ENHANCED DAMAGE MODELLING OF QUASI-BRITTLE AND FATIGUE FRACTURE – COMPUTATIONAL ASPECTS

R.H.J. Peerlings^{*}, W.A.M. Brekelmans^{*}, R. de Borst[†], and M.G.D. Geers^{*‡}

*Faculty of Mechanical Engineering, Eindhoven University of Technology P.O. Box 513, 5600 MB Eindhoven, The Netherlands E-mail: ronp@wfw.wtb.tue.nl, marcel@wfw.wtb.tue.nl, geers@wfw.wtb.tue.nl

[†]Koiter Institute Delft / Faculty of Aerospace Engineering, Delft University of Technology P.O. Box 5058, 2600 GB Delft, The Netherlands E-mail: R.deBorst@lr.tudelft.nl

> [‡]Department of Civil Engineering, Royal Military Academy Avenue Renaissance 30, 1000 Brussels, Belgium E-mail: geers@cobo.rma.ac.be

Key words: Damage mechanics, fracture, fatigue, finite element method, localisation.

Abstract. Continuum damage mechanics can be used to model the initiation and growth of cracks. However, finite element analyses using standard fatigue damage formulations exhibit an extreme sensitivity to the spatial discretisation of the problem. Nonlocal and gradient damage formulations do not exhibit this mesh sensitivity. But the nonlocality or gradient terms in these models require some modifications of standard finite element algorithms for damage mechanics. In particular, care must be taken that the continuum representation of a crack is separated from the remaining material by applying the correct boundary conditions and remeshing the problem domain. For high-cycle fatigue analyses, special time integration schemes are needed to limit the computational effort involved. With these enhancements, reliable and mesh objective finite element analyses of crack initiation and growth become feasible, as is demonstrated by an application in metal fatigue.

1 Introduction

Fracture of engineering components is often preceded by considerable changes in the microstructure of the material they are made of. Accurate failure predictions can only be obtained if this microstructural damage is taken into account in the fracture modelling. This requirement has led to the development of so-called local approaches to fracture, in which fracture is regarded as the ultimate consequence of the material degradation process [1, 2]. In these methods, the degradation is often modelled using continuum damage mechanics [2–4]. Continuum damage theory introduces a set of field variables (damage variables) which explicitly describe the local loss of material integrity. A crack is represented by that part of the material domain in which the damage has become critical, i.e., where the material cannot sustain stress anymore. Redistribution of stresses results in the concentration of deformation and damage in this process zone which determines in which direction and at which rate the crack will propagate. Crack initiation and growth thus follow naturally from the standard continuum mechanics theory, instead of from separate fracture criteria.

In the early stages of their development, local approaches to fracture were considered particularly attractive from a computational standpoint. Damage formulations can be fitted into nonlinear finite element algorithms and implemented in simulation codes with relative ease and they do not seem to require the special discretisation and remeshing techniques used in numerical fracture mechanics. It has since been found, however, that finite element solutions of standard damage problems often do not seem to converge upon mesh refinement [5, 6]. As a matter of fact, they do converge to a solution, but this solution is physically meaningless as a consequence of the inability of the modelling to properly describe the physical phenomena that take place [5, 7, 8]. This can be understood if one realises that the concept of a continuous damage variable presumes a certain local homogeneity – or at least smoothness – of the microstructural damage distribution. But the development of damage localises in a surface while the surrounding material remains unaffected. This localisation of damage is in contradiction with the supposed smoothness of the damage field and thus affects the physical relevance of the model.

A range of extensions to the conventional damage and plasticity models have been proposed in order to regularise the localisation of deformation. Among them, the most viable is perhaps the class of nonlocal and gradient models. Both approaches introduce spatial interaction terms in the constitutive model, either using integral (nonlocal) relations [5,9,10] or gradients of some constitutive variable [11–15]. The additional terms have a smoothing effect on the deformation (and damage) fields, and thus preclude localisation in a surface. It is the aim of this contribution to investigate the impact of nonlocality and gradient terms on the computational treatment of damage mechanics for fracture problems. Elasticity based damage will serve as the framework of these developments because of its conceptual simplicity. After a brief introduction of the nonlocal and gradient damage models, their finite element implementation is discussed in detail. Special attention is given to the time integration of damage growth and the numerical treatment of the continuum damage representation of a crack. Finally, results are presented for the case of high-cycle metal fatigue.

2 Constitutive modelling

The basis for our developments is formed by an isotropic, quasi-brittle damage model, in which a scalar damage variable D degrades the elastic stiffness. The classical stressstrain relation for this type of models reads [16]:

$$\sigma_{ij} = (1 - D) C_{ijkl} \varepsilon_{kl}, \tag{1}$$

where Einstein's summation convention has been used and σ_{ij} (i, j = 1, 2, 3) denote the Cauchy stresses, C_{ijkl} the standard elastic constants and ε_{kl} (k, l = 1, 2, 3) the linear strains. The damage variable D satisfies $0 \leq D \leq 1$. A value of D = 0 represents the initial, undamaged material with the virgin stiffness; D = 1 represents a state of complete loss of stiffness. It is this critical state D = 1 which is used to model a crack as a zone of material that can no longer transfer stresses. Relation (1) does not account for permanent deformations. This means that applications of the model are limited to phenomena in which plastic deformations remain negligible. Two important fracture mechanisms which satisfy this condition are quasi-brittle fracture and high-cycle fatigue. Both phenomena will be treated here in the same elasticity-based damage framework.

