinity of both for the torque M_C and for the tangent force component F_{\parallel} . But, for the normal friction force component, corresponded Pade approximation, naturally, became of the second order.

$$\begin{aligned}
 M_C &= M_0 \left(\frac{u}{u+mv} + 2\pi \left((2\mu_1 v^2 - \mu_2) u I_3 + \mu_1 u^3 I_5 \right) \right), \quad \frac{1}{m} = \frac{v}{M_0} \frac{\partial M_C}{\partial u} \Big|_{u=0} \\
 F_{\parallel} &= F_0 \left(\frac{v}{v+au} + 2\pi \left((\mu_1 v^3 - \mu_2 v) I_1 + 2\mu_1 v u^2 I_3 \right) \right), \quad \frac{1}{a} = \frac{u}{F_0} \frac{\partial F_{\parallel}}{\partial v} \Big|_{v=0} \\
 F_{\perp} &= \frac{\mu F_0 u v}{(u+bv)(v+au)}, \quad \frac{1}{b} = \frac{v}{\mu F_0} \frac{\partial F_{\perp}}{\partial u} \Big|_{u=0}
 \end{aligned} \tag{9}$$

The linear-fractional Pade' approximations (9) preserve the values of the functions $F_{\parallel}(u, v)$, $F_{\perp}(u, v)$ and $M_C(u, v)$ at zero, as well as their behavior and the behavior of their first derivatives at infinity. But model of this type cannot completely preserve the values of all first partial derivatives of these functions at zero. To obtain a correct description of the behavior of the first derivatives at zero, it is required to use the second-order Pade' approximations, and then the coupled model of sliding and spinning friction takes the form

$$\begin{aligned}
 M_C &= M_0 \left(\frac{u^2 + muv}{v^2 + muv + u^2} + 2\pi \left((2\mu_1 v^2 - \mu_2) u I_3 + \mu_1 u^3 I_5 \right) \right), \quad m = \frac{v}{M_0} \frac{\partial M_C}{\partial u} \Big|_{u=0} \\
 F_{\parallel} &= F_0 \left(\frac{v^2 + auv}{v^2 + auv + u^2} + 2\pi \left((\mu_1 v^3 - \mu_2 v) I_1 + 2\mu_1 v u^2 I_3 \right) \right), \quad a = \frac{u}{F_0} \frac{\partial F_{\parallel}}{\partial v} \Big|_{v=0} \\
 F_{\perp} &= \frac{\mu F_0 u v}{(u+bu)(v+au)}, \quad \frac{1}{b} = \frac{v}{\mu F_0} \frac{\partial F_{\perp}}{\partial u} \Big|_{u=0}
 \end{aligned} \tag{10}$$

The second-order model (10) completely satisfies all properties 1–7 of the exact integral models (8). But, for the majority of the problems of dynamics, it is sufficient to use the first order model (9). The second-order model (10) is required for a more precise qualitative analysis, for example, for determining the boundaries of the stagnant region and the motion stopping time.

The approximations (9) and (10) hold for positive values of u and v . They can be easily generalized to the case of arbitrary (in sign) velocities u and v by a formal change by absolute values in the denominators of the corresponding expressions.

The use of the friction models based on the Pade' expansions allows one to avoid calculations of multiple integrals over the contact spot, which significantly simplifies their use in problems of dynamics.

The approximate models preserve all properties of the models based on the exact integral expressions and correctly describe the behaviour of the net vector and torque of the friction forces and their first derivatives at zero and infinity. Moreover, the models coefficients can be identified from experiments [8]. Consequently, the models based on Pade expansions may be considered as phenomenological models of the combined dry friction.

CONCLUSIONS

It is developed a dynamically coupled integral dry friction model. It is shown that the distortion in the symmetry of the normal contact stresses distribution in the case of circular contact sites results to the appearance of the friction force component directed along the normal to the trajectory of the mass center of the rubbed solids and, consequently, the mass center trajectory is inclined from the straight line.

To escape the double integrals calculation in the motion equations, the exact integral expressions are replaced by appropriate Pade expansions. Models based on Pade expansions may be considered as phenomenological models of the combined dry friction because their coefficients can be defined from the experiments.

REFERENCES

- [1] Contensou, P. Couplage Entre Frottement de Glissement et Frottement de Pivotement Dans la Th eorie de la Toupie. *In Kreiselprobleme Gyrodynamics: IUTAM Symp. Celerina* (1962):201-216.
- [2] Zhuravlev, V.Ph. The Model of Dry Friction in the Problem of the Rolling of Rigid Bodies. *Prikl. Mat. Mekh.* (1998) **62**(5):762–767.
- [3] Kireenkov, A.A. On the Motion of a Homogeneous Rotating Disk along a Plane in the Case of Combined Friction. *Izv. Akad. Nauk. Mekh. Tverd. Tela.* (2002) **1**:60–67.
- [4] Zhuravlev, V.Ph. and Kireenkov, A.A. Pade Expansions in the Two-Dimensional Model of Coulomb Friction. *Izv. Akad. Nauk. Mekh. Tverd. Tela.* (2005) **2**: 3–14.
- [5] A.A.Kireenkov, A.A. Generalized two dimensional model of the sliding and spinning friction. *Doklady Physics.* (2010), **55**(4):482—486.
- [6] Goryacheva, I.G. *Mechanics of friction interaction*, NAUKA, Moscow, (2001).
- [7] Ivanov, A.P. A dynamically consistent model of the contact stresses in the plane motion of a rigid body. *J. Appl. Math. Mech.* (2009) **73**(2):134-144.
- [8] A.A. Kireenkov A.A. and Semendyaev S.V. and Filatov V.V. Experimental Study of Coupled Two-Dimensional Models of Sliding and Spinning Friction. *Izv. Akad. Nauk. Mekh. Tverd. Tela.* (2010) **6**:192-202

METHOD FOR ESTIMATING PARAMETERS OF COUPLED PROBLEM OF INTERACTION OF GAS FLOWS LOADED BY SOLID PARTICLES WITH SOLIDS

A.V.NENAROKOMOV*, O.M.ALIFANOV*, E.A.ARTIUKHINE†, I.V.REPIN* AND
TITOV D.M.*

* Dept. of Space System Engineering, Moscow Aviation Institute,
4 Volokolamskoe Sh.Moscow, 125993, Russia
email: mai@mai.ru, web page: <http://www.mai.ru>

† CREST UMR CNRS 6000
University of Franche-Comte, Belfort, France

Abstract. The present paper outlines a method of processing and analysing experimental data based on the methodology of the solution of inverse heat transfer problems. An algorithm and the results of the computational and experimental study of heat transfer in the vicinity of the critical point of a specimen in a high-enthalpy particles-loaded flow are presented. One of the main difficulties here is how to determine coefficients of the mathematical model, which provide its adequacy to real processes. Direct measurement of most characteristics of heat transfer is usually impossible, and their theoretical estimates are often far from being true and often contradictory. That is why, the unknown parameters of the heat-balance equation at the external moving boundary of the specimen are determined from the inverse problem of heat transfer, which is solved by the method of iterative regularization. The results of experimental data processing for the interaction of particles-loaded flows with plane surfaces of the cylindrical specimen are also presented as well as the optimal experiment design problems for corresponded experiments.

NOMENCLATURE

b	thickness of a specimen
f	experimental measurements
g	increment of unknown function
J	minimized (residual) functional
M	number of temperature measurements
\bar{P}	unknown (desired) parameters vector
γ	descent step

- δ measurements error
- θ sensitivity function
- σ deviation of measurements

1 INTRODUCTION

With the development of new materials and processes, identification of system parameters has been the prime goal for understanding and defining these systems. Under development is an approach to a study of high temperature thermal processes, based on the principles of identification of non-linear systems with distributed parameters. One of the main difficulties is how to determine the coefficients of a model, which is adequate to real processes. Methods based on solving boundary inverse heat conduction problems are also widely used in experimental investigations of the thermal interaction between solids and the environment.

Presently, heat and mass transfer in heterogeneous media are being studied intensively [1], [2], [3]. This interest is associated with the important practical applications of the results of these investigations in aerospace technology, nuclear power engineering, turbine manufacture, chemical technology, and other fields.

The present work outlines a method of processing and analyzing experimental data based on the methodology of the solution of inverse heat transfer problems. The experiments are conducted in a gas-dynamic stand specially designed for modeling particles-loaded flows. Solid particles are introduced into the gas flow through a special particle source (Figure 1). The uniformity of the particle distribution over the flow cross-section and the steady flow rate of the particles during the experiment are ensured by a special supply system. The particle velocities are calculated as described in [4]. Experimental investigation of the thermal interaction of the particles-loaded flow with the material is conducted at a special calorimetric module (Figure 2).

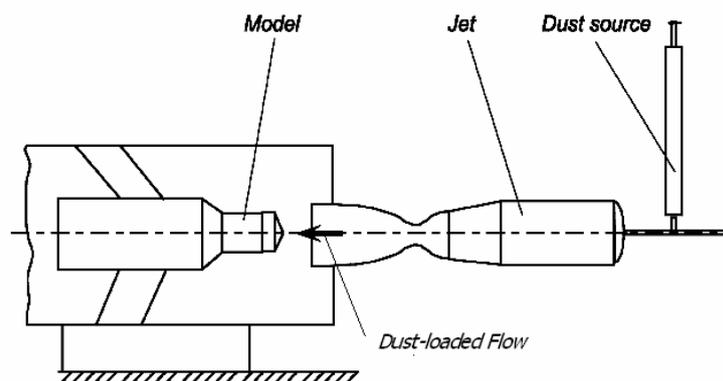


Figure 1: Experimental facilities

The structure of the specimen's models (Figure 3) permits the use of a one-dimensional mathematical model of thermal conduction. The heat transfer in the calorimeter is described by a homogeneous heat-conduction equation

$$C(T) \frac{\partial T}{\partial \tau} = \frac{\partial}{\partial x} \left(\lambda(T) \frac{\partial T}{\partial x} \right) \quad (1)$$

$$T = (x, \tau), \quad x \in (0, b(\tau)), \quad \tau \in (\tau_{\min}, \tau_{\max}]$$

where

$$b(\tau) = b(\tau_{\min}) - \int_{\tau_{\min}}^{\tau} V_{er} d\tau$$

is the coordinate of the specimen external surface heated by a two-phase flow and undergoing erosion, and V_{er} is the linear rate of erosion. τ_{\min}, τ_{\max} are the times at which the experiment begins and ends; $C(T)$ is the volume heat capacity; λ is the thermal conductivity. The initial temperature and boundary condition at the internal boundary are known, and take the form

$$T(x, \tau_{\min}) = T_0(x), \quad x \in [0, b(\tau_{\min})] \quad (2)$$

$$T(0, \tau) = T_1(\tau), \quad \tau \in (\tau_{\min}, \tau_{\max}] \quad (3)$$

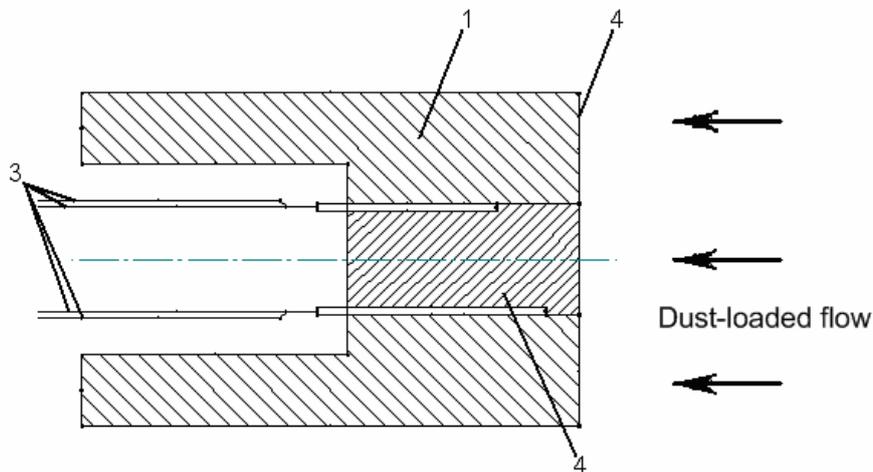


Figure 2: Experimental module: 1 – model, 2 – sensor, 3 – thermocouples, 4 – model surface

At the external boundary, (exposed to the particles-loaded flow) the following conditions are considered

$$-\lambda(T)\left(\frac{\partial T}{\partial x}\right)_\omega = q_2(\tau), \quad \tau \in (\tau_{\min}, \tau_{\max}] \quad (4)$$

where

$$\left(\frac{\partial T}{\partial x}\right)_\omega = \frac{\partial T(b(\tau), \tau)}{\partial x}$$

$$q_2(\tau) = H \text{ (parameters of particles - loaded flow)}$$

and at this boundary ($x = b(\tau)$) the condition of heat balance can be considered in the form [5]

$$-\lambda(T)\left(\frac{\partial T}{\partial x}\right)_\omega = q_{conv} + q_{turb} + q_r + q_{acc} \quad (5)$$

where q_2 is the heat flux that penetrates to the model in the vicinity of the critical point, q_{conv} is the external convective heat flux, q_{turb} is the heat flux resulting from additional turbulence caused by solid particles, q_r is the additional heat flux resulting from an increase in the surface roughness, as a result of the contact with solid particles, and q_{acc} is the heat flux generated as a result of the accommodation of kinetic energy of solid particles at the specimen surface.

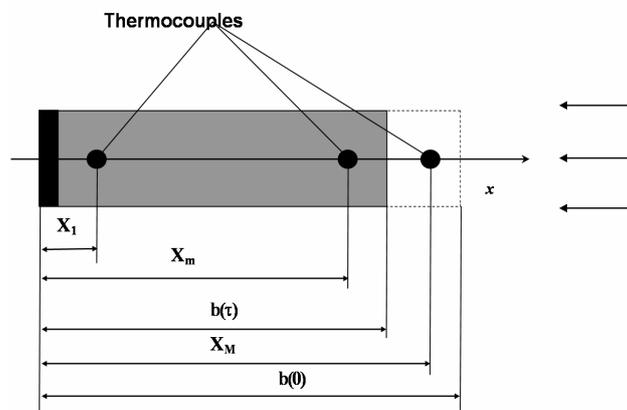


Figure 3: Scheme of sensor.

Relations for q_{conv} and q_{acc} have the form

$$q_{conv} = \frac{\alpha}{C_p} (J_e - J_\omega) \quad (6)$$

$$q_{acc} = a_{acc} \frac{G_p V_p^2}{2} \quad (7)$$

where α is the heat transfer coefficient; J_e and J_ω are the enthalpy of the flow at the recovery temperature and the temperature of the external surface of the model, respectively; a_{acc} is the accommodation coefficient; G_p is the mass rate of the solid phase; and V_p is the rate of the particles.

The effect of additional turbulence can be taken into account in the form [5]

$$q_{turb} = \frac{\alpha}{C_p} (J_e - J_\omega) a_1 \left(\frac{G_p + G_{er}}{G_g} \right)^{n_1} \quad (8)$$

where a_1 and n_1 are non-dimensional coefficients, G_g is the mass rate of the injected gaseous phase, and G_{er} is the mass rate of erosion of the model, defined by

$$G_{er} = \rho V_{er} \quad (9)$$

where ρ is the density of the material of the model and $n_1 = 1$ for a particle concentration of less than 1%.

The increase in the intensity of heat transfer due to the surface roughness is associated with the Reynolds criterion and the relation between value h of the surface roughness and thickness θ of the loss of momentum [6]

$$q_r = \frac{\alpha}{C_p} (J_e - J_\omega) f_r \frac{h}{\theta} \text{Re}_0^{0.5} \quad (10)$$

where f_r is the roughness coefficient. According to [6], the thickness of the loss of momentum is defined by the equation

$$\theta = 0.245 \sqrt{\frac{\nu_e}{\beta}} \left(1.4 - 0.4 \frac{T_\omega}{T_e} \right) \quad (11)$$

where ν_e is the kinematics viscosity of the gas and

$$\beta = \frac{4V_g}{3\pi R_T} \sqrt{\bar{\rho}(2 - \bar{\rho})} \quad (12)$$

is the gradient of the gas in the vicinity of the critical point. Here, V_g is the velocity of the gas flow, R_T is the radius of the specimen, and $\bar{\rho} = \rho_g / \rho_2$ is the ratio of the density of the gas in the flow to its density behind the shock wave. This ratio and the Reynolds number are given

by the formulas

$$\bar{\rho} = \frac{(k-1)}{(k+1)} + \frac{2}{(k+1)} \frac{1}{M_\infty^2} \quad (13)$$

$$\text{Re}_0 = \frac{\rho_g V_g R_T}{\mu_g} \quad (14)$$

where k is the adiabatic exponent and μ_g is the dynamic viscosity of the gas. Under these assumptions, the mathematical model of nonsteady-state mass transfer caused by the interaction of materials with two-phase flows is covered by equations (1)-(14).

In general, an inverse problem formulated this way has no unique solution. To provide the uniqueness of the solution, we proposed the simultaneous analysis of the data of several nonstationary experiments under different conditions of thermal interaction of the specimens with particles-loaded flow. The method based on the variations of particle concentration in the forward flow is most appropriate for changing the loading mode.

The characteristics $\alpha/C_p, a_1, f_r, a_{acc}$ are assumed to be constant in the mathematical model described by equations (1)-(14). In this case, the heat-balance equation (4) can be written in the form

$$-\lambda(T) \left(\frac{\partial T}{\partial x} \right)_\omega = H(\bar{P}, T_\omega(\tau), \tau) \quad (15)$$

where $\bar{P} = \{P_1, P_2, P_3, P_4\}^T$ is the vector of unknown characteristics whose components are defined by the relations $P_1 = \alpha/C_p, P_2 = a_1, P_3 = f_r, P_4 = a_{acc}$ and H is a function of known form. As a result, the inverse problem, which consists in determining the characteristics of heat transfer on the surface of a material, is formulated as follows. It is necessary to determine the vector of unknown parameters satisfying the mathematical model (1)-(15), using the data on additional internal temperature measurements

$$T_n(X_{m,n}, \tau) = f_{m,n}(\tau), \quad m = 1, \dots, M_n, \quad n = 1, \dots, N \quad (16)$$

where n is the number of the experiment, N is the total number of simultaneously analysed experiments, and M_n is the number of thermocouples in the n -th experiment. The functions $b_n(\tau), C(T), \lambda(T), \tau_{\min}^n, \tau_{\max}^n, T_{0,n}(x)$ and $T_{1,n}(\tau)$, as well as the form of the function H , are known.

2 IDENTIFICATION OF THE MATHEMATICAL MODEL

The algorithm used to solve the inverse problem is constructed on the basis of the gradient

method of minimization of the mean square functional of discrepancy (see Alifanov, Artyukhin et al. (1995)). In the case under consideration, the discrepancy functional is composed for the entire body of simultaneously analysed experiments and has the form

$$J(\bar{P}) = \sum_{n=1}^N \sum_{m=1}^{M_n} \int_{\tau_{\min}^n}^{\tau_{\max}^n} (T_n(X_{m,n}, \tau) - f_{m,n}(\tau))^2 d\tau \quad (17)$$

where $T_n(x, \tau)$ is the solution of the boundary-value problem equations (1)-(15) for the n -th experiment.

The iteration process of successive approximations of the unknown vector \bar{P} is constructed in accordance with the following procedure:

1. The initial approximation of the vector of required parameters is preset:

$$\bar{P}^0 = \{P_1^0, P_2^0, P_3^0, P_4^0\}^T.$$

2. The value of the required vector at the next iteration is calculated according to the formula

$$\bar{P}^{s+1} = \bar{P}^s + \Delta\bar{P}^s, \quad s = 0, 1, \dots \quad (18)$$

where s is the iteration number. The increment $\Delta\bar{P}^s$ is determined from the condition

$$\frac{\partial J(\bar{P}^s + \Delta\bar{P}^s)}{\partial \Delta\bar{P}^s} = 0 \quad (19)$$

3. The condition $J\{\bar{P}\} = \delta^2$ of stopping of the iterative process is verified, where δ^2 is the integral error of temperature measurements:

$$\delta^2 = \sum_{n=1}^N \sum_{m=1}^{M_n} \int_{\tau_{\min}^n}^{\tau_{\max}^n} \sigma_{m,n}(\tau)^2 d\tau \quad (20)$$

where $\sigma_{m,n}(\tau)$ is the deviation of $\{m, n\}$ -th measurement. If this condition is satisfied, then the iterative process is terminated. In the opposite case, the procedure of successive approximations is continued.

Following the approach suggested to calculate increments of the desired vector $\Delta\bar{P}^s$, the minimized functional (17) at the $(s+1)$ -th iteration can be presented as

$$J(\bar{P}^{s+1}) = \sum_{n=1}^N \sum_{m=1}^{M_n} \int_{\tau_{\min}^n}^{\tau_{\max}^n} (T_n^s(X_{m,n}, \tau) + \Delta T_n^s(X_{m,n}, \tau) + o(\Delta P^2) - f_{m,n}(\tau))^2 d\tau \quad (21)$$

where

$$\Delta T_n^s(X_{m,n}, \tau) = \sum_{k=1}^4 \Delta T_{n,k}^s(X_{m,n}, \tau) = \sum_{k=1}^4 \frac{\partial T_n^s(X_{m,n}, \tau)}{\partial P_k} \Delta P_k = \sum_{k=1}^4 \theta_{n,k}(X_{m,n}, \tau) \Delta P_k \quad (22)$$

Here the temperature $T_n^s(X_{m,n}, \tau)$ can be determined from the solution of the direct problem (1)-(15) at $\bar{P} = \bar{P}^s$, and $\theta_{n,k}(x, \tau)$, $k = 1, 2, 3, 4$ satisfy the following set of equations

$$C(T) \frac{\partial \theta_{n,k}}{\partial \tau} = \frac{\partial}{\partial x} \left(\lambda(T) \frac{\partial \theta_{n,k}}{\partial x} \right) + \frac{\partial T}{\partial x} \frac{d\lambda}{dT} \frac{\partial \theta_{n,k}}{\partial x} + \left(\frac{\partial^2 T}{\partial x^2} \frac{d\lambda}{dT} + \frac{d^2 \lambda}{dT^2} \left(\frac{\partial T}{\partial x} \right)^2 - \frac{\partial T}{\partial \tau} \frac{dC}{dT} \right) \theta_{n,k} \quad (23)$$

$$\theta_{n,k} = \theta_{n,k}(x, \tau), \quad x \in (0, b^n(\tau)), \quad n = 1, 2, \dots, N, \quad \tau \in (\tau_{\min}^n, \tau_{\max}^n]$$

$$\theta_{n,k}(x, \tau_{\min}^n) = 0, \quad x \in [0, b^n(\tau_{\min}^n)], \quad n = 1, 2, \dots, N \quad (24)$$

$$\theta_{n,k}(0, \tau) = 0, \quad \tau \in (\tau_{\min}^n, \tau_{\max}^n] \quad (25)$$

$$\lambda(T) \frac{\partial \theta_{n,k}(b(\tau), \tau)}{\partial x} - \frac{\partial T}{\partial x} \frac{d\lambda}{dT} \theta_{n,k}(b(\tau), \tau) - \frac{\partial H}{\partial T} \theta_{n,k}(b(\tau), \tau) - \frac{\partial H}{\partial P_k} = 0, \quad (26)$$

$$\tau \in (\tau_{\min}^n, \tau_{\max}^n]$$

Then, using equation (22), a system of linear algebraic equations can be obtained

$$\sum_{l=1}^4 \Delta P_l \sum_{n=1}^N \int_{\tau_{\min}^n}^{\tau_{\max}^n} \theta_{n,l}(X_{m,n}, \tau) \theta_{n,k}(X_{m,n}, \tau) d\tau = \sum_{n=1}^N \sum_{m=1}^{M_n} \int_{\tau_{\min}^n}^{\tau_{\max}^n} \theta_{n,k}(X_{m,n}, \tau) (T_n(X_{m,n}, \tau) - f_{m,n}(\tau)) d\tau, \quad k = 1, 2, 3, 4 \quad (27)$$

from which the increments $\bar{\Delta P}$ can be determined.

3 EXPERIMENTAL APPROVING

In implementing the algorithm described, the value problem is solved by a finite difference method on an implicit four-point scheme. In the numerical solution the direct nonlinear problem is treated by iteration in the coefficients. The approximation of all three boundary value problems is carried out on one and same difference grid, making it possible to achieve error matching. Below we provide results of handling the experimental data obtained during the four-point process of an experimental device. The tests differed from each other by the mass discharge values of solid phase. The particle diameter of the solid phase was $250 \mu m$. The rate of the solid phase particles was $V_p = 1083$ m/sec, the gas rate was $V_g = 1797$ m/sec,

and the mass discharge of gas was $G_g=1490 \text{ kg/(m}^2/\text{sec)}$. Variation of the length of the specimen $b(\tau)$ as a result of erosional destruction in the course of the experiment is shown in Figure 4. The remaining characteristics are shown in Table 1.

Cylindrical specimen of radius 40mm, prepared from copper, had different initial thickness $b(0)$ (Figure 4), as well as an amount and coordinates of thermocouple devices in the transducers, while one of the thermocouples was located on the rear surface, and the remaining ones – at the internal points of the transducers, located in the vicinity of the critical point of the model. The results of thermocouple measurements in all specimens are shown in Figure 3. In solving the inverse problem the experimental data obtained in the six tests were analyzed simultaneously. The temperature value calculated in this case at the points of the thermocouple devices are also given in Figure 5. The results of processing the experimental data are presented in Table 2.

Table 1:Experimental Data

No, of tests	G_p , kg/(m ² /sec)	T_0 , K	V_{er} , kg/(m ² /sec)
1	5.1	1621	2.51
2	6.9	1593	3.52
3	0	1581	0.0
4	7.2	1612	3.91
5	7.8	1579	4.14
6	0	1637	0.0

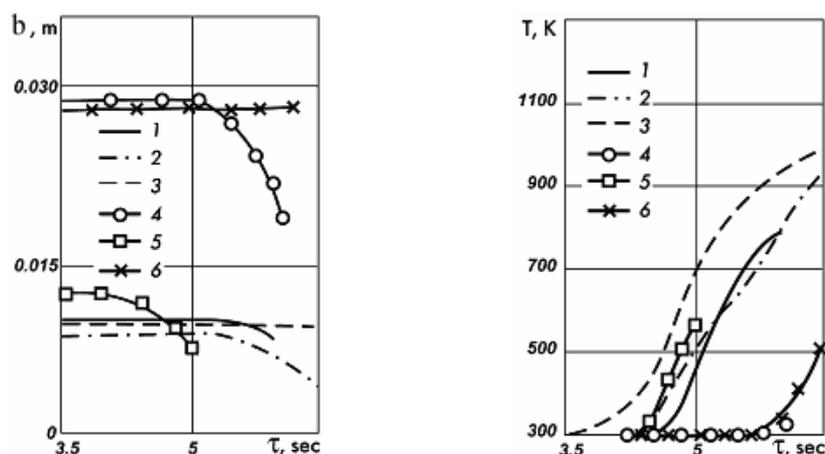


Figure 4: Erosional destruction of specimens and temperature at the left boundaries (6 tests).

Table 2 :Estimated Parameters

Parameters	Estimates
$\frac{\alpha}{C_p}$	2.42 kg/(m ² sec)
a_{acc}	0.71
a_1	45.53
f_r	0.0371

The result of experimental data processing reveals a reasonably good agreement between the measured and calculated temperature values (Figure 5 and Table 3). Note that the following constrains were imposed on the required parameters while solving the inverse problem $\frac{\alpha}{C_p} > 0$, $a_1 \geq 0$, $f_r \geq 0$, and $0 \leq a_{acc} \leq 1$.

Table 3: Estimating errors

No, of tests	J, K^2	$\Delta T_{max}, K$
1	0.28*10 ⁵	22.51
2	0.9*10 ⁴	33.52
3	0.16*10 ⁴	27.03
4	0.12*10 ⁵	23.91
5	0.2*10 ⁵	34.14
6	0.35*10 ⁵	35.31

Determining the function q_2 in the boundary condition in (15) from the results of temperature measurement at several internal points of the specimen $X_m, m=1, \dots, M$, is a well-known boundary inverse heat conduction problem [5]. This problem is solved for each experiment using the method presented in [7] and [8]. The heat fluxes q_2 and temperature of the external calorimeter surface T_w (Figure 6) indicate that, with increase in mass concentration of the incoming particles to 1.1% the heat flux reaching the material at the "cold" wall is twice that in the case of a flow with no particles.

CONCLUSION

In this paper, a sensitivity method formulation is presented for the solving of the inverse problem of interaction materials with particles-loaded high-enthalpy flows. The objective was to investigate the influence of different factors on the intensity of heat transfer. The

implementation of the optimization problem of the inverse formulation is carried out by Newton's method, while the direct and sensitivity problems in each iteration are solved by the finite difference method. The obtained results should be considered to be another step toward the construction of adequate mathematical models that describe the interaction of materials with particles-loaded flows. Further parameter studies and tests on more complicated cases remain to be done in the future to examine more complex mathematical models of heat transfer.

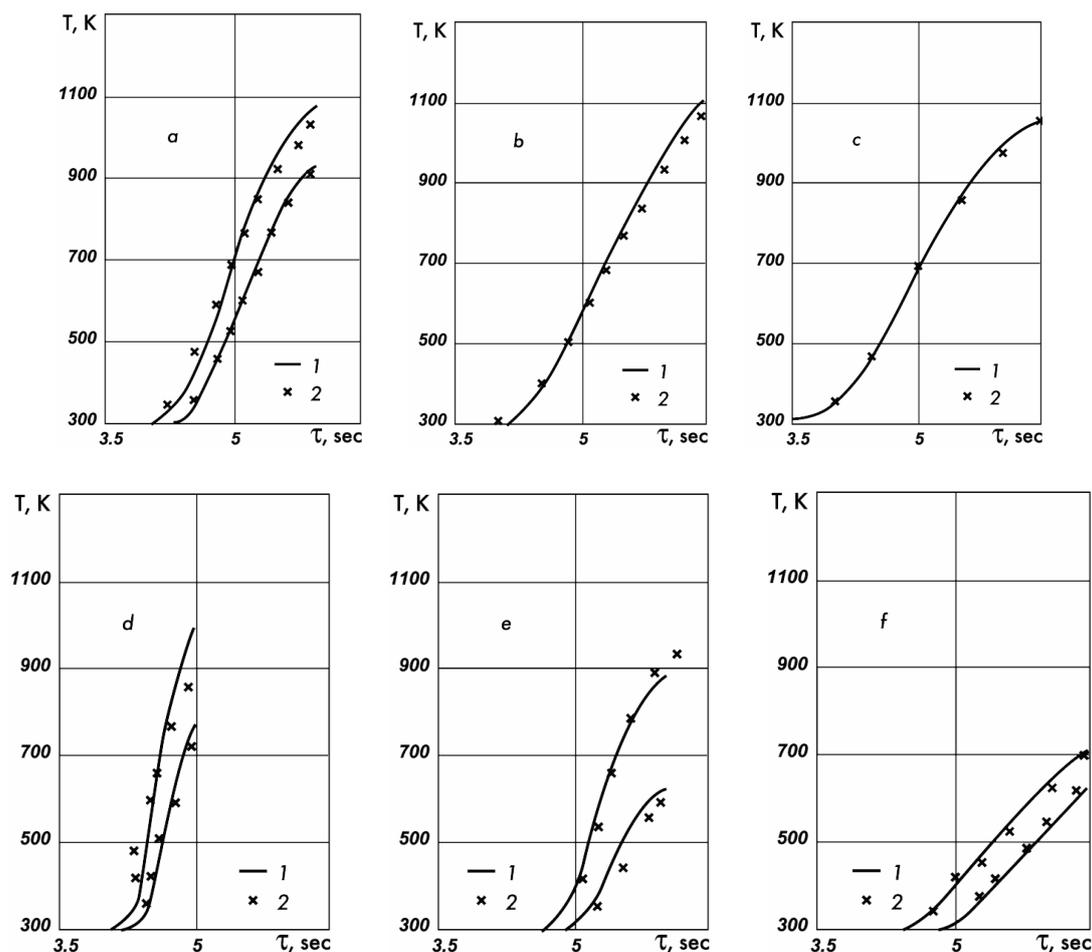


Figure 5: Experimental (1) and calculated (2) temperatures.
 a - test 1, b - test 2, c - test 3, d - test 4, e - test 5, f - test 6.

REFERENCES

- [1] Flecner, W.A. and Watson R.H. *Convective Heating in Dust-Loaded Hypersonic Flow*, AIAA Paper, N.0761.(1973)

- [2] Bakum, B.L. and Komarova G.S. *The Effect of the Dust in the Working Stream of Hypersonic Wind Tunnels on The Results of Heat- Transfer Measurement*, *J. of Eng. Physics*, Vol. 21, pp. 1361-1363. (1963)
- [3] Vasin, A.V., Mikhatulin, D.S., Polezhaev, Yu.V. *Determination of Thethermal State of Material Subjected to Erosive Degradation*, *Journal of Engineering Physics*, Vol. 52, pp. 150-155. (1987)
- [4] Artyukhin, E.A., Killikh, V.E., Nenarokomov, A.V. and Repin, I.V. *Investigation Of Thermal Interaction Of A Material With Two-Phase Flows By The Inverse Problem Methods*, *High Temperature*, Vol. 28, pp. 94-99. (1990)
- [5] Alifanov, O.M., Artyukhin, E.A. and Rumyantsev S.V. *Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Problems*, Begell House, New York/Wallinford (UK). (1995).
- [6] Alifanov, O.M. and Repin, I.V. *Investigation of Heat Transfer in Heterogeneous Flows Through Method of Inverse Problems*, *High Temperature*, Vol. 32, pp. 78-83. (1994)
- [7] Budnik, S.A. and Nenarokomov, A.V. *Optimum Planning of Measurements in Determining the Characteristics of Heat Loading of Bodies with Movable Boundaries*, *High Temperature*, Vol. 35, pp. 453-457. (1997)
- [8] Nenarokomov, A.V., Alifanov, O.M., Artyukhin, E.A. and Repin, I.V. *A Study of Convective Heat Fluxes for Material Interacting with Dust-Loaded Flows by Inverse Problems Method*, *International Journal of Thermal Sciences*, Vol. 43, pp. 825-831. (2004)
- [9] Ritchie, G. S. *Nonlinear Dynamic Characteristics of Finite Journal Bearing*, *Trans. ASME, J. Lub. Tech.*, Vol. 1, No. 3, pp. 375-376. (1983)

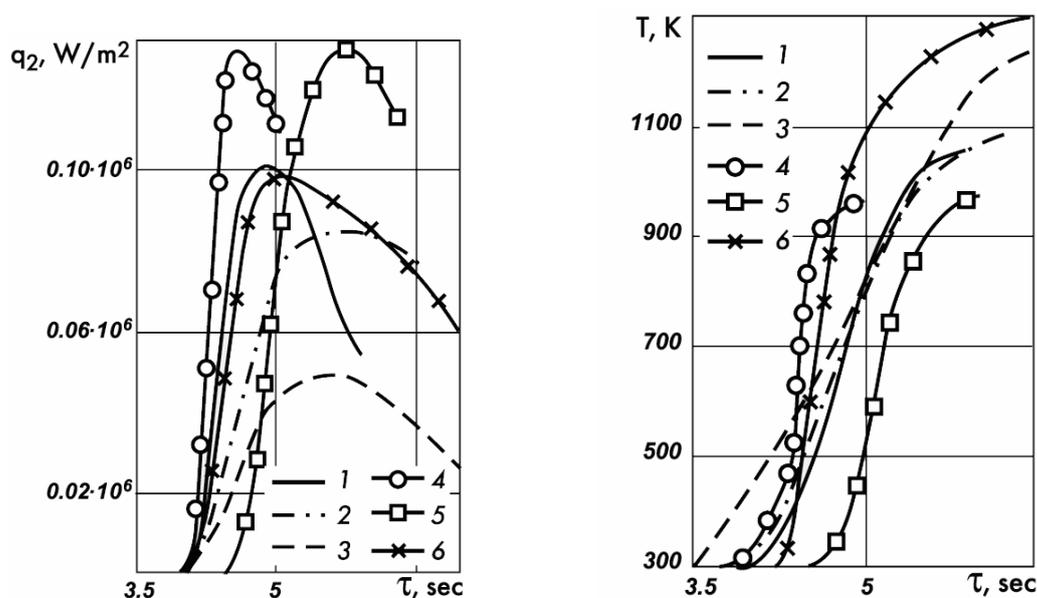


Figure 6: Reconstruction of heat fluxes and surface temperatures (6 tests)

MODELLING OF SIMULTANEOUS INDUCTION HARDENING IN MONOLITHIC FORMULATION

PAVEL KARBAN^{*}, IVO DOLEŽEL[†] AND BOHUŠ ULRYCH^{*}

^{*} Faculty of Electrical Engineering
University of West Bohemia
Univerzitní 16, 306 14 Plzeň
e-mail: {karban, ulrych}@kte.zcu.cz

[†] Faculty of Electrical Engineering
Czech Technical University
Technická 2, 166 27 Praha 6
email: dolezel@fel.cvut.cz

Key words: Induction Hardening, Nonlinear Coupled Problem, Monolithic Formulation, Numerical Analysis, Electromagnetic Field, Temperature Field.

Abstract. Induction hardening of an axisymmetric steel workpiece is modeled. This evolutionary and highly nonlinear process is solved by a fully adaptive higher-order finite element method. Numerical computations are realized in the monolithic formulation, using own code Hermes. All nonlinear dependencies of material properties on temperature are respected. The methodology is illustrated by a typical example whose results are discussed.

1 INTRODUCTION

Hardening of steel bodies is a widely used metallurgical process whose purpose is to bring about local changes in the crystalline structure of its surface layers resulting in their higher hardness. The part of the body to be hardened is first heated somewhat above temperature A_{c3} when steel forms a uniform austenite structure. Then, after eventual equalization of temperatures, the body must intensively be cooled by a suitable quenchant. The result is harder, but more brittle martensite structure of the hardened part. The structure of steel in internal layers remains unchanged.

From the physical viewpoint, modeling of this process is still represents a challenge because it represents a strongly nonlinear and evolutionary coupled problem. It includes generation of magnetic field, production of the Joule losses in the processed object, its heating and consequent cooling accompanied by metallurgical changes in its surface structure. Papers that are aimed at the relevant simulations are rather rare [1–4], and usually respect only the nonlinearity of the magnetic permeability.

The authors present a numerical solution to the problem that takes into account all material nonlinearities (most of the material properties are considered as nonlinear functions of the temperature). The solution itself is performed by a fully adaptive higher-order finite element method in the monolithic formulation and realized by own code Hermes [5].

2 FORMULATION OF THE PROBLEM

Consider an arbitrary axisymmetric steel body, whose surface is to be locally hardened, see Fig. 1. The hardened part is placed in the inductor connected to a current source providing harmonic current of amplitude I and frequency f . These parameters have to be sufficient enough for heating the surface to the prescribed temperature in a reasonable time.

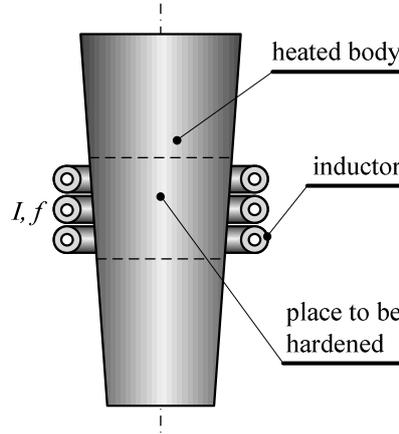


Figure 1: Hardening of axisymmetric bodies

The task is to model the process with the following aims:

- to find the time evolution of the surface temperature in the hardened region,
- to find the time evolution of the cooling process and its velocity, and
- to determine the distribution of the resultant hardness.

3 CONTINUOUS MATHEMATICAL MODEL OF THE PROCESS

The mathematical model of the problem is given by two nonlinear and nonstationary partial differential equations (PDEs) describing distribution of the magnetic and temperature fields. These equations have to be supplemented with correct boundary conditions.

Magnetic field is described by the well-known parabolic equation for magnetic vector potential \mathbf{A} in the form

$$\operatorname{curl}\left(\frac{1}{\mu}\operatorname{curl}\mathbf{A}\right)+\gamma\frac{\partial\mathbf{A}}{\partial t}=\mathbf{J}_{\text{ext}}, \quad (1)$$

where μ denotes the magnetic permeability, γ the electric conductivity and \mathbf{J}_{ext} the vector of the external harmonic current density in the inductor.

But solution to equation (1) is, in this case, practically unfeasible. The reason consists in the deep disproportion between the frequency f (on the order of kHz) of the field current I and time of heating t_h (seconds or tens of seconds). That is why the model was somewhat simplified using the assumption that the magnetic field is harmonic. In such a case it can be described by the Helmholtz equation for the phasor $\underline{\mathbf{A}}$ of the magnetic vector potential \mathbf{A} [6]

$$\operatorname{curl}(\operatorname{curl}\underline{\mathbf{A}})+j\cdot\omega\gamma\mu\underline{\mathbf{A}}=\mu\underline{\mathbf{J}}_{\text{ext}}, \quad (2)$$

where ω is the angular frequency. But the magnetic permeability of ferromagnetic parts is always assigned to the local value of magnetic flux density. Its computation is, in such a case, based on an iterative method.

The conditions along the axis of the device and artificial boundary placed at a sufficient distance from the system are of the Dirichlet type ($\underline{A} = \underline{0}$).

Temperature field is described by the heat transfer equation [7]

$$\operatorname{div}(\lambda \cdot \operatorname{grad}T) = \rho c_p \cdot \frac{\partial T}{\partial t} - p, \quad (3)$$

where λ is the thermal conductivity, ρ the mass density and c the specific heat (all of these parameters are temperature-dependent functions). Finally, symbol p denotes the time average internal volumetric sources of heat that generally consist of the volume Joule losses p_J due to eddy currents and magnetization losses p_m . Thus, we can put

$$p = p_J + p_m, \quad (4)$$

where

$$p_J = \frac{|\underline{J}_{\text{eddy}}|^2}{\gamma}, \quad \underline{J}_{\text{eddy}} = \mathbf{j} \cdot \omega \gamma \underline{A}, \quad (5)$$

while p_m (provided that they are considered) are determined from the known measured loss dependence $p_m = p_m(|\underline{B}|)$ for the used steel (magnetic flux density \underline{B} in every element in this model is harmonic).

During the process of cooling the inductor is switched off, so that the term determining the internal volumetric losses p in (2) vanishes.

The boundary conditions take into account convection and radiation, but their particular application depends on the case solved.

All material parameters (μ , γ , λ , ρc_p) occurring in (2) and (3) are nonlinear functions of temperature. As the temperature rise of the heated body in the course of the process ranges from 800–1000 °C, these nonlinearities cannot be neglected, because the error of computation could reach an unacceptable value.

4 NUMERICAL SOLUTION

The numerical solution of the problem is performed by a fully adaptive higher-order finite element method (*hp*-FEM). It is a modern version of the finite element method combining finite elements of variable size (h) and polynomial degree (p) in order to obtain fast exponential convergence [8].

The automatic adaptation of *hp*-meshes significantly differs from the adaptivity in standard methods of this kind. This implies that traditional error estimates (one number per element) do not provide enough information to guide *hp*-adaptivity. One needs a better knowledge of the distribution of the error function derived from the difference between the exact solution u

and approximate solution u' . This error function may be expressed, for example, by the H^1 norm

$$\|e\|_{H^1} = \left| \int_{\Omega} (\delta^2 + (\text{grad } \delta) \cdot (\text{grad } \delta)) d\Omega \right|^{1/2} \quad (6)$$

where $\delta = |u - u'|$ is the local error of solution obtained in a particular step of the hp -process and Ω denotes the definition area of the problem. In principle, error δ could be obtained from the estimates of higher derivatives, but this approach is not very practical. Usually it is easier to use a reference solution, i.e., an approximation u_{ref} , which is at least by one order more accurate than u' . The hp -adaptivity is then guided by an a posteriori error estimate of the form $\delta = u_{\text{ref}} - u'$. More details on automatic hp -adaptivity on meshes with arbitrary level hanging nodes can be found in [9].

At each time level, optimal meshes are obtained automatically by independent adaptive processes. They change in time as the solution changes, respecting different features of particular fields. This is possible as a result of our own multi-mesh technique that allows us solving multiphysics problems monolithically, even though each physical field is discretized on a geometrically different mesh. Thus, our approach leads to a significant reduction of the size of the discrete problem and speeds up the whole computation [10, 11]. In practical computations all physical fields are solved simultaneously, in a monolithic formulation

Our own numerical software Hermes2D [5] was used for the computation. It is capable of all the features mentioned above, such as the higher-order finite element method, automatic adaptivity on hp -meshes or assembling the stiffness matrix on geometrically different meshes.

5 ILLUSTRATIVE EXAMPLE

The methodology will be illustrated by an example of induction hardening of the chuck head of a box-column drilling machine. Its arrangement is depicted in Fig. 2. The chuck is made of steel NZ3. The aim of the process is to harden just the low thin part of the body.

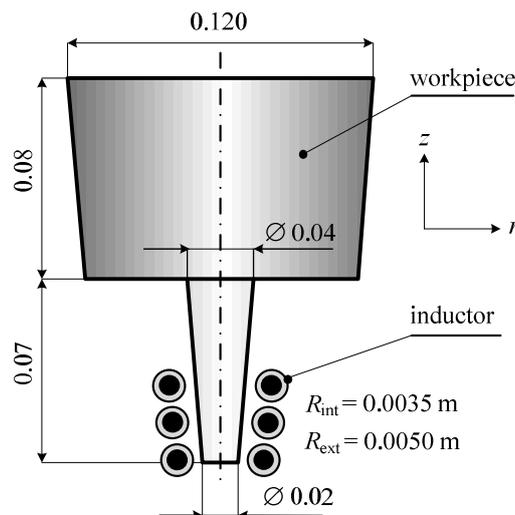


Figure 2: Hardening of axisymmetric bodies (all dimensions in mm)

The inductor (made of a massive hollow copper conductor of circular cross section cooled by water) carries current of density $J_{\text{ext}} = 2 \times 10^7 \text{ A/m}^2$ and frequency $f = 10 \text{ kHz}$. In the course of heating, the process is realized in air of temperature $T_{\text{ext}} = 20 \text{ }^\circ\text{C}$ and the generalized coefficient of convective heat transfer (respecting also radiation) $\alpha_{\text{air}} = 20 \text{ W/m}^2\text{K}$. In the process of cooling by spraying water of temperature $T_{\text{wat}} = 10 \text{ }^\circ\text{C}$, $\alpha_{\text{cool}} = 1200 \text{ W/m}^2\text{K}$. The hardening temperature A_{c3} for the considered steel is $873 \text{ }^\circ\text{C}$ and the final temperature after cooling characterized by full martensite structure of the hardened layers $T_M = 100 \text{ }^\circ\text{C}$. The most important nonlinear characteristics of steel NZ3 are depicted in Figs. 3–7.

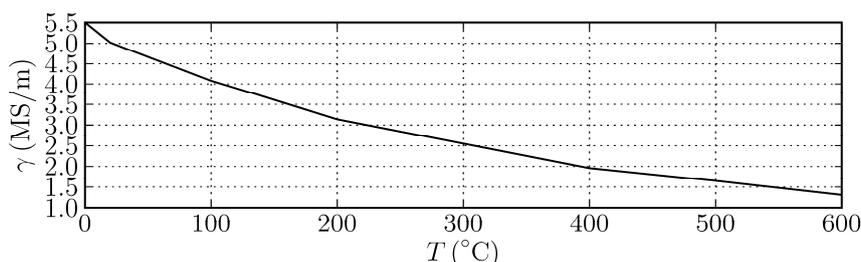


Figure 3: Temperature dependence of electrical conductivity γ (steel NZ3)

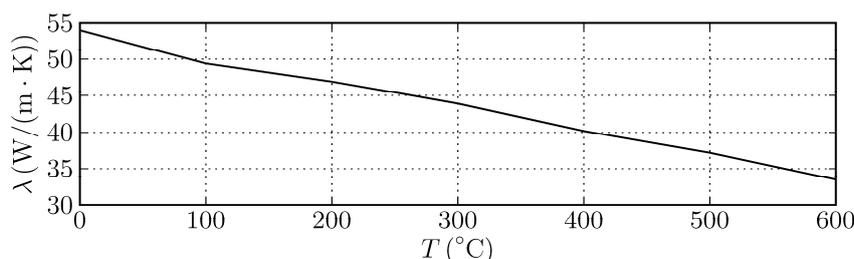


Figure 4: Temperature dependence of thermal conductivity λ (steel NZ3)

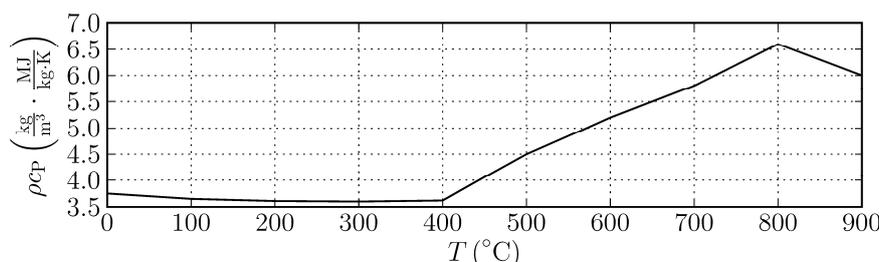


Figure 5: Temperature dependence of heat capacity ρ_{c_p} (steel NZ3)

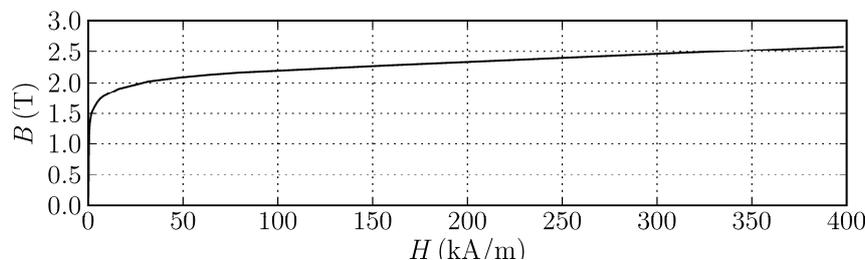


Figure 6: Magnetization characteristic of steel NZ3 (for $T = 20 \text{ }^\circ\text{C}$)

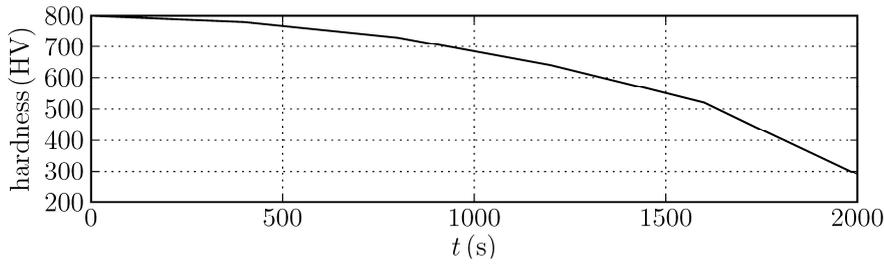


Figure 7: Hardness in HV as a function of the time of cooling

Magnetic permeability μ_r is a function of magnetic flux density and temperature. For every type of steel it must be found experimentally, which usually represents a serious difficulty. We introduce, therefore, an assumption that $\mu_r(B, T) = \mu_r(B, T_0) \cdot \varphi(T)$, where $\mu_r(B, T_0)$ is the dependence of relative permeability on magnetic flux density B at a given temperature T_0 (for example, $T_0 = 20^\circ\text{C}$) and function φ is given by the relation

$$\begin{aligned} \text{for } T_0 \leq T \leq T_C \quad \varphi(T) &= a - bT^2, \\ \text{for } T_C \leq T \quad \varphi(T) &= \frac{1}{\mu_r(B, T_0)}. \end{aligned}$$

Here

$$a = \frac{\mu_r(B, T_0)T_C^2 - T_0^2}{\mu_r(B, T_0)(T_C^2 - T_0^2)}, \quad b = \frac{\mu_r(B, T_0) - 1}{\mu_r(B, T_0)(T_C^2 - T_0^2)}$$

and T_C is the Curie temperature. For steel NZ3 $T_C = 800^\circ\text{C}$ and the function $\varphi(T)$ is depicted in Fig. 8.

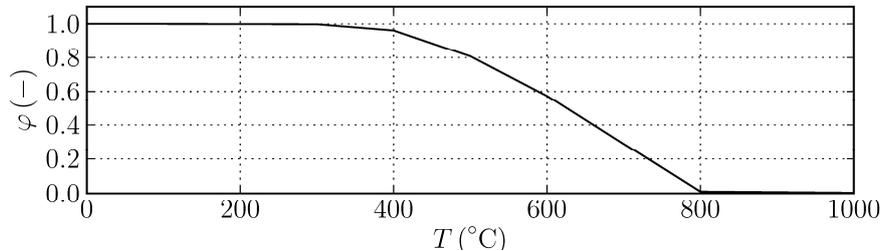


Figure 8: Function $\varphi(T)$ for the temperature correction of magnetic permeability

The discretization mesh for electromagnetic field was composed of about 27000 elements mostly of the second and third orders, while the mesh for temperature field (that was calculated only in the workpiece and turns of the inductor) about 20 000 elements, mostly of the third and fourth orders. After several tests we accepted the time step $\Delta t = 0.25$ s. The complete computations of the process ranging to 78 s (when the surface temperature drops to the value of $T_M = 100^\circ\text{C}$) took about 3 hours.

Figure 9 shows the discretization mesh in the region of the part to be hardened including the three turns of the inductor.

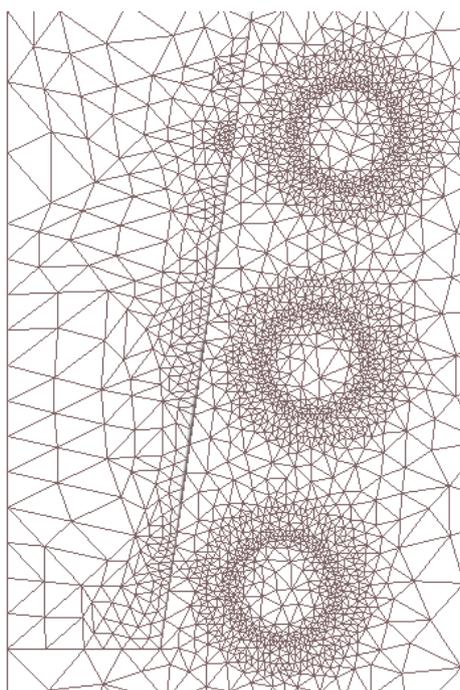


Figure 9: A part of the discretization mesh in the hardened region

Figure 10 shows the surface temperature of the hardened part (measured from the bottom edge of the body in Fig. 2) at several time levels. It is obvious that the austenitization temperature A_{c3} was only exceeded along the first 25 mm of the body.

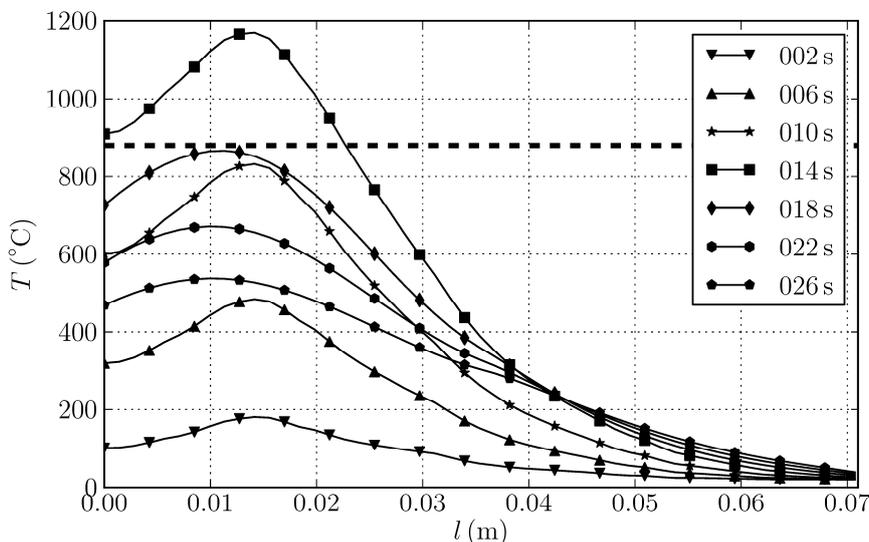


Figure 10: Distribution of the surface temperature of the hardened part (from the bottom edge) at several time levels

Figure 11 depicts the time evolution of temperature at selected point on the surface of the body. After 13 seconds we can see a sharp peak brought about by ending of the process of heating (the inductor is switched off) and immediate start of the process of intensive cooling.

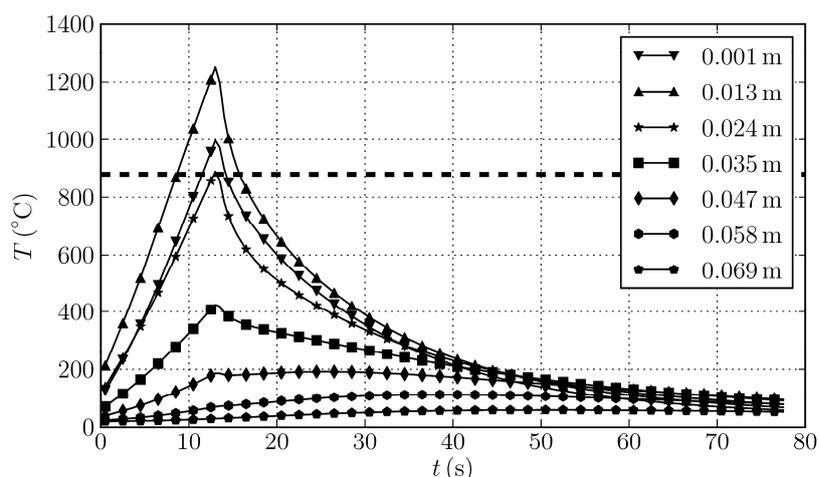


Figure 11. Time evolution of the surface temperatures at selected points of the workpiece measured from its bottom

Finally, Fig. 12 shows the distribution of temperature in the body after 10 s of heating.

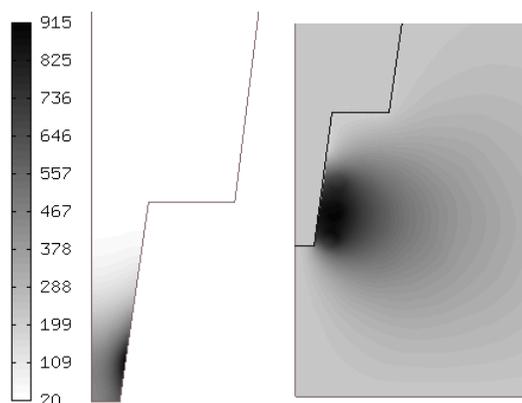


Figure 12. Distribution of the temperature in the workpiece after 10 s of heating

The time of cooling the surface from the austenite temperature A_{c3} to the martensite temperature T_M is about 64 s. For this value we obtain from Fig. 7 the resultant hardness about 785 HV.

6 CONCLUSION

The numerical results are realistic and correspond to typical values obtained by experiments. The computations are relatively fast (tens of minutes), but the group plans their further acceleration using parallelization of selected parts of the code (assembly of the stiffness matrix etc.).

7 ACKNOWLEDGMENT

This work was supported by the European Regional Development Fund and Ministry of Education, Youth and Sports of the Czech Republic under the project No.

CZ.1.05/2.1.00/03.0094: Regional Innovation Centre for Electrical Engineering (RICE), by Grant project GACR P102/11/0498 and project MSMT KONTAKT MEB051041.

REFERENCES

- [1] Zgraja, J. Simulation of induction hardening of flat surfaces of moving massive elements. *International Journal of Materials and Product Technology* (2007) **29**, No.1–4: 103–123.
- [2] Cajner, F., Smoljan, B. and Landek, D. Computer simulation of induction hardening. *Journal of Materials Processing Technology* (2004) **157–158**: 55–60.
- [3] Karban, P. and Donátová, M. Continual induction hardening of steel bodies. *Math. Comput. Simul.* (2010) **80**: 1771–1782.
- [4] Candeo, A., Bocher, P. and Dughiero, F. Multiphysics modeling of induction hardening of ring gears for the aerospace industry. *Proc. CEFC (Electromagnetic Field Computation)* (2010), Chicago, USA, DOI 10.1109/CEFC.2010.5481552.
- [5] <http://hpfem.org/>.
- [6] Stratton, A. J. *Electromagnetic theory*. McGraw Hill, NY, 2007.
- [7] Holman, J. P. *Heat transfer*. McGraw Hill, NY, 2002.
- [8] Šolín, P., Segeth, K. and Doležel, I., *Higher-order finite element methods*, Chapman & Hall/CRC Press, FL, 2003.
- [9] Šolín, P., Červený, J. and Doležel, I. Arbitrary-level hanging nodes and automatic adaptivity in the hp-FEM. *Math. Comput. Simul.* (2008) **77**: 117–132.
- [10] Šolín, P., Červený, J., Dubcová, L. and Andrš, D. Monolithic discretization of linear thermoelasticity problems via adaptive multimesh hp-FEM. *J. Comput. Appl. Math.* (2010) **234**, No. 7: 2350–2357.
- [11] Šolín, J., Dubcová, L. and Kruis, J. Adaptive hp-FEM with dynamical meshes for transient heat and moisture transfer problems. *J. Comput. Appl. Math.* (2010) **233**, No. 12: 3103–3112.

NUMERICAL SIMULATION OF AILERON BUZZ USING AN ADAPTIVE-GRID COMPRESSIBLE FLOW SOLVER FOR DYNAMIC MESHES

Giuseppe Forestieri, Alberto Guardone, Dario Isola, Filippo Marulli and
Giuseppe Quaranta

Dipartimento di Ingegneria Aerospaziale,
Politecnico di Milano,
Via La Masa 34, 20156, Milano, Italy.
e-mail: alberto.guardone@polimi.it, www.aero.polimi.it/flowmesh

Key words: Arbitrary Lagrangian Eulerian, Adaptive mesh, Fluid Structure Interaction, Aileron Buzz

Abstract. The paper presents numerical results from a novel scheme for the solution of the flow equations in two dimensional domains by an Arbitrary Lagrangian Eulerian formulation able to cope with deforming and adaptive two dimensional grids without recurring to any explicit interpolation scheme. The method is applied to the investigation of a classical transonic aeroelastic instability phenomenon: the aileron buzz. By resorting to deforming and adaptive grids, the method allows to highlight the dependency of the aeroelastic stability boundaries on the mesh spacing.

1 INTRODUCTION

The investigation of aeroelastic stability boundaries by means of Fluid Structure Interaction (FSI) analysis is becoming very popular for the preliminary and verification phases of new aircraft design [2, 14], and it is currently denominated Computational Aeroelasticity (CA). By resorting to CFD models based on Euler or Reynolds-Averaged Navier-Stokes (RANS) equations, these approaches avoid any unduly simplification in the computation of fluid flow unsteady forces, allowing to keep into account also the effect of shock-waves, and flow separations, usually neglected by classical potential approaches.

To obtain reliable results, appropriate meshes of the fluid domain must be used. What should be a reliable mesh for a static fluid flow simulation around an aircraft could be considered a question that received an answer in the literature. The same cannot be said if unsteady flow simulations in transonic flow fields are considered, as those that are of paramount importance for CA cases. The paper investigates the dependency of the stability boundary from the grid spacing showing how simulation based on time-

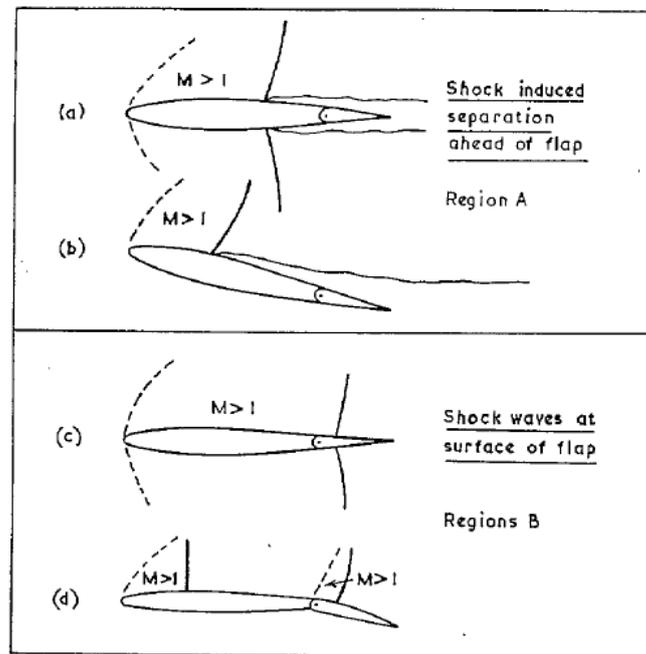


Figure 1: Aerodynamic conditions for which buzz may occur; taken from Ref. [11].

adaptive mesh may improve significantly the prediction capability of numerical approaches to aeroelasticity.

In the present work it has been decided to start tackling this problem by analyzing the dependency from the grid of the numerically evaluated stability boundary of a simple transonic aeroelastic problem characterized by a single structural degrees of freedom: the aileron buzz. This is an instability involving the interaction of a single structural degree of freedom, associated to the aileron rotation about its hinge, with unsteady aerodynamic forces caused by strong shock waves dwelling close to the hinge axis. The instability may evolve into self-sustained Limit Cycle Oscillations [10, 11]. Two principal classes of buzz have been shown during experimental campaigns. Following the classification proposed by Lambourne [10] (see figure 1), the Type A is caused by interaction of the shock-waves with the boundary layer, while Type B results from to the interaction of the shock-waves with the aileron movement without significant intervention of the boundary layer. This latter case is commonly denominated non-classical. Type A is characterized by shock waves positioned ahead of the hinge line, while for Type B phenomena the shock waves are at the flap surface, so they appear at higher Mach numbers than Type A. If a numerical approach is adopted to study the buzz problem, the Euler equations can be considered a good description of the fluid behaviour only for the Type B phenomena since no interaction with the boundary layer is necessary to capture the unsteady forces. Consequently, this paper investigates the Type B buzz taking as a reference the experimental work done

by Lambourne [10]. For numerical analysis of Type A phenomena see [3] and references therein.

The simulation of the buzz phenomenon requires one to perform unsteady simulations of the flowfield while the computational domain is continuously changing its shape to account for the flap movement. So Arbitrary Lagrangian Eulerian (ALE) formulation joined with appropriate mesh update scheme must be considered. The innovative two-dimensional numerical scheme for the compressible CFD equations on dynamic meshes proposed in [9, 13] has been used. It allows to perform computations on moving meshes with adaptation, which is required to preserve the mesh spacing for large boundary displacement.

The paper is organized as follow. The second section briefly presents the numerical approach used for the analysis of fluid flow equations with movable and adaptive grids. The third section shows the detail of the FSI models together with the partitioned strategy used to integrates the two-domain problem. Finally, the numerical results section shows the result obtained with different grids with variable refinement levels, and those obtained by using the adaptive approach.

2 FINITE VOLUME ALE SCHEME

The governing equations for a compressible inviscid fluid in two spatial dimensions are provided by the well-known Euler equations in an Arbitrary Lagrangian Eulerian (ALE) framework [6], namely,

$$\frac{d}{dt} \int_{\mathcal{C}(t)} \mathbf{u} + \oint_{\partial\mathcal{C}(t)} [\mathbf{f}(\mathbf{u}) - \mathbf{u} \mathbf{v}] \cdot \mathbf{n} = 0, \quad \forall \mathcal{C}(t) \subseteq \Omega(t), \quad (1)$$

completed by suitable initial and boundary conditions [8]. In Eq. (1) $\mathbf{u} = (\rho, \mathbf{m}, E^t)^T$, is the vector unknown of the density ρ , momentum vector \mathbf{m} , and total energy per unit volume E^t and the flux function is defined as

$$\mathbf{f}(\mathbf{u}) = (\mathbf{m}, \mathbf{m} \otimes \mathbf{m} / \rho + P(\mathbf{u}) \mathbf{l}, [E^t + P(\mathbf{u})] \mathbf{m} / \rho)^T$$

where \mathbf{l} is the 2×2 identity matrix and P is the pressure. The term $\mathbf{u}(\mathbf{v} \cdot \mathbf{n})$ takes into account the flux contribution due to the movement of the boundary of the control volume $\partial\mathcal{C}(t)$ with normal vector $\mathbf{n}(t)$.

A standard node-centered finite volume scheme is used to discretize the governing equation. [9] As shown in figure 2, a non overlapping set of cells \mathcal{C}_i is taken to discretize Ω and the unknown \mathbf{u} is approximated over \mathcal{C}_i by its average value $\mathbf{u}_i = \mathbf{u}_i(t)$. Eq. (1) can be rewritten

$$\frac{d}{dt} [V_i \mathbf{u}_i] = \sum_{k \in \mathcal{K}_{i,\neq}} \Phi_{ik}(\mathbf{u}_i, \mathbf{u}_k, \nu_{ik}, \boldsymbol{\eta}_{ik}) + \Phi_i^\partial(\mathbf{u}_i, \nu_i, \boldsymbol{\xi}_i) \quad \forall i \in \mathcal{K} \quad (2)$$

where Φ_{ik} and Φ_i^∂ are the domain and boundary numerical fluxes, $\mathcal{K}_{i,\neq}$ is the set of nodes connected to i by an edge, $\boldsymbol{\eta}_{ik}$ and $\boldsymbol{\xi}_i$ are the normal vectors integrated along the interface

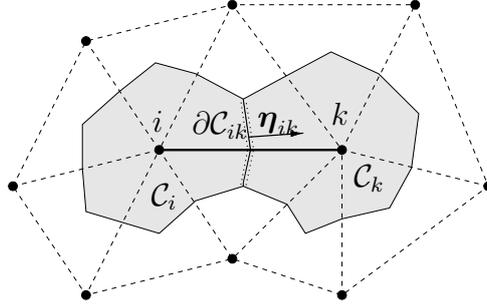


Figure 2: Edge associated with the finite volume interface $\partial\mathcal{C}_{ik} = \partial\mathcal{C}_i \cap \partial\mathcal{C}_k$ and metric vector $\boldsymbol{\eta}_{ik}$ (integrated normal) in two spatial dimensions. The two shaded regions are the finite volumes \mathcal{C}_i and \mathcal{C}_k ; dashed lines indicate the underlying triangulation.

and ν_{ik} and ν_i are the integrated normal velocities. A high-resolution expression for the integrated numerical flux [15] is used in the present work based on the Total Variation Diminishing (TVD) approach.

In order to ensure the conservativity of the scheme the well known Geometric Conservation Law [5] has to be satisfied. This is often achieved by choosing the integrated interface velocities as the time derivative of the volume swept by the interface, namely

$$\frac{dV_{i,ik}}{dt} = \int_{\partial\mathcal{C}_{ik}} \mathbf{v} \cdot \mathbf{n} \quad \text{and} \quad \frac{dV_{i,\partial}}{dt} = \int_{\partial\mathcal{C}_i \cap \partial\Omega} \mathbf{v} \cdot \mathbf{n}. \quad (3)$$

Eq. (2) is solved for the fluid variables \mathbf{u} at time level $n + 1$ implicitly by means of standard integration techniques. The Jacobian matrix is computed by resorting to a first order approximation of the integrated fluxes and a dual time-stepping technique [16]. Second and third order Backward Differentiation Formula (BDF) schemes are adopted to approximate the time derivatives.

Following [12] and [9] the numerical scheme outline above is used together with mesh adaptation techniques. The local modifications of the topology of the grid occurring during the adaptation step, e.g. edge-swapping and node insertion/deletion, are interpreted as a continuous deformation of the finite volumes associated to the grid. The interface velocities given by Eq. (3) are thus computed accounting for the distortion of the finite volumes caused by the modifications in grid topology. Such approach allows to compute the solution onto the new, adapted, grid simply integrating Eq. (2) without any explicit interpolation step. To ensure the conservativity of the resulting scheme additional flux contributions must be taken into account for every removed edge [12] and additional conservation equations must be integrated for every removed node [9]. Such additional fluxes and equations can be dropped after a given number of time steps depending on the time-integration scheme adopted, e.g. two for a BDF2 and three a BDF3, since their contribution is identically equal to zero.

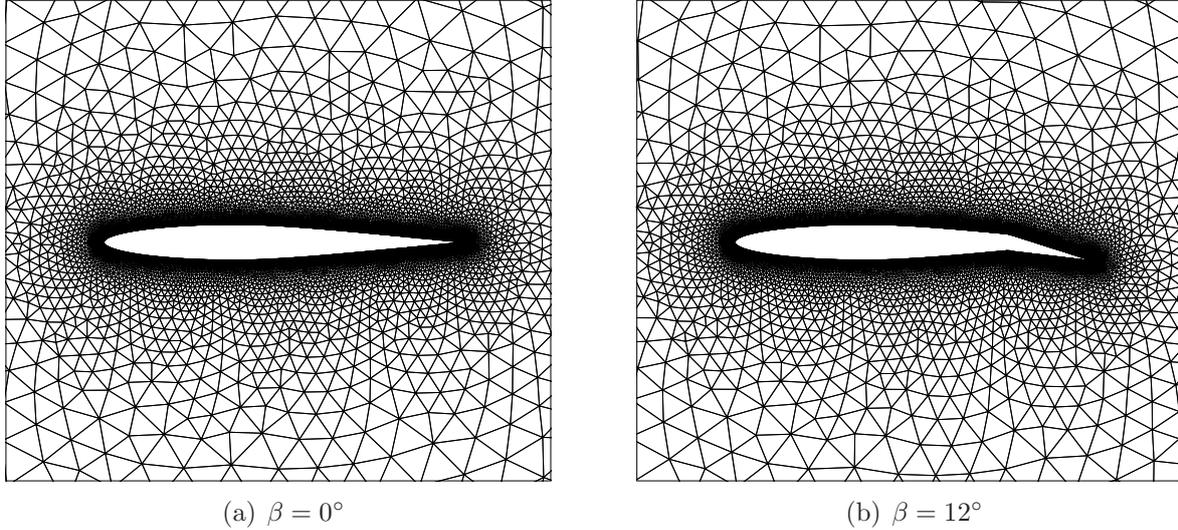


Figure 3: Numerical grid for two flap rotations.

3 FLUID-STRUCTURE INTERACTION MODELING FOR THE AILERON BUZZ

The structural model used to describe the aileron motion is a simple one degree of freedom equation expressing equilibrium of moments about the aileron hinge, namely

$$I\ddot{\beta} = M, \quad (4)$$

where I is the inertia moment of the flap around the hinge and M is the aerodynamic moment about the hinge line. Given the value of the flap rotation β at a given time t^{n+1} , the position of the grid nodes belonging to the boundary is updated and the inner nodes are displaced resorting to a mesh deformation algorithm based on the elastic analogy that preserves the good quality of elements close to the airfoil [12, 4].

The direct time integration of the fluid-structure interaction problem is tackled using a partitioned loosely coupled algorithm. Both aerodynamic and structural systems are integrated using an implicit scheme, thus achieving linear stability for any value of time-step Δt . In particular, Eq. 2 has been integrated using a second order accurate BDF scheme, while a predictor-corrector method derived from Crank-Nicholson [7] has been adopted for the structural subsystem. The latter scheme is here briefly outlined.

1. The known values of flap angle, flap angular velocity and aerodynamic loads at the time n are used to prediction of the structural state at time $n + 1$, i.e.

$$\begin{cases} \beta_p^{n+1} = \beta^n + \Delta t \dot{\beta}^n + \frac{\Delta t^2}{2I} M^n, \\ \dot{\beta}_p^{n+1} = \dot{\beta}^n + \frac{\Delta t}{I} M^n. \end{cases}$$

2. The predicted predicted structural state, β_p^{n+1} and $\dot{\beta}_p^{n+1}$ is used to compute the new mesh and subsequently the aerodynamic loads, $M^{n+1} = M(\beta_p^{n+1}, \dot{\beta}_p^{n+1})$.
3. The predicted loads are used to correct the value of the structural state at t^{n+1} , i.e.

$$\begin{cases} \beta^{n+1} = \beta^n + \Delta t \dot{\beta}^n + \frac{\Delta t^2}{2I} \frac{M^{n+1} + M^n}{2}, \\ \dot{\beta}^{n+1} = \dot{\beta}^n + \frac{\Delta t}{I} \frac{M^{n+1} + M^n}{2}. \end{cases}$$

4 NUMERICAL RESULTS

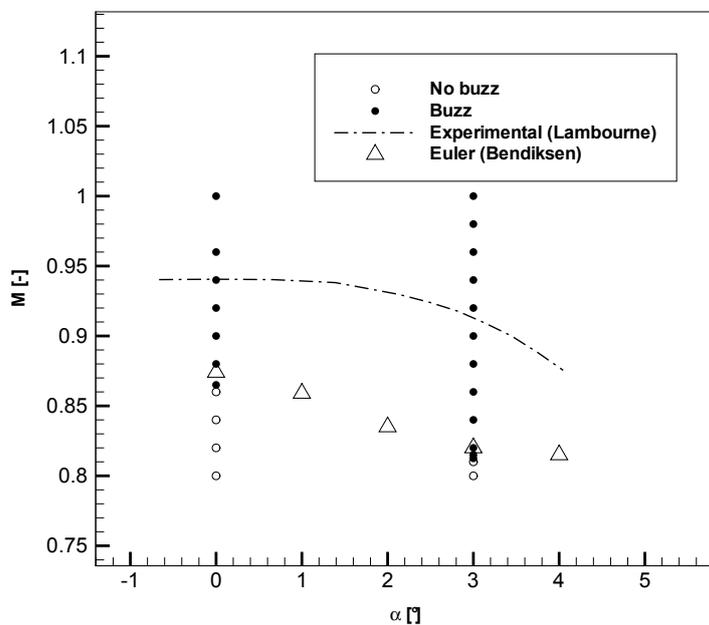


Figure 4: Comparison of experimental and numerical stability boundaries of the Mach-angle-of attack plane.

Aileron buzz is now examined to study the suitability of the proposed approach to investigate aeroelastic phenomena in two dimensional cases. The prerogative is to assess non-classical aileron buzz (type B using Lambourne classification [10]) over a range of transonic Mach numbers. Tests are conducted on a RAE 102 typical section model clamped on its mass center to avoid pitch or plunge movements and to allow only flap rotation around its hinge. Flap-chord/chord ratio is 25% and non-dimensional frequency parameter is $f = 0.063$ approximately (which corresponds to a reduced frequency of $k = \omega c/V_\infty = 0.1$). A circular domain with a radius of 20 chords is chosen to avoid

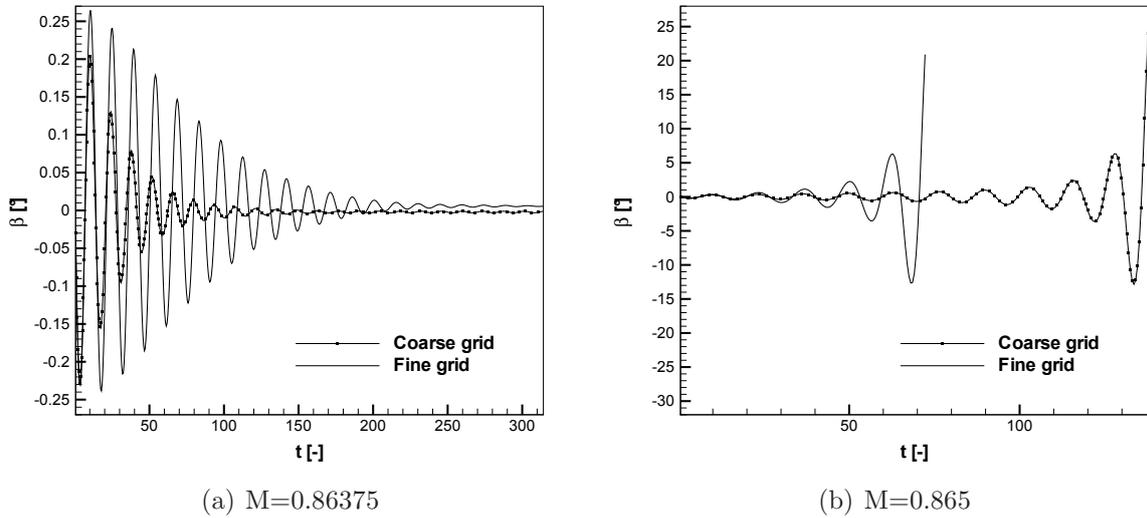


Figure 5: Comparison of flap rotation transient at $\alpha = 0^\circ$ on different grids : (a) stable responses; (b) unstable responses.

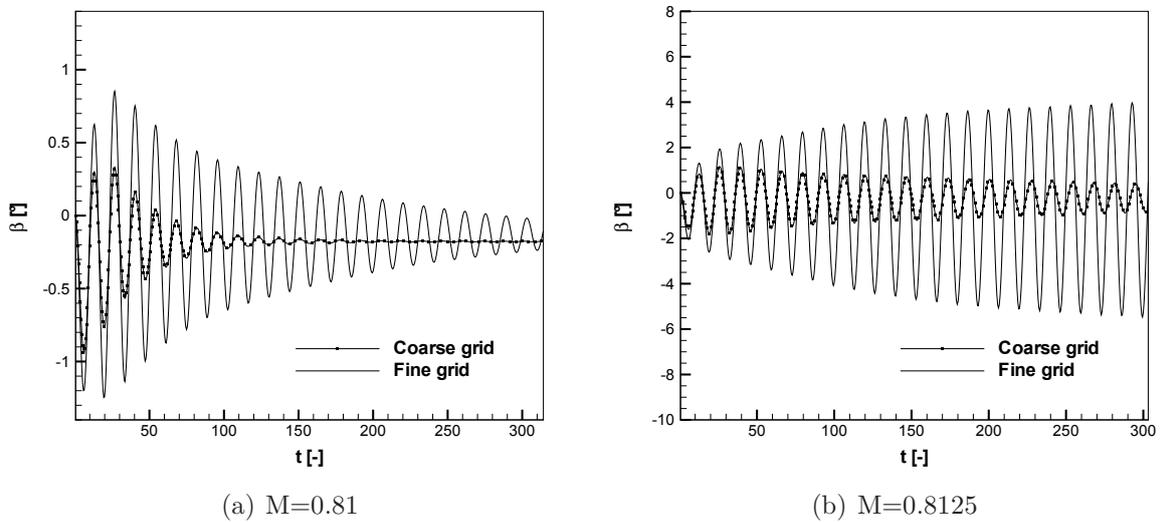


Figure 6: Comparison of flap rotation transient at $\alpha = 3^\circ$ on different grids : (a) stable response; (b) unstable/stable responses.

far-field boundary conditions interferences on the unsteady phenomenon. The result obtained is the stability boundary in the Mach-angle-of-attack plane. Since RAE 102 is a symmetrical airfoil, tests at $\alpha = 0^\circ$ are conducted imposing a non-dimensional initial

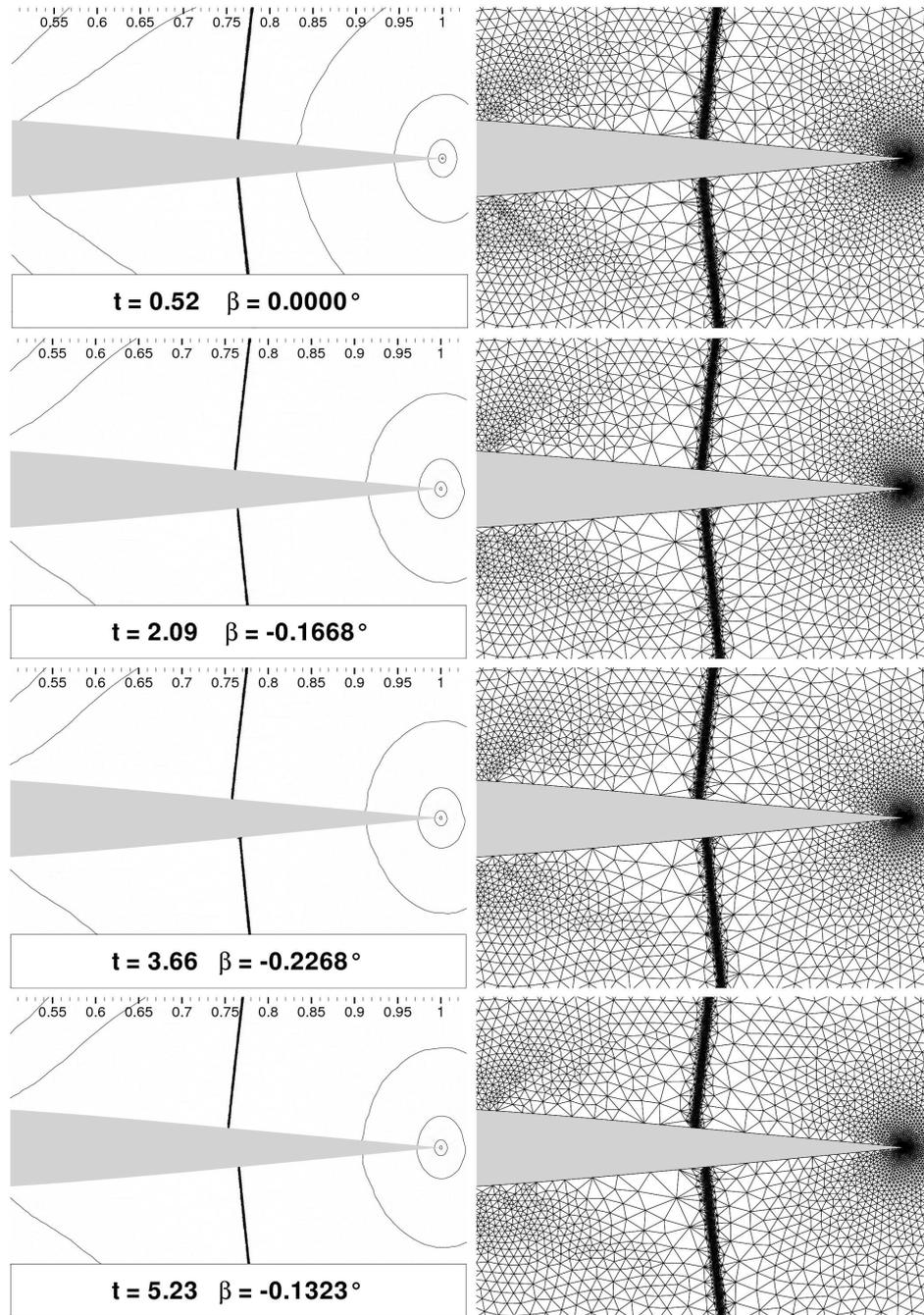


Figure 7: Snapshots of adapted grids and Mach number contour during a cycle of an unstable aileron buzz case ($M=0.865, \alpha = 0^\circ$).

angular velocity of flap around the hinge different from zero ($\dot{\beta} = -10^{-3}$). Differently, computations with an angle of attack $\alpha = 3^\circ$ don't need an initial perturbation thank to

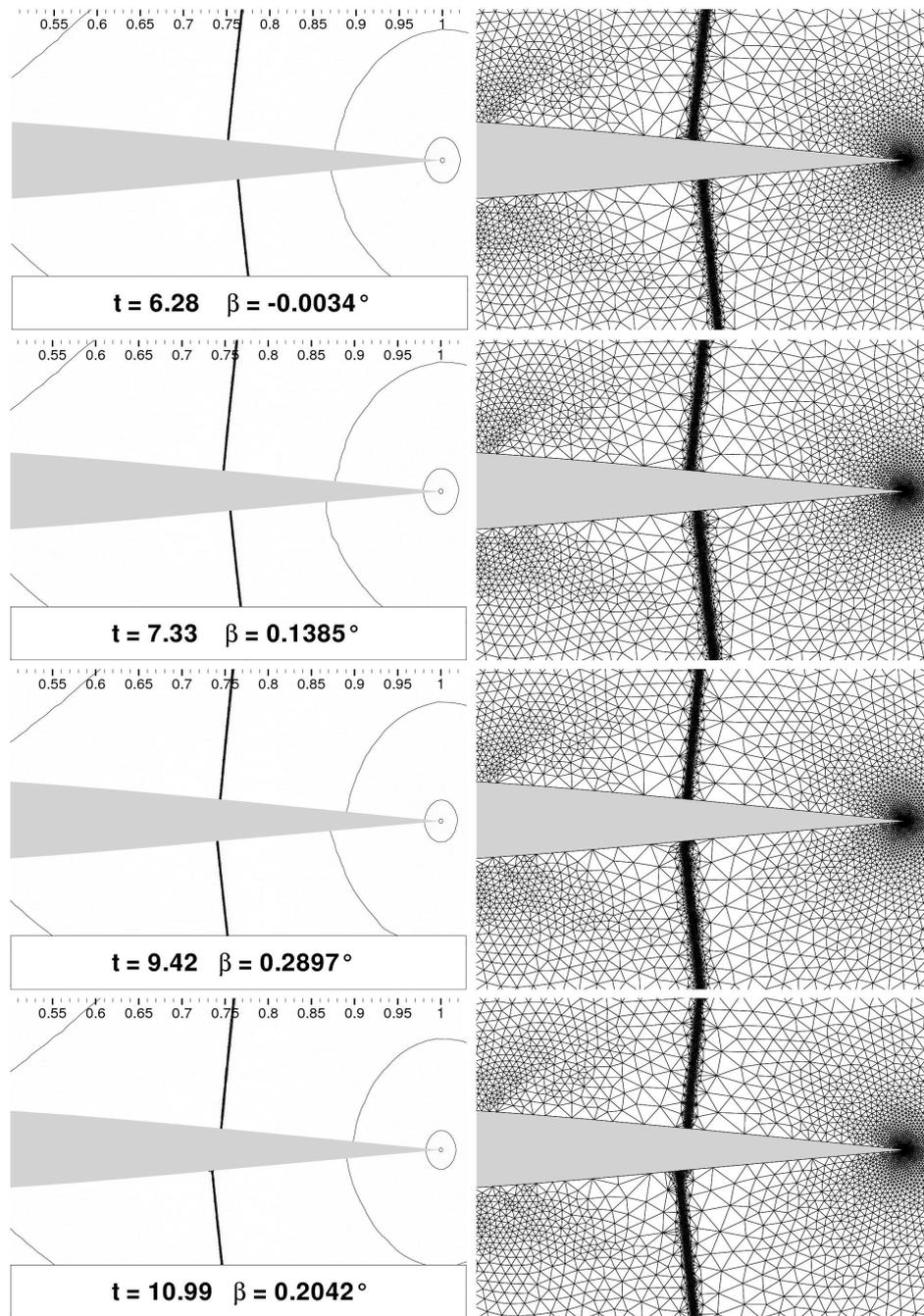


Figure 8: Snapshots of adapted grids and Mach number contour during a cycle of an unstable aileron buzz case ($M=0.865, \alpha = 0^\circ$).

a sufficient initial value of flap hinge moment coefficient $C_m = \frac{M}{1/2\rho V^2 c^2}$.