The second law of thermodynamics requires that $D \ge 0$, i.e., that the damage variable can only increase. This growth of damage is related to the development of the deformation. A scalar equivalent strain measure $\tilde{\varepsilon}$ is introduced for this purpose, which quantifies the local deformation state in the material in terms of its effect on damage. In local damage models the damage growth can be related directly to the evolution of this equivalent strain $\tilde{\varepsilon}$. In the nonlocal and gradient damage formulations, however, a nonlocal equivalent strain $\bar{\varepsilon}$ enters the relationship between deformation and damage. The way in which this nonlocal equivalent strain is connected to its local counterpart differs between the enhanced models and will be discussed below for each of them.

Whether damage growth is possible is decided on the basis of a loading function in terms of $\bar{\varepsilon}$:

$$f(\bar{\varepsilon},\kappa) = \bar{\varepsilon} - \kappa. \tag{2}$$

The equation f = 0 defines a loading surface in strain space. For strain states within the loading surface (f < 0) there is no growth of damage and the material behaviour is elastic. The damage variable can only increase when the equivalent strain reaches the threshold value κ , i.e., when $f \ge 0$. Quasi-brittle damage can now be modelled by assuming that the strain state always remains on or within the loading surface ($f \leq 0$, cf. elastoplasticity). This means that the threshold variable κ must satisfy the Kuhn-Tucker relations

$$f\dot{\kappa} = 0, \qquad f \le 0, \qquad \dot{\kappa} \ge 0, \tag{3}$$

which must be supplemented by an initial value $\kappa = \kappa_0$ in order to define the initial elastic domain [16, 17]. If, on the other hand, the loading surface is kept fixed, i.e., if κ is kept fixed at $\kappa = \kappa_0$, the same modelling can be used to describe fatigue failure. Strain states beyond the loading surface (f > 0) are then allowed and the damage variable can only increase when f > 0 and $\dot{f} \ge 0$ [17–19].

When the appropriate conditions for damage growth are satisfied, the damage rate is governed both in quasi-brittle and fatigue damage by an evolution law which reads in its most general form

$$\dot{D} = g(D,\bar{\varepsilon})\,\dot{\bar{\varepsilon}},\tag{4}$$

where a superimposed dot denotes differentiation with respect to time. The dependence of the damage growth rate on the equivalent strain rate is linear in order to avoid rate effects. Obviously, the expressions for the evolution function $g(D, \tilde{\varepsilon})$ and the equivalent strain $\tilde{\varepsilon}$ will be different for quasi-brittle damage and fatigue; examples are given by [17].

2.1 Nonlocal model

In the nonlocal damage model the nonlocal equivalent strain $\bar{\varepsilon}$ in a point **x** is defined as a weighted average of the local equivalent strain $\tilde{\varepsilon}$ in the entire problem domain Ω [9,20]:

$$\bar{\varepsilon}(\mathbf{x}) = \frac{1}{\Psi(\mathbf{x})} \int_{\Omega} \psi(\mathbf{y}; \mathbf{x}) \tilde{\varepsilon}(\mathbf{y}) \,\mathrm{d}\Omega,\tag{5}$$

where \mathbf{y} denotes the position of the infinitesimal volume $d\Omega$. The factor $1/\Psi(\mathbf{x})$, with $\Psi(\mathbf{x})$ defined by

$$\Psi(\mathbf{x}) = \int_{\Omega} \psi(\mathbf{y}; \mathbf{x}) \,\mathrm{d}\Omega,\tag{6}$$

scales $\bar{\varepsilon}$ such that it equals $\tilde{\varepsilon}$ for homogeneous strain states. The weight function $\psi(\mathbf{y}; \mathbf{x})$ is assumed to be homogeneous and isotropic, i.e., it depends only on the distance $\rho = |\mathbf{x} - \mathbf{y}|$. It is usually defined as the Gauss distribution:

$$\psi(\rho) = \frac{1}{(2\pi)^{3/2} l^3} \exp\left[-\frac{\rho^2}{2l^2}\right].$$
(7)

The length parameter l determines the volume which contributes significantly to the nonlocal equivalent strain and must therefore be related to the scale of the microstructure.

2.2 Gradient formulation

For sufficiently smooth $\tilde{\varepsilon}$ -fields, the integral relation (5) can be rewritten in terms of gradients of $\tilde{\varepsilon}$ by expanding $\tilde{\varepsilon}(\mathbf{y})$ into a Taylor series. After some further mathematical manipulation and truncation of terms of order four and higher, the following gradient approximation of (5) can be obtained [21,22]:

$$\bar{\varepsilon} - c\nabla^2 \bar{\varepsilon} = \tilde{\varepsilon},\tag{8}$$

where the Laplacian ∇^2 is defined by $\nabla^2 = \sum_i \partial^2 / \partial x_i^2$ and the coefficient *c* is given by $c = \frac{1}{2}l^2$. For a unique solution a boundary condition for $\bar{\varepsilon}$ must be given at the entire boundary Γ of Ω . A homogeneous natural boundary condition

$$\frac{\partial \bar{\varepsilon}}{\partial n} = 0 \tag{9}$$

is usually adopted for this purpose. With this boundary condition, $\bar{\varepsilon}$ equals $\tilde{\varepsilon}$ for homogeneous deformations and the gradient approximation is thus consistent with the nonlocal relation (5) in this respect.

Instead of being defined explicitly in terms of its local counterpart $\tilde{\varepsilon}$, the nonlocal strain $\bar{\varepsilon}$ is now defined as the solution of the boundary value problem consisting of (8) and (9). However, an expression equivalent to (5) can be obtained for it by formally solving the boundary value problem [17,23]:

$$\bar{\varepsilon}(\mathbf{x}) = \int_{\Omega} G(\mathbf{y}; \mathbf{x}) \tilde{\varepsilon}(\mathbf{y}) \,\mathrm{d}\Omega,\tag{10}$$

where $G(\mathbf{y}; \mathbf{x})$ denotes the Green's function associated to the boundary value problem (8),(9). This means that the gradient damage model based on the differential equation (8) can be regarded as a special case of the class of (integral) nonlocal models, in which the weight function $\psi(\mathbf{y}; \mathbf{x})$ equals the Green's function $G(\mathbf{y}; \mathbf{x})$. Accordingly, the behaviour of the gradient enhanced damage formulation is at least qualitatively equivalent to that of the integral nonlocal model, see Peerlings et al. [23].

3 Finite element implementation

As a result of the partial equivalence of the integral nonlocal damage model defined by relation (5) and the gradient formulation based on (8), their numerical treatment is also largely parallel. A major difference, however, exists in the way the nonlocal equivalent strain is computed. In the nonlocal approach the $\bar{\varepsilon}$ -field is defined by the averaging operator (5) and a set of integro-differential equations is thus obtained. Reliable finite element algorithms have been developed to solve these equations [24,25], but their implementation in standard nonlinear finite element frameworks is not straightforward. The gradient formulation seems slightly more attractive from this point of view, since it only requires one additional partial differential equation to be solved simultaneously with the equilibrium equations. This section concentrates on the implementation of the gradient model, but many of the arguments made apply equally well to the nonlocal model. Where relevant, differences with the nonlocal approach are pointed out.

The partial differential equations of the equilibrium problem are first discretised in space by a finite element interpolation. The time discretisation of the problem follows by dividing the loading history in a finite number of time increments and integrating the growth of damage within these increments. The resulting set of nonlinear algebraic equations is solved in each increment by an iterative process. These three steps are first discussed below assuming that the damage is noncritical everywhere in the body, i.e., that no crack has been initiated yet. The extension to crack growth is discussed in Section 3.4.

3.1 Spatial Discretisation

The discrete form of the equilibrium equations follows from the standard transition to the weak form and a Galerkin discretisation of the displacements by $\mathbf{u} = \mathbf{N}\mathbf{a}$:

$$\int_{\Omega} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma} \, \mathrm{d}\Omega = \int_{\Gamma} \mathbf{N}^{\mathrm{T}} \mathbf{t} \, \mathrm{d}\Gamma, \tag{11}$$

where the matrices **N** and **B** contain the displacement interpolation functions and their derivatives, respectively, and the column matrices $\boldsymbol{\sigma}$ and **t** contain the Cauchy stresses and boundary tractions. Body forces have not been taken into account, but can be added in the standard fashion. Similar to the equilibrium equations, the weak form of equation (8) can be discretised as [22]

$$\int_{\Omega} \left(c \, \bar{\mathbf{B}}^{\mathrm{T}} \bar{\mathbf{B}} + \bar{\mathbf{N}}^{\mathrm{T}} \bar{\mathbf{N}} \right) \mathrm{d}\Omega \, \mathbf{e} = \int_{\Omega} \bar{\mathbf{N}}^{\mathrm{T}} \tilde{\varepsilon} \, \mathrm{d}\Omega, \tag{12}$$

with $\bar{\mathbf{N}}$ and $\bar{\mathbf{B}}$ containing the interpolation functions of the nonlocal equivalent strain $\bar{\varepsilon}$ and their derivatives and \mathbf{e} being the column with nodal values of $\bar{\varepsilon}$.

The interpolations functions in $\mathbf{N}, \mathbf{\bar{N}}$ have to satisfy only the standard, \mathcal{C}_0 -continuity requirements and need not be of the same order. Indeed, it has been found that using interpolation polynomials of the same order may result in stress oscillations. Their origin is illustrated in Figure 1 for a one-dimensional problem with piecewise linear interpolations of the axial displacement u and the nonlocal equivalent strain $\bar{\varepsilon}$. The linearity of the displacements within each element results in the local strain ε being piecewise constant (Figure 1(a)). Since the damage rate depends nonlinearly but continuously on the nonlocal equivalent strain, which is piecewise linear, D is continuous and nearly linear within the elements for sufficiently small elements (Figure 1(b)). The combination of the constant local strain and varying damage in the stress-strain relation (1) leads to a stress distribution as shown in Figure 1(c). The stress gradient within the elements is set by the



Figure 1: Stress oscillations caused by a linear-linear finite element interpolation.

damage gradient and the strain; oscillations are therefore observed particularly in regions with high gradients of the damage variable and high strains. Refining the discretisation does not lead to a decrease of the stress gradients. Indeed, it may even result in an increase because high damage and displacement gradients can be described more accurately with a finer mesh. Similar stress oscillations have also been observed in the integral nonlocal model by Jirásek [25].