Numerical computations are conducted at different Mach numbers ($0.8 < M < 1.0$) with two different angles of attack ($\alpha = 0^\circ, \alpha = 3^\circ$) in order to define numerical boundary stability. Moreover two different grids are used to perform a sensitivity of the phenomenon with respect to the computational discretization spacing; a coarse grid with 10 396 nodes and 19 745 elements and a fine computational grid with 20 845 nodes and 40 482 elements. All computations are performed using the second-order BDF time integration scheme with $\Delta t = 0.5235$ and a non-dimensional simulation time $T = 314.15$.

Figure 4 shows the comparison of the computed stability boundary on the Mach-angle-of-attack plane with the experimental one. Euler equations model produces a result far away from the experimental model instability [10]. In fact, since viscous effects are not considered in Euler equations, a lower stability boundary is obtained. Boundary layer presence implies an increase airfoil thickness for inviscid flow providing an higher stability boundary in terms of Mach number. On the other hand, numerical results are in fairly good agreement with inviscid computations by Bendiksen [1].

In figure 5 is shown the response in terms of flap rotation β for $\alpha = 0^\circ$ across the stability boundary (respectively $M=0.86375$ and $M=0.865$) using two different grid spacings. In general, a reduced grid spacing determines a more negative aerodynamic damping effect, both for stable and unstable cases. This is due to numerical viscosity increase with volume cell. Flap rotation responses for $\alpha = 3^\circ$ are shown in figure 6. In this case numerical stability boundary has been evaluated about $M=0.8125$. In fact, across this value of Mach number a different behaviour can be noticed in changing grid space. The fine grid gives an unstable response while the coarse grid a stable one. This is in agreement with the increase of aerodynamic damping effect with cell volume.

Figure 7 and figure 8 show the details of some subsequent flow-fields around the flap during an unstable response at $M = 0.865, a = 0^\circ$. Numerical computation has been performed using a *grid adaptation* technique to improve the capturing of shocks movements on airfoil and flap surface. Mesh adaptation has been driven by an error indicator based on the Hessian matrix and the gradient of the Mach number over the whole computation domain. Initial grid at $t = 0^-$ has been created using 21 625 nodes and 42 747 elements. When the flap starts moving upward the upper shock moves forward and lower its strength while the lower shock does the opposite.

5 CONCLUSIONS

A novel conservative adaptive-grid for the Euler equations was applied to the computation of the stability boundary for the non-classical buzz problem for the RAE 102 airfoil. The proposed technique does not require cross-grid reinterpolation. As expected, for a given angle of attack, numerical simulations of the buzz phenomena on the denser grid are found to indicate a lower stability limit for the Mach number with respect to coarse grid computations. Remarkably enough, adapted grid computations are found to be less accurate in predicting the location of the stability boundary with respect to the coarse and dense grid ones, which is higher for the adapted case. This is believed to be related

to the adaptation sensors, which allows to gather grid points close to the shock wave but it is not sufficient to increase the accuracy in the smooth regions that contribute to the overall value of the hinge momentum. Current research activities are devoted to further study suitable sensors to improve the adaptation process in terms of the evaluation of the stability limit.

REFERENCES

- [1] O.O. Bendiksen. Nonclassical aileron buzz in transonic flow. *34th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference*, pages (AIAA Paper 93-1479), April 1993.
- [2] Robert M. Bennet and John W. Edwards. An overview of recent developments in computational aeroelasticity. In *Proceedings of the 29th AIAA Fluid Dynamics Conference*, Albuquerque, NM, June 15-18 1998.
- [3] Edoardo Biganzoli and Giuseppe Quaranta. Nonlinear reduced order models for aileron buzz. In *International Forum on Aeroelasticity and Structural Dynamics (IFASD 2009)*, Seattle, WA, USA, 21-25 June 2009.
- [4] Luca Cavagna, Giuseppe Quaranta, and Paolo Mantegazza. Application of Navier-Stokes simulations for aeroelastic assessment in transonic regime. *Computers & Structures*, 85(11-14):818-832, 2007.
- [5] J. Donea. A Taylor-Galerkin method for convective transport problems. *Int. J. Numer. Meth. Eng.*, 20:101-119, 1984.
- [6] J. Donea, A. Huerta, J.-Ph. Ponthot, and A. Rodríguez-Ferran. Arbitrary lagrangian-Eulerian methods. In R. Stein, E. de Borst, and T.J.R. Hughes, editors, *The Encyclopedia of Computational Mechanics*, volume 1, chapter 14, pages 413-437. Wiley, 2004.
- [7] M. B. Giles. Stability analysis of a Galerkin/runge-Kutta navier-stokes discretization on unstructured tetrahedral grids. *J. Comput. Phys.*, 132:201-214, 1997.
- [8] E. Godlewski and P. A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*. Springer-Verlag, New York, 1994.
- [9] Dario Isola, Alberto Guardone, and Giuseppe Quaranta. An ALE scheme without interpolation for moving domain with adaptive grids. In *AIAA 40th Fluid Dynamics Conference and Exhibit*, Chicago, IL, USA, 28 June-1 July 2010.
- [10] N.C. Lambourne. Flutter in one degree of freedom. In *Manual on Aeroelasticity, Part V*, chapter 5. AGARD, 1961.

- [11] N.C. Lambourne. Control–surface buzz. Reports and Memoranda 3364, Aeronautical Research Council, London, UK, May 1964.
- [12] D. Muffo, G. Quaranta, and A. Guardone. Compressible fluid-flow ale formulation on changing topology meshes for aeroelastic simulations. In *Proceedings of the 26th ICAS Congress*, Anchorage, Alaska, September 14–19 2008.
- [13] Giuseppe Quaranta, Dario Isola, and Alberto Guardone. Numerical simulation of the opening of aerodynamic control surfaces with two-dimensional unstructured adaptive meshes. In *5th European Conference on Computational Fluid Dynamics - ECCOMAS CFD 2010*, Lisbon, Portugal, 14–17 June 2010.
- [14] David M. Schuster, D. D. Liu, and Lawrence J. Huttsell. Computational aeroelasticity: Success, progress, challenge. *Journal of Aircraft*, 40(5):843–856, 2003.
- [15] B. van Leer. Towards the ultimate conservative difference scheme II. Monotonicity and conservation combined in a second order scheme. *J. Comput. Phys.*, 14:361–370, 1974.
- [16] V. Venkatakrishnan and D. J. Mavriplis. Implicit method for the computation of unsteady flows on unstructured grids. *J. Comput. Phys.*, 127:380–397, 1996.

NUMERICAL SIMULATIONS OF PARTICLES IN A SHEAR FLOW

NICOLAS VERDON*, PATRICE LAURE*, ALINE LEFEBVRE-LEPOT† AND
LAURENT LOBRY‡

* Laboratoire J.-A. Dieudonné, CNRS UMR 6621, Université de Nice-Sophia Antipolis
Parc Valrose, 06108 Nice Cedex 02, France
Email: {nicolas.verdon,patrice.laure}@unice.fr

† Centre de Mathématiques Appliquées (CMAP), Ecole Polytechnique
Route de Saclay, 91128 Palaiseau, France
Email: aline.lefevre@polytechnique.edu

‡ Laboratoire de Physique de la Matière Condensée (LPMC), Université de Nice-Sophia Antipolis
Parc Valrose, 06108 Nice Cedex 02, France
Email: laurent.lobry@unice.fr

Key words: Suspension, contact, viscous contact model, fluid-structure interaction, immersed domain method

Abstract. In this paper, we present an immersed domain approach coupled with a viscous contact model for studying the rheological behaviour of dense suspensions in a shear flow. We here demonstrate the importance of contact modelling as well as the choice of the boundary conditions on the macroscopic properties of the suspension.

1 INTRODUCTION

In the field of materials forming, as well as in many other industrial fields, determining the rheological behaviour of dense suspensions remains of great importance. Lots of different computational methods can be used to handle such fluid-structure interactions but in this work, we propose to study the behaviour of solid particles in a shear flow with the immersed domain method. First introduced in the late 90's by Glowinski *et al.* [1], this kind of methods encounters an increasing success in fluid-structure of multiphase problems because it allows to treat these problems with an Eulerian approach on the whole domain.

In order to be able to study accurately problems with a large amount of particles, we focused first our analysis to the description of the contact for a few particles in a shear flow. To avoid particle overlapping during displacement of particles, we implemented contact models such as an inelastic collisions model and the viscous contact model introduced by Maury [2] and intensively studied by Lefebvre [3] for granular applications.

We then focused our analysis on the choice of the boundary conditions for describing dense suspensions in a representative way. We present in details the way of extending the computational domain and we insist on the description of sliding biperiodic conditions of Lee-Edwards [4]. In the last part of this article, we present some representative results that confirm the different choices we have done in this study. First, the importance of the contact modelling will be pointed out with a 3D example of 13 particles in a shear flow. Finally, some examples of suspensions will also demonstrate the rheological modifications due to the choice of boundary conditions.

2 NUMERICAL MODEL

In this part, we briefly describe the main features of the numerical model that has been used for this study. The importance of the contact modelling is therefore explained at the end of the section.

2.1 The immersed domain approach

The immersed domain method is achieved by splitting the computational domain Ω into two subdomains Ω_f and Ω_s for respectively the fluid and solid parts (see Figure 1). In the case of multiple particles, the solid domain is the union of domains corresponding to each particle, namely $\Omega_s = \bigcup_i^N \Omega_{s_i}$ for N particles.

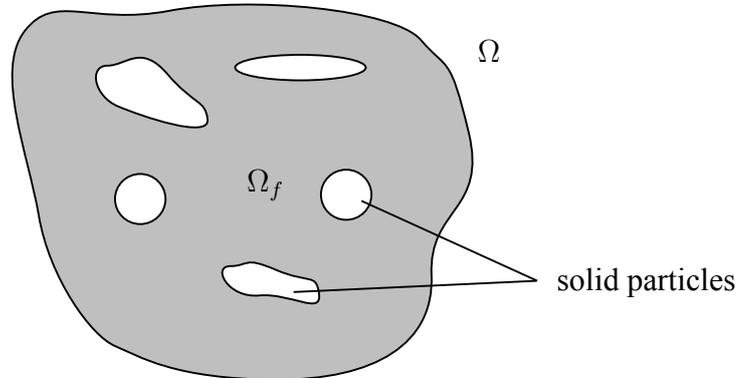


Figure 1: Schematic representation of computational domain

The interface Γ_s between the two phases is described by the zero isosurface of a distance function:

$$\alpha(\mathbf{x}, t) = \begin{cases} > 0 & \text{if } \mathbf{x} \in \Omega_s \\ < 0 & \text{if } \mathbf{x} \notin \Omega_s \end{cases} \quad (1)$$

which allows to define a "smooth" characteristic function:

$$\mathbb{I}(x, t) = \begin{cases} = 1 & \text{if } \alpha(x) > e \\ = \frac{\alpha}{e} & \text{if } 0 < \alpha(x) < e \\ = 0 & \text{if } \alpha < 0 \end{cases} \quad (2)$$

where e the mixing thickness depends on the mesh size around the interface. In addition, the viscosity η will be defined thanks to a mixing relation:

$$\eta = \mathbb{I}\eta_s + (\mathbb{E} - \mathbb{I})\eta_f \quad (3)$$

where η_f is the fluid viscosity and $\eta_s \simeq r\eta_f$ the solid viscosity (or penalisation factor) usually taken much bigger than η_f ($r \approx 10^3$).

2.2 Governing equations

Neglecting inertia and gravity, the fluid-solid problem can be written with the following set of equations:

$$\begin{cases} \nabla \cdot \sigma & = 0 \\ \nabla \cdot \mathbf{u} & = 0 \\ \sigma & = 2\eta_f \dot{\epsilon}(\mathbf{u}) - p\mathbb{E} \\ [[\mathbf{u}]]_{\Gamma_s} & = 0 \\ [[\sigma \cdot \mathbf{n}]]_{\Gamma_s} & = 0 \\ \mathbf{u} & = \mathbf{u}_\Gamma \quad \text{on the external boundary } \Gamma \end{cases} \quad (4)$$

where \mathbf{u} is the fluid velocity, $\dot{\epsilon}(\mathbf{v})$ the rate of strain tensor, σ the stress tensor, p the pressure, η_f the fluid viscosity (the symbol $[[f]]_{\Gamma_s}$ means the jump of scalar f across the interface Γ_s). Patankar *et al.* [5] have proposed to extend the above Stokes equation to the solid domain thanks to a Lagrange multiplier by using the rigidity condition $\dot{\epsilon}(\mathbf{v}) = 0$ on Ω_s . In this way, the motion in solid domain Ω_s corresponds to a fluid motion with an additional stress field. This is equivalent to take the stress tensor σ inside the solid domain of the form

$$\sigma = 2\eta_s \dot{\epsilon}(\mathbf{u}) - p\mathbb{E} + \dot{\epsilon}(\lambda) \quad (5)$$

2.3 Weak formulation of the FSI problem

The equation (5) allows us to write the following weak formulation over the whole computational domain Ω , where Dirichlet boundary conditions are imposed:

Find (\mathbf{u}, p, λ) such that $\forall (\mathbf{v}, q, \mu) \in \mathcal{H}^1(\Omega) \times \mathcal{L}_0^2(\Omega) \times \mathcal{H}^1(\Omega_s)$:

$$\begin{cases} 0 & = \int_{\Omega} 2\eta \dot{\epsilon}(\mathbf{u}) : \dot{\epsilon}(\mathbf{v}) d\Omega - \int_{\Omega} p \nabla \cdot \mathbf{v} d\Omega + \int_{\Omega_s} \dot{\epsilon}(\lambda) : \dot{\epsilon}(\mathbf{v}) d\Omega \\ 0 & = \int_{\Omega} q \nabla \cdot \mathbf{u} d\Omega \\ 0 & = \int_{\Omega_s} \dot{\epsilon}(\mu) : \dot{\epsilon}(\mathbf{v}) d\Omega \end{cases} \quad (6)$$

This formulation corresponds to an augmented Lagrangian where λ is the Lagrange multiplier and η_s the penalisation factor. This is solved by an Uzawa algorithm [6].

2.4 Contact modelling

The particles displacement is then achieved by using the velocity of the fluid, solution of (6), and a lagrangian approach. Namely the displacement \mathbf{X}_i of particle i at time $t^{n+1} = t^n + \Delta t$ is computed with the second order Adams-Bashfort scheme:

$$\mathbf{X}_i(t^{n+1}) = \mathbf{X}_i(t^n) + \frac{\Delta t}{2} [3\mathbf{u}(\mathbf{X}_i(t^n), t^n) - \mathbf{u}(\mathbf{X}_i(t^{n-1}), t^{n-1})] \quad (7)$$

It has been observed that overlapping can occur during this step, mathematically the distance D_{ij} between two particles i and j can become negative. In order to avoid this non-physical behaviour, we used the viscous contact model that has been introduced by Maury [2] and Lefebvre [3]. This model can be seen as a predictor–corrector model based on the action of the lubrication force. Indeed, the velocity field \mathbf{u}^* obtained by solving Equation (6) does not take into account contacts. Hence, this *predicted velocity* will be corrected in a way that it satisfies the non-overlapping condition, namely by solving:

$$\frac{1}{2} |\mathbf{u}^n - \mathbf{u}^*|^2 = \min_{\mathbf{v} \in K(\mathbf{X}^n, \gamma_{ij}^n)} \frac{1}{2} |\mathbf{v} - \mathbf{u}^*|^2 \quad (8)$$

where K is the space of admissible velocity defined by

$$K(\mathbf{X}^n, \gamma_{ij}^n) = \left\{ \mathbf{v} \left| \begin{array}{l} D_{ij}^n + \Delta t(\mathbf{v}_j - \mathbf{v}_i) \cdot \mathbf{e}_{ij}^n \geq 0 \text{ if } \gamma_{ij}^n = 0 \\ D_{ij}^n + \Delta t(\mathbf{v}_j - \mathbf{v}_i) \cdot \mathbf{e}_{ij}^n \leq 0 \text{ if } \gamma_{ij}^n < 0 \end{array} \right. \right\} \quad (9)$$

The contact between particles is then described by a new variable γ_{ij} which can be seen as a microscopic distance between particles i and j :

$$\gamma_{ij}^n \begin{cases} < 0 \text{ if there is contact between particles } i \text{ and } j \\ = 0 \text{ else} \end{cases} \quad (10)$$

Let us define the functional J as follows:

$$J(\mathbf{v}) = \frac{1}{2} |\mathbf{v} - \mathbf{u}^*|^2 = \frac{1}{2} \mathbf{M}\mathbf{v} \cdot \mathbf{v} - \mathbf{M}\mathbf{u}^* \cdot \mathbf{v} \quad (11)$$

where $\mathbf{M} = \text{diag}(\dots, m_i, \dots, m_j, \dots)$ is the mass matrix. Then the Lagrangian of $J(\mathbf{u})$ for two particles i and j has the following form:

$$\begin{aligned} \mathcal{L}(\mathbf{v}, \lambda_{ij}^\pm) &= J(\mathbf{v}) - \lambda_{ij}^+ (D_{ij} + \Delta t(\mathbf{v}_j - \mathbf{v}_i) \cdot \mathbf{e}_{ij}) \\ &\quad - \lambda_{ij}^- (-D_{ij} - \Delta t(\mathbf{v}_j - \mathbf{v}_i) \cdot \mathbf{e}_{ij}) \end{aligned} \quad (12)$$

where λ_{ij}^+ and λ_{ij}^- are the Lagrange multipliers associated to the contact constraints. Finally, by solving $\frac{\partial \mathcal{L}}{\partial \mathbf{v}_i} = 0$ and $\frac{\partial \mathcal{L}}{\partial \mathbf{v}_j} = 0$, we obtain the two following relations:

$$\begin{cases} m_i \mathbf{u}_i^n = m_i \mathbf{u}_i^* - (\lambda_{ij}^+ - \lambda_{ij}^-) \Delta t \mathbf{e}_{ij} \\ m_j \mathbf{u}_j^n = m_j \mathbf{u}_j^* + (\lambda_{ij}^+ - \lambda_{ij}^-) \Delta t \mathbf{e}_{ij} \end{cases} \quad (13)$$

under the conditions :

$$\lambda_{ij}^\pm (D_{ij} + \Delta t (\mathbf{u}_j^n - \mathbf{u}_i^n) \cdot \mathbf{e}_{ij}) = 0 \quad \text{with} \quad \lambda_{ij}^+ \geq 0 \quad \text{and} \quad \lambda_{ij}^- \leq 0 \quad (14)$$

which is also solved with an Uzawa procedure.

Then the correction step of the viscous contact model involves the Fundamental Principle of the Dynamics where the lubrication forces are taken into account. The evolution of the contact parameter γ_{ij} can be obtain after some basic calculus, see [7] for more details:

$$\frac{d\gamma_{ij}}{dt} = -\frac{1}{a^2} \lambda_{ij} \quad \text{with} \quad \lambda_{ij} = \lambda_{ij}^+ - \lambda_{ij}^- \quad (15)$$

Note that the inelastic collisions model can be obtained from these expressions by imposing $\gamma_{ij} = 0$, that is to say $\lambda_{ij} > 0$.

3 DESCRIPTION OF THE BOUNDARY CONDITIONS

When increasing the concentration of solid particles in a suspension, we notice that the choice of the boundary conditions becomes of great importance. Indeed, the number of contacts between particles is also increased and hence the boundary conditions (BC) can affect the rheology of the suspension. In this section, we present the different choices that can be made by focusing particularly on the Lee-Edwards' biperiodic conditions.

3.1 Representative elementary volume

In the numerical study of suspension, it is important to work on a suitable representative elementary volume (REV). Ideally, one would like to be able to know the behaviour of suspensions in an infinite domain, practically by working on a very large computational domain Ω_t with boundaries extremely far from the domain of interest Ω_{REV} , as depicted in figure 2. Unfortunately, this kind of approach is nowadays very computationally expensive, especially for 3D simulations, that is why we have to deal with the boudary conditions in order to have a representative domain. For low concentrated suspension, it is common to work on a small domain but if the concentration increases, it is no more representative. Indeed, in Ω_{REV} the influence of the particles from the domain $\Omega_t \cap \overline{\Omega}_{REV}$ are taken into account whereas in a small domain of size Ω_{REV} , the particles inside the domains do not see particles from outside.

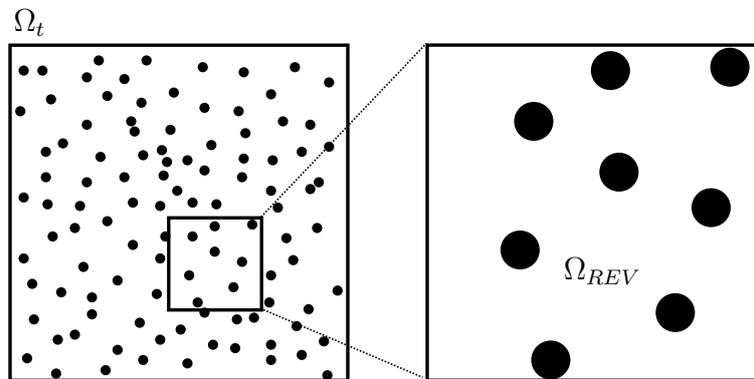


Figure 2: Ideal computational domain and Representative Elementary Volume

3.2 Extension of the computational domain

In this article, we focused our work on a shear flow in a cavity, as represented in figure 3. The velocities on the upper and lower boundaries are hence imposed in order to get the desired

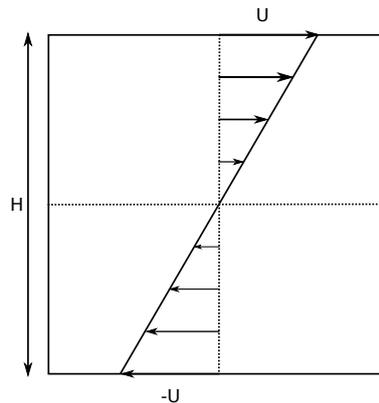


Figure 3: Configuration of the shear flow

shear rate whereas the vertical velocities in the lateral walls are equal to zero. As explained in the previous section, we mainly use small computational domains Ω_0 like 4(a) for dilute suspensions. In this case, we impose periodic conditions on the horizontal directions. That means that each particle that goes outside the domain through vertical walls re-enters in the domain from the other side. However, because of the zero vertical velocity imposed in these walls, the velocities can be badly estimated. Indeed, even if the motion of particles is periodic, the computed velocity is not periodic, that is why we proposed to extend the computational domain in the x -direction as depicted in figure 4(b). Thanks to this extension, the particles near a boundary see those from the opposite side.

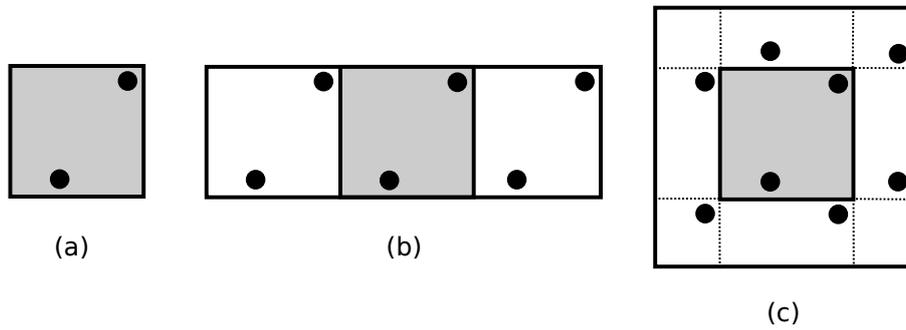


Figure 4: Extension of the computational domain : (a) original domain CL (0,0), (b) domain extended in the x -direction CL (1,0), (c) domain with sliding biperiodic boundary conditions CL (1/2,1/2)

In practice, we can extend the computational domain as much as we want and k_x and k_y are the two parameters which characterise the extension of the computational domain in respectively x and y - direction. For example, in the configuration 4(b), $k_x = 1$ and $k_y = 0$. The domain is extended in x -direction of length H in both sides of the original domain so that the new domain is then three times larger than the original domain Ω_0 . The way of duplicating particles follows these rules:

$$\text{if } x_\alpha + H \in [H, H + k_{x_\alpha}H], \text{ then } x_\alpha^+ = x_\alpha + H, \quad \text{where } x_\alpha = x \text{ or } y \quad (16)$$

$$\text{if } x_\alpha - H \in [H - k_{x_\alpha}H, 0], \text{ then } x_\alpha^- = x_\alpha - H, \quad \text{where } x_\alpha = x \text{ or } y \quad (17)$$

and figure 5 schematically represents the duplication of two particles for two different extensions of the domain: the first 5(a) for an extension with $k_x = 1/2$ and the second 5(b) with $k_x = 1$ ($k_y = 0$ for both of them). With this modification of the boundary conditions, we

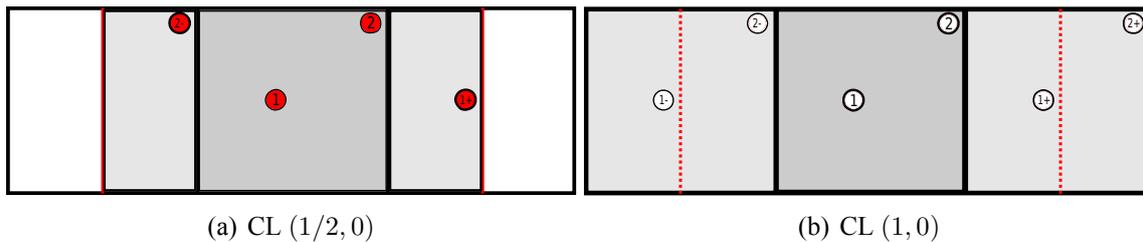


Figure 5: Schematic description of particles' duplication

decrease the influence of the vertical walls on the computation of the velocity. We can precise that only particles inside the original domain are moved. The new particles are introduced just for improving the computation of the velocity field and their positions are determined only through geometric considerations. In order to limit the influence of the horizontal walls, we also carry out biperiodic boundary conditions such as presented in figure 4(c). The methodology is described more precisely in the next section.

3.3 Sliding biperiodic boundary conditions

Here we describe the changes induced by the Lee-Edwards' biperiodic boundary conditions [4, 8]. The new configuration is presented in figure 6 where we introduced the following nota-

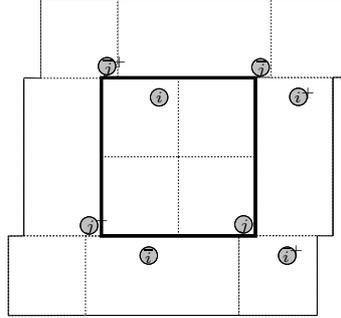


Figure 6: Two particles in a shear flow with biperiodic boundary conditions

tions : i and j design two physical particles, i^\pm and j^\pm are the x -periodic images, as described in last section, \bar{i} and \bar{j} design the images by the biperiodic boundary conditions, and at last \tilde{i}^\pm and \tilde{j}^\pm are the x -periodic images of particles \bar{i} and \bar{j} . In order to keep the shear condition on the upper and lower boundaries of the domain of interest, it is necessary to consider *sliding* boundary conditions as depicted in figure 6. The displacement velocity of the sliding boundary conditions is computed from the definition of the shear rate $\dot{\gamma} = \frac{2U}{H}$; the upper and lower additional domains have hence to move with the velocity:

$$\mathbf{U} = \pm \frac{\dot{\gamma}H}{2} \cdot \mathbf{e}_x \quad (18)$$

so that the positions of the duplicated particles are given by:

$$\bar{\mathbf{x}}_\alpha^{n+1} = \bar{\mathbf{x}}_\alpha^n \pm \frac{\dot{\gamma}}{2} H \Delta t \cdot \mathbf{e}_x \quad \alpha = i, j \quad (19)$$

3.4 Computation of the distances between particles

Let us now consider the positions and the velocities of the two particles known at times t^n and t^{n-1} . Using the Adams-Bashforth scheme, we are able to compute the positions of physical particles at time t^{n+1} as follows:

$$\mathbf{x}_\alpha^{n+1} = \mathbf{x}_\alpha^n + \frac{\Delta t}{2} (3\mathbf{u}_\alpha^n - \mathbf{u}_\alpha^{n-1}) \quad \alpha = i, j \quad (20)$$

and the positions of the corresponding images are computed with the velocity of physical particles:

$$\mathbf{x}_\alpha^{\pm n+1} = \mathbf{x}_\alpha^{\pm n} + \frac{\Delta t}{2} (3\mathbf{u}_\alpha^n - \mathbf{u}_\alpha^{n-1}) \quad \text{for the images in } x \text{ - direction,} \quad \alpha = i, j \quad (21)$$

$$\bar{\mathbf{x}}_\alpha^{n+1} = \bar{\mathbf{x}}_\alpha^n + \frac{\Delta t}{2} (3\mathbf{u}_\alpha^n - \mathbf{u}_\alpha^{n-1}) \quad \text{for the biperiodic images} \quad \alpha = i, j \quad (22)$$

In the gluey contact model, the condition that must be fulfilled in order to avoid the overlapping is the following:

$$d = \min(d_{ij}, d_{i\bar{j}}, d_{ij^\pm}) \geq 0 \quad (23)$$

So we have to compute the distances between the particles. Using (19), (21) and (22), we obtain the following relationships:

$$d_{ij} = \|\mathbf{x}_i \mathbf{x}_j\|^n + \frac{\Delta t}{2} ((3\mathbf{u}_j^n - \mathbf{u}_j^{n-1}) - (3\mathbf{u}_i^n - \mathbf{u}_i^{n-1})) \cdot \mathbf{e}_{ij} \quad (24)$$

$$d_{i\bar{j}} = \|\mathbf{x}_i \bar{\mathbf{x}}_j\|^n + \frac{\Delta t}{2} ((3\mathbf{u}_j^n - \mathbf{u}_j^{n-1}) - (3\mathbf{u}_i^n - \mathbf{u}_i^{n-1}) \pm \dot{\gamma} H \cdot \mathbf{e}_x) \cdot \mathbf{e}_{ij} \quad (25)$$

$$d_{ij^\pm} = \|\mathbf{x}_i \mathbf{x}_j^\pm\|^n + \frac{\Delta t}{2} ((3\mathbf{u}_j^n - \mathbf{u}_j^{n-1}) - (3\mathbf{u}_i^n - \mathbf{u}_i^{n-1})) \cdot \mathbf{e}_{ij} \quad (26)$$

In the numerical procedure, the distances between physical particles and all images are computed in order to take the minimum value for the contact model.

4 NUMERICAL RHEOLOGY

As explained in the introduction, determining rheological properties of dense suspensions remains important for many applications. For this purpose, we present here the way of calculating the macroscopic variables from our computations. Let first note $\langle \mathbf{X} \rangle_\Omega$ the mean value of \mathbf{X} at time t over Ω . We have:

$$\langle \mathbf{X} \rangle_\Omega = \frac{1}{\Omega} \int_\Omega \mathbf{X}(\mathbf{x}) \, d\Omega \quad (27)$$

By taking the mean value of the stress tensor σ using (27), one gets:

$$\langle \sigma \rangle_\Omega = \langle \sigma_f \rangle_\Omega + \langle \sigma_s \rangle_\Omega \quad \text{with} \quad \begin{cases} \langle \sigma_s \rangle_\Omega = 2\eta_s \langle \dot{\epsilon}(\mathbf{u}) \rangle_{\Omega_s} - \langle p \rangle_{\Omega_s} \mathbf{I}_d + \langle \lambda \rangle_{\Omega_s} \\ \langle \sigma_f \rangle_\Omega = 2\eta_f \langle \dot{\epsilon}(\mathbf{u}) \rangle_{\Omega_f} - \langle p \rangle_{\Omega_f} \mathbf{I}_d \end{cases} \quad (28)$$

where $\langle \sigma_s \rangle_\Omega$ and $\langle \sigma_f \rangle_\Omega$ are respectively the solid and fluid stress tensors. Theoretically, the xy -component of the mean stress tensor in the suspension is given by:

$$\langle \sigma_{xy} \rangle_\Omega = 2\eta_{eff} \langle \dot{\epsilon}(\mathbf{u})_{xy} \rangle_{\Omega_f} \quad (29)$$

which allows us to write the effective viscosity as follows:

$$\eta_{eff} = \frac{\langle \sigma_{xy} \rangle_\Omega}{\langle \dot{\gamma} \rangle_\Omega} \quad (30)$$

5 RESULTS

In previous papers [7, 9] we analysed the reversibility of Stokes equations with two particles and we studied academic configurations such as three 2D particles in a shear flow. In the following, we will focus on more complicated configurations.

5.1 Influence of computational parameters

In order to obtain relevant results for analysing the rheology of dense suspensions, it is necessary to check the importance of some computational parameters. Namely, we present here the influence of both the penalisation and the number of Uzawa iterations for the convergence of the 2D Stokes flow, and the influence of the mesh size. The example that has been used for this purpose is the 2D suspension of concentration $c \approx 0.24$.

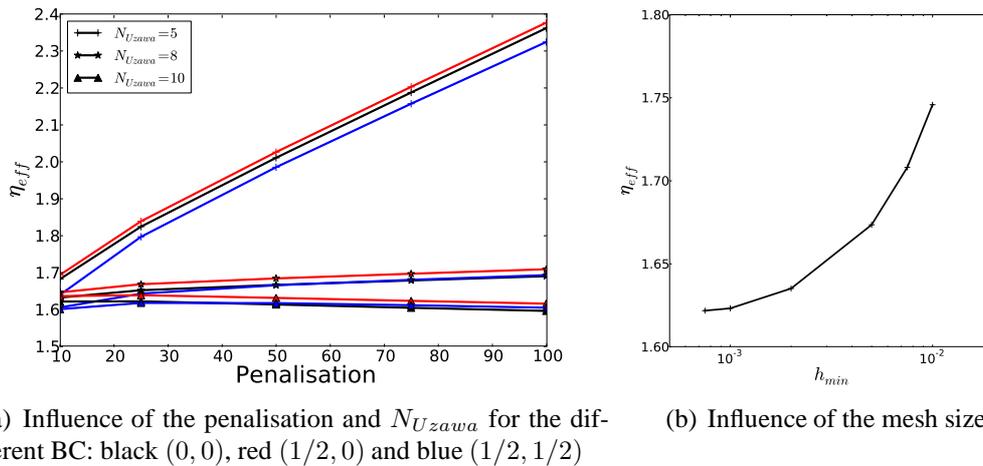


Figure 7: Influence of the computational parameters on the effective viscosity η_{eff}

The results of the different tests are summed up in figure 7. On figure 7(a) one clearly observes that only 5 Uzawa iterations are not sufficient for obtaining the correct effective viscosity, whereas 8 and 10 Uzawa iterations lead to almost the same viscosity, when taking the penalisation enough small. The same observations can be made for the three different boundary conditions. We also tested the influence of the mesh size on the effective viscosity by fixing all meshing parameters, and then gradually decreasing the value of the minimum authorized size of element, called h_{min} . The lower h_{min} is, the finer the mesh is, which is supposed to enhance the quality of the results. The results of this test are reported on figure 7(b), where we notice that it exists an asymptotic value of the viscosity when refining the mesh. In the following, and for CPU time reasons, we will chose a penalisation of 10, 8 Uzawa iterations and $h_{min} = 10^{-3}$.

5.2 Influence of the contact model

The example presented here is the motion of 13 particles of radius $a = 0.05$ in a 3D shear flow. Initially, the particles are located very close from each other (compacity close to 0.74) in order to have the more contacts possible. Figure 8 shows the evolution of the particles trajectories in the (x, y) plane after $t = 12.5s$. This figure emphasizes the differences induced by choice of the contact model: whereas the particles remain stuck during the whole computation

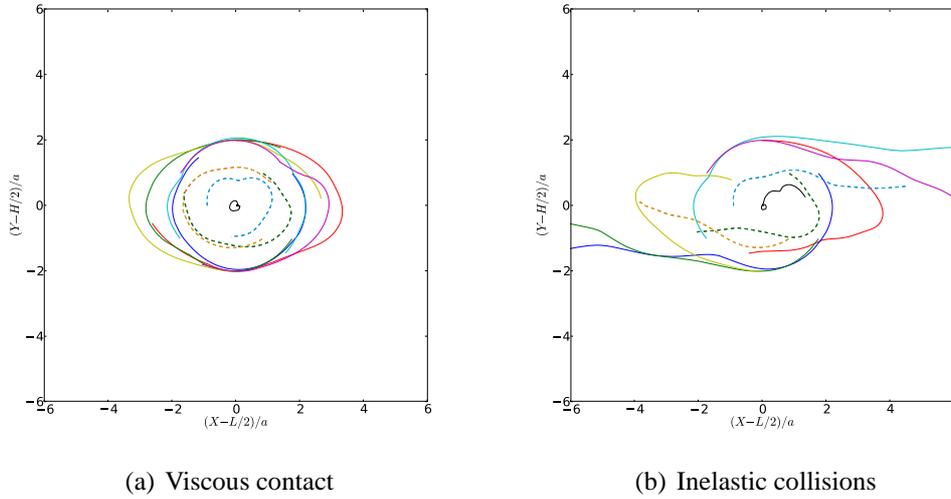


Figure 8: Influence of the contact modelling for the 13-particles case: trajectories of the particles

with the viscous contact model, after a few time steps the cluster of particles disappears with the inelastic collisions model and the particles move away. This is due to the physical nature of the contact models. For the inelastic collisions, the value λ of the Lagrange multiplier can be only positive which means that it acts like a repulsive force between particles. Otherwise, in the viscous model, λ can be either negative or positive, which indicates it acts like a lubrication force: an attractive force is exercised on the particles when they go away from each other.

5.3 Influence of the boundary conditions

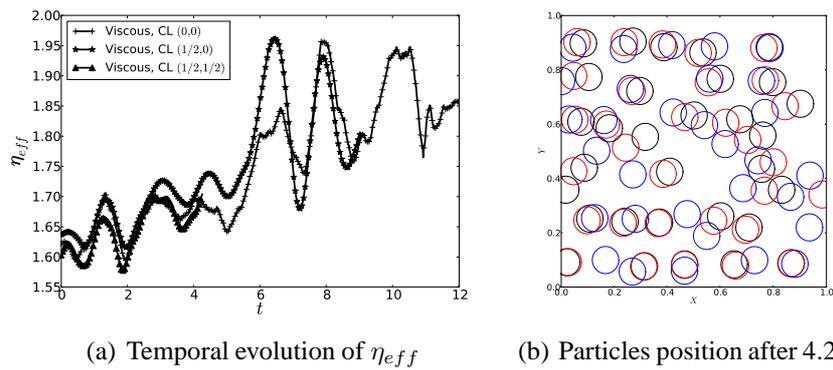


Figure 9: Influence of the boundary conditions for the 2D suspension of concentration $c \approx 0.24$

In this section we study the rheological behaviour of a 2D suspension of concentration

$c \approx 0.24$ with computations for the three different boundary conditions. Figure 9(a) presents the time evolution of the effective viscosity whereas figure 9(b) shows the position of the particles at $t = 4.25s$. These results point out that boundary conditions can significantly affect the trajectories of the particles and hence the rheology of the suspension. Indeed, the biperiodic conditions limit the influence of the upper and lower walls, so that after 4.25s the effective viscosity is lower ($\eta_{eff} \approx 1.64$) than for the two other boundary conditions for which it is almost the same value, $\eta_{eff} \approx 1.66$. At last, the computed values are far from empirical ones ($\sim 2.1 - 2.5$), which emphasizes the importance of 3D effects that are neglected in this study.

6 CONCLUSIONS

Throughout this article, we pointed out the importance of contact modelling as well as the choice of well-suited boundary conditions. The 13-particles example shows a real big difference in the motion of the particles between the two implemented contact models whereas the 2D suspension example insists on the importance of the boundary conditions. The identification of the rheological parameters indicates that biperiodic boundary conditions limit wall effects that could be an advantage for future computations of dense suspensions. Nevertheless, we repeat here that these are just preliminary results, and in the future 3D computations are mandatory for understanding the behaviour of suspensions.

REFERENCES

- [1] R. Glowinski, T.-W. Pan, T.I. Halsa, and D.D. Joseph. A distributed lagrange multiplier/fictitious domain method for particulate flows. *Int. J. Multiphase Flows*, 25:755–794, 1999.
- [2] B. Maury. A gluey particle model. *ESAIM: Proc.*, 18:133–142, 2007.
- [3] A. Lefebvre. Numerical simulation of gluey particles. *M2AN*, 43:53–80, 2009.
- [4] A.W. Lees and S.F. Edwards. The computer study of transport process under extreme condition. *J. Phys. C: Solid State Phys.*, 5:1921–1928, 1972.
- [5] N.A. Patankar, P. Singh, D.D. Joseph, R. Glowinski, and T.-W. Pan. A new formulation of the distributed lagrange multiplier/fictitious domain method for particulate flows. *Int. J. of Multiphase flow*, 26:1509–1524, 2000.
- [6] P. Laure, G. Beaume, O. Basset, L. Silva, and T. Coupez. Numerical methods for solid particles in particulate flow simulations. *European J. Comp. Mechanics*, 16:365–383, 2007.
- [7] N. Verdon, A. Lefebvre-Lepot, L. Lobry, and P. Laure. Contact problems for particles in a shear flow. *European J. Comp. Mechanics*, 19:513–531, 2010.
- [8] W.R. Hwang, M.A. Hulsen, and H. E. H. Meijer. Direct simulation of particle suspensions in sliding bi-periodic frames. *J. Comp. Physics*, 194:742–772, 2004.
- [9] N. Verdon, G. Beaume, A. Lefebvre-Lepot, L. Lobry, and P. Laure. Immersed finite element method for direct numerical simulation of particle suspension in a shear flow. *J. Comp. Physics*, submitted, 2010.

REDUCED-ORDER MODELING OF PARAMETRIZED FINITE ELEMENT SOLUTIONS BY THE POD-ISAT TECHNIQUE. APPLICATION TO AIRCRAFT AIR CONTROL SYSTEMS

Dung Bui*, Mohamed Hamdaoui* and Florian De Vuyst†

* Mathematics Applied to Systems Lab,
Ecole Centrale Paris,
Grande Voie des Vignes, Châtenay-Malabry cedex, France
e-mail: dung.bui@ecp.fr, mohamed.hamdaoui@m4x.org, <http://www.mas.ecp.fr/>

† Center of Applied Mathematics and their Applications (CMLA)
Ecole Normale Supérieure de Cachan,
61, avenue du Président Wilson 94235 Cachan cedex
e-mail: devuyst@cmla.ens-cachan.fr, <http://www.cmla.ens-cachan.fr/>

Key words: Multiphysics Problems, Reduced-Order Modeling (ROM), nonintrusive, In-Situ-Adaptive-Tabulation (ISAT), Proper Orthogonal Decomposition, Aircraft air control system.

Abstract. A combined Proper Orthogonal Decomposition (POD) + In Situ Adaptive Tabulation (ISAT) is proposed for the representation of parameter-dependent solutions of coupled partial differential equations (PDE). The method is tested on a coupled fluid-thermal problem: the design of a simplified aircraft air control system. Furthermore, the control of the method's accuracy is discussed, leading to the metamodeling of the residual itself. The presented POD-ISAT approach provides by its flexibility and robustness an appropriate representation of the solutions for different use cases (sensitivity analysis, optimization, etc.)

1 Introduction

Computational tools are today a success factor in Engineering Design. Finite Elements or Finite Volumes codes become very efficient in the evaluation of criteria at given design points. However, the 'full' exploration of the design space is still a difficult task because of the curse of dimensionality and the weak computing performance available for such applications. Alternative solutions are the use of meta-models or low-order optimal bases that show both accuracy and computational efficiency for particular classes of problems (elliptic problems). Proper Orthogonal Decomposition (POD) [1], Reduced

Basis Method (RBM) [2, 3], LATIN methodology [4] or Proper Generalized Decompositions (PGD) [5] are among the most known computational approaches for dimensionality reduction of PDE solutions. But there are still big issues such as the case of hyperbolic problems, convection-dominated problems, strongly coupled multiphysics problems, high-dimensional design spaces, etc.

The present paper deals with the design of a simplified aircraft air control system depending on inflow and exterior conditions. Navier-Stokes equations are coupled with a thermal equation by means of a buoyancy force (Boussinesq approximation). In this work, an innovating approach called POD-ISAT is presented. It combines Proper Orthogonal Decomposition for the representation of the spatial fields and In Situ Adaptive Tabulation (ISAT) for the local representation of the solution in the design space. This leads to a set of local reduced-order models whose fidelity is controlled by means of trust regions (TR). In Pope's ISAT model [6, 7], ellipsoids of accuracy (EOA) are used and adapted during the learning process of the table. Here we rather use a threshold criterion on a residual. The whole algorithm is detailed in the paper and numerical results show the efficiency of the approach.

2 Mathematical setting

We are interested in the modelling of stationary air circulation and heating conditions in an aircraft cabin. For the sake of simplicity, the flow is supposed two-dimensional and the domain of interest is the cross-section of the fuselage (see figure 1). The air is seen as an incompressible fluid but we take into account buoyancy Archimedes forces due to air heating. So the stationary Navier-Stokes equations with the Boussinesq approximation are considered. At the right hand side of the Navier-Stokes momentum equation (2) appears a buoyancy term depending on the gravity \mathbf{g} and the temperature deviation $(T - T_0)$ from the nominal temperature T_0 . The Navier-Stokes equations are coupled, through this buoyancy term, with a thermal equation that governs the evolution of the temperature of the fluid (equations (1)-(3)). The coefficient κ is the thermal diffusivity of the air.

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega, \tag{1}$$

$$\mathbf{u} \cdot \nabla \mathbf{u} - \nu \Delta \mathbf{u} + \nabla p = \mathbf{g} (1 - \alpha(T - T_0)) \quad \text{in } \Omega, \tag{2}$$

$$\mathbf{u} \cdot \nabla T - \nabla \cdot (\kappa \nabla T) = 0 \quad \text{in } \Omega, \tag{3}$$

In realistic conditions, the reference length L is 1 m, the characteristic speed U is 1 m/s and the kinematic viscosity of the air at 300K is $1.57 \cdot 10^{-5} \text{ m}^2/\text{s}$ so that the Reynolds number is equal to

$$Re = \frac{LU}{\nu} \approx 6.37 \cdot 10^4.$$

The flow regime is turbulent, but, for the sake of simplicity, we do not take into account any turbulence model. Moreover the thermal diffusivity of air at 300K and 1 atm is

$2.22 \cdot 10^{-5} \text{ m}^2/\text{S}$, thus the Péclet number is

$$Pe = \frac{LU}{\kappa} \approx 4.52 \cdot 10^4.$$

Let us consider now the boundary conditions. The cabin boundary is denoted Γ . It is

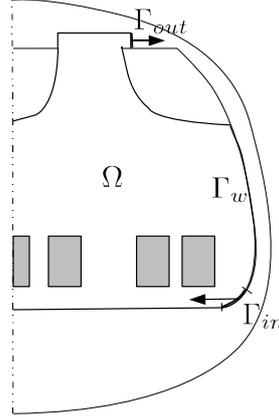


Figure 1: Spatial geometry and domain boundaries (half domain with symmetry axis)

divided into three parts: the inflow boundary Γ_{in} , the outflow Γ_{out} and the wall boundary Γ_w . For the fluid, no slip boundary conditions is used on Γ_w , velocity is imposed at the inflow and constant pressure is given at the outflow:

$$\mathbf{u} = 0 \text{ on } \Gamma_w, \mathbf{u} = \mathbf{u}_{in} \text{ on } \Gamma_{in} \text{ and } p = 0 \text{ on } \Gamma_{out}. \quad (4)$$

For thermal boundary conditions, we use Dirichlet boundary conditions on Γ_{in} with imposed inflow temperature T_{in} . The heat loss at the walls is expressed by inhomogeneous Fourier boundary conditions. The boundary heat flux may depend on the difference between the wall temperature and the exterior temperature. Finally, homogeneous Neumann boundary conditions are written at the outflow:

$$T = T_{in} \text{ on } \Gamma_{in}, \frac{\partial T}{\partial n} = 0 \text{ on } \Gamma_{out}, \kappa \frac{\partial T}{\partial n} = \Phi(T - T_{ext}) \text{ on } \Gamma_w. \quad (5)$$

Possibly, if interior boundaries are defined (like seats for example), then homogeneous Neumann boundary conditions are imposed. The whole system is non linear and the dominating phenomenon is the convection (because of the large Reynolds and Péclet numbers). It is assumed that the domain boundaries are Lipschitz continuous, $\mathbf{u}_{in}, T_{in} \in L^2(\Gamma_{in})$, $\Phi \in C^\infty$, so that (\mathbf{u}, p, T) are searched in $U_{u_{in}} \times L^2(\Omega) \times X_{T_{in}}$, where

$$U_w = \{ \mathbf{v} \in [H^1(\Omega)]^2, \mathbf{v} = 0 \text{ on } \Gamma_w, \mathbf{v} = \mathbf{w} \text{ on } \Gamma_{in} \},$$

$$X_w = \{ \tau \in H^1(\Omega), \tau = w \text{ on } \Gamma_{in} \}.$$

According to some approximate candidates $\tilde{\mathbf{u}} \in U_{u_{in}}$, $\tilde{p} \in L^2(\Omega)$ and $\tilde{T} \in X_{T_{in}}$, one can define a residual functional relative to the test functions $\mathbf{v} \in U_0$, $q \in L^2(\Omega)$ and $\tau \in X_0$:

$$\begin{aligned} R(\tilde{\mathbf{u}}, \tilde{p}, \tilde{T}, \mathbf{v}, q, \tau) &= \int_{\Omega} (\tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}}) \cdot \mathbf{v} \, dx + \int_{\Gamma_{out}} \frac{\partial \tilde{\mathbf{u}}}{\partial \mathbf{n}} \cdot \mathbf{v} \, d\sigma - \int_{\Omega} \tilde{p} \nabla \cdot \mathbf{v} \, dx \\ &- \int_{\Omega} (1 - \alpha(\tilde{T} - T_0)) \mathbf{g} \cdot \mathbf{v} \, dx + \int_{\Omega} \nabla \cdot \tilde{\mathbf{u}} q \, dx + \int_{\Omega} (\tilde{\mathbf{u}} \cdot \nabla \tilde{T}) \tau \, dx \\ &- \int_{\Omega} \kappa \nabla \tilde{T} \cdot \nabla \tau \, dx - \int_{\Gamma_w} \Phi(\tilde{T} - T_{ext}) \tau \, d\sigma \quad \forall (\mathbf{v}, q, \tau) \in U_0 \times L^2(\Omega) \times X_0. \end{aligned} \quad (6)$$

Parameters There are many parameters in the problem for which we would like to know the solutions of (1)-(5). Each parameter may vary in an interval. Among the parameter of interest, let us mention the exterior temperature T_{ext} (that naturally varies during the cruise), the inflow speed u_{in} , the inflow temperature T_{in} . Using dimensionless parameters $\theta_i \in [0, 1]$, $i = 1, \dots, p$, we are looking for the family of fluid-thermal solutions $(\mathbf{u}^\theta = \mathbf{u}(\theta, \cdot), T^\theta = T(\theta, \cdot))_{\theta \in [0,1]^p}$.

3 ROM methodology: general aspects and related works

A Reduced Order Model (ROM) for partial differential equations consists of a low-order representation of the solution by help of an 'optimal' basis and possibly an adaptivity and enrichment process. Considering for example the parameterized temperature field $\mathbf{u}^\theta = \mathbf{u}(\cdot, \theta)$, reduced-order are searched in the form

$$\mathbf{u}^\theta(x) = \mathbf{u}^{lift,\theta}(x) + \sum_{k=1}^K a_k(\theta) \Psi^k(x). \quad (7)$$

The function $\mathbf{u}^{lift,\theta}$ is a lifting function aimed at satisfying some boundary conditions (especially Dirichlet BC), possible depending on θ but quite easy to compute (for example the solution of a linear Stokes Problem). The family $(\Psi^k)_{k=1,\dots,K}$ is the 'optimal' basis. The truncation rank K is expected to be rather small, let us say 10. The expansion coefficients $a_k(\theta)$ are functions depending on the vector parameter $\theta \in [0, 1]^p$. In a ROM methodology, there are two main steps: the design of the basis functions Ψ^k and the learning process of the $a_k(\theta)$.

In the POD snapshot approach [1, 8], some snapshot fields $(\mathbf{u}^i)_{i=1,\dots,N}$ are computed according to a Design of Computer Experiment (DoCE). Then, the POD basis spawns the best linear subspace able to represent the snapshot solutions:

$$\min_{\substack{(\Psi^1, \dots, \Psi^K) \\ (\Psi^k, \Psi^\ell) = \delta_{k\ell}, 1 \leq k \leq \ell \leq N}} \frac{1}{2} \sum_{i=1}^N \left\| \mathbf{u}^i - \sum_{k=1}^K (\mathbf{u}^i, \Psi^k) \Psi^k \right\|^2. \quad (8)$$

In the POD methodology, the (Ψ^k) are said to be an empirical basis because of the empirical choice of the snapshot set (see [9] for a recent analysis on the optimal location of the snapshots). Reduced Basis Methods or RBM [2, 3] are more rigorous approaches where the basis is enriched during an iterative learning process. At a given iteration (k) involving k modes, a $(k + 1)$ th mode Ψ^{k+1} is searched as a best corrector direction corresponding to the the worst case location in the parameter domain. This is a kind of 'min-max' algorithm. RBM involves easy-to-compute accuracy estimators; we refer to the literature ([10, 11, 12]) for this issue. Because of the iterative enrichment process, RBM belongs to the family of greedy algorithms. Let us emphasize that the RBM analysis framework is restricted to elliptic problems. Another and recent approach which knows an increasing interest is the Proper Generalized Decomposition or PGD pioneered by Ladevèze and Chinesta and since extended and used in many fields of applications ([5, 13, 14, 15, 16, 17]). PGD is also a greedy algorithm where the variables are separated. From a level- k model $\tilde{\mathbf{u}}^{(k)}(\theta, \cdot)$, a higher-fidelity model $\tilde{\mathbf{u}}^{(k+1)}(\theta, \cdot)$ in the form

$$\tilde{\mathbf{u}}^{(k+1)}(\theta, \cdot) = \tilde{\mathbf{u}}^{(k)}(\theta, \cdot) + a_1^{(k+1)}(\theta_1) a_2^{(k+1)}(\theta_2) \dots a_p^{(k+1)}(\theta_p) \Psi^{(k+1)}(x), \quad (9)$$

where the one-dimensional functions $a_1^{(k+1)}(\theta_1)$, $a_2^{(k+1)}(\theta_2) \dots a_p^{(k+1)}(\theta_p)$ and the spatial model $\Psi^{(k+1)}(x)$ are searched in an optimal way, for example by a variational principle and a Galerkin projection, see references [5, 18] for more details. Although very promising, PDG still needs investigation especially for parameter problems. It is unclear from the numerical analysis point of view what is the truncation rank K for a given error criterion. Moreover, PGD for the moment is an intrusive approach, what can be a shortcoming in a practical industrial context. PGD also needs more developments in the case of coupled problems. In what follows, we are going to present the POD-ISAT methodology which is an easy-to-implement non-intrusive approach that can be used in an industrial context.

4 The POD-ISAT algorithm

The determination of POD modes by the method of snapshots [1] relies on the computation of some accurate finite element solutions. A very popular physics-based meta-modeling technique, namely the POD-Galerkin approach, consists in carrying out the approximation on the full Finite Element vector fields using POD modes and Galerkin projection [19]. Nevertheless, the main drawback of this method is its intrusive feature: the computational code has to be modified in order to develop a ROM. Besides, the resulting system of ODEs can become unstable or chaotic, misrepresenting the physics [20]. In the present work, we expose a non-intrusive reduced order model based on the combination of the POD method with the famous ISAT algorithm [6, 7].

4.1 Design of Computer Experiment(DoCE)

The parameter space sampling has an important impact on the metamodels accuracy. Commonly used DoCE procedure include Latin Hypercube Sampling (LHS), U-

designs [21] and Lattice Design [22]. A comparison of these methods [23], showed that the Lattice Design method outperforms the two other methods regarding the minimum distance criterion. Let N_{exp} be the number of design sites in the parameter space chosen according to a lattice design procedure. After computing the exact solutions (e.g. the temperature fields) for these design sites by a FE code, an initial snapshot set $\mathcal{S}^{N_{exp}}$ is formed as follows:

$$\mathcal{S}^{N_{exp}} = \{ \mathbf{u}_0^i, i = 1, \dots, N_{exp} \}. \quad (10)$$

4.2 The ISAT algorithm

The purpose of the ISAT algorithm is to tabulate a function $f(x)$, where x and f are of dimension n_x and n_f , respectively. Given a query, x_q , ISAT returns $f^a(x_q)$, an approximation to $f(x_q)$. An essential aspect of ISAT is that the table is built up, not in a pre-processing stage, but in situ (or "on line") as the simulation is being performed.

4.3 Local form of the POD-ISAT ROM

Assume that the current query parameter θ^i becomes a new entry for the table, say the i^{th} entry. The corresponding solution \mathbf{u}^i (e.g. temperature field) is computed by the FE model. Then a local reduced order model is built up at θ^i . This local model consists of an approximate model $\tilde{\mathbf{u}}_{(i)}(\theta)$ and a region of accuracy (ROA), denoted by $\mathcal{E}(i)$. The initial snapshot set $\mathcal{S}^{N_{exp}}$ is used to build a local snapshot set of N_{local} solutions that are centred with respect to \mathbf{u}^i as follows:

$$\mathcal{S}_{(i)}^{N_{local}} = \{ \mathbf{u}_0^j - \mathbf{u}^i, j = 1, \dots, N_{local} \}. \quad (11)$$

Then $K \leq N_{local}$ local POD modes $\Psi_i^k, k \in [1, \dots, N_{local}]$ are computed using this local snapshot set and the local reduced order model $\tilde{\mathbf{u}}_{(i)}^\theta$ reads as follows:

$$\tilde{\mathbf{u}}_{(i)}^\theta(x) = \mathbf{u}_{(i)}^{lift, \theta}(x) + \sum_{k=1}^K a_{(i)}^k(\theta) \Psi_{(i)}^k(x). \quad (12)$$

where the local POD coefficients $a_{(i)}^k$ depend on the parameters θ et θ^i .

4.4 Trust region

For $\theta = \theta^i$, provided that $\forall k \in [1, \dots, K], a_k(\theta^i) = 0$, we have $\tilde{\mathbf{u}}(\cdot, \theta^i) = \mathbf{u}(\cdot, \theta^i)$. Since the POD coefficients are supposed to be continuous, there is a region $\mathcal{E}(i)$, called trust region, such that the residual field satisfies

$$\forall \theta \in \mathcal{E}(i) \quad \|R(\tilde{\mathbf{u}}_{(i)}(\theta))\|_{\mathcal{L}^2}^2 \leq \varepsilon_{tol}^2,$$

with $\varepsilon_{tol} \ll 1$. The trust region $\mathcal{E}(i)$ is approximated by an ellipsoid [6] in \mathbb{R}^p which is defined by $\frac{p^2+p}{2}$ unknown coefficients. The trust region is built by finding out $M > \frac{p^2+p}{2}$ different θ^* next to θ^i such as $|||R(\theta^*)|||_{\mathcal{L}^2}^2 - \varepsilon_{tol}^2|$ is minimal. To tackle this problem, θ^*

is searched in the form $\theta^* = \theta^i + \alpha^* h$, with $\alpha^* \in \mathbb{R}$ and $h \in \mathbb{R}^p$ a fixed unit vector. By choosing M different vectors h , M vectors θ_h^* are computed in parallel on a multi-core machine as the M minimization runs are independent. Then, the M vectors θ_h^* are used to determine an ellipsoid of accuracy (EOA) using Ellipsoidal Toolbox ¹. The size of the EOA depends obviously on the choice of ε_{tol} .

4.5 POD-ISAT algorithm

Initially, the POD coefficients $a_{(i)}^k(\theta)$, $k \in [1, \dots, K]$ (see (12)) are not known, however we have:

$$a_{(i)}^k(\theta^j) = (\mathbf{u}(x, \theta^j), \mathbf{\Psi}_{(i)}^k(x)); k = 1, \dots, K; j = 1, \dots, N_{local}. \quad (13)$$

Using (13), the coefficients $a_{(i)}^k(\theta)$ can be interpolated or approximated by standard robust methods (Moving Least Square (MLS) [24, 25], artificial neural networks (ANN) [26], radial basis functions (RBF) [27] or Kriging approaches).

In this paper, the coefficients $a_{(i)}^k(\theta)$ are determined by minimizing the \mathcal{L}^2 norm of the residual:

$$(a_{(i)}^1, \dots, a_{(i)}^K)(\theta) = \arg \min_{(a_{(i)}^1, \dots, a_{(i)}^K)} \frac{1}{2} \left\| R \left(\mathbf{u}_{(i)}^{lift, \theta}(x) + \sum_{k=1}^K a_{(i)}^k(\theta) \mathbf{\Psi}_{(i)}^k(x) \right) \right\|_{\mathcal{L}^2}^2. \quad (14)$$

The optimization problem (14) is formulated in a low-dimensional space and is easy to solve. As initial guesses $a^k(\theta^i) = 0$, $k \in [1, \dots, K]$ are used, but one can also use interpolated POD coefficients.

Once the $\mathcal{E}(i)$ for θ^i is determined, a sampling $(\theta^j)_{j=1, \dots, M} \in \mathcal{E}(i)$ is generated and the POD coefficients $a_k(\theta^j)$, $k \in [1, \dots, K]$, $j \in [1, \dots, M]$ are computed by minimizing the residual (see (14)). Then, using these coefficients, a kriging interpolation model of $a_k(\theta)$ is built. Thus, the local ROM at point θ^i is completely defined by the EOA characteristics and the kriging interpolation model which are added to the table which is enriched adaptively allowing to cover the design domain [7]. The POD-ISAT algorithm is summarized in Fig 2.

5 Numerical results

Table 1 shows all the control parameters that impact the quality of the POD-ISAT ROM. The most important is ε_{tol} which controls the threshold error of the ROM. Some exact solutions computed with the FE model are plotted in figure 3. The significant CPU costs are outlined in Table 2. The speed-up, defined as the ratio of the time to evaluate a solution with the FE code by the time needed to compute a solution with the ROM, is $1000/0.6 \sim 1700$, which shows the efficiency of the POD-ISAT algorithm. A 100 random

¹<http://www.mathworks.com/matlabcentral/fileexchange/21936-ellipsoidal-toolbox-et>

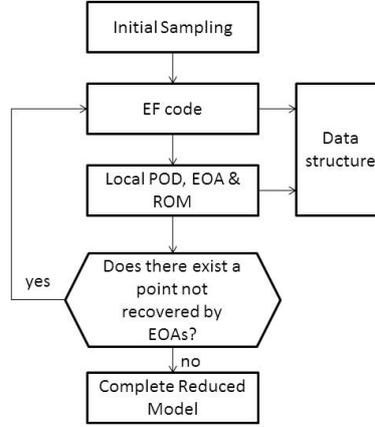


Figure 2: Schematic of the POD-ISAT algorithm

Table 1: Properties of reduced model

N_{exp}	60	Initial DoCE Size
ε_{tol}^2	$5 \cdot 10^{-10}$	Tolerance
N_{local}	10	Number of nearest neighbours for local POD
K	5	Number of POD modes used
p	3	Number of parameters

parameters vectors $\theta^i, i = 1, \dots, 100$ are drawn in the parameter space. Then, the ROM is compared with the FE model using the following error formula:

$$relative\ error = \frac{\|\tilde{\mathbf{u}}(\theta) - \mathbf{u}(\theta)\|_{\mathcal{L}^2}}{\|\mathbf{u}(\theta)\|_{\mathcal{L}^2}}. \quad (15)$$

In Figure 4, where each color represents an EOA, we can see that the mean relative error is about 0.3% and that the maximum relative error is attained for the solution corresponding to θ^{12} (1.2%). Furthermore, θ^{12} , θ^{70} , θ^{81} and θ^{57} are in the same EOA (they have the same color), which shows that their EOA was the worst defined because it is too large and inaccurate. Figure 5 shows the temperature fields of θ^{12} obtained by both reference model and reduced model. In order to better see the difference between both temperature fields, the difference in absolute value is plotted in Fig. 6.

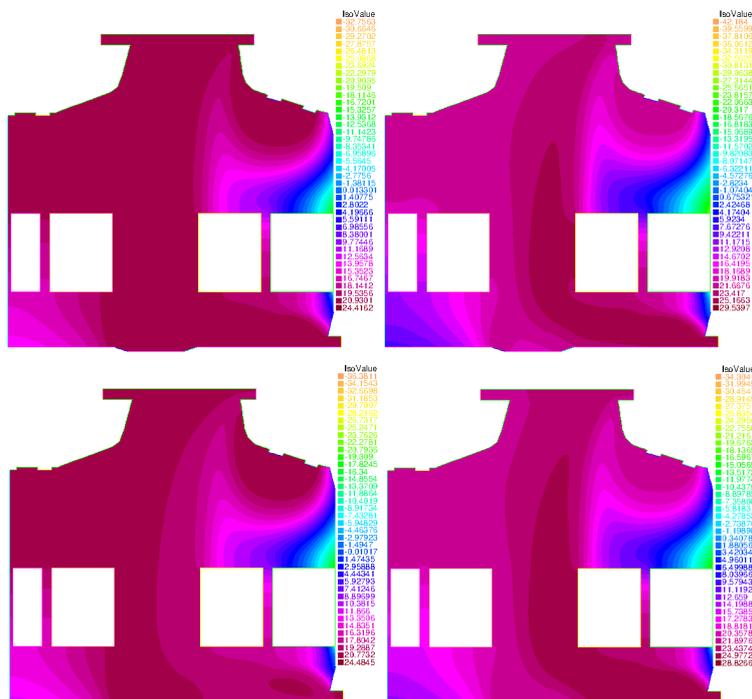


Figure 3: Reference FE solutions for different parameter samples.

	Online	Outline
EOA Building		1200 sec
Kriging model		300 sec
FE computation of $u(x, \theta)$		1000 sec
Retrieve	0.17 sec	
FE computation of $u^{lift}(x, \theta)$	0.43 sec	

Table 2: CPU costs

6 Concluding remarks

In this paper a non intrusive adaptive ROM combining POD and ISAT has been presented on the design of an aircraft control system. The numerical results show that the ROM is both efficient and accurate. However, the EOA can be inaccurately defined and as pointed out by [7] this problem can be tackled by adding an ellipsoid of inaccuracy or EOI around an EOA to control the inaccuracy of the EOA. Furthermore, the ε_{tol} controls the error on the residual and not on the solution itself which is not convenient,

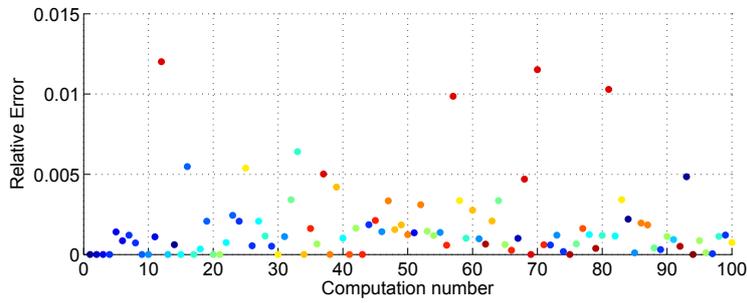


Figure 4: Relative errors for 100 random computations

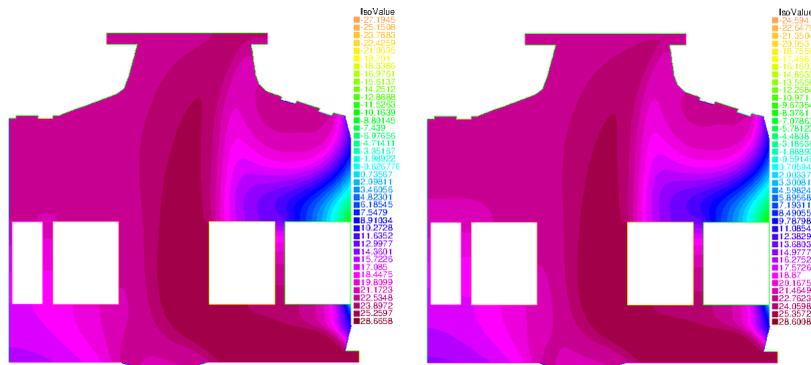


Figure 5: Solution of parameter θ^{12} for both reference model(left) and reduced model(right)

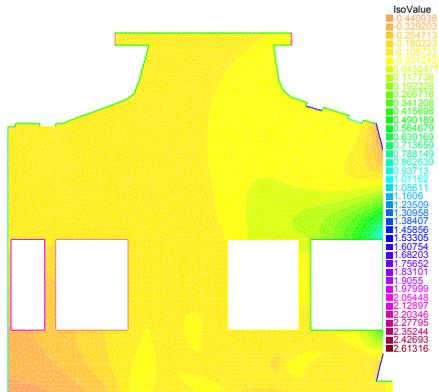


Figure 6: The error field for sample θ^{12} between the reference model and reduced model ($^{\circ}\text{C}$)

which points out the need to consider different approaches to define the error.

Acknowledgements. This work is supported by the industrial platform 'Complex System Design Lab' CSDL, Pôle de Compétitivité System@tic, Paris Région Ile-de-France.

REFERENCES

- [1] G Berkooz, P Holmes, and JL Lumley. The proper orthogonal decomposition in the analysis of turbulent flows. *Annual review of fluid mechanics*, 25:539–575, 1993.
- [2] Yvon Maday and Einar M. Ronquist. A reduced-basis element method. *Comptes Rendus Mathématique*, 335(2):195 – 200, 2002.
- [3] G. Rozza, D. Huynh, and A. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15:229–275, 2007. 10.1007/s11831-008-9019-9.
- [4] Ladevèze Pierre. *Nonlinear Computational Structural Mechanics. New Approaches and Non-Incremental Methods of Calculation*. Mech. Eng. Series. Springer-verlag, 2009.
- [5] A. Ammar, B. Mokdad, F. Chinesta, and R. Keunings. A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. *Journal of Non-Newtonian Fluid Mechanics*, 139(3):153 – 176, 2006.
- [6] S.B. Pope. Computationally efficient implementation of combustion chemistry using in situ adaptive tabulation. *Combustion Theory and Modelling*, 1(1):41–63, 1997.
- [7] Liuyan Lu and Stephen B. Pope. An improved algorithm for in situ adaptive tabulation. *Journal of Computational Physics*, 228(2):361 – 386, 2009.
- [8] F. De Vuyst. *Multidisciplinary Design Optimization in Computational Mechanics*, chapter PDE Metamodeling using Principal Component Analysis. Wiley ISTE, April 2010.
- [9] Kunisch, Karl and Volkwein, Stefan. Optimal snapshot location for computing pod basis functions. *ESAIM: M2AN*, 44(3):509–529, 2010.
- [10] K. Veroy and A. T. Patera. Certified real-time solution of the parametrized steady incompressible Navier-Stokes equations: rigorous reduced-basis a posteriori error bounds. *International Journal for Numerical Methods in Fluids*, 47(8-9):773–788, 2005.
- [11] Gianluigi Rozza and Karen Veroy. On the stability of the reduced basis method for Stokes equations in parametrized domains. *CMAME*, 196(7):1244 – 1260, 2007.
- [12] Simone Deparis. Reduced basis error bound computation of parameter-dependent Navier-Stokes equations by the natural norm approach. *SIAM J. Num. A.*, 46(4):2039–2067, 2008.
- [13] A. Dumon, C. Allery, and A. Ammar. Proper general decomposition (pgd) for the resolution of Navier-Stokes equations. *Journal of Computational Physics*, 230(4):1387 – 1407, 2011.
- [14] F. Chinesta, A. Ammar, A. Leygue, and R. Keunings. An overview of the proper generalized decomposition with applications in computational rheology. *Journal of Non-Newtonian Fluid Mechanics*, In Press, Corrected Proof:–, 2011.

- [15] G. Bonithon, P. Joyot, F. Chinesta, and P. Villon. Non-incremental boundary element discretization of parabolic models based on the use of the proper generalized decompositions. *Engineering Analysis with Boundary Elements*, 35(1):2 – 17, 2011.
- [16] A. Leygue and E. Verron. A first step towards the use of proper general decomposition method for structural optimization. *ACME*, 17:465–472, 2010. 10.1007/s11831-010-9052-3.
- [17] M. Beringhier, M. Gueguen, and J. Grandidier. Solution of strongly coupled multiphysics problems using space-time separated representations. application to thermoviscoelasticity. *Arch. of Comp. Methods in Engineering*, 17:393–401, 2010. 10.1007/s11831-010-9050-5.
- [18] Anthony Nouy. A priori model reduction through proper generalized decomposition for solving time-dependent partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 199(23-24):1603 – 1626, 2010.
- [19] K. Kunisch and S. Volkwein. Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J. on Numerical Analysis*, 40(2):492–515, 2002.
- [20] A. Iollo, S. Lanteri, and J.-A. Désidéri. Stability properties of pod-galerkin approximations for the compressible Navier-Stokes equations. *Theoretical and Computational Fluid Dynamics*, 13:377–396, 2000. 10.1007/s001620050119.
- [21] B. Tang. Orthogonal array-based latin hypercubes. *J. Am. St. Asso.*, (88):13921397, 1993.
- [22] P. Winker K.T. Fang, D.K.J. Lin and Y. Zhang. Uniform previous termdesignnext term: theory and application. *Technometrics*, (42):237248, 2000.
- [23] Dizza Bursztyn and David M. Steinberg. Comparison of designs for computer experiments. *Journal of Statistical Planning and Inference*, 136(3):1103 – 1119, 2006.
- [24] P. Lancaster and K. Salkauskas. Surfaces Generated by Moving Least Squares Methods. *Mathematics of Computation*, 37(155):141–158, 1981.
- [25] Piotr Breitkopf, Hakim Naceur, Alain Rassineux, and Pierre Villon. Moving least squares response surface approximation: Formulation and metal forming applications. *Computers & Structures*, 83(17-18):1411 – 1428, 2005. Advances in Meshfree Methods.
- [26] Dreyfus Gérard. *Neural networks: methodology and applications*. Editions Eyrolles, 2005.
- [27] C. Audouze, F. De Vuyst, and P. B. Nair. Reduced-order modeling of parameterized pdes using time-space-parameter PCA. *IJNME*, 80(8):1025–1057, 2009.

SIMULATION OF RESIDUAL STRESSES IN AN INDUCTION HARDENED ROLL

CLEMENS GROTH^{*}, LUDWIG HELLENTHAL[†]

^{*} CADFEM GmbH
Hannover office
Pelikanstraße 13, 30177 Hannover, Germany
e-mail: cgroth@cadfem.de, www.cadfem.de

[†] WALZEN IRLE GmbH
Hüttenweg 5, 57250 Netphen-Deuz, Germany
email: hellenthal@walzenirle.com, www.walzenirle.com

Key words: Induction, microstructure, transformation induced plasticity, heat treatment, induction hardening, weld simulation

Abstract. Hardening processes are directed towards improving the metallurgical conditions of steel parts. The thermal process is directly coupled with mechanical consequences due to the thermal elongation and plastic mechanical behaviour. The metallurgical transformation of the steel results in material data changes that further complicate the simulation of the process. A simulation application is presented where the heat is induced by an electrical inductor. The part under consideration is a calender roll for a paper production site. The inductor encloses the roll and travels along the length. A coolant flow is applied directly behind the heated zone. Due to the induction, electrical and magnetical effects are significant for the heat source distribution and have to be included in the simulation. An overview over the coupling effects is presented. The simulation steps are directly and sequentially coupled. The result of the simulation is the temperature distribution over time and space and the residual stress distribution. The simulation results are compared with experimental results of residual stresses. The coincidence is satisfactory.

1 INTRODUCTION

The structure under consideration is a roll for the plastics industry. Rolls like the one mentioned here are used pairwise to produce very thin plastic sheets. Due to the intended use of the roll the surface must be manufactured regarding extreme accuracy requirements. This applies to the surface roughness as well as to other dimensions like excentricity and cylindricity under cold and hot (200°C) conditions. Additionally, the operating loads include a significant bending moment and heat expansion, which leads to tension stresses at the roll surface. A compression residual stress can help to neutralize these operating stresses and thus reduce the possibility of crack initiations.

The size and cost of the roll is significant. The simulation is very advantageous to

investigate the heat treatment process and parameter variations beforehand.

2 THE STRUCTURE

The roll outer diameter is 900 mm, the axial length exceeds 2000 mm. After the forging and annealing processes the roll outer surfaces are pre machined closely to final dimensions. The roll as it is used for the hardening process can be regarded as being massive. The final construction has internal drill holes and openings.

The material is steel comparable to 1.6959. It is a ferritic steel that shows microstructural changes when heated above temperatures of 800°C.

3 THE HARDENING PROCESS

The objective of the hardening process is to achieve a hardness level at the roll surface which is needed to fulfill adequate life cycle and wear requirements. In addition to this it is intended to achieve a certain residual stress level to neutralize the operating stresses.

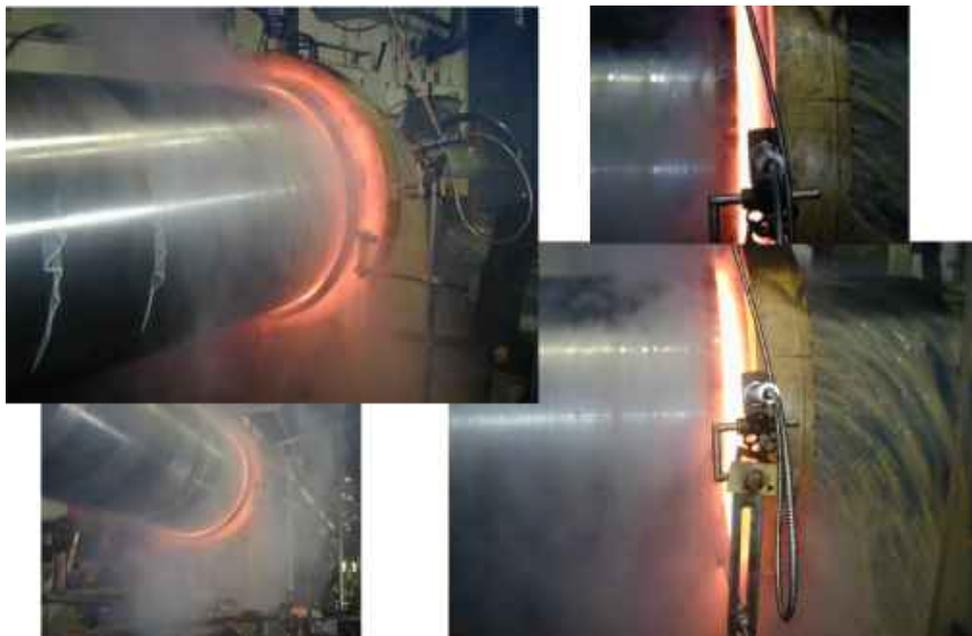


Fig.1 The roll during the hardening process

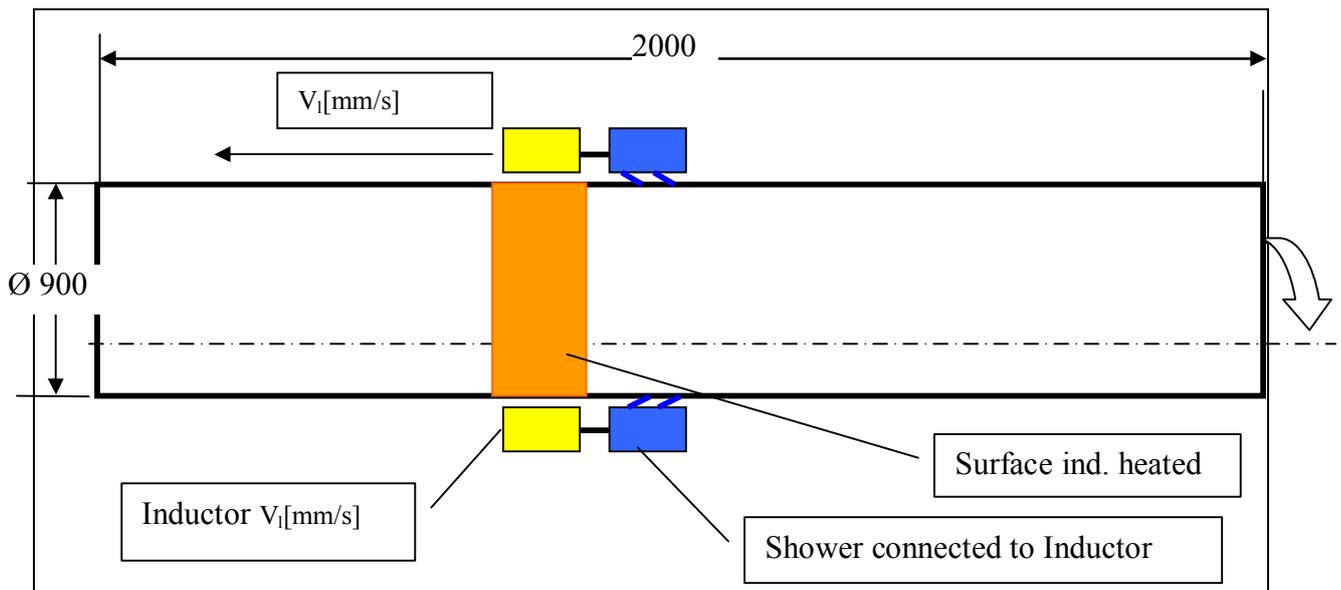


Fig.2 The roll during the hardening process

This hardening process is to be simulated. Economic considerations require the number of variations to be minimized as far as possible. The production of a roll is extremely cost and time consuming. Any variations of process parameters have to be simulated in advance and the results thoroughly evaluated.

The hardening process as shown in Fig. 1 and 2 was performed with the roll as described in chapter 2. The simulation task was to idealize this hardening process and to compare the experimental and simulation results. This comparison was used to qualify the simulation as a decision basis for future variations of the process.

The hardening is done by heating the roll material by means of an induction ring surrounding the roll. An alternating electric current flows through the induction ring and induces a magnetic field around the ring. The induction ring is a hollow rectangular copper section. The magnetic field induces electric currents in the surface region of the roll material. The current losses result in a heat-up. With the appropriate feed rate V_1 of the axial movement of the ring the maximal temperature at the roll surface reaches values of close to 1000°C . The inductor ring is slowly moved axially along the roll axis. Following the inductor ring is a water jet channel which sprays a watery coolant on the roll surface. During this process the roll is rotated slowly so that all circumferential sections are treated identically. Surface measurements are used to adjust the process parameters so that the surface temperature is close to 1000°C .

4 PHYSICAL INTERACTIONS

An overview of the physical interactions is given in Fig. 3. The alternating electric current results in an alternating magnetic field (Magnetisches Feld) around the inductor ring. This

magnetic field induces electric currents in the metallic roll material (Elektrisches Feld). This electric-magnetic field interaction requires the surrounding air and far field regions to be included in the model.

The thermal field (Temperatur-Feld) is coupled bidirectional. The electric-to-thermal coupling includes the current energy losses in the roll material which act as thermal heat source terms. The reverse thermal-to-electric coupling includes the influences of the temperatures on the material data. The microstructural behavior (Gefügeumwandlung, micro structure transformation) influences all of these field effects. The steel used for the roll is a ferritic steel which changes to austenitic material above a certain boundary temperature and shows a reverse change when a certain boundary temperature (usually lower than the first mentioned boundary temperature) is passed. The material behavior is described by CCT diagram. These material changes influence the material data significantly, especially the magnetic behavior is influenced extremely.

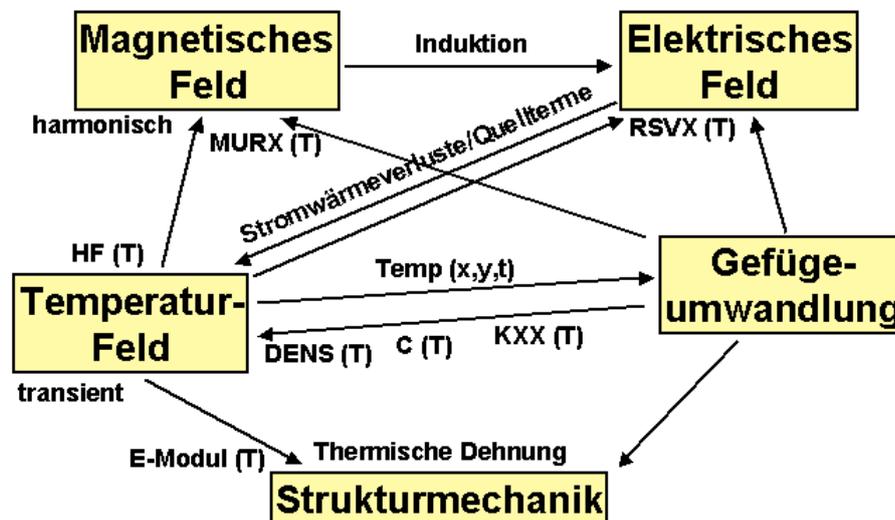


Fig.3 Physical interactions

The mechanical simulation (Strukturmechanik) is coupled uni-directional. The mechanical material values depend on the thermal field and on the microstructural changes, for example the module of elasticity and also the thermal expansion. The thermal expansion is provided as dilatogram graphs, see Fig.6. The graphs are common means to measure the microstructural changes and to document the material behavior.

5 ELECTRIC-MAGNETIC-THERMAL SIMULATION

For the electric-magnetic field a harmonic simulation is used to idealize the alternating current source. The model which was used is shown in Fig. 3. This figure shows the inductor in the centered position along the roll axis. The mesh is axisymmetric and consists of the roll material, the inductor ring cross section and the surrounding air region. Infinite elements are spread over the outer air region boundary to care for the far field effect. To simulate the

movement of the inductor ring along the axis this model is used for a sequence of configurations where the inductor plus the surrounding air are moved axially and coupled to the appropriate wall surface mesh nodes. This sequence is repeated for several positions and thus provides a sequence of field distributions for discrete time points. Parallel to this electromagnetic field analysis a transient thermal field is simulated where the electric energy losses are used as heat source terms. Both analyses exchange the heat source terms and the temperature values (for material data adjustment) sequentially resulting in a kind of "explicit" coupling.

The microstructural changes were taken into account by means of an APDL (ANSYS parametric design language) algorithm that represents the CCT diagram information. Temperatures, temperature gradients and bounding temperature limits are considered.

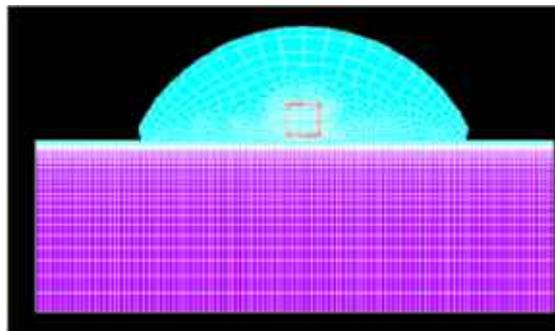


Fig. 3: Model configuration and FEM mesh

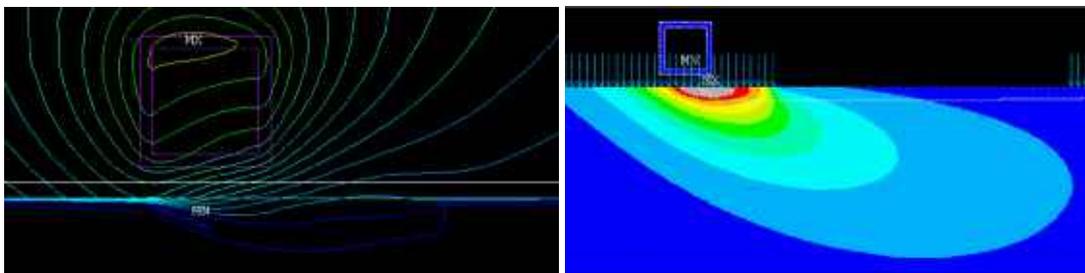


Fig. 4: Magnetic field and temperature distribution at halfway inductor position

6 THE VIRTUAL WELD SIMULATION TOOL (VWS)

The VWS virtual weld simulation tool is a software add-on produced by CADFEM GmbH and partners, funded by BMBF. The tool uses the ANSYS core product as basis and benefits from the numerous capabilities of this general purpose program. The VWS weld simulation tool provides the special features that are necessary to simulate welding processes of any kind. First applications focused on tailored blanks and laser welds. The features of the VWS tool also cover the special requirements of applications like welds with other heat sources (electron beam, friction stir), welds for shipbuilding, commercial vehicles, airplane panels, or heat treatments.

One of the VWS tool features which was used for the induction hardening of the roll was the STAAZ method. This method uses material data representing the thermal expansion of the material with microstructural changes. Common methods to quantify the thermal expansion are dilatograms showing the elongation of a specimen during a heat-up and cool-down cycle. Usually these dilatogram results are used to design the CCT diagram of the material. The STAAZ method requires a set of dilatograms to be provided covering different cases of value *tripel* with

- maximum temperature during the process,
- austenitization time (time span in the austenitic regime) and
- cooling gradient (time between 800°C and 500°C, $t_{8/5}$).

The VWS tool uses the dilatogram information and interpolates the thermal expansion in each element according to the local value *tripel* as extracted from the thermal analysis. The advantage of the STAAZ method is that the primary experimental results are used directly without intermediate steps like designing CCT diagram, extracting analytical function parameters or interpolating in the CCT diagram or other analysis steps. The disadvantage of the STAAZ method is that no information about material fractions like austenite, bainite, martensite or others can be achieved.

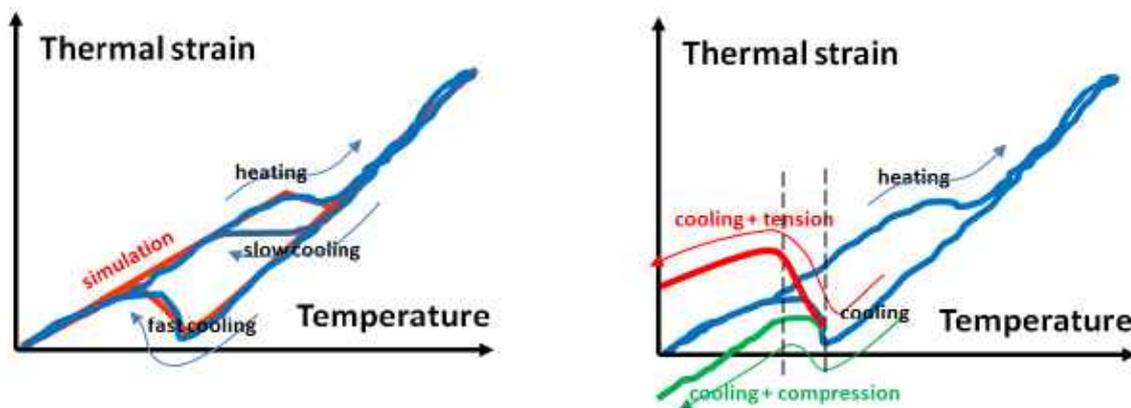


Fig. 5: Typical dilatogram (left: cooling rate, right: TRIP transformation induced plasticity)

Fig. 5 (left) shows a dilatogram result with a variation of cooling rates (fast cooling, slow cooling). When the experiment is done without mechanical stresses (usually the test equipment controls the stress to be zero), no significant residual plastic strains can be observed at the end of the load cycle. The application described here was used to study the influence of the TRIP effect. The transformation induced plasticity, TRIP, is a significant plastic strain contribution which can be observed when a mechanical stress acts during the reverse microstructural change from the austenitic to the ferritic phase. When a mechanical stress is applied, a plastic strain occurs which resembles a parallel shift of the graph, see Fig. 5 (right). This effect is neglected in many cases, because the material data are not available, additional experimental runs are required and its influence on the simulation results is still in discussion.

7 MECHANICAL SIMULATION

The results of the thermal analysis were used for the subsequent mechanical analysis. For this purpose a special postprocessing was run that extracts the significant information of each element of the roll model. This information consists of three values including

- the local maximum temperature during the process,
- the austenitization time (time span in the austenitic regime) and
- the cooling gradient (time between 800°C and 500°C, $t_{8/5}$).

This value triplet was stored for each element for further processing in the VWS tool.

The mechanical simulation was performed with a reduced mesh representing the axial center of the roll. This reduced mesh consisted of one single line of elements from the inner diameter to the outer diameter. This reduced mesh was used to get ideal results for the center portion of the roll without taking end effects into account. The reduced mesh was constrained in axial direction so that plane sections were enforced, with axial displacements suppressed on the right side and axial displacements coupled on the left side. For the material a viso-plastic stress-strain law was chosen which is included in the VWS tool.

The temperatures of the thermal analysis were applied and the STAAZ method activated to idealize the thermal expansion. When the heat-up and cool-down cycle were passed and uniform room temperature was reached, the residual stress distribution was recorded.

Fig. 6 shows the residual stress distribution along the radial direction. The left side of each curve represents the position at the material surface of the roll (outer diameter). The right end of each curve represents a position 40 mm inside the material in radial direction. Both curves show a compression residual stress at the surface which changes to a tension stress at a moderate depth.

Fig. 6 (left) shows the residual stress distribution neglecting the TRIP effect. Fig. 6 (right) displays the residual stress distribution including the TRIP effect.

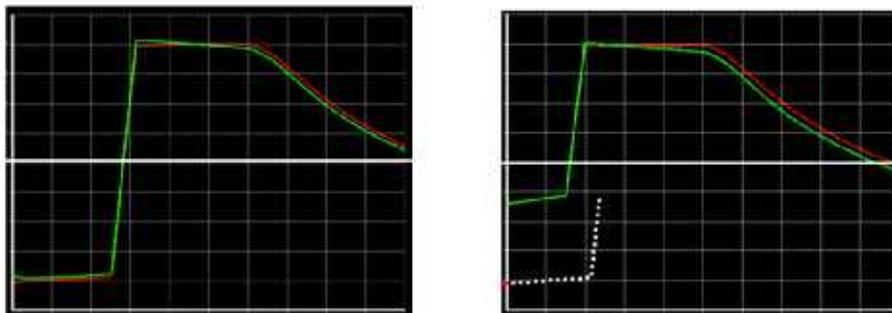


Fig. 6: Residual stress profile neglecting (left side) and including (right side) the TRIP effect

After these initial results were available, the simulation was repeated with slightly adjusted material and process data.

8 RESIDUAL HARDENING STRESS MEASUREMENTS

The residual stress measurements were made by the following methods:

- a) X-Ray stress Analyzer (=> surface 0.05 mm deep results)
- b) Bore hole method (=> surface 0.1-2 mm deep results)
- c) Ring-Core-Method (=> surface 2 – 25 mm deep results)
- d) 2-D Measurement from roll end
- e) SACHS Method (=> roll center bore to 50 mm below the surface)

We compared the results of all 5 methods and found best matching results in our case with the Ring-Core method. This method gives good data for the surface and the most reliable data for the section of interest 0 to 25 mm depth into the surface. The 2-D end plate test also yielded good results but it always requires an additional numerical calculation.

9 COMPARISON OF SIMULATION AND MEASUREMENT RESULTS

The following diagram, Fig. 7, shows the simulated stress distribution (including TRIP) and the measured values measured with the Ring-Core method. The comparison is satisfactory.

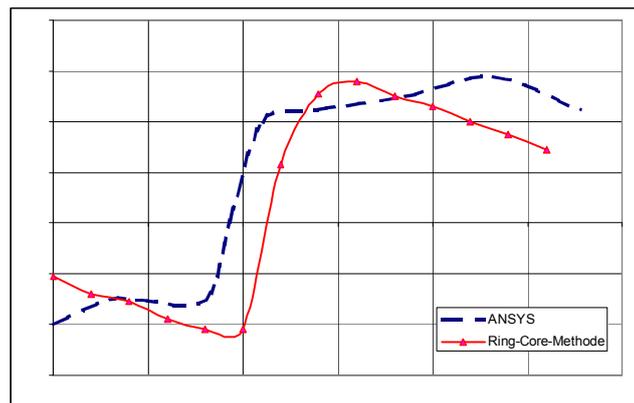


Fig. 7: Resulting residual stress distribution, simulation vrs measurements

This method now allows to optimize the hardening parameters to get the desired optimised residual stress distribution which is needed for the special operating conditions of a roll.

REFERENCES

- [1] R. Ossenbrink and V. Michailov, “Simulation der Gefügewandlung mit der STAAZ-Methode”, 23. *CADFEM Users Meeting*, Bonn, 2.6.7, (2005).

SOLUTION OF DENDRITIC GROWTH IN A BINARY ALLOY BY A NOVEL POINT AUTOMATA METHOD

A.Z. LORBIECKA* AND B. ŠARLER†

* University of Nova Gorica
Vipavska 13, Rožna Dolina
Nova Gorica, Slovenia
e-mail: alorbiecka@ung.si

† University of Nova Gorica
Vipavska 13, Rožna Dolina
Nova Gorica, Slovenia
email: bozidar.sarler@ung.si

Key words: Cellular Automata (CA), Point Automata (PA), Dendritic Growth, Heat and Mass Transfer.

Abstract. The aim of this paper is simulation of thermally induced liquid-solid dendritic growth in a binary alloy (Fe-0.6%C) steel in two dimensions by a coupled deterministic continuum mechanics heat and species transfer model and a stochastic localized phase change kinetics model that takes into account the undercooling, curvature, kinetic, and thermodynamic anisotropy. The stochastic model receives temperature and concentration information from the deterministic model and the deterministic heat and species diffusion equations receive the solid fraction information from the stochastic model. The heat and species transfer models are solved on a regular grid by the standard explicit Finite Difference Method (FDM). The phase-change kinetics model is solved by the novel Point Automata (PA) approach. The PA method was developed and introduced [1,2] in order to circumvent the mesh anisotropy problem, associated with the classical Cellular Automata (CA) method. Dendritic structures are in the CA approach sensitive on the relative angle between the cell structure and the preferential crystal growth direction which is not physical. The CA approach used in the paper for reference comparison is established on quadratic cells and Neumann neighborhood. The PA approach is established on randomly distributed points and neighborhood configuration, similar as appears in meshless methods. Both methods provide same results in case of regular PA node arrangements and neighborhood configuration with five points. A comparison between both stochastic approaches has been made with respect to dendritic growth with different orientations of crystallographic angles. It is demonstrated that the new PA method can cope with dendritic growth of a binary alloy in any direction which is not the case with the CA method.

1 INTRODUCTION

Numerical modeling is having an increasingly important role in studies of dendritic growth

during solidification [3]. The simulation of microstructure based on the CA technique can reproduce most of the dendritic features observed experimentally with the acceptable computational efficiency. However, simple CA models are not capable of reproducing the typical growth features of dendrites when the primary branches do not coincide with the preferential mesh orientations. The reason for that is that the simple CA models suffer from the strong impact of the mesh orientation. It does not matter which crystallographic orientation will be chosen, the CA will always shift the dendrite with respect to the grid axis. The crystallographic orientation axes of different dendrites have in general different angles with respect to the coordinate system. We use a novel Point Automata (PA) method in this paper. It follows the CA concept and is able to solve the mentioned crystallographic orientation problem. A basic feature of this method is the random distribution of the nodes in the domain instead of using regular cells, which leads to different distances between the nodes and different neighborhood configurations for each of them. This new concept was first proposed by Janssens for modeling the recrystallization [4,5]. The first results of the dendritic growth with various orientations based on the PA method have been developed in [1] for pure metals. The PA algorithm, developed in the present paper, is able to obtain the dendritic morphology of solidifying binary alloy (Fe-0.6% C steel is taken as an example), by solving the heat and mass transfer equations, coupled with the solid fraction field through the calculations of the crystal growth velocity, interface curvature, thermodynamic and kinetic anisotropy, respectively. Previous classical CA solutions of the dendritic growth of the Fe-0.6% C steel are demonstrated in [6,7]. The present paper is structured in the following way: the governing equations of the heat and mass transfer model are defined first, followed by the description of the stochastic model. The solution of temperature, concentration field and solid fraction is explained afterwards. Finally, the numerical results of FDM-PA method are shown and compared with the results of the FDM-CA method.

2 MODEL DESCRIPTION

Consider a two dimensional domain Ω with boundary Γ filled with a binary phase change material which consists of at least two phases, solid and liquid, separated by an interfacial region.

2.1 Heat and species transport

2.1.1 Heat diffusion

The following heat transport equation is solved first:

$$\frac{\partial}{\partial t}(\rho h) = \nabla \cdot (\lambda \nabla T) \quad (1)$$

where ρ , h , λ , T represent material density, specific enthalpy, thermal conductivity and temperature, respectively. The specific enthalpy is constituted as $h = c_p T + f_l L$, where c_p , L , f_l represent the specific heat, the latent heat and liquid fraction, respectively. All material

properties are assumed constant for simulation simplicity. We search for the temperature at time $t_0 + \Delta t$ by assuming the initial conditions:

$$T(\mathbf{p}, t_0) = T_0(\mathbf{p}); \mathbf{p} \in \Omega; f_s(\mathbf{p}, t_0) = f_{s0}(\mathbf{p}); \mathbf{p} \in \Omega \quad (2)$$

(where \mathbf{p} represents the position vector) and Neumann boundary conditions:

$$\frac{\partial T}{\partial \mathbf{n}}(\mathbf{p}, t) = F(\mathbf{p}, t); \mathbf{p} \in \Gamma, t_0 < t \leq t_0 + \Delta t \quad (3)$$

where \mathbf{n} represents the normal on Γ and T_0, f_{s0}, F represent known function.

2.1.2 Species diffusion

The solution of the heat transfer equation is followed by the solution of the solute transfer equation. The governing equation for the solute transfer in both solid and liquid phases is formulated in terms of mixture concentration [8]:

$$c = f_s c_s + (1 - f_s) c_l \quad (4)$$

$$\frac{\partial c}{\partial t} = \nabla \cdot (D \nabla c) - c(1 - k_p) \frac{\partial f_s}{\partial t} \quad (5)$$

where D stands for solute diffusion coefficient, defined as $D = f_s D_s + (1 - f_s) D_l$ with D_s and D_l defining the solute diffusion coefficients in solid and liquid, respectively. It is assumed that the concentrations of solid and liquid at the interface are in equilibrium i.e. $c_s = k_p c_l$ where k_p, c_s and c_l are the partition coefficient, concentration in the solid and liquid phase, respectively. We search for the concentration c at time $t_0 + \Delta t$ by assuming the initial and Neumann boundary conditions:

$$c(\mathbf{p}, t_0) = c_0(\mathbf{p}); \mathbf{p} \in \Omega \quad (6)$$

$$\frac{\partial c}{\partial \mathbf{n}}(\mathbf{p}, t) = F(\mathbf{p}, t); \mathbf{p} \in \Gamma, t_0 < t \leq t_0 + \Delta t \quad (7)$$

2.2 Solid fraction calculations

2.2.1 Interface undercooling

The phase change situation can be achieved by undercooling a liquid below its liquid temperature. When a solid seed is placed in such an undercooled melt, solidification will be initiated. Due to crystal anisotropy and perturbations in the system, the growth of the solid

from the seed will not be uniform and an equiaxed dendritic crystal will form. The solid liquid interface is undercooled to the temperature T_f defined as [9]:

$$T_f = T_l + m(c_l - c_0) - \Gamma K \quad (8)$$

where Γ , K , m are the Gibbs-Thomson coefficient, the interface curvature and the liquidus slope, respectively.

2.2.2 Growth velocity

The growth process is driven by the local undercooling. The interface growth velocity is given by the classical sharp model [10]:

$$V_g^*(\mathbf{p}, t) = \mu_K (T_f - T(\mathbf{p}, t)); \mathbf{p} \in \Gamma_{s,l} \quad (9)$$

where V_g^* , μ_K , $\Gamma_{s,l}$ are the growth velocity, interface kinetics coefficient and the solid liquid interface, respectively. Dendrites always grow in the specific crystallographic orientations. Therefore it is necessary to consider anisotropy in either the interfacial kinetics or surface energy (or both). The present model accounts for the anisotropy in both kinetics.

2.2.3 Thermodynamic anisotropy

The Gibbs-Thomson coefficient can be evaluated by taking into account the thermodynamic anisotropy related to the crystal orientation and type as follows:

$$\Gamma = \bar{\Gamma} \left[1 - \delta_t \cos \left[S(\theta - \theta_{def}) \right] \right] \quad (10)$$

where S , θ , θ_{def} , δ_t , $\bar{\Gamma}$ represent factors which control the number of preferential directions of the material's anisotropy ($S = 0$ for the isotropic case, $S = 4$ for four fold anisotropy and so on), growth angle (angle between the y coordinate and the line that connects the centre of the mass of the dendrite and point at $\Gamma_{s,l}$), the preferential crystallographic orientation, thermodynamic anisotropy coefficient and the average Gibbs-Thomson coefficient, respectively.

2.2.4 Kinetic anisotropy

The crystal growth velocity is calculated according to the crystal orientation by taking into the consideration the crystal growth direction θ and the preferred orientation θ_{def} . The crystal growth velocity follows:

$$V = V_g^*(\mathbf{p}, t) \left[1 + \delta_k \cos \left(S(\theta - \theta_{def}) \right) \right] \quad (11)$$

where δ_k represents the degree of the kinetic anisotropy.

3 COUPLING

The represented numerical model consists of two schemes: FDM for evaluating the heat and mass transport (Section 2.1) and a novel PA method for simulating the phase change kinetics (Section 2.2). The calculation domain is divided into FDM nodes used for the calculation of the temperature and solute profiles and random PA points to calculate the solid fraction field. Each four regularly spaced FDM nodes include one random point between them (or a CA cell located in the center between them), see Figure 1. The number of points in FDM mesh in x and y directions is N . At one time step, heat and solute fields are first calculated for the FDM nodes. These values need to be interpolated from the FDM nodes to PA points (or CA cells in case of the CA method). Then, based on the obtained profiles, the local undercooling and the growth velocity of the interface are calculated using Equations 8 and 9. Afterwards, a new solid fraction f_s for PA points (or CA cells in case of CA method) is calculated. At the end of the time step are the temperature field and solute field are updated, based on the new solid fraction profile. The procedure is repeated in the next time step by using the calculated temperature, concentration, and solid fraction fields as initial data.

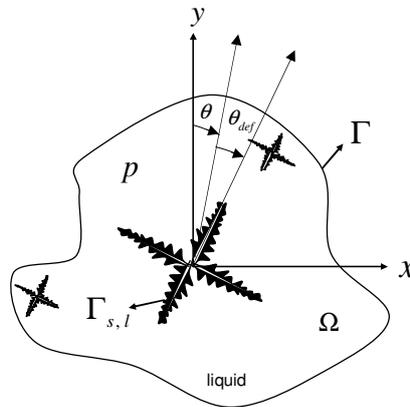


Figure 1: Scheme of the dendritic growth

3.1 Solution of temperature and concentration field

The solution of the temperature field is performed by the simple explicit FDM scheme through the following discretised equation:

$$T_{i,j} = T_{0i,j} + \frac{\Delta t \lambda}{\rho c_p} \left(\left[(T_{0i-1,j} - 2T_{0i,j} + T_{0i+1,j}) / (\Delta x^2) \right] + \left[(T_{0i,j-1} - 2T_{0i,j} + T_{0i,j+1}) / (\Delta y^2) \right] \right) + \frac{L}{c_p} (f_{si,j} - f_{0si,j}) \quad (12)$$

for $i = 2, 3, \dots, N-1$ and $j = 2, 3, \dots, N-1$, where Δt , $f_{0si,j}$, $T_{0i,j}$, $T_{0i+1,j}$, $T_{0i-1,j}$, $T_{0i,j+1}$, $T_{0i,j-1}$, ρ , h , λ represent the time step, initial solid fraction, initial temperature in the FDM central, east, west, north and south nodes, material density, specific enthalpy and thermal conductivity, respectively. The solution of the concentration field is performed by the simple explicit FDM scheme through the following discretised equation:

$$c_{i,j} = c_{0i,j} + \Delta t D \left[\left(c_{0i-1,j} - 2c_{0i,j} + c_{0i+1,j} \right) / (\Delta x^2) \right] + \left[\left(c_{0i,j-1} - 2c_{0i,j} + c_{0i,j+1} \right) / (\Delta y^2) \right] - c_{0i,j} \left[(1-k_p) (f_{si,j} - f_{0si,j}) \right] \quad (13)$$

for $i = 2, 3, \dots, N-1$ and $j = 2, 3, \dots, N-1$, where Δt , $f_{0si,j}$, $c_{0i,j}$, $c_{0i+1,j}$, $c_{0i-1,j}$, $c_{0i,j+1}$, $c_{0i,j-1}$ are the time step, initial solid fraction, initial concentration in the FDM central, east, west, north and south nodes, respectively. The obtained values of concentration on regular FDM grid are in each time step transferred to random PA grid (or regular CA grid). The rejected solute amount $\Delta c = c_l - k_p c_l$ is added to the liquid points in the surrounding neighbors which fall into the circle of $R_{C-H} = 20a$, where a represents the typical mesh distance for each cell (point) separately. Thus, the overall solute in the domain can be kept consistent. A detailed description of FDM-PA (FDM-CA) transfer of temperature and PA-FDM (CA-FDM) transfer of solid fraction from the regular to the random grids and vice versa is elaborated in [1]. The same algorithms are used in the present paper.

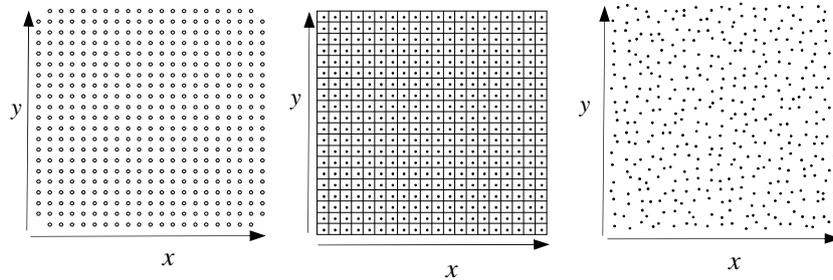


Figure 2: Scheme of space discretization: (left) FDM nodes with $N = 21$, (middle), CA cells with $n = 20$, (right) PA nodes with $n = 20$

3.2 Interface curvature calculations

The interface curvature is approximated by the counting cell procedure developed by Sasikumar and Sreenivasan [11]. The expression for PA is derived from the expression of the CA method by assuming the average node distance \bar{a} instead of the regular node distance a .

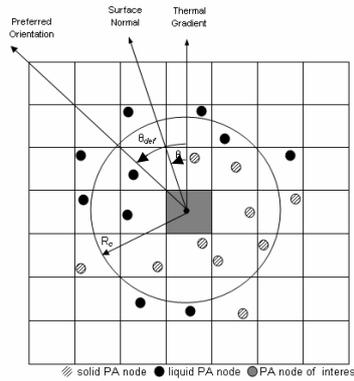


Figure 3: A scheme of curvature calculation in the PA method with $R_c = 2a$ (example with: $N_{sPA} = 7$ and $N_{lPA} = 11$)

The expression for curvature K is given by the formula:

$$K = \frac{1}{a} \left(1 - \frac{2N_{sPA}}{N_{tPA}} \right) \quad (14)$$

where N_{sPA} and N_{tPA} are the number of random points whose centres fall inside the circle of assumed radius R_c and the total number of points whose centres fall inside the circle, respectively.

Some fluctuations need to be introduced into the calculations in order to avoid the symmetric shape of the dendrite in the conventional CA approach. Thermal noises are usually presented by putting the random fluctuations into the calculations of latent heat, undercooling temperature or velocity [12]. In the present study it is not necessary to put any thermal fluctuations in the PA method. The random node arrangements in the PA method replace the thermal fluctuations of the CA method.

In the CA approach only the closest, Moore neighbourhood configuration [3] has been analysed. In the PA method, solid grows with respect to the ‘neighbourhood’ configuration which is now associated with the position of the neighbouring PA nodes which fall into a circle with radius R_H [2]. The radius of neighbourhood should be kept at a minimum of 1.5 μm in case of $a = 0.5 \mu\text{m}$. For smaller values the dendritic shapes become distorted and the preferred orientations are lost.

4 NUMERICAL EXAMPLE

4.1 Definition of test case

A dendritic growth of Fe-0.6% C steel solidifying into an undercooled melt is simulated. The square computational domain with length 350 μm is divided into 701x701 FDM nodes and 700x700 randomly located PA points. Each four FDM nodes involve one PA node (randomly located between the four FDM nodes). At the beginning of the simulation nucleus (es) is assigned with the preferential orientation(s) and start to growth with respect to above algorithm. The initial solid concentration for point is assumed to be $c_s = k_p c_0$ with temperature $T_l = 1490^\circ\text{C}$ at $c_0 = 0.6 \%$ wt. The nominal parameters used in all presented simulations are listed in Table 1. Varied data are presented in Table 2. In the numerical simulation the reference CA method has been used. The details of this method and numerical implementation are elaborated in [1].

4.2 Results

The numerical examples in the present paper are solved by the FDM based temperature and concentration calculations and PA based solid fraction calculations. Our testing was primarily focused on the growth of the dendrite at different orientations by the novel PA method coupled with the heat and mass transfer calculations. To achieve the same length for primary and secondary dendrite arms in PA method as in the CA method, an empirical factor, which

multiplies the calculated velocity in the PA method, has to be introduced. It can be shown that putting a factor of 1.25 in the growth velocity calculations for PA, the branches will have the same length in both methods. Detailed discussion of model parameters was elaborated in [1]. The following numerical examples are shown:

- Case 1 and Case 2 represent the dendritic growth process simulated for Fe-0.6% C steel for undercooling of $\Delta T = 250$ °C by the FDM-CA and FDM-PA methods. The calculation time of both methods is practically the same (Figure 5).
- From Case 3 to 4 the dendritic growth process is simulated for Fe-0.6% C steel for two different preferential orientations $\theta_{def} = 15^\circ$ and $\theta_{def} = 32^\circ$ by the FDM-PA method (Figure 6).
- Case 5 represents six dendrites growing simultaneously simulated by the FDM-PA method.
- On Figure 4 the concentration profiles along the primary dendrite arms obtained by the FDM-CA method are depicted (for Case 1). (Figure 7).

Table 1: Nominal parameters used in the calculations for Fe 0.6% C

Symbol	Value	Unit
ρ	7300	kg/m ³
T_l	1490	°C
λ	30	W/mK
c_p	800	J/kgK
L	2.7×10^{-5}	J/kg
D_s	5.0×10^{-10}	m ² /s
D_l	2.0×10^{-9}	m ² /s
c_0	0.6	%
k_p	0.34	1
Δt	7.65×10^{-10}	s
$\bar{\Gamma}$	1.9×10^{-7}	Km
δ_k	0.75	1
S	4	1
R_c	1.5	μm
R_H	2	μm
μ_K	0.2	m/sK
l	350	μm
n	700	PA nodes/ CA cells
N	701	FDM nodes

Table 2: Varied parameters in different cases

Case	Method	angle	ΔT
CASE 1	FDM-CA	$\theta_{def} = 0^\circ$	$\Delta T = 250^\circ \text{C}$
CASE 2	FDM-CA	$\theta_{def} = 0^\circ$	$\Delta T = 250^\circ \text{C}$
CASE 3	FDM-PA	$\theta_{def} = 15^\circ$	$\Delta T = 150^\circ \text{C}$
CASE 4	FDM-PA	$\theta_{def} = 32^\circ$	$\Delta T = 150^\circ \text{C}$
CASE 5	FDM-PA	$4^\circ, 14^\circ, 33^\circ, 28^\circ, 43^\circ, 8^\circ, 0^\circ$	$\Delta T = 150^\circ \text{C}$

5 CONCLUSIONS

In this paper the temperature and mass transfer equations are coupled with the novel PA method to calculate the solid fraction field in the dendritic growth. Advantages of the developed PA method are:

- No need for mesh generation or polygonisation. Only the node arrangement has to be generated, but without any geometrical connection between the nodes.
- In the new PA method is the microstructure evolution solved with respect to the location of the points (not polygons) on the computational domain.
- The random grid PA method allows rotating dendrites in any direction since it has a limited anisotropy of the node arrangements.
- PA method offers a simple and powerful approach of CA type simulations. It is shown that both methods are able to qualitatively and quantitatively model a diverse range of solidification phenomena in almost the same calculation time.
- The dimension of the neighborhood radius and generation of the random node arrangement has to be chosen carefully in order to be able to rotate the dendrite.
- Straightforward node refinement possibility.
- Straightforward extension to 3-D.

ACKNOWLEDGEMENT

This paper forms a part of the project J2-0099 Multiscale Modelling of Liquid-Solid Processes. Financial support from Slovenian Grant Agency is gratefully acknowledged.

REFERENCES

- [1] Lorbiecka, A.Z. and Šarler, B. Simulation of dendritic growth with different orientation by using the point automata method. *Computers, Materials, Continua* (2010) **18**:69-104.
- [2] Lorbiecka, A.Z. and Šarler, B. A sensitivity study of grain growth model for prediction of ECT and CET transformations in continuous casting of steel. *Materials Science Forum* (2010) **649**: 373-378.

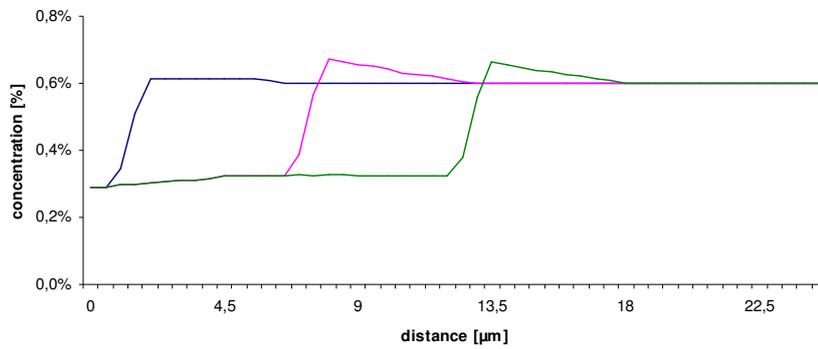


Figure 4: Concentration profiles along the primary dendrites arms obtained by the FDM-PA method for different solidification time: 0.15×10^{-6} [s] (blue line), 1.37×10^{-6} [s] (pink line) and 2.4×10^{-6} [s] (green line)

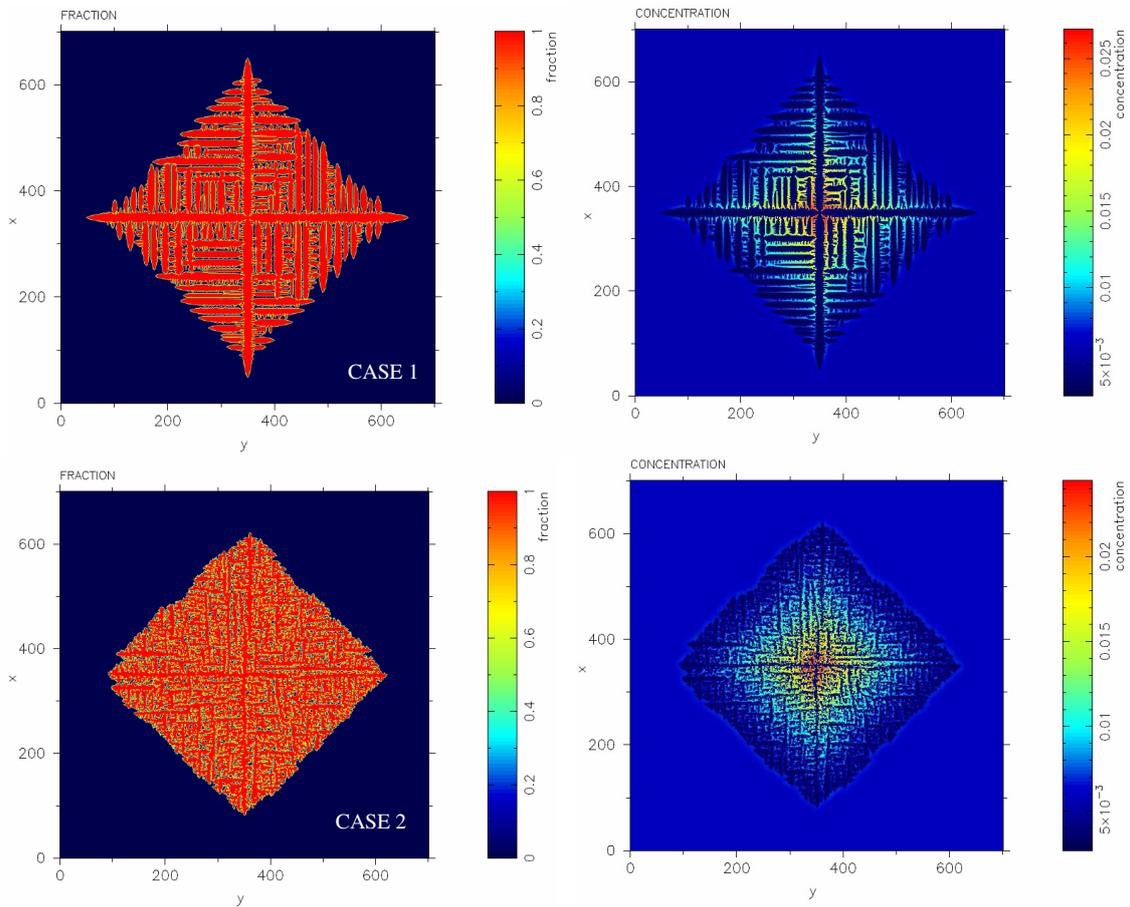


Figure 5: Solid fraction and concentration fields simulated by FDM-CA method (Case 1) and FDM-PA method (Case 2) for undercooling temperature $\Delta T = 250^\circ \text{C}$ after 0.76×10^{-5} [s]

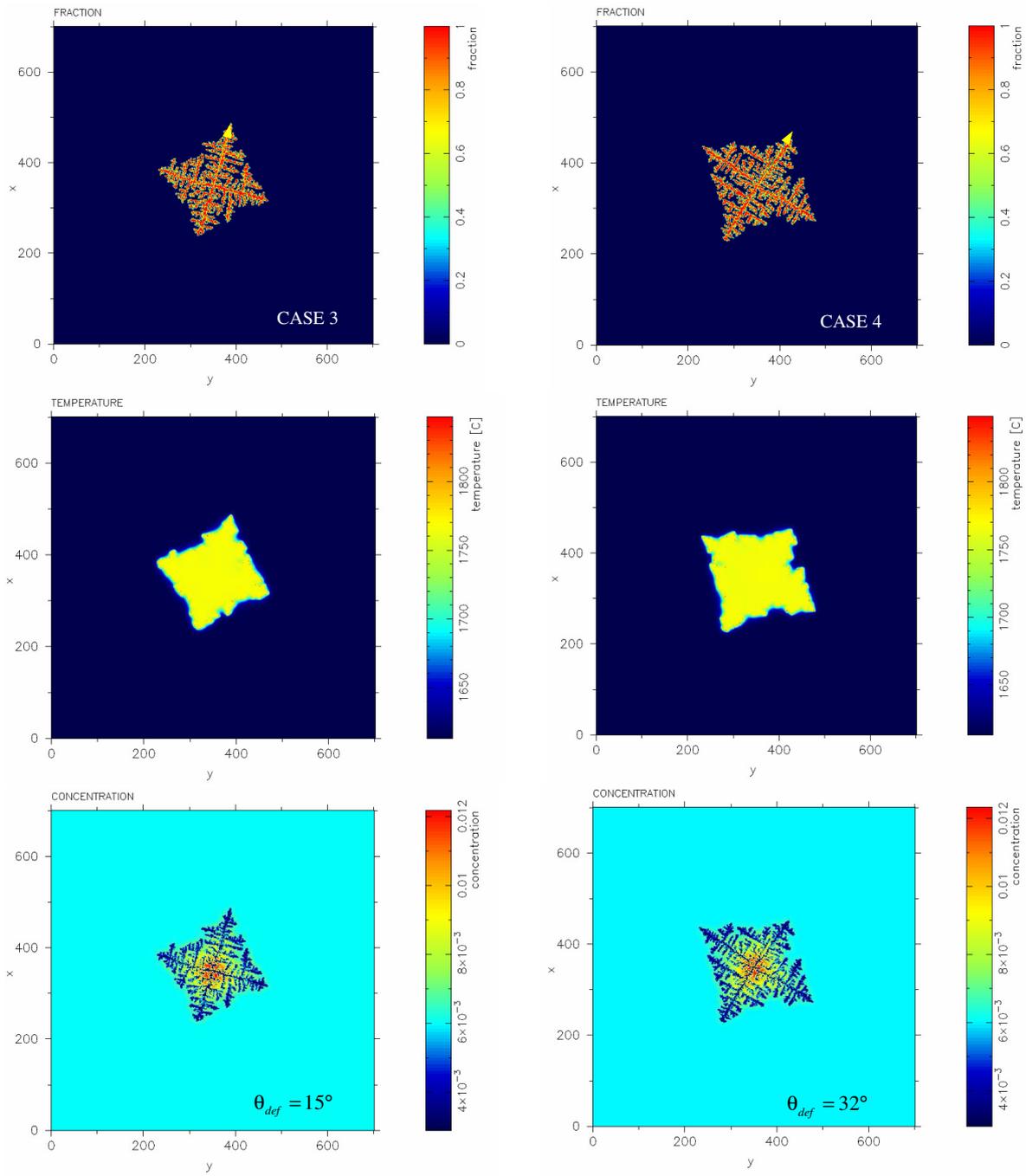


Figure 6: Solid fraction, temperature and concentration fields simulated by FDM-PA method for two different preferential orientations: Case 3 ($\theta_{def} = 15^\circ$) and Case 4 ($\theta_{def} = 32^\circ$) after 1.53×10^{-5} [s]

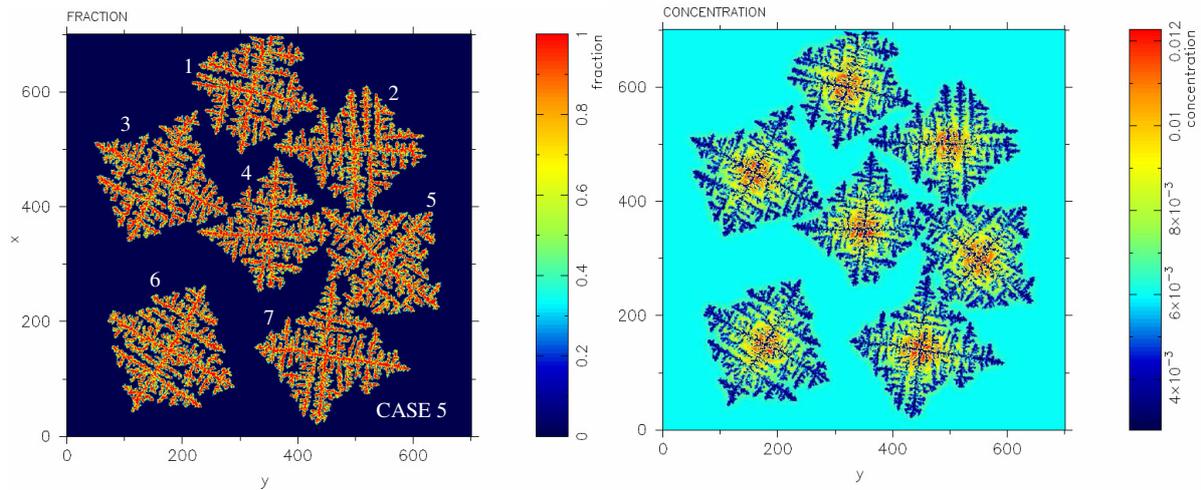


Figure 7: Seven dendrites growing simultaneously simulated by the FDM-PA method (Case 5) for 14° (1), 0° (2), 28° (3), 4° (4), 43° (5), 33° (6), 8° (7) preferred orientations after 1.53×10^{-5} [s]

- [3] Nastac, L. *Modelling and Simulation of Microstructure Evolution in Solidifying Alloys*. Kluwer Academic Publishers (2004).
- [4] Janssens, K.G.F. Three dimensional, space-time coupled cellular automata for the simulation of recrystallization and grain growth. *Modelling and Simulation in Materials Science and Engineering* (2003) **11**: 157-171.
- [5] Janssens, K.G.F. An introductory review of cellular automata modelling of moving grain boundaries in polycrystalline materials. *Mathematics and Computers in Simulations* (2010) **80**: 1361-1381.
- [6] Nastac, L. Numerical modeling of solidification and segregation patterns in cast alloys. *Acta Metallurgica* (1999) **47**: 4252-4262.
- [7] Daming, L., Ruo, L. and Zhang, P. A new coupled model for alloy solidification. *Science in China Ser. A. Mathematics* (2004) **47**: 41-52.
- [8] Zhu, M.Z., Dai, T., Lee, S.Y. and Hong, C.P. Modeling of solutal dendritic growth with melt convection. *Science Direct* (2008) **55**: 1620-1628.
- [9] Saito, Y., Goldbeck-Wood, G. and Muller-Krumbhaar, H. Numerical simulation of dendritic growth. *Physical Review* (1988) **33**: 2148-2157.
- [10] Zhu, M.F. and Hong, C.P. A modified cellular automata model for the simulation of dendritic growth in solidification of alloys. *ISIJ International* (2001) **5**: 436-445.
- [11] Sasikumar, R. and Sreenivasan, R. Two-dimensional simulation of dendrite morphology. *Acta metall. Mater* (1994) **42**: 2381-2386.
- [12] Voller, V.R. An enthalpy method for modelling dendritic growth in a binary alloy. *International Journal of Heat and Mass Transfer* (2008) **52**: 823-834.

SOLUTION OF INTEGRAL EQUATION IN SCATTERING ANALYSIS OF CONDUCTING BODIES OF REVOLUTION BY MOM WITH FIRST TYPE ELLIPTIC INTEGRALS

VIDAL C. F. V. P. * AND RESENDE U. C. †

* Dept. Electric Engineering (PPGEL)
Centro Federal de Educação Tecnológica e Minas Gerais CEFET-MG
Av. Amazonas, 7675, Belo Horizonte, MG, Brazil CEP: 35510-000
e-mail: camila@tormax.com.br, www.cefetmg.br

* Dept Engineering
UNA- MG
Av. Afonso Vaz de Melo, n 456, Belo Horizonte, MG, Brazil CEP: 36640-070
e-mail: camila.vidal@una.br, www.una.br

† Dept. Electric Engineering (PPGEL)
Centro Federal de Educação Tecnológica e Minas Gerais CEFET-MG
Av. Amazonas, 7675, Belo Horizonte, MG, Brazil CEP: 35510-000
email: resendeursula@des.cefetmg.br, www.cefetmg.br

Key words: Electromagnetic Scattering, Electric Field Integral Equation (EFIELD), Elliptic Integral, Method of moments.

Abstract. The electromagnetic scattering in conducted bodies is of great study in the literature once its importance on practical modern engineering applications. The use of MoM with the integral equations are one of the ways to solve the electric field (EFIE) and magnetic field (MFIE) equations and in this process, these integrals present some singularities. In this work the analysis of integral equation, for situations where the singularities are not present, is conducted using Gaussian Quadrature. For cases where singularities are present, a technique to remove these singularities is presented. In this technique, the singular part of the integrand is reduced to a first type elliptic integral and evaluated in a closer form or replaced by an accuracy approximation.

1 INTRODUCTION

The effect known as electromagnetic scattering has been widely studied in the actuality once its importance on practical modern engineering applications is getting more present in electronic and telecommunications equipments, in broadband transmissions, antennas design [1,2], biomedical [3] and electromagnetic coupling applications.

There are different techniques used to solve problems of electromagnetic scattering such as: FFT, FEM[4], FDTD[5,6], Meshless [7,8] and MoM[9,10] where, this last one is a known numerical technique to solve EFIE and MFIE associated with numerical analysis of the electromagnetic scattering by bodies of revolution. However, MoM solution requires a great computational time, for axial symmetric bodies (bodies of revolution –BOR) the computation

requirement is reduced once its analysis is conducted only in two dimensions. In electromagnetic scattering analysis by BOR, the MoM solution leads to complex integral equations that cannot be solved by analytical process and it presents very severe singularities.

In order to simplify the integral equations complexities and remove its singularities is necessary to develop techniques and good approximations that can eliminate this issue [11, 12]. Many techniques however with good accuracy can lead to long computational time and it is even more significant when a complex and large structure is under analysis.

In this work, EFIE scattering integral is analyzed under a conducting closed BOR. The Gaussian Quadrature is applied when the singularity is not present and first type elliptic integral is used to eliminate the singularity by a simpler and computational faster approximation.

2 INTEGRAL EQUATIONS

EFIE has being satisfactorily used in much electromagnetic scattering analysis, however, for solving the current under conducting closed BOR, it leads to resonance issued results. Harrington [13] satisfactorily demonstrated that it is possible to combine MFIE and EFIE in order to have a linear combination of both integral equations and eliminate the resonance problem. This method named Combined Field Equation (CFIE) is applied only to conducting BOR. For dielectric bodies, different linear combination between MFIE and EFIE can be done such as Mueller and PMCWHT methods [14, 15].

This work analyses electromagnetic scattering on closed conducting BOR (spheres) using CFIE equations to eliminate resonance problems. The generatrix is divided using segments in order to apply MoM. The superficial currents were modeled using triangular basis functions (TBF's) defined over two consecutive segments. The current under the surface is represented in \hat{t} (along the BOR generatrix) and $\hat{\phi}$ (azimuthally). Using these basis triangular functions with CFIE integral equations the following equation is defined:

$$I = \int_{\phi=0}^{2\pi} \int_{\alpha=-1}^1 \int_{\alpha'=-1}^1 \left(\frac{ab}{c}\right) G(\varphi) \frac{e^{-jKR}}{R^d} d\alpha' d\alpha d\varphi, \quad (1)$$

where $R = |\mathbf{r} - \mathbf{r}'|$ is the distance between source and observation points. a may be equal to 1, α' or α'^2 , b may be equal to 1, α or α^2 , c may be equal to 1, ρ, ρ' or $\rho\rho'$ and the exponent $d=1$ or 3. The coordinates t , and t' are parameterized as functions of variables α or α' and $G(\varphi)$ can be defined by [16]:

$$G_1(t, t') = 2 \int_0^\pi \sin^2 \left(\frac{\varphi}{2}\right) \cos m\varphi G_H d\varphi, \quad (2)$$

$$G_2(t, t') = \int_0^\pi \cos m\varphi \cos m\varphi G_H d\varphi, \quad (3)$$

$$G_3(t, t') = \int_0^\pi \sin m\varphi \sin m\varphi G_H d\varphi, \quad (4)$$

$$G_4(t, t') = 2 \int_0^\pi \sin^2 \left(\frac{\varphi}{2}\right) \cos m\varphi G_E d\varphi, \quad (5)$$

$$G_5(t, t') = \int_0^\pi \cos m\varphi \cos m\varphi G_E d\varphi, \quad (6)$$

$$G_6(t, t') = \int_0^\pi \sin m\varphi \sin m\varphi G_E d\varphi, \quad (7)$$

$$G_7(t, t') = \int_0^\pi \cos m\varphi G_E d\varphi = G_4(t, t') + G_5(t, t'), \quad (8)$$

where Green's function are:

$$G_E = \frac{e^{-jKR}}{kR}, \tag{9}$$

$$G_H = \left[\frac{1+jk_1R}{(k_1R)^2} \right] G_E. \tag{10}$$

In this work, the singularity removal technique was only used on the integrals associated with EFIE (equations: 5,6,7,8). In (equations: 2,3,4) Gaussian Quadrature was used and the integrals were evaluated using a different segmentation between source and observer. When the source and observer segments are coincident the singularity problem on the EFIE associated Integrals are present only in $G_5(t, t')$ and this equation can be rewritten as:

$$G_5(t, t') = G_{51}(t, t') + G_{52}(t, t'), \tag{11}$$

where:

$$G_{51}(t, t') = \int_0^\pi \left(\cos m\varphi \cos m\varphi \frac{e^{-jKR}}{kR} - \frac{1}{R} \right) d\varphi, \tag{12}$$

$$G_{52}(t, t') = \int_0^\pi \frac{\varphi}{R} d\varphi. \tag{13}$$

The integral G_{51} has no singularity and can be solved using Gaussian Quadrature. The integral G_{52} remains with a singular part and it is treated using elliptic integrals:

$$R_1 = \sqrt{((\rho - \rho')^2 + (z - z')^2)}, \tag{14}$$

$$\beta_1 = \frac{2\sqrt{\rho\rho'}}{R_1}, \tag{15}$$

$$G_{52}(t, t') = \int_0^\pi \frac{d\varphi}{R_1 \sqrt{1 + \beta_1 \sin^2 \frac{\varphi}{2}}}, \tag{16}$$

using elliptic integral definition:

$$\kappa(u) = \int_0^{\frac{\pi}{2}} \frac{d\zeta}{\sqrt{1 - u^2 \sin^2 \zeta}}, \tag{17}$$

$G_{52}(t, t')$ can be rewritten as:

$$G_{52}(t, t') = \int_0^{\frac{\pi}{2}} \frac{d\zeta}{R_2 \sqrt{1 - \beta_2 \sin^2 \frac{\zeta}{2}}}, \tag{18}$$

where:

$$R_2 = \sqrt{((\rho + \rho')^2 + (z - z')^2)}, \tag{19}$$

$$\beta_2 = \frac{2\sqrt{\rho\rho'}}{R_2}, \tag{20}$$

The integral in φ was solved but, $\kappa(u)$ remains singular for $u=1$ ($\beta_2 = 1$), what occurs in $\rho = \rho'$ and $z = z'$ and is equivalent to $t = t'$. The remaining integral is singular and the behavior of $\kappa(u)$ for $t \rightarrow t'$ is defined by [17]:

$$\mathcal{F}(\beta_2, \pi/2) \xrightarrow{\lim(\rho \rightarrow \rho' \text{ and } z \rightarrow z')} \frac{\log(4) + \log(R_2) - \log(R_1)}{2\rho'} \tag{20}$$

The singularity remains in $\log(R_1)$ and it is treated numerically:

$$\int \log(R_1) = \left(\alpha + \frac{T_0}{T} \right) * \log((T\alpha + T_0)^2 + T_1) - 2\alpha + \frac{2\sqrt{T_1}}{T} \operatorname{atan} \left(\frac{T\alpha + T_0}{\sqrt{T_1}} \right), \tag{21}$$

where

$$T_0 = (\rho_m' - \rho)^2 \sin u + (z_m' - z) \cos u, \tag{22}$$

$$T_1 = \left(((\rho_m' - \rho)^2 + (z_m' - z)^2)^2 - ((\rho_m' - \rho)^2 \sin u + (z_m' - z) \cos u)^2 \right). \tag{23}$$

When TBF's are used, other integral that must be solved is:

$$\int \alpha \log(R1) = \frac{1}{2} * \log(T\alpha' + T_0) \left(\alpha'^2 - \frac{T_0}{T^2} + \frac{T_1}{T^2} \right) - \frac{1}{2} \alpha'^2 + \frac{\alpha' T_0}{T} - \frac{2}{T^2} \sqrt{T_1} * \operatorname{atan} \left(\left(\frac{T_0 \alpha' + T_1}{\sqrt{T_1}} \right) * T_0 \right). \tag{24}$$

However, when the segment intercepts the symmetry axes the $1/\rho$ in the integral (equation 20) leads to precision issues once ρ will tend to zero. In this case, the numerical solution will be solved by an arithmetic mean [18,19,20]. This approximation was formally presented by Laden and used by Legendre in 1788 and where, thru this calculus, there was possible to find a very precised and rapidly convergent method. It is defined as:

$$a_0 = a \tag{25}$$

$$b_0 = b, \tag{26}$$

$$a_n = \frac{a_{n-1} + b_{n-1}}{2}, \tag{27}$$

$$b_n = \sqrt{a_{n-1} b_{n-1}}, \tag{28}$$

$$c_n = \sqrt{a_n^2 - b_n^2}. \tag{29}$$

3 - COMPUTATIONAL TEST ANALYSIS

In order to validate the proposed solution, some testes were done integrals (equation 6) where the electromagnetic scattering is analyzed under a perfect conducting BOR (sphere) with radius of 1λ (where λ is the wave length on vacuum) and divided in 53 segments. The FIGURES (1-4) shows the logarithmic calculated errors for each G_5 of the proposed method compared with [16]. In cases where the singularity is present more Gaussian Quadrature segmentation are used. The integral G_5 is presented in four different equations (1) such as : $G_{5aa}(a = 1, b = 1)$, $G_{5ab}(a = 0, b = 1)$, $G_{5ba}(a = 1, b = 0)$, $G_{5bb}(a = 0, b = 0)$.

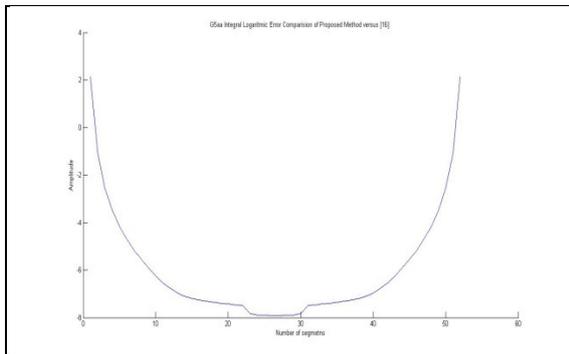


Figure 1: Logarithmic error for G_{5aa} of the proposed method compared with [19]

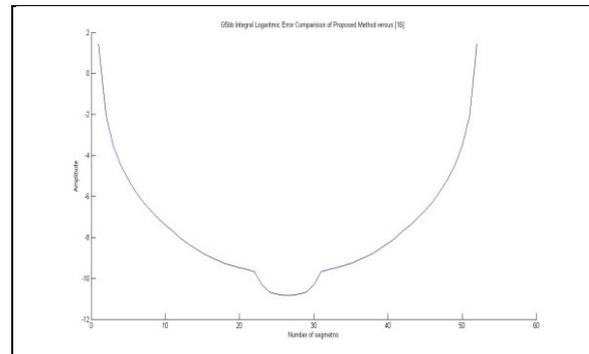


Figure 2: Logarithmic error for G_{5bb} of the proposed method compared with [19]

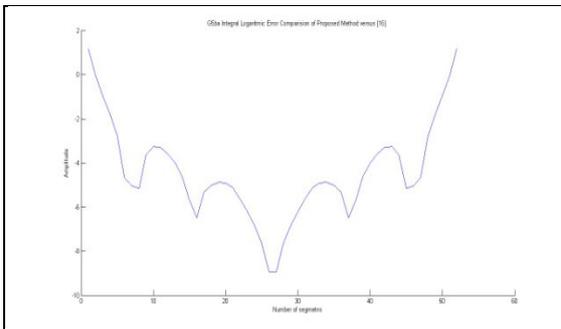


Figure 3: Logarithmic error for G_{5ba} of the proposed method compared with [19]

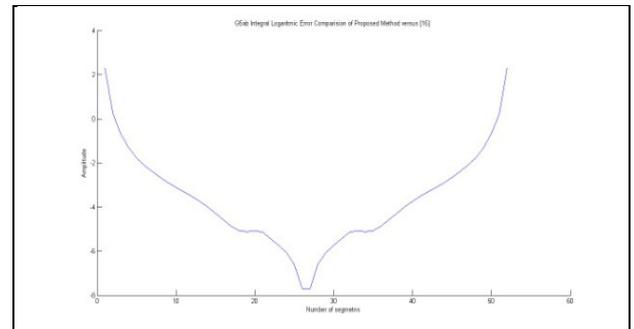


Figure 4: Logarithmic error for G_{5ab} of the proposed method compared with [19]

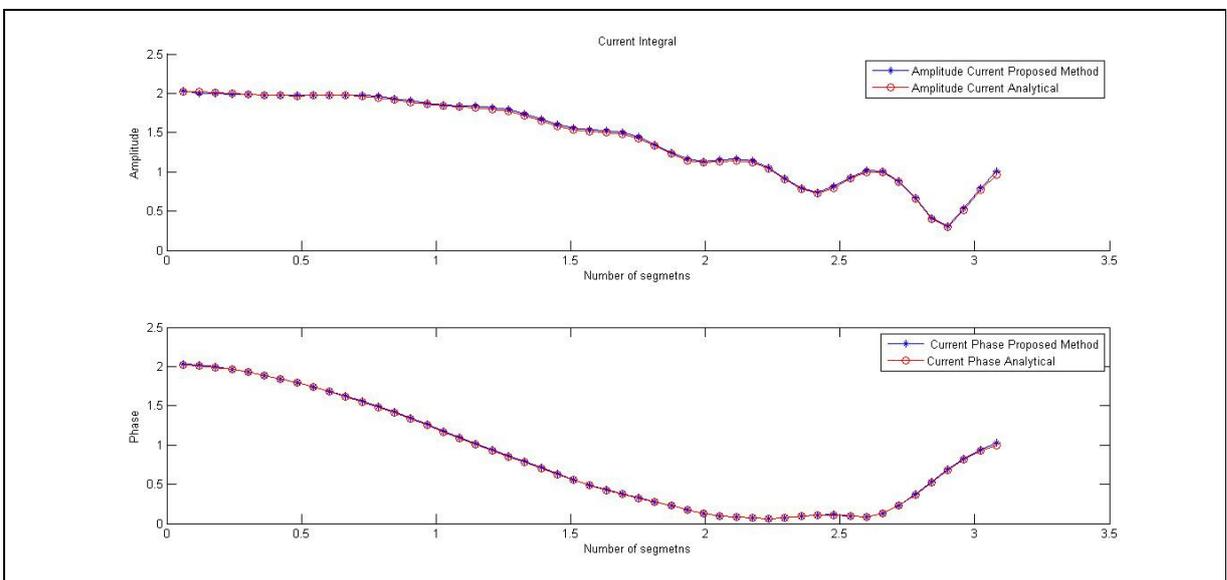


Figure 5: Amplitude and Phase of Current under Sphere BOR comparison of Analytic Solution versus Proposed Method

As can be observed in FIGURES (1-4) analysis, in the great majority of the segments the result agrees with high precision with [16]. The result diverges only in regions where the observation and source segment touches the symmetry axes, however, the results obtained with the current under the conducting BOR under analysis on FIGURE (5) shows that if compared with analytic solution obtained thru MIE series, the proposed technique represents a good approximation.

5 CONCLUSIONS

The studies of integral singularities extraction techniques are really important for the analysis of the electromagnetic scattering once its elimination can result in more accuracy results and simpler calculus.

The technique demonstrated to have a much simpler computational implementation if compared with other techniques and have considerably good approximations.

More can be studied regarding the use of this integral to composed bodies and more complex structures, in order to have faster results and as accuracy as other more elaborated techniques.

REFERENCES

- [1] DAVID T. AUUCKLAND, A. R. F. H. *Electromagnetic Transmission Through a Filled Slit in a Conducting Plane of Finite Thickness*, TE Case, v. MTT-26, July 1978. ISSN 7.
- [2] KORADA UMASHANKAR AND ALLEN TAFLOVE. *A Novel Method to Analyze Electromagnetic Scattering of Complex Objects*, v. EMC-24, NOVEMBER 1982. ISSN 4.
- [3] SADIKU, M. N. O., *Numerical Techniques in electromagnetics with Matlab*. Boca Raton, Florida: Taylor & Francis Group, LLC, 2009. ISBN 13:978-1-4200-6309-7.
- [4] JOHN L. VOLAKIS, A. C. A. L. C. K. *Finite Element Method for Electromagnetics*. New York: Willey Interscience, 1998. ISBN 0-7803-3425-6.
- [5] C. D. MOSS, F. L. T. Y. E. Y. A. J. A. K. *Finite-Difference Time-Domain Simulation of Scattering From Objects in Continuous Random Media*, v. 40, JANUARY 2002. ISSN 1.
- [6] F. L. TEIXEIRA, W. C. C. M. S. M. L. O. A. T. W. *Finite-difference time-domain simulation of ground penetrating radar on dispersive, inhomogeneous, and conductive soils*, v. vol. 36, Nov. 1998.
- [7] WILLIAMS L. NICOMEDES, R. C. M. A. F. J. S. M. *An integral meshless-based approach in electromagnetic scattering*. The International Journal for Computation and Mathematics in Electrical and Electronic Engineering (COMPEL), 2010.
- [8] M. S. TONG, W. C. C. *Meshless Evaluation of Domain Integrals for Solving Electromagnetic Integral Equations*. URSI International Symposium on Electromagnetic Theory: [s.n.]. 2010.
- [9] LI, J.-Y. L. A. L.-W. *Electromagnetic Scattering by a Mixture of Conducting and Dielectric Objects: Analysis using Method of Moments*. IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, v. 53, n. 2, MARCH 2004. ISSN 0018-9545/04.
- [10] HARRINGTON, R. F. *Matrix Methods for Field Problems*. PROCEEDINGS OF THE IEEE, 55, FEBRUARY, 1967.
- [11] ERGÜL, L. G. A. Ö. *Singularity of the Magnetic-Field Integral Equation and its Extraction*. IEEE ANTENNAS AND WIRELESS PROPAGATION LETTERS, 4, 2005.
- [12] SHUNG WU-LEE, J. B. C. L. L. A. G. A. D. *Singularity in Green's Function and Its Numerical Evaluation*. IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION, AP-28, n. 3, MAY 1980.
- [13] HARRINGTON, J. R. M. E. R. F. *H-FIELD, E-FIELD and combined field solutions*

- for bodies of revolution*. Syracuse University. Syracuse, p. 45. 1977. (Technical Report, TR-77-2).
- [14] MAXWELL, J. C. *A Treatise on Electricity and Magnetism*. New York: Oxford, 2002.
- [15] HARRINGTON, J. R. M. A. R. F. *Thesis of doctored: "Electromagnetic scattering from a Homogeneous Body of Revolution"*. Syracuse University. [S.l.]. 1977. (TR-77-10).
- [16] RESENDE, Ú. D. C. *Análise de antenas refletoras circularmente simétricas com a presença de corpos dielétricos*. Belo Horizonte, p. 266. 2007. Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais.
- [17] TEIXEIRA, F. L. **Aplicação do método dos momentos para a análise de refletores circularmente simétricos**. PUC RJ. Rio de Janeiro, p. 149 páginas. 1995.
- [18] SUPPES, C. G. C. A. P. *Approximations for the period of the simple pendulum based on the arithmetic-geometric mean*. American Association of Physics Teachers, 12, 2008.
- [19] HUAITANG CHEN, A. H. Z. *Extended Jacobin elliptic function method and its applications*. J. Appl. Math. & Computing, v. 10, p. 119-130, 2002. ISSN 1-2.
- [20] GOOD, R. H. *Elliptic integrals, the forgotten functions*. EUROPEAN JOURNAL OF PHYSICS, n. 22, p. 119-126, 2001.
- KENNAUGH, E. A. C. R. *The use of impulse response in electromagnetic scattering problems*. IRE International Convention Record. [S.l.]: [s.n.]. March, 1958. p. 72 - 77

SOME ASPECTS OF DYNAMIC COMPUTATIONAL MODELLING OF DIRECT CURRENT PLASMA ARC PHENOMENA

QUINN G. REYNOLDS^{*} AND B. DAYA REDDY[†]

^{*} Pyrometallurgy Division
Mintek
Private Bag X3015, Randburg 2125, South Africa
e-mail: quinnr@mintek.co.za, <http://www.mintek.co.za/Pyromet>

[†] Centre for Research in Computational and Applied Mechanics (CERECAM)
University of Cape Town
Private Bag X3, Rondebosch 7701, South Africa
email: daya.reddy@uct.ac.za, <http://www.cerecam.uct.ac.za>

Key words: Direct Current Furnace, Plasma Arcs, Dynamic Modelling

1 INTRODUCTION

Direct current arc furnaces see considerable use in modern industrial melting and smelting processes. Pyrometallurgical applications for this type of furnace are wide-ranging, and include commodities such as Ferrochrome, Ferronickel, Cobalt, Zinc, Magnesium, Titanium Dioxide, Platinum-group metals¹, and others.

Central to the operation of such furnaces is the direct current plasma arc, a sustained high temperature jet of ionised gas which is formed between the end of one or more graphite electrodes and the bath of molten process material below. Passage of electric current through the arc inputs energy and maintains the high temperatures necessary for ionisation via ohmic heating. This is balanced by various mechanisms of energy loss from the arc, including volumetric radiation and convection to the molten bath surface below. Much of this energy is delivered to a localised area directly beneath the arc, making it a very efficient means of heating the process material.

Flow of plasma in the arc column is driven strongly by electromagnetic Lorentz forces resulting from the constriction of the conduction channel in the vicinity of the electrode. This constriction causes the arc to draw in gas from the surroundings and accelerate it away from the electrode surface, toward the molten bath below (the Maecker effect²).

Much research has been conducted in the area of numerical modelling of arc phenomena, starting with Szekely and co-workers³ and becoming increasingly more sophisticated with the advent of better software, property data, and increased computing capability. However, the majority of arc modelling efforts concentrate on steady-state, axisymmetric systems. While valuable from an engineering standpoint these models are not able to describe any transient behaviour exhibited by the arc, or any evolution of the shape and structure of the arc which breaks the symmetry imposed by the model. Both of these aspects are important for a deeper understanding of direct current plasma arc behaviour.

2 MATHEMATICAL MODEL

The direct current plasma arc is a coupled multiphysics system, requiring distributed parameter models of fluid flow, energy transfer, and electromagnetics. All these fields must be approximated by mathematical expressions and solved simultaneously in order to achieve an overall model of the arc.

As the arc is a high-velocity gas jet, it is fundamentally a fluid flow problem. For simplicity and in order to focus on the dynamic aspects of the arc behaviour, incompressible flow with constant physical properties is assumed. In this case the Navier-Stokes and continuity equations governing fluid flow reduce to:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} + \nabla p = \nu \nabla^2 \mathbf{v} + \frac{\mathbf{j} \times \mathbf{B}}{\rho} \quad (1)$$

$$\nabla \cdot \mathbf{v} = 0 \quad (2)$$

In these equations, \mathbf{v} is the fluid velocity vector, p is the reduced pressure (P/ρ), ν is the dynamic viscosity, ρ is the density, \mathbf{j} is the current density vector, and \mathbf{B} is the magnetic field vector. The momentum source term is the Lorentz force, which arises due to the interaction of magnetic and electric fields in the plasma. This provides strong coupling between the velocity and electromagnetic fields in the model.

In order to make (1) more amenable to numerical treatment, the gauge method formulation of E and Liu⁴ is used. By defining auxiliary vector field \mathbf{a} and gauge variable θ such that:

$$\mathbf{a} = \mathbf{v} + \nabla \theta \quad (3)$$

$$p = \frac{\partial \theta}{\partial t} - \nu \nabla^2 \theta \quad (4)$$

and substituting into (1) and (2) produces:

$$\frac{\partial \mathbf{a}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = \nu \nabla^2 \mathbf{a} + \frac{\mathbf{j} \times \mathbf{B}}{\rho} \quad (5)$$

$$\nabla^2 \theta = \nabla \cdot \mathbf{a} \quad (6)$$

\mathbf{v} is recovered from the additional fields by a simple calculation using (3). A key strength of the gauge method over other primitive variable formulations is that boundary conditions for both \mathbf{a} and θ may be unambiguously specified by using the gauge freedom. For example, a non-slip wall condition ($\mathbf{v} = \mathbf{0}$ at boundary) may be written as:

$$\frac{\partial \theta}{\partial n} = 0, \mathbf{a} \cdot \mathbf{n} = 0, \mathbf{a} \cdot \boldsymbol{\tau} = \frac{\partial \theta}{\partial \boldsymbol{\tau}} \quad (7)$$

Energy transfer in the direct current plasma arc model is treated using several simplifying assumptions. Firstly, the plasma fluid is considered to be in local thermodynamic equilibrium (LTE)⁵, as a result of which a single temperature can be used to characterise the energy content of the material. LTE requires that the plasma be optically thin, dominated by reversible collision reaction processes, and exhibit small local gradients of plasma properties. The thermal plasmas found in direct current arc furnaces generally meet these conditions, although deviations can occur near to cold surfaces. In addition to LTE the physical properties of the plasma are assumed to be constant, as for the flow equations. With this, the energy conservation equation becomes:

$$\frac{\partial T}{\partial t} + \mathbf{v} \cdot \nabla T = \alpha \nabla^2 T + \frac{\mathbf{j} \cdot \mathbf{j}}{\sigma \rho C_P} - \frac{Q_R}{\rho C_P} \quad (8)$$

Here, T is the plasma temperature, α is its thermal diffusivity, C_P is its heat capacity, σ is its electrical conductivity, and Q_R is the volumetric rate of radiative energy loss. Both σ and Q_R are strong functions of temperature and plasma composition, as shown for an air plasma^{5,6} in Figures 1 and 2. Temperature is coupled to the velocity field via the convection term, and to the electromagnetic fields by the ohmic heating source term.

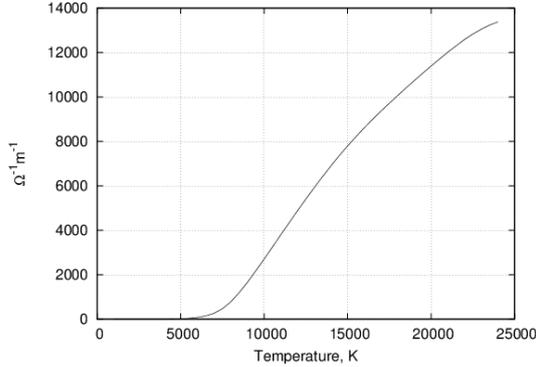


Figure 1: σ of air as a function of temperature

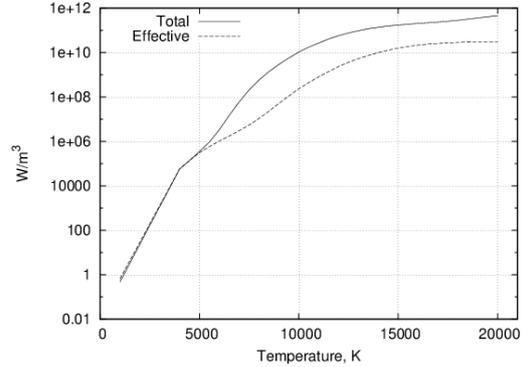


Figure 2: Q_R of air as a function of temperature

The electromagnetic field components in the plasma arc model are modelled using Maxwell's equations, simplified for the case of a neutral medium. Since the time scales involved in plasma arc motion are much longer than those relevant to the transient terms in Maxwell's laws (in which field propagation occurs at the speed of light), the electrostatic and magnetostatic versions of the laws are used. The peak velocities in the arc jet are also assumed to be low enough to ignore the induced current due to fluid motion. With these approximations, we have from charge continuity:

$$\mathbf{j} = -\sigma \nabla \phi \quad (9)$$

$$\nabla \cdot (\sigma \nabla \phi) = 0 \quad (10)$$

In these relationships, ϕ is the electric potential field. Boundary conditions for the electric field are a combination of Dirichlet (specified potential, for eg. molten bath surface at ground = 0) and Neumann (specified current, for eg. electrically insulating surfaces, or fixed current densities at arc root) types.

The magnetic field is calculated from the current density distribution, once it is known. Ampere and Gauss's laws govern the magnetic field behaviour:

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{j} \quad (11)$$

$$\nabla \cdot \mathbf{B} = 0 \quad (12)$$

These may then be reformulated using the magnetic vector potential, \mathbf{A} , such that:

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (13)$$

Substituting into (11) and applying the Coulomb gauge for \mathbf{A} then gives:

$$\nabla^2 \mathbf{A} = -\mu_0 \mathbf{j} \quad (14)$$

This expression is more amenable to numerical treatment, as it decouples the vector

components of the magnetic field into three independent Poisson equations. Boundary conditions are imposed on (14) using the assumption that the boundary surfaces are magnetically insulating, only permitting lines of magnetic flux to pass into or out of the calculation region on parts of the boundary through which current is flowing:

$$\mathbf{A} \cdot \boldsymbol{\tau} = 0, \frac{\partial(\mathbf{A} \cdot \mathbf{n})}{\partial n} = 0 \quad (15)$$

It is interesting to note that as a result of the electromagnetic relationships, the Lorentz force term in (1) can be shown to be irrotational in two-dimensional cartesian coordinates. This can pose problems⁷ for solutions of the plasma arc model in 2D. Fortunately these can be avoided by using a simple modification in the numerical treatment of (14), considering the calculation region for the magnetic field as though it has a small finite thickness in the third dimension. This results in the expression for (14) changing to the following, for 2D cartesian problems only:

$$\nabla^2 \mathbf{A} - \frac{2\mathbf{A}}{\delta L^2} = -\mu_0 \mathbf{j} \quad (16)$$

The size of the correction factor δL is specified in accordance with the dimensions of the arc column at its base, that is, where it attaches to the surface of the graphite electrode. This ensures that the magnitude and functional form of the magnetic field in the vicinity of the arc root, where the Lorentz force takes on its largest values, is accurate and physically realistic.

The calculation region and boundaries for the direct current plasma arc model are shown in Figures 3 and 4. A rectangular region immediately surrounding the arc in the central area of the furnace is modelled, consisting of the gas space between the tip of the graphite electrode and the top surface of the molten bath.

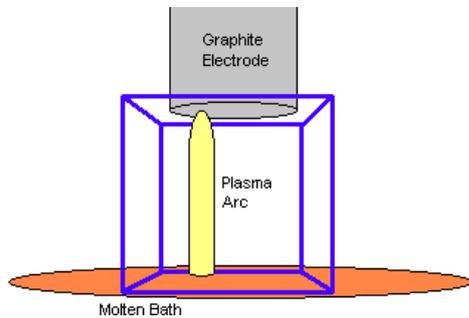


Figure 3: Solution region for arc models

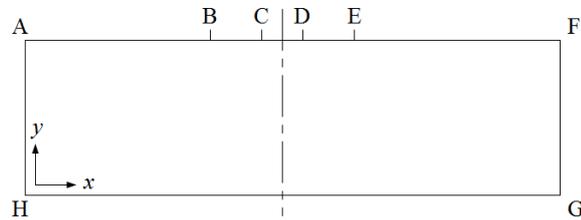


Figure 4: Region geometry (for 2D models)

In Figure 4, the origin of a cartesian coordinate system is located at H. The molten bath anode is located at surface GH, and the surface of the graphite electrode is section BE. The cathode spot, which forms the root of the arc column, is at CD. The remaining boundaries are treated as solid walls, forming a closed system.

The boundary conditions for the plasma arc model (in 2D) are shown in Table 1. The current density at the root of the arc on the surface of the graphite electrode, j_k , is governed by thermionic emission from the hot electrode, and has been determined experimentally⁸ to be of the order of 3.5×10^7 A/m². Together with a specified total current, this determines the dimensions of the cathode spot at CD.

Table 1: Boundary conditions for plasma arc model

	CD	BC & CE	AH & FG	AB & EF	GH
\mathbf{v}	$v_x = v_y = 0$	$v_x = v_y = 0$	$v_x = v_y = 0$	$v_x = v_y = 0$	$v_x = v_y = 0$
θ	$\frac{\partial \theta}{\partial y} = 0$	$\frac{\partial \theta}{\partial y} = 0$	$\frac{\partial \theta}{\partial x} = 0$	$\frac{\partial \theta}{\partial y} = 0$	$\frac{\partial \theta}{\partial y} = 0$
\mathbf{a}	$a_y = 0, a_x = \frac{\partial \theta}{\partial x}$	$a_y = 0, a_x = \frac{\partial \theta}{\partial x}$	$a_x = 0, a_y = \frac{\partial \theta}{\partial y}$	$a_y = 0, a_x = \frac{\partial \theta}{\partial x}$	$a_y = 0, a_x = \frac{\partial \theta}{\partial x}$
T	$T = T_E$	$T = T_E$	$T = T_W$	$T = T_W$	$T = T_A$
ϕ	$\frac{\partial \phi}{\partial y} = -\frac{j_k}{\sigma}$	$\frac{\partial \phi}{\partial y} = 0$	$\frac{\partial \phi}{\partial x} = 0$	$\frac{\partial \phi}{\partial y} = 0$	$\phi = 0$
\mathbf{A}	$A_x = \frac{\partial A_y}{\partial y} = 0$	$A_x = \frac{\partial A_y}{\partial y} = 0$	$A_y = \frac{\partial A_x}{\partial x} = 0$	$A_x = \frac{\partial A_y}{\partial y} = 0$	$A_x = \frac{\partial A_y}{\partial y} = 0$

The boundaries in the model are specified as having constant temperatures T_E , T_W , and T_A . These are the temperatures at the electrode, surrounding walls, and molten bath surface respectively. It is not physically unrealistic to have these surfaces at constant temperature, since in general the temperatures in the arc column are high enough to cause vapourisation of the materials they are made of; this process would be expected to hold the surfaces at the temperature of the phase change.

3 NUMERICAL METHOD

The numerical treatment of the direct current plasma arc model was approached with the aim of studying qualitative, time-dependent evolution of arc systems in mind. High spatial and temporal resolution as well as high performance were desirable, and led to the selection of finite difference methods together with explicit forward time-stepping techniques for the discretisation of the governing equations.

Structured cartesian grids with constant grid spacing were used in both 2D and 3D in an attempt to impose as little symmetry on the problem as possible, and to improve numerical performance. All spatial derivative terms are discretised using second-order centered differences, and a staggered grid is employed for certain fields such as \mathbf{a}/v and \mathbf{j} .

The time dependence of the governing equations is treated using a combination of methods. The flow equations are stepped forward in time using an explicit fourth order Runge-Kutta (RK4) algorithm. This imparts stability even at high Reynolds numbers⁹, and results in relatively low computational cost per time step. Assuming ∇_h and Δ_h are the discrete equivalents of the Nabla and Laplace operators respectively, this may be written as:

$$\frac{\mathbf{a}_1 - \mathbf{a}_t}{\frac{1}{2}\delta t} + (\mathbf{v}_t \cdot \nabla_h) \mathbf{v}_t = \nu \Delta_h \mathbf{a}_t + \frac{\mathbf{j}_t \times \mathbf{B}_t}{\rho}, \quad \Delta_h \theta = \nabla_h \cdot \mathbf{a}_1, \quad \mathbf{v}_1 = \mathbf{a}_1 - \nabla_h \theta \quad (17a)$$

$$\frac{\mathbf{a}_2 - \mathbf{a}_t}{\frac{1}{2}\delta t} + (\mathbf{v}_1 \cdot \nabla_h) \mathbf{v}_1 = \nu \Delta_h \mathbf{a}_1 + \frac{\mathbf{j}_1 \times \mathbf{B}_1}{\rho}, \quad \Delta_h \theta = \nabla_h \cdot \mathbf{a}_2, \quad \mathbf{v}_2 = \mathbf{a}_2 - \nabla_h \theta \quad (17b)$$

$$\frac{\mathbf{a}_3 - \mathbf{a}_t}{\delta x} + (\mathbf{v}_2 \cdot \nabla_h) \mathbf{v}_2 = \nu \Delta_h \mathbf{a}_2 + \frac{\mathbf{j}_2 \times \mathbf{B}_2}{\rho}, \quad \Delta_h \theta = \nabla_h \cdot \mathbf{a}_3, \quad \mathbf{v}_3 = \mathbf{a}_3 - \nabla_h \theta \quad (17c)$$

$$\frac{\mathbf{a}_4 - \mathbf{a}_t}{\delta x} + (\mathbf{v}_3 \cdot \nabla_h) \mathbf{v}_3 = \nu \Delta_h \mathbf{a}_3 + \frac{\mathbf{j}_3 \times \mathbf{B}_3}{\rho} \quad (17d)$$

$$\mathbf{a}_{t+\delta x} = \frac{1}{3} \mathbf{a}_1 + \frac{2}{3} \mathbf{a}_2 + \frac{1}{3} \mathbf{a}_3 + \frac{1}{6} \mathbf{a}_4 - \frac{1}{2} \mathbf{a}_t, \quad \Delta_h \theta = \nabla_h \cdot \mathbf{a}_{t+\delta x}, \quad \mathbf{v}_{t+\delta x} = \mathbf{a}_{t+\delta x} - \nabla_h \theta \quad (17e)$$

The major computational cost of this algorithm is four solutions of the Poisson equation, one for the θ variable at each fractional time step. Due to the regular, structured nature of the discretisation grid and the homogeneous boundary conditions for θ , fast spectral transform methods can be applied to achieve rapid direct Poisson solves.

Adaptive time stepping is applied as the flow field evolves. Assuming the size of the grid spacing in the model to be δx , this is calculated according to the constraints⁹:

$$\delta x < \frac{\delta x^2}{4\nu}, \quad \delta x < \frac{\delta x}{|\mathbf{v}|_{\max}} \quad (18)$$

Due to the very strong self-coupling that occurs in the energy conservation equation (8), it requires additional numerical treatment in order to perform in a stable manner. The convection/diffusion terms are separated from the source terms using first-order operator splitting; since the source terms contain no spatial derivatives, they may then be calculated in a semi-implicit way without generating significant computational overhead. The remaining terms are then treated using RK4 explicit time stepping, as for the momentum equations. Using an intermediate variable T^* , the source term fractional time step becomes:

$$\frac{T^* - T_t}{\delta x} = \frac{\mathbf{j}_t \cdot \mathbf{j}_t}{\sigma(T^*) \rho C_p} - \frac{Q_R(T^*)}{\rho C_p} \quad (19)$$

(19) may then be solved for T^* using the known relationships between temperature and the plasma conductivity, and volumetric radiation loss. This expression is not fully implicit, as the temperature can additionally affect the current density \mathbf{j} via the conductivity, but the computational cost of recalculating the electromagnetic fields iteratively does not justify the small additional stability gains to be had.

Once T^* is known, it is used as the starting point for the calculation of the convection/diffusion terms. This is illustrated in (20) using explicit forward-Euler time stepping, although in practice RK4 is used to retain the stability features described earlier:

$$\frac{T_{t+\delta x} - T^*}{\delta x} + \mathbf{v} \cdot \nabla_h T^* = \alpha \Delta_h T^* \quad (20)$$

As a result of the use of RK4 for the convection/diffusion components of the energy balance equation, an additional time step size constraint is introduced:

$$\delta x < \frac{\delta x^2}{4\alpha} \quad (21)$$

This is used together with (18) to calculate the adaptive time step size during execution of the algorithm.

Solution of the pseudo-steady state electromagnetic field equations is required at each time step in the direct current plasma arc model. These equations have no time dependence, and must therefore be solved implicitly. The electric field equation (10) is discretised as:

$$\nabla_h \cdot (\sigma \nabla_h \phi) = 0 \quad (22)$$

The electrical conductivity is calculated from the most recent temperature field. The values of σ are required on the staggered grid, and so where interpolation between the grids is necessary, linear or bilinear methods are used. (22) is a non-constant coefficient problem and cannot be efficiently tackled with fast spectral methods; it is instead solved using a standard geometric multigrid algorithm. As the ϕ field is recalculated at least once every time step, and the time steps in the model are generally very small, a good initial estimate of the field is always available (the values from the previous step). Iterative multigrid with a preset number of cycles and relaxation steps was therefore chosen over the full multigrid algorithm.

Once the electric potential field is known, the current density field j may be calculated directly from it using the discrete version of (9).

The discretised versions of Poisson equations (14) and (16) governing the magnetic vector potential field A have constant coefficients and homogeneous boundary conditions, and are solved using rapid fast-spectral-method based solvers once the current density is known. The magnetic field is then calculated by taking the discrete curl of A .

The computational cost of the electromagnetic field calculation is high, comparable to the entire time step calculation for the remaining variables. They are therefore only recalculated once per time step (as opposed to once per RK4 stage). Together with the operator splitting of the energy equation this has the effect of reducing the temporal accuracy of the overall algorithm to first-order, although numerical testing suggests that the stability remains largely unaffected.

An overview of the algorithm used in the numerical solution of the direct current plasma arc model is shown in Figure 4.

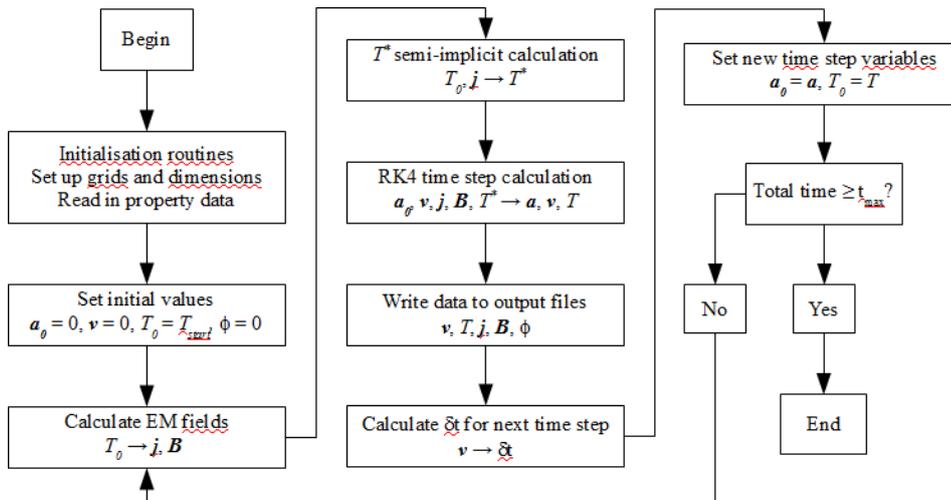


Figure 4: Direct current plasma arc model flowsheet

Implementation of the algorithm was performed using ANSI C. Outer loops of computationally intensive parts of the algorithm (particularly the fast spectral Poisson solvers and multigrid solver) were parallelised using OpenMP threading, providing increased performance on SMP machines. The Poisson solvers make extensive use of fast sine and cosine transforms provided by FFTW¹⁰. Extensive memory usage analysis and optimisation

was conducted to limit the memory footprint of the algorithm. Data sampling methods were implemented in order to reduce the quantity of file I/O during operation of the algorithm, which would otherwise be a rate-limiting step.

The implementation has been compiled and tested on a variety of computing platforms to date, including Intel (Pentium, Core2 and Xeon using gcc and icc compilers), IBM (P690 using xlc_r compilers), and Sun (SPARC M9000 using suncc compilers). Various flavours of Linux as well as Solaris have been successfully used as the host operating systems.

4 MODELLING RESULTS – 2D

Some results from an example case for the modelling of the direct current plasma arc system in two dimensions are presented below. The plasma gas used is air, for which physical property information is readily available. The parameters used for the model are shown in Table 2.

Table 2: Parameters used for 2D arc model

Region dimensions	0.2 m x 0.05 m	Electrode width	0.05 m
ν	0.005040 m ² /s	Current	500 A
ρ	0.02593 kg/m ³	T_W, T_A, T_E	2000K, 3000K, 4100K
α	0.01256 m ² /s	Grid dimensions	1024 x 256
C_P	9420 J/kg.K	Model time	10 ms

Initial conditions of zero velocity and constant temperature of 10000K are applied throughout the calculation region.

At early stages of the simulation, the arc model forms a strong jet of plasma material directed away from the arc attachment spot on the electrode surface. The jet is initially stable and symmetric, however, this behaviour is not robust. Oscillations near to the arc root begin from approximately 0.6 ms resulting in the formation of strong vortices, which travel down the length of the arc column and break up the structure. The oscillating vortex production causes the arc to take on a somewhat sinusoidal shape.



Figure 5: Temperature profile at 0.17 ms



Figure 6: Temperature profile at 0.50 ms



Figure 7: Temperature profile at 0.60 ms



Figure 8: Temperature profile at 0.70 ms

The process of arc jet formation and breakdown is shown in Figures 5 to 8. The

temperature profile is shown at various times, with the temperature scale ranging from 2000K (white) to 15000K (black).

Further into the simulation the arc jet becomes increasingly turbulent and chaotic in motion, forming a number of interesting structures as a result of the coupling between the flow field and the electromagnetic forces acting on it. One such phenomenon is the presence of persistent and highly-mobile arcs on the molten bath surface (which serves as anode in these models), shown in Figures 9 and 10. These are spontaneous emergent phenomena as the boundary conditions at this surface are entirely uniform, and have been confirmed experimentally using high speed photography¹¹.



Figure 9: Temperature profile at 7.80 ms, showing anode arcs and arc structure

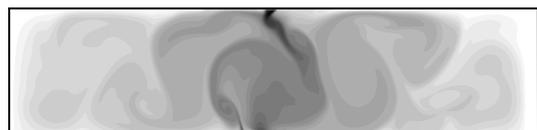


Figure 10: Temperature profile at 9.98 ms, showing anode arcs and arc structure

5 MODELLING RESULTS – 3D

The direct current plasma arc model is easily extended to three dimensions. This improves the spatial accuracy of the magnetic field and more accurately captures the qualitative geometry and behaviour of the arc, which is inherently a three-dimensional phenomenon. In the 3D models, the electrode surfaces and arc attachment spots are modelled as circular and may be located anywhere on the upper surface of the rectangular calculation region.

The 3D model is particularly well suited to the study of multiple arc systems. Some results from an example of this, a small twin-cathode furnace design with two electrodes and two arcs both carrying current in the same direction down to the molten bath anode, are shown below. Table 3 shows the parameters used for this model (where not given, parameters are identical to those given in Table 2).

Table 3: Parameters used for 3D arc model

Region dimensions	0.2m x 0.1m x 0.05m	Number of arcs	2
Grid dimensions	384 x 192 x 96	Current, arc 1	250 A
Arc separation	0.04 m	Current, arc 2	250 A

As in the 2D case, the twin-arc system initially forms stable arc jets directed away from the arc attachment spots. The onset of transient motion and decay of the pseudo-stable state occurs at approximately 1.5 ms. The nature of the subsequent motion is however considerably different, with the lower arc current causing the arc columns to settle into regular oscillations driven by precession of the arc jet around the attachment points on the electrode surfaces. This results in a dynamic helical structure forming within each arc column¹¹. The macroscopic behaviour of the system is also interesting to observe, as the arcs attract each other by virtue of the current they carry. This attractive force draws the arc jets toward the centerline of the region, causing significant deflection of the arc columns in the process.

The three-dimensional temperature field is sampled by taking the maximum value along the y axis, producing a projected 2D temperature field in the x-z plane. These are shown for the early stages of the motion in Figures 11 to 14. The temperature scale runs between 2000K (white) and 15000K (black).



Figure 11: Projected temperature field at 0.5 ms

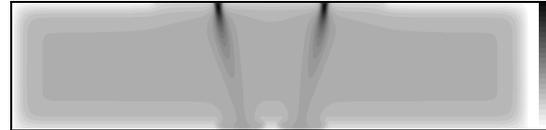


Figure 12: Projected temperature field at 1.0 ms

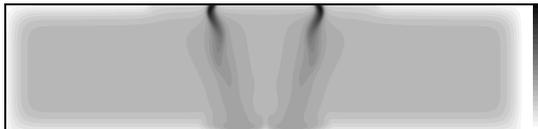


Figure 13: Projected temperature field at 1.5 ms



Figure 14: Projected temperature field at 2.0 ms

The full 3D fields for electric potential and temperature at the end of the simulation are shown in Figures 15 and 16. The scale ranges are 0 – 200 V and 2000 – 15000 K respectively.