Although the stress field of Figure 1(c) is a perfectly valid solution of the weak equilibrium problem, the stress oscillations may lead to serious misinterpretations and should therefore be avoided. An obvious way to do so is by using quadratic instead of linear polynomials for the displacements, so that the strain field is no longer constant and agrees with the linear nonlocal equivalent strain. Particularly when combined with a reduced Gauss integration for the equilibrium equations (11) and a full Gauss integration for equation (12), this discretisation has been found to be effective and efficient. A second possibility is to retain the linear interpolations of the displacements and nonlocal equivalent strain, but to define the damage variable as uniform in each element. The strain and damage are then both piecewise constant, so that the stress is also piecewise constant. This approach is preferred in crack growth problems, see Section 3.4.

3.2 Temporal discretisation

The time discretisation of the equilibrium problem follows by dividing the loading history into a finite number of time intervals and requiring the discrete balances (11) and (12) to be satisfied only at the end of each interval. Assuming the complete deformation and damage state to be known at time t, the problem then reduces to finding the displacements and nonlocal strains which satisfy (11) and (12) at the end of a time increment Δt . For this purpose the damage growth rate must be integrated from t to $t + \Delta t$. In the quasi-brittle damage model, this integration can be carried out in closed form, resulting in a direct relation between the damage variable $D^{t+\Delta t}$ and \mathbf{e} [17, 22].

In the fatigue model the damage rate cannot be integrated analytically and an approximation must therefore be used. The damage variable at $t + \Delta t$ can formally be written as

$$D(t + \Delta t) = D(t) + \int_{t}^{t + \Delta t} \dot{D}(\tau) \,\mathrm{d}\tau.$$
(13)

The standard procedure would now be to approximate the right-hand side of this expression by an integration rule, e.g., by the trapezoidal rule

$$\int_{t}^{t+\Delta t} \dot{D}(\tau) \,\mathrm{d}\tau \approx \frac{1}{2} \left(\dot{D}^{t} + \dot{D}^{t+\Delta t} \right) \Delta t.$$
(14)

This discrete approximation is accurate when the damage rate varies slowly within the time increment. Under cyclic loading conditions this implies that each loading cycle must be interpolated with a number, say O(10), of increments. The total number of increments needed to simulate the entire fatigue fracture process would then be of the order of ten times the fatigue life. For high-cycle fatigue simulations, involving fatigue lives of 10^5 cycles and more, this would clearly become impractical. For such analyses there is a need for an approximate integration which allows to span a large number of cycles within each time increment. The loss of resolution which is inevitable in such a procedure is acceptable since it is usually not the precise, small growth of damage within each loading cycle which



Figure 2: Nonlocal equivalent strain cycles.

is relevant, but rather the more substantial changes in the damage field resulting from larger numbers of cycles.

The evolution of $\bar{\varepsilon}$ during an interval $(t, t + \Delta t)$ which comprises a number of loading cycles is shown schematically in Figure 2. It has been assumed in this figure that the loading is proportional and fully reversed. Each loading cycle results in two maxima of $\bar{\varepsilon}$, one corresponding to the tensile part of the cycle and one to the compressive part. Both maxima, which are denoted by $\bar{\varepsilon}_{a1}$ and $\bar{\varepsilon}_{a2}$, follow the envelope $\bar{\varepsilon}_{a}$. The integration of the damage rate according to (13) still holds for the situation of Figure 2. However, the integral in the right-hand side of (13) can no longer be directly approximated by (14) because $\bar{\varepsilon}$ and $\dot{\bar{\varepsilon}}$, and thus also \dot{D} , fluctuate strongly within the increment. Instead of directly using an integration rule, the integral is formally written as a sum of integrals on the cycles within the increment Δt :

$$D^{t+\Delta t} = D^t + \sum_{n=N}^{N+\Delta N} \int_{t_n}^{t_n+\Delta t_n} \dot{D} \,\mathrm{d}\tau, \qquad (15)$$

where t_n and $t_n + \Delta t_n$ correspond to the beginning and end of cycle n, respectively, and N and $N + \Delta N$ are the number of cycles at time t and $t + \Delta t$. Taking into account that damage growth is possible only if $\bar{\varepsilon} \geq \kappa$ and $\dot{f} \equiv \dot{\varepsilon} \geq 0$ and using evolution law (4), relation (15) can be rewritten as

$$D^{N+\Delta N} = D^N + \sum_{n=N}^{N+\Delta N} \left(\int_{\kappa}^{\bar{\varepsilon}_{a1}} g(D,\bar{\varepsilon}) \,\mathrm{d}\bar{\varepsilon} + \int_{\kappa}^{\bar{\varepsilon}_{a2}} g(D,\bar{\varepsilon}) \,\mathrm{d}\bar{\varepsilon} \right), \tag{16}$$

where it has been assumed that $\bar{\varepsilon}_{a1}, \bar{\varepsilon}_{a2} \geq \kappa$. Obviously, cycles in which $\bar{\varepsilon}_a < \kappa$ (and thus $\bar{\varepsilon}_{a1} < \kappa$ and $\bar{\varepsilon}_{a2} < \kappa$) do not contribute to the growth of damage. In contrast to $\bar{\varepsilon}$ and

 \dot{D} , the cycle amplitudes $\bar{\varepsilon}_{a1}$ and $\bar{\varepsilon}_{a2}$ as well as the damage variable D can be assumed to vary slowly within the increment. Thus, similar to the integration rule (14), the sum over the cycle numbers in (16) can be approximated by the average of the integrals evaluated at the beginning and end of the increment multiplied by the number of cycles in the increment, ΔN :

$$D^{N+\Delta N} = D^N + \frac{1}{2} \left(G(D^N, \bar{\varepsilon}^N_{\mathbf{a}}) + G(D^{N+\Delta N}, \bar{\varepsilon}^{N+\Delta N}_{\mathbf{a}}) \right) \Delta N, \tag{17}$$

with $G(D, \bar{\varepsilon}_{\mathbf{a}})$ defined by

$$G(D,\bar{\varepsilon}_{\mathbf{a}}) = 2 \int_{\kappa}^{\bar{\varepsilon}_{\mathbf{a}}} g(D,\bar{\varepsilon}) \,\mathrm{d}\bar{\varepsilon},\tag{18}$$

while $G(D, \bar{\varepsilon}_{a}) = 0$ if $\bar{\varepsilon}_{a} < \kappa$ or if D = 1.