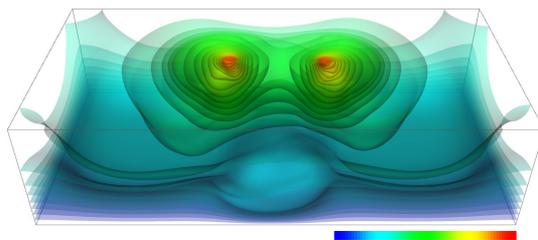


Figure 15: Electric potential field at 10 ms

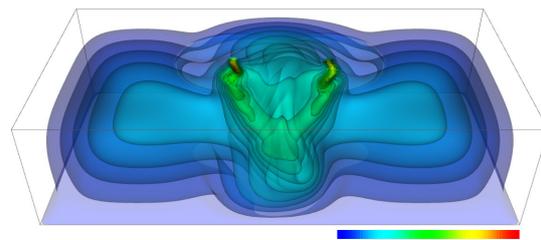


Figure 16: Temperature field at 10 ms

5.1 Transition effects with current

It is interesting to observe changes in the transient behaviour of the direct current plasma arc model as the various parameters are altered. One key parameter is the current carried by the arc(s), as this is an important operating variable for industrial direct-current arc furnaces.

The effect at low current was examined by changing the “Current, arc 1” variable in Table 3. Three different cases using values of 50 A, 100 A and 150 A were tested by completing simulation runs of the 3D model. The resulting projected temperature fields at the end of each simulation are shown in Figures 17 to 19, with scale range from 2000K (white) to 15000K (black). It can be seen that the low-current arcs, at left in each case, exhibit very different structure depending on the current. This is borne out in the temporal behaviour – Figure 20 shows the evolution of the arc voltage, which is defined as the local maximum value of the electric potential field in the vicinity of the arc attachment spot on the electrode surface, for the left-hand arc in each of the low current cases. There is a marked change in the transient arc behaviour between 100 and 150 A; at lower values, the arc voltage appears to remain approximately at steady state with only slow variations in time, while at higher currents, the

voltage begins to oscillate at a constant frequency. This reflects a change in the transient structure of the arc column, from a steady jet to a helical structure formed by precession of the jet around the arc attachment spot.



Figure 17: Projected temperature field, left arc at 50 A, right arc at 250 A



Figure 18: Projected temperature field, left arc at 100 A, right arc at 250 A

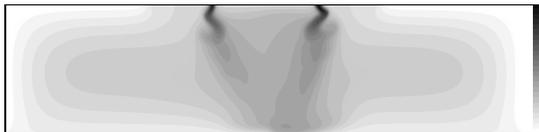


Figure 19: Projected temperature field, left arc at 150 A, right arc at 250 A

In order to compare the nature of the transient behaviour of arcs in the model at higher currents, two additional cases of the twin-arc 3D model were tested. In both cases the “Arc separation” variable in Table 3 was set at 0.1 m, placing the arcs well apart to reduce their interaction. The first case was run with the arc currents at 250 A each, and the second was run with the currents at 500 A each. The second case also used a higher-resolution numerical grid, at 768 x 384 x 192 in size. A sample of the temporal behaviour of the arc voltage for the left-hand arc in each case is shown in Figure 21. The increased current is seen to produce considerably more more irregular and chaotic behaviour.

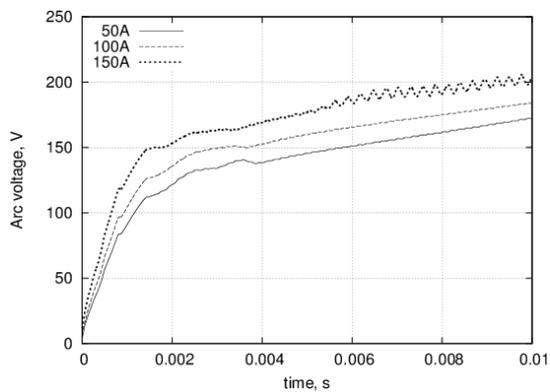


Figure 20: Evolution of arc voltage at low currents

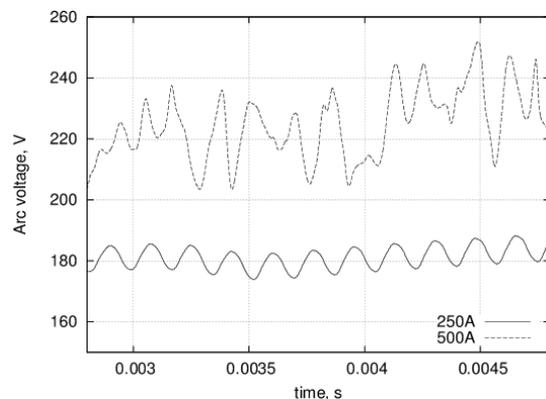


Figure 21: Evolution of arc voltage at higher currents

Multiple transition effects are therefore possible with variation of the current parameter in the direct current plasma arc model. Low currents produce steady arc jets, moderate currents produce regular oscillations in the system, and finally higher currents can produce erratic, complex motion. Similar transition effects have been observed experimentally using high speed photography¹¹, however much work remains to be done in this area.

6 CONCLUSIONS

The development of a dynamic model of direct current plasma arc systems of relevance to metallurgical furnace design has been largely successful. A novel solver algorithm using finite difference methods, explicit time-stepping, and rapid solver techniques has been produced, and is capable of scaling to large problem sizes on modest computing resources. The solver retains reasonable spatial and temporal accuracy while also possessing good stability properties.

Selected results from the model in both 2D and 3D show a variety of transient phenomena, including symmetry breaking in the early stages of arc formation, emergent behaviour such as spontaneous anode arc formation, and distinct transition effects with system parameters such as arc current.

Further avenues of research should include extension of the model to variable physical properties, induced current terms in the electromagnetic field equations, and compressible flow. Investigation of two-temperature plasma fluid models is also advised.

7 ACKNOWLEDGEMENTS

This paper is published by permission of Mintek. The assistance of the CSIR/Meraka Center for High Performance computing and the University of Cape Town is appreciated.

REFERENCES

- [1] Jones, R.T. Towards commercialisation of Mintek's ConRoast process for platinum smelting, *Proceedings of the 48th Annual Conference of Metallurgists of CIM* (2009) 159-168
- [2] Maecker, H. Plasmastromungen in Lichtbogen infolge Eigenmagnetische Kompression, *Zeitschrift fur Physik* (1955) **141**:198-216
- [3] Ushio, M., Szekely J. and Chang C.W. Mathematical modelling of flow field and heat transfer in high-current arc discharge, *Ironmaking and Steelmaking* (1981) **6**:279-286
- [4] E, W. and Liu, J.G. Gauge method for incompressible flows, *Communications in Mathematical Sciences* (2003) Vol. 1 **2**:317-332
- [5] Boulos, M.I., Fauchais, P. and Pfender, E. *Thermal plasmas: fundamentals and applications*. Plenum Press, Vol. 1 (1994)
- [6] Naghizadeh-Kashani, Y., Cressault, Y. and Gleizes, A. Net emission coefficient of air thermal plasmas, *Journal of Physics D: Applied Physics* (2002) **35**:2925-2934
- [7] Mutschke, G. and Bund, A. On the 3D character of the magnetohydrodynamic effect during metal electrodeposition in cuboid cells, *Electrochemistry Communications* (2008) **10**:597-601
- [8] Bowman, B. Properties of arcs in DC furnaces, *Proceedings of the 52nd Electric Furnace Conference* (1995) 111-120
- [9] E, W. and Liu, J.G. Vorticity boundary condition and related issues for finite difference schemes, *Journal of Computational Physics* (1996) **124**:368-382
- [10] <http://www.fft.w.org>
- [11] Reynolds, Q.G, Jones, R.T. and Reddy, B.D. Mathematical and computational modelling of the dynamic behaviour of direct current plasma arcs, *Proceedings of the 12th International Ferroalloys Congress* (2010) 803-814

THE INFLUENCE OF STEM SURFACE IN MICROMOBILITY AND CEMENT BONE STRESSES

A. RAMOS, C. RELVAS, A. COMPLETO AND J. A. SIMÕES

Biomechanics Research Group, Department of Mechanical Engineering,
University of Aveiro, 3810-193
Aveiro, Portugal
e-mail: a.ramos@ua.pt, web page: <http://www.mec.ua.pt>

Key words: hip arthroplasty, numerical models, micromobility, stress distribution, cement bone

Abstract. Cemented hip prostheses have produced excellent clinical results and THR is at that moment one of the most successful surgical techniques, with good success rates. Cemented fixation represents 87% of the total number of replacements according to the Swedish orthopedic register. Therefore, it is important to reduce revisions and understand why revision happens. The purpose of this study was to access the micro mobility and stresses developed in a cemented hip replacement. An in vitro cemented Lubinus SPII stem replacement was performed in synthetic femurs and sectioned. Section 5 was analyzed after fatigue test and was observed to be the most critical in crack incidence. A CAD model of this section was built considering bone and cement boundary geometry. The finite element model was built and the influences of different interface conditions of the cement interfaces (bone and stem) were analyzed. The interface stiffness associated with stresses for interface failure was used to simulate different surface roughness and time after surgery. The surface roughness associated with the interface strength did not present significant influence relatively to cement interface stresses and micro mobility of the stem. The type if interface changes the stress and strain distribution of bone and the most severe factor is friction at the cement/bone interfaces. The cement/bone interface debonding increase the bone strains and suggests pain.

1 INTRODUCTION

Cemented hip prostheses have produced excellent clinical results and THR is at that moment one of the most successful surgical techniques, with good success rates [1]. The success rate of all implants has increased and for 10 years is of the order of 93.5% (1991-96) [2]. The Swedish orthopedic register describes an increase number of revisions and present 13.6% of all revisions. For this reason it is important to reduce revisions and understand why these happen. An important reason for implant revision is aseptic loosening and biomechanical factors. The mechanism of aseptic loosening is not completely understood and is multifactorial according to some studies [3]. Mechanical factors are fundamental to trigger

the aseptic loosening process of a cemented implant. Other factors are type of cement and surface finish that have been correlated with long term failure scenarios. The fatigue failure of the cement mantle has been identified as a possible loosening mechanism. In other studies, failure has been associated with the thickness of the cement mantle [4, 5]. Other indicator associated to the long term failure is the micro motion and migration of the stem. Some experimental in vitro studies have correlated these parameters with long term results. The fatigue process results from normally activities with repeated loads [6, 7]. In some studies, micro motions between interfaces have been associated with the failure process with the formation of debris in the interface, promoting the creation of fibrous tissue [8, 9]. Aseptic loosening is one of the most important failure scenarios, being the cement damage accumulation an indicator of potential mechanical failure.

The purpose of this study was to access the micro mobility and stresses developed in Lubinus SPII stem hip replacement. A finite element model was built based on sections obtained image capture (section 5). Four stems of a Lubinus SPII were implanted in synthetic femurs and loaded and tested in fatigue, using a moment from stair climbing loading profile during one million cycles. The mechanical properties of bone were the same of those used by other studies, and different friction coefficients and interface stiffness were simulated. This study gives insight to identify the influence of the interface stiffness on the cement stress and bone strains.

2 MATERIAL AND METHODS

2.1 Experimental model

The experimental procedure involved the replacement of synthetic femurs with Lubinus SPII stems. Figure 1 presents the stem geometry of the Lubinus SPII system. The Lubinus SPII hip is one of the most used in Sweden, accounting for almost 57% of the total number of prostheses applied in 2007 [2]. It appeared in the market recently and at the moment presents a success rate of 98% for 10 years in 25.620 cases and 87.2% for 16 years [1].



Figure 1: Lubinus SPII cemented stem.

Annaratone et al. [10] refer that the Lubinus SPII is "one of the best" in the market of cemented stems. In vitro cemented hip replacements were performed using synthetic femurs (large left femur, mod. 3403, Pacific Research Labs, Vashon Island, WA, USA). These femurs have proven to be suitable for static and dynamic experimental simulations [11, 12]. The stems were implanted into synthetic femurs by an experienced orthopedic surgeon with adequate surgical instrumentation and CMW3® cement bone was used.

The loading applied allowed the combination of bending moments, a torsion moment and axial loads. The load was applied with the femur positioned at 11° on frontal plane and 9° on the sagittal plane, as specified by the ISO 7206 standard and used in other studies [6, 7].

The load applied was of sinusoidal type, which replicates loads induced by stair climbing and is considered the most severe [12]. The loading is approximately representative of 24 years of stair climbing, considering a hip reaction force of 2250 N for the intact femur. A maximum load and a minimum load resulted from the loading (table 1) configuration used after the adequate correction relatively to the intact femur head center was changed (intact and implanted femur) [13]. The femur was fixed at the distal condyles and one million load cycles at a frequency of 2Hz were applied through a pneumatic device.

Table 1: Characteristics of the Lubinus SPII stem.

	Lubinus SPII
Size	LII (left)
Material	CrCo alloy
Head size	28 mm
Surface roughness (Ra)	0.98 (± 0.02)
Maximum load (N)	1732 (± 88)
Surface roughness (Ra)	681 (± 94)

At the end of the fatigue testing procedure all specimens were transversely cutted in 11 sections. The first section was considered tangent to the collar. The cutting process was made using a high speed disc. The specimens were inspected with a non destructive technique using dye penetrating liquid [8].

2.2 Numerical model

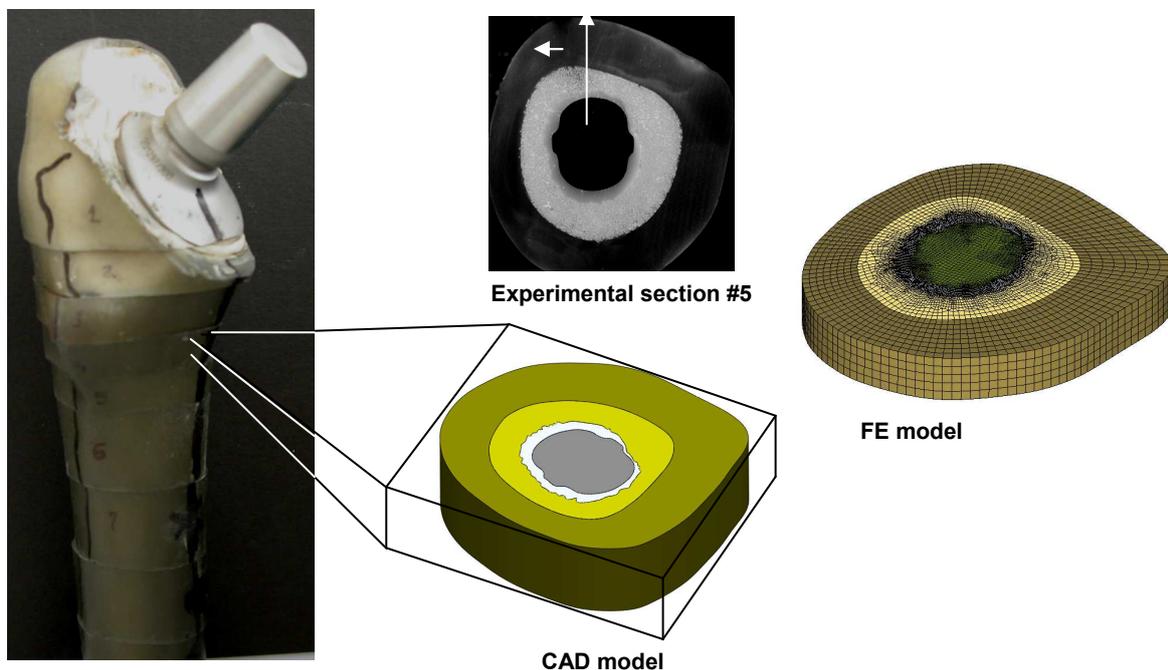
The finite element model was built based on the in vitro femur replacements. The CAD model was generated as illustrated in figure 3. The image was scanned with a 10 μm per pixel and boundary conditions defined between cortical bone, cancellous bone, cement and stem. The load applied was a moment with a 0.79 N.m as used in other studies and referred to be the most critical [12].

The thickness of the cement layer was non uniform, which is biomechanically relevant [14, 15]. The numerical model was constructed by hexagonal elements of 8 nodes.

Table 2: Material properties of the numerical model.

	Young Modulus (MPa)	Poisson Coefficient
Cortical bone	17400	0.3
Cancel bone	280	0.28
Cement	3000	0.3
CrCo	210000	0.3

The finite element mesh was performed using Hyperworks V10, (Altair) software and solver MSc Marc used. The material properties were assumed to be linear elastic and properties supplied by Sawbones® (table 2) were considered. The FEM model contained 35695 elements and 46825 nodes (figure 3). The surface roughness measured was $0.98 \mu\text{m}$.

**Figure 3:** Procedure to build the CAD and FE model

The surface roughness is important to define the interface stiffness between stem-cement bone-cement interfaces. Zelle et al [16] discuss the importance of the normal stress in the interface which depends on the surface roughness. For the stem studied the authors presented a maximum normal stress of 0.50 MPa. Jui-Pin and Fu-Chai [17] defined the interface strength as 0.49 to 9.95 MPa for surface roughness from 0.89 to 2.76 respectively.

The different models analyzed are defined in table 3. Different boundary conditions at each interface were simulated to consider the influence of time after surgery. The interface with glue represents a bonded interface without separation. Model #2 and #3 presents possible stem-cement interface separation. Model #4 simulates contact conditions of both interfaces in the last step of femur replacement. The contact control was defined as Coulomb friction with

a glue condition in the first step. . The two interface stiffness was simulated considering 0.5 MPa and 6.7 MPa until interface separation [18, 19].

Table 3: FEM boundary conditions.

	Interface stem/cement (μ / MPa)	Interface bone/cement (μ / MPa)
Model #1	glued	glued
Model #2	0.5/ 0.5	glued
Model #3	0.5/ 6.7	glued
Model #4	0.15/ 0	0.5/ 0

3 RESULTS

3.1 Experimental results

After the in vitro fatigue tests the femur replacements were sectioned and analyzed. Figure 4 presents the cracks distribution in the cement mantle. More cracks in the cement-bone interface of section #5 were observed. At medial aspect the stem cement interface presented more cracks in section #5. The most critical medial and posterior aspect was observed for section #5. For the lateral aspect, section #4 was the most critical in cement bone interface.

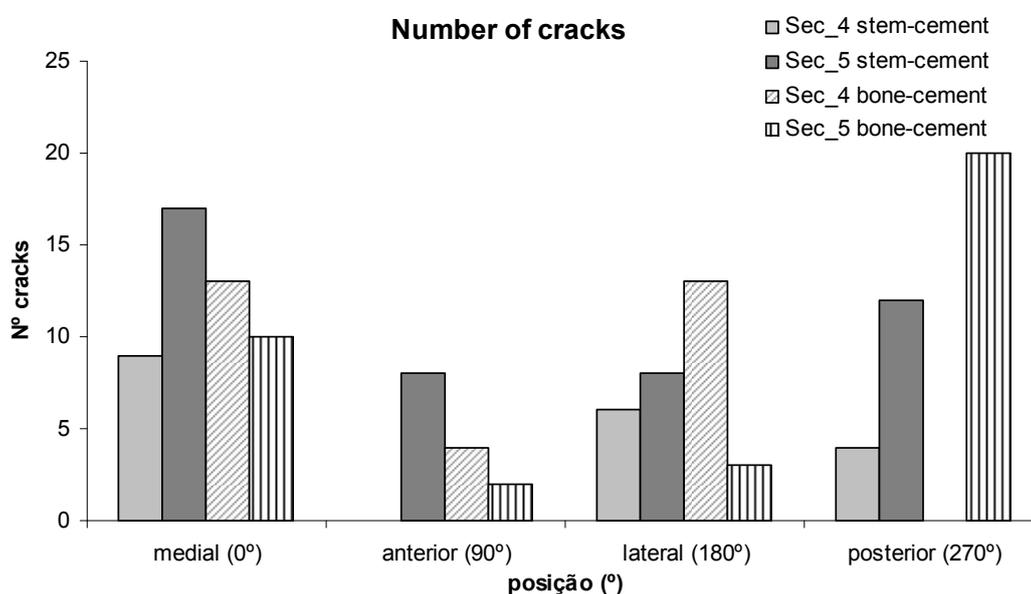


Figure 4: Number of cracks for section #4 and #5.

4.2 Numerical results

The displacement distributions in the three models are presented in figure 5 and 6. Model

#2 and #3 present similar results, showing small influence of the interface stiffness from 0.5 MPa (#2) to 6.7 MPa (#3). The results presents a similar behavior (relative displacement), only in the posterior aspect presents the same displacement.

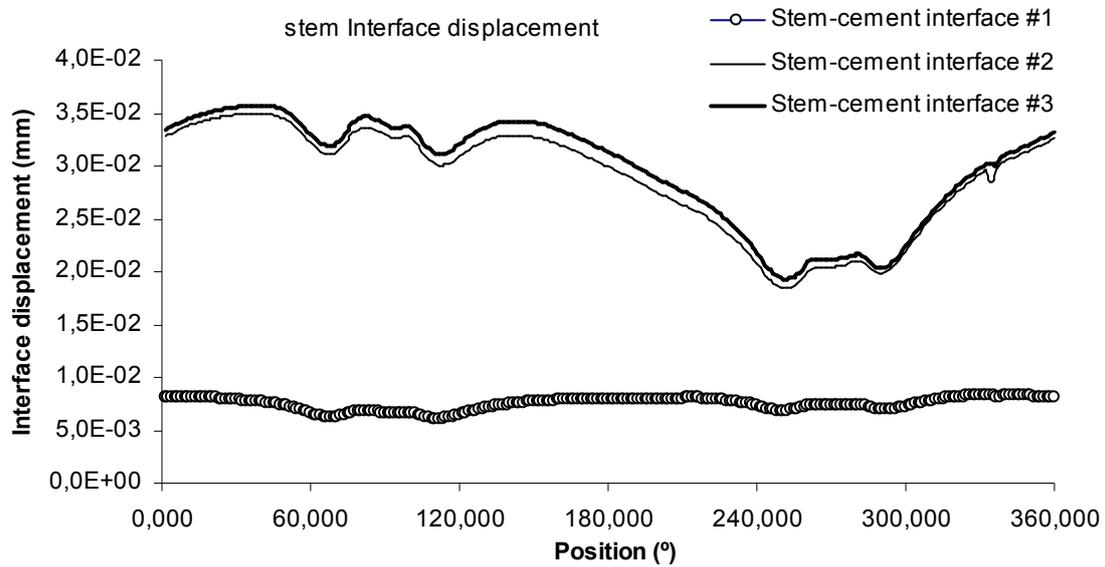


Figure 5: Interface displacements for different interface conditions.

The bonded interfaces (#1) presented the most uniform distribution. The interface stem-cement with contact presented the most concentrated distribution in the anterior e posterior aspects. These models show the variation of displacements with durability of time of replacement.

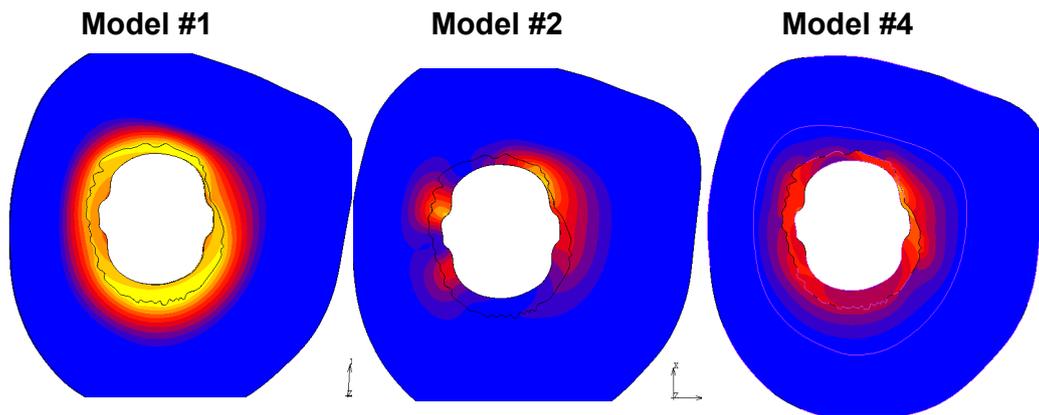


Figure 6: Interface displacements for different interface conditions.

The micro mobility in the stem-cement interface for models #2 and #3 are very similar (figure 7). This could be explained by the high moment applied and contact conditions do not influence the micro mobility. For the two interfaces simulating debonding (#4) we observed that the micro mobility in stem-cement interface decrease.

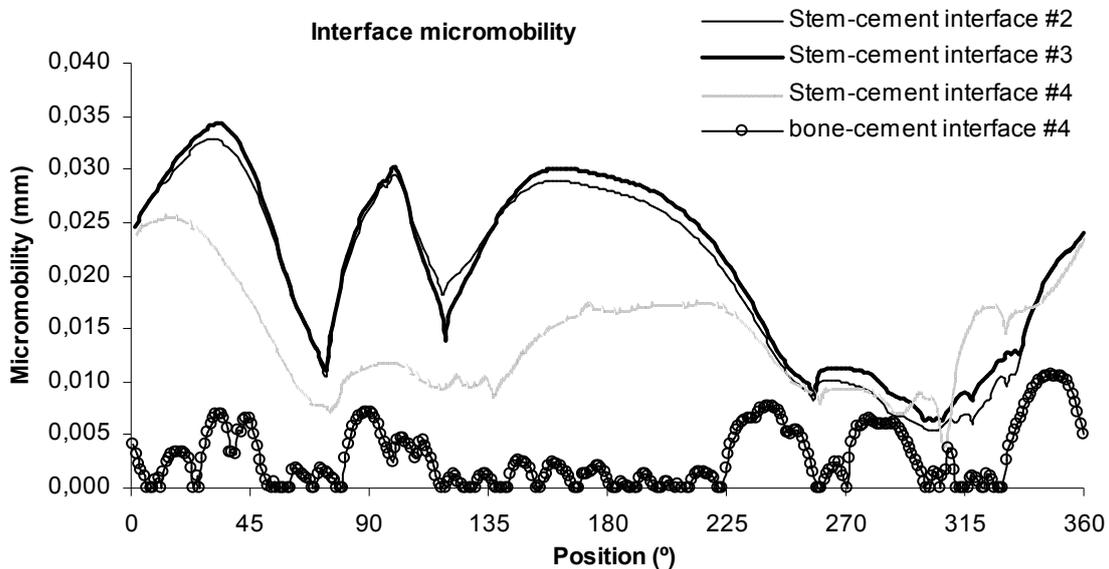


Figure 7: Maximum and minimum principal bone strains at the cement-bone interface.

The cement stresses were critical in the medial/posterior and lateral/anterior sides of the femur. The most critical conditions were observed for models #2 and #3 in some contact points. The bonded interface in model #1 presented the lowest principal stress. The maximum principal stress in the cement/stem interface was 39.6 MPa (#3), and in the interface cement/bone was 22.7 MPa (#3).

The maximum and minimum principal strains in the bone interface were most critical in model #4, with a maximum strain of 20000 μ strain. The maximum principal strain for model #1 with bonded interfaces was 4700 μ strain. Figure 8 presents the strain distribution for the three models. The anterior and posterior aspects presented the highest principal strain values. Not taking into account the posterior aspect, model #2 and #3 presented similar strain distributions. These values are in agreement of experimental crack results in posterior aspect at cement-bone interface.

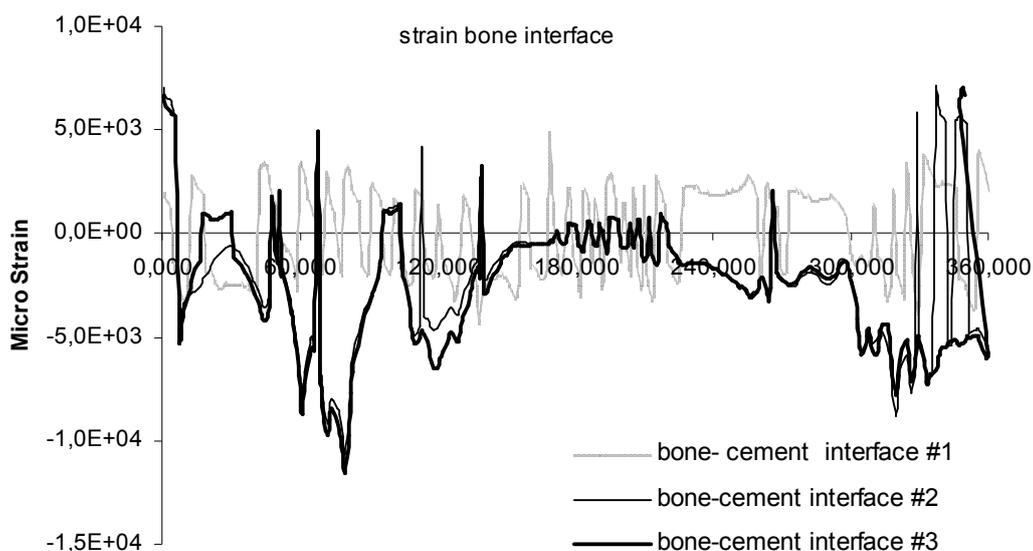


Figure 8: Maximum and minimum principal bone strains at the cement-bone interface.

CONCLUSIONS

- The loading configuration is an important issue to obtain representative results. The moment applied represents loads due to stair climbing and is the most severe;
- The interface stiffness does not present significant different results in terms of mobility, but changes the maximum stress in the cement interface up to 30%;
- The analysis of the sections of an implanted femur can be correlated with experimental results;
- The medial and lateral aspects present the most critical bone stresses;
- The micro mobility in the stem cement interface is critical in the first pos implanted period, because some debonding in the cement-bone interface occur the micro mobility decrease in stem interface. But debonding in bone interface increase the bone strain distribution and promote the pain.

REFERENCES

- [1] Kärrholm, J., Garellick, G., Rogmark, C., Herberts, P., 2007, Swedish Hip Arthroplasty Register Annual Report 2007.
- [2] Malchau H, Herberts P, Eisler T, 2008, The Swedish total hip replacement register. J Bone Joint Surg; 84A (Suppl 2):2—20.
- [3] Gravius, S., Wirtz, D. C., Siebert, C. H., Andereya, St., Mueller-Rath, Maus, U., Mumme, T., 2008, In vitro interface and cement mantle analysis of different femur stem designs, Journal of Biomechanics (2008), 41 (9), 2021-2028.

- [4] Fisher, D., A., Tsang, A. C., Paydar, N., Milions, S. and Turner, C., H., Cement-mantle thickness affects cement strains in total hip replacement. *Journal of Biomechanics* 30 (1997), 1173-1177.
- [5] Cristofolini, L., P. Erani, P., Savigni, P., Grupp, T., Otto, T., Viceconti, M. 2007, "Increased long-term failure risk associated with excessively thin cement mantle in cemented hip arthroplasty: A comparative in vitro study." *Clinical Biomechanics*, 22(4): 410-421.
- [6] A. Ramos, J.A. Simoes, The influence of cement mantle thickness and stem geometry on fatigue damage in two different cemented hip femoral prostheses, *Journal of Biomechanics*, Volume 42, Issue 15, 13 November 2009.
- [7] A. Ramos, J.A. Simoes, In vitro fatigue crack analysis of the Lubinus SPII cemented hip stem, *Engineering Failure Analysis*, Volume 16, Issue 4.
- [8] Anjuli R Cherukuri, Mark A. Miller, Amos Race, Timothy H. Izant, Kenneth A Mann, A wax barrier to simulate bone resorption for pre-clinical laboratory models of cemented total hip replacements, *Journal of Biomechanics*, Volume 43, 14, 2010, Pp 2855-2857
- [9] Waide, W., Cristofolini, L., Stolk, J., Verdonschot, N., Boogaard, G. J., Toni, A. , 2004, Modeling the fibrous tissue layer in cemented hip replacements: experimental and finite element methods. *Journal of Biomechanics* 37, pp: 13 - 26.
- [10] Annaratone, G., Surace, F. M., Salerno, P., Regis, G. F., 2000, Survival analysis op the cemented SPII stem, *Journal Orthopaedic Traumatol*, 1, 41 – 45.
- [11] Cristofolini, L, Viceconti, M., Cappello, A., Toni, A., 1996, Mechanical Validation of Whole bone composite femur models. *Journal of Biomechanics* 29(4), 525-535.
- [12] A. Race, M.A. Miller, K.A. Mann, Novel methods to study functional loading micromechanics at the stem-cement and cement-bone interface in cemented femoral hip replacements, *Journal of Biomechanics*, Volume 43, Issue 4, 3 March 2010, Pp. 788-791.
- [13] Ramos, A., Fonseca, F. e Simões, J. A., 2006, Simulation of Physiological Loading in Total Hip replacements”, *J. Biomechanical Engineering* 138: 579 – 587.
- [14] Jasty, M., Maloney, W. J., Bragdon, C. R., O’Connor, D. O., Haire, T. and Harris, W. H., 1991, The initiation of failure in cemented femoral components of hip arthroplasties. *Journal of Bone and Joint Surgery*, 73B, 551-558.
- [15] Race, A., Miller, M. A., David, C. A., Kenneth, A. M., 2003, Early cement damage around a femoral stem is concentrated at the cement/bone interface, *J. of Biomech* 36: 189-496.
- [16] J. Zelle, D. Janssen, S. Peeters, C. Brouwer, N. Verdonschot, Mixed-mode failure strength of implant-cement interface specimens with varying surface roughness, *Journal of Biomechanics*, Volume 44, Issue 4, 24 February 2011, Pages 780-783.
- [17] Jui-Pin Hung, Fu-Chai Chang, Computational modeling of debonding behavior at the bone/cement interface with experimental validation, *Materials Science and Engineering: C*, Volume 30, Issue 3, 6 April 2010, Pages 445-453
- [18] Daan Waanders, Dennis Janssen, Kenneth A. Mann, Nico Verdonschot, The behavior of the micro-mechanical cement-bone interface affects the cement failure in total hip replacement, *Journal of Biomechanics*, Volume 44, Issue 2, Bone Tissue: Hierarchical Simulations for Clinical Applications, 11 January 2011, Pages 228-234,
- [19] Daan Waanders, Dennis Janssen, Kenneth A. Mann, Nico Verdonschot, The effect of cement creep and cement fatigue damage on the micromechanics of the cement-bone

interface, Journal of Biomechanics, Volume 43, Issue 15, 16 November 2010, Pages 3028-3034

COEFFICIENT IDENTIFICATION FOR SHIP MANOEUVRING SIMULATION MODEL BASED ON OPTIMIZATION TECHNIQUES

K.T. TRAN*, A. OUAHSINE*, H. NACEUR**, F. HISSEL[†] AND A.
POURPLANCHE[†]

*University of Technology of Compigne (UTC)
Roberval Laboratory UMR-CNRS 6253
BP 20529, 60206, Compigne, France
e-mail: kxanh-toan.tran@utc.fr, ouahsine@utc.fr/

**University of Valenciennes and Hainaut-Cambresis (UVHC)
Laboratory of LAMIH, UMR-CNRS 8530
Le Mont Houy, 59313 Valenciennes, France
e-mail: hakim.naceur@univ-valenciennes.fr

[†]Research Center of Maritime and Fluvial Technique (CETMEF)
No 2, Boulevard Gambetta, 60321 Compigne, France
e-mail: Francois.Hissel@developpement-durable.gouv.fr,
Alain.Pourplanche@developpement-durable.gouv.fr

Key words: Ship Manoeuvring Simulation, Ship Hydrodynamic Coefficients, Optimization Techniques

Abstract. The frequency of ship grounding and collisions led to model the various factors involved during the course of a maneuvering ship. Among these factors are the ship hydrodynamic forces, the ship propulsion forces and forces due to the environmental conditions and the effects of confined water. This paper presents an approach for the identification of coefficients for a free-running ship. We elaborated a ship manoeuvring simulation model with a numerical procedure based on the coupling of optimization techniques and a ship motion simulation model. To identify the hydrodynamic coefficients, an automatic approach is proposed with two main steps: firstly, a sensitivity analysis to identify the most sensitive coefficients; secondly, optimization techniques to calculate their optimal value. Our model has been validated by using experimental data of Esso Bernicia Tanker (190000dwt) for the Turning Circle Test.

1 INTRODUCTION

Many equations of ship motion have been studied to describe the external forces acting on a ship. These external forces or hydrodynamic forces are function of many hydrodynamic coefficients. Abkowitz [8] proposed a hydrodynamic force model based on the multiple polynomial equations of ship manoeuvring variables from Taylor expansion. In this model, the derivatives of hydrodynamic forces with respect to each manoeuvring variables are named the hydrodynamic coefficients. Since, many mathematical models of ship motion have proposed to identify the hydrodynamic coefficients [1], [13], [14], [5],[4], [6]. Several methods identify the hydrodynamic coefficients, such as the captive model test, strip theory, empirical formulae, computational fluid dynamics (CFD), system identification (SI), optimization technique, etc. Among them the system identification and optimization technique are practical and widely used to avoid the scaling effect between the real ship and the scale model [10].

In respect of the system identification (SI), Hwang [12] applied the state augment of Extended Kalman Filtering (EKF) technique to identify the dynamic system of a manoeuvring ship. The slender body theory is utilized to explain the intrinsic nature of cancelation effect for dynamics of ship motion.

Recently, H.K. Yoon and K.P. Rhee [10] applied the Estimation-Before-Modeling (EBM) technique for hydrodynamic coefficients identification. EBM is the two-step method, based on the Extended Kalman Filtering (EKF) technique and Modified-Bryson-Frazier (MBF) smoother to estimate motion variables, hydrodynamic forces, speed and direction of current.

Regarding to the optimization technique, M. Viviani et al. [11] applied the numerical optimization techniques to hydrodynamic coefficients identification from standard manoeuvres (specified by IMO) for a series of twin-screw ships. From regression formulae based on existing ship model-test data, they developed hydrodynamic coefficients identification for non-conventional ships exceeding the parametric range of the experimental data base. By means of a sensitivity analysis, they identified the 5 most sensitive coefficients that their influence is much stronger than remaining coefficients. The objective function represents discrepancies between simulated and experimental manoeuvres, evaluated a sum of relative errors of a series of macroscopic manoeuvring parameters in Turning Circle test and ZigZag test. Optimization procedure is carried out by applying a Multi Objective Genetic Algorithm (MOGA).

To reach a more accuracy of ship simulation trajectory based on the optimization techniques, we have focused analyzing the sensibility of hydrodynamics coefficients in the alternative manoeuvring tests. The objective is to find more the sensible hydrodynamic coefficients, that means the number of identified coefficient is increased. In this study, we carried out an optimization procedure for 10 most sensible coefficients.

2 SHIP MANOEUVERING SIMULATION MODEL

2.1 Ship motions and manoeuvring tests

In the present study, the ship's motions were simulated in 3DOF, are presented in Fig. 1, including surge (along OX), sway (along O_0Y_0) and yaw (along O_0Z_0). The moving coordinate frame $GX_0Y_0Z_0$ is fixed to the ship's gravity center (G) is called the body-fixed reference frame or ship-fixed frame. The coordinate frame $OXYZ$ is called the earth-fixed reference frame [7].

where: $U = ui + vj$: ship velocity; i, j : unit vectors along the GX_0 and GY_0 ; ψ : heading angle; $r = \frac{d\psi}{dt}$: yaw angular velocity

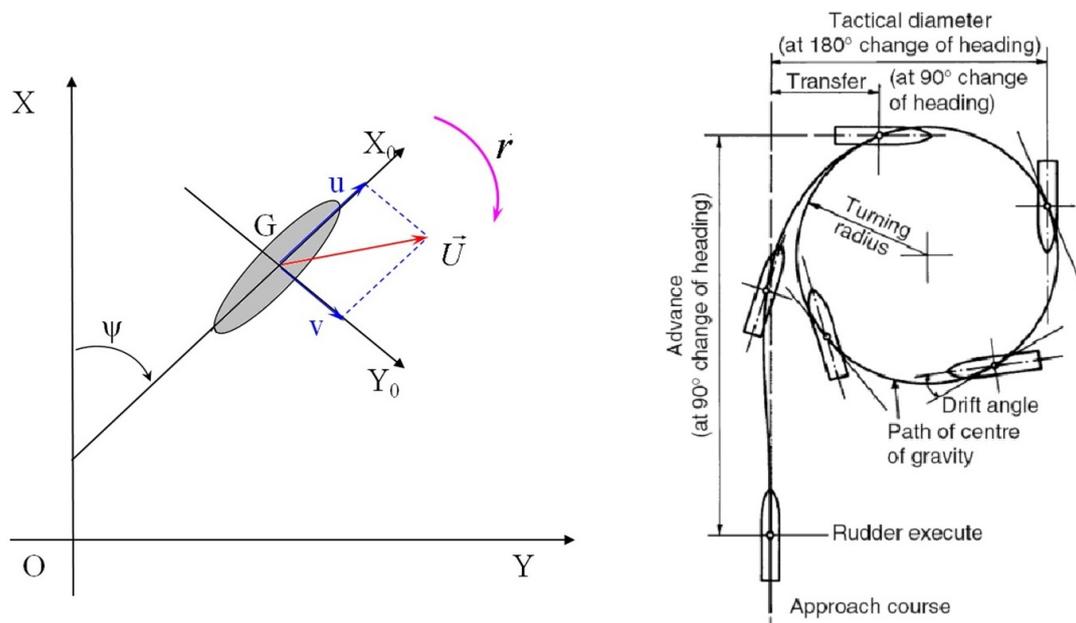


Figure 1: Ship motions in 3DOF and Turning Circle test

The main manoeuvring tests of ship are recommended by the Manoeuvring Trial Code of ITTC and the IMO [9], including: Turning circle test, Spiral manoeuvres, Pull-out manoeuvre, Zigzag manoeuvre, Stopping trial, Hard rudder test and Man-overboard manoeuvre.

In the present study, due to limited number of experimental data, only the Turning Circle test was applied to validate for the Esso Bernicia Tanker (Esso 190000dwt) model.

Turning circle test[9] Starting from straight motion at constant speed, the rudder is turned at maximum speed to an angle δ and kept at this angle, until the ship has performed a turning circle of at least 540° . The essential information obtained from this

Parameter	Value	Unit
Length between perpendicular (L_{pp})	304.8	m
Beam (B)	47.17	m
Draft to design waterline (T)	18.46	m
Displacement (∇)	220000	m^3
L_{pp}/B	6.46	-
B/T	2.56	-
Block coefficient (C_B)	0.83	-
Design speed (U_0)	16	knots
Nominal propeller (n)	80	rpm

Table 1: Parameters of Esso Bernicia model

manoeuvre consists of (Fig. 1): Tactical diameter, Maximum advance, Transfer at 90^0 change of heading, Times to change heading 90^0 and 180 and Transfer loss of steady speed.

2.2 The Esso Bernicia Tanker (Esso 190000dwt) model

Parameters of the Esso Bernicia Tanker model Mathematical models describing the maneuverability of large tanker in deep and confined waters are found by Van Berlekom and Goddard (1972). One of these models is the Esso Bernicia Tanker (Esso 19000 dwt) [7], with the ship parameters is presented in Tab. 1.

3 DOF motion equations of the Esso Bernicia Tanker model The 3DOF equations of ship motion in Bis-System are given by Van Berlekom and Goddard[7]:

$$\begin{aligned} \dot{u} - vr &= g.X'' \\ \dot{v} + ur &= g.Y'' \\ (L.k_z'')^2 \dot{r} &= g.L.N'' \end{aligned} \quad (1)$$

with:

$$\begin{aligned} g.X'' &= X''_{\dot{u}} \cdot \dot{u} + \frac{1}{L} \cdot X''_{|u|} \cdot u |u| + \frac{1}{L} \cdot X''_{vr} \cdot vr + \frac{1}{L} \cdot X''_{vv} \cdot v |v| + \frac{1}{L} \cdot X''_{c|c|\delta\delta} \cdot c |c| \delta^2 + \quad (2) \\ &+ \frac{1}{L} \cdot X''_{c|c|\beta\delta} \cdot c |c| \beta \delta + g.T (1 - t) + \frac{1}{L} \cdot X''_{\dot{u}\xi} \cdot \dot{u} \xi + \frac{1}{L} \cdot X''_{|u|\xi} \cdot |u| \xi + \frac{1}{L} \cdot X''_{vr\xi} \cdot vr \xi + \frac{1}{L} \cdot X''_{vv\xi\xi} \cdot v^2 \xi^2 \\ g.Y'' &= Y''_{\dot{v}} \cdot \dot{v} + \frac{1}{L} \cdot Y''_{uv} \cdot uv + \frac{1}{L} \cdot Y''_{|v|v} \cdot |v| v + \frac{1}{L} \cdot Y''_{|c|c\delta} \cdot |c| c \delta + \frac{1}{L} \cdot Y''_{ur} \cdot ur + \frac{1}{L} \cdot Y''_{|c|c|\beta|\delta} \cdot |c| c |\beta| \beta \delta + \quad (3) \\ &+ Y''_T \cdot g.T + \frac{1}{L} \cdot Y''_{ur\xi} \cdot ur \xi + \frac{1}{L} \cdot Y''_{uv\xi} \cdot uv \xi + Y''_{\dot{v}\xi} \cdot \dot{v} \xi + \frac{1}{L} \cdot Y''_{|v|v\xi} \cdot |v| v \xi + \frac{1}{L} \cdot Y''_{|c|c|\beta|\delta\xi} \cdot |c| c |\beta| \beta \delta \xi \\ g.L.N'' &= \frac{1}{L^2} \cdot N''_{\dot{r}} \cdot \dot{r} + \frac{1}{L^2} \cdot N''_{uv} \cdot uv + \frac{1}{L} \cdot N''_{|v|r} \cdot |v| r + \frac{1}{L^2} \cdot N''_{|c|c\delta} \cdot |c| c \delta + \quad (4) \\ &+ \frac{1}{L} \cdot N''_{ur} \cdot ur + \frac{1}{L^2} \cdot N''_{|c|c|\beta|\delta} \cdot |c| c |\beta| \beta \delta + \frac{1}{L} \cdot N''_T \cdot g.T + \frac{1}{L} \cdot N''_{ur\xi} \cdot ur \xi + \\ &+ \frac{1}{L^2} \cdot N''_{uv\xi} \cdot uv \xi + N''_{\dot{r}\xi} \cdot \dot{r} \xi + \frac{1}{L} \cdot N''_{vr\xi} \cdot vr \xi + \frac{1}{L^2} \cdot N''_{|c|c|\beta|\delta|\xi} \cdot |c| c |\beta| \beta |\delta| \xi \end{aligned}$$

where $X^{\parallel}, Y^{\parallel}, N^{\parallel}$ are the non-dimensional forces and moments

$k_z^{\parallel} = \frac{1}{L} \sqrt{\frac{I_z}{m}}$ is the non-dimensional radius of gyration

$X_{\dot{u}}^{\parallel}, X_{|u|}^{\parallel}, \dots, Y_{\dot{v}}^{\parallel}, Y_{uv}^{\parallel}, \dots, N_{\dot{r}}^{\parallel}, N_{uv}^{\parallel}, \dots, N_{|c|\beta|\delta|\xi}^{\parallel}$ are the non-dimensional derivatives of ship hydrodynamic coefficients in Bis-System, which will be identified by optimization techniques.

3 OPTIMIZATION ALGORITHMS

In the present study, the multi-variable optimization problems [15] was applied for ship hydrodynamic coefficient identification, derives in 2 options:

- The constrained optimization problem with the SQP method: Minimize the objective function $Obj = f(\alpha)$ for $\alpha \in E^n$, subject to $\alpha_{min} \leq \alpha \leq \alpha_{max}$.

where

$\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$ is the vector of ship hydrodynamic coefficients or the vector of variables, n is the number of ship hydrodynamic coefficients,

$\alpha_{min}, \alpha_{max}$ are the minimum and maximum values of estimated hydrodynamic coefficients, are also the inequality constraints of optimization problem.

- The unconstrained optimization problem with the Simplex method and BFGS method: Minimize the objective function $Obj = f(\alpha)$ for $\alpha \in E^n$. where $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$ is the vector of ship hydrodynamic coefficients or the vector of variables, n is the number of ship hydrodynamic coefficients.

Objective of ship coefficient identification is to identify the optimal hydrodynamic coefficients so that the ship's computed trajectory (simulated trajectory) approximates the ship's experimental trajectory. The deviation between computed trajectory and experimental trajectory needs to be minimized. So the form of objective function Obj is below:

$$Obj = \sqrt{\sum_{i=1}^N \Delta S_i^2} \tag{5}$$

$$\Delta S_i^2 = (x_i^{cal} - x_i^{exp})^2 + (y_i^{cal} - y_i^{exp})^2 \tag{6}$$

where

N is the number of coupled points to be approximated,

(x_i^{cal}, y_i^{cal}) are the coordinates of the point number i on the ship's computed trajectory,

(x_i^{exp}, y_i^{exp}) are the coordinates of the point number i on the ship's experimental trajectory,

ΔS_i^2 is the square of distance discrepancy between the coupled point number i on the ship's computed trajectory and on the ship's experimental trajectory, is also the function of ship hydrodynamic coefficients $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$, is presented in Fig. 4.

Input data	Value	Unit
Initial ship's position (x_0, y_0)	(0,0)	m
Initial heading angle (ψ_0)	0	deg
Initial advance velocity of ship (U_0)	5.3	m/s
Initial of rudder angle (δ_0)	0	deg
Maximum rotation velocity of rudder $(\dot{\delta}_{max})$	2.7	deg/s
Initial shaft velocity (n_0)	80	rpm
Shaft velocity command (n_c)	80	rpm
Rudder command (to port side) (δ_c)	-35	deg

Table 2: Input data for Turning Circle test of Esso Bernicia model

4 NUMERICAL OPTIMIZATION PROCEDURE

The numerical procedure is presented in Fig. 2. In this article, only emphasizing two main steps:

- (i) Sensitivity analysis to identify the most sensitive coefficients.
- (ii) Optimization techniques to calculate the optimal value of these coefficients.

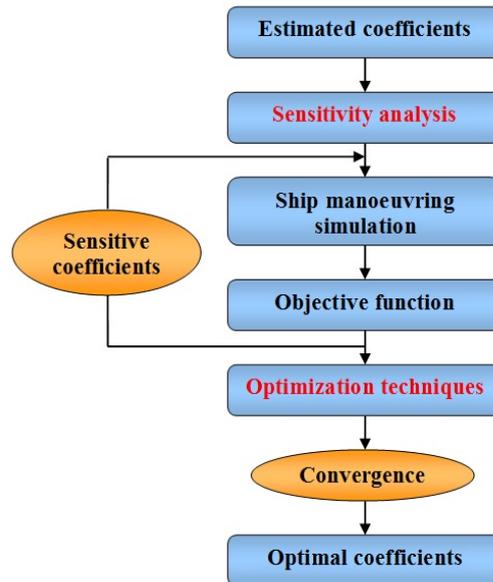


Figure 2: Flowchart of the numerical procedure

In the present study, the numerical model was validated for Esso Bernicia Tanker model in Turning Circle test with the input data presented in Tab. 2.

The numerical procedure starts from all estimated hydrodynamic coefficients in the ship motion equations (Eq. 3, 4 and 5), which will be analyzed by sensitivity analysis. The

analyzing method is to compare the gradient of objective function $Fobj$ while varying the values of each coefficient α_i , so to find coefficients which influence more strongly the gradient of objective function. The 10 most sensitive coefficients (stronger influence coefficients) were chosen among 35 coefficients in ship motion equation with the following condition:

$$\left| \frac{\partial Fobj}{\partial \alpha_i} \right| \geq 0.1 \tag{7}$$

Next, the 10 most sensitive coefficients (optimization variables) are applied the optimization techniques, consisting of SQP, BFGS and Simplex methods. The objective function is normalized by:

$$Fobj^j(j) = \frac{Fobj(j)}{Fobj(j=0)} \tag{8}$$

where j is the iteration number of optimization procedure, $j = 0$ is the first iteration, so that the value of $Fobj^j(j)$ will be reduced from 1 to an approximate value of 0.

The numerical procedure is repeated until convergence of objective function and variable is reached. Chosen objective function tolerance is 10^{-4} , and the optimization variable one is 10^{-4} .

5 NUMERICAL RESULTS

5.1 Ship trajectory simulation and sensitivity analysis before optimization

Computed and experimental trajectories of ship and sensitivity analysis of ship hydrodynamic coefficients in Turning Circle test (c.f. the experimental data of Esso Bernicia Tanker model [2]) are presented in Fig. 3.

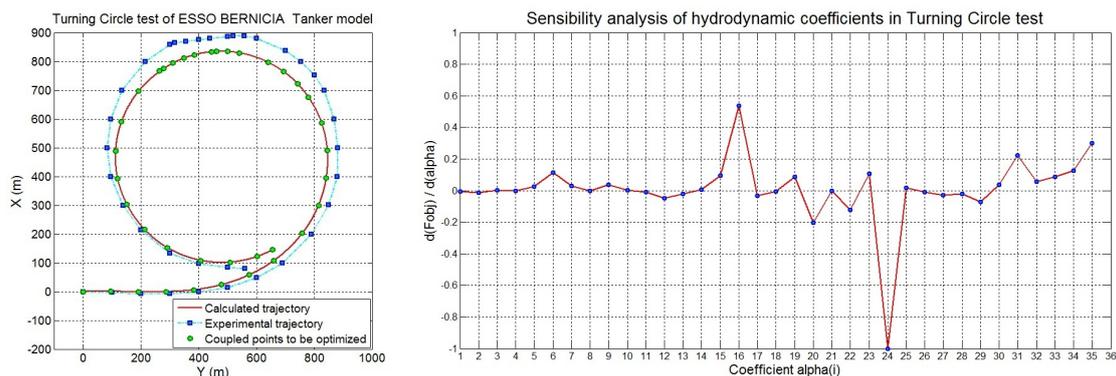


Figure 3: Ship trajectory and sensitivity analysis of ship hydrodynamic coefficients in Turning Circle test

Schematic presentation and value of the deviation between computed and experimental trajectories are presented in Fig. 4.

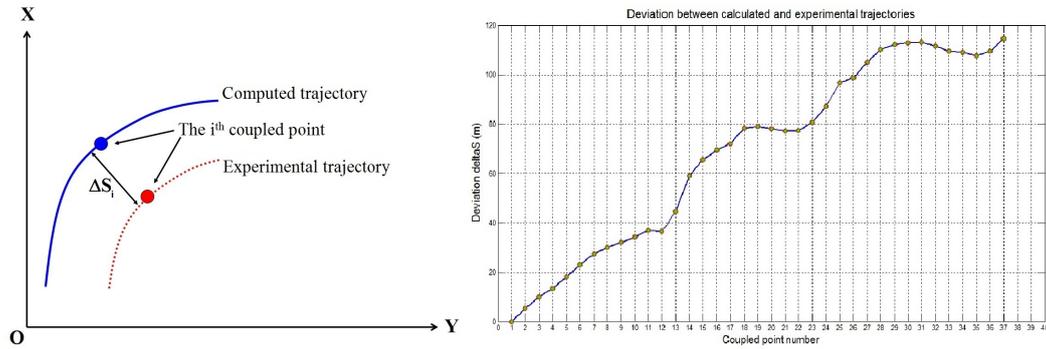


Figure 4: Schematic presentation of the deviation between computed and experimental trajectories

No	Coefficient	Value	No	Coefficient	Value
6	N_T^{II}	-0.02	23	$X_{c c \beta\delta}^{II}$	0.152
15	$Y_{ur\xi}^{II}$	0.182	24	$N_{ c c\delta}^{II}$	-0.098
16	$N_{ur\xi}^{II}$	-0.047	31	N_{ur}^{II}	-0.207
20	$Y_{ c c\delta}^{II}$	0.208	34	$X_{ u u\xi}^{II}$	-0.0061
22	$N_{uv\xi}^{II}$	-0.241	35	$X_{c c \delta\delta}^{II}$	-0.093

Table 3: The 10 most sensitive coefficients of Esso Bernicia Tanker model

Average optimal deviation was calculated as below: $\Delta S_{average} = \frac{\sum_{i=1}^N \Delta S_i}{N} = 68.0(m)$
 The 10 most sensitive coefficients table in Turning Circle test (as condition 7) is presented in Tab. 3.

5.2 Numerical simulation and coefficient identification after optimization

The optimization solution obtained by SQP, BFGS and Simplex methods is presented in Tab. 4

Resolution of 10 optimal hydrodynamic coefficients obtained by SQP, BFGS and Simplex methods is presented in Tab. 5

Method	SQP	BFGS	Simplex
Tolerance of objective function	10^{-4}	10^{-4}	10^{-4}
Tolerance of optimization variables	10^{-4}	10^{-4}	10^{-4}
Number of iterations	29	9	254
Minimum objective function	0.084	0.120	0.085
Average deviation of optimal trajectory (m)			
$\Delta S_{average} = \frac{\sum_{i=1}^N \Delta S_i}{N}$	5.8	8.0	5.8

Table 4: The optimization solutions obtained by SQP, BFGS and Simplex methods

Var	Coeff	Est	Opt (SQP)	Dev(%) (SQP)	Opt (BFGS)	Dev(%) (BFGS)	Opt (Simplex)	Dev(%) (Simplex)
x(1)	N_T''	-0.02	-0.0240	-16.7	-0.0207	-3.5	-0.0184	8.0
x(2)	$Y_{ur\xi}''$	0.182	0.1598	-13.9	0.1822	0.1	0.2113	16.1
x(3)	$N_{ur\xi}''$	-0.047	-0.0416	13.0	-0.0533	-13.4	-0.0462	1.7
x(4)	$Y_{ c c\delta}''$	0.208	0.1761	-18.1	0.2052	-1.3	0.1904	-8.5
x(5)	$N_{uv\xi}''$	-0.241	-0.2823	-14.6	-0.2400	0.4	-0.2329	3.4
x(6)	$X_{c c \beta\delta}''$	0.152	0.1684	9.7	0.1519	-0.1	0.1902	25.1
x(7)	$N_{ c c\delta}''$	-0.098	-0.0805	21.7	-0.0942	3.9	-0.0820	16.3
x(8)	N_{ur}''	-0.207	-0.2105	-1.7	-0.2096	-1.3	-0.1862	10.0
x(9)	$X_{ u u\xi}''$	-0.0061	-0.0073	-16.4	-0.0065	-6.6	-0.0061	0.0
x(10)	$X_{c c \delta\delta}''$	-0.093	-0.1000	-7.0	-0.0936	-0.6	-0.1100	-18.3

Table 5: The 10 optimal hydrodynamic coefficients (Var:Variable, Coeff:Coefficient, Est:Estimation, Opt:Optimization, Dev:Deviation)

Optimal trajectory and evolution of objective function obtained by SQP, BFGS and Simplex methods are presented in Fig. 5, Fig. 6, and Fig. 7.

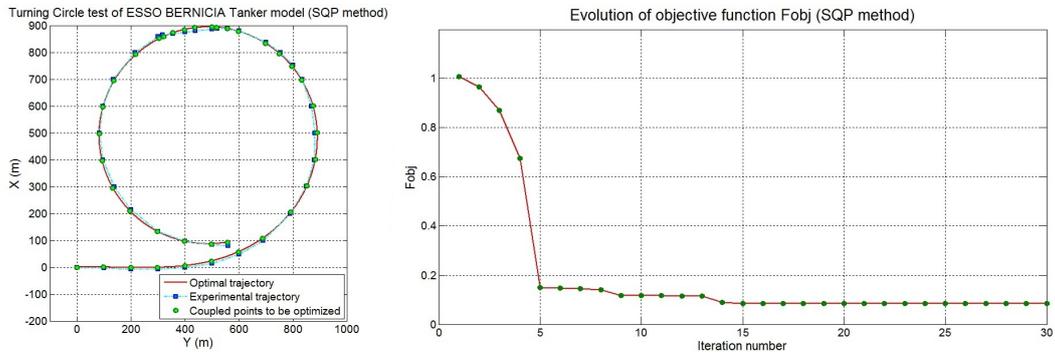


Figure 5: Optimal trajectory and evolution of objective function in Turning Circle test (SQP method)

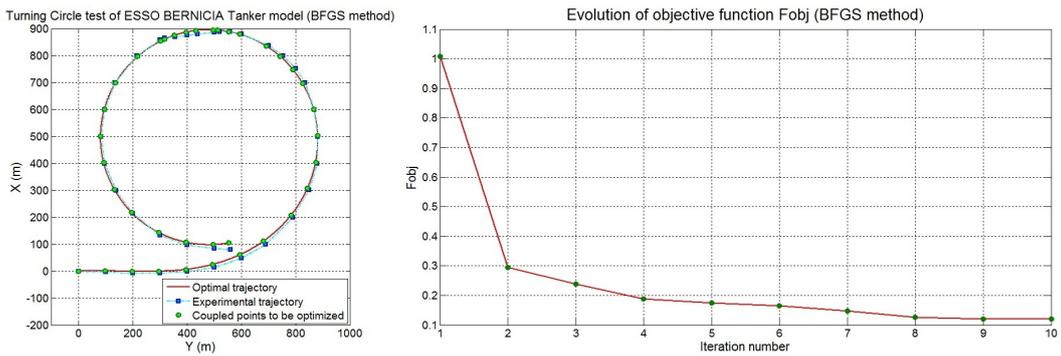


Figure 6: Optimal trajectory and evolution of objective function in Turning Circle test (BFGS method)

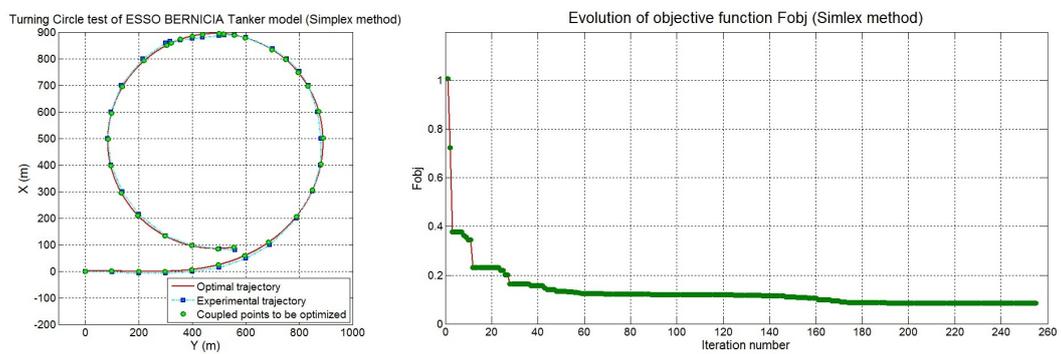


Figure 7: Optimal trajectory and evolution of objective function in Turning Circle test (Simplex method)

As it can be seen in Tab. 4, in the optimization results of SQP method, $F_{obj_{min}}$ is minimum, $|\Delta S_{average}|$ is also minimum and number of iterations is acceptable. So among these methods, SQP method is the robust method in case of Turning Circle test.

6 CONCLUSIONS

- In the present study, a ship manoeuvring simulation model was elaborated with a numerical procedure based on the coupling of optimization techniques and ship motion simulation. To identify the hydrodynamic coefficients, an automatic approach is proposed with two main steps: firstly, a sensitivity analysis to identify the most sensitive coefficients; secondly, optimization techniques to calculate their optimal value.
- The sensitivity analysis of ship hydrodynamic coefficients is based on the analysis of the gradient of objective function while varying the values of each coefficient, so as to find the most sensitive coefficients. In the step of optimization techniques, the form of objective function was developed and the optimization methods were applied. The optimization techniques are carried out in 2 options of multi-variable optimization problem: The unconstrained optimization problem applying the Simplex method and BFGS method; The constrained optimization problem applying the SQP method.
- Our ship manoeuvring simulation model was validated by using experimental data of Esso Bernicia Tanker model (190000dwt) for the Turning Circle test. The coefficient identification was carried out successfully with a good optimization result. Applying the SQP method to approximate the computed and experimental trajectories, then an averaged optimal discrepancy of 5.8m is obtained. This discrepancy is reduced from a value of 68.0m (before optimization) with a converge was reached after 29 iterations.

REFERENCES

- [1] Y. Hollocou, D.V.Q.A Nguyen, A. Kanschine and H. Thuillier, *Prsentation d'un Modle Mathematique de Navire*, Service des Phases et Balises, (1977).
- [2] Y. Hollocou et S.H. Lam, *Modlisation de volution d'un navire en manoeuvre*, Centre d'Etudes Techniques Maritimes Et Fluvials (CETMEF), (2008).
- [3] M. Vayssade and A. Pourplanche, *A piloting SIMulator for maritime and fluvial NAVigation: SimNav*, Lecture Notes in Computer Science, vol. 2840, pp. 672-676, (2003).
- [4] R. Skjetne. *The Maneuvering Problem*, NTNU PhD Thesis, Norway, 2005.
- [5] S. Sutulo, L. Moreira, C. G. Soares, *Mathematical models for ship path prediction in manoeuvring simulation systems*, Ocean Engineering 29, pp. 1-19, (2002).
- [6] K. Mc Taggart, *Simulation of Hydrodynamic Forces and Motions for a Freely Manoeuvring Ship in a Seaway*, DRDC Atlantic TM 2005-071, (2005).
- [7] Thor I.Fossen, *Guidance and Control of Ocean Vehicles*, John Wiley&Sons, (1994).
- [8] E.V. Levis, *Principles of Naval Architecture*, SNAME, (1989).
- [9] V.Bertram, *Practical Ship Hydrodynamics*, Butterworth-Heinemann, (2000).
- [10] H.K. Yoon et K.P. Rhee, *Identification of hydrodynamic coefficients in ship manoeuvring equations of motion by Estimation-Before-Modeling technique*, Ocean Engineering 30, pp. 2379-2404, (2003).
- [11] M. Viviani, C.P. Bonvino, R. Depascale, F. Conti and M. Soave, *Identification of Hydrodynamic Coefficients from Standard Manoeuvres for a Series of Twin-Screw Ships*, 2nd International Conference on Marine Research and Transportation (ICMRT '07), Italia, pp. 99-108, (2007).
- [12] W.Y. Hwang, *Application of System Identification to Ship Maneuvering*, MIT PhD Thesis, (1980).
- [13] S. Inoue, M. Hirano and K. Kijima, *Hydrodynamic Derivatives on Ship Manoeuvring*, Ship Building, Marine Technology Monthly, Vol. 28, No. 321, pp.112-125, (1981).
- [14] S. Inoue, M. Hirano, K. Kijima and J. Takashina, *A Pratical Calculation Method of Ship Maneuvering Motion*, Ship Building, Marine Technology Monthly, Vol. 28, No. 325, pp.207-222, (1981).
- [15] A. Antoniou and W.S. Lu, *Practical Optimization: Algorithms and Enginnering Applications*, Springer, (2007).

EVOLUTIVE, ITERATIVE AERODYNAMICAL GLOBAL OPTIMIZATION, WITH WEAK AERODYNAMICS/ STRUCTURE INTERACTIONS

A. NASTASE*

* Aerodynamik des Fluges, RWTH-Aachen
Templergraben 55, 52062 Aachen, Germany
e-mail: nastase@lafaero.rwth-aachen.de

Key words: Global Shape Optimization, Enlarged Variational Method, Coupling Aerodynamics/Structure, Full Navier-Stokes Partial-Differential Equations, Multiphysics Problems, Hybrid Numerical Solutions.

Summary. *The aerodynamical global optimization (GO) of the shape of a flying configuration (FC) leads to an extended variational problem with free boundaries. An own evolutive iterative optimum-optimorum (OO) theory was developed in order to solve this enlarged variational problem inside of a class of admissible FCs, defined by some chosen properties. An inviscid, GO shape of FC is used only in its first step of iteration. This shape is checked by using hybrid numerical solutions of the Navier-Stokes layer (NSL). The friction drag coefficient is determined and a control of the inviscid GO shape of FC, for the structure point of view, is performed. New and/or modified initial constraints can occur, due to the interactions aerodynamics/ structure. In the second step of the iterative OO-theory, migrations in the drag functional (which is now the total drag) and in the constraints are performed.*

1 INTRODUCTION

The enlarged variational problem with free boundaries concerns the performing of the aerodynamical GO shape of the outer surface of a FC and also of its planform, in order to obtain a minimum drag, at cruising Mach number and to fulfill some constraints. An own OO-theory was developed in order to solve this enlarged variational problem with free boundaries, inside of a chosen class of admissible FCs, defined by some chosen properties. According to the OO-theory, two FCs belong to the same class, if their surfaces are piecewise approximated through homogeneous polynomials in two variables of the same chosen maximal degrees, the planforms are polygons, which can be related through affine transformations and the shapes of the FCs fulfill the same constraints. The parameters of global optimization are the coefficients of the polynomial expansions of the surface and the similarity parameters of the planform of the FC. Further, a lower limit hypersurface of the drag functional $C_d^{(i)}$, as function of the similarity parameters ν_i of the planform is defined, namely:

$$(C_d^{(i)})_{opt} = f(v_1, v_2, \dots, v_n) \quad . \quad (1)$$

Each point of this hypersurface is obtained by solving a classical variational problem with given boundaries (i.e. a given set of similarity parameters). The position of the minimum of this hypersurface, which is numerically determined, gives us the best set of the similarity parameters and the FC's optimal shape, which corresponds to this set, is at the same time the global optimized FC's shape of the class. The OO-theory was applied to the effective determination of the inviscid GO shape of some FCs.

The variational problem concerns the determination of the GO shape of the structure located inside of the aerodynamical GO shape in order to obtain a minimum weight and to fulfill some constraints in order to assure the necessary stiffness and to limit the magnitude of deformation. Between the aerodynamical and the structural optimizations it exists a certain degree of freedom, which suggests weak interactions.

The evolutive, iterative OO-theory of the author is proposed for this purpose. The OO-theory uses the inviscid GO shape of FC only in the first step of the iteration, as in ¹. This shape is checked by using hybrid numerical solutions for the three-dimensional partial-differential equations (PDEs) of the Navier-Stokes layer, which allows the computation of friction drag coefficient. This shape is also controlled for the structure point of view, in its longitudinal central section and in some transversal sections (especially located in the rear part of FC). The requirements of the structure stiffness can lead to additional and/or to modification of some initial constraints.

In the second step of iteration, the new functional is the total drag functional and a mutation in the constraints is performed. The deformation of the FC is computed by using solutions of the Sophie Germain PDE. Additional constraints, due to the limitation of the magnitude of deformation, can occur. The iteration is continued, until the maximal shape difference in two consecutive steps of iterations is neglectable.

2 THE INVISCID GLOBAL OPTIMIZATION OF THE SHAPE OF AN INTEGRATED WING-FUSELAGE, IN SUPERSONIC FLOW

The wing-fuselage FC is here considered like an equivalent discontinuous wing alone, which is fitted with two artificial ridges along the junction lines wing/fuselage. If, additionally, the wing and the fuselage have the same mean surface and the same tangent plane, in each point of their junction lines, the equivalent wing of the wing-fuselage configuration is here called integrated wing-fuselage. Its mean surface $Z(x_1, x_2)$ is supposed to be continuous, but, for the sake of generality, the thickness distributions $Z^*(x_1, x_2)$ on the lateral sides, corresponding to the wing and $\bar{Z}'(x_1, x_2)$ on the central part, corresponding to the fuselage zone, are different. The computation is made in a dimensionless system of coordinates $\tilde{O} \tilde{x}_1 \tilde{x}_2 \tilde{x}_3$, as in ¹, ², it is:

$$\tilde{x}_1 = \frac{x_1}{h_1}, \quad \tilde{x}_2 = \frac{x_2}{\ell_1}, \quad \tilde{x}_3 = \frac{x_3}{h_1} \quad . \quad (2a-c)$$

The downwashes w and w^* of the thin and the thick-symmetrical components of the thick, lifting FC are supposed to be expressed in form of superpositions of homogeneous polynomials in two variables, with free coefficients, it is:

- for the thin component of FC:

$$w \equiv \tilde{w} = \sum_{m=1}^N \tilde{x}_1^{m-1} \sum_{k=0}^{m-1} \tilde{w}_{m-k-1,k} |\tilde{y}|^k, \quad (3)$$

- and, for the thick-symmetrical component of FC on the wing and of the fuselage zone:

$$w^* \equiv \tilde{w}^* = \sum_{m=1}^N \tilde{x}_1^{m-1} \sum_{k=0}^{m-1} \tilde{w}_{m-k-1,k}^* |\tilde{y}|^k, \quad w^* \equiv \bar{w}^* = \sum_{m=1}^N \tilde{x}_1^{m-1} \sum_{k=0}^{m-1} \bar{w}_{m-k-1,k}^* |\tilde{y}|^k. \quad (4a,b)$$

The start solutions for the inviscid GO of the shape of the integrated wing-fuselage FC are the three-dimensional analytical hyperbolical potential solutions, for the axial disturbance velocities u and u^* of the author, by using of the hydrodynamic analogy of Carafoli, written in the integrated form, which contain good suited minimal singularities only along the singular lines like: subsonic leading edges, junction line wing-fuselage and, eventually, a central ridge, as in ¹, namely, for the thin and thick-symmetrical components of the wing-fuselage FC with subsonic leading edges, fitted eventually with a central ridge

$$u \equiv \ell \tilde{u} = \ell \sum_{n=1}^N \tilde{x}_1^{n-1} \left\{ \sum_{q=0}^{n-1} \tilde{G}_{nq} \tilde{y}^q \left(\cosh^{-1} S_1' + (-1)^q \cosh^{-1} S_2' \right) + \sum_{q=0}^{E\left(\frac{n}{2}\right)} \frac{\tilde{A}_{n,2q} \tilde{y}^{2q}}{\sqrt{1-\tilde{y}^2}} + \sum_{q=1}^{E\left(\frac{n-1}{2}\right)} \tilde{C}_{n,2q} \tilde{y}^{2q} \cosh^{-1} \sqrt{\frac{1}{\tilde{y}^2}} \right\}, \quad (5)$$

$$u^* \equiv \ell \tilde{u}^* = \ell \sum_{n=1}^N \tilde{x}_1^{n-1} \left\{ \sum_{q=0}^{n-1} \tilde{H}_{nq}^* \tilde{y}^q \left(\cosh^{-1} M_1 + (-1)^q \cosh^{-1} M_2 \right) + \sum_{q=0}^{E\left(\frac{n-2}{2}\right)} \tilde{D}_{n,2q}^* \tilde{y}^{2q} \sqrt{1-\nu^2 \tilde{y}^2} + \sum_{q=1}^{E\left(\frac{n-1}{2}\right)} \tilde{C}_{n,2q}^* \tilde{y}^{2q} \cosh^{-1} \sqrt{\frac{1}{\nu^2 \tilde{y}^2}} + \sum_{q=0}^{n-1} \tilde{G}_{nq}^* \tilde{y}^q \left(\cosh^{-1} S_1 + (-1)^q \cosh^{-1} S_2 \right) \right\}. \quad (6)$$

Hereby M_1 , M_2 , S_1 and S_2 are of the form:

$$\begin{aligned}
M_1 &= \sqrt{\frac{(1+\nu)(1-\nu\tilde{y})}{2\nu(1-\tilde{y})}} \quad , \quad M_2 = \sqrt{\frac{(1+\nu)(1+\nu\tilde{y})}{2\nu(1+\tilde{y})}} \quad , \quad S_1 = \sqrt{\frac{(1+\tilde{\nu})(1-\nu\tilde{y})}{2(\tilde{\nu}-\nu\tilde{y})}} \quad , \\
S_2 &= \sqrt{\frac{(1+\tilde{\nu})(1+\nu\tilde{y})}{2(\tilde{\nu}+\nu\tilde{y})}} \quad . \quad (7a-d)
\end{aligned}$$

The free similarity parameters of the planforms of the wing and of the fuselage zone are $\nu = B\ell$ and $\tilde{\nu} = Bc$ ($B = \sqrt{M_\infty^2 - 1}$, $\ell = \ell_1/h_1$, $c = c_1/h_1$) and M_∞ , ℓ , ℓ_1 , h_1 , c and c_1 are the cruising Mach number, the dimensionless span, the maximal half-span and the maximal depth of the wing, the dimensionless span and the maximal half-span of the fuselage zone. The coefficients of the axial disturbance velocities of the thin and of the thick-symmetrical components of the thick, lifting FC are related to the corresponding downwashes through linear and homogeneous relations, deduced by using the compatibility relations of Germain and the condition on the Mach cone at the apex of the FC, as in ¹. These three-dimensional, analytical hyperbolic potential solutions are used for the determination of the inviscid GO shapes of two integrated wing-fuselage models, namely Fadet I and Fadet II, optimized at cruising Mach numbers $M_\infty = 2.2$ and, respectively, $M_\infty = 3$. and for the generation of the proposed hybrid numerical solutions for the NSL's PDE. The inviscid GO shape of FC is used as first step and the proposed hybrid numerical solutions for the NSL are used, up the first computational checking and up the second step of iteration of the evolutive, iterative OO-theory, as in ¹.

3 THE INVISCID, GLOBAL OPTIMIZATION OF THE SHAPE OF THE MODEL FADET I, VIA OPTIMUM-OPTIMORUM THEORY

The determination of the GO shape of the model Fadet I, via OO-theory, is presented here, as exemplification.

The parameters of optimization are the coefficients of the downwashes \tilde{w}_{ij} , \tilde{w}_{ij}^* and \bar{w}_{ij}^* and the similarity parameters of the planform of the entire FC and of the planform of the fuselage ν and, respectively, $\bar{\nu}$. The quotient $k = \bar{\nu}/\nu$, which depends on the purpose of the FC, is taken constant, during the optimization process. For a given value of the similarity parameter of the planform of the FC, the optimization of the shapes of its thin and thick-symmetrical components can be separately treated.

The initial constraints for the thin FC component are: given lift and pitching moment coefficients and also the own introduced Kutta condition on leading edge in order to cancel the induced drag and to avoid the contournement of flow on leading edge, at cruise. The initial constraints for the thick-symmetrical FC component are: the given relative volumes of the fuselage and of entire FC, the closing condition on leading edges and also the own introduced integration conditions along the junction lines wing/fuselage.

For a given value of the similarity parameter of the planform of the FC, the corresponding optimal values of the coefficients of downwashes are analytical, uniquely determined by solving of two linear algebraic systems, as in ^{1, 2}.

If the similarity parameter of the planform of FC is sequentially varied, a lower limit line of the inviscid drag functional of optimal FCs, as function of this similarity parameter, is obtained by solving the variational problems for the corresponding values of the similarity parameter (for FCs with subsonic leading edges is $0 < \nu < 1$). The position of the minimum of this limit line gives the optimal value of the similarity parameter $\nu = \nu_{opt}$ and the corresponding optimal FC is, at the same time, the global optimized FC of the class.

The GO shape of the fully-integrated wing-fuselage model Fadet I, designed by the author, for the cruising Mach number $M_\infty = 2.2$, obtained by using the OO-theory, is presented in the (Fig. 1) and looks, in transversal sections, like a flying bird!

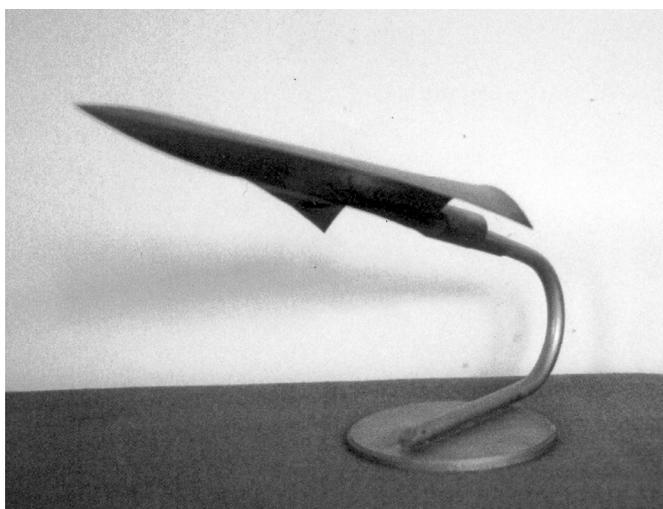


Figure 1: The global optimized shape of the fully-integrated wing-fuselage model Fadet I

The aerodynamical characteristics of the integrated, GO shape of the model Fadet I were checked in the trisonic wind tunnel of the DLR-Cologne, in the frame of the author's research project, sponsored by the DFG. Very good agreements between the theoretical and experimental results, obtained by using inviscid analytical, hyperbolic potential solutions for its lift and pitching moment coefficients, are presented in the (Fig. 2a,b) and (Fig. 2c,d) and between the theoretical predicted distribution of pressure coefficient with experimental results, in the longitudinal central section of the upper side of model Fadet I, at the angles of attack $\alpha = -8^\circ, 0^\circ, 8^\circ$ and for the range of Mach numbers $M_\infty = (1.25 \div 2.4)$ are obtained, as in ¹. For this range of Mach numbers, the model Fadet I has subsonic leading edges. These good agreements between experimental and theoretical predicted values of lift, pitching moment and pressure obtained by using hyperbolic analytical potential solutions for the axial disturbance velocities on FC, at moderate angles of attack, proposed here as in ¹, lead to the following important remarks:

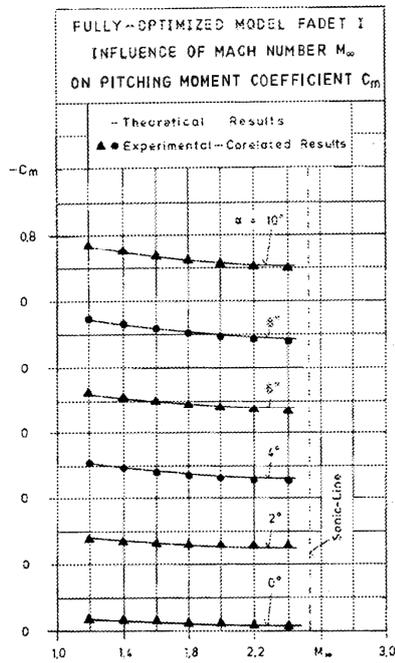
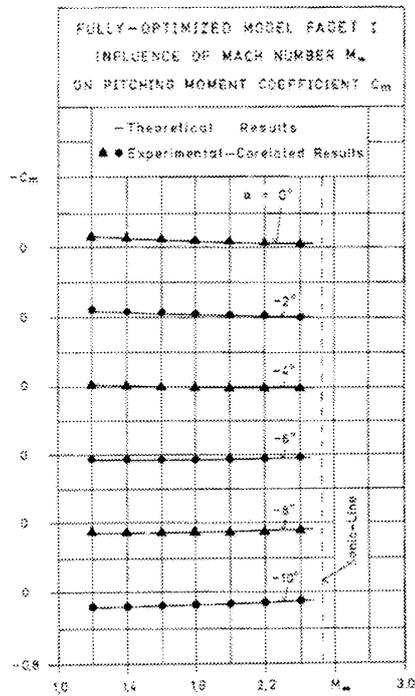
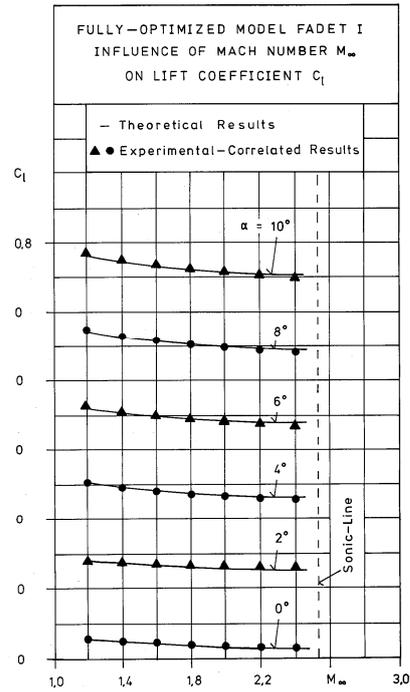
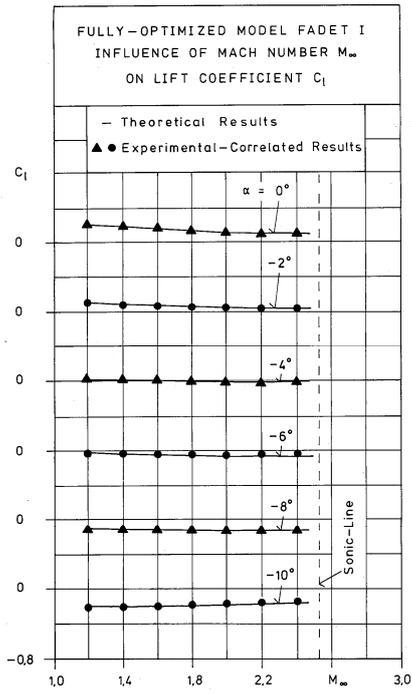


Fig. 2a-d: The agreement between the correlated experimental values and the theoretical predicted distributions of the lift and pitching moment coefficients of the global optimized model Fadet I

- these proposed solutions for the axial disturbance velocities obtained by using balanced minimal singularities only along singular lines and the corresponding developed software, for the computation of the above coefficients, are confirmed;
- the flow is laminar, as supposed here, and remains attached in supersonic flow, for larger range of angles of attack than in subsonic flow;
- if the FC is flattened enough, the economical supersonic flight with characteristic surface as supposed here, instead of flight with shock wave surface, is confirmed;
- the influence of friction upon the above coefficients is neglectable;
- these solutions contain useful informations over the correct behaviors of the flow in the vicinity of the singular lines and can be seen as asymptotical solutions, at the NSL's edge..

Further, hybrid NSL's solutions are proposed here, especially for the computation of the total drag coefficient, including friction and for the computation of the other aerodynamical characteristics, at higher angles of attack.

4 HYBRID NUMERICAL SOLUTIONS FOR NAVIER-STOKES LAYER

The proposed numerical solutions for the NSL's PDEs use the three-dimensional hyperbolic analytical potential solutions given above two times, namely: at the NSL's edge and in the structure of their velocities components, which are products between the analytical solutions of the same FC and polynomials with free coefficients. A spectral coordinate η is defined, as in [1], namely:

$$\eta = (x_3 - Z(x_1, x_2)) / \delta(x_1, x_2) \quad . \quad (8)$$

Hereby $Z(x_1, x_2)$ is the equation of the surface of the flattened FC and $\delta(x_1, x_2)$ is the thickness distribution of NSL. The spectral forms of the axial, lateral and vertical velocity's components u_δ , v_δ and w_δ , the density function $R = \ln \rho$ and the absolute temperature T are here proposed:

$$u_\delta = u_e \sum_{i=1}^N u_i \eta^i, \quad v_\delta = v_e \sum_{i=1}^N v_i \eta^i, \quad w_\delta = w_e \sum_{i=1}^N w_i \eta^i,$$

$$R = R_w + (R_e - R_w) \sum_{i=1}^N r_i \eta^i, \quad T = T_w + (T_e - T_w) \sum_{i=1}^N t_i \eta^i \quad . \quad (9a-e)$$

Here R_w and T_w are the given values of R and T at the wall, u_e , v_e , w_e , R_e and T_e are the values of u , v , w , R and T at the NSL's edge, obtained from the outer inviscid hyperbolic potential flow, and u_i , v_i , w_i , r_i and t_i are their free spectral coefficients, which are determined by fulfilling of the NSL's PDEs. The physical equation of ideal gas for the pressure p and an exponential law of the viscosity μ versus T are used:

$$p \equiv R_g \rho T = R_g e^R T, \quad \mu = \mu_\infty \left[\frac{T}{T_\infty} \right]^{n_1} \quad . \quad (10a,b)$$

Here are: R_g and T_∞ the universal gas constant and the absolute temperature of the undisturbed flow and n_1 is the viscosity exponent. The spectral forms (9a-e) automatically satisfy the boundary conditions at wall ($\eta=0$). Seven free coefficients of the velocity's components are used to satisfy the boundary conditions at the NSL's edge.

By using of a logarithmic density function R and of the collocation method in the continuity PDE, it was possible to determine the free spectral coefficients r_i as function of the free spectral coefficients of the velocity's components by solving of a linear algebraic system.

The temperature PDE and the collocation method are used to determine the free spectral coefficients t_i of the absolute temperature T as function only of the free spectral coefficients of the velocity's components by solving of a transcendental algebraic system. A splitting of the NSL's PDEs is realized, due to the use of the logarithmic density function.

The free spectral coefficients of the velocity's components are determined by using the impulse PDEs, which are treated iteratively, as in ².

The hybrid discontinuous numerical NSL's solutions, proposed here, have the following important properties:

- their viscous/inviscid coupling, at the NSL's edge, does not need interface;
- they have correct last behaviors;
- they have correct jumps along the singular lines and the singularities are chosen by using the principle of minimal singularities and these singularities are balanced;
- they are accurate, because the partial derivatives of the velocity's components are exactly computed;
- they are split and therefore the speed up of the computation time occurs;
- additionally, for hyperbolical PDEs the condition of their characteristical surface is automatically fulfilled.

The hybrid NSL's solutions are used for the computational checking of the formerly determined inviscid GO shape of FC and up the second step of the iterative OO-theory, presented below.

5 THE ITERATIVE OPTIMUM-OPTIMORUM THEORY, WITH WEAK INTERACTIONS AERODYNAMICS/STRUCTURE

Further, an iterative OO-theory is proposed, which uses the hyperbolical potential solutions as start solutions for the optimization and the inviscid GO shape of FC, only in the first step of its iteration. An intermediate computational checking of the inviscid GO shape of the FC is made with own zonal spectral viscous solvers, for the three-dimensional NSL. The friction drag coefficient $C_d^{(f)}$ of the FC is determined and represents about 27% of the total drag of Fadet I, at cruising Mach number. The inviscid GO shape of FC is checked also for the structural point of view. Weak interactions aerodynamics/structure are here proposed. Additional or modified constraints, introduced in order to control the camber, twist and thickness distributions of the inviscid GO shape of FC, for structure reasons, in the central and, especially, in the rear part of FC are proposed. A limitation of the twist can be necessary for FCs, which are flying at higher supersonic Mach number and it can be realized by imposing the constraint of pressure equalization along the subsonic leading edges at a lower supersonic

Mach number than the cruising one. The thickness distribution can be augmented in the rear part, by artificial prolongation of the depth and, after the optimization of the thickness distribution, the initial depth is restored by cutting and elimination of the prolongation.

Up the second step of optimization, the predicted inviscid GO shape of the FC is reshaped, by including all the constraints in the variational problem and of the friction drag coefficient in the drag functional, as presented in the (Fig. 3). The iteration is continued until the maximal difference of the GO shape of the FC in two consecutive steps of iteration is neglectable.

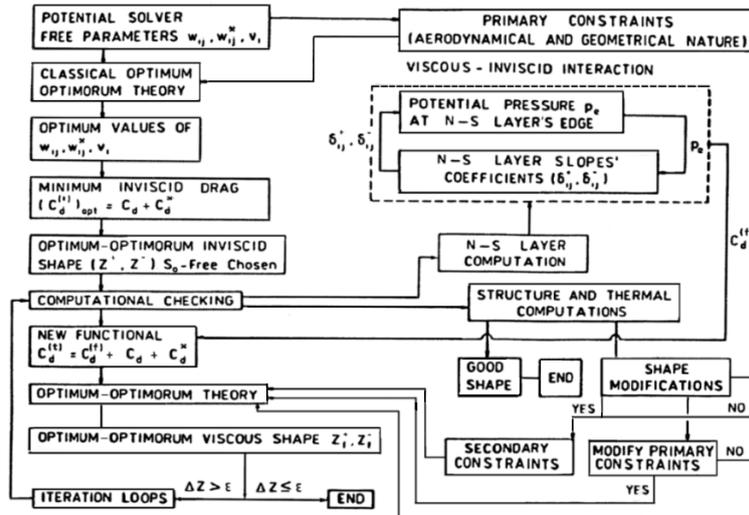


Figure 3: The evolutionary, iterative optimum-optimorum theory

6 CONCLUSIONS

The hybrid solutions for the NSL, proposed here, are matched with the outer inviscid hyperbolic potential flow, at the NSL's edge. The viscous/inviscid interaction is realized without needing an interface.

The weak interaction aerodynamics/structure is realized by introduction of secondary constraints and by modifying the primary constraints of the aerodynamical GO of the FC's shape, according to the structure needs.

The deformation of the structure, due to the distributions of weight and to the aerodynamical pressure, can be computed by using new solutions of the Sophie Germain PDE. The final viscous GO shape of the FC is obtained by subtracting the distribution of deformation.

REFERENCES

- [1] A. Nastase, *Computation of supersonic flow over flying configurations*. Elsevier, Oxford, UK (2007).
- [2] A. Nastase, Improved hybrid solutions for the Navier-Stokes layer, *Proceedings of ARA Congress, Sibiu*, (2009), p. 39-42.

MODEL ORDER REDUCTION OF SYSTEMS FOR ACTIVE VIBRATION AND NOISE CONTROL

MATTHIAS KURCH*, HEIKO ATZRODT*, VALERIO CARLI*, OLIVER HEUSS* AND JAN MOHRING†

* Fraunhofer Institute for Structural Durability and System Reliability LBF
Bartningstr. 47, 64289 Darmstadt, Germany
e-mail: matthias.kurch@lbf.fraunhofer.de, www.lbf.fraunhofer.de

† Fraunhofer Institute for Industrial Mathematics ITWM
Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany
e-mail: jan.mohring@itwm.fraunhofer.de, www.itwm.fraunhofer.de

Key words: Model Order Reduction, Multiphysics Problem, Finite Element Method, Experimental Verification

Abstract. This paper presents a framework for the design of active vibration control (AVC) and active noise control (ANC) systems. The framework is composed of a finite element (FE) model, model order reduction (MOR) methods and software for system-level simulations. The finite element method (FEM) is used to develop an experimentally verified model of the coupled structural, acoustic and piezoelectric problem. This model serves as an example for the study and discussion of modal and moment matching based MOR approaches. Finally, a reduced model is used to design an active control approach, which proves the feasibility of the framework.

1 INTRODUCTION

Many of today's structures, designed to reduce weight and material costs are more susceptible to vibrations than traditionally-designed structures. This lightweight design may cause increased noise, vibration and fatigue problems. The vibration and the sound radiation are often reduced by installation of active vibration and noise control. Piezoelectric sensors and actuators are linked with a controller that adapts itself to changing operating conditions. To be effective, there is a need to study different designs using numerical simulations in order to evaluate adaptive control strategies and feasible sensor or actuator concepts.

The FEM is a well established tool to set up customizable models of structures. It is then possible to create configurable dynamic models for the mechanical as well as the acoustic domains. Researchers proposed finite element formulations for piezoelectric [1]

and acoustic problems which are implemented in commercial FEM codes by now. To predict the real dynamic behaviour of the structure, it is effective to perform an experimental modal analysis (EMA). Based on these results, the FEM model can be tuned towards its real behavior. Unfortunately, the FEM models do not satisfy the requirements of the subsequent computer aided design of control systems which is carried out in the time domain. The dimension of these FEM models can be so large that time integration becomes inefficient or even prohibitive. An approach to solve this issue is by the application of model order reduction [2] methods. These techniques approximate the dynamical model by one of a smaller dimension while preserving its input-output behavior. After a description of the demonstration object used for this paper, the set-up of its FE model is explained in Section 2. The model is verified using measured data. The test preparation and the results are presented in Section 3. Among the model order reduction approaches, a modal technique [3] for unsymmetric system matrices and a moment matching method via Krylov subspaces [4] by means of the Arnoldi process are introduced in Section 4. In the remainder of this paper, tests are executed in order to compare the performance of these methods with each other. Selected components of the acoustic box are reduced and the performances, as well as the results, are compared in Section 5. After this preprocessing, the reduced model is imported into the simulation software MATLAB/Simulink. The updated and reduced model is used to implement a control approach in order to show its capability. Within the MATLAB/Simulink environment, the interaction of structure, actuators, sensors and controller is optimized (Sec. 6) until the magnitude of vibration or sound radiation is minimal. Finally, the work is concluded and further research is outlined.

2 THE FINITE ELEMENT MODEL

In order to study and test the sound transmission, to develop reduction methods for sound radiation [5], an acoustic demonstrator was manufactured at the LOEWE-Zentrum AdRIA. The demonstrator consists of a cuboid box (Fig. 1a) with sound-reflecting walls. The top of the box is covered by a clamped elastic aluminium plate (Fig. 1b). The box is stiff, compared to its cover, and sound transmission through the box is negligible. Preliminary studies showed that this assumption holds true up to 500 Hz. It is possible to study and test acoustical behavior and smart structure systems for noise reduction in a frequency range from 0 to 500 Hz.

For this demonstration object FEM models were set up using the FE package ANSYS 12.1. Assuming that the acoustic fluid inside the box is incompressible, inviscid and that there is no mean flow of the fluid and density and pressure are constant throughout the fluid, the acoustic cavity was discretized using the 3-D acoustic element FLUID30. The element has eight corner nodes with four degrees of freedom (DOF) per node. These are the translations in the nodal x-, y- and z-directions and the pressure. The walls of the box are not modelled, because they are assumed to be stiff have no sound transmission. This effect is simulated when no absorption at the boundary is applied and the nodal



Figure 1: AdRIA acoustic box: a) general view, b) top view.

translation of the elements in the cavity is deactivated. The fluid-structure interaction (FSI) is induced by a layer of coupling elements between the cavity and the plate. Considering the partial differential equations (PDE) of this acoustics fluid-structure coupling, the discretization of this equation by means of the FEM yields a system of N ordinary differential equations:

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{D}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{B}^{\text{inf}}\mathbf{f}(t) \quad (1a)$$

$$\mathbf{y}(t) = \mathbf{B}_1^{\text{out}}\mathbf{q}(t) + \mathbf{B}_2^{\text{out}}\dot{\mathbf{q}}(t). \quad (1b)$$

where $\mathbf{M}(t)$, $\mathbf{D}(t)$, $\mathbf{K}(t)$ are the system matrices, $\mathbf{B}^{\text{inf}}\mathbf{f}(t)$ are the loads, and $\mathbf{q}(t) = (\mathbf{u}(t) \ \mathbf{p}(t))^T$ is a vector of unknown degree of freedom, where $\mathbf{u}(t)$ is the mechanical displacement and $\mathbf{p}(t)$ is the acoustic pressure. For the sake of simplicity, the time-dependence of the variables will be dropped from further calculation. The mass, damping and stiffness matrix \mathbf{M} , \mathbf{D} , \mathbf{K} are assembled as follows:

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{uu} & \mathbf{0} \\ \mathbf{M}_{pu}^{fs} & \mathbf{M}_{pp} \end{pmatrix}, \mathbf{D} = \begin{pmatrix} \mathbf{D}_{uu} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{pp} \end{pmatrix}, \mathbf{K} = \begin{pmatrix} \mathbf{K}_{uu} & \mathbf{K}_{up}^{fs} \\ \mathbf{0} & \mathbf{K}_{pp} \end{pmatrix}. \quad (2)$$

The mass matrix \mathbf{M} and the stiffness matrix \mathbf{K} are both unsymmetric [6]. The loads are $\mathbf{B}^{\text{inf}}\mathbf{f} = (\mathbf{F} \ \mathbf{0})^T$. When piezoelectric transducers are applied to the aluminum plate of the box, the FE formulation have to be extended. Adding the governing equations and the linear piezoelectric material law to the formulation yields [1]:

$$\mathbf{M} = \begin{pmatrix} \mathbf{M}_{uu} & \mathbf{0} & \mathbf{0} \\ \mathbf{M}_{pu}^{fs} & \mathbf{M}_{pp} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{D} = \begin{pmatrix} \mathbf{D}_{uu} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_{pp} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{D}_{\phi\phi} \end{pmatrix} \mathbf{K} = \begin{pmatrix} \mathbf{K}_{uu} & \mathbf{K}_{up}^{fs} & \mathbf{K}_{u\phi} \\ \mathbf{0} & \mathbf{K}_{pp} & \mathbf{0} \\ \mathbf{K}_{u\phi}^T & \mathbf{0} & -\mathbf{K}_{\phi\phi} \end{pmatrix}. \quad (3)$$

The mass, damping, and stiffness matrices from ANSYS 12.1 are non-symmetric and/or singular. The loads and the degree of freedom vector are assembled $\mathbf{B}^{\text{inf}}\mathbf{f} = (\mathbf{F} \ \mathbf{0} \ \mathbf{Q})^T$ and $\mathbf{q} = (\mathbf{u} \ \mathbf{p} \ \phi)^T$ respectively. The matrix properties make demands on the MOR algorithms, which are discussed later in this paper (see Sec. 4).

To study the MOR techniques, three FEM models were set up. First, a pure mechanical

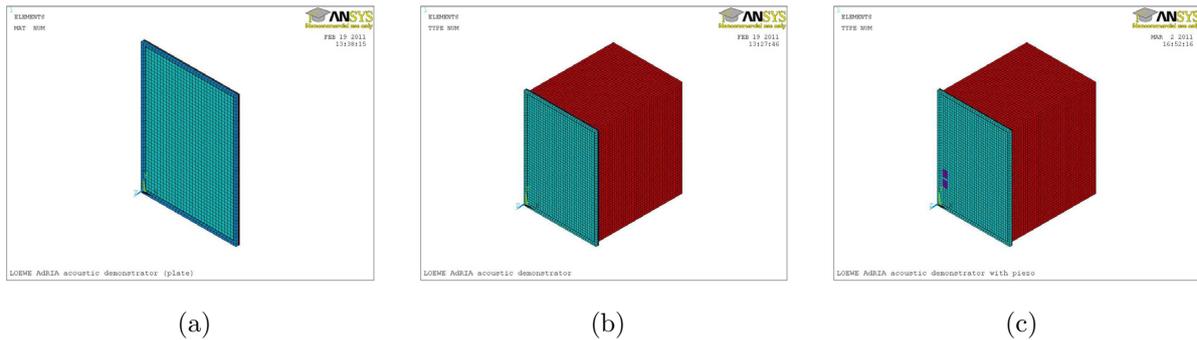


Figure 2: FEM models: a) plate, b) plate and cavity c) plate, cavity and piezo patches.

model of the aluminium plate was realized (Fig. 2a). The plate was discretized using eight nodes structural elements that exhibit linear displacement behavior. Certainly quadratic elements are more recommended due to their better strain approximation. However these set-up may require contact elements at the interface to the fluid in subsequent analyses, because no quadratic fluid elements are available. However such elements complicate the model but provide no additional intelligence. The linear elements were used accordingly. The support of the plate which consists of a frame and the rubber sealing (Fig. 1b) was modeled by means of linear elements as well.

For the analysis of the coupled structural acoustic behaviour the FE model was extended. This acousto-mechanical model is depicted in Fig. 2b. Assuming that the walls of the box are rigid (up to 500 Hz only), the volume of the cavity was modelled. The volume was discretized using FLUID30 elements which features fluid medium behaviour and the interface in fluid/structure interaction problems. The third model is identical to the latter, however piezoceramic patch actuators (Fig. 2c) were added at the left side of the plate in order to design active noise and vibration control.

3 EXPERIMENTAL VERIFICATION OF THE MODEL

In order to validate the coupled vibro-acoustical model, two EMA tests were carried out. In the first test the FE model without the acoustic cavity was validated using measured data which were available from previous studies [7]. The purpose of this was to find a feasible model for the plate clamping. Assuming that the clamping is roughly similar to a fixed support, the Young's modulus of the discretized sealing was adjusted [8] until the averaged relative error of the measured and calculated eigenfrequencies became 2.4 %. The second EMA was used to validate the coupled mechano-acoustical model. The quantities to be measured for the experimental vibro-acoustical modal analysis are the excitation of the structure and fluid as well as the displacement, the velocity or the acceleration responses. A force applied to the mechanical part of the structure or a defined volume displacement to the acoustical fluid are feasible excitations. In the present case,

the excitation was realized by applying a point force with an electromechanical shaker (see Fig. 1b). The driving point spectra were captured with an impedance sensor, which allows the acquisition of force and acceleration simultaneously. The structure responses were measured with a laser Doppler vibrometer (LDV) at 1276 points of the plate, which are nearly coincident with the nodes of the FE mesh. Finally, the acoustic responses were recorded using a microphone inside the lower right corner of the cavity. A frequency domain multiple degree of freedom (MDOF) analysis using the PolyMAX algorithm, which leads to a modal model of the coupled system was conducted. There exist 15 structural and 13 acoustical modes in the frequency range up to 500 Hz. The 4,1 mode and the mode at 279.10 Hz were not detected using this set-up. The results of the measurement were used to tune the FEM model towards the real eigenfrequencies (Table 1). The relative

Table 1: Eigenfrequencies of the acousto-mechanical system.

No.	mode shape	f_{EMA} [Hz]	damp	f_{FEM} [Hz]	rel. error
1	1,1	61.63	1.71	60.51	1.824
2	1,2	98.839	1.2	98.09	0.757
3	2,1	144.888	0.89	140.61	2.953
4	1,3	165.332	0.97	163.23	1.271
5	2,2	182.849	0.81	182.15	0.382
6		197.449	0.14	198.76	-0.664
7		224.981	0.15	226.01	-0.457
8	2,3	246.434	0.8	247.16	-0.295
9	1,4	256.418	0.8	252.49	1.532
10	3,1	278.922	0.31	269.33	3.439
11				279.10	
12		297.964	0.21	299.97	-0.673
⋮	⋮	⋮	⋮	⋮	⋮
29	4,2	493.171	1.14	505.71	-2.543

error is almost less than 1%. For the comparison of experimental and numerical results, the modal assurance criterion (MAC) [9] of the plate displacement is depicted in Fig. 3. For the 1st to the 5th, the 8th, the 9th and the 17th mode, the MAC value is 100% which implies very good correlation. The 6th, the 7th and the 16th mode are acoustic resonances, therefore a MAC value of 100% is not mandatory. When the 12th and 13th, the 14th and 15th, or 18th and 19th mode is considered, one can see that their mode shapes are similar or even equal because of the structural-acoustic coupling. However, the correlation of the FEM model and the experiment are good and the model is valid for further studies.

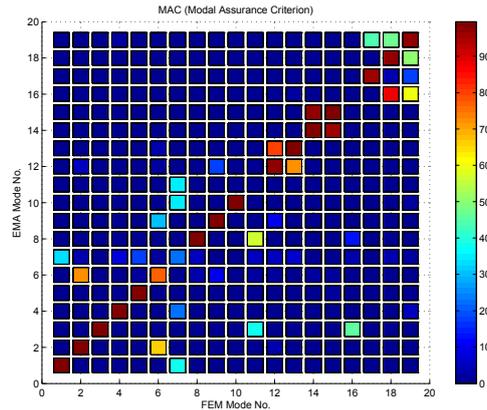


Figure 3: 2-D presentation of MAC Values.

4 MODEL ORDER REDUCTION

A common approach [2] to model reduction is to find a transformation \mathbf{V}_n to a low-dimensional subspace. This transformation $\mathbf{q} = \mathbf{V}_n \mathbf{q}_n + \varepsilon$ should approximate the system behavior accurately within an error bound and project Eq. 1b onto that subspace. Therefore the reduced system becomes:

$$\mathbf{M}_n \ddot{\mathbf{q}}_n + \mathbf{D}_n \dot{\mathbf{q}}_n + \mathbf{K}_n \mathbf{q}_n = \mathbf{B}_n^{in} \mathbf{f} \quad (4a)$$

$$\mathbf{y} = \mathbf{B}_{1,n}^{out} \mathbf{q} + \mathbf{B}_{2,n}^{out} \dot{\mathbf{q}}, \quad (4b)$$

where $\mathbf{M}_n = \mathbf{V}_n^T \mathbf{M} \mathbf{V}_n$, $\mathbf{D}_n = \mathbf{V}_n^T \mathbf{D} \mathbf{V}_n$, $\mathbf{K}_n = \mathbf{V}_n^T \mathbf{K} \mathbf{V}_n$, and $\mathbf{B}_n^{in} = \mathbf{V}_n^T \mathbf{B}^{in}$. The transformation matrix \mathbf{V}_n can be determined by different methods. For this paper, the component mode synthesis (CMS) [10], a moment matching method via Krylov subspaces [11] and the modal reduction of non-symmetric systems [3] were studied.

4.1 Component Mode Synthesis

The component mode synthesis (CMS) was first proposed by Hurty [12] and further developed by Craig and Bampton [10]. The method has been developed with the purpose of analysing a complex structure as an assembly of less complex sub-structures. After reduction of the size of each sub-structure, all reduced models are then assembled into the global model, which has a much smaller size compared to the physical model. Considering a sub-structure A the physical DOFs are partitioned into boundary DOFs \mathbf{q}_b^A and internal DOFs \mathbf{q}_i^A . The latter set is reduced by replacing it with the vector \mathbf{p}_N of the generalised modal coordinates. The transformation matrix \mathbf{V}_n for A is given by:

$$\mathbf{q}^A = \mathbf{V}_n \mathbf{p}^A = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{\Phi}_C & \mathbf{\Phi}_N \end{pmatrix} \begin{pmatrix} \mathbf{q}_b^A \\ \mathbf{p}_N^A \end{pmatrix} \quad (5)$$

In Eq. 5, $\mathbf{\Phi}_C$ is the matrix of the constraint modes of the sub-structure A . The matrix $\mathbf{\Phi}_N$ represents a truncated set of normal modes computed from the internal DOFs of A

when all boundary nodes kept fixed. This procedure is only applicable provided that the sub-structures system matrices are symmetric and positive semidefinite. Considering the discrete undamped equation of motion for the sub-structure A , after substitution of \mathbf{q}_b^A the sub-system matrices \mathbf{M}^A , \mathbf{K}^A and the load mapping $\mathbf{B}^{A,\text{in}}$ are transformed (Eq. 4b) into a low-dimensional subspace. The reduced model is generated, when this method is applied to different sub-structures of the demonstration object, like the plate, the cover frame or the cavity, and these are assembled.

4.2 Moment Matching

The concept of the projection-based moment matching approach is to find a projection matrix \mathbf{V}_n so that the leading term of a Taylor series expansion of the transfer function matches for the reduced and the original system. An efficient method for engineering applications is moment matching via Krylov subspaces by means of either the Arnoldi or the Lanczos process. For this paper, a first order Krylov subspace was studied. Therefore The second order system can be converted into a descriptor first-order state space system of size $2N$:

$$\mathbf{C}\dot{\mathbf{x}} + \mathbf{G}\mathbf{x} = \mathbf{B}\mathbf{u} \quad (6a)$$

$$\mathbf{y} = \mathbf{L}\mathbf{x}. \quad (6b)$$

The state vector \mathbf{x} is the concatenation of the first and second time derivative $\mathbf{x} = (\dot{\mathbf{q}}\mathbf{q})^T$ and the matrices \mathbf{C} , \mathbf{G} , \mathbf{B} , and \mathbf{L} were assembled as follows:

$$\mathbf{C} := \begin{pmatrix} \mathbf{0} & \mathbf{X} \\ \mathbf{M} & \mathbf{D} \end{pmatrix}, \mathbf{G} := \begin{pmatrix} -\mathbf{X} & \mathbf{0} \\ \mathbf{0} & \mathbf{K} \end{pmatrix}, \mathbf{B} := \begin{pmatrix} \mathbf{0} \\ \mathbf{B}^{\text{in}} \end{pmatrix}, \mathbf{L} := \begin{pmatrix} \mathbf{B}_2^{\text{out}} & \mathbf{B}_1^{\text{out}} \end{pmatrix}, \quad (7)$$

where \mathbf{X} is an arbitrary regular matrix. In order to obtain symmetric matrices \mathbf{C} and \mathbf{G} the matrix \mathbf{X} is often set to $\mathbf{X} = -\mathbf{K}$ or $\mathbf{X} = \mathbf{M}$ in the case that \mathbf{M} and \mathbf{K} are symmetric. For this section the descriptor representation was chosen, because the matrices \mathbf{C} and \mathbf{G} are the input parameters of the utilized 1st order Krylov subspace method. The transfer function $\mathbf{H}(s)$ is developed by applying the Laplace transformation to (6) and eliminating the Laplace transform $\hat{\mathbf{x}}$ of \mathbf{x} which results:

$$\mathbf{H}(s) = \frac{\hat{\mathbf{y}}(s)}{\hat{\mathbf{u}}(s)} = \mathbf{L}(\mathbf{G} + s\mathbf{C})^{-1}\mathbf{B}. \quad (8)$$

The transfer function of the reduced-order model of size n that approximates the input-output behaviour of (8) is given by:

$$\mathbf{H}_n(s) = \mathbf{L}_n(\mathbf{G}_n + s\mathbf{C}_n)^{-1}\mathbf{B}_n. \quad (9)$$

The concept of the projection-based moment matching approach is to find a projection matrix \mathbf{V}_n so that the leading term of a Taylor series expansion of $\mathbf{H}(s)$ and $\mathbf{H}_n(s)$ are

matched. This can be done by means of a block Krylov-subspace method. Based on this, the transformation matrix \mathbf{V}_n may be obtained by execution of the Arnoldi process listed in Algorithm 1 presented in [11]. Therefore, the matrices of the reduced-order model are defined as follows:

$$\mathbf{G}_n := \mathbf{V}_n^T \mathbf{G} \mathbf{V}_n, \mathbf{C}_n := \mathbf{V}_n^T \mathbf{C} \mathbf{V}_n, \mathbf{B}_n := \mathbf{V}_n^T \mathbf{B}, \mathbf{L}_n := \mathbf{L} \mathbf{V}_n. \quad (10)$$

The aforementioned algorithm was implemented using the uBLAS C++ template class library and the parallel sparse direct solver MUMPS 4.8 in the MORAS software.

4.3 Modal reduction of non-symmetric systems

Classical modal reduction can be applied only to symmetric positive semidefinite systems with Rayleigh damping. Obviously, none of these properties is satisfied by the present system (3). Therefore, we have developed a generalized modal approach, which is based on the first order representation (6), projects the system on both, left and right eigenspaces, and accounts for higher order modes by static correction. More precisely, we construct reduced versions of (6):

$$\begin{aligned} \mathbf{C}_n \dot{\mathbf{x}} + \mathbf{G}_n \mathbf{x} &= \mathbf{B}_n \mathbf{u} \\ \mathbf{y} &= \mathbf{L}_{n,1} \mathbf{x} + \mathbf{L}_{n,2} \mathbf{u} \end{aligned} \quad (11)$$

using the following algorithm:

$$\begin{aligned} &\text{Choose shift } s_0 \text{ not being eigenvalue.} \\ &\text{Set } \mathbf{T} = -(\mathbf{G} + s_0 \mathbf{C}), \mathbf{F} = \mathbf{T}^{-1} \mathbf{B}, \mathbf{H} = \mathbf{T}^{-1} \mathbf{C}. \\ &\text{Compute incomplete Schur factorizations of desired size} \\ &\mathbf{H} \mathbf{V} = \mathbf{V} \mathbf{S}, \mathbf{H}^* \mathbf{W} = \mathbf{W} \tilde{\mathbf{S}}. \text{ With } \mathbf{J} = \mathbf{W}^* \mathbf{V} \text{ set} \\ &\mathbf{C}_n = \mathbf{J} \mathbf{S}, \mathbf{G}_n = -(\mathbf{J} + s_0 \mathbf{C}_n), \mathbf{B}_n = \mathbf{W}^* \mathbf{F}, \mathbf{L}_{n,1} = \mathbf{L} \mathbf{V} \\ &\mathbf{L}_{n,2} = \mathbf{L}_{n,1} \mathbf{J}^{-1} \mathbf{B}_n - \mathbf{L} \mathbf{F}. \end{aligned} \quad (12)$$

In the Fraunhofer model reduction toolbox (MRT) [3], we use ARPACK [13] and the LU-decomposition of $s_0^2 \mathbf{M} + s_0 \mathbf{D} + \mathbf{K}$ in order to compute the Schur factorizations in an efficient way.

5 NUMERICAL RESULTS

Generally, the CMS is applicable to acousto-mechanical problems [14] but not yet implemented in ANSYS. For this reason, the performance of the CMS, MORAS and MRT was compared when the FE model (11420 DOFs) of the pure mechanical plate is reduced to a 60 DOF first order system. The transfer functions (Fig. 4a) and the relative error computed against the full ANSYS model (Fig. 4b) are depicted in decibel scale. The best results for this model were achieved when using the Krylov subspace method (Fig. 4b).

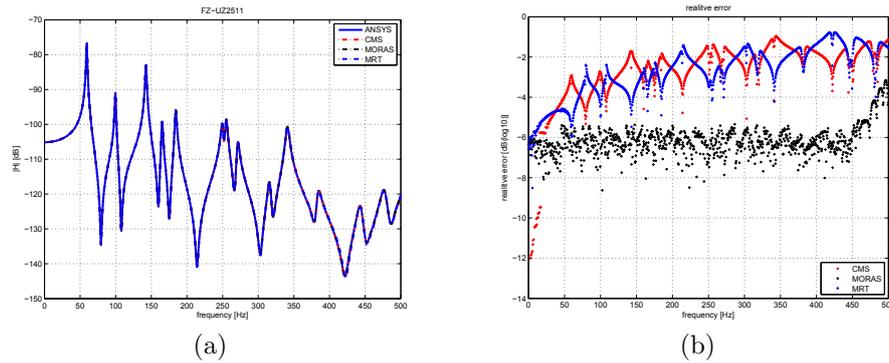


Figure 4: a) Transfer function of the mechanical model, b) error plot.

But considering the usual uncertainties of engineering applications, all MOR methods produce suitable results. Secondly, acousto-mechanical model (69020 DOFs) was reduced to 120 DOFs using moment matching via Krylov subspaces. In order to check the dynamic response the system was excited by a z-direction force in the lower right corner. The mechanical transfer function (Fig. 5a) and the acoustical transfer function (Fig. 5b) show a strong correlation up to 500 Hz, however when this method is applied the system becomes unstable, with poles appearing in the left plane of the pole plot (Fig. 5c). Proof

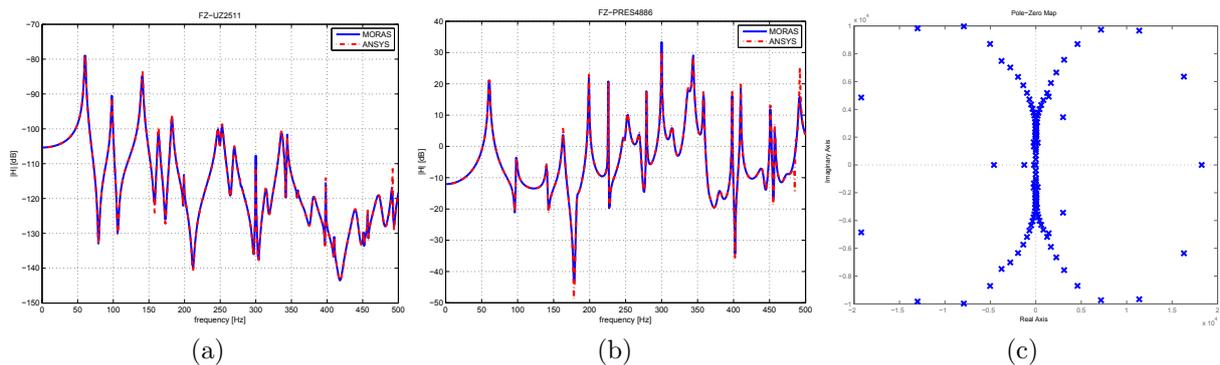


Figure 5: a), b) Transfer function of the acousto-mechanical model, c) pole map.

of this instability is presented by [11]. This paper proves that passivity and stability is guaranteed only if $\mathbf{G} + \mathbf{G}^T \geq 0$ and $\mathbf{C} = \mathbf{C}^T \geq 0$ are positive semidefinite and if matrix pencil $\mathbf{G} + s\mathbf{C}$ is regular. But the matrix \mathbf{C} never becomes symmetric because of the unsymmetric mass matrix \mathbf{M} [8] (Eq. 7 and Eq. 2). To overcome this issue, the modal reduction of non-symmetric systems was developed and this method was also applied to the model. The test used for the Krylov subspace method was executed using the Fraunhofer MRT code. The results are plotted in Fig. 6. The transfer functions are consistent up to 350 Hz and the systems remains stable (Fig. 6c). As a result, this model can be

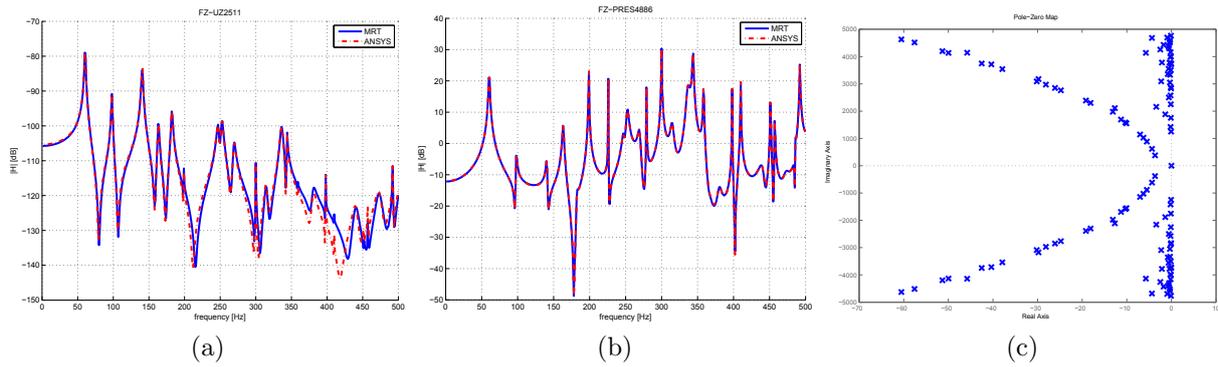


Figure 6: a), b) Transfer function of the acousto-mechanical model, c) pole map.

used for the controller implementation.

6 CONTROLLER IMPLEMENTATION

To show the capability of the reduced model for the development of active vibration control and active noise control, a system-level simulation was developed [8]. A FEM model with plate, cavity and piezo patches (98310 DOFs) was reduced to 60 DOFs and a six-mode positive position feedback (PPF) was implemented. The MATLAB/Simulink model, depicted in Fig. 7a, indicates the reduced system and the controller. The system was excited by a z-direction force impulse and the response recorded when the controller was switched on and off. With the active controller, the displacement response is decreased 7 dB, 8 dB, 14 dB, 3 dB and 4 dB for the 1st, the 2nd, the 3rd, the 4th and the 5th mode, respectively (Fig. 5a).

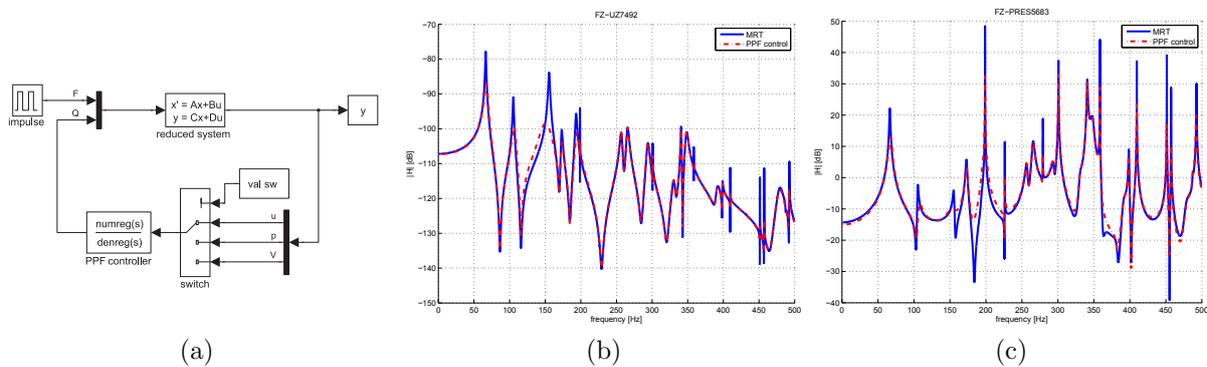


Figure 7: a) Simulink model, transfer function: b) force-displacement, c) force-pressure.

7 CONCLUSION

This paper has considered the MOR of the AdRIA acoustic demonstrator for design of active vibration control. Three different model order reduction techniques were evaluated, however only the modal reduction of non-symmetric systems was able to produce stable reduced models of the acoustic demonstration object. This model was used to set up a system-level simulation, which demonstrates the feasibility of this tool chain for the design of active vibration and active noise control.

This proposed framework enables researchers to efficiently model, simulate and study active structures, including acoustic cavities, with attached actuators and sensors. The MOR of the large FE model speeds up the simulations process, which helps to save significant time and costs.

The demonstration object used in this paper was covered by a plane plate. Currently, one-way and two-way curved shells covering acoustic cavities [5] are studied in order to minimize the of sound radiation of active structures. Therefore, parametric-reduced models of coupled mechanical, acoustic and electrical smart structure would be beneficial. Researcher [3, 15] have proposed promising approaches based on the interpolation of the system matrices. Research in the application of these methods to the demonstration object covered by curved shells will be the focus of future work.

8 ACKNOWLEDGMENT

The results presented in this study were developed within the framework of the LOEWE-Zentrum AdRIA (Adaptronic-Research, Innovation, Application), coordinated by Fraunhofer LBF and funded by the government of the German Federal State of Hesse. Their support is greatly acknowledged. We also thank Christoph Tamm, Oliver Janda and Necati Mercan for their constant support and for many helpful and stimulating discussions.

REFERENCES

- [1] H. Allik and T. Hughes, “Finite element method for piezoelectric vibration,” *International Journal for Numerical Methods in Engineering*, vol. 2(2), pp. 151–157, 1970.
- [2] A. C. Antoulas, *Approximation of Large-scale Dynamical Systems*. Advances in Design and Control, Society for Industrial & Applied Mathematics, U.S., 2005.
- [3] J. Mohring, A. Wirsén, J. Stoev, S. Lefteriu, M. Kurch, and N. Mercan, “Parametric model reduction of systems for active noise control,” in *6th International Congress on Sound and Vibration (ICSV16), Proceedings of*, 2009.
- [4] M. Kurch, H. Atzrodt, F. Kartzow, L. Schewe, and O. Janda, “On model order reduction for parameter optimisation of vibration absorbers,” in *Proceeding of the 10th*

International Conference on Recent Advances in Structural Dynamics, RASD2010, 12-14 July, 2010.

- [5] J. Tschesche, C. Thyges, J. Bös, H. Hanselka, and S. Herold, “Application and analysis of structural acoustic optimization techniques in smart structure product development and systems engineering,” in *8th ASMO UK/ISSMO Conference on Engineering Design Optimization*, Proceedings, Juli 2010.
- [6] ANSYS, Inc., *Theory Reference for the Mechanical APDL and Mechanical Applications, Release 12.1.*
- [7] “9. milestone report ’validiertes lastenheft für die technologiebereiche’,” tech. rep., LOEWE-Zentrum AdRIA (Adaptronic-Research, Innovation, Application), 2010. (in German).
- [8] C. Tamm, “A comparison of model order reduction methods for vibro-acoustic systems with piezoelectric transducers,” Master’s thesis, Technische Universität Darmstadt, 2010. (in German).
- [9] R. J. Allemang, “The modal assurance criterion – twenty years of use and abuse,” *The Noise and Vibration Control Magazine*, vol. 8, pp. 14 – 21, 2003.
- [10] R. J. Craig and M. Bampton, “Coupling of substructures for dynamic analysis,” *AIAA Journal*, vol. Vol. 6, no. 7, pp. pp. 1313–1319, 1968.
- [11] R. W. Freund, “Krylov-subspace methods for reduced-order modeling in circuit simulation,” *Journal of Computational and Applied Mathematics*, vol. 123, no. 1-2, pp. 395–421, 2000.
- [12] W. Hurty, “Dynamic analysis of structural systems using component modes,” *AIAA Journal*, vol. 3, no. 4, pp. 678–685, 1965.
- [13] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK Users Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*. SIAM, 1998.
- [14] J. Herrmann, M. Maess, and L. Gaul, “Efficient substructuring techniques for the investigation of fluid-filled piping systems,” in *Proceedings of IMAC XXVII, Orlando, FL (USA)*, 2009.
- [15] H. Panzer, J. Mohring, R. Eid, and B. Lohmann, “Parametric model order reduction by matrix interpolation,” *at - Automatisierungstechnik*, vol. 58, pp. 475–484, Aug. 2010.

SENSOR NETWORK OPTIMIZATION FOR DAMAGE DETECTION ON ALUMINIUM STIFFENED HELICOPTER PANELS

C. SBARUFATTI^{*}, A. MANES[§] AND M. GIGLIO[†]

^{*} Politecnico di Milano (Dipartimento di Meccanica)
Via la Masa 1, 20156 Milano, Italy

E-mail: claudio.sbarufatti@mail.polimi.it - Web page: <http://www.mecc.polimi.it>

[§] Politecnico di Milano (Dipartimento di Meccanica)
Via la Masa 1, 20156 Milano, Italy

E-mail: andrea.manes@polimi.it - Web page: <http://www.mecc.polimi.it>

[†] Politecnico di Milano (Dipartimento di Meccanica)
Via la Masa 1, 20156 Milano, Italy

E-mail: marco.giglio@polimi.it - Web page: <http://www.mecc.polimi.it>

Key words: Artificial Neural Network, network optimization, structural diagnosis, crack, panel.

Abstract. Health and Usage Monitoring Systems (HUMS) has received considerable attention from the helicopter community in recent years with the declared aim to increase flight safety, increase mission reliability, extend duration of life limited components and of course reduce the maintenance costs. The latter is about 25 per cent of the direct operating cost of the helicopter, thus playing an important role especially in the case of the ageing aircrafts. In particular, with respect to helicopter fuselages, only some attempts were carried out to monitor directly on-line the damage accumulation and propagation during life. In this field, and in particular in the military applications, an integrated and reliable system for monitoring the damage in the fuselage and for evaluating the time inspections and remaining life (prognosis) is missing. However, because of the presence of many vibratory loads, the diagnosis of helicopter structures is very critical. From one hand, a very large number of sensors would be needed for a robust appreciation of the structural health, from the other hand the industrialization of the product brings the need for a low impact over the existing structures, or toward a reduction in the allowed amount of sensors. As a result, comes the importance for an optimization of the sensor network, with the aim to find out the regions inside the structure which are the most sensible to a damage and at the same time robust to noise. The aim of the present work is to define a methodology for optimising the sensors position inside an helicopter fuselage panel in order to obtain the best compromise between the simplicity and the robustness of a sensor network. In particular, a Finite Element (FE) model will be used to create a database of various damages inside the structure, thus consequently optimising the network sensitivity to any damage. The evaluation of the network performances is provided when some realistic noise [1,2] is added to the FE calculation.

1 INTRODUCTION

Structural Health Monitoring has received a lot of attention from the aerospace industry, where the costs related to maintenance and machine stops are very high and safety is a primary issue. In particular, vibrations play an important role in helicopters, where structural defects like cracks could propagate very fast causing unexpected failures of components. To date, fatigue behavior is controlled with a deep analysis in the design phase, governed by damage tolerant regulations, and with a clear schedule of inspections during life, which reduce a lot the availability of the machines, with a cost that is around the 25% of the operating effort.

The coupling of Finite Element Methods (FEM) with Artificial Neural Networks (ANN) has proved to be promising in the frame of Structural Health Monitoring applications, concerning aerospace industries as well as civil structural ones. The main advantage is that FEM allows for a low cost knowledge generation, upon which it is possible to optimize the ANN parameters. For a deeper introduction to the two tools, the interested reader could refer to [3,4]. Though many parameterized damages could be modeled with FE, it is however important to keep in mind that some test experiments are needed in order to validate the models as well as to calibrate the damage sensitive parameters and to appreciate the best direction for a network optimization. For instance, in [1] some data were presented about dynamic crack propagation tests in helicopter fuselage panels, and it has been possible to appreciate the extent of real world dispersion with respect to FEM crack propagation calculations, giving it as input to the current work. In fact, to deliver a critical product such as an SHM system, the most important point is to associate a reliability and availability to it, thus calculating the feasibility either from the economic or safety points of view. In fact, especially if the system has to interact with the pilot to suggest a certain flight profile according to the structural health of the machine, the False Alarm parameter should be reduced at a minimum. Different requirements arise if the system has to communicate with the maintenance center, where measured data are stored and processed by maintenance engineers in order to set up the proper maintenance schedule, concept at the basis of the Condition Based Maintenance (CBM). The former approach is by far more critical, thus requiring also higher performances, especially in terms of reliability.

Inside this work, the training data are provided by FE analysis, having the advantage of allowing a large range of damage parameters to be analyzed. However, if in FE analysis there is no limit on the spatial resolution of the data which is obtained (either strains or mode shapes), in reality the number of sensors available will be limited and this will create restrictions on data resolution. As a result, it would be necessary to optimize the number and location of sensors for each given problem.

The sensor network should thus be optimized in order to maximize the efficiency of the SHM, compromising between the increasing cost of a dense network and the lowering of performances by decreasing the number of sensors. An approach for sensor network optimization is proposed inside this paper, trying to find the best sensor position and number for a stiffened helicopter panel SHM. The aim is to estimate the amount of sensors necessary to obtain a certain performance, maybe required from regulations or from economic analysis. Given the repetitive structure of the typical aeronautical stiffened panel analyzed, the attention is not focused on the optimization algorithm but on performance results. However, the

methodology is valid also for more complex cases where symmetries and repetitions couldn't be exploited, for which the coupling with more complex optimization techniques like those reported in literature would be useful. A detailed survey of recent works on sensor placement is given in [5], focusing the attention on combinatorial optimization, such as Genetic Algorithms (GA) and Simulated Annealing (SA), which proved to be particularly efficient. Again, in [6] the information entropy parameter is proposed as a measure upon which to minimize the uncertainties in the estimation of model parameters, however mostly suited for a Bayesian Network. Furthermore, in [7] the optimization of optical Fiber Bragg Grating sensor locations inside a composite repair patch is proposed, with the aim to determine the laminates between which the fibers should be embedded as well as the actual crack sensitivity of the sensor.

2 THE DIAGNOSTIC SYSTEM

In figure 1 the organization of the diagnostic system is reported. As suggested in [5] and reported in [8], one could distinguish many steps or levels, each one inferring over a particular aspect of structural diagnostic. In particular, [5] recommends 4 layers, named Anomaly Detection (alarm generation), Localization, Quantification and Prognosis. However, an additional knowledge layer is needed, that is to say Damage Recognition, or the ability of the algorithm to distinguish the correct damage the structure undergoes. This is mostly important for the Quantification phase, as a wrong estimation of the damage type could lead to wrong assumption in Layer 3, 4 and 5. In particular, as introduced before, helicopters are critical machines because of fatigue (connected with crack damage type) as well as low velocity impacts due to the harsh environment where they operate (impact damage type). Moreover, military helicopters are subjected to bullet impacts, once more connected with a different damage pattern.

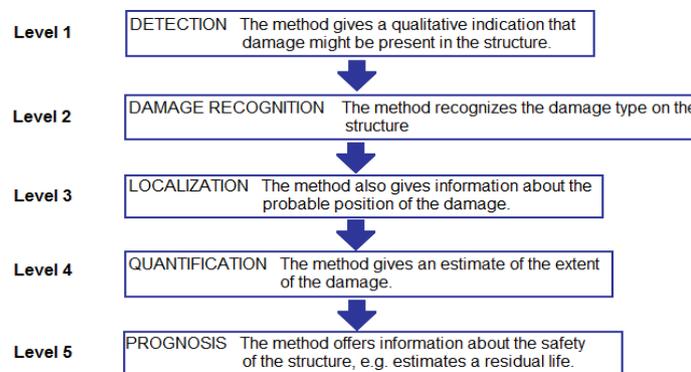


Figure 1: Flow Chart representing the organization of information inside the diagnostic system

An important thing to be decided is whether to unify two or more levels of diagnostic inside just one ANN with multiple output. According to some ANN performance analysis, using a separate ANN to infer over each single diagnostic layer appears to enhance the performances. This can be explained through the assumption that, when optimizing the ANN to solve one layer, each synapses weight is calibrated for that layer information. Obviously, it

is important to remember that, given that the performance is measured through the percentage of right assumptions, the reliability of the coupled system has to be compared with the product of the uncoupled reliabilities.

3 PROBLEM DEFINITION

A FE model of a stiffened panel has been created in ABAQUS 6.9, consisting of a typical aerospace structure, referring in particular to helicopter fuselage design. Some geometry information are reported in figure 2, while for a detailed explanation of the model the interested reader could refer to [8]. Two crack damages have been modeled, one consisting of a rivet crack, starting from one of the rivet connections inside the panel, while the other consisting in a bay crack, randomly positioned in a bay to simulate a crack propagating from an accidental damage. A summary of the problem statement is reported in table 1.

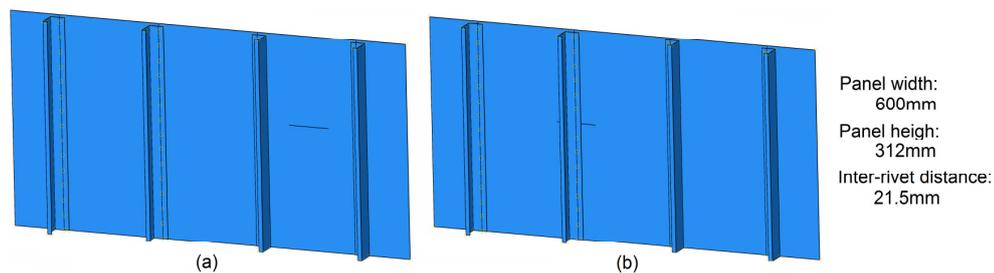


Figure 2: Panel geometry, with model for (a) bay skin crack and (b) rivet skin crack

Table 1: Summary of problem statement

Item	Description	Notes
Crack Types	Rivet Crack	The most common case
	Bay Crack	Simulating crack starting from accidental damage
Crack Length (for ANN training)	Constant, 60 mm	Inside this paper, a constant crack length is chosen to train the ANN, then evaluating the performances for different crack length.
Crack Angle	Constant, perpendicular to stringers	The loading on the panel is directed parallel to stringers
Crack Cases for ANN	80 cases for each bay	Total $80 \times 3 = 240$ cracks on the panel
	1 case for each rivet	Total $15 \times 4 = 60$ cracks
Allowed Sensor Positions	Skin	Along a path in the centre of the bay parallel to stringers
	Stringer	Along a path on the stringer next to rivets parallel to stringers
Sensor constraints	Skin/Stringer	The number of sensor is constant among all the stringers. The number of sensor is constant among all the bays.

The sensor network design will be based hereafter upon the optimization for Level 1 and 2 diagnostics. In fact, according to figure 1, Level 1 information appears to be the most critical point, as it influences the activation of all the consecutive level checks. In fact, the practical

implementation of the on-board system should rely on the alarm generator, which has to produce the least number of False Alarms (damage doesn't exist but it is detected) and Missed Event (damage exists but it is not detected). Level 1 decision will be modeled by means of a *Pattern Recognition ANN* algorithm, inside the MATLAB environment. Only when the anomaly is detected, all the consecutive steps should be activated. In particular, the need for a damage type discrimination has been advanced above. Inside the present work, the Level 2 check has to distinguish between 2 damage types that are practically similar, as they both refer to a skin crack. However, the system feasibility will be demonstrated. For a more practical and useful solution one could try to model the rupture of a stringer, thus discerning it with respect to skin crack damages. Again, Level 2 decision will be modeled by means of a *Pattern Recognition ANN* algorithm, inside the MATLAB environment.

4.1 Performance indicators

It is important to define the performance indicators of the ANN. For this purpose, 3 parameters have been identified as the most indicative, that is to say, the Probability of damage Detection (PoD), the Probability of False Alarms (PFA) and the Probability of Wrong Assumptions (PWA):

$$PoD = \frac{Detected_cracks}{Total_crack_cases} \quad (1)$$

$$PME = 1 - PoD \quad (2)$$

$$PFA = \frac{Correct_alarms}{Generated_alarms} \quad (3)$$

$$PWA = \frac{Wrong_Decisions}{Total_Nr._of_analysed_cases} \quad (4)$$

In particular, the parameter defined in Eq. (3) is able to synthesize the information gained through PoD and PFA. For that reason, PWA will be used in the next chapters to reason about the sensor number selection.

4.2 Noise extent appreciation

As introduced in Ch. 1, during the training phase, FE calculations for strains over the structure will be given as input to the *Pattern Recognition ANN*, while the output will be a binary variable indicating whether the damage is present or not. However, it is important to consider that, during the algorithm testing or, more generally, while using the properly weighted algorithm in a real environment, the input data will be noised, at least hopefully with a mean value near to the FEM prediction. In [1] some propagation tests were performed over the same structure geometry, thus appreciating the extent of the variability of strain measurements over the structure in function of the crack length (all the cracks were artificially initiated at the same location in a typical panel geometry).

The main output from [1] was that two classes of sensor positions could be established,

that is to say sensors on skin or on stringers, with sensing direction parallel to maximum principal stresses. The sensors located on the skin presented a larger deviation from FEM with respect to those placed on the stringers. On the other hand, skin appeared as more sensitive to crack presence, which means that the strain parameter variation per unit of crack length increase was larger with respect to what happened for stringer sensor. However, being the noise robustness more influent with respect to the losses in sensitivity, stringer located sensors behaved better than skin sensors in crack hypothesis testing. The extreme variability of measures coming from skin sensors might be due to compressive stresses perpendicular to stringers generated in the bay skin, similar to buckling effects, very case to case dependent. On the other hand, stringers are designed to transmit a the load in a certain direction, thus acting as a filter with respect to noises coming from all the other directions. In table 2 is reported the extent of noise percentage superposed to FEM calculated inputs during the ANN TESTING phase.

Another point to be discussed is whether to introduce or not the noise also in the ANN TRAINING phase. In fact, the proper ANN optimization would stop before starting to learn the noise coming from the input data, thus maintaining the property of generalization. However, if the network training is done with deterministic data (FEM based), practically the algorithm doesn't recognize that the same defect could produce a certain distribution of possible strain outputs, instead of just one value. The proposal inside the current work is to introduce a certain amount of noise also in the training phase of the neural network. This has been done by creating 25 replies for each strain measure associated to each damage case, sampling from a Gaussian distribution with mean equal to the FEM prediction and variance connected to the percentage variation expressed in table 2 for the training phase (the choice for the noise model has to be carefully justified, as explained later on). In particular, the performance evaluation for sensor network modeling has been repeated for 3 noise amounts applied in the training phase, as reported again in table 2.

Table 2: Noise Estimation for ANN training and testing

Training Phase	Sensor on Skin	0%, 6%, 10%
	Sensor on Stringer	0%, 4%, 8%
Testing Phase	Sensor on Skin	15%
	Sensor on Stringer	7.5%

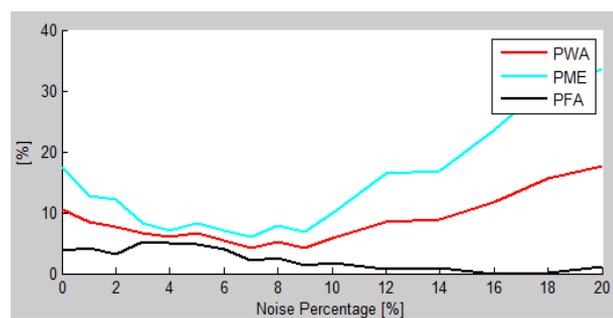


Figure 3: Probability of False Alarm (PFA), Probability of Missing Event (PME) and Probability of total Wrong Decision as a function of noise added in the ANN training

The noise amount to be introduced in the training phase should be estimated upon appreciation of the actual noise present on the real case structure (Testing Phase Noise). In figure 3 the performance variation is reported in function of the noise added to FE data in the training phase. It can be noticed that a beneficial effect in performances is obtained by adding noise up to a certain value, below which the algorithm starts to decrease its detection capabilities.

Finally, the choice for a Gaussian distribution for noise modeling has to be justified according to the assumption that the causes of that noise are also Normal distributed. The following are the main reasons for real data dispersion with respect to FEM prediction on strain measures:

- Manufacturing process
- Material non-uniformities
- Non Linearity (different from damage)
- Environmental influences
- Sensor locations
- Crack angles

Under the hypothesis of material uniformity and linearity, allowing a 1-2% variation in material properties, it will be reflected in the same percentage uncertainties on the FEM prediction. In figure 4 it is possible to appreciate the effect on the strain FEM outputs due to a $\pm 5.5^\circ$ variation in crack angle (corresponding to $\pm 5\%$ of the 90° target angle with respect to stringer direction). In case of perfect Gaussian error propagation, the lines relative to $\pm 5.5^\circ$ angle variation should be mirrored with respect to 0% difference horizontal line. This is not happening which means that, though the angle dispersion could be Normal distributed, its noise propagation at sensor measure level won't still be Normal distributed. However, it rarely exceeded the 2% threshold variation in strain field output.

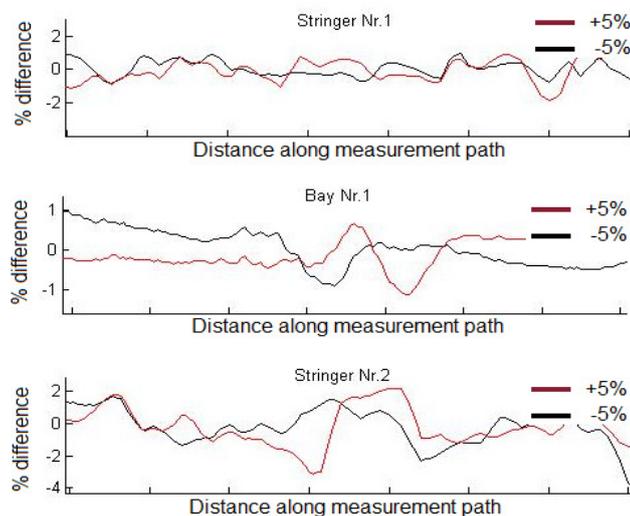


Figure 4: Uncertainty in FEM strain measures estimation for $\pm 5.5^\circ$ variation of crack angle w.r.t. nominal crack angle (90°)

5 RESULTS AND CONSIDERATIONS

According to the problem statement discussed above, some sensitivity plots have been produced in order to evaluate the performances of the ANN in function of the PWA parameter, as defined in Ch. 4.2. For each sensor number combination inside the plots, the ANN (pattern recognition) performance has been evaluated trying 5 different hidden layer numbers and repeating the network training 4 times, in order to select the best trial. This because, by varying the number of hidden layers, the structure of the ANN is changed, and thus its performances. Moreover, the optimization process itself isn't deterministic, thus producing a different ANN for each training trial. For a better understanding on the structure of ANN the interested reader could refer to [3,8].

5.1 Case 1 (a, b)

Problem	Level 1: PATTERN RECOGNITION (Distinguish damaged from undamaged panel)		
Noise in training set	Sensor on Skin (a,b,c)	0%	10%
	Sensor on Stringer (a,b,c)	0%	8%
Noise in testing set	Sensor on Skin	15%	
	Sensor on Stringer	7.5%	
Crack Length for training	60mm		
Crack Length for testing	60mm		
Crack Angle	Perpendicular to stringer		
Crack Type	Either rivet crack or bay crack		

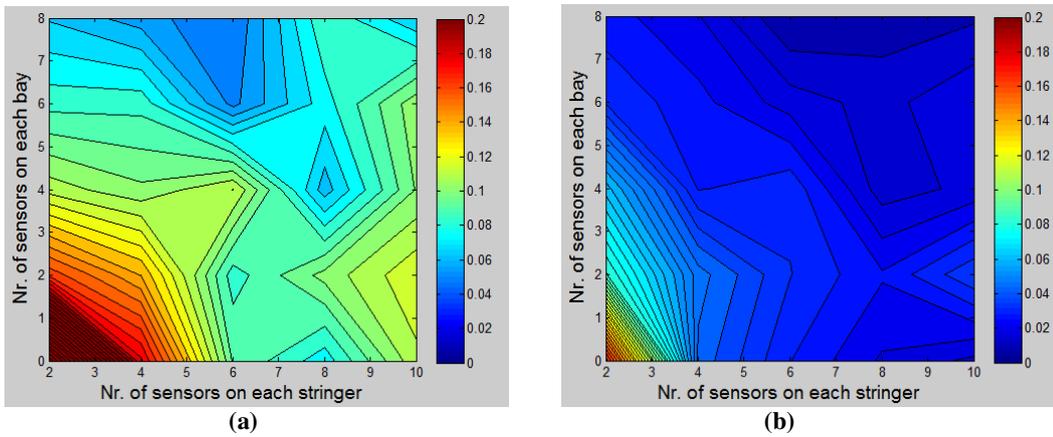


Figure 5: Case 1 sensitivity plots for PWA with (a) no noise in ANN training (b) noised training set

5.2 Case 2 (a, b)

Problem	Level 1: PATTERN RECOGNITION (Distinguish damaged from undamaged panel)		
Noise in training set	Sensor on Skin (a,b,c)	0%	10%
	Sensor on Stringer (a,b,c)	0%	8%
Noise in testing set	Sensor on Skin	15%	
	Sensor on Stringer	7.5%	
Crack Length for training	60mm		
Crack Length for testing	100mm		
Crack Angle	Perpendicular to stringer		
Crack Type	Either rivet crack or bay crack		

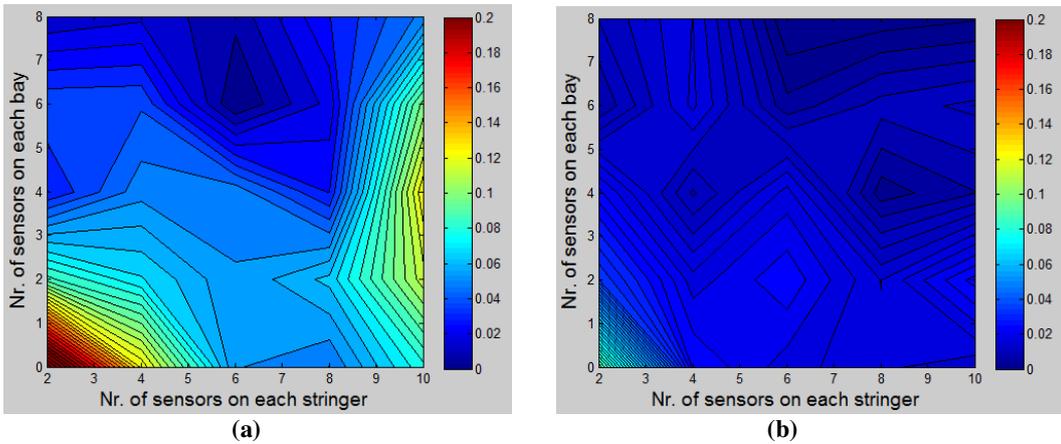


Figure 6: Case 2 sensitivity plots for PWA with (a) no noise in ANN training (b) noised training set

5.3 Case 3 (a, b)

Problem	Level 1: PATTERN RECOGNITION (Distinguish damaged from undamaged panel)		
Noise in training set	Sensor on Skin (a,b,c)	0%	10%
	Sensor on Stringer (a,b,c)	0%	8%
Noise in testing set	Sensor on Skin	15%	
	Sensor on Stringer	7.5%	
Crack Length for training	60mm		
Crack Length for testing	40mm		
Crack Angle	Perpendicular to stringer		
Crack Type	Either rivet crack or bay crack		

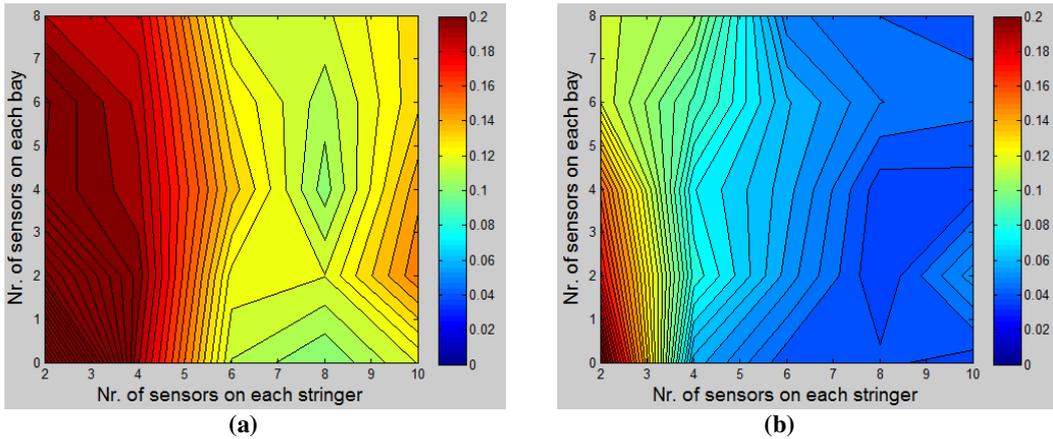


Figure 7: Case 3 sensitivity plots for PWA with (a) no noise in ANN training (b) noised training set

5.4 Case 4 (a, b)

Problem	Level 2: PATTERN RECOGNITION (Distinguish skin crack from rivet crack)		
Noise in training set	Sensor on Skin (a,b,c)	0%	10%
	Sensor on Stringer (a,b,c)	0%	8%
Noise in testing set	Sensor on Skin	15%	
	Sensor on Stringer	7.5%	
Crack Length for training	60mm		
Crack Length for testing	60mm		
Crack Angle	Perpendicular to stringer		
Crack Type	Either rivet crack or bay crack		

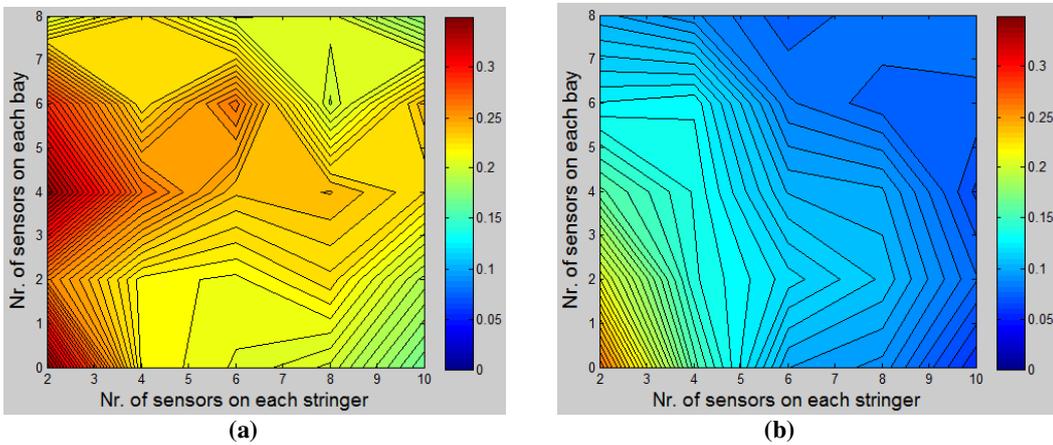


Figure 8: Case 4 sensitivity plots for PWA with (a) no noise in ANN training (b) noised training set

5.5 Considerations

In figure 5 the ANN performances in terms of PWA are plotted as a function of the sensor combination over the structure. The first thing that should be noticed is the improvement in detection capabilities because of the introduction of a certain level of noise inside the training phase (figure 5b). In addition one could notice that, while for a lower number of stringer sensors it could be worth to acquire some skin measurements, when increasing the sensor number on the stringer it becomes practically negligible the additional information gained with skin bay measures. On the other hand, the information gained by adding one more stringer sensor is higher with respect to that of a skin bay sensor, maybe due to the noise model, however in agreement with experimental test conducted in [1, 2].

If results reported in figure 5 refer to a case where crack length for training (60 mm) coincide with that for testing (60 mm), in figure 6 and 7 the testing crack length has been respectively moved to 100 mm and 40 mm. As expected, the detection performances got worse when detecting cracks smaller than those for training and rather better when moving in the opposite direction. A similar information can be retrieved from figure 9, where the changes in ANN capabilities as a function of sensor number and crack length are clearly appreciable for the cases with no sensors on the skin but varying sensors on the stringers (figure 9a) and 6 sensors on each stringers and varying number of skin sensors (figure 9b). It is however demonstrated (at least for stiffened panels) that designing a training set with a threshold crack length allows to detect also longer cracks with a reliability at least equal to the design point.

Finally, in figure 8 the capabilities on damage recognition for Level 2 diagnostic is reported as a function of sensor number combinations. Again it is clear that the information gained with a stringer sensor is by far more effective than the one of a skin located sensor. It is also possible to appreciate that sometimes adding additional sensors on the skin could worsen the ANN performances.

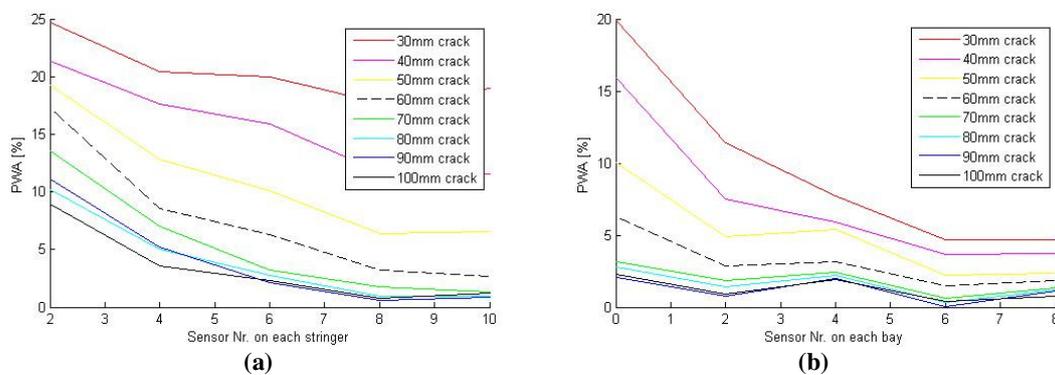


Figure 9: (a) PWA as a function of sensor Nr. on each stringer (with no skin sensors) for different crack length in testing phase and (b) PWA as a function of sensor Nr. on each bay (with 6 sensors on each stringer) for different crack length in testing phase. Training has been done with 60mm cracks (dashed line)

6 CONCLUSIONS

According to some preliminary studies on helicopter fuselage damage tolerant design [1, 2], it has been possible to appreciate the extent of variability in strain measures for stiffened panel like structures subjected to growing damages. The information retrieved has been used inside this work to enhance the performances of an ANN, thus selecting the best sensor arrangement in terms of numbering and position, according to the capabilities of different configurations, expressed as PWA, PME and PFA. It was demonstrated the performance increase due to the superposition of even a small noise to the deterministic training set coming from FE simulations. Moreover, it was confirmed [1] the best performance of a stringer sensor with respect to skin bay measure points. Finally, the configuration suggested for the current case test is a 6-sensor-per-stringer/320mm-stringer. The choice can be taken as a good compromise between the additional information that could come from one more sensor in stringer/skin and the feasibility of the SHM hardware. However, according to the new sensor technologies based upon Fibre Bragg Grating (FBG) systems, even more sensors could be distributed along the stringer direction thanks to the possibility for multiplexing.

REFERENCES

- [1] C. Sbarufatti, A. Manes and M. Giglio, *Probability of detection and false alarms for metallic aerospace panel health monitoring*. Proc. 7th Int. Conf. on CM & MFPT, BINDT, (2010).
- [2] M. Giglio and A. Manes, *Crack propagation on helicopter panel: experimental test and analysis*. Engineering fracture mechanics, Vol. 75, pp. 866-879, (2008).
- [3] C. Bishop, *Neural Networks and Pattern Recognition*, Oxford University Press, (1995)
- [4] O.C. Zienkiewicz, R.L. Taylor and J.Z. Zhu, *The finite Element Method: Its Basis and Fundamentals*, 6th Ed., Butterworth-Heinemann
- [5] K. Worden and A.P. Burrows, *Optimal sensor placement for fault detection*, Engineering Structures 23 (2001) 885–901
- [6] C. Papadimitriou, *Optimal sensor placement methodology for parametric identification of structural systems*, Journal of Sound and Vibration 278 (2004) 923–947
- [7] G. Tsamasphyros, N. Furnarakis, G. Kanderakis and Z. Marioli-Riga, *Optimization of Embedded Optical Sensor Location in Composite Repairs*, Applied Composite Materials **10**: 129–140, 2003.
- [8] C. Sbarufatti, A. Manes and M. Giglio, *Sensor network optimization for damage detection on aluminum stiffened helicopter panels*, Proc. Coupled Problems in Science and Engineering (Coupled 2011)

SOLVING THE DILEMMA OF CONTRADICTORY GOALS

OLAF FROMMANN*

* Institute for Aerospace Technology
University of Applied Sciences Bremen
Flughafenallee 10, 28199 Bremen, Germany
e-mail: olaf.frommann@hs-bremen.de

Key words: Multidisciplinary / Multi-criteria Problems, Optimization, Objective Function, Fuzzy Logic, Pareto Front, Contradictory Goals.

Abstract. Coupled problems in engineering inevitably lead to contradictory goals for single quality criteria. Applying numerical optimization to find the best solution requires the definition of an objective function based upon these criteria as a measure of merit for the whole design. It will be shown that the usual approach of a weighted sum is improper and unreliable. Also the consideration of the complete Pareto front is not only no solution to this problem, it usually requires a not maintainable amount of calculations. Therefore, a new approach for the definition of objective functions is proposed, which solves the dilemma of contradictory goals by finding an appropriate compromise based upon the intuition.

1 INTRODUCTION

Coupled problems, and multidisciplinary optimization is one of them, often lead to contradictions with respect to the goals. These contradictions do not only occur in multidisciplinary cases, e.g. low aerodynamic drag and a wing as thick as possible for an aircraft, but also in simpler cases like multipoint designs. This results in the problem to obtain an overall appropriate measure of merit for the design, instantiated by a so called objective function. Usually, this function is set up as a weighted sum of all conceived criteria. While this approach seems to be simple and obvious, it comprises a severe amount of arbitrariness concerning the choice of usually applied weighting factors. Furthermore, if constraints for some criteria have to be considered, assuming the application of optimization algorithms which are not capable to handle them directly, this is only possible by the introduction of so called penalty functions in order to worsen the objective function value artificially in case of the violation of constraints. These penalty functions are usually quadratic, i.e. nonlinear, with the effect of an unforeseeable distortion of the solution space topography. After all, it can additionally be shown that not all possible combination of optimal criteria is possible with this approach.

An alternative approach is the consideration of the so called Pareto front. This constitutes the non-dominating solutions, meaning that the improvement of one criterion leads to a degradation of the other. Historically, this is the usual way considering criteria when evolutionary or genetic algorithms are applied, because the huge amount of objective function evaluations simply provides this front. On the other hand, there exists no distinct solution.

Users need to choose one solution out of all Pareto optimal ones, leading again to some kind of arbitrariness. Furthermore, if this may be possible in the case of only two criteria, it will become impossible for more than three, because the Pareto front becomes a hyper-dimensional surface, which cannot be imagined by humans. To handle this problem, so called decision making tools respectively algorithms have been developed, but there exists no prove for their correctness or applicability. Therefore, a new approach for an objective function, based upon the principles of Fuzzy Logic [4], is proposed, which simplifies the decision finding process significantly and leading to a distinct measure of merit for each solution. In contrast to the mentioned decision finding tools, this constitutes a real objective function.

In the following, starting with a simple example to demonstrate the difficulty in setting up an appropriate objective function based upon weighted sums, and to understand the meaning of the solution space topography, the proposed new approach will be discussed in relation to the above mentioned current procedures. Finally, a practical application to an eight criteria optimization problem will show the simplicity and advantages compared to using weighted sums.

2 SOLUTION SPACE TOPOGRAPHIES

Numerical Optimization algorithms usually need some kind of exact measure for each single solution in order to vary the design variables for an improvement. This is accomplished by setting up an objective function. However, a function means defining a surface in hyper-dimensional space. Having two criteria leads to a surface in three dimensions like a landscape. Within this landscape one can find the highest and lowest points, which are the optima. This means, by defining an objective function, the location of the optimum is also predefined, although unknown a priori. The following simple example shows this relation for only two criteria, making it possible to visualize the landscape. Having more than two criteria worsens the problem and leads to unsolvable arbitrariness.

The basis of this optimization problem is a beam with length L shown in Fig. 1, loaded by a force P and having the design parameters width w and height h . The goal is to maximize the benefit in terms of the usable length L while minimizing the costs represented by the volume V , calculated from equation (1). Assuming some kind of allowable stress $\sigma_{allowed}$ of the material, which should not be exceeded by the maximum stress occurring at the left mounting location, the resulting maximal length can be calculated based upon the geometric moment of inertia from equation (2).

$$V = L w h \quad (1)$$

$$L = \sigma_{allowed} \cdot w \cdot h^2 / (6 \cdot P) \quad (2)$$

$$\min f = w_1 V - w_2 L \quad (3)$$

The corresponding objective function, built as a weighted sum with the weighting factors w_1 and w_2 , is shown in equation (3). However, it is unknown which value these factors should have. In order to demonstrate the influence of this choice, three slightly different combinations, shown in Tab. 1, were used to calculate the solution space from a full factorial search. The corresponding topographies are depicted in Fig. 2.

Table 1: Weighting Combinations for the Beam Objective Function

	Case 1	Case 2	Case 3
V	50%	60%	40%
L	50%	40%	60%

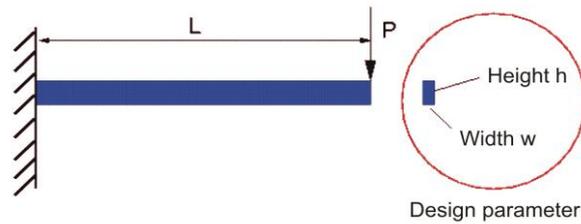


Figure 1: Beam optimization example

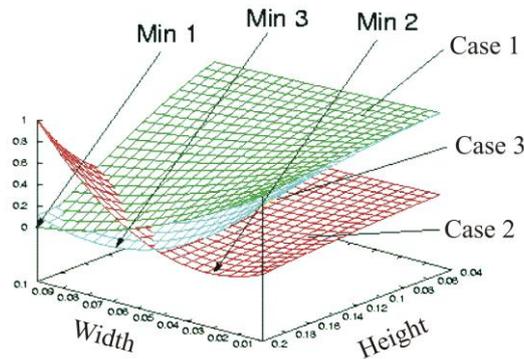


Figure 2: Beam solution spaces for different weighting factors

Different objective functions with varying weighting factors, more or less arbitrarily chosen, obviously lead to different results. Changing the weighting factors respectively the objective function means, changing the shape of the solution space, predetermining the possible optima with the result of different solutions. While this example is very simple, in practice the topographies look much more complex and non-smooth like in Fig. 3. Even in this two-dimensional case for the aerodynamic drag of an airfoil, calculated by an iterative code based upon the full potential theory, for the design parameters camber and thickness, it is obvious, that finding an optimum is not trivial within this landscape. The plateau and the peaks are due to not converging solutions, something that is very likely to occur during an optimization.

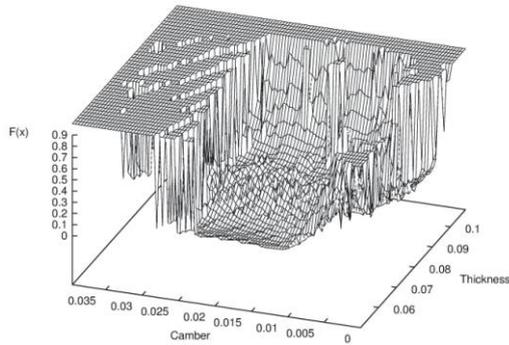


Figure 3: Solution space topography for the aerodynamic drag of an airfoil, depending on camber and thickness, calculated by a code based upon the full potential theory

3 PARETO FRONTS AND WEIGHTED SUM OBJECTIVE FUNCTIONS

In the case of multi-objective optimization no single objective can be maximized or minimized without the compromise of deteriorating another. This is known as the so called Pareto front, which includes all these solutions. To get a hands-on example, the following two objectives may serve. For this simple example we suppose, that one objective will be minimized with a decreasing design parameter, while the other one improves with an increasing one. A possibility for such a scenario are the functions from equation (4) and (5) within the interval of [0:1], Fig. 4.

$$f_1 = 1 - 2x + x^2 \tag{4}$$

$$f_2 = \sqrt{x} \tag{5}$$

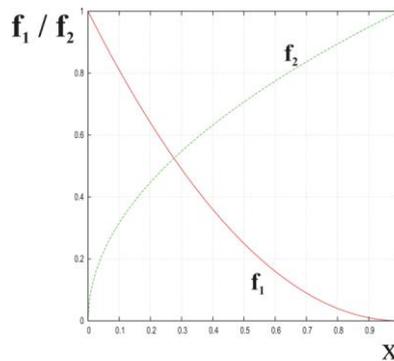


Figure 4: Competing objectives from equation (4) and (5) within the interval [0:1]

Plotting function f_2 depending on f_1 yields the Pareto front, Fig. 5, i.e. all possible compromise solutions, which in this case exist of a convex and a concave area. Now we assume to find one of these compromises by the application of a weighted sum, equation (6). Trying to incorporate this into Fig. 5 requires transposing it to equation (7). This is a linear

equation with a negative gradient of w_1/w_2 . Fig. 6 shows several lines with different slopes. A minimization of f means a parallel translation to the left. An equal weighting leads to the bottom right point of the front, while choosing $w_1 > w_2$ some points within the convex region can be reached. The boundary in the right picture of Fig. 6 is found for $f_1 = 1/9 = 0.111$, while the convex part ends at $f_1 = 0.444$ where the curvature changes. This means, no point on the concave part of this Pareto front and also many points on the convex part, overall about 75%, are reachable by this kind of objective function.

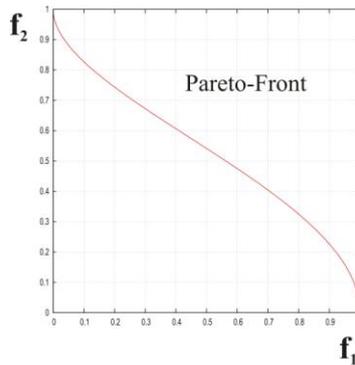


Figure 5: Pareto front for the objectives from equation (4) and (5) within the interval [0:1]

$$\min f = w_1 f_1 + w_2 f_2 \tag{6}$$

$$f_2 = f/w_2 - f_1 w_1/w_2 \tag{7}$$

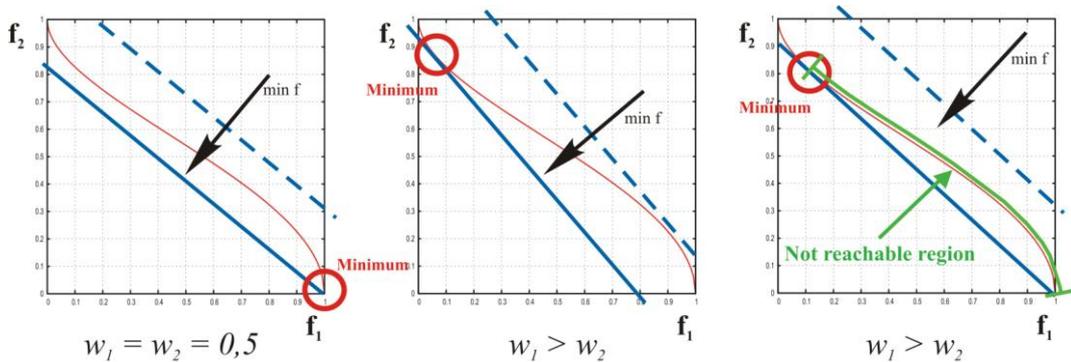


Figure 6: Pareto front compared to weighted sum objective functions

Therefore, weighted sums cannot be an appropriate approach for the description of optimization goals. Nobody ever can tell, whether the hyper-dimensional Pareto front consists of concave parts, which may be missed. Thus, another way of defining optimization goals is necessary.

4 TRANSFORMING APPRAISALS INTO DISTINCT MEASURES

The most difficult part when applying Numerical Optimization is to find an appropriate

measure of merit for each single design, i.e. to constitute the objective function. This is particularly true for cases with multiple and in general competing criteria. It seems to be self-evident to decide whether something is good or bad, but unfortunately things are never only good or bad in an absolute sense. They are more or less good or bad. What looks like a problem, adheres the solution in itself. In workaday life people need to find answers to questions like “Does a 27 year old person belong to the group of about 30 year old people?”. The astonishingly simple answer is “To some extent”, and the only question left is how to quantify this answer. Here the theory of Fuzzy Logic from Zadeh [4] comes into play. Assuming that one knows the theoretical best solution and is able to quantify an actual solution compared to that, it would be possible to specify its quality. This is the basis of Fuzzy Logic objective functions (FLOFs).

Applying Fuzzy Logic to objective functions consists of three parts:

1. Based upon the assumption that desirable, tolerable and unacceptable solutions exist, the determination of a relative quantity concerning the membership to each of these classes
2. Setting up logical rules describing the interdependencies of criteria qualities, e.g. if *criterion 1* is desirable and *criterion 2* is unacceptable, then the overall solution is unacceptable
3. Based upon the aforementioned logical rules finding some kind of compromise

The first part consists, in the terminology of Fuzzy Logic, of the constitution of so called membership functions, describing the degree of membership. In this case, a linear dependency is sufficient and yields, for the length and volume of the above example of the beam defined within the Software *CAOne*[®] [1], classes depicted in Fig.7. The only necessary information is an actual respectively current or starting value of the criterion and the direction for improvement.

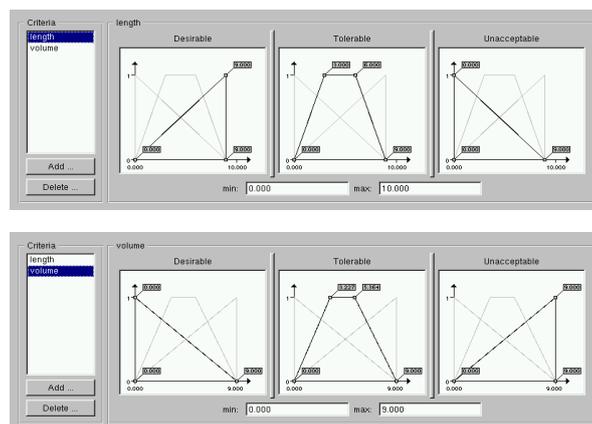


Figure 7: Classes for desirable, tolerable and unacceptable solutions for the length and volume of the beam

The second part, setting up logical rules for calculating the consequences, may look like in Fig. 8. The reader may note that it is not necessary to declare all possible rules. From experience, only the ones that are of interest should be defined. Also, the obvious rules like “If

the volume is desirable than the solution is desirable” are already available within the aforementioned software. The result of each of these rules in case of an AND-condition is, comparable to binary logic, the intersection of the considered classes, meaning the minimum of the two grades of membership.

Rules						
	IF	IS	CONDITION	IF	IS	THEN
0	volume	Desirable	AND	length	Desirable	Desirable
1	volume	Desirable	AND	length	Tolerable	Tolerable
2	volume	Desirable	AND	length	Unacceptable	Unacceptable
3	volume	Tolerable	AND	length	Desirable	Tolerable
4	volume	Tolerable	AND	length	Tolerable	Tolerable
5	volume	Tolerable	AND	length	Unacceptable	Unacceptable
6	volume	Unacceptable	AND	length	Desirable	Unacceptable
7	volume	Unacceptable	AND	length	Tolerable	Unacceptable
8	volume	Unacceptable	AND	length	Unacceptable	Unacceptable

Figure 8: Logical rules for the length and volume of the beam

The last step consists of the so called defuzzification, which means calculating an overall measure of merit for the results of the consequences. This is accomplished by calculating a compromise, again using membership functions for desirable, tolerable and unacceptable results. Fig. 9 shows the applied approach. Mathematically, it is again a weighted sum for all consequences, with the weight being the center of gravity (also called Center of Moment CoM method) of the classes respectively triangles shown in equation (8), which is the one that any optimization strategy will minimize. Here, the memberships of the consequences are denoted with s_k while x_{ck} means the location of the CoM of the respective solution class, i.e. $1/6$ for the desirable, $1/2$ for the tolerable and $5/6$ for the unacceptable class.

$$f = \frac{\sum_k x_{s_k} s_k}{\sum_k s_k} \tag{8}$$

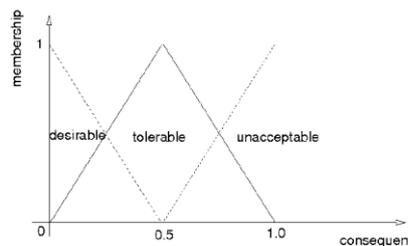


Figure 9: Classes for the desirable, tolerable and unacceptable solutions

Applying this approach to the beam example yields the solution space shown in Fig. 10. In this case it becomes obvious that the goal was to improve both criteria to the same extent, which was not clear before by applying just a weighted sum with unknown weights.

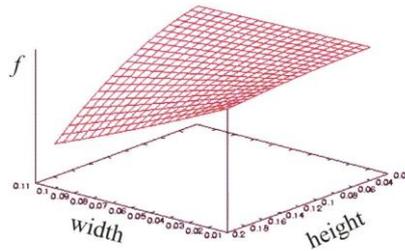


Figure 10: Solution space topography for the beam example applying a fuzzy logic objective function

5 UNDERSTANDING THE FUZZY LOGIC APPROACH

The first step in using FLOFs is the constitution of membership functions. In the case of linear dependencies this becomes a linear equation, e.g. $s_{Vd} = 1 - V/9$ and $s_{Vu} = V/9$ for the desirable or unacceptable classes of the volume in Fig. 7. Concentrating only on the desirable solutions would transform a concave Pareto front into a convex one, but also vice versa. Considering just the mere desirable solutions, without the tolerable and unacceptable ones and without any interdependency between the criteria, would yield a simple weighted sum following equation (8), which is not comparable to equation (6). Thus, it is obvious that the additional classes and the logic rules are the key ingredient for this approach. As an example, the introduction of the rule “*If Volume is desirable and Length is desirable than the solution is desirable*” would lead to equation (9), where s_V and s_L constitute the grades of membership for the desirable, tolerable and unacceptable classes respectively, while x_d , x_t and x_u are the above mentioned values for the CoM-method.

$$f = \frac{x_d (s_{Vd} + s_{Ld} + \min(s_{Vd}, s_{Ld})) + x_t (s_{Vt} + s_{Lt}) + x_u (s_{Vu} + s_{Lu})}{s_{Vd} + s_{Ld} + s_{Vt} + s_{Lt} + s_{Vu} + s_{Lu}} \quad (9)$$

Keeping in mind that for s_V and s_L the functions for the memberships have to be introduced, and here we are only considering one single rule leading to one *min*-term (AND-condition), it becomes clear that these functions do not represent a linear line like in Fig. 6. Of course, this function consists of linear parts, but at least the division by the membership functions leads to a nonlinear curve. This makes it possible to reach even the concave parts of Pareto fronts.

In order to demonstrate this capability we will get back to the example of section 3. Let the functions f_1 and f_2 from equations (4) and (5) represent two qualities or criteria for an optimization problem, which tries to minimize both at the same time. Furthermore, it is supposed that initial solutions exist, e.g. $f_1 = 0.9$ and $f_2 = 0.3$, and any improvement is welcome. This would yield the membership functions in Fig. 11. Applying absolutely no rule, which is similar to define the obvious one like “*If f_1 is desirable and f_2 is desirable than the solution is desirable*”, yields a result within the convex region, Fig. 12, but already beyond the boundary of the weighted sum. On the other hand, just introducing the rule “*If f_1 is desirable and f_2 is unacceptable than the solution is unacceptable*” finds the result on the

concave part of the Pareto front, Fig. 13. This solution is impossible to find using a standard weighted sum, although it is a valid one. Fig. 14 shows the development of weighted sum based objective functions for different weighting factors in comparison to the Fuzzy Logic approach. The unreachable region is tremendously large.