Relation (17) is a nonlinear equation in terms of the damage variable $D^{N+\Delta N}$. Solving this equation iteratively can be avoided by using Heun's method, i.e., by replacing $D^{N+\Delta N}$ in the right-hand side of (17) by a predictor value D^{p} based on a forward Euler step:

$$D^{\mathbf{p}} = D^{N} + G(D^{N}, \bar{\varepsilon}^{N}_{\mathbf{a}}) \Delta N, \qquad (19)$$

$$D^{N+\Delta N} = D^N + \frac{1}{2} \left(G(D^N, \bar{\varepsilon}^N_{\mathbf{a}}) + G(D^{\mathbf{p}}, \bar{\varepsilon}^{N+\Delta N}_{\mathbf{a}}) \right) \Delta N.$$
(20)

Since the incremental damage growth is given by relations (19) and (20) in terms of the nonlocal equivalent strain envelope $\bar{\varepsilon}_{a}$, the equilibrium problem must be solved for this envelope. This means that the unknowns of the problem must be interpreted as the momentary amplitudes rather than actual values.

The accuracy of the numerically integrated damage growth increases when a smaller step size ΔN is used in the time integration. The step size needed to meet the desired overall accuracy is set by the strongly progressive damage growth which usually occurs near the end of the fatigue life. For reasons of efficiency, however, larger step sizes may be used where the damage variable varies relatively slowly (usually in the early stages of the damage process) without compromising the overall accuracy of the analysis. Accuracy and efficiency can be balanced by adapting the step size to the momentary development of damage. This can be done in an objective way by estimating the error which will be made in computing the damage growth during the time increment. Practical fatigue damage growth relations are strongly progressive. This implies that the solution fields associated to them diverge and thus that solutions are intrinsically unstable in the sense that small errors are amplified and may thus become relatively large as the damage growth progresses. Note that in numerical terms this is a matter of accuracy rather than stability. The inherited error due to the amplified propagation of errors is believed to be more critical than the truncation error made by using (17) and is therefore used as a criterion for the selection of the increment size.

In order to assess the inherited error at the end of an increment ΔN , it is assumed that the computed value of the damage variable at N, denoted as \tilde{D}^N , contains a small error δ : $\tilde{D}^N = D^N + \delta$. Substitution of this value in (17) gives the computed damage at the end of the increment as

$$\tilde{D}^{N+\Delta N} = \tilde{D}^N + \frac{1}{2} \left(G(\tilde{D}^N, \bar{\varepsilon}^N_{\rm a}) + G(\tilde{D}^{N+\Delta N}, \bar{\varepsilon}^{N+\Delta N}_{\rm a}) \right) \Delta N.$$
(21)

Linearising this relation with respect to δ and ΔN results in a first-order approximation of the inherited error at $N + \Delta N$:

$$\tilde{D}^{N+\Delta N} - D^{N+\Delta N} \approx \left(1 + \frac{\partial G}{\partial D} \Delta N\right) \delta.$$
(22)

The increment size is now selected such that the second term in the amplification factor equals a predefined constant η :

$$\Delta N = \frac{\eta}{\partial G/\partial D}.$$
(23)

In principle, the derivative $\partial G/\partial D$ in (23) can be evaluated at any time within the interval $(N, N + \Delta N)$. In practice, however, it is evaluated at time N, so that the step size can be fixed at the beginning of the increment. Furthermore, lower and upper bounds are imposed upon the value given by (23) in order to prevent excessively small or large cycle increments.

3.3 Iterative scheme

The spatial and time discretisation reduce the equilibrium problem to a set of coupled, nonlinear algebraic equations, which can be written as

$$\mathbf{f}_{ai} = \mathbf{f}_{ae},\tag{24}$$

$$\mathbf{K}_{\rm ee}\mathbf{e} - \mathbf{f}_{\rm e} = \mathbf{0},\tag{25}$$

where

$$\mathbf{f}_{ai} = \int_{\Omega} \mathbf{B}^{\mathrm{T}} \boldsymbol{\sigma} \, \mathrm{d}\Omega, \qquad \qquad \mathbf{f}_{ae} = \int_{\Gamma} \mathbf{N}^{\mathrm{T}} \mathbf{t} \, \mathrm{d}\Gamma, \qquad (26)$$

$$\mathbf{K}_{\rm ee} = \int_{\Omega} \left(c \, \bar{\mathbf{B}}^{\rm T} \bar{\mathbf{B}} + \bar{\mathbf{N}}^{\rm T} \bar{\mathbf{N}} \right) \, \mathrm{d}\Omega, \qquad \mathbf{f}_{\rm e} = \int_{\Omega} \bar{\mathbf{N}}^{\rm T} \tilde{\varepsilon} \, \mathrm{d}\Omega.$$
(27)

A full Newton-Raphson scheme is used to solve the set of equations (24), (25) iteratively at the increment level. For this purpose, relations (24) and (25) are rewritten for iteration i + 1 as

$$\delta \mathbf{f}_{\mathrm{ai}} = \mathbf{f}_{\mathrm{ae}} - \mathbf{f}_{\mathrm{ai}}^{i},\tag{28}$$

$$\mathbf{K}_{\rm ee}\delta\mathbf{e} - \delta\mathbf{f}_{\rm e} = \mathbf{f}_{\rm e}^{i} - \mathbf{K}_{\rm ee}\mathbf{e}^{i},\tag{29}$$

where \mathbf{e} , \mathbf{f}_{ai} and \mathbf{f}_{e} have been written as the sum of their value in the previous iteration \mathbf{e}^{i} , \mathbf{f}_{ai}^{i} , \mathbf{f}_{e}^{i} and iterative corrections $\delta \mathbf{e}$, $\delta \mathbf{f}_{ai}$, $\delta \mathbf{f}_{e}$. The variation of \mathbf{f}_{e} can be linearised using

$$\delta \tilde{\varepsilon} = \left(\frac{\partial \tilde{\varepsilon}}{\partial \boldsymbol{\varepsilon}}\right)^{\mathrm{T}} \mathbf{B} \, \delta \mathbf{a},\tag{30}$$

where the derivative of the equivalent strain must be evaluated for $\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}^{i}$. Similarly, changes of the internal nodal forces, $\delta \mathbf{f}_{ai}$, are linearised by

$$\delta \boldsymbol{\sigma} = (1 - D^i) \mathbf{C} \mathbf{B} \,\delta \mathbf{a} - \delta D \,\mathbf{C} \,\mathbf{B} \,\mathbf{a}^i,\tag{31}$$

where the matrix \mathbf{C} contains the elastic constants. If the conditions for damage growth are satisfied, the iterative change of the damage variable follows for the fatigue model from linearising relation (20) as

$$\delta D = q \,\bar{\mathbf{N}} \,\delta \mathbf{e},\tag{32}$$

with q defined by

$$q = \frac{1}{2} \frac{\partial G}{\partial \bar{\varepsilon}} \Delta N, \tag{33}$$

where the index a of $\bar{\varepsilon}_{a}$ has been dropped and $\partial G/\partial \bar{\varepsilon}$ must be evaluated for $D = D^{p}$. For the quasi-brittle model relation (32) remains valid, albeit with a different definition of q [17,22]. In both cases, the second term in (31) vanishes when the conditions for damage growth are not met.

Using (30), (31) and (32) in (28), (29) results in the set of linear equations

$$\begin{bmatrix} \mathbf{K}_{\mathrm{aa}} & \mathbf{K}_{\mathrm{ae}} \\ \mathbf{K}_{\mathrm{ea}} & \mathbf{K}_{\mathrm{ee}} \end{bmatrix} \begin{bmatrix} \delta \mathbf{a} \\ \delta \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_{\mathrm{ae}} - \mathbf{f}_{\mathrm{ai}}^{i} \\ \mathbf{f}_{\mathrm{e}}^{i} - \mathbf{K}_{\mathrm{ee}} \mathbf{e}^{i} \end{bmatrix},$$
(34)

with

$$\mathbf{K}_{\mathrm{aa}} = \int_{\Omega} \mathbf{B}^{\mathrm{T}} (1 - D^{i}) \mathbf{C} \, \mathbf{B} \, \mathrm{d}\Omega, \qquad \mathbf{K}_{\mathrm{ae}} = -\int_{\Omega} \mathbf{B}^{\mathrm{T}} \mathbf{C} \, \boldsymbol{\varepsilon}^{i} q \, \bar{\mathbf{N}} \, \mathrm{d}\Omega, \tag{35}$$

$$\mathbf{K}_{\rm ea} = -\int_{\Omega} \bar{\mathbf{N}}^{\rm T} \left(\frac{\partial \tilde{\varepsilon}}{\partial \varepsilon} \right)^{\rm T} \mathbf{B} \, \mathrm{d}\Omega. \tag{36}$$

As usual, this set of equations is to be solved repeatedly, until the residue on the nonlinear equations (24),(25) has become sufficiently small. Note that the tangential stiffness matrix in (34) is fully consistent. It is nonsymmetric and has a slightly higher optimum bandwidth than that for the local model because additional degrees of freedom have been introduced. For the nonlocal model, these extra degrees of freedom do not appear, but the optimum bandwidth of the consistent stiffness matrix is nevertheless higher as a result of the nonlocality.