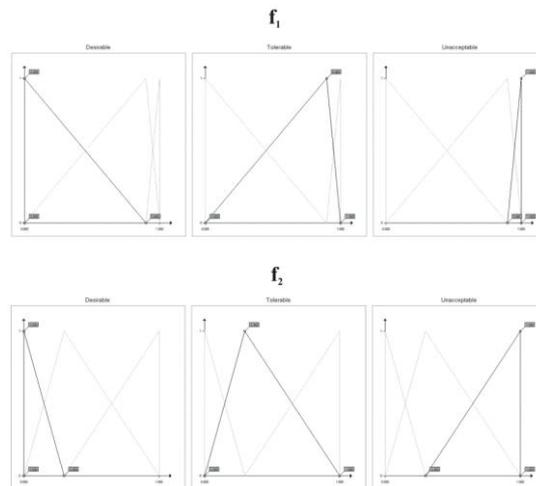


Figure 11: Membership functions

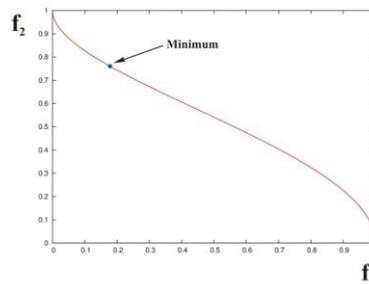


Figure 12: Minimum without any rules

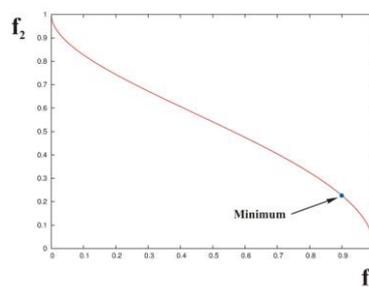


Figure 13: Minimum with one simple rule

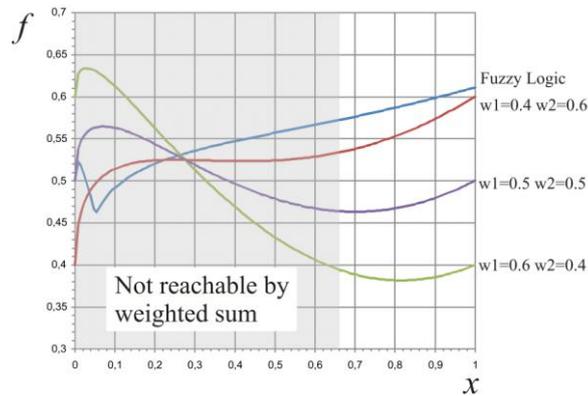


Figure 14: Comparison of different objective functions

Another effect of the application of this approach is the smoothing of the solution space topography. Fig. 15 shows on the left side a solution space which is characterized by peaks, plateaus and shallow gradients. Calculating the solution space based upon the proposed Fuzzy Logic approach yields the right picture, which shows an extremely smoothed topography. From the point of view of an optimization algorithm it is much simpler to find the optimum within this topography, particularly because the minimum has been intensified. Details concerning this can be found in [3]. This leads to a fewer number of necessary iterations and therefore additionally saves computing time.

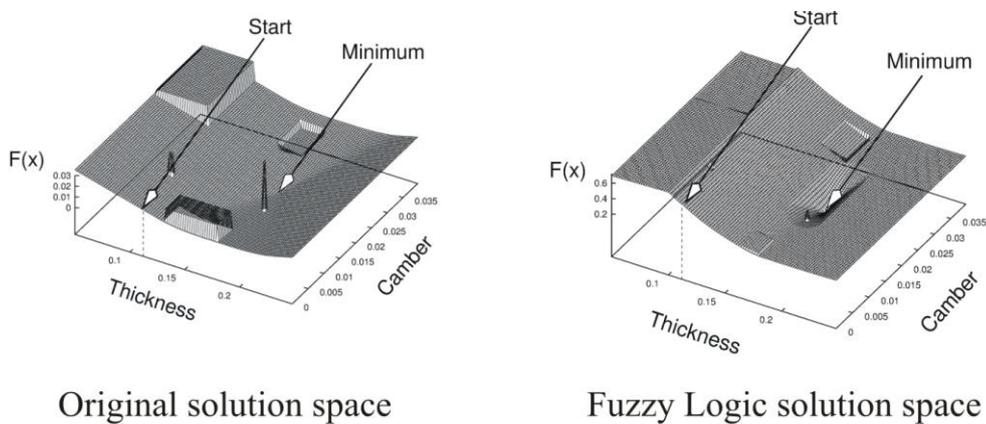


Figure 15: Smoothing of solution space topographies

6 MULTI-CRITERIA EXAMPLE

In order to show the advantages of the proposed approach in practice, a multi-criteria example was chosen. It consists of the design of an aircraft wing for a transonic transport aircraft. This wing is set up from three lofted airfoils, each described by five design parameters, Fig. 16. The design criteria are as follows:

1. Minimum aerodynamic drag in three design points (Mach numbers / lift coefficients)
2. Minimum structural weight (depending on airfoil / wing thickness, contradictory to point 1)
3. Sufficient fuel volume for a flight mission (wing volume, contradiction to point 1)
4. Best possible low-speed properties (airfoil properties contradictory to point 1)

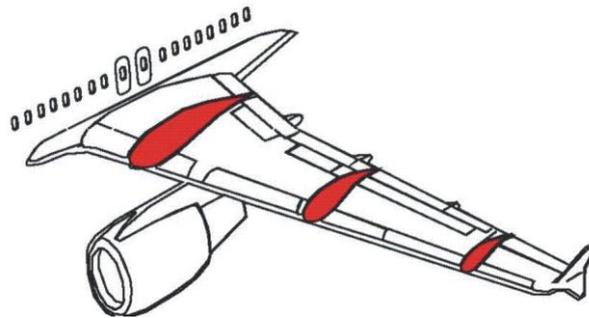


Figure 16: Aircraft wing and airfoils

These criteria are formulated in the usual linguistic manner in order to demonstrate the advantage of the FLOF approach. The modeling of all the qualities is not part of this work and can be found along with every other detail and result in [2]. All together, the designer is faced with eight criteria to be assessed (three airfoil drags, the wing weight, the fuel volume and three airfoil lifts). Now it is up to the reader to give eight weighting factors for an objective function based upon a weighted sum. Obviously, this is not possible and results in a trial and error process. For this example, three different weighting combinations were chosen from experience. On the other hand, if we have an initial solution and values for each criterion, then we easily can describe the direction of improvement and quality interdependencies. The result of an optimization based upon the three weighting combinations and a FLOF is depicted in Fig. 17. The objective function values are normalized in order to make them comparable. Clearly, the Fuzzy Logic approach reaches the minimum in less iterations and the solution is better (lower values), while as expected the weighted sums lead to different results depending on the weighting values due to the deformation of the solution space topography. The FLOF leads to a decrease of the wing weight from $91t$ to $60t$ compared to about $83t$ for the weighted sums. The same superior improvement can be found for all other criteria [2]. All together, applying the Fuzzy Logic for objective functions leads to an acceleration in convergence and better results compared to the usual approach. Above all, it is much simpler to set up and much more reliable in the sense of transforming appraisals into an exact mathematical measure of merit.

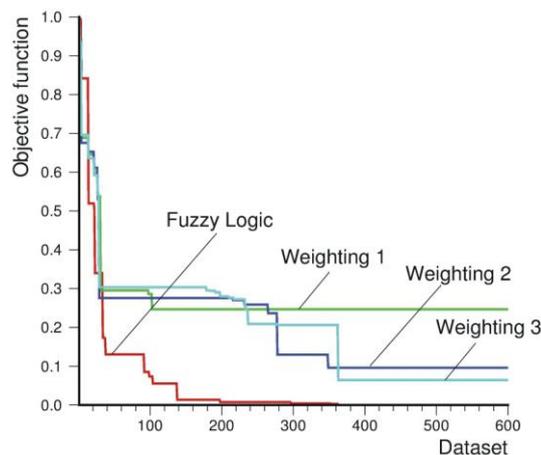


Figure 17: Convergence history for different objective functions

7 SUMMARY

Coupled problems in engineering naturally lead to contradictory criteria for each design quality and need to be assessed in relation to each other for a single goal. This is accomplished by setting up an objective function, which spans a hyper-dimensional surface in n -dimensional space. Usual approaches using weighted sums are improper, because they suffer from arbitrary choices for the weighting factors with the corresponding change in the solution space topography and, therefore, the definition of the possible optimum. Furthermore, it is impossible to describe any conceivable combination of criteria with this approach. The alternative of a so called Pareto-optimization leads to an assessment of all non dominating solutions, which is extremely difficult in the case of more than three criteria and, furthermore, shifts the procedure of the formation of opinion downstream. A new approach based upon Fuzzy Logic has been proposed, which solves all the mentioned problems and additionally yields a simple and reliable definition of design goals, even in very complicated cases. As a side effect the solution space topography is smoothed and leads to a fewer number of iterations, thus saving time and money.

REFERENCES

- [1] CAOtec Software GmbH, *CAOne[®] Users guide and tutorial*. www.caotec.com, (2011)
- [2] Frommann, O. *Bewertung multipler und gegensätzlicher Qualitätskriterien in Multidisziplinärer Optimierung*. Proc. Deutscher Luft- und Raumfahrtkongress, Leibzig (2000).
- [3] Frommann, O. *Conflicting Criteria Handling in Multiobjective Optimization Using the Principles of Fuzzy Logic*. Applied Aerodynamics Conference, Albuquerque, New Mexico, AIAA Paper 98--2730, (1998)
- [4] Zadeh, L.A. *Fuzzy Sets*. Information and Control, 8, 338/352, (1965)

A FINITE ELEMENT MODELLING OF THERMO-HYDRO-MECHANICAL BEHAVIOUR AND NUMERICAL SIMULATIONS OF PROGRESSING SPALLING FRONT

M.T. PHAN*, F. MEFTAH¹, S. RIGOBERT², P. AUTUORI³, C.
LENGLET⁴, S. DAL PONT⁵

*Laboratoire de Modélisation et Simulation Multi Echelle (MSME)
Université Paris-Est Marne-la-Vallée
5, bd Descartes, Champs sur Marne 77454 Marne-la-Vallée, France
e-mail: minh-tuyen.phan@univ-paris-est.fr

¹Laboratoire Mécanique et Matériaux du Genie Civil
Université Cergy-Pontoise
Neuville-sur-Oise, France
e-mail: fekri.meftah@ucergy.fr

²Laboratoire Central des Ponts et Chaussées
58 boulevard Lefebvre 75732 Paris, France
e-mail: stephane.rigobert@lcpc.fr

³Laboratoire Central des Ponts et Chaussées
58 boulevard Lefebvre 75732 Paris, France
e-mail: dalpont@lcpc.fr

⁴Bureau d'Etude, Bouygues Travaux publics
Challenger, 1 avenue Eugène Freyssinet, Guyancourt, 78061 Saint Quentin en Yvelines, France
e-mail: p.autuori@bouygues-construction.com

⁵Centre d'études des Tunnels
25 avenue François Mitterrand, 69674 Bron, France
e-mail: celine.lenglet@developpementdurable.gouv.fr

Key words: Thermo-hydro-mechanical model, Finite Elements, High Temperature, Spalling, Concrete

Abstract. This paper presents a coupled thermo-hydro-mechanical (THM) model enriched with a buckling-type criterion for progressive spalling. In the first part of the paper, a general fully coupled multi-phase THM model describing the behaviour of concrete at moderate and high temperatures is presented. Then the spalling criterion and its numerical implementation in the framework of the finite element method are presented.

Finally, a simple 1D numerical example will illustrate the effectiveness of the implemented numerical approach.

1 INTRODUCTION

In the past few years, major tunnel fires (Channel 1996, Ekeberg 1996, Mont-Blanc and Tauern 1999, Kaprun 2000, Gleinalm 2001 and St Gotthard 2001) has caused fatalities, severe traffic restrictions as well as important economic losses. In extreme conditions such as those typical occurring during a fire (i.e. temperatures exceeding 1200°C for considerable time spans), concrete experiences a drastic decrease of its performances due to degradation processes that are induced by several coupled thermo-hydro-chemo-mechanical (THCM) phenomena. Pore pressure build-up, restrained thermal dilatation, cement paste to aggregate incompatibility, thermal de-cohesion, dehydration . . . are just some main physical phenomena affecting concrete performances and which may cause spalling. Progressive concrete spalling occurring during a fire is a physical process of the breakdown of surface layers which flake into small pebble-like pieces in response to THM stresses.

In this contribution, a THM finite element model is enriched with a buckling-type criterion for progressive spalling. The thermo-hygral part of the THM model is based on the three fluid approach for partially saturated media [5]. The approach is developed by writing the relevant balance equations for the constituents at the pore scale and by up-scaling these equations to the macroscopic scale, taking into account thermodynamic constraints. The mechanical part is derived within the framework of poro-mechanics coupled to damage and softening plasticity [1]. The final model, after introduction of the constitutive equations, consists of a mass balance equation for the dry air, a mass balance equation for the fluid phases (water and vapour), a mass balance equation for the solid phase, an energy balance and mechanical equilibrium equations for whole porous medium.

The spalling criterion is an iterative post-processing, within each time step, of stress, pore pressure, damage and current strength fields that are obtained from THM solution. In this criterion, effective (acting on the solid phase) pore pressures may give rise to micro-cracking that percolate to form a flake whose size is related to the maximum aggregate size. A buckling analysis is then performed on the delimited flake (with damaged stiffness) while subjected to compressive stresses. The efficiency of the implemented numerical model in capturing initiation and progression of a spalling front is finally illustrated by an example.

2 THERMO-HYGRAL MODEL

The thermo-hygral model is described by a set of balance equations completed by an appropriated set of constitutive relationships describing the behaviour of concrete at moderate and high temperatures. The main physical phenomena such as vapour diffusion, liquid water flow due to pressure gradients and capillary effects, dehydration [4], evapora-

tion and condensation phenomena [9] are taken into account. The governing equations of the model are finally given in terms of the chosen state variables: the capillary pressure p_c , the gas pressure p_g and the temperature T . Hence, mass balance equations write, for solid matrix:

$$\frac{\partial m_s}{\partial t} = \dot{m}_{dehyd} \quad (1)$$

liquid water:

$$\frac{\partial m_l}{\partial t} + \nabla \cdot (m_l \mathbf{v}_{l-s}) = -\dot{m}_{vap} - \dot{m}_{dehyd} \quad (2)$$

vapour:

$$\frac{\partial m_v}{\partial t} + \nabla \cdot (m_v \mathbf{v}_{g-s}) + \nabla \cdot (m_v \mathbf{v}_{v-g}) = \dot{m}_{vap} \quad (3)$$

dry air:

$$\frac{\partial m_a}{\partial t} + \nabla \cdot (m_a \mathbf{v}_{g-s}) + \nabla \cdot (m_a \mathbf{v}_{a-g}) = 0 \quad (4)$$

where m_x is the mass per unit volume of porous medium of each constituent ($x = s, l, v, a$):

$$m_s = (1 - \phi) \rho_s, \quad m_l = \rho_l S_l \phi, \quad m_v = \rho_v (1 - S_l) \phi, \quad m_a = \rho_a (1 - S_l) \phi \quad (5)$$

in which ρ_x is the corresponding density, ϕ is the porosity, S_l is the degree of saturation, \dot{m}_{vap} is the evaporation mass rate and \dot{m}_{dehyd} is the dehydration mass rate.

Besides, the energy balance equation of the whole medium is:

$$\begin{aligned} \rho C_p \frac{\partial T}{\partial t} - K \left(C_l \frac{\rho_l k_{rl}}{\mu_l} (\nabla p_g - \nabla p_c) + C_g \frac{\rho_g k_{rg}}{\mu_g} \nabla p_g \right) \cdot \nabla T - \nabla \cdot (\lambda \nabla T) \\ = -H_{vap} \dot{m}_{vap} + H_{dehyd} \dot{m}_{dehyd} \end{aligned} \quad (6)$$

where C_p is heat capacity, \mathbf{q} is the heat flux, H_{vap} is the enthalpy of vaporization and H_{dehyd} is the enthalpy of dehydration.

3 MECHANICAL MODEL

The mechanical behaviour of the matrix is described by an elasto-plastic approach coupled to damage. The constitutive law reads:

$$\sigma = (1 - D_{tc})(1 - D_m) \tilde{\sigma} + b p_s \delta \quad (7)$$

where σ is the apparent stress tensor, $\tilde{\sigma}$ is the effective stress tensor [7], $b = 1 - (1 - b_0)[1 - (1 - D_{tc})(1 - D_m)]$ is the Biot's coefficient of damaged porous medium [6], p_s is the pore pressure acting on the solid phase and δ is the second order unit tensor.

This equation takes into account the thermo-chemical process due to degradation and the mechanical damage. Two variables of thermo-chemical damage D_{tc} and mechanical damage D_m are therefore introduced.

Furthermore, the effective strain tensor $\tilde{\varepsilon}$ is given by the decomposition of total strain tensor ε to the elastic strain tensor ε_e , the plastic strain tensor ε_p describing crack opening, the transient creep component which is the additional strain observed when heating occurs with a concomitant sustained applied load ε_{tc} [10] [11] and the free thermal strain tensor ε_{th} .

Thus the effective stress tensor is given by:

$$\tilde{\sigma} = \mathbf{E} : (\varepsilon - \varepsilon_p - \varepsilon_{th} - \varepsilon_{tc}) \quad (8)$$

The final mechanical constitutive equation then writes:

$$\sigma = (1 - D_{tc})(1 - D_m)\mathbf{E} : (\varepsilon - \varepsilon_p - \varepsilon_{th} - \varepsilon_{tc}) + bp_s\delta \quad (9)$$

4 SPALLING CRITERIA AND SIMULATION OF SPALLING FRONT

4.1 Spalling criteria

The spalling phenomenon is assumed to be the result of a combination of two main processes: the first one is the thermo-hygral process related to the build-up of pore pressures acting on the solid phase; the second one is the thermo-mechanical process which is associated to the thermal dilatation gradient.

The first process is related to the mass transfer of liquid phases (liquid water, vapour and dry air) [8]. When temperature increases, strong vapour gradients are generated causing moisture to migrate both towards the heated surface and inwards, towards the colder layers: this latter results in vapour condensation. This phenomenon, often referred to as "moisture clog", forms a quasi-saturated layer that acts as an impermeable wall for gases, resulting in gas pressure built-up (Fig:1-b). This pressure may give rise to micro-cracks that propagate forming a flake whose size is related to the maximum aggregate size.

The first spalling criterion is then stated by comparing the pore pressure and the tensile strength:

$$F_{cri1} = f_t(T) - bp_s < 0 \quad (10)$$

where F_{cri1} is the value of the first spalling criterion, f_t is the tensile strength which depends on the temperature, b is the Biot's coefficient and p_s is the pore pressure acting to the solid phase:

$$p_s = p_g - \chi p_c = p_g - S_l p_c \quad (11)$$

in which χ is the Bishop's coefficient.

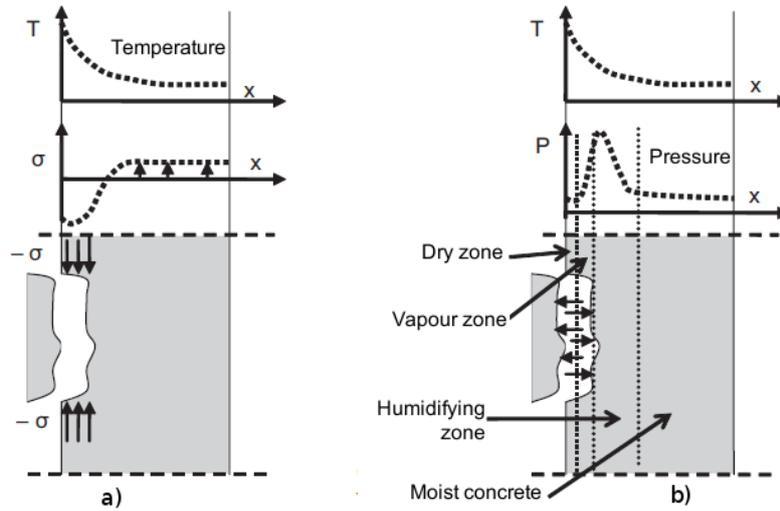


Figure 1: Spalling mechanism: a) Thermal dilatation [3] b) Pore pressure [2]

The thermo-mechanical process is directly associated to the temperature field. As temperature increases, the temperature gradients generate a thermal dilatation which engenders compressive stresses in the direction parallel to the heated face (Fig:1-a). An additional external compressive load parallel to the heated face can intensify this effect [8, 12].

If the compressive stresses induced by the thermal gradients are strong enough (and the first spalling criterion is satisfied), the external layer buckles.

The second spalling criterion writes:

$$\begin{aligned}
 F_{cri2} &= \sigma_{cr} - \langle \sigma \rangle & (12) \\
 &= \frac{\pi^2 EI}{Ah} - \langle \sigma \rangle \\
 &= \frac{\pi^2 e^2}{12h^2} (1 - \langle D \rangle) E_0 - \langle \sigma \rangle
 \end{aligned}$$

where F_{cri2} is the value of the second spalling criterion, σ_{cr} is the Euler's criterion stress, σ is the compressive stresses (in equation 7) in the direction parallel to the heated face, A is the area of spalling zone, h is the length of spalling zone, e is the depth of the spalling zone, E is the stiffness of material, $\langle \cdot \rangle$ is the average operation and D is the damage of material.

4.2 NUMERICAL SIMULATION OF SPALLING

The spalling criterion is an iterative post-processing of stress, pore pressure, damage and current strength fields that are obtained from the THM solution.

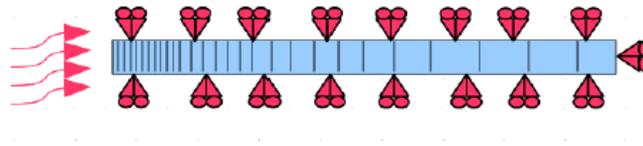


Figure 2: Studied configuration

The simulation of a progressing spalling front within one time step is then verified by using the two spalling criteria presented in the previous section. If the criteria are both verified and concrete spalls, the associated FE mesh is deactivated (no remeshing is required) and boundary conditions are transferred to the interface between the spalled zone and the intact layer of elements.

The THM model and the spalling criteria have been introduced into CESAR, a finite element code developed by Laboratoire Des Ponts et Chaussées (LCPC).

5 NUMERICAL SIMULATION

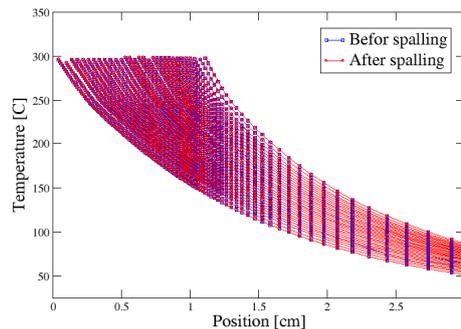


Figure 3: Temperature evolutions in different time steps

In this section, the presented THM modelling and the spalling phenomenon are illustrated by a simple numerical 1D example. The concrete specimen ($2mm \times 10cm$) is exposed to an ISO Fire and a Neumann type boundary condition on p_c and p_g on the left side of the specimen (Fig 2) is applied. In order to promote the spalling phenomenon, in this (purely numerical) example the spalling criteria related to the gas pressure is considered as satisfied (i.e. the tensile strength is artificially set to zero $f_t = 0$). In figures 3-4 the evolutions of capillary pressure, gas pressure and temperature are given. We observe that the specimen spalls many times from $t = 485s$ to $t = 760s$. The depth of the spalling zone is $12mm$ and the spalling velocity variate from $1.8mm/min$ to $5mm/min$. The figure 5 represent the evolution of capillary pressure in the first $2mm$ of the specimen. At the first spalling, the heated surface layer of 4 elements is deactivated. Then the boundary

condition is transferred successfully to the the surface between external spalled layer and inner layer. This boundary condition causes the increase of capillary pressure (Fig 4-5) and the decrease of gas pressure (Fig 4)

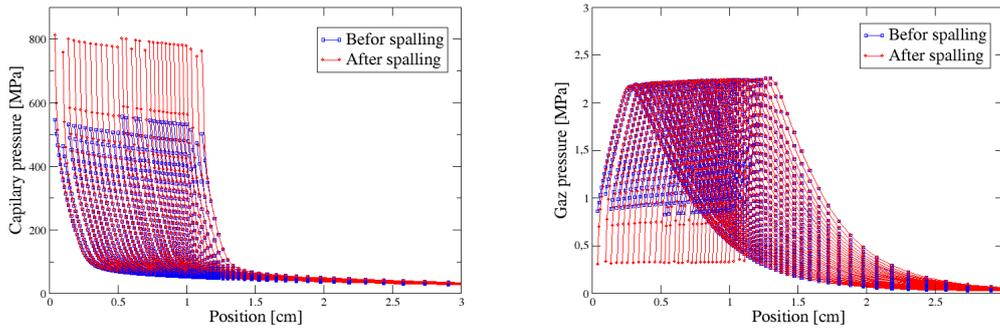


Figure 4: Capillary pressure evolutions in different time steps

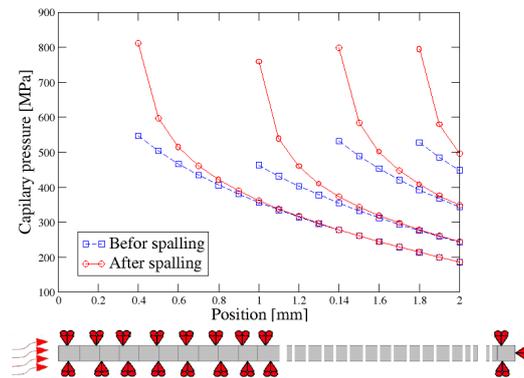


Figure 5: Capillary pressure evolutions in different time steps of 2mm external layer

6 CONCLUSION

In this paper, a fully coupled THM model, enriched with a buckling-type criterion for progressive spalling, has been presented. Spalling is taken into account by a combination of two criteria based on pore pressure and thermal expansion. The feasibility of the proposed approach is illustrated by means of a simple, 1D example illustrating the propagation of a spalling front. The presented approach constitutes a general framework for describing the behaviour of concrete when extreme loading conditions (such as those typical of a fire in a tunnel) occur.

REFERENCES

- [1] A. Ali, *Modélisation et simulation du comportement du béton sous hautes températures par une approche thermo-hygro-mécanique couplée. Application des situations accidentelles* PhD thesis, Université Marne La Vallée, 2004.
- [2] Y. Anderberg *Spalling phenomena of HPC*. International workshop on Fire Performance of High Strength Concrete, Maryland, NIST Special Publication 919, February 13-14, 1997
- [3] Z.P. Bazant *Analysis of pore pressure, thermal stresses and fracture in rapidly heated concrete*, International workshop on Fire Performance of High Strength Concrete, Maryland, NIST Special Publication 919, February 13-14, 1997
- [4] S. Dal Pont, A. Ehrlacher *Numerical and experimental analysis of chemical dehydration, heat and mass transfer in a concrete hollow cylinder submitted to high temperatures*. Int.J.Heat and mass transfer, Vol 1, 4, 135-147, 2004.
- [5] S. Dal Pont, S. Durand, B.A. Schrefler *A multiphase thermo-hydro-mechanical model for concrete at high temperatures? Finite element implementation and validation under LOCA load*. Int.J.Heat and mass transfer, Vol 237, 22, 2137-50, 2007.
- [6] S. Grasberger and G. Meschke, *Thermo-hydro-mechanical degradation of concrete: Form coupled 3D material modelling to durability oriented multifield structural analyses*, Materials and Structures, Vol 37, 244-256, 2004
- [7] L.M. Kachanov, *Introduction to continuum damage mechanics*, Martinus Nijhoff Publishers, 1986
- [8] P. Kalifa F. D. Menneteau and D Quenard *Spalling and pore pressure in HPC at high temperatures*, Cement and Concrete Research 30, 1915-1927
- [9] R.W Lewis, B.A. Schrefler, *The finite element method in the static and dynamic deformation and consolidation of porous media* Wiley, 1998.
- [10] H. Sabeur, F. Meftah, H. Colina, G. Plateret *Correlation between transient creep of concrete and its dehydration*, Magazine of concrete research, Vol 60(3), 157-163, 2008
- [11] H. Sabeur, F. Meftah *Dehydration creep of concrete at high temperatures*, Materials and Structure, Vol 41(1), 17-30, 2008
- [12] W.Z. Zheng, X.M. Hou n, D.S. Shi and M.X. Xu *Experimental study on concrete spalling in prestressed slabs subjected to fire* Fire Safety Journal 45, 283-297, 2010

AERO-THERMO-MECHANICAL COUPLING FOR FLAME-WALL INTERACTION

B. BAQUE*, **M.-P. ERRERA****, **A. ROOS***,
F. FEYEL*, **E. LAROCHE*****, **D. DONJAT*****

*ONERA - Department of Metallic Materials and Structures
BP 72 - 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France
e-mail: benedicte.baque@onera.fr, web page: <http://www.onera.fr/>

**ONERA - Department of Computational Fluid Dynamics and Aeroacoustics
BP 72 - 29 avenue de la Division Leclerc, 92322 Châtillon Cedex, France

***ONERA - Department of Aerodynamics and Energetics Modeling
BP 74025 - 2 avenue Edouard Belin, 31055 Toulouse Cedex 4, France

Key words: aero-thermo-mechanical algorithm, heat transfer, thermal strain

Abstract. This paper investigates a flame-wall interaction consisting of a premixed flame impinging on a metallic plate. This is a coupled problem as the heat transfer from the flame increases the temperature of the plate and bends it, which in turn modifies the shape of the flame. This study aims at designing an aero-thermo-mechanical coupling between both codes CEDRE (Computational Fluid Dynamics) and Z-SeT (computational solid mechanics and heat conduction) to simulate this complex system. Numerical results for aero-thermal coupling are compared with experimental data.

1 INTRODUCTION

Numerical simulations of real-world engineering systems involve a large number of physical phenomena. In the case of fluid-structure interaction with heat transfer such as the one occurring in a flame-wall interaction, many phenomena are strongly coupled: convection in the fluid, heat conduction through the solid and its deformation. Separate simulations of Computational Fluid Dynamics (CFD), computational solid mechanics and conduction generally do not give accurate solutions due to the assumptions at boundaries (such as heat flux, temperature, pressure or position) made in the separate calculations.

In the fluid-solid coupling, information provided by each model is complementary. The external coupling approach has the advantage to build upon the experience put into the separate solvers and to use the most appropriate methods in each discipline. It is therefore



Figure 1: Flame-wall interaction setup (ONERA Toulouse).

an attractive solution, but a robust and efficient algorithm between the fluid and the solid media is required.

The goal of this paper is to analyze a turbulent premixed flame impinging on a wall. The numerical problem is challenging due to the great time disparities of the physical models in interaction. Furthermore, a procedure has to be defined over a long period of time, such as the one of the experiment (300 s), with respect to the chemical and aerodynamical characteristic times. Thus, the numerical coupled strategy must provide reasonable solutions at acceptable computing efforts.

In this paper, a numerical approach is validated by comparison to experimental data. The external coupling involves three independent solvers (CEDRE for fluid, and Z-SeT for structure and heat conduction) and needs an efficient algorithm to predict the transient aero-thermo-mechanical process. The coupling strategy is responsible for determining which variables are relevant in the multi-physics interaction and which ones must be exchanged at the interfaces. The resulting algorithm carries out the temporal exchanges and is closely related to the physics to be studied.

2 FLAME-WALL INTERACTION

The experiment has been performed at the ONERA Toulouse (see Roinard *et al.* [4]). A Bunsen burner produces a premixed propane-air flame, with an equivalence ratio of $r = 1.2$ (fuel excess) and an ejection velocity of $V_0 = 2,4$ m/s. At ejection, it corresponds to a Reynolds number Re of 9000, which means that the flame is turbulent. The maximum temperature inside the flame is about 1700°C . Two flame to wall distances H (i.e. between the exit of the burner nozzle and the plate) are tested successively: 54 and 108 mm, for a nozzle diameter $D = 41$ mm. The sample wall consists of metallic plates made of two different aeronautical alloys (TA6V and INCO600), with dimensions of 400×400 mm and a thickness of 3 mm. The experiment lasts 300 s. The transient evolution of the system is recorded for temperatures on the upper plate surface (infra-red thermography), as well as its position (spectroscopic photogrammetry) and the flame front location (ultra-violet camera). An envelope flame is observed: hot burnt gases are trapped between the

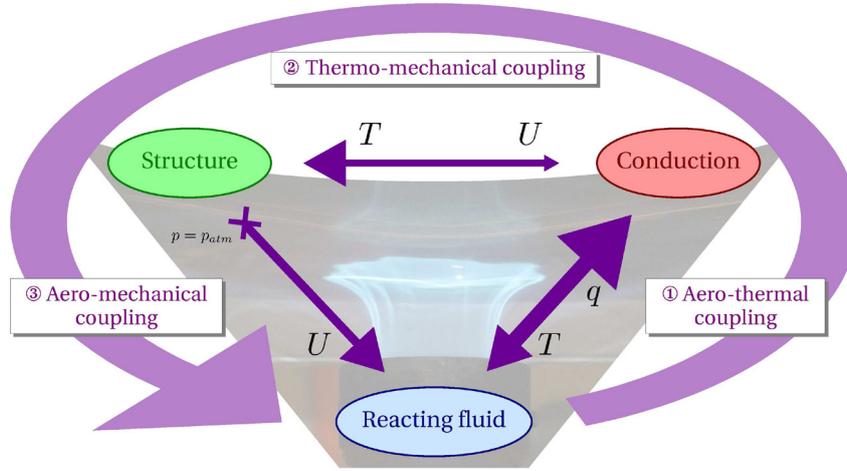


Figure 2: Mutual influences of the different physics in flame-wall interaction.

chemical reaction zone and the wall, which implies a maximal heat transfer at the center of the plate.

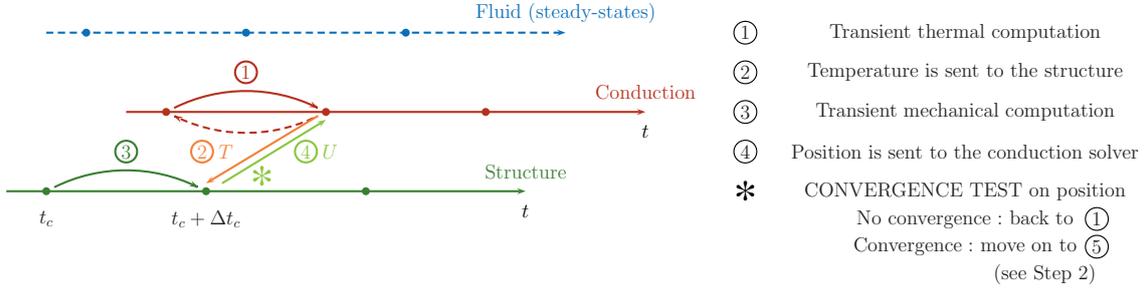
Mutual influences and relative emphasis of every physics in the flame-wall interaction are presented in Figure 2. The flame stimulates the coupled system through the transfer of a convective heat flux q (see Milson & Chigier [3]). Thus, the aero-thermal coupling (Conjugate Heat Transfer) is predominant. In reaction, the temperature T increases through the plate, which bends under dilatation and creep (thermo-mechanical coupling). The evolution of the position of the plate U causes a change of the shape of the flame (aero-mechanical coupling), which modifies the flux profile. All along the experiment, the flame pressure p is equivalent to the atmospheric pressure p_{atm} . As the variation of this pressure is negligible, the system can be simplified.

3 AERO-THERMO-MECHANICAL COUPLING

The resolution of the combustion equations is very time-consuming. As a result, the whole computation time is mainly determined by CFD. Consequently, the coupling strategy needs to minimize the fluid computation time. Thus, the algorithm in Figure 3 shows a minimum number of fluid calls. Even if the aero-thermal coupling is predominant in the flame-wall interaction, updating the flow with the best prediction of the plate position seems to be cheaper in terms of iterations to update the deforming fluid mesh. For simplicity, the algorithm has been split into two parts.

The first part consists in performing the thermo-mechanical coupling. After the transient conduction computation (step ①), the temperature T at the interface is sent to the structure (step ②), which determines the new plate position U at this coupling instant (steps ③ and ④). An iterative process is generally needed to reach temperature and

Step 1 : Thermo-mechanical coupling



Step 2 : Aero-thermal coupling with position update

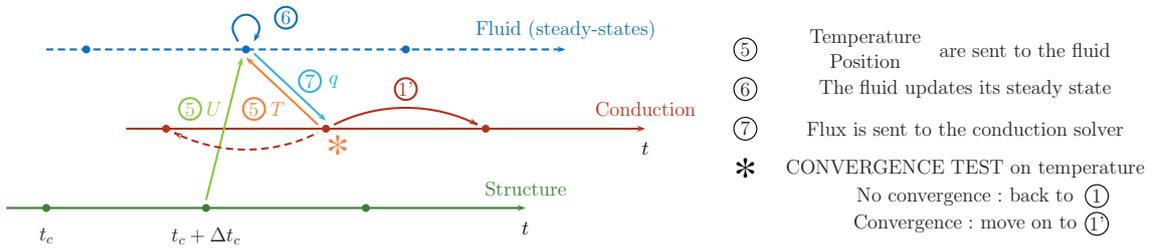


Figure 3: Aero-thermo-mechanical coupling algorithm.

position equilibria.

As for the second part, once the two fields (T, U) are converged at $t_c + \Delta t_c$, they are simultaneously sent to the fluid (step ⑤). Thus the fluid mesh can be updated according to the new plate position before computing a new steady-state flow (step ⑥). The aero-mechanical coupling is limited to the update of the plate position (one-way coupling). Then, the fluid sends the heat flux q to the conduction solver (step ⑦).

Because of the significant discrepancy of characteristic times between fluid and solid, a simplified aero-thermal coupling based on a Conjugate Heat Transfer (see Baqué *et al.* [1]) is used: the flow solution may be considered as a sequence of steady states, whereas in the solid the fields evolve in a fully transient manner. In order to reach temperature and flux equilibria, a fixed point procedure - which includes the thermo-mechanical coupling loop - is used at every coupling time step.

4 RESULTS FOR AERO-THERMAL COUPLING

The aero-thermal coupling is carried out in an axisymmetric configuration.

The turbulence of the flame is processed by a RANS model (k- ω). The reacting fluid computation produces a cool central core flame (Figure 4), which differs from the experiment: cool unburnt gases are in contact with the material at the center without heat transfer to the plate. A heat flux peak is reached at the impingement point of the flame front (around $x = 5$ cm).

Foat *et al.* [2] experimentally showed that the transition between an envelope flame and a cool central core flame is a quasi-instantaneous process. These two flames can occur

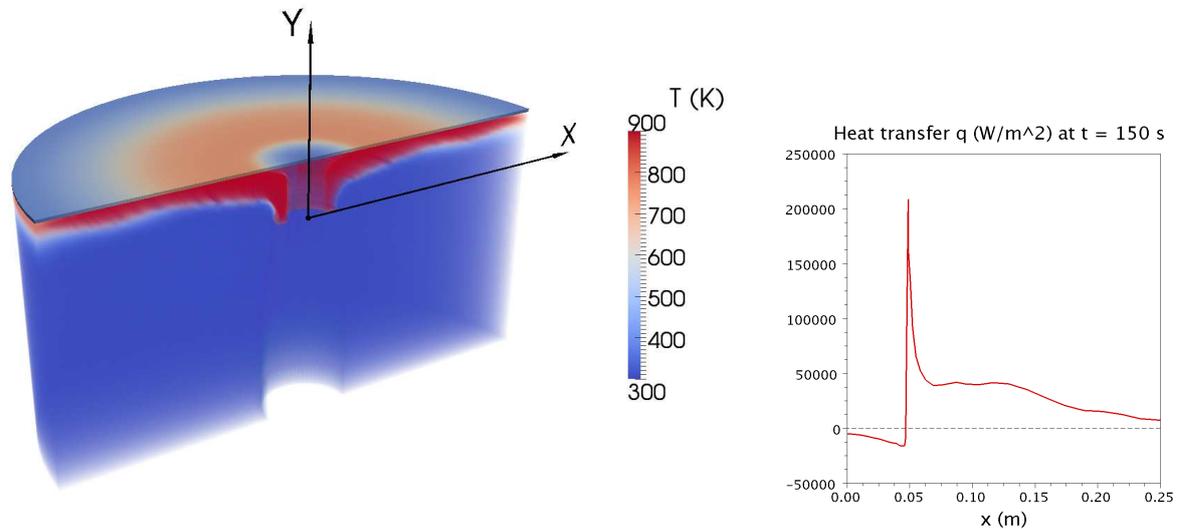


Figure 4: Temperature fields in fluid and solid and flux profile at the wall at $t = 150$ s.

at the same equivalence ratio (when the reactant mixture is fuel-rich, see Zhang & Bray [5]) depending on whether the fuel flow is increasing or decreasing. Indeed, once a cool central core has been formed it is difficult for atmospheric air to diffuse through the hot burnt gases zone, as the predominant direction of the flow is away from the stagnation zone.

As for the solid part, a radiative flux (plate emission) and a natural convective flux are set as thermal boundary conditions at the back of the plate. The external lateral side of the plate is an adiabatic wall (no heat transfer). The results in Figure 4 refer to the computation with a plate made of INCO600. In this aero-thermal computation, the deformation of the plate is not considered.

The coupling time step is $\Delta t_c = 30$ s in order to perform only ten couplings during the entire computation. The computation encountered a flame instability after $t = 150$ s.

All along the computation, the conductive heat transfer through the plate increases the temperature at the center of the plate. This implies that heat enters the boundary layer flow inside the cool central core of the flame from the plate. Moreover, the impingement point gets slowly closer to the center of the flame. Combined, these two phenomena may cause a change of the flame shape into an envelope as the one observed in the experiment, which might explain the instability of the aero-thermal calculation.

5 CONCLUSIONS

The flame-wall interaction problem has been studied on axisymmetric domains for aero-thermal coupling. As the procedure follows a logical step-by-step approach, the next computation will be the axisymmetric thermo-mechanical coupling. The influence of the

flame will be represented by a heat transfer coefficient h associated to the adiabatic temperature at the front of the plate. Furthermore, creep generates plastification inside the plate. Indeed, a permanent deformation of the plate is experimentally observed after the flame extinction. This will also be taken into account in the mechanical model.

Extension to a full aero-thermo-mechanical coupling is currently underway in 3D sectorial domains and will be considered in future work.

Analogies between non-reacting jets and reacting jets (flames) with impingement were noted by Milson & Chigier [3]. In order to save computation time, the aero-thermo-mechanical algorithm will be first validated on an equivalent case of a hot jet, which produces a bell-shaped flux.

REFERENCES

- [1] Baqué, B., Errera, M.-P. and Roos, A., Coupling procedure for modelling transient Conjugate Heat Transfer. *Proceedings of Coupled Problems* (2009).
- [2] Foat, T., Yap, K.P. and Zhang, Y., The Visualization and Mapping of Turbulent Premixed Impinging Flames. *Combustion and Flame* (2001) **125**:839-851.
- [3] Milson, A. and Chigier, N.A., Studies of Methane and Methane-Air Flames Impinging on a Cold Plate. *Combustion and Flame* (1973) **21**:295-305.
- [4] Roinard, H., Micheli, F., Reulet, P. and Millan, P., Transferts thermiques lors de l'interaction flamme paroi. *Proceedings of Société Française de Thermique* (2007).
- [5] Zhang, Y. and Bray, K.N.C., Characterization of Impinging Jet Flames. *Combustion and Flame* (1999) **116**:671-674.

APPLICATION OF A VISCOPLASTIC DAMAGE MODEL FOR THE FAILURE PREDICTION OF REGENERATIVELY COOLED NOZZLE STRUCTURES

VIVIAN TINI*, IVAYLO N. VLADIMIROV AND STEFANIE REESE

*Institute of Applied Mechanics
RWTH Aachen University
Mies-van-der-Rohe Straße, 52074 Aachen, Germany
e-mail: vivian.tini@rwth-aachen.de, www.ifam.rwth-aachen.de

Key words: Regeneratively Cooled Nozzle, Combustion Chamber, Dog-house Effect, Lifetime Prediction, Viscoplastic-Damage Model, Thermomechanical Analysis

Abstract. Regeneratively cooled nozzle structures belong to the critical components of a space shuttle main engine. The cooling channel wall in the combustion chamber is subjected to extreme cyclic thermomechanical loadings which eventually lead to the damage of the wall, well known as the "dog-house" effect. A material model for the purpose of reliable lifetime prediction is being developed. The model shall describe the material behaviour under hardening conditions as well as the superimposed effect of fatigue which occurs due to cyclic loadings.

Motivated by extending the classical rheological model for elastoplasticity with Armstrong-Frederick kinematic hardening, a viscoplastic model is formulated in the small strain regime. The coupling with damage is performed using the well-known concept of effective stress and the principle of strain equivalence. Parameter identification on the basis of experimental results for the high temperature copper alloy NARloy-Z, which is one of the typical cooling channel liner materials, is performed. The applicability of the model will be shown by means of sequentially coupled thermomechanical analyses.

1 INTRODUCTION

The regeneratively cooled rocket thrust chamber is one of the most critical components determining the performance of a Reusable Launch Vehicle (RLV). Over the years research in the field of structural computation has been performed with the goal to optimize the overall performance of rocket engines.

Finite element analyses have been implemented to simulate the cooling channel response under in-service conditions [1-8]. These previous studies mainly focused on the

application of a viscoplastic material model to describe the dog-house failure mode [1, 2, 8], as well as important factors influencing the lifetime of the combustion chamber wall [3, 5]. With the exception of the work of Schwarz et al. [8], only 2D structural analyses were performed and the material modelling did not include the effect of damage. The work of Schwarz incorporates aging as well as the crack closure effect to describe the dog-house effect.

Failure prediction of the thrust chamber structure requires complete description of the cooling channel response under extreme operating conditions. There are many factors to be considered such as the chemical reactions during the combustion process, the fluid flow in the thrust chamber and the cooling channels and the heat transfer between the fluid flow and the chamber wall structure. This makes up a complex system. For reliable lifetime prediction purposes fluid-structure-interaction analyses are indeed required.

This paper will focus on the application of a viscoplastic damage model in structural thermomechanical analyses. The thermal loads are applied in a heat transfer analysis to obtain the temperature field history of the combustion chamber segment. The temperature history is then given as input for subsequent structural analyses, where the cyclic evolution of damage is to be seen.

In the following sections the development of the model based on the well known rheological model of Armstrong-Frederick kinematic hardening will be elaborated. The material parameters are obtained by fitting tensile test experimental data performed for NARloy-Z copper alloy. Finally results of the implementation of the viscoplastic damage model for the thermomechanical analyses will be shown.

2 THE VISCOPLASTIC DAMAGE MODEL

2.1 Continuum mechanical approach by a rheological model

First of all, we extend the classical rheological model for elastoplasticity with Armstrong-Frederick kinematic hardening shown in Fig. 1(a) to include rate-dependent effects. This is done by adding a dashpot element with the viscosity parameter η as shown in Fig. 1(b). In the small strain regime, both models employ an additive decomposition of the total strain ε into elastic and inelastic parts, i.e. $\varepsilon = \varepsilon_e + \varepsilon_p$ and $\varepsilon_p = \varepsilon_{pe} + \varepsilon_{pi}$.

The spring stiffness constants are represented by the parameters E and E_h . The serially connected dashpot in Fig. 1(a) with the “pseudo”-viscosity parameter $\eta_h = E_h/(\dot{\lambda}b)$ is not a dashpot in the true sense, however by choosing the right value of the spring constant E_h and the dimensionless parameter b , the nonlinear effect of kinematic hardening can be represented. Here $\dot{\lambda}$ is the plastic multiplier.

Coupling with damage is shown schematically in Fig. 1(b) using dotted lines for the elastic spring. The spring constant E_d is defined as

$$E_d = E(1 - D) \tag{1}$$

where D is the scalar variable for isotropic damage (Kachanov 1958) [9]. The wellknown concept of effective stress (Rabotnov 1968)[9] is applied, therefore the effective stress takes

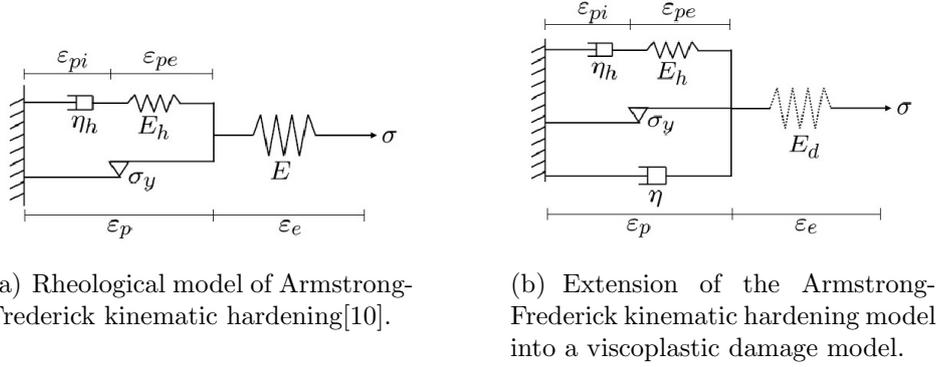


Figure 1: The classical rheological model of Armstrong-Frederick kinematic hardening and its extension into a viscoplastic damage model.

the form

$$\tilde{\sigma} = \frac{\sigma}{(1 - D)} \quad (2)$$

In this viscoplastic damage model $\dot{\lambda}$ is computed using the Perzyna formulation, $\dot{\lambda} = \langle \Phi \rangle / \eta$, where η is the viscosity parameter. The symbol $\langle \cdot \rangle$ defines the Macauley brackets i.e. $\langle x \rangle = \frac{x + |x|}{2}$.

2.2 Three dimensional extension within small strains formulation

Based on the rheological model and the principle of strain equivalence, the constitutive equations are generalized for three dimensions. Assuming elastic isotropic response, the stress-strain relationship can be written as $\boldsymbol{\sigma} = \tilde{\mathbf{C}}_1 [\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}_p]$, $\tilde{\mathbf{C}}_1 = (1 - D) \mathbf{C}_1$, where \mathbf{C}_1 is the fourth order elasticity tensor. The back stress tensor is defined as $\mathbf{X} = \mathbf{C}_2 [\boldsymbol{\varepsilon}_{pe}]$, where \mathbf{C}_2 is another elasticity tensor. The components of both elasticity tensors depend on the elastic moduli E_1 , E_2 and the Poisson's ratios ν_1 , ν_2 , respectively.

Applying J_2 flow theory and nonlinear isotropic hardening, the loading function is taken from the von Mises condition:

$$\Phi = \left\| \tilde{\boldsymbol{\sigma}}^D - \mathbf{X}^D \right\| - \sqrt{\frac{2}{3}} (\sigma_y + R) \quad (3)$$

where $(\cdot)^D$ represents the deviatoric term of the tensor. The isotropic hardening function is defined as

$$R = Q_0 (1 - \exp(-\kappa \alpha)) \quad (4)$$

where α is the internal variable for isotropic hardening, Q_0 and κ are material parameters. The evolution equations of the plastic strain $\boldsymbol{\varepsilon}_p$, the inelastic part of the plastic strain $\boldsymbol{\varepsilon}_{pi}$,

and the accumulated plastic strain α are given as follows:

$$\dot{\epsilon}_p = \frac{\dot{\lambda}}{1-D} \frac{\tilde{\sigma}^D - \mathbf{X}^D}{\|\tilde{\sigma}^D - \mathbf{X}^D\|}, \quad \dot{\epsilon}_{pi} = \dot{\lambda} b \boldsymbol{\epsilon}_{pe}^D, \quad \dot{\alpha} = \sqrt{\frac{2}{3}} \dot{\lambda} \quad (5)$$

The isotropic damage variable D is assumed to evolve in the following way over time:

$$\dot{D} = \sqrt{\frac{2}{3}} \frac{\dot{\lambda}}{1-D} \left(\frac{Y}{S}\right)^k H_{(\alpha-p_D)}, \quad Y = \frac{1}{2} \boldsymbol{\epsilon}_e \cdot \mathbf{C}_1 [\boldsymbol{\epsilon}_e] \quad (6)$$

Here p_D is the damage threshold and Y is the strain energy density release rate. The step function H is zero for $\alpha < p_D$ and is equal to one for $\alpha \geq p_D$. The rate of damage evolution is influenced by the material parameters S and k . The plastic multiplier is defined to take the form

$$\dot{\lambda} = \frac{\langle \bar{\Phi}^m \rangle}{\eta}, \quad \bar{\Phi} = \frac{\|\tilde{\sigma}^D - \mathbf{X}^D\|}{\sqrt{\frac{2}{3}} (\sigma_y + R)} - 1 \quad (7)$$

where m is a material parameter. In total the model has 13 parameters: $E_1, \nu_1, E_2, \nu_2, \sigma_y, Q_0, \kappa, b, S, k, p_D, \eta, m$. The Poisson's ratios ν_1 and ν_2 are assumed to be constant with respect to temperature. Other parameters depend on temperature.

3 FITTING WITH EXPERIMENTAL DATA

Mechanical and physical properties of the NARloy-Z alloy, such as the modulus of elasticity, the Poisson's ratio, the density, specific heat, thermal conductivity, and the coefficient of thermal expansion are obtained from the experimental report of six candidate rocket materials prepared by Esposito et al. for the NASA Lewis Research Center [11]. The yield stress and the kinematic hardening parameters are defined by fitting the model response to tensile stress-strain curves at 27.6 K, 294.3 K, 533.1 K and 810.9 K. The temperature dependency of these parameters is defined to be linear.

For the material parameters and the model response see Table 3 and Fig. 3. The values of these parameters at other temperatures are obtained by linear interpolation. The experimental data are represented by different symbols (see "exp 27.6", "exp 294.3", "exp 593.1", "exp 810.9"). The simulation results are plotted as solid lines. It is assumed that damage has not taken place yet. Furthermore, viscous effects are neglected. The tensile tests were performed at a strain rate of 0.002 s^{-1} .

Identification of the damage parameters p_D, S, k and the parameters η and m requires tensile tests until rupture as well as relaxation tests at different temperatures. Unfortunately these data are barely available and experimental work is beyond the scope of the current project.

σ_y	E_2	b	θ
[MPa]	[GPa]	[-]	[K]
190	7.47	155	27.6
158	7.84	150	294.3
130	7.49	170	533.1
90	7.71	225	810.9

Table 1: The value of the yield stress and the kinematic hardening parameters fitted at the test temperatures.

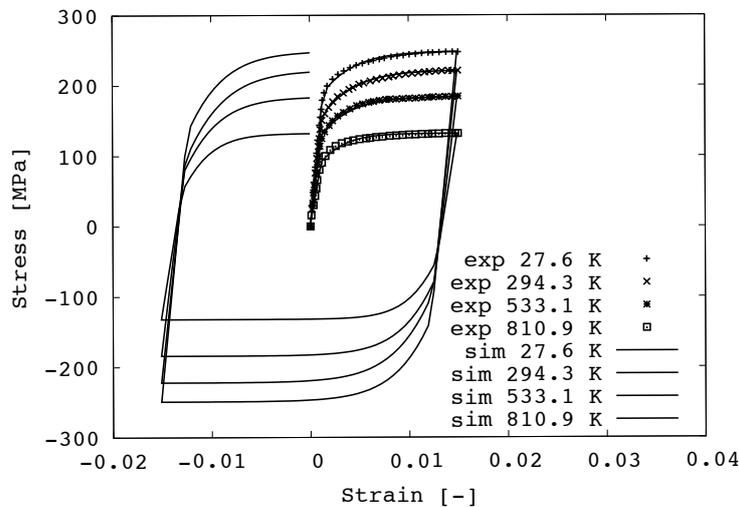


Figure 2: Tensile tests on NARloy-Z. Experimental results are taken from [11] and shown by the different symbols, computational results are shown by the solid lines.

4 THERMOMECHANICAL ANALYSES OF THE COMBUSTION CHAMBER SEGMENT

The viscoplastic damage model presented is discretized using the implicit backward Euler scheme and is implemented as a user subroutine UMAT into the Abaqus finite element analysis software. First of all, a transient thermal analysis of the combustion chamber segment is performed to obtain the temperature field of the entire chamber segment. The temperature history is then used as input for the user subroutine in the static analyses.

4.1 Transient thermal analysis

Fig. 3(a) shows the schematic cross section of a typical combustion chamber. The outer wall is made out of nickel alloy. The cooling channel walls are made out of copper alloy. The red and blue areas show where the hot gas and the coolant flow.

The geometry of the modelled segment is the quasi Vulcain geometry following the

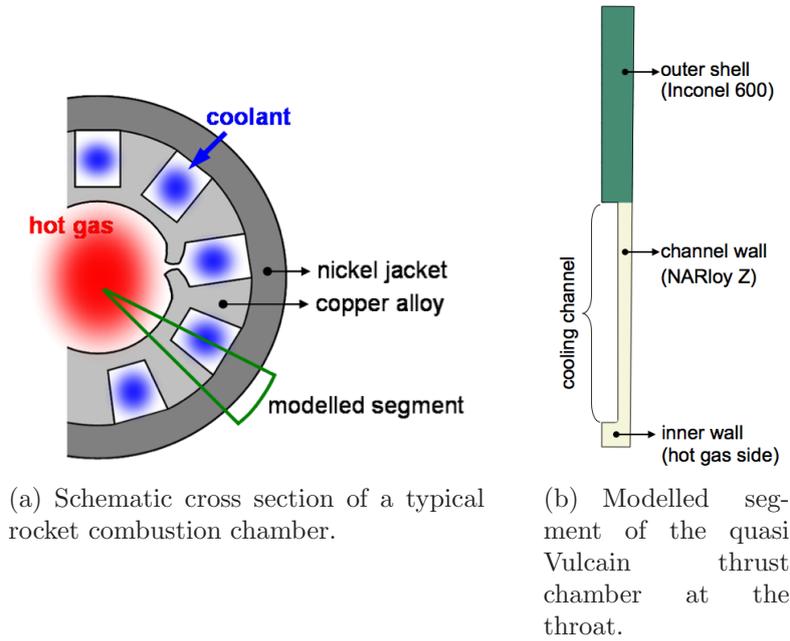


Figure 3: Schematic cross section of a combustion chamber and the modelled segment.

Phase	Time [s]	T_{hotgas} [K]	$T_{coolant}$ [K]
Pre-cooling	0 - 2	40	40
Hot run	3 - 603	950	40
Post-cooling	604 - 605	40	40
Relaxation	605 - 620	293.15	-

Table 2: Thermal cycle applied for the transient thermal analysis.

work of Kuhl et al. [12]. Convective thermal boundary conditions are employed at the inner and outer radii as well as in the cooling channel similar to the work of Riccius et al. [4]. The left and right sides have zero flux boundary conditions to ensure symmetry of the thermal field. The thermal cycle described in Table 2 is applied in the analysis. Fig. 4 shows some snapshots of the temperature field resulting from the transient thermal analysis.

4.2 Static analyses

The resulting temperature distribution from the thermal analysis is used as input for subsequent 3D static analyses. The goal of the static analysis is to see the feasibility of the implemented viscoplastic damage model to describe the dog-house failure mode qualitatively. The pressure cycle in Table 3 is applied as load in the static analyses.

For the static analyses 8-node brick elements with reduced integration were applied.

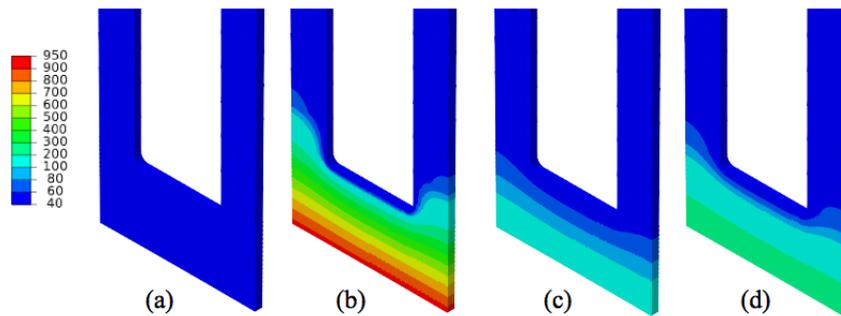


Figure 4: Temperature distribution of the combustion chamber segment at different phases of the assumed operational cycle: (a) pre cooling (b) hot run (c) post cooling (d) relaxation.

Phase	Time [s]	P_{hotgas} [MPa]	$P_{coolant}$ [MPa]
Pre-cooling	0 - 2	0	2
Hot run	3 - 603	10	14.5
Post-cooling	604 - 605	0	2
Relaxation	605 - 620	0	0

Table 3: Pressure cycle applied for the static analyses.

The critical area of interest is the cooling channel wall at the hot gas side. Two different meshes were used. The coarser and the finer mesh have in total 1338 and 4014 elements respectively. Fig. 5 shows the discretization of the hot gas side wall using both meshes. The corner of the cooling channel passage is rounded with 0.1 mm fillet radius.

The damage contours obtained from the analyses are shown in Fig. 6. The computation using Mesh 1 could be performed up to 16 cycles. Applying Mesh 2 the computation could be performed up to 13 cycles. As the damage increases, the local iteration within the user subroutine requires smaller and smaller time steps. The corresponding cyclic damage evolutions over time are shown in Fig. 7. Here it can be seen that with the coarser mesh, the damage evolves slightly faster, in comparison to the one obtained using the finer mesh. As mentioned before, the analyses were performed neglecting viscous effects. This might be the reason of the mesh dependency of the result. This phenomenon is well known to occur in elastoplasticity coupled with damage due to the change of the type of the differential equations to be solved as damage evolves.

Improvement on the robustness of the computation shall enable further computations, such that no convergence problem would occur as the number of elements with very low stiffness due to damage grows. Fig. 8(a) shows the predicted end shape of the cooling channel segment. The grey area represents the area mostly affected by damage. Fig. 8(b) shows the cross section of a combustion chamber made out of the NARloy-Z copper alloy. It is mentioned by Hannum et al. [13] that low cycle fatigue has stronger influence on the dog-house failure mode occurring within a NARloy-Z combustion chamber. It can be

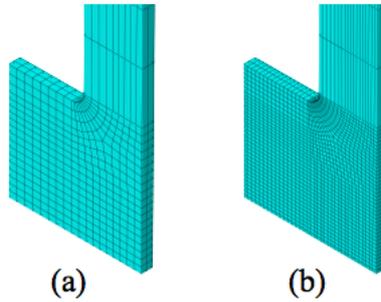


Figure 5: (a) Mesh 1 with 1338 elements (b) Mesh 2 with 4014 elements

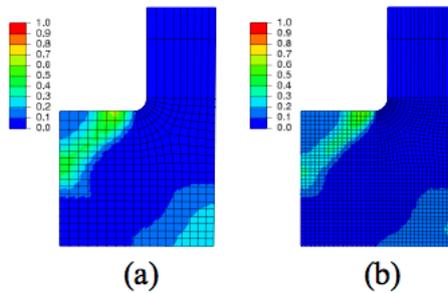


Figure 6: Damage contour obtained using the viscoplastic damage model implemented in UMAT: (a) Mesh 1 at the 16th cycle. (b) Mesh 2 at the 13th cycle.

seen that the hot gas side wall does not bulge so much towards the chamber. On the contrary, in an OFHC combustion chamber, it has been observed that the hot gas side wall becomes significantly thinner and bulges more towards the interior of the chamber. It is mentioned that crack occurs after the occurrence of a necking phenomenon. For both materials, the typical dog-house shape of the cooling channels are to be seen. However depending on the liner material, the failure cause can be different.

5 CONCLUSIONS

In this work the application of a viscoplastic damage model for the failure prediction of a typical rocket combustion chamber wall made from NARloy-Z has been presented. Further improvement on the robustness of the computation, as well as considering temperature dependency of the damage and viscosity parameters shall enable better failure prediction.

6 ACKNOWLEDGEMENT

Financial support of the German Research Foundation (DFG) within the project CRC/Transregio 40 Fundamental Technologies for the Development of Future Space-

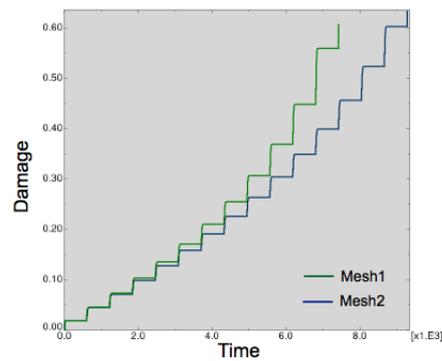


Figure 7: The cyclic evolution of the damage over time.

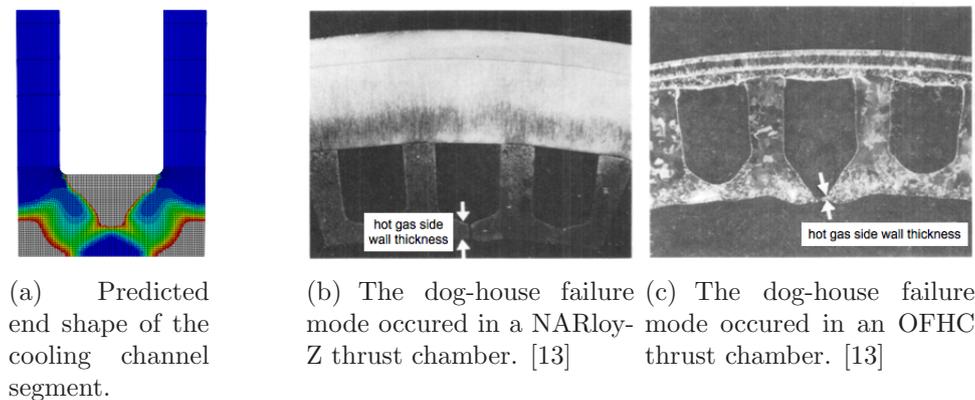


Figure 8: The predicted end shape of the channel segment from the simulation and the dog-house failure mode.

Transport-System Components under High Thermal and Mechanical Loads is gratefully acknowledged.

REFERENCES

- [1] Vinod K. Arya. Viscoplastic analysis of an experimental cylindrical thrust chamber liner. *AIAA Journal*, 30(3):781–789, March 1992.
- [2] D.T. Butler. Jr, Jacob Aboudi, and Marek-Jerzy Pindera. Role of the material constitutive model in simulating the reusable launch vehicle thrust cell liner response. *Journal of Aerospace Engineering*, 18(1):28–41, 2005.
- [3] Jörg R. Riccius, Malte R. Hilsenbeck, and Oskar J. Haidn. Optimization of geometric parameters of cryogenic liquid rocket combustion chambers. In *AIAA 2001-3408*, 1–11, AIAA, 2001.

- [4] Jörg R. Riccius and E. B. Zametaev. Stationary and dynamic thermal analyses of cryogenic liquid rocket combustion chamber walls. In *38th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, 2002.
- [5] Jörg R. Riccius, Oskar J. Haidn, and Evgeny B. Zametaev. Influence of time dependent effects on the estimated life time of liquid rocket combustion chamber walls. In *40th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, 2004.
- [6] Jörg R. Riccius, Evgeny B. Zametaev, Oskar J. Haidn, and Gaelle de Boisvilliers. Comparison of 2d and 3d structural fe-analyses of lre combustion chamber walls. In *42th AIAA/ASME/SAE/ASEE Joint Propulsion Conference*, 2006.
- [7] Jörg R. Riccius, Evgeny B. Zametaev, Oskar J. Haidn, and Christian Gogu. Lre chamber wall optimization using plane strain and generalized plane strain models. In *42nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, 2006.
- [8] W. Schwarz, S. Schwub, K. Quering, D. Wiedmann, H. W. Höppel, and M. Göken. Life prediction of thermally highly loaded components: Modelling of the damage process of a rocket combustion chamber hot wall. In *2010 Space Propulsion Conference*, 2010.
- [9] J. Lemaitre and R. Desmorat. *Engineering Damage Mechanics*. Springer-Verlag, Berlin Heidelberg, 2005.
- [10] I. N. Vladimirov, M. P. Pietryga, and S. Reese. On the modeling of nonlinear kinematic hardening at finite strains with application to springback - comparison of time integration algorithms. *International Journal of Numerical Methods in Engineering*, **75**:1–28, 2008.
- [11] John J. Esposito and Ronald F. Zabora Thrust Chamber Life Prediction Volume I - Mechanical and Physical Properties of High Performance Rocket Nozzle Materials NASA Lewis Research Center, 1975.
- [12] D. Kuhl and Jörg Riccius and O.J. Haidn Thermomechanical Analysis and Optimization of Cryogenic Liquid Rocket Engines *Journal of Propulsion and Power*, **18**:835–846, 2002.
- [13] Ned P. Hannum and Harold G. Price, Jr. Some Effects of Thermal-Cycle-Induced Deformation in Rocket Thrust Chambers NASA Langley Research Center, 1981.

COUPLED ANALYSIS OF TRANSPORT PROCESSES AND MECHANICAL BEHAVIOUR OF CONCRETE AT HIGH TEMPERATURES

F. CRAMER*, L. OSTERMANN*, U. KOWALSKY* D. DINKLER*

*Institute for Structural Analysis
TU Braunschweig, 38106 Braunschweig, Germany
e-mail: f.cramer@tu-bs.de, www.statik.tu-bs.de

Key words: Concrete, High Temperatures, Thermo-chemical Damage, Mechanical Behaviour, Non-local Damage, Structural Analysis

Abstract. A model for the numerical analysis of hygro-thermal and mechanical behaviour of concrete at high temperatures including transport, reaction and deformation processes is presented. For the description of transport and material behaviour the model comprises the balance equations for mass and enthalpy and the constitutive equations for concrete at high temperatures. The degradation of the cement minerals C-S-H and portlandite is described by two chemical reactions. The consideration of thermo-chemical damage in the classical non-local damage theory is discussed. Finally, the finite element formulation for solving the coupled balance equations regarding the primary variables displacement, temperature and relative humidity is presented and numerical results of a 3-D frame corner are discussed.

1 INTRODUCTION

Several serious fires in multi-story buildings and tunnels have been leading to increased research interest in material behaviour under extremely high temperatures. Especially in buildings made of concrete, the most frequently reported building material, fire may effect the strength of construction and reduces the load-carrying capacity.

Concrete is a composite material and consists of several phases. The solid phase comprises of aggregates and cement paste. Here, the cement paste is composed of Calcium-Silicate-Hydrates (C-S-H) and portlandite (CH) and the aggregates are made up of siliceous sand and gravel. The liquid phase consists of capillary water and physically bounded water, dry air and vapour provide the gas phase. High temperatures rising up to 1000 K in the case of fire lead to evaporation of capillaries and absorbed water with concurrent increase of the pore pressure in areas close to surfaces. Above the critical temperature of water, $T_{cr} = 647.3 K$, water only exists in form of gas, i.e. vapour.

Additionally a chemical reaction in the cement paste leads to degradation of the cement minerals, vapour is released and the porosity of the cement stone increases. As the result of evaporation and dehydration the rising pore pressure and thermal strains lead to damage which often manifests in the form of spalling.

2 BALANCE EQUATION

For the description of transport and reaction phenomena the balance equations of vapour and water and the conservation of energy are taken into account. In the proposed model the effect of dry air is neglected, thence for the gas phase only a vapour balance is necessary. The following balance equations are formulated for the volume average on macroscopic level

$$\frac{\partial}{\partial t} (S_g n \rho_v) + \nabla \cdot \mathbf{q}_v = s_{v \leftrightarrow w} + s_{v, deh} \quad . \quad (1)$$

The first term of the vapour balance covers the local mass and depends on the porosity n , the saturation of the gas phase S_g and the density of vapour ρ_v . The source terms $s_{v \leftrightarrow w}$ and $s_{v, deh}$ represent sources from phase changes between liquid and gas and the freed vapour from the dehydration reaction. The water balance equation is formulated equivalently with saturation of the liquid phase S_w , density of water ρ_w and the phase change $s_{w \leftrightarrow v}$. The second term in equation (1) represents the pressure driven flow. Here a Darcy law is used for water and vapour. With the assumption of an instant phase change of water and vapour, the summation of water and vapour balances leads to the general water balance equation

$$\frac{\partial}{\partial t} (S_g n \rho_v) + \nabla \cdot \mathbf{q}_v + \frac{\partial}{\partial t} (S_w n \rho_w) + \nabla \cdot \mathbf{q}_w = s_{v, deh} \quad . \quad (2)$$

In context of modelling concrete under high temperatures the formulation of the conservation equation of energy leads to the conservation of heat which depends on the heat capacities C_i of each phase, flows, the source terms for evaporation or condensation and the heat of dehydration

$$\left(\sum C_i \bar{\rho}_i \right) \frac{\partial T}{\partial t} + \nabla \cdot \mathbf{q}_T = s_{v \leftrightarrow w} \cdot \Delta H_{v \leftrightarrow w} + \Delta H_{deh} \quad . \quad (3)$$

The first term of Equation (3), the local change of heat, depends on temperature T and density of the components related to the concrete volume $\bar{\rho}_i$. The next term describes the heat flow which is divided in a convective and a conductive flow following the Fourier law. The convective part of heat flow is described by the product of the respective flow velocity, density and heat capacity. The heat of evaporation and condensation is the product of change of pore water and latent heat of water which depends on temperature. The last term gives the heat of chemical reaction and may be determined as the product of dehydration rate multiplied by the specific reaction heat. For coupling heat and mass

transfer with mechanical behaviour in addition the balance of linear momentum

$$\nabla \cdot \boldsymbol{\sigma} + \rho_c \mathbf{g} = 0 \quad (4)$$

with Cauchy stress $\boldsymbol{\sigma}$, density of concrete ρ_c and gravitation \mathbf{g} is formulated. The density of concrete ensues from the density of the components in consideration of their volume contents.

3 CONSTITUTIVE EQUATIONS

In addition to the balance equations constitutive models for transport, chemical reactions and the mechanical behaviour are necessary which relate the primary variables relative humidity φ , temperature T and displacement vector \mathbf{u} with other variables. They are assembled in the next section.

3.1 Heat and mass transport

The transport processes in general may be divided into convection and diffusion. The convective transport is described by Darcy's law and reads for the liquid phase as

$$\mathbf{q}_w = \rho_w \cdot \frac{\mathbf{k} \cdot k_{rw}}{\mu_w} \cdot (-\nabla p_w + \rho_w \mathbf{g}). \quad (5)$$

The water pressure p_w which can be expressed in terms of the gas pressure and the capillary pressure $p_w = p_g - p_c$ is the driving force for the transport in the liquid phase, thence the last term is significantly smaller. Further parameters are the intrinsic permeability tensor \mathbf{k} , which is depending on the degree of dehydration and the grade of damage, the relative permeability k_{rw} and the dynamic viscosity of water μ_w which is a function of temperature. Analogously the vapour flow is determined by respective parameters. The relative permeabilities k_{rw} and k_{rg} depend only on saturation [1] and are given in Figure 1. As a result of evaporation of pore water and dehydration vapour is generated, what locally leads to an increase of vapour pressure and therewith a higher vapour concentration. The resulting diffusion is described by the diffusion model after Fick

$$\mathbf{q}_v = -\mathbf{D} \cdot \nabla \rho_v. \quad (6)$$

The diffusion coefficient tensor \mathbf{D} incorporates the influence of pore structure and depends on the saturation, temperature and pressure as well.

The saturation of water described by p_c - S_w -relation is received from experimental results [1]. Furthermore, saturation depends on temperature which influences the surface tension of water. Above the critical temperature only vapour exists which is described by the ideal gas law.

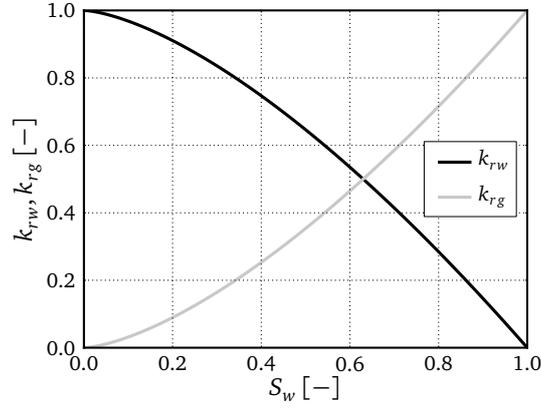


Figure 1: Relative permeability of liquid and gas phase

3.2 Chemical reaction

Besides the transport dehydration is a determining process at high temperatures. During dehydration water, which the concrete contains in form of pore water or chemically bounded water in the crystals, is released and the cement paste is decomposed. Figure 2 shows the degree of dehydration. Here, degradation of portlandite and C-S-H phases

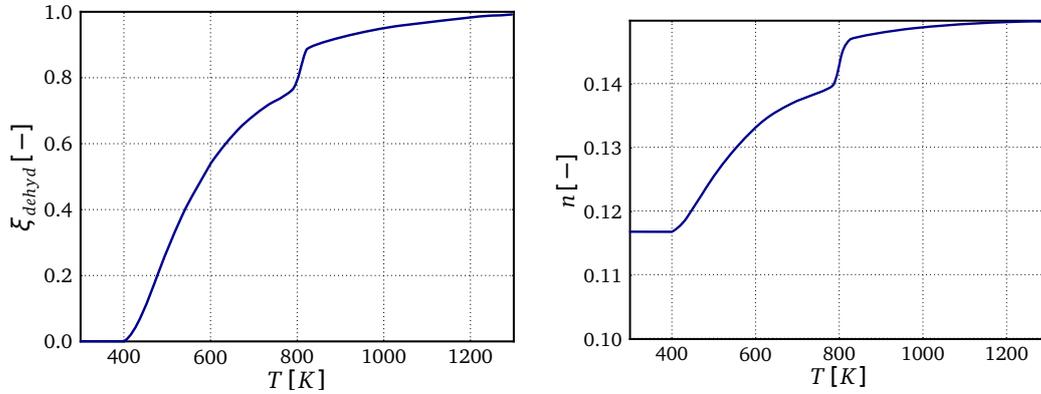


Figure 2: Degree of dehydration and porosity

may be considered by two chemical reactions, which are leading to additional sources in the heat balance and in the vapour balance [6]. Due to the decomposed cement paste the porosity n increases, see Figure 2, which leads to a significant change in permeability. The enormous ascent of the curves in the temperature range 800-850 K is due to the reaction of the portlandite phase.

After the dehydration process the micro structure is not destructed completely, therefore the scalar chemical damage D_{chem} is defined with the linear relationship $D_{chem} = 0.5 \cdot \xi_{dehyd}$.

3.3 Mechanical behaviour

The stress state at a material point is bounded by a failure surface consisting of a Rankine criterion for tensile principal stresses and of a Drucker-Prager criterion for compressive stresses. Due to the composite failure surface, Koiter's flow rule for multi-surface plasticity [5] with the internal variables λ_R and λ_{DP} is used to compute the inelastic strain tensor for stress states at the surface.

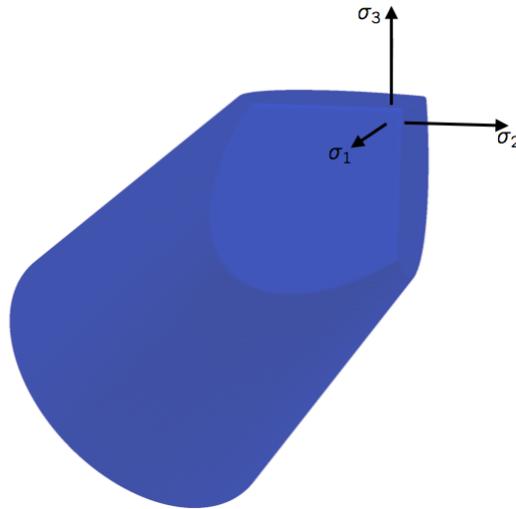


Figure 3: Segmented failure surface for tension and compression

Due to the fact that in classical damage models the results depends on the discretisation parameters the non-local damage theory is used. The isotropic damage parameters D_R and D_{DP} are described with an exponential approach by Peerlings [7] and are affected by the non-local variables $\bar{\lambda}_R$ for tension and $\bar{\lambda}_{DP}$ for compression. The total mechanical damage which is defined by

$$D_{mech} = 1 - (1 - D_R) \cdot (1 - D_{DP}) \quad (7)$$

allows an individual damage evolution for concrete in uniaxial and multiaxial stress states.

3.4 Coupling of heat and mass transport with the mechanical behaviour

For the coupling of heat and mass transport with the mechanical behaviour it is essential to consider the pore pressure. Often an effective stress tensor

$$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + b p_{pore} \mathbf{I} \quad (8)$$

is employed, which is the sum of the Cauchy stress tensor and the product of effective pore pressure time the Biot number b times the unit tensor [4]. Baroghel-Bouny et al. define the effective pore pressure as the of the sum of the pressures of liquid and gas

phase [1]. Against this Gawin et al. propose a definition of the effective pore pressure as the difference of gas pressure and the capillary pressure [2]. On the other hand pore pressure can be considered by using an equivalent strain

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^{el} + \boldsymbol{\varepsilon}^{in} + \boldsymbol{\varepsilon}^{hom} \quad . \quad (9)$$

The elastic strain $\boldsymbol{\varepsilon}^{el}$ and the inelastic strain $\boldsymbol{\varepsilon}^{in}$ result from external load or displacement constraints, whereas the homogenised strain $\boldsymbol{\varepsilon}^{hom}$ is effected by pore pressure within the cement paste, thermal strain of aggregates and cement, and chemical and hygral shrinkage of the cement paste.

Furthermore a combination of mechanical damage D_{mech} and thermo-chemical damage D_{chem} for the description of the total damage D is necessary. Therefore the product approach

$$D = 1 - (1 - D_{chem}) \cdot (1 - D_{mech}) \quad (10)$$

is used. The total damage D is included by the effective stress concept in the material model.

4 STRUCTURAL ANALYSIS

The balance equations are discretised in space by a finite element formulation. The primary variables relative humidity, temperature and displacement are expressed by their nodal values. The discretised form of the model equations reads as the non-linear coupled equations system

$$\mathbf{C}_{ij}(\mathbf{x}) \frac{\partial \mathbf{x}}{\partial t} + \mathbf{K}_{ij}(\mathbf{x}) \mathbf{x} = \mathbf{f}_i(\mathbf{x}). \quad (11)$$

The non-linear coefficient matrix \mathbf{C} describes the local changes of variables with respect to time, \mathbf{K} the fluxes. For the discretisation in time an implicit Euler scheme is used. The system of non-linear equations is solved by the Newton-Raphson procedure.

For numerical analysis of the process of heating concrete, a 3-D frame corner with heat impact at three surfaces is considered taking the symmetry of the structure into account. The dimension of the cross-section area is 30x30 cm². The exposure of the structure to fire is modelled by means of the volume averaged flux boundary condition as described in [3]. The initial conditions of temperature and relative humidity are $T_0 = 293.15 \text{ K}$ and $\varphi_0 = 0.6$, the temperature of the fire T_∞ follows the Standard-Fire characteristics.

Figure 4 depicts results of numerical analysis after 10 min. for temperature, degree of dehydration, relative humidity and vapour pressure. During the heat impact the temperature increases and diffuses into the interior. The highest temperature and therefore the greatest damage occurs between the claimed surfaces. At the same time, water evaporates rapidly which induces enormous pressure increase in the gas phase and causes a moisture clog which can be identified by local increase in relative humidity. The highest vapour pressure from evaporation and dehydration appears behind the edges near to the surfaces. Here spalling occurs first.

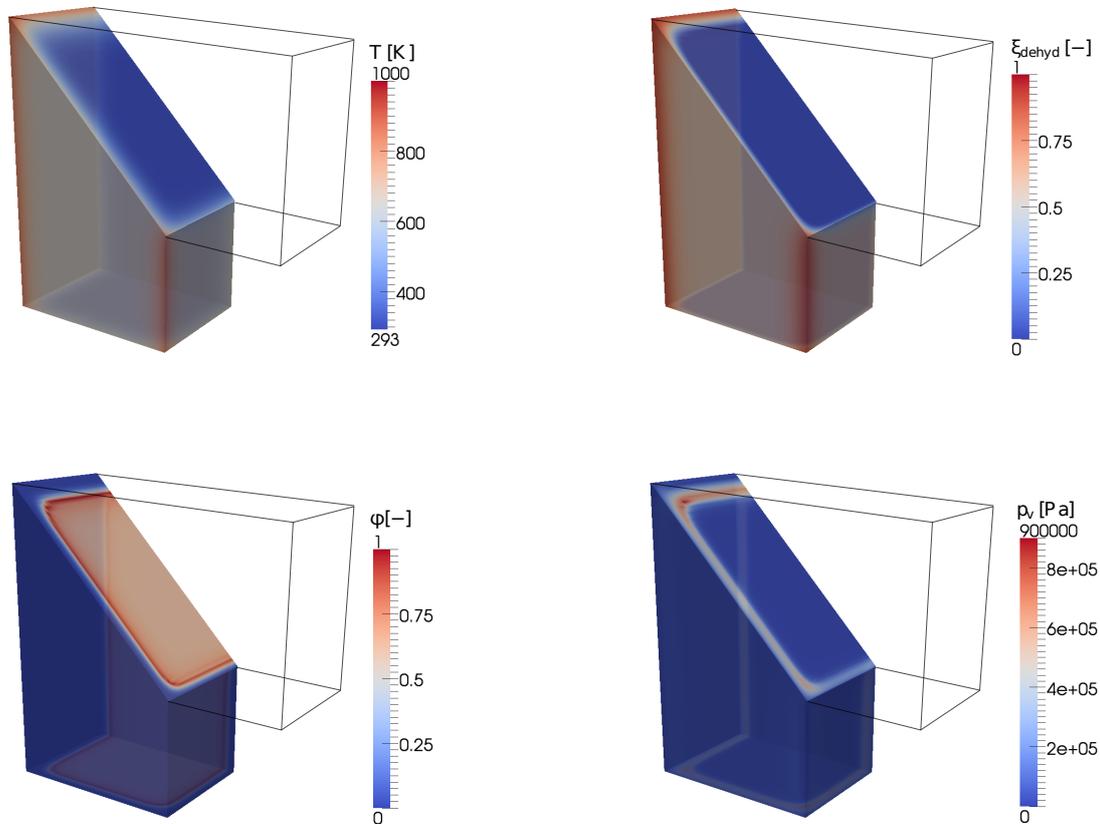


Figure 4: Numerical analysis of a 3-D frame corner - distributions of temperature, degree of dehydration, relative humidity and vapour pressure after 10 min.

5 CONCLUSIONS

A mathematical model for the numerical analysis of hygro-thermal behaviour of concrete structures subjected to high temperatures, including the developing of temperatures and vapour pressure and the damage evolution is presented. The macroscopic balance equations for the conservation of heat, mass of water and vapour and the linear momentum are formulated and important constitutive relations are proposed. The thermo-chemical damage in concrete under high temperatures is considered by non-local damage theory. Finally results of numerical analysis of a 3-D frame corner are discussed. The presented model allows a more detailed understanding of material degradation of concrete at high temperatures. The decrease of load-carrying capacity after fire events may be predicted more realistically.

REFERENCES

- [1] V. Baroghel-Bouny, M. Mainguy, T. Lassabatere and O. Coussy, Characterization and identification of equilibrium and transfer moisture properties for ordinary and high-performance cementitious materials. *Cement and Concrete Research* (1999) **29**:1225–1238.
- [2] D. Gawin, C.E. Majorana and B.A. Schrefler, Numerical analysis of hygro-thermal behaviour and damage of concrete at high temperature. *Mechanics of cohesive-frictional materials* (1999) **4**:37–74.
- [3] D. Gawin, F. Pesavento, and B.A. Schrefler, Modelling of hygro-thermal behaviour of concrete at high temperature with thermo-chemical and mechanical material degradation. *Comp. Meth. Appl. Mech. Engrg.* (2003) **192**:1731–1771 .
- [4] W.G. Gray and B.A. Schrefler, Thermodynamic approach to effective stress in partially saturated porous media. *European Journal of Mechanics-A/Solids* (2001) **20**:521–538.
- [5] W.T. Koiter, General theorems for elastic-plastic solids. *Progress in Solid Mechanics* (1960) 167–221.
- [6] L. Ostermann, and D. Dinkler, Numerical Simulation of the transport and the mechanical behaviour of concrete at high temperatures. *Proceedings EURO:TUN 2007*, Vienna, Austria (2007).
- [7] R.H.J. Peerlings, R. De Borst, W.A.M. Brekelmans and M.G.D. Geers, Gradient-enhanced damage modelling of concrete fracture. *Mechanics of Cohesive-Frictional Materials* (1998) **3**:323–342.

COUPLED SIMULATION OF PROCESS-INTEGRATED POWDER COATING BY RADIAL AXIAL ROLLING OF RINGS

REZA KEBRIAIEI*, JAN FRISCHKORN and STEFANIE REESE

*Institute of Applied Mechanics
RWTH Aachen University
Mies-van-der-Rohe-Straße 1, 52074 Aachen, Germany
e-mail: Reza.Kebriaei@rwth-aachen.de, web page: <http://www.ifam.rwth-aachen.de/>

Key words: Ring Rolling, Finite Element Method, Powder Compaction, Process Control

Abstract. The process-integrated powder coating by radial axial rolling of rings represents a new hybrid production technique that combines the diameter increase of ring-shaped work pieces with the application of functional outer layers [1]. The hot rolling of the ring is strongly influenced by a coupled temperature field since the forming resistance of the ring has to be kept under a certain level. The fact that the coating is produced by powder metallurgy brings up another problem: In conventional ring-rolling processes the infeed of the axial and the radial rollers in order to increase the ring's diameter are controlled by the assumption of volume consistency [6]. But this is not true any more for a ring exhibiting a compressible layer. Thus attention has to be drawn to a different coupling scheme of the roller infeed with the current shape of the ring.

The paper deals with the finite element (FE) modelling of this new process. A parameterized 2D model is used to examine the influence of several parameters on the residual porosity in the layer. Therefore different geometric parameters are examined as well as the influence of the roller and the ring geometry. It will be shown that the implementation of a PID control unit is essential to ensure the roundness of the final ring shape [5]. At the end the simulation results are illustrated.

1 INTRODUCTION

One of the important branches of metal forming processes which originally developed in the mid-19th century is the ring-rolling process. This process is used to manufacture seamless rings that are precisely dimensioned and have circumferential grain flow [2]. In comparison to alternative methods (e.g. flat rolling of sheets and the rolling of long

products) the process of ring rolling is relatively short and simple, due to reduced material handling requirements. In many applications the rolled ring has to be furnished with a wear resistant functional layer around it which could be found for rollers in crushing and briquetting mills used in mineral industry.

Modern techniques for the application of such coatings are the hot-isostatic pressing (HIP), flame spraying and build-up welding. The most effective concepts for wear resistant coatings are based on powder metallurgically produced metal matrix composites (MMC) consist of a high alloyed steel matrix with embedded hard phases (e.g. borides and carbides) [4]. The applied material should have a sufficient toughness in order to avoid sudden fracture under stress. The produced MMCs reach wear resistances up to 50 times higher than the products manufactured by melt metallurgy with no residual porosity, however there exist some disadvantages. The HIP process that is used to apply the compound layer and the rolled ring is a limiting factor for the producible ring size. The largest HIP plants can house rings with a diameter of about 1.5 m. Moreover, the HIP process itself is not energy efficient.

In the development of the new hybrid production process introduced by [1], the main focal point is the integration of the application of the powder metallurgical coatings to large ring shaped parts into radial-axial ring-rolling which overcomes the mentioned limitations. The versatile use of the finite element method (FEM) is crucial for the design of this new process and to investigate important quantities such as e.g. the relative density of the layer. Therefore, the main focus in this paper is put on the parameterized model in order to examine the effects of several geometric parameters (e.g. ring radius, wall thickness and roller radii) on the residual porosity in the layer and the behaviour of the stress in the rolling gaps.

One major problem which is essential for a stable process is the control of the guide rollers [7]. This becomes even more important if the rings consist of a compressible layer. In contrast to solid rings the plastic deformation in the layer is not isochoric. Therefore, new techniques should be applied taking into account the compressibility of the layers.

The present article is structured as follows. In Section 2 the integration of the compaction process of the layer material into ring-rolling process is described. An appropriate material model for the description of the layer material is given in Section 3. Section 4 deals with the FE simulation of the ring-rolling process. The excitation of the guide rollers and their significant influence on ring roundness and ring-rolling stability are discussed in Section 5. Section 6 includes investigated results and the validation of the applied methods. Section 7 is devoted to the concluding discussion.

2 POWDER COMPACTION PROCESS

The standard ring-rolling process starts at temperatures between $1000^{\circ}C$ and $1200^{\circ}C$. A pre-punched blank is formed in the radial and the axial pass. The rotation of the main roller by angular velocity and displacement of the mandrel on the one hand, the friction between ring and rollers on the other hand lead to a continuous decrease of the

ring cross-section and a growth of its diameter. The mentioned set-up is illustrated in FIGURE 1.

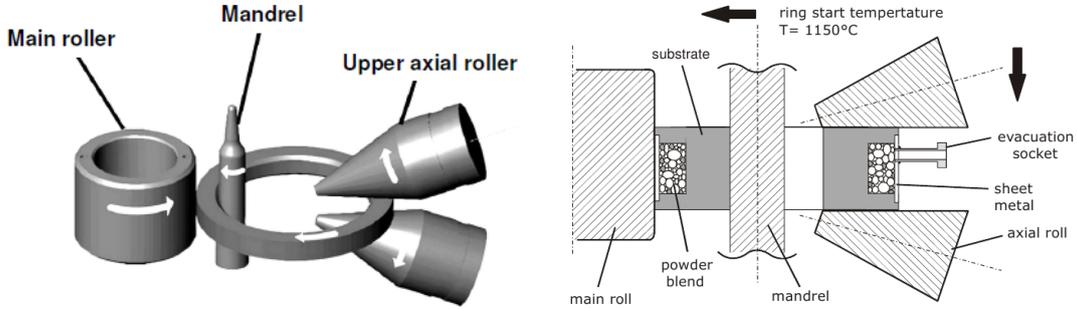


Figure 1: The ring-rolling system (left) evolution of the ring diameter of a capsuled Ring (right) [1]

In the conventional process the layer material is encapsulated by a sheet metal in the rolled ring. Afterwards the layer material is compacted in the HIP plant. In contrast, in the new process the layer material is integrated in the unrolled ring which has been depicted in the right part of FIGURE 1. The heating of the ring to the rolling temperature is used to trigger a sintering of the layer material which results in the formation of sintering necks between the powder particles. After that the ring rolling process is started which leads to a noticeable increase of the diameter and further compaction of the material.

3 MATERIAL MODEL FOR COMPRESSIBLE LAYER

The model of the compressible layer material has been developed by a finite strain elasto-plastic material formulation, in which the rate dependency of the material should be taken into account due to the high temperature [3]. The deformation of a line element $d\mathbf{X}$ during the motion is given by the deformation gradient

$$\mathbf{F} = \frac{\partial \chi(\mathbf{X}, t)}{\partial \mathbf{X}} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} = \text{Grad } \mathbf{x} \quad (1)$$

in which $\mathbf{x} = \chi(\mathbf{X}, t)$ represents the spatial position of a particle in the current configuration Ω at time t . $\mathbf{C}_i = \mathbf{F}_i^T \mathbf{F}_i$ is the inelastic right Cauchy-Green tensor and \mathbf{b}_e given by $\mathbf{b}_e = \mathbf{F} \mathbf{C}_i^{-1} \mathbf{F}^T$. We suppose the free energy per mass ψ could be additively decomposed into two parts of ψ_s and ψ_p in which ψ_s arises from the energy stored in the solid skeleton and the free surface energy ψ_p is due to the porosity of the solid. Therefore the free energy per mass and its material time derivative are

$$\psi = \psi_e(\mathbf{b}_e, \theta) + \psi_i(\kappa, \theta) + \psi_t(\theta) + \psi_p(\rho^r) \quad (2)$$

$$\dot{\psi} = \frac{\partial \psi_e}{\partial \mathbf{b}_e} \cdot \dot{\mathbf{b}}_e + \frac{\partial \psi_i}{\partial \kappa} \dot{\kappa} + \frac{\partial \psi}{\partial \theta} \dot{\theta} + \frac{\psi_p}{\partial \rho^r} \dot{\rho}^r \quad (3)$$

In the above equations ψ_e represents the elastic strain energy, ψ_i is the inelastic energy due to isotropic hardening and ψ_t is the thermal energy. By using the Lie time derivative $\mathcal{L}_\nu \mathbf{b}_e = \mathbf{F} \dot{\mathbf{C}}_i^{-1} \mathbf{F}^T$ and the symmetric property of \mathbf{b}_e we obtain $(\mathbf{F} \mathbf{C}_i^{-1} \mathbf{F}^T)^\dot{=} = 2\text{sym}(\mathbf{l}_{\mathbf{b}_e}) + \mathcal{L}_\nu \mathbf{b}_e$. Using the coaxiality of \mathbf{b}_e and $\partial\psi_e/\partial\mathbf{b}_e$ which is fulfilled by choosing ψ_e as an isotropic tensor function leads to the symmetry of $(\partial\psi_e/\partial\mathbf{b}_e)\mathbf{b}_e$. We finally arrive at

$$\frac{\partial\psi_e}{\partial\mathbf{b}_e} \dot{\mathbf{b}}_e = 2 \frac{\partial\psi_e}{\partial\mathbf{b}_e} \mathbf{b}_e \cdot \mathbf{d} - 2 \frac{\partial\psi_e}{\partial\mathbf{b}_e} \mathbf{b}_e \cdot \left(-\frac{1}{2} \mathcal{L}_\nu \mathbf{b}_e \mathbf{b}_e^{-1}\right) \quad (4)$$

Exploiting the assumption of small elastic volume changes $\det \mathbf{F}_i$ leads to

$$\det \mathbf{F}_i \approx \frac{\rho_0^r}{\rho^r} \quad (5)$$

and¹

$$\dot{\rho}^r = -\rho^r \text{tr} \mathbf{d}_i = -\rho^r \mathbf{I} \cdot \left(-\frac{1}{2} \mathcal{L}_\nu \mathbf{b}_e \mathbf{b}_e^{-1}\right) \quad (6)$$

Let us insert (2), (3), (5) into the entropy inequality in the Clausius-Duhem form² and write

$$\begin{aligned} & (\boldsymbol{\tau} - 2\rho_0 \frac{\partial\psi_e}{\partial\mathbf{b}_e} \mathbf{b}_e) \cdot \mathbf{d} - \rho_0 \left(\frac{\partial\psi}{\partial\theta} + \eta\right) \dot{\theta} \\ & + \rho_0 \left(2 \frac{\partial\psi_e}{\partial\mathbf{b}_e} \mathbf{b}_e + \frac{\partial\psi_p}{\partial\rho^r} \rho^r \mathbf{I}\right) \cdot \left(-\frac{1}{2} \mathcal{L}_\nu \mathbf{b}_e \mathbf{b}_e^{-1}\right) - \rho_0 \frac{\partial\psi_i}{\partial\kappa} \dot{\kappa} \geq 0 \end{aligned} \quad (7)$$

which yields the constitutive equations

$$\boldsymbol{\tau} = 2\rho_0 \frac{\partial\psi_e}{\partial\mathbf{b}_e} \mathbf{b}_e, \quad \eta = -\frac{\partial\psi}{\partial\theta} \quad (8)$$

for the Kirchhoff stress tensor and the entropy, respectively. Additionally we define the so-called effective Kirchhoff stress tensor as

$$\hat{\boldsymbol{\tau}} = \boldsymbol{\tau} - \tau_s \mathbf{I} \quad (9)$$

It contains a positive scalar quantity τ_s called sintering stress. The sintering stress represents the driving force that leads to a reduction of the porosity even in the absence of external loads which is essential in order to model sintering effects. Considering the isotropic hardening governed by drag stress R ,

¹In this part $\mathbf{d}_i := \text{sym}(\dot{\mathbf{F}}_i \mathbf{F}_i^{-1})$ is the inelastic part of the rate of deformation tensor.

²The non-isothermal case of the entropy inequality in the clausius-Duhem form is

$$-\rho_0(\dot{\psi} + \eta\dot{\theta}) + \boldsymbol{\tau} \cdot \mathbf{d} - \frac{1}{\theta} J \mathbf{q} \cdot \text{grad } \theta$$

$$\tau_s = -\rho_0 \frac{\partial \psi_p}{\partial \rho^r} \rho^r, \quad R = -\rho_0 \frac{\partial \psi_i}{\partial \kappa} \quad (10)$$

the residual inequality is given by

$$\hat{\boldsymbol{\tau}} \cdot \left(-\frac{1}{2} \mathcal{L}_\nu \mathbf{b}_e \mathbf{b}_e^{-1}\right) + R \dot{\kappa} \geq 0 \quad (11)$$

The evolution equations for the internal variables

$$-\frac{1}{2} \mathcal{L}_\nu \mathbf{b}_e \mathbf{b}_e^{-1} = \dot{\gamma} \frac{\partial \phi}{\partial \hat{\boldsymbol{\tau}}}, \quad \dot{\kappa} = \dot{\gamma} \frac{\partial \phi}{\partial R} \quad (12)$$

can be derived from a pressure sensitive yield potential ϕ which would be derived as

$$\phi = \delta_\phi \left(\frac{\tau_{y0}}{\tau_{y0} - R}\right)^2 \frac{3}{2} \|\hat{\boldsymbol{\tau}}'\|^2 + \eta_\phi \left(-\frac{1}{3} \text{tr} \hat{\boldsymbol{\tau}}\right)^2 - \tau_{y0}^2 \quad (13)$$

τ_{y0} denotes the initial uniaxial Kirchhoff yield stress. The weighting functions are chosen such that the influence of hydrostatic stress states vanishes if the relative density tends to 1. These weight functions are given as

$$\delta_\phi = 1 + \delta_f (1 - \rho^r)^{\delta_e}, \quad \eta_\phi = \eta_f (1 - \rho^r)^{\eta_e} \quad (14)$$

The parameters $\delta_f, \delta_e, \eta_f$ and η_e control the densification behavior. Therefore, they should be defined in such a way that the modeled values and experimental results coincide with each other. To investigate the densification behaviour of the applied materials the closed die experiment has been applied [4].

In the closed die test the cylindrical specimen with the length of 3 mm and the diameter of 3 mm is considered (FIGURE 2). The specimen is heated in the dilatometer under inductive vacuum atmosphere mode. In order to achieve homogeneous warming of the specimen the temperature will be fixed for about five minutes. After that the specimen will be deformed. Finally the cooling of the specimen to the ambient temperature is carried out by nitrogen deterrence. A press die made of Al_2O_3 is used to apply the deformation and a temperature range of $900^\circ C - 1100^\circ C$ is chosen in this experiment.

In FIGURE 3 the comparison of the experiment (symbols) and simulated (lines) results are illustrated for the different temperatures. The simulations have been carried out at the material point level. Therefore, the deformation and stress state are defined by

$$\mathbf{F} = \begin{bmatrix} \frac{L}{L_0} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \boldsymbol{\sigma} = \begin{bmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix}$$

In FIGURE 3 both of the curves describe σ_{11} with respect to F_{11} . The left diagram is related to the hard work steel (material No. : 1.2344, EN standard: X40CrMoV5-1) and the right one related to the cold work steel (material No. : 1.2380, EN standard: X220CrMo13-4). Both of these materials are applied in industrial applications as MMC functional layers. The figure shows that the correspondence between experiment and simulation is very good.

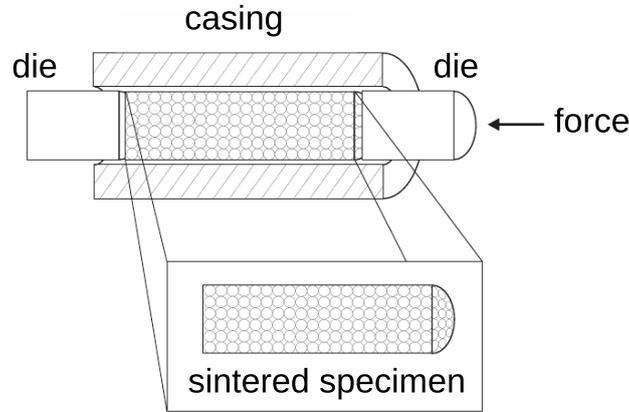


Figure 2: Closed die test with obstruction of the transversal flow [4]

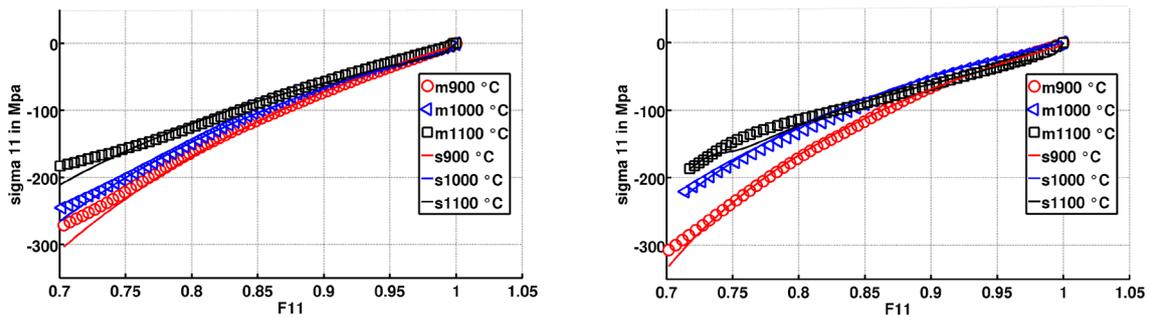


Figure 3: Comparison of the experimental (symbols) and the simulated (lines) results in closed die compaction test

4 FINITE ELEMENT SIMULATION OF RING-ROLLING PROCESS

In this section the finite element model of radial-axial ring-rolling is constructed under the ABAQUS software environment. Due to the large computational effort a parameterized two-dimensional finite element model has been established. We use in this analysis a 4-node bilinear finite element formulation based on reduced integration with hourglass control. It includes displacement and temperature degrees-of-freedom. In order to convergence problems frequently experienced by using implicit procedure an explicit finite element method is applied. The material model described above has been implemented by means of the user interface VUMAT. All simulations shown in this paper are presented at temperatures between 900°C and 1100°C. We focus in this simulation on scrutinizing the influence of geometric parameters (e.g. ring radius, wall thickness, and rigid rollers-radiuses) on the residual porosity and the behaviour of the stress in the rolling gap. The geometric parameters are given in the left part of FIGURE 4 (r_{MR} - radius of main roller,

r_{MA} - radius of mandrel, r_{GR} - radius of guide roller, r_{SU} - outer radius of substrate ring, t_{SU} - thickness of substrate ring, t_{LA} - layer thickness, t_{EC} - thickness of encapsulation). The model can be created automatically by specifying these parameters in the ABAQUS input file.

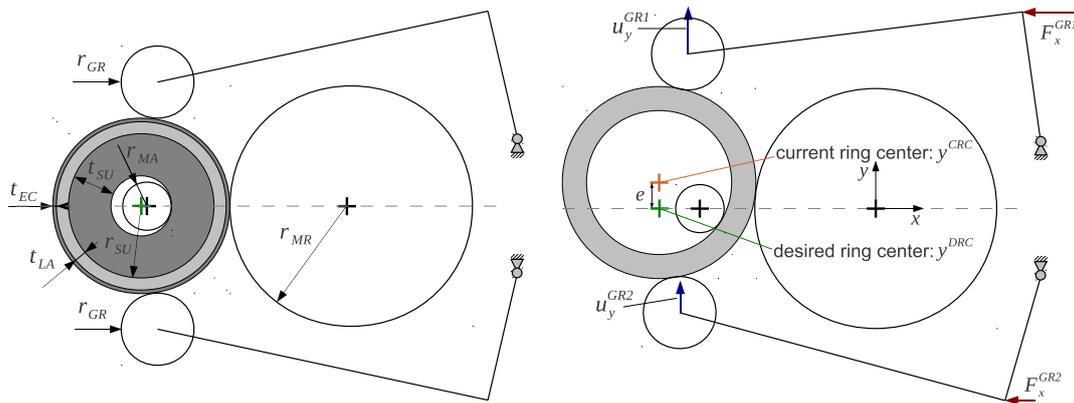


Figure 4: Geometric parameters of the FE ring-rolling model (left), sensor and actuator quantities of the control system (right)

Regarding the small alteration of height in this process the plane strain assumption has been chosen. Of course this is reached in an incremental manner within the real process (increase of the height in the radial rolling gap, reduction of the height in the axial rolling gap). Consequently, the stress state predicted by the simulation will differ from the one that is present in reality. On the other hand FE simulations of this process applying the plane stress assumption have revealed that this assumption is less appropriate. In this case the material flow in thickness direction is overpredicted and the ring growth is significantly underestimated.

5 EXCITATION OF THE GUIDE ROLLERS

Guide rollers have a significant effect on ring roundness and ring-rolling stability. The basic radial design of ring-rolling mills supports the ring only at the roll bite [8]. This leads to the instability of the system. The applied momentum due to the friction applied to the ring in the radial rolling gap leads to an oscillation of the ring and subsequently to a non-circular shape. Therefore, the excitation of the guide rollers that are used to stabilize the position of the ring is an important task. Huge amount of load on the guide rollers causes bad deformation and distortion of the ring and low amount of load tends to instability of the system. In order to determine an appropriate amount of load on guide rollers a control system has been implemented into the FE model. For this purpose the user interface VUAMP has been used.

Within VUAMP the guide roller displacements in y-direction (see right part of FIGURE 4) can be obtained as sensor values in order to compute the deviation of the current ring

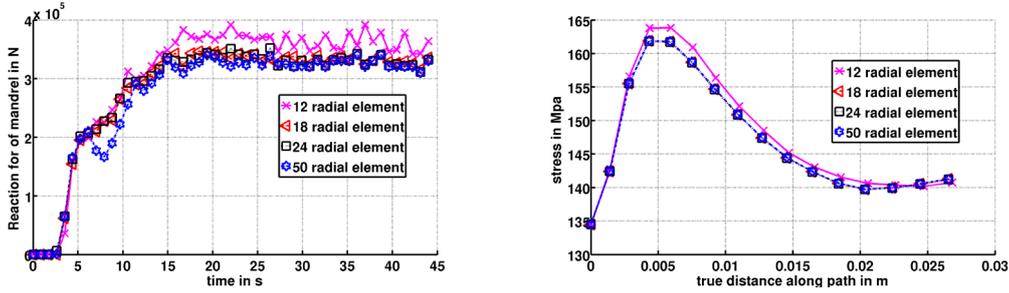


Figure 5: Reaction force of mandrel (left), mises stress on the illustrated path B (right)

center position with respect to the desired position. With this information the forces acting on the guide rollers are computed by a proportional-differential-integral (PID) control unit that has been coded in VUAMP. Then the computed forces are applied on the guide rollers in order to shift the ring back to the required position. With regard to the quantities illustrated in the right part of FIGURE 4 the PID controller minimizes an error given by $e = |y^{DRC} - y^{CRC}|$ with $y_{CRC} = (u_y^{GR1} + u_y^{GR2})/2$. The change of the guide roller forces is then computed by

$$\Delta F_x(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt} \quad (15)$$

For the coefficients K_p , K_i and K_d appropriate values have to be chosen to get a stable position of the ring and to avoid a large steady state error. In a time discrete computational scheme the guide roller forces $F_{x,i+1}^{GR1}$ and $F_{x,i+1}^{GR2}$ of the time increment $i + 1$ are computed by

$$F_{x,i+1}^{GR1} = F_{x,i}^{GR1} - \text{sgn}(y^{DRC} - y_i^{DRC}) \frac{\Delta F_x}{2}, \quad F_{x,i+1}^{GR2} = F_{x,i}^{GR2} + \text{sgn}(y^{DRC} - y_i^{DRC}) \frac{\Delta F_x}{2} \quad (16)$$

In comparison with alternative methods the application of the PID control unit leads to a smaller rise time, less oscillations as well as a much smaller steady state error.

6 RESULTS

In order to work with an appropriate discretization, the convergence of the results with increasing mesh density has been studied. To achieve this aim, two methods are considered. At first the reaction force of the mandrel is deliberated and secondly the mises stress along the illustrated path B in FIGURE 5 is considered. With respect to the FIGURE 5 we can conclude that with more than 18 radial elements the solution is not altered by further mesh refinement.

Next, the stability of the system and the influence of the PID controller on the system should be considered. To see the influence of PID controller on the system two strategies have been chosen. The oscillation of the ring center reveals the preciseness of the applied loads on the guide rollers. To this end, four symmetric points are chosen around the ring

such that the angle between every two consecutive points is 90° . The mean displacement of these four points is equal to the displacement of the ring center. FIGURE 6 (first picture) shows the oscillation of the ring center in Y direction. Regarding to FIGURE 6 (first picture) the small displacement of the ring center could be seen which discloses the proper functionality of the PID controller. Secondly the roundness of the ring has been measured. To see the ring roundness during the process, the average of relative deviation from the mean radius of the ring has been considered. With respect to the nodes at the circumference of the ring (outer side) the current radius of the ring r_i is equal to the initial length of the ring r_0 plus the displacement of the respected node Δr_i^t .

$$\Delta r_i^t = \sqrt{u_{x_i}^t{}^2 + u_{y_i}^t{}^2} \tag{17}$$

$$r_i = r_0 + \Delta r_i^t \tag{18}$$

Therefore, the mean radius of the ring radius is equal to³

$$r_m^t = \frac{1}{n_p} \sum_{i=1}^{i=n_p} r_0 + \Delta r_i^t \tag{19}$$

Consequently the relative deviation from the mean radius of the ring could be computed by

$$errr^t = \frac{1}{n_p} \sum_{i=1}^{i=n_p} \frac{|r_m^t - r_i^t|}{r_m^t} \tag{20}$$

From FIGURE 6 (second picture) it can be concluded that the relative deviation of the ring radius from the mean radius is less than 1 %. This means that the roundness defect of the ring is very small.

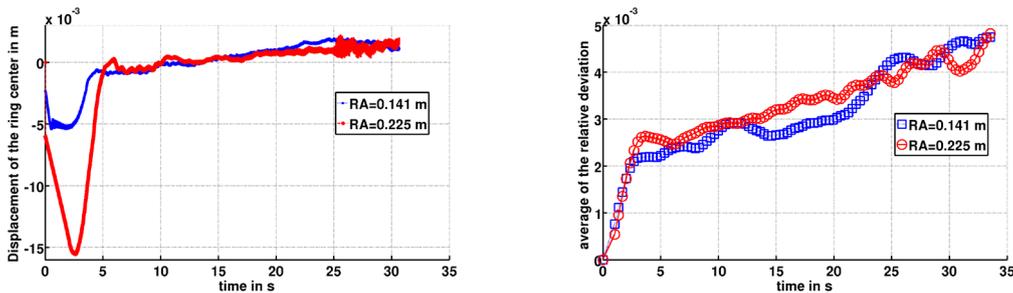


Figure 6: Oscillation of the ring center (left), average of the relative deviation from mean radius of the ring (right)

Afterwards, the behaviour of a compound ring is compared to the one of a solid ring during the rolling process. Therefore we measure the diameter evolution and the rate of

³ n_p is the number of tangential nodes

this evolution throughout the rolling process. We study one solid and compound ring with radius of 0.141 m. For the compound ring the layer thickness of 0.0185 m and a substrate thickness of 0.029 m has been chosen. Therefore, the total thickness is 0.0475 m for both rings. The encapsulation has not been considered in this simulation. Regarding the left part of FIGURE 7 a strong reduction of the ring diameter in the first rolling stage can be observed for the compound ring in contrast to the solid ring. The growth of the compound ring begins after about four seconds. In the FIGURE 7 (right) the rate of diameter evolution with respect to time can be seen.

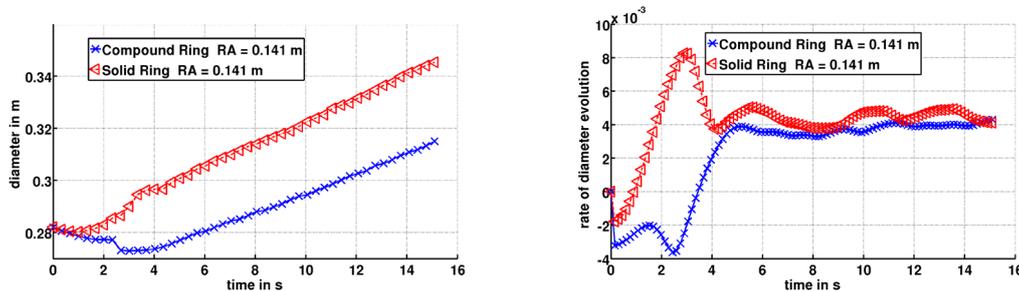


Figure 7: Evolution of the ring diameter for solid and compound ring with respect to the rolling time (left) rate of diameter evolution for solid and compound ring with respect to the rolling time (right)

In order to judge the compaction of the layer, the relative density is considered. FIGURE 8 shows the relative density at point A over time in different systems. The ratio of the main roller radius with respect to the mandrel radius is defined as k . The diagram shows that with $k = 1$ the higher value for relative density could be obtained.

At the end for the both mentioned material described in section 3 the distribution of the relative density inside the layer at a process time of 4 seconds when the ring starts to grow is at the path B from point A considered. With the initial relative density of about 0.7 for the both material, it could be seen that the layer include cold work steel is compacted worse than hard work steel (FIGURE 9). Experimental investigations showed that in some cases the ring growth starts at a lower relative density. This behaviour depends on the difference in the yield stress of the substrate with respect to the one of the solid layer material.

7 CONCLUSION

In this paper a parameterized FE ring-rolling model has been presented that is applicable to the simulation of process-integrated powder coating by radial axial rolling of rings. In addition, a material model that is able to describe the compaction of the powdery layer material has been described. With respect to the experimental evaluations, two different material models are applied and the agreement between simulation and experiment has been shown.

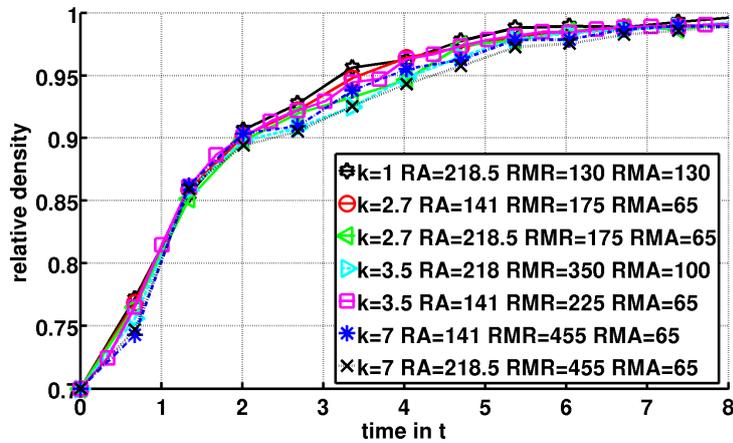


Figure 8: Comparison of relative density in different systems

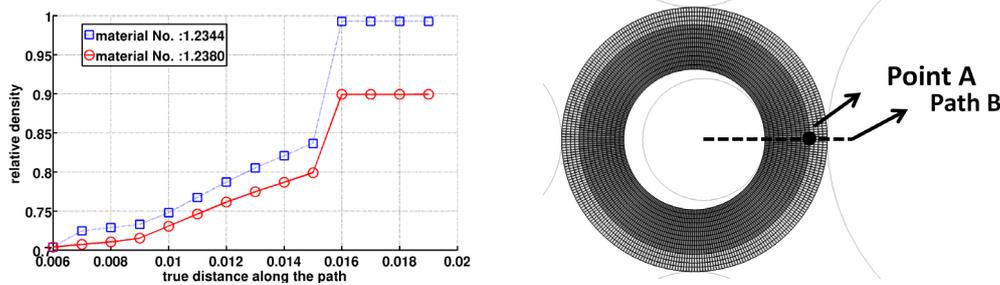


Figure 9: Distribution of the relative density (SDV2) in the layer with the material No.:1.2344, material No.:1.2380 (left), considered path in the rolling gap (right)

The FE model represents a basis to investigate the influence of several geometry parameters on the ring deformation during the rolling and the residual porosity within the layer. Therefore, in order to prevent the locking effects and the mesh dependency of solution the convergence of the selected meshes are studied. Afterwards, the quantities such as e.g. relative density and the diameter increase of the ring are considered and the behavior of a compound ring compared to the one of a solid rings during rolling is demonstrated by simulation results.

By changing the geometric parameters to obtain a stable system which leads to a round final ring shape, it has been shown that a control system is necessary. Such a control system has been programmed and implemented via a user interface.

8 ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of this work by the Volkswagen Foundation under reference number I/81 247 Ak within the initiative 'Innovative Methods for the Manufacturing of Multifunctional surfaces'. All the experimental results presented in

this paper have been performed by the Chair of Materials Technology and the Chair of Production systems at the Ruhr University of Bochum (Germany). Therefore we would like to thank Dr.-Ing. Heiko Moll (Chair of Materials Technology) and Dipl.-Ing. Tubias Husmann for providing the experimental data.

REFERENCES

- [1] H. Moll, W. Theisen, R. Hammelmann, H. Meier "Prozessintegriertes Pulverbeschichten durch Radial-Axial Ringwalzen", *Materialwissenschaft und Werkstofftechnik*, Vol. 38/6, pp. 459, (2007).
- [2] H. Berns and S. Koch, "Influence of abrasive particles on wear mechanism and wear resistance in sliding abrasion tests at elevated temperatures," in *Wear*, edited by B. La, 233-235, 1999, pp. 424-430.
- [3] H. Berns and S. Koch, "Influence of abrasive particles on wear mechanism and wear resistance in sliding abrasion tests at elevated temperatures," in *Wear*, edited by M. D. F. et al., AIP Conference Proceedings 505, American Institute of Physics, New York, 1999, pp. 651-654.
- [4] H. Moll, "Werkstofftechnische Betrachtung des Pulverbeschichtens durch Walzen," *Dissertation*, Ruhr University of Bochum, Germany, (2009)
- [5] J.M. Allwood and A.E. Tekkaya and T.F. Stanistreet, "The Development of Ring Rolling Technology", *Steel research*, Vol. 76, pp. 111, (2005).
- [6] L. Lanyun and H. Yang and L. Guo and Z. Sun, "A control method of guide rolls in 3D-FE simulations of ring rolling," *Materials Processing Technology*, Vol. 205/1-3, pp. 99-110, (2008).
- [7] M. P. Brown, and K. Austin, *The New Physique*, Publisher Name, Publisher City, 2000, pp. 212-213.
- [8] M. P. Brown, and K. Austin, *Appl. Phys. Letters* **85**, 2503-2504 (2000).
- [9] R. Wang, "Title of Chapter," in *Classic Physiques*, edited by R. B. Hamil, Publisher Name, Publisher City, 2000, pp. 212-213.

EVALUATION AND IMPROVEMENT OF THE THM MODELLING CAPABILITIES FOR ROCK SALT REPOSITORIES

ALEXANDRA PUDEWILLS

Institut für Nukleare Entsorgung (INE), Karlsruher Institut für Technologie (KIT)
Campus Nord, 76344 Eggenstein - Leopoldshafen, Germany
e-mail: alexandra.pudewills@kit.edu, www.kit.edu

Key words: Coupled Problems, THM Modeling, Rock Salt, Laboratory Tests, Nuclear Waste Repositories.

Abstract. *This paper provides a summary description of the selected results obtained in the frame of the THERESA- project cosponsored by the European Commission (EC). The numerical modeling of coupled thermal-hydraulic-mechanical (THM) processes with impact on repository long-term safety was focused on simulation of a number of representative laboratory experiments on rock salt samples. The scope of these calculations was the validation of the actual capabilities of the constitutive model and to identify needs for further improvements of the model. The measured development of volumetric strains and permeability in the samples during loading process was compared posterior to calculation results.*

1 INTRODUCTION

The objectives of the European Community project THERESA ("Coupled thermal-hydrological-mechanical-chemical (THMC) processes for application in repository safety assessment") [1] are to develop, verify and improve the modelling capabilities and capacities of mathematical models and computer codes for coupled processes for use in performance assessment of the long-term safety of nuclear waste repositories in crystalline rocks and rock salt. The Work Package 3 (WP3) of this project focused on the evaluation and improvement of numerical modelling capabilities for assessing the performance and safety of nuclear waste repositories in rock salt, with particular regard to the long-term evolution of the excavation damaged zone (EDZ), considering thermal-hydraulic-mechanical processes.

Furthermore in this Work Package triaxial compressive test on rock salt samples and a large-scale test on a hollow salt cylinder were performed and then considered for an unconventional benchmark analysis. The evolution of the THM processes occurring during these tests was calculated using the finite element code ADINA [2]. A new viscoplastic constitutive model for rock salt that can describe the volumetric strain (dilatancy) and the damage of the rock has been proposed and implemented in this code. The rock damage (i.e. micro cracks or fractures) was judged by criteria for shear and tensile fracture and by the dilation criterion. The main attention focused on the simulation of strains and permeability

development the rock salt samples. Subsequently, the calculated strain rates were compared to experimental data.

2 CONSTITUTIVE MODEL FOR ROCK SALT

The constitutive model proposed in this project is based on the assumption of small strains, where the total strain rate, $\dot{\epsilon}_{tot}$ is split into elastic and viscoplastic parts as follows:

$$\dot{\epsilon}_{tot} = \dot{\epsilon}_{el} + \dot{\epsilon}_{vp} \quad (1)$$

$\dot{\epsilon}_{el}$ elastic strain rate tensor

$\dot{\epsilon}_{vp}$ viscoplastic strain rate tensor

The elastic behaviour is assumed to be time-independent. Furthermore, the viscoplastic strain rate tensor is decomposed into a viscoplastic strain rate tensor by constant volume and a viscoplastic strain rate tensor due to damage that considers the volume change, such as dilatancy or compaction of the material:

$$\dot{\epsilon}_{vp} = \dot{\epsilon}_{vp}^c + \dot{\epsilon}_{vp}^d \quad (2)$$

$\dot{\epsilon}_{vp}^c$ viscoplastic strain rate without volume change

$\dot{\epsilon}_{vp}^d$ viscoplastic strain rate due to damage which describes a volumetric strain

For each viscoplastic strain rate, an associated flow rule is used:

$$\dot{\epsilon}_{vp} = \gamma \langle \Phi (F(\sigma)) \rangle \partial F / \partial \sigma \quad (3)$$

where

$\gamma = a_1 \exp(-a_2 / T)$ is the fluidity parameter,

a_1 and a_2 are material constants and T is temperature;

The term $\Phi (F)$ denotes a monotonic function of the yield function (F). The meaning of the brackets $\langle \rangle$ is as follows:

$$\begin{aligned} \langle \Phi (F) \rangle &= 0 && \text{if } F \leq 0 \\ \langle \Phi (F) \rangle &= \Phi (F) && \text{if } F > 0 \end{aligned} \quad (4)$$

The function $\Phi (F)$ is defined as:

$$\Phi (F) = (F - F_0)^m \quad (5)$$

where m is an arbitrary constant and F_0 is the uniaxial yield stress and set to zero for instance. For our viscoplastic model, the functions F^c and F^d are defined as follows:

$$F^c = q^2 \quad (\text{without volume change}) \quad (6)$$

$$F^d = n_1 p^2 + n_2 q^2 \quad (7)$$

where

p is the mean stress and q is the standard stress deviator,

n_1, n_2 are material functions of the volumetric strain, ϵ_{vol} ,

and expressed as:

$$n_1 = c_1 (q^2/p^2 - c_2 (\eta_0 + \epsilon_{vol}) / (1 + \epsilon_{vol})) \quad (8)$$

$$n_2 = 1 - c_3 \cdot n_1 p^2 / q^2 \quad (9)$$

with c_1, c_2 , and c_3 being material constants to be evaluated by laboratory tests. In the present approach η_0 is the initial porosity of the undamaged rock salt.

This viscoplastic material model for damage is based on the mathematical formulation proposed by Hein [3] for granular materials, such as crushed salt, and was implemented in the finite-element code ADINA [4]. Separate criteria for shear and tensile fracture and a compression-dilation boundary [5, 6] are available to judge the damage of rock salt (i.e. micro-cracks or fractures):

- Shear stress criterion for compression

$$\tau_f \geq b |\sigma_m|^p \tag{10}$$

where

- τ_f : predicted shear stress at failure
- σ_m : mean stress
- b and p are fitting parameters.

- Tension-induced failure is assume if the max. principal stress exceeds a tension of 1MPa.

- Compression-dilation boundary

$$\tau_{cd} \geq f_1 \sigma - f_2 \sigma_m^2 \tag{11}$$

where f_1 and f_2 are fitting parameters.

The permeability of the damage rock salt is correlated to the dilatant volumetric strain according to the following preliminary function reported in the literature [7]:

$$k = A \cdot \varepsilon_{vol}^B \tag{12}$$

where

- k : permeability of rock salt;
- A and B are material parameters given by different authors and types of the rock salt.

Recently, a new permeability model for rock salt has been proposed by [8] and is presented below together with the material parameters. This model represents a relation between the mean stress and rock porosity:

$$k = \frac{k_{tp}}{\left(\frac{\phi}{\phi_{tp}}\right)^{-p_1} + \left(\frac{\phi}{\phi_{tp}}\right)^{-p_2}} \tag{13}$$

with

$$k_{tp} = a_k \cdot \exp(-p_c \cdot b_k)$$

$$\phi_{tp} = a_\phi \cdot \exp(-p_c \cdot b_\phi)$$

3 MODEL CALIBRATION

In order to adjust the material parameters and to demonstrate the applicability of the material model to describe the dilatant volumetric strain of rock salt, a number of different triaxial laboratory tests were investigated numerically. The calculated strain rates were compared to experimental data. The influence of different material parameters on the numerical results was studied and described in Reference [9]. In the framework of the

THERESA project some new numerical analyses of various laboratory transient-creep experiments [10] have been performed. A test conducted on rock salt from the Asse mine for which the volumetric strain rate was measured has been numerically analysed. The cylindrical specimen had a diameter of 86.5 mm, a length of 175.0 mm, and a density of 2160 kg/m³ at the beginning of the test and was subjected to an axial compression with controlled strain rates of 10⁻⁷ 1/s and a horizontal confining pressure of 1MPa. In this analysis it was assumed that the dilation of the samples starts immediately after loading (i. e. the short time elastic compaction of the sample was not modelled). A comparison of the measured and calculated strain-stress curves, as well as the development of volumetric strains, are presented in the Figures 1 and 2, respectively.

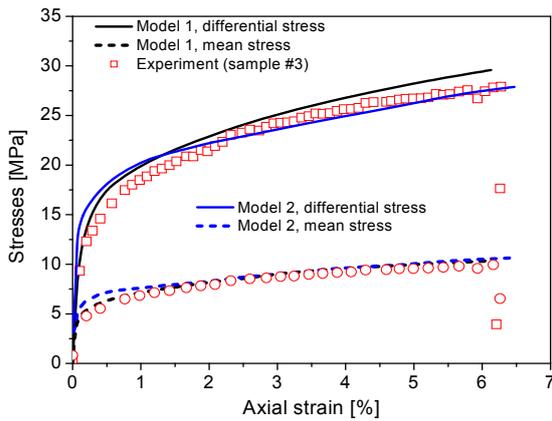


Figure 1: Development of stresses as a function of axial strain; Model 1, with damage parameters fitted on strength tests (strain rate of 1E-05 1/s), Model 2, new parameters set.

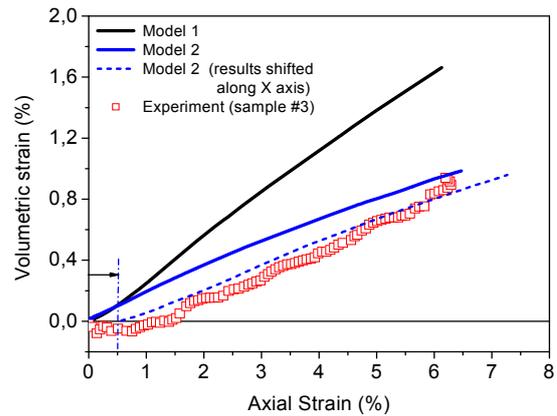


Figure 2: Comparison of measured and calculated volumetric strains for two sets of parameters.

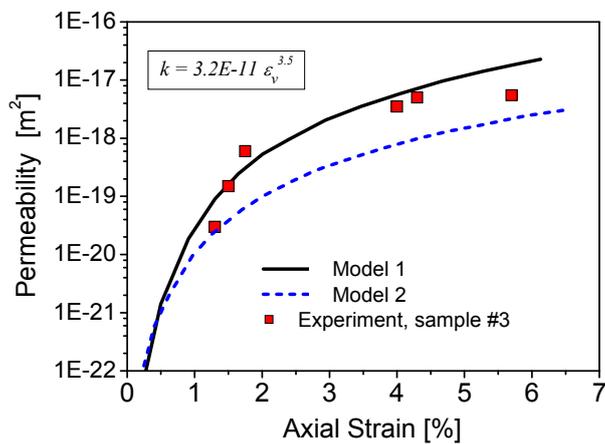


Figure 3: Permeability of the sample during the transient creep test and the calculated evolution of the permeability for two sets of parameters of the calculation results along the horizontal axis lead to a good overall correlation to measurements (Model 2, dashed curve).

Using the calibrated parameters on earlier laboratory transient creep tests the model predicts the development of effective and mean stresses quite well, but the volumetric creep strain over-estimates the actual measurements (Model 1). With a minor adjustment of the damage parameters it can be recognised that the results obtained are in better agreement to the experimental data (Model 2).

The porosity-permeability relation described by the equation (12) was used to calculate the development of permeability of the sample. Figure 3 shows the permeability of rock salt calculated from the obtained volumetric strain in comparison with the laboratory measurements. In Tab. 1 are the parameters obtained from the model calibration summarized.

Table 1: Material parameters for rock salt

Properties	Parameters
Thermo-elastic properties	$E = 27 \text{ GPa}; \nu = 0.25; \alpha = 4.2\text{E-}05 \text{ 1/K}$
Transient creep	$a_0 = 0.018; a_1 = 240; a_2 = 0.112;$ $Q_c / RT = 6495$
Viscoplastic damage	$m = 2.25; c_1 = 0.3; c_2 = 400;$ $c_3 = 2; \eta_0 = 0.02\%$
Hydraulic properties Equations (12) and (13)	$A = 3.2\text{E-}11; B = 3.5$
	$p_1 = 4; p_2 = 1; a_k = 4.27\text{E-}14;$ $b_k = 1.26; a_\phi = 0.0263; b_\phi = 0.3093$

4 LABORATORY TEST CASE

In order to demonstrate the capabilities and suitability of the calibrated models to calculate EDZ evolution and reconsolidation of the damaged salt, a laboratory benchmark test was designed and conducted [10]. A schematic representation of the triaxial test cell is shown in Figure 4. The test was performed in two steps (test case TC2A and TC2B, respectively). After two load steps at 11 MPa and 19 MPa deviatoric stress of the sample without permeability increase, the sample was removed from the apparatus and inspected. Afterwards, the sample was installed again and the experiment was restarted with several loading steps, a reconsolidation phase and a phase of elevated temperature. The evolution of the stresses during the second phase (TC2B) is presented in Figure 5. More details are given in the THERESA laboratory test report [10].

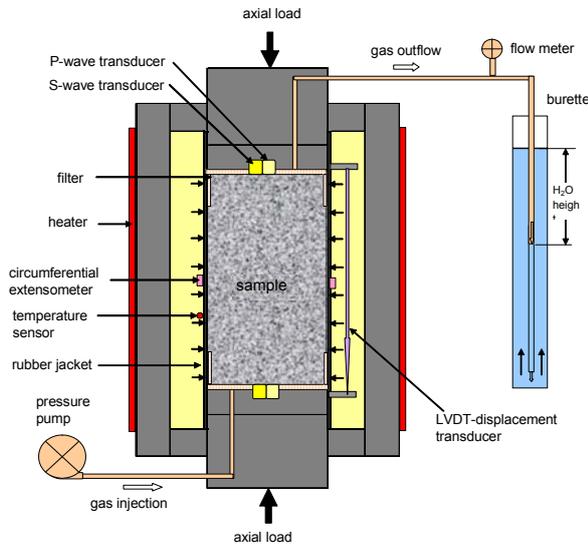


Figure 4: Schematic assembly of a salt sample in the triaxial cell [10].

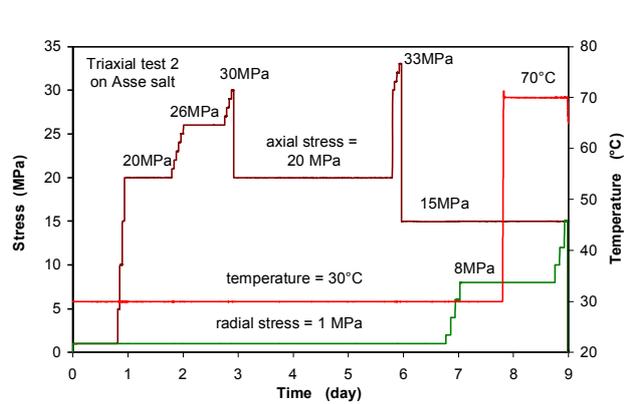


Figure 5: Load conditions applied to the benchmark sample in phase TC2B [10].

5 NUMERICAL SIMULATION

A triaxial compressive test (TC2B) conducted on Asse rock salt was proposed for the benchmark analyzes. The evolution of the THM processes occurring during those tests was simulated using the ADINA code. The attention was focused on the simulation of strains and permeability development in the rock salt samples. Subsequently, the calculated strain rates were compared to experimental data. The main objectives of the benchmark calculations are the validation of the actual capabilities of the constitutive model and to identify needs for further improvements of the model.

5.1 Description of the model and boundary conditions

The model geometry used for the analysis is the same as that of the experimental cylindrical rock salt sample with a diameter of 100 mm and a height of 190 mm. The initial and boundary conditions are assumed as follows:

- Thermal: The whole sample was kept at room temperature of 30°C for about 8 days after this period the temperature was rapidly increased to 70°C and kept constant until the test ends.
- Hydraulic: Boundaries are impervious except for the both top boundaries to permit gas in- and outflow, where gas pressure at the bottom of 0.5 MPa and the atmosphere pressure at top were imposed.

- Mechanical: At the top boundary a time dependent vertical stress and at the outer cylinder surface a radial stress was applied. The bottom surface of the model was fixed in the normal and horizontal directions.

Details of the loadings are given in Figure 5 and the laboratory tests report [10]. The parameters adopted in the simulation coincide with those assumed in the previous model calibration and are summarized in Tab. 1.

5.2 Calculation results and discussions

The comparison of the calculated strains with the strains measured is shown in Figure 6 on which the symbols are the experimental data and the continuous lines are the response of the model. There are some differences during the second creeping stage due to the fast increase of the vertical stress from one to 20 MPa. During the next moderate increase and decrease of the loading stages there is a better agreement between the experimental and predicted strains. In our case the calculated salt permeability using both available relations shows differences in the first loading stages (Figure 7) because of the assumed initial porosity of the salt sample as near zero ($\eta_0=0.0002$). Taking into account that the initial porosity of the sample was about 0.2% and the corresponding permeability of about 10^{-19} m^2 , a new simulation will match quite well with the measured behaviour. The increasing temperature and also the confining pressure have different influences on the actual results. Unfortunately, it must be mentioned that at the moment there are too few experimental results available to be sure that our assumptions are correct.

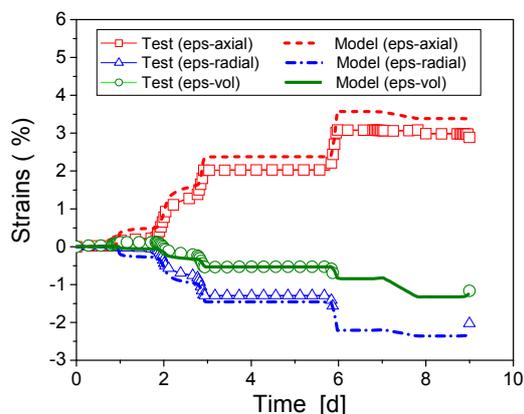


Figure 6: Comparison of measured and predicted axial, radial and volumetric strains.

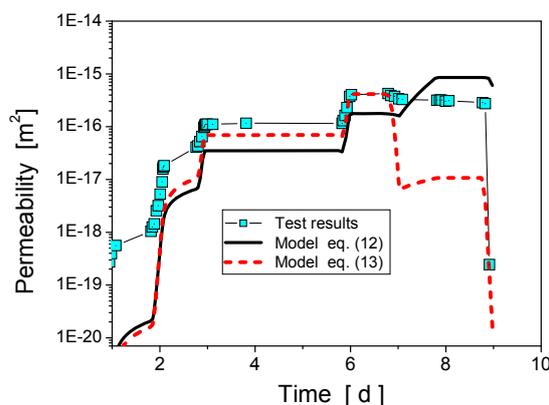


Figure 7: Permeability of the sample during the transient creep test TC2B and the calculated evolution of the permeability with two different relations.

The distributions of void volume fraction and strains in the model at the time $t = 8.3$ days are shown in Figure 8. Due to the experimental boundary condition at the upper and lower surfaces of the sample, large strain gradients and the failure of rock salt were calculated. This is not surprising as the case under study is characterized by small dimensions and a rapid change in stress loading and unloading. The effect will also influence the hydraulic behaviour of the sample material at these locations. It must be pointed out that the numerical model used

is a simplified hydro-mechanical formulation based on Darcy' law and a saturated porous media which are questionable with regard to gas flow through a viscous rock at high temperature.

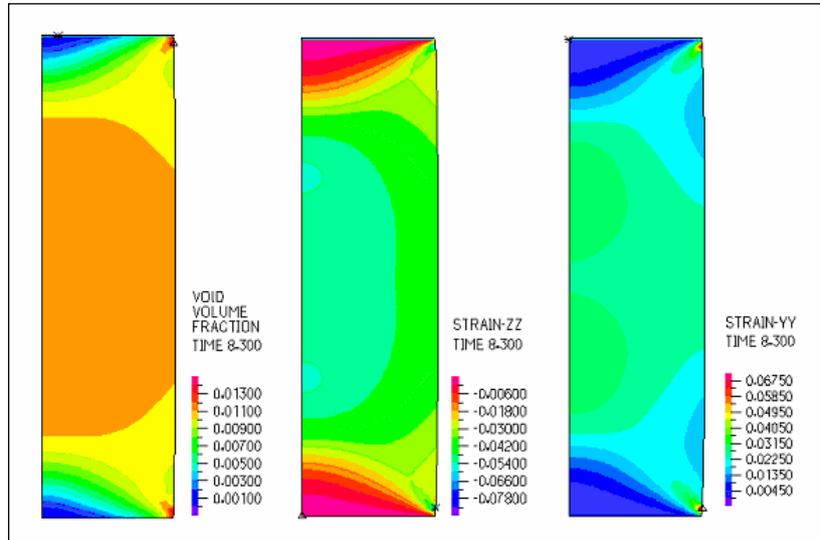


Figure 8: The spatial distributions of the void volume fraction, axial and horizontal strains after 8.3 days.

12 CONCLUSIONS

The numerical simulation results of the laboratory benchmark tests, performed on rock salt have shown a quite good agreement with experimental data. The mathematical formulation of the constitutive model is simple and can be easily implemented in different numerical codes. The model is suitable to describe the main hydro-mechanical behaviour of the rock salt such as transient creep, volumetric strain and material damage. In the case of the sealing (re-consolidation) of the damaged salt, there is still a considerable difference between the model and experimental data. A future improvement of the proposed model is required.

ACKNOWLEDGEMENTS: The THERESA project was co-funded by EC under contract No. FP6-036458.

11 REFERENCES

- [1] THERESA project: "Description of Work", Annex 1 to the project contract FP6-036458, 2007.
- [2] Adina R & D Inc., ADINA (Automatic Dynamic Incremental Nonlinear Analysis), Report ARD 01-9, Watertown, MA, US, 2008.
- [3] Hein, H. J.: Ein Stoffgesetz zur Beschreibung des thermomechanischen Verhaltens von Salzgranulat, Dissertation, RWTH Aachen, 1991.

- [4] Pudewills, A., Krauss, M.: Implementation of a viscoplastic model for crushed salt in the ADINA programme, *Computers and Structures*, vol. 72, pp. 293-99, 1999.
- [5] Hunsche, U.E. (1992). Failure behaviour of rock salt around underground cavities. In *Proceedings of the 7th International Symposium on Salt, Kyoto, 1992*. Amsterdam: Elsevier.
- [6] Cristescu N. and Hunsche, U.: *Time effects in rock mechanics*, John Wiley & Sons, 1998.
- [7] Bechthold, W.; Rothfuchs, T.; Poley, A.; Ghoreychi, M.; Heusermann, S.; Gens, A., and Olivella, S.: *Backfilling and sealing of underground repositories for radioactive waste in salt (BAMBUS project): final report*, Luxemburg: Office for the Official Publications of the European Communities, 1999.
- [8] Heemann, U., Heusermann, S.: *Theoretical and Experimental Investigation on Stresses and Permeability in the BAMBUS Project*. DisTec 2004, Int. Conf. on Radioactive Waste Disposal, April 26-28, Berlin, 2004.
- [9] Pudewills, A.: *Numerical modelling of the long-term evolution of EDZ: development of material models, implementation in finite-element codes, and validation*. FZKA-7185, Karlsruhe: Forschungszentrum Karlsruhe, 2005.
- [10] Zhang, C.-L, K. Wiczorek, and T. Rothfuchs: *THERESA Project, Work Package 3, Deliverable D7, Laboratory Benchmark Tests on Rock Salt*. Brussels: European Commission, 2009.

MASS, HEAT AND MOMENTUM TRANSFER IN NATURAL DRAFT WET COOLING TOWER WITH FLUE GAS DISCHARGE

ADAM F. KLIMANEK* AND RYSZARD A. BIALECKI*

*Institute of Thermal Technology
Silesian University of Technology
ul. Konarskiego 22, 44-100 Gliwice, Poland
e-mail: ryszard.bialecki@polsl.pl, www.itc.polsl.pl

Key words: CFD, Coupled Heat, Mass and Momentum Transfer, Cooling Towers

Abstract. The paper presents CFD simulation results of a natural draught wet-cooling tower (NDWCT) with flue gas discharge. The problem considered is mixing of the flue gases with the rising plume and possible corrosion of the tower shell due to acid condensate. A previously developed CFD model of a NDWCT has been used in the analysis. No wind conditions have been assumed and the results have shown that under this condition the corrosion is unlikely to occur.

1 INTRODUCTION

Cooling towers are devices used to cool industrial water. In conventional coal-fired power plants, they cool the water coming from the condenser of the steam turbine. The aim of the tower is to keep the cooled water temperature low, since this temperature affects the efficiency of the power plant and thus the fuel consumption and pollutant emissions. The most frequently encountered natural draught cooling towers are massive structures whose height reaches up to 200 m [1, 2]. The water to be cooled is sprayed on the top of a heat and mass exchanger called a fill (packing) flowing downwards in a form of a thin film. Evaporation and heat exchange between the film and the air are responsible for the cooling effect of the water. Free-fall rain zone below the fill, allows airflow into the tower. Finally the droplets traveling downward the rain zone are collected in the water basin from where the water is pumped back to the condenser. Increased humidity and (and usually higher temperature) reduce the density of the air in the tower and form the draught of the tower. From the viewpoint of mass, heat and momentum transfer the three zones: spray, fill and rain are the most important regions of the cooling tower. In Figure 1 the three transfer zones are presented. Two groups of models are used to simulate the operation of natural draught towers:

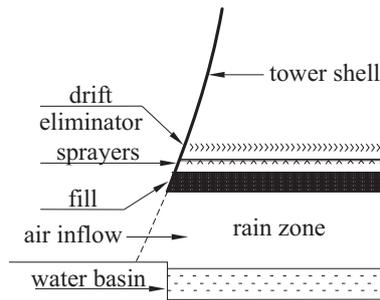


Figure 1: The heat, mass and momentum transfer zones in the cooling tower

- Zero or one-dimensional models (global models) used in both design calculations and performance tests. These models work on average flow and thermal parameters of cooling towers.
- Two or three-dimensional models making use of the Computational Fluid Dynamics (CFD) codes. These models can be used in both design calculations and performance tests, however since they are more computationally demanding, they are used when the local flow and thermal parameters of cooling towers need to be determined.

Combustion exhaust gases cleaned in wet type desulphurization plants are too cold (50 - 80 °C) to be introduced without preheating into the chimney of a power plant. These gases can either be reheated and directed to the chimney or introduced into the cooling tower and dispersed in the atmosphere with the plume. The latter solution is extremely interesting due to the lack of a reheating system, which reduces the investment cost of a newly built desulphurization plant. When this approach is used in newly built power plants, the stack does not need to be erected at all. Figure 2 shows a simplified diagram of a power plant equipped with the desulphurization plant transferring the cleaned gases to the cooling tower.

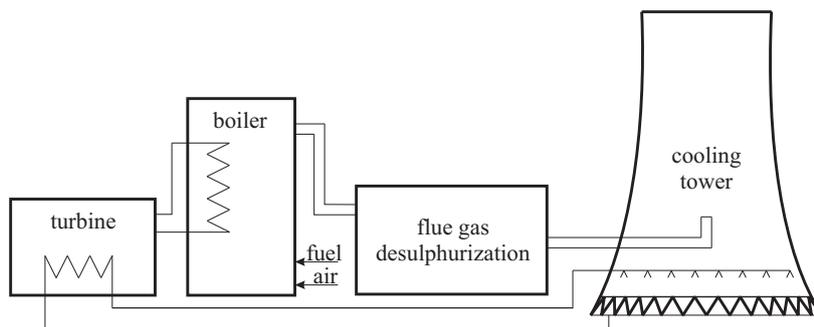


Figure 2: Schematic diagram of a power plant with flue gas discharge

The first plant using the flue gas discharge into the cooling tower was built in Völklingen, Germany in 1982 as a model (prototype) version of this solution [3]. Introduction of more strict regulations regarding SO₂ emissions required installation of desulphurization plants in both existing and newly built power plants in order to meet this regulations. In the last 30 years a couple dozen of retrofitted (e.g. RWE Niederaussem, Germany; Vattenfal Jämschwalde and Boxberg, Germany; PKE Jaworzno III, Poland) and newly built plants (e.g. RWE Niederaussem, Germany [1, 2]; PGE Belchatów, Poland; PKE Lagisza, Poland) used the concept of flue gas discharge into the cooling tower [4].

Despite the known aforementioned advantages, the problem of corrosion of the inner part of the tower shell due to acid condensate from the flue gas is of concern. This can be solved by coating the internal part of the shell by acid resistant material e.g. epoxy resin, however it generates additional cost. Also questions arose how the introduction of flue gases affect the cooling efficiency of the tower and how well the flue gases are mixed with the plume? There is also a risk that under wind conditions the flue gases would not mix properly with the rising plume. They could then flow along the tower shell and the remaining sulphur oxides would be the reason of corrosion of the tower shell if not protected. Another question is how the flue gases are dispersed in the atmosphere? This question is specifically important under strong wind conditions. Under normal and mild wind conditions the towers are known to disperse the gases even better than stacks [4]. The first applications of flue gas discharge into the cooling tower used specially designed mixers to increase diffusion of the flue gas in the plume [4]. In new cooling towers, however, the inflow is through a gas duct placed centrally a few meters above the fill.

To address these questions a previously developed and validated CFD model of a natural draught wet-cooling tower [5, 6, 7] has been adapted to take into account the flue gas discharge into the tower. In this paper, as a preliminary calculations, only the mixing effect of the introduced gases with the rising plume is examined.

2 CFD MODEL OF THE COOLING TOWER

The object under consideration is a 120 m high natural draught wet-cooling tower built in PKE Jaworzno III power plant in Poland. The flue gases are introduced in the center of the tower 25.86 m above the ground level by 7 m diameter duct as shown schematically in Figure 3 (left). The duct is introduced to the tower at the level of 18 m through an opening in the tower shell. The geometry of the model encompasses both the tower and the surrounding atmosphere comprising a 300 m high cylinder of 200 m diameter, as shown in Figure 3 (right). The model of the tower has been developed using a commercial CFD code ANSYS Fluent. The coupled heat, mass and momentum transfer of the multiphase flow in the cooling tower is solved by the use of built-in functions of the CFD code and additional models developed by the authors of this paper. The developed models are incorporated via the User Defined Functions mechanism. The need of special treatment of the heat, mass and momentum exchange in the spray, fill and rain zones is induced by the large scale difference of the surrounding atmosphere ($O(10^3)$ m) and

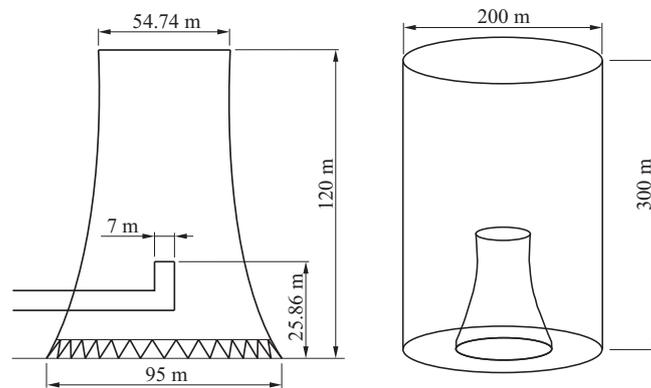


Figure 3: Main dimensions of the cooling tower (left) and the computational domain (right)

the internal equipment of the tower ($O(10^{-2})$ m). Since the modeling approaches were already published in [5, 6, 7] they will not be presented here in detail but will be briefly discussed.

The heat and mass transfer in the spray and fill zones are expressed in terms of two point boundary value problem [7, 8]. For boundary conditions that are determined from the CFD code, results of the equations in the fill are the distributions of mass and heat sources. These sources are plugged into appropriate transport equations of the CFD code. The momentum transfer in the spray and fill zone is accounted for by treating them as a porous medium. Substantial acceleration of the computations is achieved by employing an original technique based on Proper Orthogonal Decomposition Radial Basis Function (POD-RBF) network [7, 9]. The final result of this technique is a vector-matrix product defining the distribution of the heat and mass sources in the fill.

The heat, mass and momentum transfer in the rain zone is simulated using the Euler-Euler multiphase model available in the Fluent code. The standard functionality of this model has been extended to account for mass transfer.

The geometry of the model has been created using the Gambit preprocessor and the generated mesh consisted of 5.12 million structured and unstructured grid cells. The mesh has been verified in terms of quality.

2.1 Governing equations

The equations solved in the CFD code are the mass, momentum and energy transport equations. Additionally species transport equations are solved for O_2 , CO_2 , SO_2 and H_2O (vapor). The concentration of N_2 is inferred from the concentrations of the remaining species. The multiphase flow in the rain zone of the cooling tower is solved using the Euler-Euler approach. This requires solution of the mass, momentum and energy transport equations for the liquid phase as well. The Standard $k - \epsilon$ turbulence model is used to close the system of equations. The option Dispersed is used to account for the turbulence of the secondary phase. The governing equations are summarized in Table 1 The governing

Table 1: Transport equations solved in Fluent

Equation	# of equations	Applied to
Continuity	2	both phases – air and gas, liquid water
Species transport	4	primary phase – O ₂ , CO ₂ , SO ₂ and H ₂ O
Momentum	6	both phases – air and gas, liquid water
Energy	2	both phases – air and gas, liquid water
Turbulence	2	primary phase – air and gas

equations are solved in a steady state with second order discretization schemes in space.

2.2 Boundary conditions

The pressure inlet and pressure outlet boundary conditions are assigned to the side boundaries and the top of the cylinder like domain, respectively. The pressure and temperature are assumed constant at that boundaries. This also infers that no wind conditions were assumed at this stage. The velocity inlet boundary condition is used for the flue gas introduction. The summary of the input data assigned at the boundaries is presented in Table 2.

Table 2: Air, water and flue gas parameters

Atmospheric air parameters		
Air temperature	T_a , °C	7.11
Relative humidity	φ , %	53.45
Pressure	p , hPa	996.1
Water parameters		
Hot water temperature	T_w , °C	29.22
Water mass flow rate	m_w , kg/s	14876
Flue gas parameters		
Gas temperature	T_g , °C	70
Gas mass flow rate	m_g , kg/s	649.2
Composition:		
O ₂	z_{O_2} , mol/mol	0.06
CO ₂	z_{CO_2} , mol/mol	0.135
H ₂ O	z_{H_2O} , mol/mol	0.08
SO ₂	z_{SO_2} , mol/mol	10^{-4}
N ₂	z_{N_2} , mol/mol	0.7249

3 RESULTS AND DISCUSSION

In order to determine the effect of flue gas mixing with the rising plume the cooling process has been simulated using the methods described above. The obtained velocity and temperature fields are presented in Figure 4. As can be seen the velocity of the flue

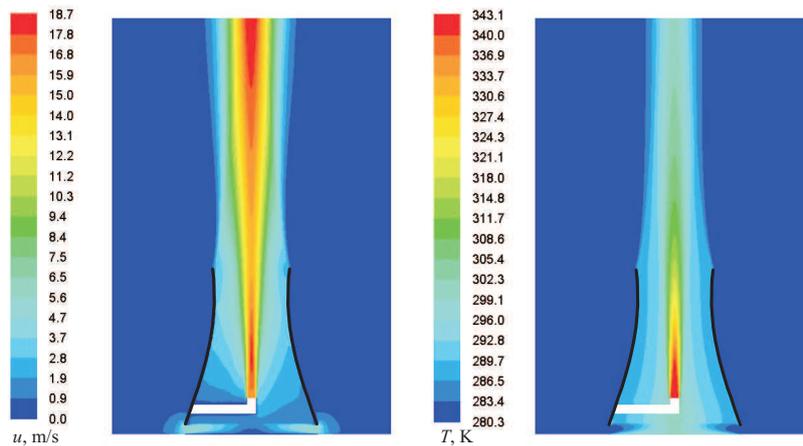


Figure 4: Contours of velocity (left) and temperature (right)

gases entering the tower dissipates quickly and then increases again within the plume. The initial raise of the plume velocity above the tower outlet, can be explained by the low pressure that is formed inside the tower. The lowest pressure is just above the fill. Due to buoyancy forces the plume rises towards higher pressure which is reducing its velocity. The rising plume accelerates to some altitude and then starts to decelerate until its inertia and buoyancy potential (difference in density of plume and air) is lost by mixing with the ambient air. This final reduction of velocity is not shown in the pictures, since the height of the domain is not large enough. The temperature of the flue gases gradually decreases, however, as can be seen from Figure 5 (right), the core of warm flue gases is encircled by air stream until approximately 210 m, where the profiles of inner and outer part of the jet become smooth. Similar behavior can be observed for concentration of the flue gas constituents. It can be seen from Figure 5 (left) that under no wind conditions the concentration of SO_2 is zero near the tower shell at any altitude. At the tower outlet the distance of the closest non-zero SO_2 concentration from the shell is ca. 12 m.

4 CONCLUSIONS

The preliminary results of simulation of the natural draught wet-cooling tower with flue gas discharge showed that under no wind conditions the mixing of the flue gases inside the tower is low enough to be sure that the gases do not touch the tower shell. This infers that no corrosion due to acid condensation from the remaining sulphur oxides will occur. It should be however stressed, that this situation may be change under wind conditions which

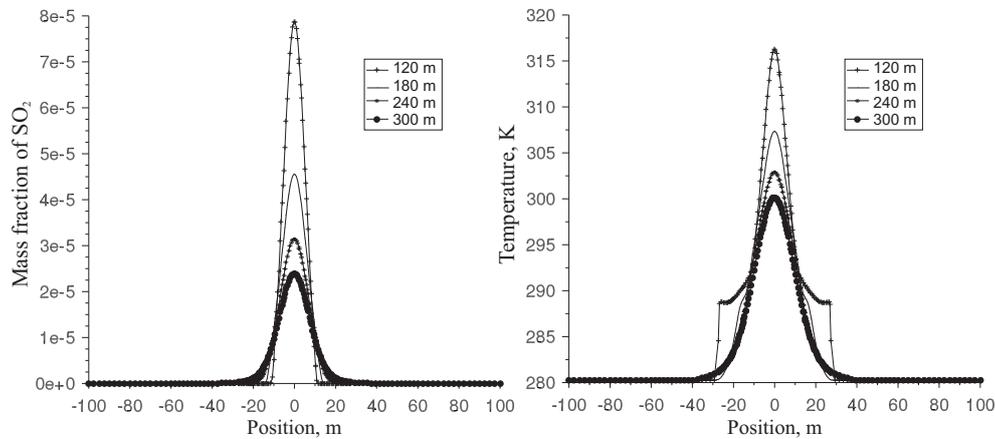


Figure 5: Profiles of SO_2 mass fraction (left) and temperature (right) at various elevations above the tower

will be the subject of our future research. The factors that may also have influence on the obtained results are the turbulence model used, and the assumptions made regarding the boundary conditions, specifically the constant temperature and pressure along the vertical boundaries.

ACKNOWLEDGEMENT

The research has been supported by a grant of the Polish Ministry of Research and Higher Education. This financial assistance is gratefully acknowledged herewith.

REFERENCES

- [1] Busch, D., Harte, R., Krätzig, W.B. and Montag, U. New natural draft cooling tower of 200 m of height. *Engineering Structures* (2002) **24**:1509-1521.
- [2] Roth, M., Gerber, R. and Niepel, A. Commissioning and operation of the world-wide largest Natural Draught Wet Cooling Tower of a lignite-fired power plant with flue gas discharge. *Natural draught cooling towers*. Mungan & Wittek (eds). Taylor & Francis Group, London, (2004).
- [3] Ernst, G., Leidinger, J.G., Natusch, K., Petzel, H.K. and Scholl G. *Kühlturm und Rauchgasentschwefelungsanlage des Modellkraftwerkes Völklingen*. Fortschritt-Berichte VDI, Reihe 15, No. 45, (1986)
- [4] Damjakob, H., Wöhler H.J. and Duddeck, G. Flue gas discharge via natural draught cooling towers - A survey. *Natural draught cooling towers*. Wittek and Krätzig (eds). Balkema, Rotterdam, (1994).
- [5] Klimanek, A., Bialecki, R.A. and Ostrowski, Z. CFD two scale model of a wet natural draft cooling tower. *Numerical Heat Transfer, Part A* (2010) **57**:119–137.

- [6] Klimanek, A. and Bialecki, R.A. A 3D CFD model of a natural draft wet-cooling tower. *Archives of Thermodynamics* (2009) **30**:119–132.
- [7] Klimanek A. *Numerical Modelling of Heat, Mass and Momentum Transfer in Natural Draft Wet-cooling Tower*. Faculty of Energy and Environmental Engineering, Silesian University of Technology, PhD Thesis, (2010).
- [8] Klimanek, A. and Bialecki, R.A. Solution of heat and mass transfer in counterflow wet-cooling tower fills. *Int. Commun. Heat Mass* (2009) **36**:547-553.
- [9] Klimanek, A., Ostrowski, Z. and Bialecki, R.A. Reduced Order Model of Heat and Mass Transfer in Wet-cooling Tower Fills, *Recent Developments in Artificial Intelligence Methods, AI-METH Series*, T. Burczyński, W. Cholewa, W. Moczulski (eds), Gliwice, (2009), 149–160.

PROCESS MODELLING OF LINEAR FRICTION WELDING (LFW) BETWEEN AA2124/SiC_p COMPOSITE AND UNREINFORCED ALLOY

X. SONG^{*}, N. BAIMPAS^{*}, S. HARDING[†] AND A. M. KORSUNSKY^{*}

^{*} Department of Engineering Science, Parks Road
University of Oxford,
Oxford, UK OX1 3PJ
e-mail: xu.song@eng.ox.ac.uk, alexander.korsunsky@eng.ox.ac.uk,
<http://www.eng.ox.ac.uk/solidmech/>

[†] Combustion Systems - Engineering
WH-45, Rolls Royce Plc
Bristol, UK BS34 7QE
email: stephen.harding@rolls-royce.com, <http://www.rolls-royce.com/>

Key words: Finite Elements (FE), Linear Friction Welding (LFW), Metal Matrix Composite (MMC), Fully coupled thermo-mechanical process modelling

Abstract. In the present study, the Linear Friction Welding (LFW) process between a bar of Metal Matrix Composite (MMC) AMC225xe (AA2124 with 25% SiC particulate reinforcement) and a bar of unreinforced base alloy was simulated using the commercial finite element package ABAQUSTM. Fully coupled implicit thermo-mechanical analysis procedure was employed, with semi-automatic re-meshing using Python scripting and output database scripting methods for extracting deformed configurations. Due to the large deformation near the weld region, multiple analyses were carried out between each re-meshing stage in order to limit the element distortion. Comparison of the simulation results with the experimental data collected during welding, and with post-weld optical section micrograph has shown satisfactory agreement.

1 INTRODUCTION

Metal Matrix Composites (MMCs) are a class of materials that typically contain reinforcement particles or fibres dispersed in a continuous metallic matrix. Over the past few decades, significant effort has been dedicated to developing MMCs with better physical and mechanical properties compared to monolithic metals. The synergy between the properties of the matrix (usually a light alloy based on Al, Ti or Mg) and the reinforcement (usually a hard ceramic oxide, carbide, or nitride) allows the MMCs to exceed the parent material's performance [1]. MMCs based on aluminium alloys form a class of attractive light-weight materials possessing a good combination of high stiffness and strength. The incorporation of stiff and hard reinforcement in the form of fibres or particles leads to a significant increase in the overall elastic modulus. Strength is also improved due to a variety of mechanisms, including grain refinement and the creation of additional obstacles to dislocation movement.

AA2124 is a wrought aluminium alloy widely used in the aerospace industry, e.g. fuselage, fuel tanks etc. It possesses a good combination of mechanical properties: relatively high stiffness and high strength, low density, and good resistance to fatigue crack growth. MMCs based on the AA2124 alloy usually use silicon carbide (SiC) particle reinforcement, due to the good interfacial bond that can be formed with the matrix.

Despite the several advantages over conventional alloys, a significant limitation to the industrial application of AMCs is posed by the problems that arise in conventional joining techniques, such as segregation and degradation (fracture) of the reinforcement phase. Recently, linear friction welding has been successfully applied to join aluminium alloy based MMCs. One of the crucial advantages of the LFW joining technique is that it avoids melting and solidification [2]: it is a solid state joining process in which the bonding of two parallel-edged components is completed by their relative reciprocating motion under the action of steady axial compressive force. During the process, significant heat is generated by the friction at the component interface, resulting in the continued displacement of plastically deformed material [3-7]. This reciprocal motion and the very large attendant strains, together with the complex interaction between the thermal and mechanical behaviour of components makes the task of modelling the process somewhat of a challenge. Previous attempts have been made to carry out Finite Element (FE) simulation of the LFW process between two components made from the same Ti alloy, leading to limited success in capturing the temperature field evolution [7, 8]. To the best of the authors' knowledge, no work has been reported to date on simulating LFW of aluminium alloys or aluminium-based composites, particularly for the case of joining two blocks with different material properties.

Below we describe the simulation setup and report the comparison of the optimised model output with observations.

2 THE FINITE ELEMENT SIMULATION

2.1 The LFW process description

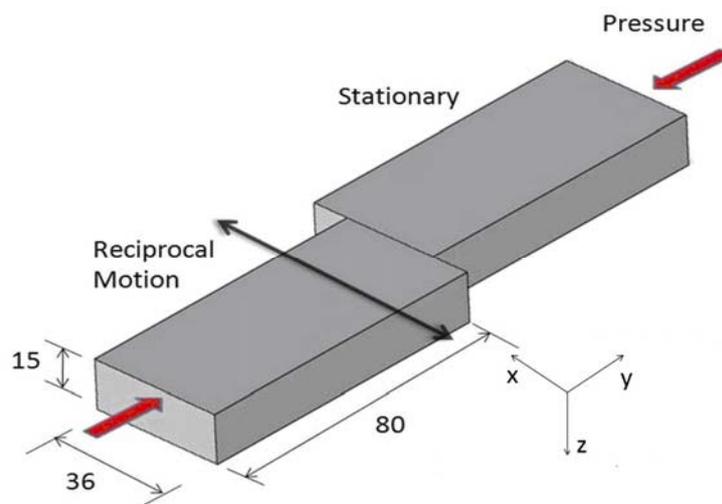


Figure 1: Illustration of the AA2124/MMC linear friction welding arrangement, and the coordinate system used.

In order to simulate the entire linear friction welding process, the detailed parameters of the welding process are required. In the present study we used the data collected during Linear Friction Welding of AA2124/AMC225xe performed at TWI (The Welding Institute, Cambridge, UK). The schematic illustration of the welding setup can be seen in Figure 1, and the details of the welding process parameters are given in Table 1 and Figure 2.

Table 1: Linear friction welding process parameters

Joint	Force (kN)	Pressure (MPa)	Frequency (Hz)	Amplitude (mm)	Burn-off (mm)	Piece Initial Length (mm)	Weldment Length (mm)	Total Upset (mm)
AA2124/AMC225xe	85	157	50	2	2	80	150.94	9.06

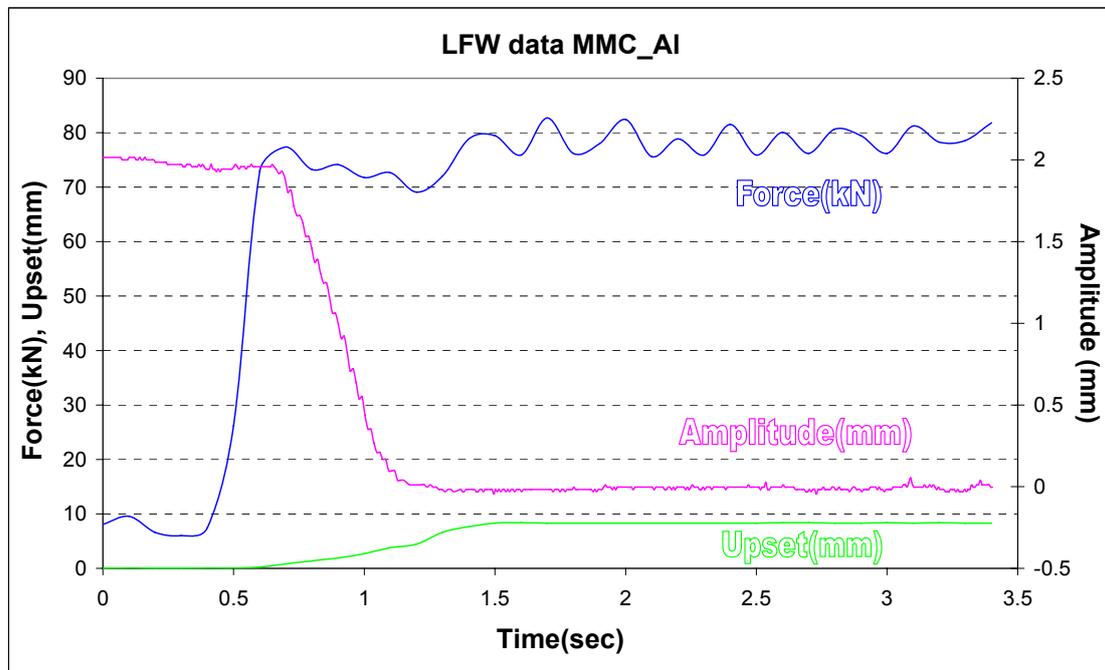


Figure 2: Plots of time history of the applied force, amplitude and upset during LFW process

In Table 1, total upset means the total axial shortening of the two components after welding, while burn-off indicates the critical initial shortening at which the LFW control system begins to reduce the oscillation amplitude. In the simulation, one bar of the assembly was maintained stationary, while the other was subjected to oscillatory movement along the y -direction whilst experiencing a compressive force in the x -direction applied at the top end of the bar. When the axial shortening reached 2mm, when the corresponding time was ~ 0.65 s in Figure 2, the oscillation amplitude was triggered to reduce in a linear ramp that reached zero in 0.5s, which also corresponded to the time interval for observation made in the course of the experiment (Figure 2).

2.2 Setup of 2D LFW simulation

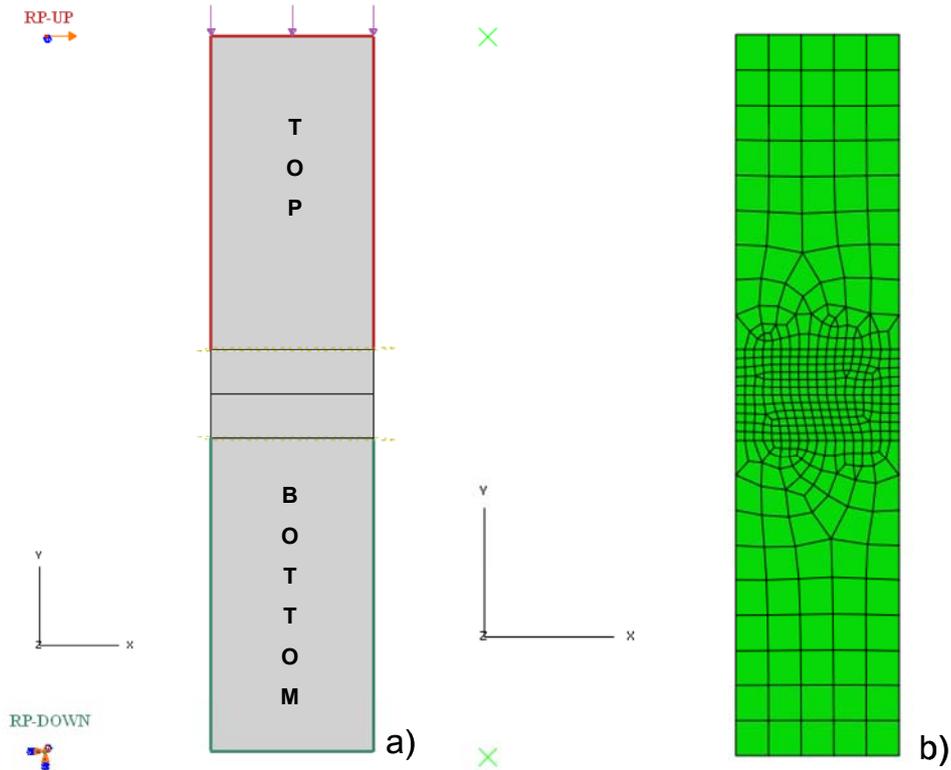


Figure 3: a) 2D model setup with boundary conditions and loads; b) 2D mesh with fine and coarse mesh regions

For the purposes of the current study, a 2D model was created with the exact in-plane dimensions of the specimens, i.e. the width of 36mm and length of 80mm for each bar (Figure 3). As illustrated in the Figure, the bottom bar was encastre, whilst the top bar had time-varying displacement boundary conditions applied to represent the oscillatory movement during the welding process. Two Reference Points (RP) were created in the model, for the following purposes: 1) to apply fixed or moving displacement boundary conditions, by linking the RPs through input command “*Equations” with the specific edges of the model, highlighted in colour in Figure 3a; 2) to act as a sensor through a user subroutine (UEL) for measuring the current weld upset. When a critical user-defined upset distance is reached or exceeded, the UEL calls the utility routine XIT to trigger re-meshing and to ensure that the elements are not excessively distorted; 3) to play the role of information channel between the input file and various subroutines at the beginning of each re-mesh analysis step, so that the total run time information can be made available for the user-defined subroutine UAMP that defines the amplitudes of oscillation and pressure. In this way, the kinematic aspects of the bars being joined can be fully monitored and controlled via the respective RPs.

The LFW process model developed in the present study follows a nonlinear, quasi-static, thermo-mechanically coupled analytical framework. Each analysis step in the simulation sequence represented a single fully coupled temperature-displacement calculation. The exact duration of the step was not known a priori, but was in fact controlled by the user element subroutine UEL.

The key aspect of the implementation that was crucial for successful completion of the simulation was the re-meshing capability. During LFW, large deformation occurs in the near the weld region, namely, local shearing, forging and flash formation. If a single mesh were used, no matter how fine, element distortion would accumulate and soon render the calculation impossible. To limit element distortion, re-meshing had to be triggered when certain criteria were fulfilled. The procedure was automated through the use of Python scripts. To capture the significant changes in the component shape, the capabilities of ABAQUS/CAE were used to extract the outer contour of the bars, re-seed the surface, and create the new mesh in the automatic mode. In each bar, the mesh was divided into two regions (see Figure 3b). Smaller elements were used in the region near the bond line, and also for the materials forming the flash.

To describe the contact conditions, two types of contact interactions were defined: weld contact and self-contact. The weld contact was described by a pair of interactions that was symmetrical, in the following sense. The first interaction defined the bottom surface of the top bar as the *master* surface and the top surface of the bottom bar as the *slave* surface. In the second interaction definition this master-slave relationship was reversed. This “balanced master-slave” arrangement ensures more accurate and stable description of the contact pressure at the weld interface and avoids “hourglass” effects. Furthermore, it was combined with a softened contact interaction description in order to promote the re-distribution of the contact pressure between nodes lying along and to both sides of the interface between the two bars being joined.

The other type of contact interaction was introduced to address the possibility of self-contact that may cause problems during re-meshing. The *Part2DGeomFrom2DMesh* command was used to generate the new, current configuration geometry. This is achieved by performing curve-fit operations, and these in turn may lead to self-intersections of the boundary, with consequent invalid part topology and meshing failure. To overcome this problem, a softened contact model was used that introduced a normal pressure even for a small separation distance (0.01 mm). The separation distance was kept as small as practical to avoid introducing non-physical assumptions into the contact behaviour.

2.3 Material properties

The material of the top and bottom bars was AMC225xe and AA2124, respectively. The inelastic deformation response of both materials was described by constitutive laws that incorporated temperature and strain rate dependence of yield stress. The strain rate dependence was defined by the Johnson-Cook law where the prevailing exponential coefficient used was $C = 0.0083$ [9]. The temperature dependence of the other relevant physical and mechanical properties was found from the ASM Metals Handbook [10]. The values used in the simulation are given in Table 2 and 3. It is worth noting that the temperature dependence of the yield stress exerts crucial influence on the LFW process, while the strain rate-dependence of the yield stress greatly influences the model convergence. For a simulation that is stable and realistic, correct definitions of the temperature and strain rate dependence of the properties of both AA2124 and AMC225xe materials are of crucial importance.

Table 2: Material properties for AA2124

Temperature (°C)	Yield Stress (MPa)	Elongation	UTS (MPa)
24	450	8%	485
150	395	10%	415
205	340	13%	365
260	240	17%	270
315	145	23%	160
370	69	35%	76
Density (kg/m)	Thermal Conductivity (W/m·K)	Specific Heat (J/kg·K)	Young's Modulus (GPa)
2770	191	882	71

Table 3: Material proerties for AMC225xe

Temperature (°C)	Yield Stress (MPa)	Elongation	UTS (MPa)
20	480	5%	650
25	480	5%	650
150	321	9%	428
200	276	18%	358
260	102	34%	200
350	48	45%	65
Density (kg/m)	Thermal Conductivity (W/m·K)	Specific Heat (J/kg·K)	Young's Modulus (GPa)
2880	150	836	115

3 RESULTS AND DISSCUSSION

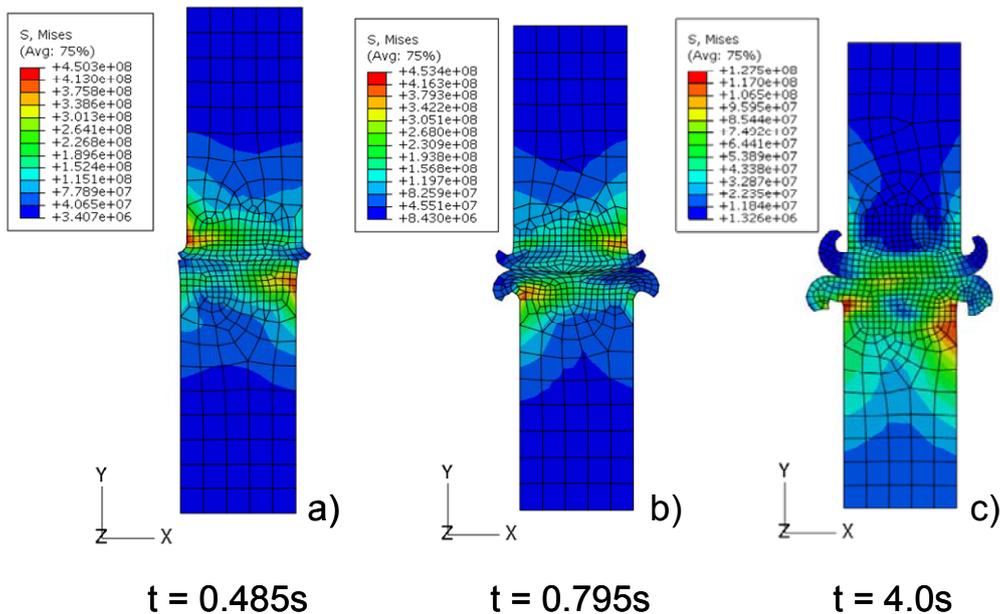


Figure 4: Von Mises stress contours in the specimen at different frame time: a) 0.485s, b) 0.795s and c) 4.0s

The LFW simulation presented here had the total duration of 4s, and burn-off time 0.85s. The oscillation amplitude was ramped down to zero at a time close to 1.15s. At 3.5s the applied load began to decrease, and at the target time of 4s the compressive load was reduced to zero.