3.4 Crack growth

A crack is represented in the damage model by a region of completely damaged material. It is important to realise that the complete loss of stiffness in this region implies that stresses are identically zero for arbitrary deformation fields. As a consequence, the equilibrium equations are meaningless in the cracked region. This can be seen by substituting the stress-strain relations (1) into the standard equilibrium equations

$$\frac{\partial \sigma_{ij}}{\partial x_i} = 0. \tag{37}$$

Using the definition of the linear strains and the right minor symmetry of the elasticity tensor (i.e., $C_{ijkl} = C_{ijlk}$) the resulting equations can be rewritten as

$$(1-D) C_{ijkl} \frac{\partial^2 u_k}{\partial x_i \partial x_l} - \frac{\partial D}{\partial x_i} C_{ijkl} \frac{\partial u_k}{\partial x_l} = 0.$$
(38)

For a given noncritical damage field $D(\mathbf{x}) < 1$, the displacement components u_k can be determined from this differential system and the corresponding kinematic and dynamic boundary conditions. In a crack however, where $D \equiv 1$, both terms in the differential equations vanish. Consequently, the differential system degenerates and the boundary value problem becomes ill-posed. The crack region must therefore be excluded from the domain of the equilibrium problem by introducing an internal boundary, on which the condition of zero tractions and, in the gradient model, condition (9) are imposed as boundary conditions. In the nonlocal model the integral in (5) must be limited to the noncritical domain and the scaling factor Ψ must be recomputed accordingly. A free boundary problem is thus obtained, in which the position of the internal boundary (the crack front and crack faces) follows from the growth of damage.

In numerical analyses the fact that the domain of the equilibrium problem changes with each increment of crack growth means that the problem discretisation must be adapted. In order to avoid frequent remeshing, numerical damage analyses are often defined on the original domain even if this domain contains a crack. The material in the crack is given a small residual stiffness, for instance by limiting the damage variable to a value which is slightly smaller than one, in order to avoid that the discrete equilibrium equations become singular. It is then argued that the stresses which are still transferred by the crack influence equilibrium only marginally if the residual stiffness is sufficiently small. This may indeed be true in local damage models, in which the large strains in the crack do not influence the surrounding material. But if this approach is followed for the nonlocal and gradient damage models, the nonlocal equivalent strain maps the (nonphysical) strains in the cracked region onto the surrounding material in which the damage variable is not (yet) critical. This does not only result in faster growth of damage in front of the crack and consequently in higher predicted crack growth rates, but also in damage growth at the faces of the crack, thus causing the thickness of the cracked region to increase in an unrealistic way. This phenomenon has been shown for instance by Geers et al. [26], who proposed to remove it by a transient gradient formulation, i.e., a gradient model with a decaying internal length scale.

These difficulties are avoided here by using a rather crude remeshing method: completely damaged elements are removed from an otherwise fixed finite element mesh. The damage variable is taken constant in each element in order to avoid partially cracked elements, since these have been found to have a negative influence on the mesh objectivity [17]. When the damage variable is critical at the end of an increment in a certain element, this element is removed from the finite element mesh. Nodes and degrees of freedom which are not connected to other elements are also removed and the set of discrete equations is resized accordingly. The increment which led to the critical damage is then recomputed starting from the converged state in the previous increment, in which the element was not yet cracked, so that the growth of damage in other elements is consistent with the current configuration.

4 Application: metal fatigue

The numerical implementation of the gradient damage model has been applied to the problem geometry of Figure 3. The thickness of the specimen is 0.5 mm. A blunt notch has been used in order to have a finite number of cycles to crack initiation, which allows to study the initiation and propagation of a crack in the same problem. The lower edge of the specimen is fixed in all directions, while fully reversed vertical displacement cycles with an amplitude of 0.0048 mm are applied to its top edge. The material data that have been used in the analyses are those of 1015 steel and have been taken from Suresh [27]. See Peerlings et al. [19] for details of the damage model.

Because of symmetry, only the top half of the specimen has been modelled. The ref-



Figure 3: Problem geometry and loading conditions of the fatigue problem (dimensions in mm).



Figure 4: Reference discretisation: (a) entire specimen and (b) refinement at the notch tip (h = 0.04 mm).

erence mesh contains a regular grid of elements with an edge length h = 0.04 mm in an area of approximately $0.65 \times 0.12 \text{ mm}^2$ at the notch tip, see Figure 4. The discretisation has been successively refined in this area to h = 0.02, 0.01 and 0.005 mm. Bilinear interpolations have been used for the displacements and the nonlocal strain and the damage variable has been assumed piecewise constant. The integration in time has been carried out with the explicit cycle-based scheme (19), (20) and the adaptive step size selection algorithm, with $\eta = 0.5$ and minimum and maximum increment sizes of 1 and 10^5 cycles.

Figure 5 shows the crack initiation and growth process as simulated with the finest of the four meshes. The stress concentration at the notch tip leads to a concentration of damage in this area. At a certain stage a crack is initiated, i.e., the damage variable becomes critical in an element which is then removed from the mesh. For continued cycling the crack grows along the symmetry axis. The crack width decreases as the damage zone which was formed before crack initiation is traversed. Beyond this damage zone the crack width becomes stationary at 0.04 mm, which is of the order of the internal length $\sqrt{c} = 0.1$ mm.

The influence of the finite element discretisation on the crack shape is shown in Figure 6, in which the final crack pattern has been plotted for the four discretisations. The coarsest mesh (Figure 6(a)) gives a rather crude approximation of the crack shape and overestimates the width of steady-state part of the crack because this width is smaller than the element size. However, the h = 0.02 and 0.01 mm meshes give a good approximation of the crack shape in the finest discretisation. The width of the steady-state part of the crack does not vary between the three finest discretisations.



Figure 5: Damage and crack growth at the notch tip in the $h = 0.005 \,\mathrm{mm}$ mesh.