The von Mises stress contours at different stages of the simulation are shown in Figure 4. The sequence provides an illustration of how the process evolves and what significant shape changes occur. They involve the “squeezing out” of the material from the weld zone and into the flash. It is seen in Figure 4a that at ~ 0.5 s significant flash generation and axial shortening (upset) begin. At the time of ~ 0.8 s, greater amount of flash was generated in both bars, with some noticeable asymmetry (right to left), and with significantly more flash produced on the AA2124 side. This remains true at 4.0s (end of the process, Figure 4c): most of the flash is seen in the unreinforced alloy side of the weldment, while only a relatively thin layer of highly curved AMC flash can also be found.

Figure 4c represents the distribution of von Mises stress within the assembly at the end of the process. At this stage the externally applied compressive force is already removed, so the residual stress state is being considered. As expected, a higher level of residual stress is found in the softer of the two bars being joined.

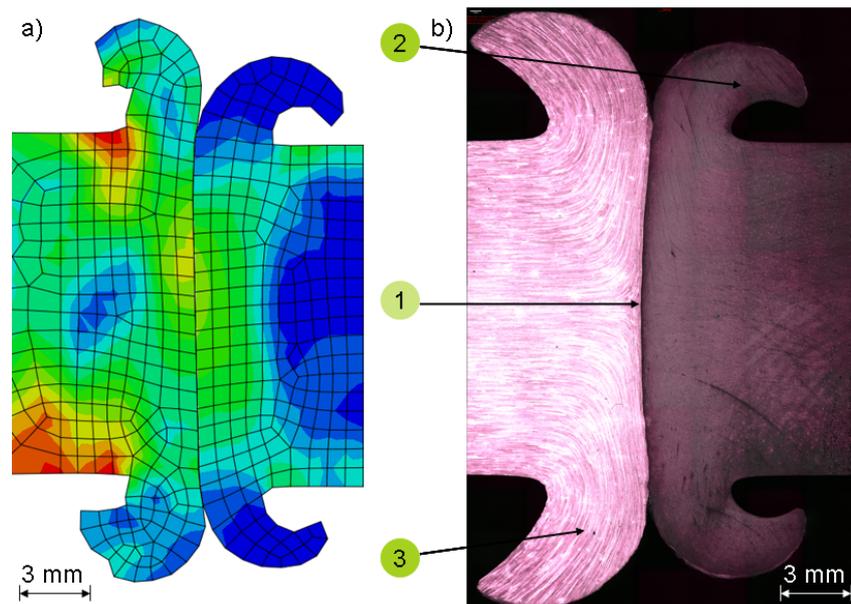


Figure 5: Flash geometry a) in the model and b) in the experiment in LFW of AA2124/AMC225xe

One possibility of validating the process model is to compare the predicted post-weld flash shape with the optical section micrograph. Figure 5 shows, on the same scale, the comparison of the flash geometry between the FE model and experiment (LFW bond region observed under an optical microscope). The shapes of the flash for the AMC225xe material obtained

from the model and experiment appear almost identical, with both the flash thickness and curvature (Region 2 in Figure 5b) captured well. The flash shape for the unreinforced alloy AA2124 obtained in the model is somewhat similar from the observation, although the flash thickness once again appears to be captured correctly (see Region 3). The model also captures correctly the gentle curvature of the bond line in Region 1, with the region occupied by the relatively hard MMC appears convex, and concave for the softer AA2124 alloy.

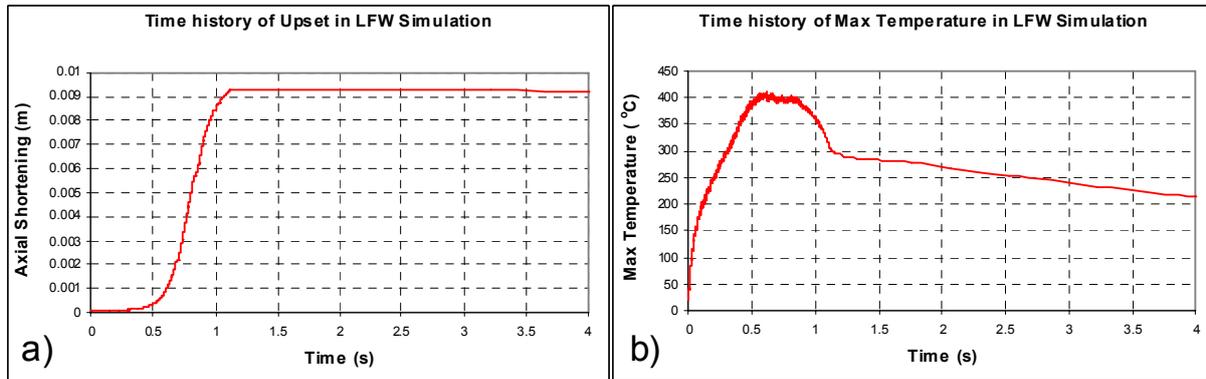


Figure 6: a) Time history of the axial shortening (upset) in the LFW simulation; b) Time history of the maximum temperature in the LFW simulation

Model post-processing was also developed in order to extract the upset and maximum temperature across the entire sequence of simulation steps. Figure 6 illustrates the time history of the model upset and maximum temperature during the LFW simulation. It is worth noting that the final upset of the model (9.36mm) is quite close to the experimental value (9.06mm). Furthermore, at the time of 0.65s, the upset predicted by the model is ~2mm, i.e. in precise agreement with the burn-off value when the oscillation amplitude started to decline. It is worth pointing out for clarity that the ramping down of oscillation amplitude in the simulation is time-triggered at $t=0.65$ s. In contrast, in the experiment it is the burn-off reaching $u=2$ mm that acts as a trigger for reducing oscillatory the amplitude.

It is also worth noting that the maximum process temperature within the model fell into the logical range, not exceeding the solidus temperature (502°C) yet close to that of the forging (400°C).

4 CONCLUSIONS

In the present study, a successful simulation is reported of the LFW process between unreinforced aluminum alloy AA2124 and Metal Matrix Composite AMC225xe that is based on the same alloy as the matrix and is reinforced with particulate SiC. The fully coupled implicit thermo-mechanical analysis procedure is described, with semi-automatic re-meshing to control element distortion. The simulation provided satisfactory agreement with the records of process parameters taken during welding, i.e. compressive force required for forging, burn-off and total upset. Further validation of predictive capability in terms of the mechanical behaviour of joints requires comparing the residual stresses from the simulation with experimental measurements carried out on the weldments e.g. by diffraction. This work is currently under way.

REFERENCES

- [1] Clyne, T.W. and Withers, P.J., *Introduction to metal matrix composites*, Cambridge University Press, (1993).
- [2] Jun, T.S., Rotundo, F., Ceschini, L., and Korsunsky, A.M. A study of residual stresses in Al/SiCp linear friction weldment by energy-dispersive neutron diffraction. *Key. Eng. Mater.* (2008) **385-387**: 517-520.
- [3] Vairis, A., and Frost, M., High frequency linear friction welding of a titanium alloy. *Wear* (1998) **217**:117-131.
- [4] Li, W.Y., Ma T.J. and Li, J.L., Numerical simulation of Linear Friction Welding of titanium alloy: Effects of processing parameters. *Mater. Design*, (2010) **31**: 497-1507.
- [5] Wanjara, P., Jahazi, M., Thermal-Phase Transformation Modelling and Neural Network Analysis of Friction Welding of Non-Circular Eutectoid Steel Components. *Metall. Mater. Trans. A*, (2005) **36**: 2149-2164.
- [6] Daymond, M.R. and Bonner, N.W., Measurement of strain in a titanium linear friction weld by neutron diffraction, *Physica B*, (2003) **325**: 130-137.
- [7] Vairis, A., and Frost, M., Modelling the linear friction welding of titanium blocks, *Mater. Sci. Engng. A*, (2000) **292**: 8-17.
- [8] Müller, S., Rettenmayr, M., Schneefeld, D., Roder O. and Fried W., FEM simulation of the linear friction welding of titanium alloys, *Comp. Mat. Sci.*, (2010) **48**: 749–758.
- [9] Schmidt, H. and Hattel, J., A local model for the thermomechanical conditions in friction stir welding, *Modelling Simul. Mater. Sci. Eng.*, (2005) **13**: 77–93
- [10] Cubberly, W.H. et al., *ASM Metals Handbook* (Ninth Edition), American Society for Metals, Vol. II., (1979).

THREE DOMAIN THERMAL AND MECHANICAL FLUID-STRUCTURE INTERACTION ANALYSIS APPLIED TO COOLED ROCKET THRUST CHAMBERS

DANIEL S. C. KOWOLLIK, MATTHIAS C. HAUPT AND PETER HORST

Institute of Aircraft Design and Lightweight Structures
Technical University Braunschweig
Herrmann-Blenk-Str. 35, 38106 Braunschweig, Germany
e-mail: d.kowollik@tu-bs.de, web page: <http://www.ifl.tu-bs.de>

Key words: FSI, CFD, TBC, Coupling, Thrust Chamber, Liquid Rocket Engine

Abstract. Regeneratively cooled combustion chamber and nozzle structures are exposed to extreme temperature gradients in space and time. One sided wall heating during the hot run generates thermomechanical loads that induce high heat fluxes on the surface and consequently high stresses inside the thin cooling channel structures. In order to address the strong interaction between the structure and the different flow fields a coupled simulation considering the thermal and mechanical interactions is desirable. The present study covers both physical couplings in a partitioned approach applied to the steady state case of a subscale thrust chamber.

Furthermore, this study will present a novel full parametric 3D modelling approach for cooled rocket thrust chambers, which is specifically designed to reduce computational expense in FSI analyses by applying non conforming symmetry conditions across coupling interfaces. The parametrization ranges from thrust chamber contour design through CAD modelling up to grid generation of the individual domains. Further extension of the parametric approach allows the analysis of thermal protection systems applied inside the combustion chamber.

1 INTRODUCTION

The regeneratively cooled rocket thrust chamber is one of the most important components determining the performance of today's launch vehicles. Over the years research has been performed with the goal to optimize the overall performance of rocket engines. Complex thermal and mechanical loading conditions cause a strong interaction between the hot combustion gas, the thin wall structure of the combustion chamber and the cryogenic liquid fuel which is pumped through the high aspect ratio cooling channels. Combustion gases reach a temperature level of about 3,600 K with a chamber pressure up to 20 MPa [1]. The highest heat flux is reached with 131–147 MW/m² at the throat region [2].

The response of the cooling channel structure can only be addressed through an adequate prediction of the in-service conditions. Chemical reactions during the combustion process, fluid flow in the thrust chamber and cooling channels, heat transfer between the involved domains and the thermomechanical coupling in the structure make up a complex system. The thermal and pressure loads generated by the hot gas and cooling channel flow are essential for the prediction of the cooling channel response.

Computer resources in the past were limited to the application of simplified models for coupled simulations. In [3] the thermal interaction between a 2D structural model and a straight 3D cooling channel segment is analysed. The interaction with the hot gas side is neglected, whereas experimental data of [4] serve as boundary condition. This simplified model is extended by a thermomechanical coupling strategy in [5]. Riccius argues in [6] about the relevance of a correct transient load cycle, whereas a 1D fluid model is deployed for quasistatic and transient thermal analyses. A 3D coupled heat transfer analysis is presented by Liu et al. in [7], where the hot gas and cooling channel structure are modelled by a finite volume scheme. The cooling channel flow is included by a 1D approach. Current industrial simulation strategies were presented by Knab et al. in [8], where a coupled heat transfer simulation is performed between the hot gas side and the coolant flow. The hot gas side is modelled by an axisymmetric multi-phase Navier-Stokes solver, whereas the cooling channel flow and the structure are analysed by a 3D conjugate heat transfer model.

The increasing computing capacity now allows us to perform a coupled 3D fluid-structure interaction analysis of the whole thrust chamber structure. A coupled steady state heat transfer analysis with a subsequent static mechanical analysis of a subscale LOX/H₂ thrust chamber is investigated in this work.

Assuming finalized combustion at the inlet, the hot gas flow is determined by means of the 3D steady RANS solver of the CFD code DLR-TAU [9]. ABAQUS FE Software is used to analyse the cooling channel response in a hot gas run at steady state.

The FSI computation environment *ifls* developed at the Institute of Aircraft Design and Lightweight Structure at the Technical University of Braunschweig will be shown to be able to perform the fully coupled 3D FSI Analysis of the thrust chamber structure.

Building parametric models for FSI Analysis is a laborious work, because each physical domain has to be carefully designed starting from CAD modelling, going through specific grid generation techniques according to the physics involved and finally defining the appropriate boundary conditions, where e.g. discretised coupling surfaces have to be extracted for interpolation purposes of the involved field and flux variables considering non-conforming grids. In this study we will present a novel object oriented full parametric 3D modelling approach for cooled rocket thrust chambers. The approach covers all above mentioned modelling steps in one software concept.

In the following sections we will present the partitioned numerical coupling concept implemented in *ifls*, outline the parametric modelling approach and demonstrate the complete coupling approach of this study on a steady state subscale thrust chamber

computation, which will be discussed.

2 NUMERICAL FSI COUPLING CONCEPT

The developed simulation environment *ifls* provides a framework to simulate coupled problems in a partitioned approach, where the coupling domains are analysed with individual codes for the aerodynamic and structural phenomena. Detailed information of the software concept implemented in *ifls* and the available state-of-the-art techniques for numerical coupling can be taken from [10]. The software environment has been tested and validated for different individual code combinations. For details, refer to the previous studies [10, 11, 12, 13].

Basic requirement for such a coupled concept is the equilibrium of energy transferred across the coupling surfaces during a single coupling time step. In the following analysis a steady state problem is solved. To ensure conservation in space, non-conforming grids need to be handled such that the transfer of the state variables is achieved in a conservative manner. Several techniques to exchange the boundary conditions on the coupling surface were proposed. A general approach to construct conservative transfer schemes is provided in [14, 15, 16], which fulfill the interface conditions in a weak formulation based on the Lagrangian multiplier technique. Based on this approach *ifls* provides several popular transfer schemes presented in [10, 17]. The transfer scheme applied in the present study is referred to as conservative interpolation or node projection scheme. In this case the nodal loads of the fluid interface are mapped to the nodes of the closest structural interface element, where the structural shape functions are evaluated at the fluid interface nodes. A schematic illustration of *ifls* iteration procedure and framework is shown in Fig. 1.

The fluid-structure interaction considered in this work is assumed to be steady state for all involved domains. A two way coupled formulation accounts for the heat transfer and stress/displacement problem between the hot gas and the cooling channel structure. In each equilibrium iteration step the structural domain is solved in a sequentially coupled thermomechanical analysis, in which first the heat flux computed on the hot gas side serves as input for the structural thermal analysis and second the structurally computed temperature distribution together with the hot gas pressure loads is applied in a subsequent static stress/displacement analysis. The deformation results of the structural analysis are transferred to the fluid domain in order to perform a grid deformation in each iteration step. Furthermore, the surface temperatures computed by the structural solver are interpolated to the fluid side serving as new input in the equilibrium algorithm.

The solution of the coupled problem is obtained by the Dirichlet-Neumann iteration. Formulating the Dirichlet problem in terms of a Schur complement one defines symbolically the fluid operator \mathcal{F} as follows:

$$\mathcal{F} \phi^{(f)}|_{\Gamma} = \phi_{,\mathbf{n}}^{(f)}|_{\Gamma}, \quad (1)$$

where ϕ represents the state variables (temperature, displacements) and $\phi_{,\mathbf{n}}$ the flux (heat flux, surface tractions) at the coupling interface Γ . The indices (s) and (f) at the state

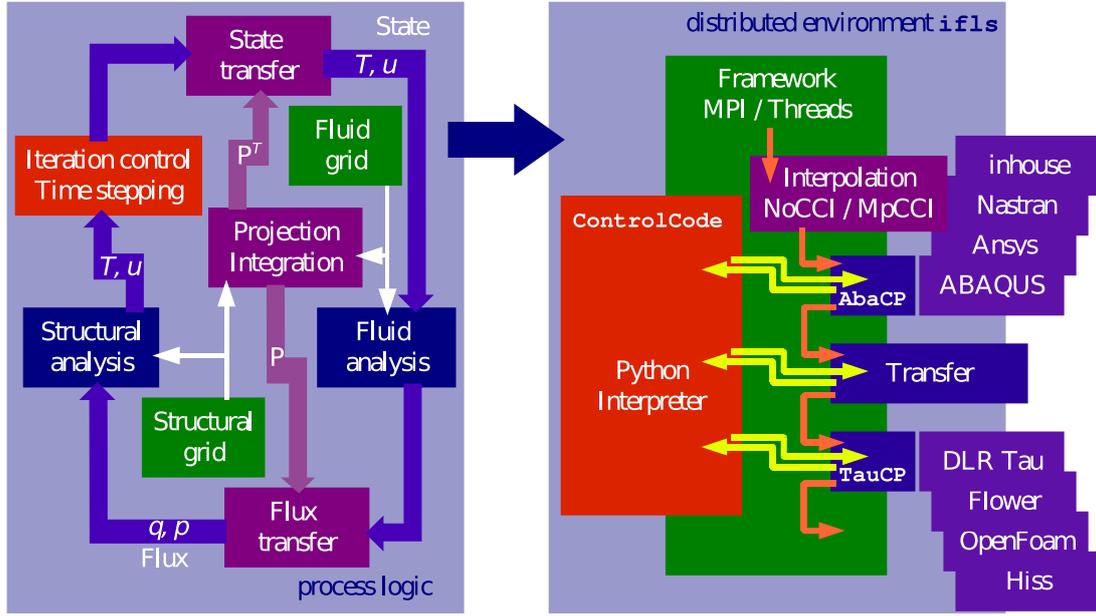


Figure 1: Process logic of the coupling algorithms and the corresponding software architecture

variables ϕ and the flux $\phi_{,n}$ represent the fluid and the structural domains, respectively. The inverse Schur complement is formulated for the structural problem, consequently the Neumann problem is solved:

$$\phi_{,n}^{(s)}|_{\Gamma} = \mathcal{S}^{-1} \phi^{(s)}|_{\Gamma} \longleftrightarrow \mathcal{S} \phi^{(s)}|_{\Gamma} = \phi_{,n}^{(s)}|_{\Gamma}. \quad (2)$$

For the classical Dirichlet-Neumann iteration the state variables are relaxed in each steady state iteration step as follows:

$$\phi_{k+1} = (1 - \omega) \phi_k + \omega \mathcal{S}^{-1} \mathcal{F} \phi_k, \quad (3)$$

with the relaxation parameter ω .

The application of a fixed ω between 0.7 and 1.0 can be assumed as standard in coupled heat transfer problems. However, an optimal computed $\omega = \omega_{opt}$ is desirable to accelerate the equilibrium iteration. Several acceleration methods, e.g. Aitken method and gradient method have been analysed in the past [18, 19].

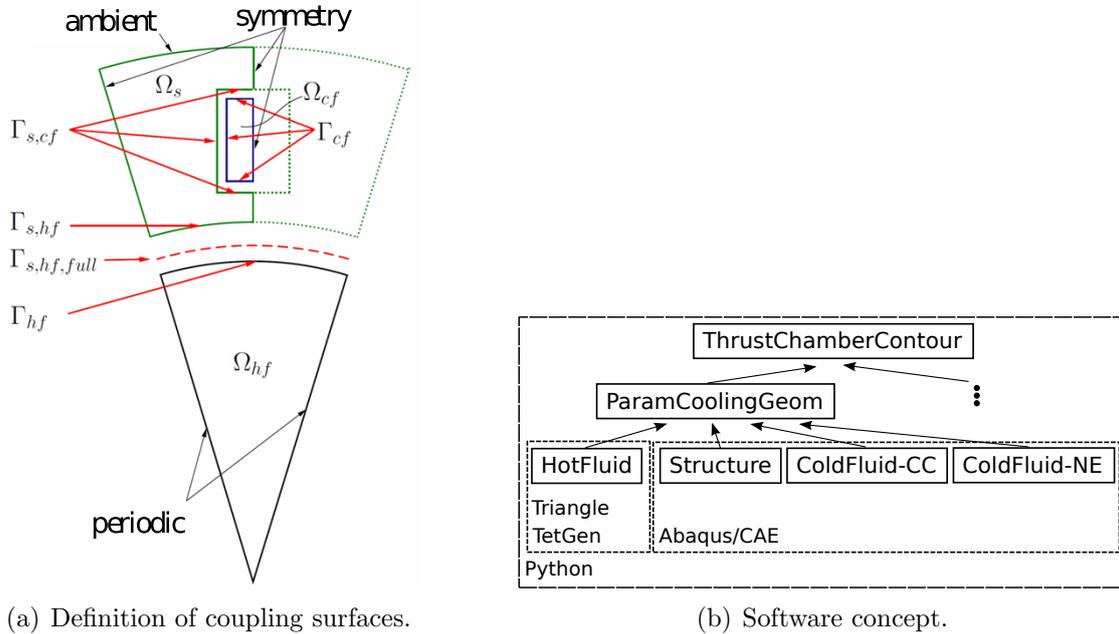
For the fully coupled problem in the present work the Aitken method is used to accelerate the fixed point iteration. The Aitken method is defined as follows:

$$\omega_{opt,k} = 1 - \mu_k, \quad (4)$$

where the Aitken coefficient μ_k is defined as

$$\mu_k = \mu_{k-1} + (\mu_{k-1} - 1) \frac{(\Delta\phi_k - \Delta\phi_{k+1})^T \Delta\phi_{k+1}}{(\Delta\phi_k - \Delta\phi_{k+1})^2} \quad \text{for } \mu_k \geq 1, \quad (5)$$

with $\Delta\phi_{k+1} = \phi_k - \phi_{k+1}$. For the first iteration loop $\mu_k = 0$.



(a) Definition of coupling surfaces.

(b) Software concept.

Figure 2: Parametrized modelling aspects of cooled rocket engines.

3 PARAMETRIC 3D MODELLING APPROACH

3D FSI analyses of cooled rocket thrust chambers are computationally expensive. Computational reduction can be achieved if symmetry conditions are used, consequently the assumption of the present parametric approach is to model half of a cooling channel segment for the combustion chamber cooling circuit. The hot gas is modelled with periodic boundary conditions, which is shown in Fig. 2(a), where the coupling surfaces are defined. The data transfer between the hot gas and structural domain is achieved through an additional coupling surface $\Gamma_{s,hf,full}$, where the state and flux variables are mirrored to satisfy integrity of the applied boundary conditions in each iteration step. Accounting for all symmetry conditions in the different computational domains allows the simulation of the 3D state by the assumption of periodic repetition.

In numerical coupling it is very time consuming to realize CAD modelling and grid generation for all involved computational domains. Validation of numerical methods is always a topic, therefore a software concept for parametrized modelling of thrust chamber designs and cooling channel setups was developed. The parametrization reaches from the thrust chamber contour and cooling channel design to the CAD modelling and finally to the grid generation of the hot gas, structure and cooling channel flow field.

Fig. 2(b) shows schematically the developed software architecture, which uses the python interface to the preprocessor Abaqus/CAE. Python scripting allows an object oriented modular and reusable framework for the different components of the parametrized models. Additional software packages like Triangle [20] and TetGen [21] enhance

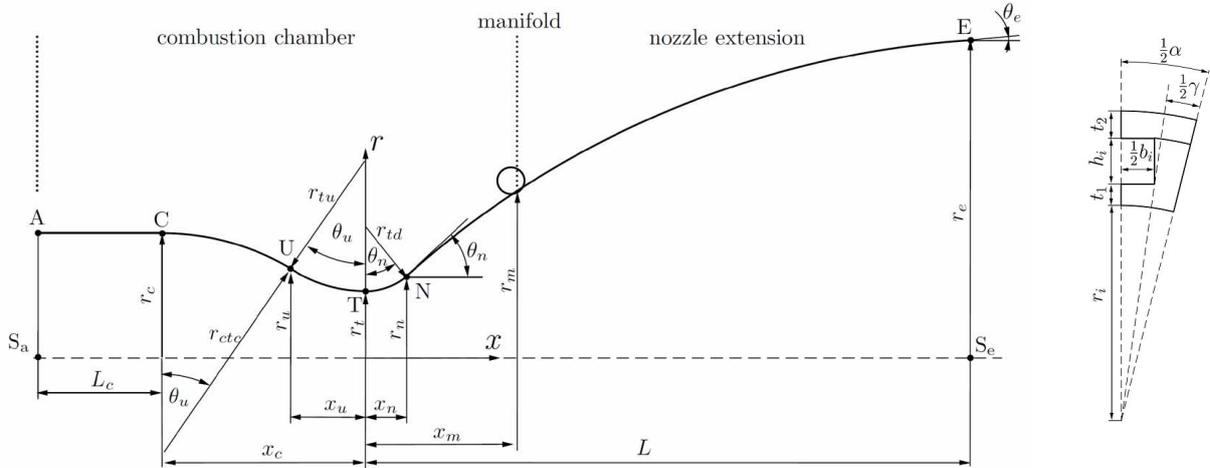


Figure 3: Parametrized model: thrust chamber contour (left); cooling channel cross section (right).

the possibilities in grid generation.

The class definition of *ThrustChamberContour* describes the complete rocket engine contour of Fig. 3 (left), where the nozzle is designed with the thrust optimized parabola (TOP) from Rao [22]:

$$\left(\frac{bx}{r_t} + \frac{r}{r_t}\right)^2 + \frac{cx}{r_t} + \frac{dr}{r_t} + e = 0. \quad (6)$$

Even modern nozzles like the Vulcain and SSME nozzle can be studied by means of a parabolic contour [23]. The nozzle contour can be described by five independent variables r_{td} , θ_n , L , r_e and θ_e . The angles θ_n and θ_e are evaluated through a bivariate spline whose input data was taken from Rao [22].

The derived class *ParamCoolingGeom* accounts for the chosen cooling circuit cross section definition shown on the right of Fig. 3. Different cooling channel setups are possible and can be derived likewise. For each cooling channel setup, one can derive classes for the different computational domains. These classes use the object oriented Abaqus/CAE interface to implement the parametrized CAD and grid models. In this study the analysed subscale thrust chamber consists of three domains structure, hot gas and the two cooling circuits, one in the combustion chamber region and one in the nozzle extension. For both cooling circuits a continuously varying cooling channel geometry can be defined for the cross section sketched in Fig. 3 (right).

4 NUMERICAL RESULTS

4.1 Cooled subscale thrust chamber

In this study a coupled FSI analysis of a LOX/H₂ subscale rocket thrust chamber is simulated. The subscale thrust chamber is defined by Astrium Space Transportation GmbH, Propulsion & Equipment. Extracted general parameters of the detailed test case

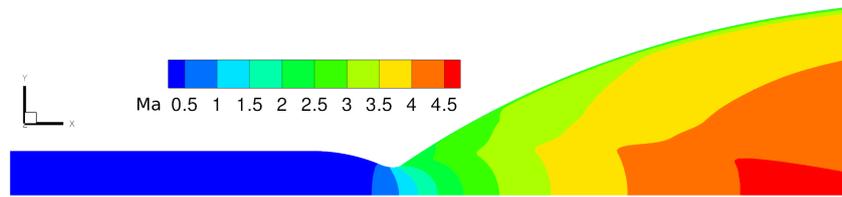


Figure 4: Mach number distribution at steady state of the coupled FSI analysis.

definition serve as input for the parametrized thrust chamber contour of this work. The test case consists of 80 cooling channels in the combustion chamber and 160 cooling channels in the nickel based nozzle extension. The combustion chamber material setup is composed of a NARloy-Z [24] liner and an INCONEL alloy 600 jacket.

On the hot gas side a 3D steady RANS analysis using ideal gas law assuming finalized combustion is performed in each Dirichlet step of the static FSI analysis. For the fluid simulation the DLR TAU-Code is used, which is an unstructured RANS solver based on the finite volume method. Reservoir pressure inflow conditions were computed with the preliminary design tool RPA (Rocket Propulsion Analysis) [25] and served as inlet conditions for the DLR TAU-Code. The temperature niveau computed with RPA reaches 3502 K at a pressure level of 9.35 MPa. At the outlet supersonic outlet conditions are applied, while an isothermal wall is defined at the coupling surface. The turbulent effects are modelled by the original version of the Spalart-Allmaras model implemented in the TAU-Code. The hybrid grid consists of 590833 grid points, 686535 tetrahedra and 785686 prisms. The dimensionless y^+ is adapted to a maximum of 0.57 at the thrust chamber throat. Fig. 4 shows the typical mach number profile of a TOP nozzle, where compression waves are generated at the intersection of the downstream arc and the parabola contour. These compression waves coalesce in an internal shock. Comparisons to Östlund approve these flow phenomena [23].

The steady state heat transfer problem of the cooling channel structure is analysed with the ABAQUS FE software. The resultant heatflux of the hot gas simulation is transferred as boundary condition on the coupling surface. In this study the cooling circuits introduced in section 3 are accounted for by the definition of constant film coefficients at the structural boundary. The film coefficient of $h_{f,CC} = 150 \text{ kW}/(\text{m}^2 \text{ K})$ is applied for the combustion chamber circuit and $h_{f,NE} = 30 \text{ kW}/(\text{m}^2 \text{ K})$ is applied for the nozzle extension circuit. For both circuits the sink temperature is defined as $T = 40 \text{ K}$. Radiation effects in the combustion chamber are not considered. Radiation to ambient is computed for the outside facing surface of the cooling channel structure.

The resultant temperature distribution of the cooling channel structure is used as input condition for the static stress/displacement analysis. The fluid solver provides the mechanical surface loads for the hot gas side. At the inlet deflections in the axial direction are suppressed. Symmetric deformation is guaranteed by applying zero deflection normal to the symmetry planes sketched in Fig. 2(a). For the static analysis a linear

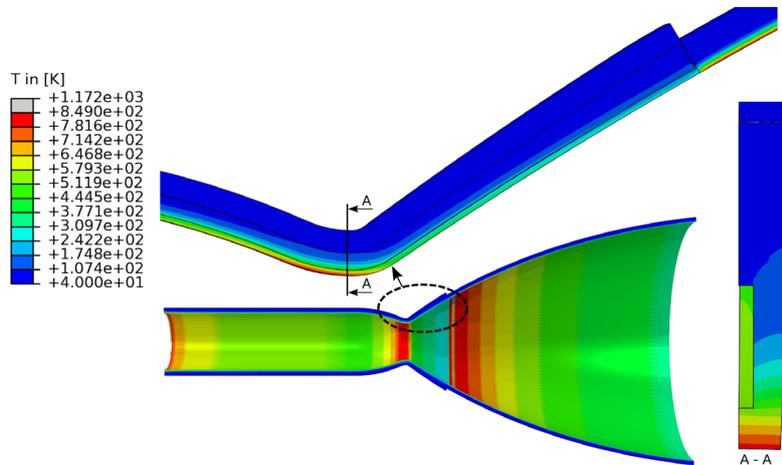


Figure 5: Structural temperature distribution of the cooled rocket engine at steady state.

elastic material model with temperature dependant material parameters is applied for the above mentioned materials. Both finite element models are modelled by 3D 8–node linear hexaeder elements.

Fig. 5 shows the converged temperature distribution of the structural domain. The high temperatures at the combustion chamber inlet can be explained by the assumption of finalized combustion at the inlet, where a homogenous inflow condition neglecting the typical injector head is applied. Realistic physical conditions are met at the throat region, where the temperature peak reaches about 849 K. The cut view at the throat region shown in Fig. 5 gives insights about the temperature distribution of the cooling channel cross section.

Another temperature peak can be identified in Fig. 6(a) at the nozzle extension just downstream of the manifold position. The nozzle extension material is completely out of INCONEL alloy 600, which has a heat conductivity one order lower than the upstream NARloy-Z. Furthermore, the assumed constant film coefficient is five times higher in the upstream cooling circuit.

The structural response of the static stress/displacement analysis is shown for the thrust chamber counter in Fig.6(b). The effect of the global deformed state on the aerothermomechanical analysis is small compared to the strong influence of the thermal interactions.

4.2 FSI analysis including thermal protection systems

The transfer of thermal barrier coatings (TBCs) to the high heat flux environment of rocket thrust chambers seems straightforward because TBCs are effectively used in power generation and other aerospace applications. Referring to thrust chambers, TBCs offer great potential to reduce heat flux, coolant temperature and pressure loss; consequently an increased chamber life is expected.

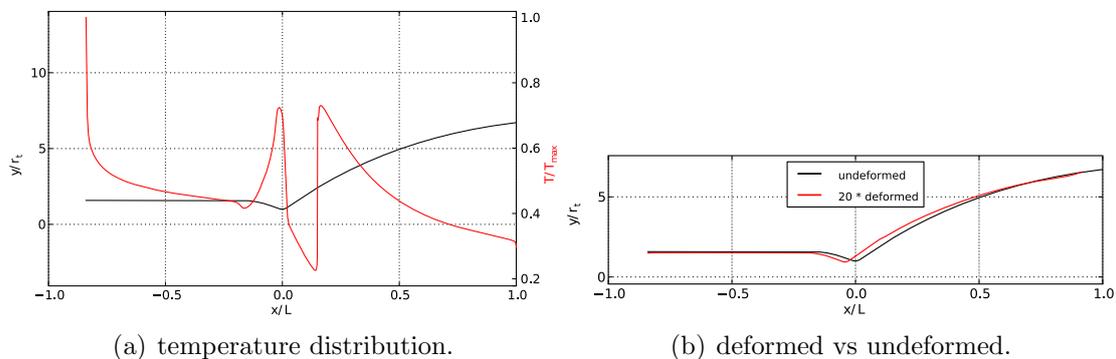


Figure 6: Results along the thrust chamber contour of a static aerothermomechanical analysis.

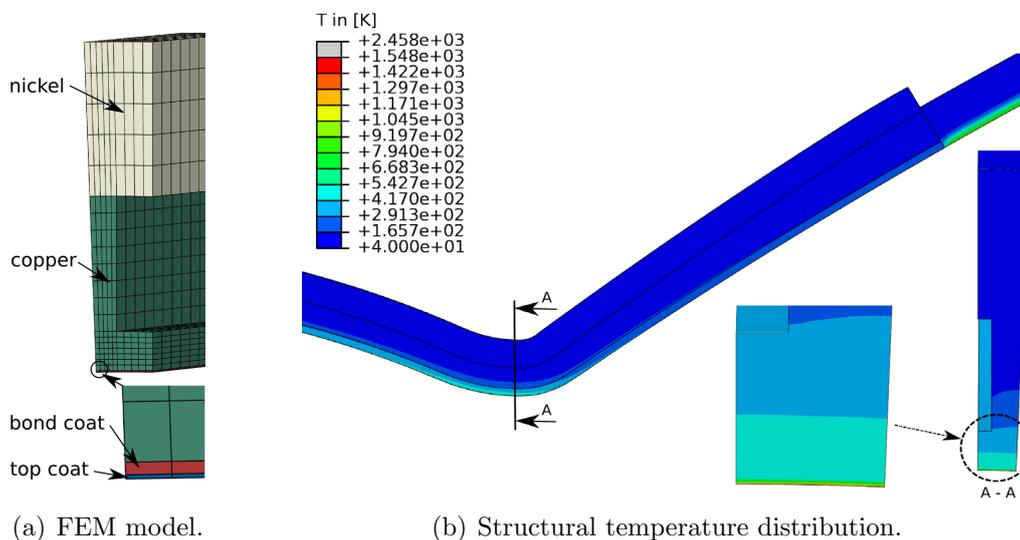


Figure 7: The cooled rocket engine at steady state including the TBC-system.

The parametrized modelling approach introduced in section 3 can be easily extended to analyse TBC systems in the context of a coupled thrust chamber analysis. In this study a standard TBC coating system consisting of a $30\mu\text{m}$ MCrAlY bond coat and a $10\mu\text{m}$ zirconia top coat is analysed. A small section of the parametrized FEM model is shown in Fig. 7(a). The thin layers of the TBC system are analyzed by 8-node continuum shells in which the copper and nickel alloy is modelled by 8-node solid elements like before. Material parameters of copper, BC and TBC are taken from [26, 27] and [24].

Fig. 7(b) shows the converged temperature distribution near the throat region of the structural domain including the TBC system. The temperature peak near the throat region reaches about 1548K. One can depict from the cut view shown in Fig. 7(b) that the temperature level of the copper alloy is drastically reduced compared to the conventional setup shown in Fig. 5.

5 CONCLUSIONS AND OUTLOOK

- Both physical couplings, thermal and mechanical interactions, have been analysed for a steady state case of a subscale thrust chamber. It was shown that the developed simulation environment ifls is capable of simulating FSI phenomena of cooled rocket thrust chambers.
- A novel parametrized 3D modelling approach for cooled rocket thrust chambers has been introduced. The approach is object oriented and easily extendable, which allows for future sensitivity analyses and quick design changes and most important, to understand in more depth the physical coupling phenomena involved.
- At last, the extensibility of the presented modelling approach was shown by a fully coupled steady state thrust chamber analysis including a TBC system applied inside the combustion chamber.
- In the ongoing research we are focused on the transient analysis of complete engine cycles at in-service conditions in order to address the limiting lifetime factors of cooled combustion chambers. In this context the cooling fluid domains analysed by RANS model will be integrated in the presented fully coupled approach.

Acknowledgments

Financial support has been provided by the German Science Foundation (Deutsche Forschungsgemeinschaft - DFG) in the framework of the Collaborative Research Centre Transregio 40. The authors would like to thank Astrium Space Transportation GmbH, Propulsion & Equipment for defining the subscale thrust chamber, which was studied in the present work.

REFERENCES

- [1] Sung, I.-K. and Anderson, W., A subscale-based rocket combustor life prediction Methodology. *AIAA, 2005-3570, 41st Joint Propulsion Conference* (2005).
- [2] Quentmeyer, R.J., Experimental fatigue life investigation of cylindrical thrust chambers. *AIAA, 77-893 13th Propulsion Conference* (1977).
- [3] Kuhl, D., Woschnak, A. and Haidn ,O.J., Coupled heat transfer and stress analysis of rocket combustion chambers. *AIAA, 98-3373, 34th Joint Propulsion Conference* (1998).
- [4] Fischer, S.C., Popp, M. and Quentmeyer, R.J., Thrust chamber cooling and heat transfer. *2nd International Symposium on Liquid Rocket Propulsion* (1995).

- [5] Kuhl, D., Riccius, J.R. and Haidn, O.J., Thermomechanical fluid-structure analysis and mathematical optimization techniques applied to cryogenic liquid rocket engine design. *4th International Symposium on Liquid Space Propulsion* (2000).
- [6] Riccius, J.R., Haidn, O.J. and Zametaev, E.B., Influence of time dependent effects on the estimated life time of liquid rocket combustion chamber walls. *AIAA, 2004-3670, 40th Joint Propulsion Conference* (2004).
- [7] Liu, Q.Y., Luke, E.A. and Cinella, P., Coupling heat transfer and fluid flow solvers for multidisciplinary simulations. *Journal of Thermophysics and Heat Transfer* (2005) **19**(4), 417–427.
- [8] Knab, O., Frey, M., Görgen, J., Maeding, C., Quering, K. and Wiedmann, D., Progress in combustion and heat transfer modelling in rocket thrust chamber applied engineering. *AIAA, 2009-5477, 45th Joint Propulsion Conference* (2009).
- [9] TAU-Code DLR, Technical documentation of the DLR TAU-Code. *Institut of Aerodynamics and Flow Technology Braunschweig* (2009).
- [10] Haupt, M.C., Niesner, R. and Unger, R., Computational Aero-Structural Coupling for Hypersonic Applications. *AIAA, 2006-3252, 9th Joint Thermophysics and Heat Transfer Conference* (2006).
- [11] Haupt, M.C., Niesner, R., Unger R. and Horst, P., Numerical Analysis of Thin-Walled Structures in Hypersonic Flow Regime. *ICAS 2008 Congress* (2008).
- [12] Haupt, M.C., Niesner, R., Horst, P., Esser, B. and Gülhan, A., Computational Analysis of a Heat Transfer Experiment Including Thermally Induced Deformations. *First Int. Conference on Comp. Meth. for Thermal Problems (ThermaComp2009)* (2009).
- [13] Haupt, M.C., Niesner, R. and Horst, P., Thermal-Mechanical Analysis of Functionally Graded Plates Subjected to Hypersonic Flow Conditions. *8th Int. Congress on Thermal Stresses* (2009).
- [14] Park, K. and Felippa, C., A Variational Principle for the Formulations of Partitioned Structural Systems. *Int. J. Numer. Meth. Engng.* (2000) **47**, 395–418.
- [15] Farhat, C., Lesionne, M. and Tallec, P.L., Load and Motion Transfer Algorithms for Fluid/Structure Interaction Problems with Non-Matching Discrete Interfaces: Momentum and Energy Conservation, Optimal Discretization and Application to Aeroelasticity. *Comp. Meth. Appl. Mech. Engng.* (1998) **157**, 95–114.
- [16] Park, K., Felippa, C. and Ohayon, R., Partitioned formulation of internal fluid-structure interaction problems by localized Lagrange multipliers. *Comp. Meth. Appl. Mech. Engng.* (2001) **190**, 2989–3007.

- [17] Unger, R., Haupt, M.C. and Horst, P., Application of Lagrange multipliers for coupled problems in fluid and structure interactions. *Computers and Structures* (2007) **85**, 796–809.
- [18] Mok, D., Wall, W. and Ramm, E., Accelerated iterative substructuring schemes for instationary fluid-structure-interaction. *First MIT Conference on Computational Fluid and Solid Mechanics* (2001) 1325–1328.
- [19] Niesner, R., Haupt, M. and Horst, P., Transient analysis methods for hypersonic applications with thermo-mechanical fluid-structure interaction. *3rd European Conference on Computational Mechanics Solids, Structures and Coupled Problems in Engineering* (2006).
- [20] Shewchuk, J., R., Delaunay refinement algorithms for triangular mesh generation. *Computational Geometry: Theory and Applications* (2002) **22**(1-3), 21–74.
- [21] Si, H., TetGen: A quality tetrahedral mesh generator and a 3d delaunay triangulator. <http://tetgen.berlios.de> (2006).
- [22] Rao, G.V.R., Exhaust nozzle contour for optimum thrust. *Jet Propulsion*, (1958) **28**(6).
- [23] Östlund, J., Flow processes in rocket engine nozzles with focus on flow separation and side-loads. *Licentiate Thesis, Royal Institute of Technology, Department of Mechanics, Stockholm* (2002).
- [24] Esposito, J.J. and Zabora, R.F., Thrust chamber life prediction volume I - mechanical and physical properties of high performance rocket nozzle material. *NASA CR 134806* (1975).
- [25] Ponomarenko, A., RPA: Design Tool for Liquid Rocket Engine Analysis. http://www.lpre.de/resources/software/RPA_en.htm (2009).
- [26] Deutsches Kupferinstitut, Kupferdatenblatt CuCr1Zr. (2005).
- [27] Freborg, A., M., Ferguson, B., L., Brindley, W., J. and Petrus, G., J., Modeling oxidation induced stresses in thermal barrier coatings. *Material Science and Engineering A* (1998) **245**, 182–190.

COUPLED REACTIVE TRANSPORT MODELING – THE PROGRAM TRANSPORT

VRATISLAV ŽABKA^{*}, JAN ŠEMBERA^{*}

^{*}Institute of new technologies and applied informatics (NTI)

Technical University of Liberec,
Studentská 2, 461 17 Liberec, Czech Republic

e-mail: vratislav.zabka@tul.cz

Key words: Column experiment, Tracing test; Groundwater modeling; Transport-reactive processes

Abstract. The contribution presents the program Transport, which serves to simulation of column transport experiments. Its main function is not to predict results of experiments but to compare influence of individual physical and chemical processes to the experiment results.

The one-dimensional advection-diffusion model is based on Finite Volume Method; it includes the triple porosity concept, sorption, retardation, and chemical reactions simulated using connected program React from The Geochemist's Workbench package or PhreeqC. Due to these geochemical programs, the user has extensive possibilities of chemistry simulation during transport. The program Transport simulates not only the processes inside the column but also preparation of entering solutions and measurement methods of outgoing solution parameters.

Important part of the contribution would be demonstration of results of simulation of real transport column experiments recently realized at the Technical University of Liberec.

1 INTRODUCTION

Modeling of transport processes is becoming increasingly important in recent years. There are many areas where it is necessary to model the transport of heat and water. This paper and the Transport program have been developed for the needs of projects dealing with study of contamination. But the modeling of transport actions is not sufficient for the contamination problem. The reactive processes are also very important, because the contaminants affect their surroundings primarily chemically. Transport and reactive components can interact, that is why we are talking about a coupled problem.

An example of impact of the reactive component to transport may be the precipitation of mineral phase. It can affect mobility of all observed components. An example of the opposite effect may be a dependence of type of ongoing reactions on flow velocity. We study those and other influences using simulations of the Transport program.

Studying of such processes in the underground, about its composition we have only limited information, is extremely difficult or almost impossible. If we want to detect patterns of events that take place under the ground, we must study them in known conditions first. In the laboratory we can split both processes and test them separately. The reactive processes

among solutions of different concentrations in flasks and transport processes using column experiments with simple solutions. Only when we can correctly simulate such processes separately, we can begin to combine them. The Transport program is used to simulate the combinations of these processes.

2 GENERAL SPECIFICATIONS

This paper aims to present the Transport program itself and design a methodology of its use for modeling of transport-reactive processes. The paper is divided into four main parts and conclusion. In the section 3 we present the Transport program and possibilities of its configuration. In this section we briefly describe the structure of calculation of the program itself as well as the structure of model problem.

An important part of the transport-reactive modeling is to understand the system from the perspective of chemical reactions. The importance of this phase of simulation will be discussed in the section 4. The Transport program itself does not calculate the reactive component of processes, but it uses the communication with geochemical software package from The Geochemist's Workbench (GWB hereafter). With this software, we can model the state of equilibrium of used solutions. This will give us knowledge of their properties and then we can compare them with changes within the transport.

A major problem in calculating the transport-reactive processes is their time-consuming intensity. In the section 5 we suggest a possible way to help. It is the Method of Contamination Front. This method is based on communication between the Transport program and the React program from GWB package. Presented method does not solve the problem of time-consuming intensity, but helps to properly set up the simulation. The resulting calculation will not be accelerated, but it is an important tool in studying the properties of processes in the stage of setting up the simulation.

The final part of the paper (section 6) will discuss the results of simulations in terms of accuracy and time-consuming intensity.

3 THE TRANSPORT PROGRAM V2.2

The Transport program is being developed at the Technical University in Liberec. It is a 1D model based on the Finite Volume Method. It can be used for simulations of column experiments, where the column is filled with a homogeneous porous rock saturated with homogeneous water solution and there are successively injected up to two different water solution. The model is adjusted so that before and behind the rock column there are so-called input and output chamber without a rock. It is well adapted for different kinds of measurements of output parameters of the solutions.

During its development our task was not to simplify the problem and to create a simple model. On the contrary, we expanded the simple model to resolve various types of experiments. The complexity of the model with more parameters emphasizes the accuracy of the modeled experiments. Then the information about the phenomena occurring in the column can be considerably more accurate.

Beside a detailed model of geometry of the column, it is difficult to improve the simulation of transport processes in any way. But large room for progress still remains in modeling of chemical reactions.

4 CHEMICAL REACTION MODELING

Following [1], while the coupling of hydrologic transport and chemical reaction models is an active area of research, the development of chemical reaction batch models has received much less attention. Whereas reactive parameters setting is more difficult than setting of transport parameters. Reactive transport program can not only compute with species concentrations. Information about other solute properties is important to include. Those properties are changing along the column experiment depending on current reactions and ambient conditions.

For example, setting of the solute and external atmosphere equilibrium is important for solute properties and inside chemical reactions. Otherwise setting of precipitation processes have an effect to solute composition and sometimes also transport properties. E. g. when column experiment takes only few days, it is not possible for hematite to precipitate; the mineral hematite is the final product of precipitation for solution including oxygen and iron but its precipitation needs at least hundreds of years and column experiments do not last as long, so we have to suppress this mineral in the thermodynamic equilibrium computations.

The Transport program includes some setups to operate with given aqueous solutions. Setting of concentrations and solute properties is an important part of operation with reaction component of the process. The Transport program user can perform complete analysis of given solutions before he starts to compute the main simulation of the column experiment. The user can also compute the equilibrium of the solute and the atmosphere. Another program option is setting of different ambient conditions in different parts along the column model. For example, the user can prohibit the access of oxygen inside of column but he can allow the calculation of equilibrium with the atmosphere in the output flask.

Coupling of reactive chemistry with transport significantly extends the computation time significantly. In this time, some trends can be observed in the effort to deal with this problem (parallel programming or development of numerical methods). Our approach to this problem is different. It is based on finding the contamination front which is almost independent of the kind of the applied software if we assume using the operator splitting method.

5 THE METHOD OF CONTAMINATION FRONT

The method of contamination front is based on reduction of the number of chemical equilibrium computations. By the “contamination front” we mean the situation where the solute concentrations significantly change in time. For example, this situation occurs when two different solutions mix because of advection. In such a situation, many chemical processes take place so it is important to compute the chemical equilibrium there. In other places of the column we can suppose that simulation of chemical processes can be omitted. The Method of Contamination Front (MCF) is based on looking for elements where the concentrations rapidly change. In this time, we are evolving two variants of the MCF: (1) testing the concentration change in time in each element, or (2) testing the concentration change along the flowpath in one time.

In the first case (variant 1), we are searching the contamination front in every single element. We wait for a significant change of concentration of the chosen species. The level of importance depends on the coefficient k (Eq. 1). If we set up the coefficient $k=0.5$, change of the specie concentration has to be larger than arithmetic mean of the past and the present

specie concentration.

$$|C_e(t) - C_e(t_{kont})| > (C_e(t) + C_e(t_{kont})) \cdot k \quad (1)$$

where C_e [mg l^{-1}] is concentration of the chosen species in the finite volume e , t_{kont} [s] is the time when the last concentration front was found, t [s] is the actual time, k [1] is the coefficient of the MCF variant 1.

In the second case (variant 2), we are searching for the contamination front according to the situation in neighboring elements. We check concentration of one species in every element using the equations (Eq. 2). Variant 2 uses only the concentrations in present time which can be an advantage for large 3D models:

$$\begin{aligned} C_- &= C_{e-1}(t) - C_e(t) \\ C_+ &= C_{e+1}(t) - C_e(t) \\ |C_- - C_+| &> |C_- + C_+| \cdot K \end{aligned} \quad (2)$$

where C_e [mg l^{-1}] is concentration of the chosen species in the finite volume e , K [1] is the coefficient of the MCF variant 2.

6 MODELING OF REAL EXPERIMENTS

In the previous three sections we have briefly discussed certain points of problems of transport-reactive processes. We will follow this introduction up with a specific example of simulation using the Transport program. In the introduction to this section we will set up a complex real problem, which we will divide into individual subtasks and present possible ways to solve them.

As it will be seen, this problem is very complex. Its solution is just at the beginning. At present, new experiments are still underway, which will help to better understand the issue. Therefore the final results of modeling of the specific task are not included at the end of the section. It is only an example of using the Transport program to simulate a particular phenomenon. Simulations of different systems may be done using the indicated procedure.

6.1 Description of the task

Currently, one of the objectives of our group is the use of iron nanoparticles for decomposition of organic contaminants in groundwater. A huge reactive surface is one of the known qualities of iron nanoparticles. Good migration property is their estimated quality. These miniature particles have the potential to penetrate even the smallest pores directly to the source of contamination. But their actual behavior is not well known, yet.

During the migration of nanoparticles through rock environment, their aggregation or sorption on the rock occurs. The exact physical and chemical principles of these phenomena are not yet sufficiently explained, or even modeled. For their explanation and for the discovery of other deviations from the behavior of larger particles we use the Transport program.

6.2 Simulation of tracer tests

One of the most important parts of the process is identifying the physical parameters of rocks and of the entire experiment. Errors made in their determination can lead to inaccurate or incorrect understanding of the characteristics of migrating nanoparticles.

The tracer test is used to determine the basic parameters of the experiment. To the column saturated with water or a solution of known properties we inject a saline solution and record the characteristics of outlet solution on the output. The most important observed parameter is the electrical conductivity; the secondary outlet is the pH of the solution. If we know the precise parameters of the column and the method of measuring the output characteristics, we can then calculate other properties of rocks. These include porosity, which can have substantial effect on the monitored processes.

Before we start modeling of the transport it is necessary to analyze the reactive components of the processes. By testing chemical properties of individual solutions and their combinations we gain important information about progress of the experiment. In the case of tracer tests using a saline solution and distilled water does not occur any complex chemical processes. From this analysis, we found the estimated value of pH.

6.3 Outputs of the experiment

The values of conductivity can be modeled very accurately. Conductivity for low concentration of the solution depends directly on the concentration of salt. Simulation of the pH of the output solution was not accurate (Figure 1). The difference of pH values in order of tenths is indeed very difficult to measure and thus very difficult to compare with the model. In our case a variation in the trend of changes in pH was observed. It was a problem as in general the trend of development of a certain quantity is more important than its specific value in geochemistry.

The measured change in pH occurred when the composition of output solution was changing. But the change was always more distinctive at the beginning of this change than supposed the model. The expected value was approached after a certain time. Using the Transport program, we tried to assess what might cause this unexpected deviation. By testing known options, we concluded that the determining phenomenon is the exchange of sodium and calcite ions at aluminosilicate clays that were present in the rock in the experiment. As can be seen from the chart, when the simulation included the sorption of sodium, the trend of the simulation corresponds much better to the measured values.

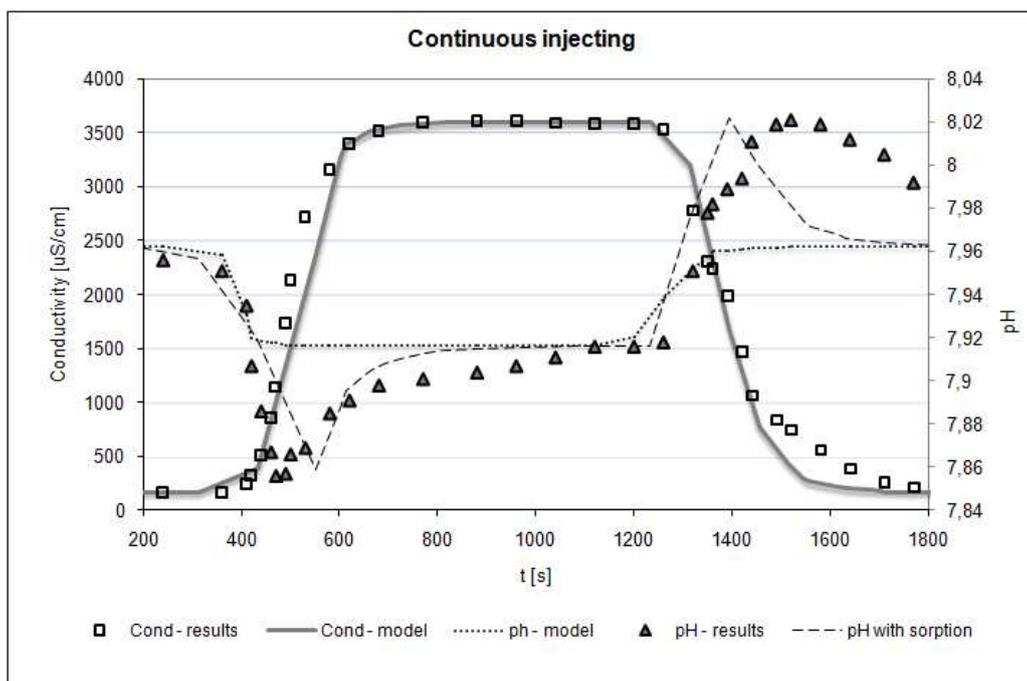


Figure 1: Continuous injecting – dependence conductivity and pH on time of the experiment.

To verify the correctness of our hypothesis, we proposed a further test experiment. Instead of rock, the column was filled with glass balls of size of tens of micrometers. Parameters of the column corresponded, but the exchange between ions in solution and ions of geological environment was excluded. The trend in pH during the experiment with balls and with rock remained unchanged, the hypothesis was thus refuted. We are still trying to attempt this phenomenon. Experiments are in progress.

6.4 Analysis of results and discussion

On this simple example, we wanted to demonstrate some of the ideas that are important in modeling of coupled problems not only in geochemistry. Also on this problem, we tested the time required for the calculation and compared it with the calculations using the method of front contamination.

The basis for modeling of coupled processes is a detailed knowledge of the various components of these processes. For this case of modeling of nanoiron migration, the first necessary is to test the transport and chemical properties of the monitored system. That is in our case the tracer test and monitoring of interactions between particles of nanoiron with different solutions and rock environment.

Although we believe that we know the observed system and we are able to describe its deviations from the normal state, it is useful to doubt our results and perform test simulations under various conditions over and over again. Only by the reactions of the system under various conditions we can learn more about it.

On the model tracer tests, we also compared the influence of the coefficient of the method of contamination front and its other settings to its time-consumption. The calculated results

were then compared with the total calculation time without using the previously mentioned method (Table 1). The comparison is only approximate, because the computing time can vary with the complexity of simulated systems. The faster calculation method can be used, the system is the simpler. For example, variant 1 of MCF with $K = 0.01$ can be used for complex actions including precipitation. And even this version is many times faster than the complete calculation.

Table 1: Examples of computation times using two variants of MCF with different coefficients. Average deviation is comparing of results of the MCF variant and complete reactive transport computation.

Variants of the method	Computation time	Average deviation [%]
Complete reactive transport	19h 16min 57s	
MCF Variant 2, $K=0.1$	9h 8min 39s	0.01
MCF Variant 2, $K=0.3$	2h 43min 32s	1.58
MCF Variant 1, $k=0.01$	44min 29s	0.21
MCF Variant 1, $k=0.1$	16min 27s	5.52

7 CONCLUSION

This paper contains a brief description of the Transport program. This program is a powerful tool for studying both simple and complex processes occurring during the column experiment. This is achieved by the wide range of setting transport and reactive parameters of the experiment. The paper also briefly describes the methodology for modeling of transport-reactive processes.

The paper also proposes a method of speeding up the work on modeling of transport-reactive processes by looking for the contamination front. Although the method itself speeds up the resulting simulation only at the cost of losing some information, it still can be a valuable tool in finding the correct parameters resulting simulation. It does not need to be limited just to the 1D problem. Currently we are working on its extending to more complex problems.

Last part of the paper is a demonstration of some simple simulation of column experiments. Their example demonstrates a possible way to work with the Transport program. The results were also compared to time simulations using the method of contamination front.

Besides introducing suitable instruments for the study of transport-reactive processes, the authors of this paper wanted to draw attention to the complexity of the modeled processes. Without proper understanding of elementary processes, it is not possible to adequately simulate complex actions that are dependent on elementary ones.

When creating models of complex systems it is necessary to simplify them. But we often certainly do not know if we can make some simplification. If it is a study of coupled problems of transport-reactive processes, the Transport program is a suitable tool for studies of these systems.

Acknowledgements This result was realized under the state subsidy of the Czech Republic within the research and development project “Advanced Remedial Technologies and Processes Center” 1M0554 - Programme of Research Centers supported by Ministry

of Education and within the research project FR-TI1/456 “Development and implementation of the tools additively modulating soil and water bioremediation” - Programme MPO-TIP supported by Ministry of Industry and Trade.

REFERENCES

- [1] Fang, Y. *Reactive chemical transport under multiphase system*. Ph.D. Dissertation, Dept. of Civil and Environmental Engineering, The Pennsylvania State University, University Park, PA, (2003)
- [2] G. Yeh, Y. Fang, F. Zhang, J. Sun, Y. Li, M. Li, and M.D. Siegel *Numerical modeling of coupled fluid flow and thermal and reactive biogeochemical transport in porous and fractured media*. *Computational Geosciences* (2009), vol. 14, pp. 149-170.
- [3] Glynn, P.D., Plummer, L.N. *Geochemistry and the understanding of ground-water systems*. *Hydrogeology Journal* (2005), 13(1), 263–287.
- [4] Prommer, H., Barry, D.A. & Zheng, C. *MODFLOW/MT3DMS-based reactive multicomponent transport modeling*. *Ground Water* (2003), 41(2), 247-257.

THERMOELASTIC STRESS ANALYSIS FOR A TUBE UNDER GENERAL MECHANOCHEMICAL CORROSION CONDITIONS

YULIA G. PRONINA

Saint-Petersburg State University
Faculty of Applied Mathematics and Control Processes
Universitetskii pr. 35, Saint-Petersburg, 198504, Russian Federation
e-mail: dr.JulyP@gmail.com, <http://eng.spbu.ru>

Key words: Elastic Tube, Pipeline, Pressure Vessel, Mechanochemical Wear, Equal-Rate Corrosion, Temperature Stress

Abstract. This work deals with the uniform surface mechanochemical corrosion of an elastic thick-walled long cylindrical tube subjected to internal and external pressure of environments at different temperatures. The rate of uniform corrosion is supposed to be linear with the stress and inversely as the exponent of time. The problem is then reduced to the first-order ordinary differential equation in a tube cross-sectional size. Analytical solutions of this equation are found. To determine the tube lifetime and the cause of its failure, the method based on various estimating functions is proposed. The algorithm for environmental contamination by corroded material is developed.

1 INTRODUCTION

Most structures are exploited being subjected to both mechanical loads and operating environments. This often causes the process of so called mechanochemical wear which is more intensive in comparison with the simple superposition of damages induced by mechanical stresses and electrochemical corrosion taken separately. The rate of corrosion depends on many factors. It is to be noted that mechanical stress does not affect the corrosiveness in neutral and weakly acid media. According to most experimental results [1, 2, 3], in other media the rate of uniform corrosion often is linear with the stress when traction increases beyond a given threshold. This depends on the stress sign and properties of the "material-medium" system. Furthermore, the corrosion rate is inversely as the exponent of time if closed oxide layer leads to the inhibition of corrosion.

This paper is concerned with equal-rate mechanochemical corrosion often observed in practice. The theoretical research in this area has been conducted by numerous authors. A comprehensive review of models and calculations for structures taking into account

corrosive wear was given e.g. in [3, 4]. Among the first works in the field there were some articles by V.M.Dolinskii, concerned with mechanochemical corrosion of a thin-walled structural members [1, 5]. Elegant mathematical apparatus was presented in the work [6] where the lifetime of a loaded pipe had been assessed under the assumption of the exponential dependence of the corrosion rate on the mean stress. Using the linear relation between the corrosion rate and the stress, the corrosive wear of a nonlinearly elastic cylinder subjected to pressure and temperature was simulated by numerical methods in the book [3]. In the articles [7, 8] the equal-rate mechanochemical corrosion of a linearly elastic thick-walled cylindrical tube subjected to a longitudinal force, internal and external pressure was discussed without taking temperature stresses into consideration. The problem was then reduced to the first-order ordinary differential equation in either tube cross-sectional sizes or the maximum principal stress as the situation requires. The purpose of this work is to study the uniform mechanochemical corrosion of a hollow cylinder under the joint action of pressure and temperature.

2 PROBLEM STATEMENT

The uniform surface corrosion of an elastic long thick-walled cylindrical tube is investigated. The tube is subjected to constant internal pressure p_r of a corrosive medium at temperature T_r and external pressure p_R of another corrosive medium at temperature T_R . The inner and outer tube radii at the initial time $t = 0$ are denoted by r_0 and R_0 ($r_0 < R_0$). The action of the ends of the cylinder is not taken into account. Changes of the tube radii are assumed to be quasi-static. The corrosion rates at the internal ($\rho = r$) and external ($\rho = R$) boundaries are given correspondingly by the expressions [2, 3]:

$$v_r = \frac{dr}{dt} = [a_r + m_r \sigma_1(r)] \exp(-bt) \exp(\beta_r [T_r - T_r^0]) \quad \text{at} \quad |\sigma_1(r)| \geq |\sigma_r^{th}|, \quad (1)$$

$$v_R = -\frac{dR}{dt} = [a_R + m_R \sigma_1(R)] \exp(-bt) \exp(\beta_R [T_R - T_R^0]) \quad \text{at} \quad |\sigma_1(R)| \geq |\sigma_R^{th}|. \quad (2)$$

Here, $a_r, a_R, m_r, m_R, b, \beta_r, \beta_R, T_r^0$ and T_R^0 are observable quantities; $a_r = v_r^0 - m_r \sigma_r^{th}$ and $a_R = v_R^0 - m_R \sigma_R^{th}$; σ_r^{th} and σ_R^{th} are the threshold stresses (as a matter of fact, which are different for traction and compression); v_r^0 and v_R^0 are the initial corrosion rates at $|\sigma_1(r)| < |\sigma_r^{th}|$ and $|\sigma_1(R)| < |\sigma_R^{th}|$, respectively; σ_1 is the maximum principal stress.

It is necessary to trace the change of the thermoelastic stresses and amount of corroded material with time t and to estimate the tube lifetime.

3 BASIC EQUATIONS

The problem of a tube under pressure and temperature has been discussed by numerous writers including G. Lamé and R. Lorenz [9, 10]. At $r = 0$, $p_R = p$ or $p_r = p_R = p$ and $T_r = T_R$, there is a homogeneous stress $\sigma_{\theta\theta} \equiv \sigma_{\rho\rho} \equiv -p$ irrespective of corrosion. In other

cases, the stress-components are expressed, by reference to cylindrical coordinates ρ, θ, z , by the equations

$$\sigma_{\theta\theta}(\rho) = \frac{(p_r + T_\sigma/2)r^2 - p_R R^2}{R^2 - r^2} + \frac{p_r - p_R + T_\sigma/2}{R^2 - r^2} \frac{r^2 R^2}{\rho^2} + T_\sigma \frac{\ln(R/\rho) - 1}{2 \ln(R/r)}, \quad (3)$$

$$\sigma_{\rho\rho}(\rho) = \frac{(p_r + T_\sigma/2)r^2 - p_R R^2}{R^2 - r^2} - \frac{p_r - p_R + T_\sigma/2}{R^2 - r^2} \frac{r^2 R^2}{\rho^2} + T_\sigma \frac{\ln(R/\rho)}{2 \ln(R/r)},$$

$$\sigma_{zz}(\rho) = \frac{p_r r^2 + (T_\sigma - p_R)R^2}{R^2 - r^2} - T_\sigma \frac{1 + 2 \ln(\rho/r)}{2 \ln(R/r)}, \quad (4)$$

where

$$T_\sigma = \frac{\alpha E}{1 - \nu} (T_R - T_r),$$

α is thermal expansion coefficient, E is Young's modulus, ν is Poisson's ratio. Formula (4) holds true for a closed cylindrical vessel with allowance for the total pressure to its bottoms.

If stress does not affect the corrosion rate, the stress-components at any time can be easily calculated by the above equations (3)–(4) for given laws of change of radii $r(t)$ and $R(t)$. This is valid for neutral and alkaline media and when load is lower than the threshold value σ^{th} . We now study mechanochemical corrosion when the conditions (1) and (2) hold true. The solution to the problem for $T_r = T_R$ and $p_r \neq p_R$ has been given in [7, 8]. We now consider other situations when $p_r \neq p_R$ and $T_r \neq T_R$.

It is evident that the maximum principal stress is the circumferential one: $\sigma_1(\rho) = \sigma_{\theta\theta}(\rho)$. This stress is then to be used in formulae (1) and (2). If $T_r = T_R$, then the greatest tension is at the inner surface: $\sigma_{\theta\theta}(r) \geq \sigma_{\theta\theta}(R)$. So we must follow the amount of $\sigma_{\theta\theta}(r)$ to determine the time t^* when $\sigma_{\theta\theta}(r)$ reaches the limiting stress. But sometimes, heat stress can compensate mechanical one. For example, $\sigma_{\theta\theta}(r) < \sigma_{\theta\theta}(R)$ when $0 < p_r - p_R < -T_\sigma$. When this is the case, we have to watch the $\sigma_{\theta\theta}(R)$. Let the greatest stress be the stress $\sigma_1(r) = \sigma_{\theta\theta}(r)$. The equation (3) gives

$$\sigma_{\theta\theta}(r) = \frac{(p_r - 2p_R + T_\sigma)\eta^2 + p_r}{\eta^2 - 1} - \frac{T_\sigma}{2 \ln \eta}, \quad (5)$$

where

$$\eta = \frac{R}{r} = \frac{R_0 - \delta_R}{r_0 + \delta_r}. \quad (6)$$

For the relation $\eta = R/r$ can not be expressed from (5), it is impossible to derive the differential equation in $\sigma_{\theta\theta}(r)$ or $\sigma_{\theta\theta}(R)$, as it was made in [7]. So let us deduce an equation in η .

If we eliminate σ_1 from the formulae (1) and (2) using (3), we can obtain the relationship

$$RM_r + rM_R = M_R \left(r_0 - \frac{B_r}{b} [\exp(-bt) - 1] \right) + M_r \left(R_0 + \frac{B_R}{b} [\exp(-bt) - 1] \right), \quad (7)$$

where

$$B_r = a_r \exp(\beta_r [T_r - T_r^0]), \quad M_r = m_r \exp(\beta_r [T_r - T_r^0]),$$

$$B_R = (a_R - m_R [p_r - p_R + T_\sigma]) \exp(\beta_R [T_R - T_R^0]), \quad M_R = m_R \exp(\beta_R [T_R - T_R^0]). \quad (8)$$

On differentiating the expression (6) with respect to t and then using (5) and (7)–(8), we can deduce the ordinary differential equation for changing η

$$\frac{d\eta}{dt} = -\frac{M_R + M_r \eta}{\exp(bt)} \frac{B_R + B_r \eta + (M_R + M_r \eta) \left[\frac{(p_r - 2p_R + T_\sigma)\eta^2 + p_r}{\eta^2 - 1} - \frac{T_\sigma}{\ln \eta^2} \right]}{M_R r_0 + M_r R_0 - \left(M_R \frac{B_r}{b} - M_r \frac{B_R}{b} \right) [\exp(-bt) - 1]}. \quad (9)$$

The initial condition to be satisfied at $t = 0$ is

$$\eta_0 = \frac{R_0}{r_0}. \quad (10)$$

4 SOLUTIONS OF THE BASIC DIFFERENTIAL EQUATION

The integral of the equation (9), satisfying the condition (10), is

$$t = -\frac{1}{b} \ln \left\{ 1 - b \frac{M_R r_0 + M_r R_0}{M_R B_r - M_r B_R} \left(\exp[(M_R B_r - M_r B_R) J(\eta)] - 1 \right) \right\},$$

where

$$J(\eta) = \int_{\eta}^{\eta_0} \frac{(\eta^2 - 1) \ln \eta^2}{M_R + M_r \eta} \left\{ [B_R + B_r \eta](\eta^2 - 1) \ln \eta^2 + [M_R + M_r \eta] \{ [(p_r - 2p_R + T_\sigma)\eta^2 + p_r] \ln \eta^2 - T_\sigma(\eta^2 - 1) \} \right\}^{-1} d\eta.$$

For one-sided corrosion, the analytical solution can be simplified. For example, at $M_r = 0$ and $B_r = 0$ (external corrosion), the result may be written in the form

$$t = -\frac{1}{b} \ln \{ 1 - bJ(\eta) \} \quad \text{at } b \neq 0, \quad (11)$$

$$t = J(\eta) \quad \text{at } b = 0,$$

where

$$J(\eta) = r_0 \int_{\eta}^{\eta_0} \frac{(\eta^2 - 1) \ln \eta^2 d\eta}{\ln \eta^2 \{B_R(\eta^2 - 1) + M_R[(p_r - 2p_R + T_\sigma)\eta^2 + p_r]\} - M_R T_\sigma(\eta^2 - 1)}.$$

For internal corrosion, at $M_R = 0$ and $B_R = 0$, the solution is of the form (11), where

$$J(\eta) = R_0 \int_{\eta}^{\eta_0} \frac{(1 - 1/\eta^2) \ln \eta^2 d\eta}{\ln \eta^2 \{B_r(\eta^2 - 1) + M_r[(p_r - 2p_R + T_\sigma)\eta^2 + p_r]\} - M_r T_\sigma(\eta^2 - 1)}.$$

The solutions of the basic differential equation give us the t -to- η corresponding. Using the equation (5), we can then calculate the stress $\sigma_{\theta\theta}(r)$ (and other stress-components) at any time t .

5 ASSESSMENT OF ENVIRONMENTAL CONTAMINATION

Using above relations, it is also possible to assess the amount of the corroded material at the inner and outer surfaces. This material is solved (evaporated) in the environments or precipitated into corrosion products (oxide film) as the case may be. The volume of the inside and the outside corroded material per unit length at any t are evaluated correspondingly by the formulae

$$V_r^p = \pi(r^2 - r_0^2), \quad V_R^p = \pi(R_0^2 - R^2),$$

where

$$r = r_0 + \delta_r = \frac{M_R r_0 + M_r R_0 - \left(M_R \frac{B_r}{b} - M_r \frac{B_R}{b}\right) [\exp(-bt) - 1]}{\eta M_r + M_R},$$

$$R = R_0 - \delta_R = \eta r,$$

η is uniquely determined by the integral curve of the differential equation (9) for the t involved.

If the internal and external environments are closed, then it is easy to calculate the impurity concentration in it at any t using above relations and chemical equations.

When the volume $V^l = \pi r_1^2$ of internal environment flows in a pipe for a known period from t_1 to t_2 , the volume of corroded material V_r^p per the environment volume (e.g., concentration of pipe material solved in the environment volume) can be estimated by the formulae

$$C = \frac{V_r^p}{V^l} 100\% = \frac{1}{2} \left(\frac{r_2^2}{r_1^2} - \frac{r_1^2}{r_1^2} \right) 100\% \quad \text{or} \quad C = 4 \frac{r_2 - r_1}{r_1 + r_2} 100\%,$$

valid for $r_2 - r_1 \ll r_1$ and $t_2 - t_1 \ll t^*$, where $r_1 = r|_{t=t_1}$, $r_2 = r|_{t=t_2}$; t^* is the predicted lifetime of a pipe.

6 LIFETIME PREDICTION

Taking into account interference of general corrosion and thermomechanical stress, lifetime of a tube may be estimated. It is obvious that pipe failure can be initiated by many causes. The life of the body is determined by the minimum time it takes for a limiting state to be achieved owing to any case: when the ultimate tensile strength is reached, upon loss of stability, upon damage accumulation (e.g., during quasi-static cyclic loading or aging) and so on. To determine the rupture source and the life of a tube it is reasonable to introduce scalar estimating functions. Following the L. Kachanov approach [11], various kinds of damage are represented by dimensionless functions varying from zero to unity (or from $-\infty$ to 1) and mounting to unity in the moments of fault related to a concrete criteria.

To assess the durability functions of the type $\Pi_s(t) = \frac{f(\sigma, \epsilon)}{f_s} \leq 1$ may be used. For the maximum normal stress criterion, we can write $\Pi_s(t_s^*) = 1$, where $\Pi_s(t) = \frac{\sigma(t)}{\sigma_s(t)}$, $\sigma_s(t)$ is the ultimate tensile strength of material that may change in time.

Another approach to determine the rupture life consists in damage accumulation assessment. For instance, according to Bailey's principle, the time to destruction t_d^* is determined by the equation $\Pi_d(t_d^*) = \int_0^{t_d^*} \frac{dt}{\tau[\sigma(t), T]} = 1$, where $\tau[\sigma, T]$ is the working life of the material under the stress σ and temperature T .

Functions to assess the stability factor may be of the form $\Pi_{cr}(t) = \frac{\sigma_{zz}(t)}{\sigma_{zz}^{cr}(t)} + \frac{\sigma_{\theta\theta}(t)}{\sigma_{\theta\theta}^{cr}(t)} \leq 1$, where σ_{zz}^{cr} is buckling stress for the tube under only longitudinal force (under no pressure), $\sigma_{\theta\theta}^{cr}$ is the upper critical stress for the tube under pressure (in the absence of axial force), expressed through the current tube sizes and mechanical quantities. It is obvious that the cylindrical shape is always stable at $\Pi_{cr}(t) \leq 0$. Stability of thin-walled shells under conditions of corrosive action have been investigated by many scientists, e.g. [12, 13].

Furthermore, failure may be determined apparently by accidental circumstances. For such assessment we can introduce estimating function as being equal to the probability, i.e. accident risk $\Pi_p(t) = P(x_1, \dots, x_n, t/y_1, \dots, y_n)$. It is to be emphasized that the unreliability function depends on the other estimating functions.

The graphs of all the estimating functions are plotted on one figure and compared with each other. The curve being the first to reach unity indicates the most probable cause of breakdown and the lifetime of an item

$$t^* = \min\{t_i^* : \Pi_i(t_i^*) = 1\}.$$

Computation results have shown that an increase in corrosion inhibition index b considerably increases the service life of the tube. At sufficiently high values of b , corrosion

can practically finish without reaching the critical state. In this case, the life is controlled by damage accumulation (decrease in the ultimate strength) or random factors.

7 CONCLUSIONS

- The problem under study has been reduced to the first-order ordinary differential equation. The analytical solution of this equation has been obtained.
- The algorithm for environmental contamination has been developed.
- To predict the life of a tube various estimating functions have been proposed.

8 ACKNOWLEDGEMENTS

This work was supported by the Russian Foundation for Basic Research, project N 10-01-00093-a.

The author also acknowledges Saint-Petersburg State University for the research grants N 9.0.165.2009 and N 9.37.129.2011.

REFERENCES

- [1] Dolinskii, V.M. Calculation of stressed pipes subjected to corrosion. *Khim. Neft. Mashinostr.* 1967, N 2: 18–21.
- [2] Pavlov, P.A., Kadyrbekov, B.A. and Kolesnikov, V.A. *Strength of steels in corrosive environments*. Nauka, Alma-Ata, (1987). [in Russian].
- [3] Naumova, G.A. and Ovchinnikov, I.G. *Strength calculations of compound frameworks and pipeline structures with taking into account corrosion damages*. Saratov State University, Saratov, (2000). [in Russian].
- [4] Ovchinnikov, I.G. and Pochtman, Yu.M. Calculation and rational design of structures subjected to corrosive wear (review). *Materials Science.* (1992) **27**, N 2: 105–116.
- [5] Dolinskii, V.M. and Sirotenko, V.A. Calculation of heat exchange systems of reduced design subjected to the effect of a corrosive medium. *Khim. Mashinostr.* 1970, N 11: 4–11.
- [6] Gutman, E.M., Zainullin, R.S., Shatalov, A.T. et al. *Strength of gas industry pipes under corrosive wear conditions*. Nedra, Moscow, (1984). [in Russian].
- [7] Pronina, Yu.G. Analytical study of general mechanochemical corrosion of the pipe under the axial force and pressure. *7th Euromech Solid Mechanics Conference (Lisbon, Portugal, September 7-11, 2009): CD-Rom Proceedings*. European Mechanics Society, APMTAC. CD-Rom\papers\Gs-sm\pap_0486_GS-SM.pdf.

- [8] Pronina, Yu.G. Estimation of the life of an elastic tube under the action of a longitudinal force and pressure under uniform surface corrosion conditions. *Russian Metallurgy (Metally)*. (2010) **2010**, N 4: 361–364.
- [9] Lamé, G. *Lecons sur la theorie mathematique de l'elasticite des corps solides*. Paris, (1852).
- [10] Lorenz, R. Temperaturspannungen in hohlzylindern. *Zeitschrift des vereines Deutscher ingenieure*. (1907) **51**: 743–747.
- [11] Kachanov, L.M. *Introduction to continuum damage mechanics*. Martinus Nijhoff Publishers, Dordrecht, Netherlands, (1986).
- [12] Karpunin, V.G., Kleshchev, S.I. and Kornishin, M.S. Endurance of plates and shells under conditions of the corrosive action of a medium. *Strength and Endurance of Structures*. Naukova Dumka, Kiev, (1980), 35–44. [in Russian].
- [13] Gutman, E.M., Haddad, J. and Bergman, R. Stability of thin-walled high-pressure vessels subjected to uniform corrosion. *Thin-Walled Structures*. (2000) **38**: 43–52.

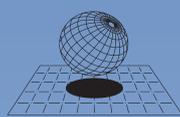
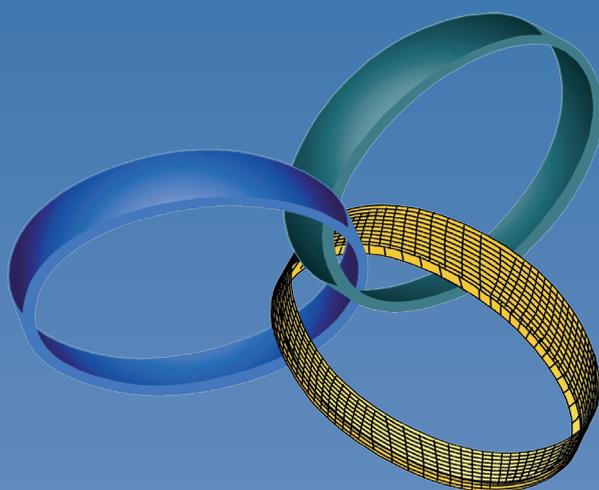
Authors Index

Adamczak C.	90	Chassaing J-C.	1007
Adnet N.	337	Chen P.R.	590
Al-Khoury R.	67	Choi H.S.	404
Alexiou T.	710	Chong O.Y.	946
Alexiou T.S.	1120	Codina R.	31
Alifanov O.M.	1155	Comi C.	55
Allery C.	919	Completo A.	1251
Alonso E.	194	Corigliano A.	448
Alotto P.	367	Coussirat M.	986
Altmann C.	355	Cramer F.	1342
Ammar A.	919	Cyron C.J.	23
Anderson K.S.	325	Dal Pont S.	1318
Arai N.	974	Dallocchio A.	413
Arairo W.	767	Datcheva M.	76
Ardito R.	448	David S.A.	1022
Artiukhine E.A.	1155	de la Fuente P.	721
Atzrodt H.	1282	de Oliveira C.	1022
Aubert S.	647	De Santis D.	289
Autuori P.	1318	De Vuyst F.	1200
Avalle M.	578	Degrande G.	934
Badia S.	31	Degroote J.	100, 934
Baek M.K.	398	Deü J-F.	802
Bahai H.	954	Dieng L.	1128
Baimpas N.	1379	Dinkler D.	837, 1342
Baqué B.	1326	Djeran-Maigre I.	767
Beer G.	301	Doležel I.	388, 1167
Behrens B-A.	277	Donjat D.	1326
Belahcen A.	472	Dubini G.	436
Belnoue J.P.	1095	Duchaine F.	826, 885
Beringhier M.	570	Dumon A.	919
Bermudez B.	996	Egusquiza E.	986
Bertarelli A.	413	Ekaterinaris J.A.	147
Bertarelli E.	436, 448, 1062	Emans M.	184
Bialecki R.A.	1371	Errera M.P.	1326
Bianchi E.	436	Fabbri A.	230
Bierwisch C.	1071	Favre L.	1033
Bijl H.	635, 779, 814, 850, 873	Ferfecki P.	424
Bíró O.	460	Ferrand P.	647
Blanc L.	647	Feyel F.	1326
Bletzinger K-U.	731	Fincato R.L.	548, 559
Boffi D.	135	Fityus S.	755
Boman R.	254	Flitz F.	925
Böttcher N.	206	Fontanals A.	986
Bouache T.	254	Fonteyn K.	472
Bruant I.	337	Forestieri G.	1176
Brummer T.	590	Frischkorn J.	1350
Bui D.	1200	Frommann O.	1306
Bursi O.S.	896	Gaillet L.	1128
Bussetta P.	242	Gardini F.	135
Canivenc R.	254	Gastaldi D.	1062
Carli V.	1282	Gastaldi L.	135
Carnelli D.	1062	Geraci G.	289
Carrera E.	623	Gicquel L.Y.	826
Carretta Y.	254	Giglio M.	1107, 1294
Casari F.	1062	Gonnet J-P.	347
Cavallini N.	135	Gonzalez F.	313
Čermák L.	424	González M.	792

Göransson P.	802	Laroche E.	1326
Görke U-J.	160, 206, 218	Laugier M.	254
Grandidier J-C.	570	Laure P.	1188
Grazina J.C.	698	Lee A.	946
Groth C.	1212	Lefebvre-Lepot A.	1188
Guardo A.	986	Legrand N.	254
Guardone A.	289, 1176	Lenglet C.	1318
Gurnett T.	313	Lindhorst K.	661
Hamdaoui M.	1200	Lobry L.	1188
Hammoud M.	570	Lolo W.	647
Harding S.	1379	Lombaert G.	934
Haupt M.C.	661, 1388	Longatte E.	1007
Hayhurst D.R.	484	Lorbiecka A.Z.	1220
Hellenthal L.	1212	Madani S.H.	954
Hernández A.	677	Magale C.	367
Hetegger M.	460	Majorana C.E.	530, 538, 548
Heuss O.	1282	Malinowski L.	301
Hillewaere J.	934	Manes A.	1107, 1294
Hinkelmann R.	90	Marotti de Sciarra F.	496
Hissel F.	1261	Martinez L.	1033
Hofmann M.	686	Marulli F.	1176
Hofstetter G.	686	Matsui K.	1045
Horst P.	1388	Matsumoto Y.	600
Imre E.	755	Mazauric V.	347
Isola D.	1176	Mazzucco G.	530, 548
Jacquemet N.	230	Meftah F.	1318
Jaindl M.	367	Meier C.	23
Jauré S.	826	Meschke G.	1083
Jia C.	896	Mohring J.	1282
Jus Y.	1007	Molinari A.	1062
Kähkönen S.	908	Mönkölä S.	908
Kampitsis A.E.	743	Montmitonnet P.	254
Kapellos G.	710, 1120	Morel Th.	885
Karban P.	1167	Moseler M.	1071
Karlis G.F.	301	Mostoufi N.	1054
Kazakidi A.	147	Müller K.	23
Kazemi-Kamyab V.	850	Munz C-D.	355
Kebriaei R.	1350	Naceur H.	1261
Kelson N.	313	Nastase A.	1273
Kim Y.S.	398, 404	Nelson D.H.	962
Kireenkov A.	1147	Nenarokomov A.V.	1155
Klimanek A.F.	1371	Nestorović T.	379
Klöppel T.	23	Neudorfer J.	355
Kok J.	313	Nguyen-Tuan L.	76
Kolditz O.	160, 206, 218	Nicolas A.	996
Koliji A.	172	Ofner G.	460
Korsunsky A.M.	1095, 1379	Okita K.	600
Korvink J.G.	436	Ono K.	600
Köstinger A.	367	Ostermann L.	1342
Kotlan V.	388	Oteo C.	721
Koudelka T.	521	Ouahsine A.	1261
Kouhia R.	472	Pablo F.	337
Kowalsky U.	1342	Panoskaltzis V.P.	506
Kowollik D.	1388	Park C-H.	218
Kreeft J.J.	873	Park I.H.	398, 404
Krimm R.	277	Pavlou S.	710, 1120
Kruis J.	521	Peguiron J.	1071
Kurch M.	1282	Pereira C.	1033
Kutschera R.	367	Pereira J.M.	230
Lagger H.	1071	Périer V.	1128
Laloui L.	172	Peroni L.	413, 578
Lang H.	460	Pertl M.	686

Petrolo M.	623	Stermecki A.	460
Phan M.T.	1318	Sternel D.C.	925
Philit M.	647	Stickle M.M.	721
Piacentini A.	885	Stock A.	355
Pierret J-C.	265	Strohmeier D.	436
Pignatelli R.	55	Strömberg N.	1136
Pinto P.L.	698	Sung T.I.	404
Planas R.	31	Suzuki M.	1045
Ponthot J-P.	242, 254	Tahmasebpoor M.M.	1054
Pourplanche A.	1261	Takagi S.	600
Poursina M.	325	Takahashi S.	974
Price C.D.	962	Takizawa K.	43, 590, 611
Price M.A.	946	Talebian M.	67
Pronina Y.	1408	Taron J.	160, 218
Proslie L.	337	Terragni F.	861
Prunier F.	767	Tezduyar T.E.	43, 590, 611
Pudewills A.	1362	Thévenin A.	885
Quaranta G.	1176	Thouverez F.	647
Quintino D.D.	1022	Tini V.	1332
Rainer S.	460	Titov D.M.	1155
Ramos A.	1251	Trajkov M.	379
Rang J.	112	Tran K.T.	1261
Rapún M-L.	861	Triantafyllou S.P.	506
Rassili A.	265	Tsakiris D.P.	147
Reddy B.D.	1239	Ulrych B.	388, 1167
Reese S.	1332, 1350	Valdés J. G.	677
Reinstädler S.	837	Vallin V.	135, 230
Relvas C.	1251	van Brummelen E.H.	124
Repin I.V.	1155	van Opstal T.M.	124
Resende U.	1232	van Zuijlen A.H.	635, 779, 814, 850, 873
Reynolds Q.G.	1239	Varello A.	623
Rigobert S.	1318	Vavourakis V.	147
Rojek J.	301	Vega J.M.	861
Rondot L.	347	Vena P.	1062
Roos A.	1326	Verdon N.	1188
Rumpler R.	802	Vidal C.	1232
Sagaut P.	1007	Vierendeels J.	100, 934
Salomoni V.A.	530, 538, 548, 559	Villa T.	1062
Sapountzakis E.J.	743	Vladimirov I.N.	1332
Šarler B.	1220	Volkwein A.	731
Sbarufatti C.	1107, 1294	von Boetticher A.	731
Scapin M.	413, 578	Vulliet L.	172
Schanz T.	76	Wager C.	277
Schiavon M.	548	Wall W.A.	23
Schneider R.	355	Wang W.	160, 206, 218
Scholcz T.	779	Wang Z.	896
Schäfer M.	925	Weghs M.	873
Selgas V.	792	Weilharter B.	460
Šembera J.	1400	Wendeler C.	731
Sherburn J.A.	962	Wissink J.	954
Simões J. A.	1251	Wong H.	230
Singh A.K.	160, 206	Wüchner R.	731
Slawson T.R.	962	Xotta G.	538, 548
Sluys L.J.	67	Yamamoto M.	1045
Soldatos D.	506	Zabala F.	194
Song X.	1379	Žabka V.	1400
Soto O.	946	Zapoměl J.	424
Sotudeh-Gharebagh R.	1054	Zarghami R.	1054
Spielman T.	43	Zhou M.	1083
Stadler L.	90	Zilian A.	837
Stephany A.	254		
Sterenberg J.J.	635		

This volume contains the Full Papers accepted for presentation at the IV International Conference on Computational Methods for Coupled Problems in Science and Engineering, COUPLED PROBLEMS 2011, island of Kos, Greece on June 20 – 22, 2011.



CIMNE[®]
International Center
for Numerical Methods in Engineering