Figure 7 shows the length of the crack, a, versus the number of loading cycles N for the four meshes. For an increasingly refined discretisation the growth curves converge to a response with a finite number of cycles to crack initiation and a finite growth rate. The converged initiation life is approximately 4210 cycles. Immediately after its initiation, the crack starts to grow at a relatively high rate. The rate of crack growth decreases as the damage zone which was formed before crack initiation is traversed until it becomes almost constant beyond this zone. This transition corresponds to the width of the crack becoming constant (Figure 5).



Figure 6: Final crack pattern in the (a) h = 0.04 mm, (b) h = 0.02 mm, (c) h = 0.01 mm and (d) h = 0.005 mm meshes.

5 Concluding remarks

The example of the previous section shows that finite element analyses of crack growth using the gradient-enhanced fatigue damage model do not suffer from the mesh sensitivity exhibited by the local damage model, i.e., the model without nonlocal or gradient terms. The crack grows at a finite rate and a positive amount of work is needed for it, instead of the instantaneous, perfectly brittle behaviour of the local model [17,19].



Figure 7: Influence of the element size on the predicted crack growth.

A key element in obtaining reliable and mesh-objective results is the treatment of the completely damaged region. In this region, which is the continuum damage representation of a crack, the equilibrium equations are no longer meaningful. It must therefore be excluded from the equilibrium boundary value problem. Numerical implementations of the enhanced damage models must reflect this separation of the cracked region from the remaining part of the continuum. This means that the spatial discretisation of the equilibrium problem must be adapted for each increment of crack growth. If this separation is not made rigorously, the damage growth rate may be overestimated and nonphysical damage growth may be predicted at the faces of the crack.

When modelling high-cycle fatigue, the large number of loading cycles renders standard integration procedures unpractical. A special time integration scheme has been developed to avoid this difficulty by following the deformation envelope rather than its detailed evolution during individual cycles. Combined with an adaptive step size algorithm, this scheme allows to simulate large numbers of cycles with an acceptable computational effort.

In the application of Section 4 the crack path is known in advance. In practical situations, however, where the location of crack initiation and the direction of crack growth are not known in advance, adaptive spatial discretisation techniques are needed to follow the free boundary which represents the crack and to accurately describe the high deformation gradients at its tip. It is interesting to note that these requirements are very similar to those of fracture mechanics and some of the techniques needed may therefore be borrowed from numerical fracture mechanics. It is believed that by adding these elements to the numerical implementation of Section 3, reliable and efficient implementations of continuum approaches to fracture become feasible in the near future.

REFERENCES

- [1] J. Lemaitre, Local approach to fracture, Eng. Fract. Mech., 25, 523–537 (1986).
- [2] J.L. Chaboche, Continuum damage mechanics. Part I General concepts; Part II Damage growth, crack initiation, and crack growth, J. Appl. Mech., 55, 59–64, 65–72 (1988).
- [3] L.M. Kachanov, On the time to failure under creep conditions, *Izv. Akad. Nauk. SSSR, Otd. Tekhn. Nauk.*, 8, 26–31 (1958). In Russian.
- [4] Y.N. Rabotnov, Creep Problems in Structural Members, North-Holland, Amsterdam (1969).
- [5] Z.P. Bažant, T. Belytschko and T.P. Chang, Continuum theory for strain-softening, J. Eng. Mech., 110, 1666–1692 (1984).
- [6] H.L. Schreyer and Z. Chen, One-dimensional softening with localization, J. Appl. Mech., 53, 791–797 (1986).
- [7] N. Triantafyllidis and E.C. Aifantis, A gradient approach to localization of deformation. I Hyperelastic materials, J. Elasticity, 16, 225–237 (1986).
- [8] R. de Borst, L.J. Sluys, H.-B. Mühlhaus and J. Pamin, Fundamental issues in finite element analysis of localization of deformation, *Eng. Comp.*, 10, 99–121 (1993).
- [9] G. Pijaudier-Cabot and Z.P. Bažant, Nonlocal damage theory, J. Eng. Mech., 113, 1512–1533 (1987).
- [10] V. Tvergaard and A. Needleman, Effects of non-local damage in porous plastic solids, Int. J. Solids Struct., 32, 1063–1077 (1995).
- [11] E.C. Aifantis, On the microstructural origin of certain inelastic models, J. Eng. Mat. Technol., 106, 326–330 (1984).
- [12] B.D. Coleman and M.L. Hodgdon, On shear bands in ductile materials, Arch. Rat. Mech. Anal., 90, 219–247 (1985).
- [13] D. Lasry and T. Belytschko, Localization limiters in transient problems, Int. J. Solids Struct., 24, 581–597 (1988).
- [14] H.-B. Mühlhaus and E.C. Aifantis, A variational principle for gradient plasticity, Int. J. Solids Struct., 28, 845–857 (1991).
- [15] R. de Borst and H.-B. Mühlhaus, Gradient-dependent plasticity: formulation and algorithmic aspects, Int. J. Num. Meth. Eng., 35, 521–539 (1992).

- [16] J. Lemaitre and J.-L. Chaboche, *Mechanics of Solid Materials*, Cambridge University Press, Cambridge (1990).
- [17] R.H.J. Peerlings, Enhanced damage modelling for fracture and fatigue, Ph.D. thesis, Eindhoven University of Technology, Eindhoven, The Netherlands (1999).
- [18] M.H.J.W. Paas, P.J.G. Schreurs and W.A.M. Brekelmans, A continuum approach to brittle and fatigue damage: theory and numerical procedures, *Int. J. Solids Struct.*, **30**, 579–599 (1993).
- [19] R.H.J. Peerlings, W.A.M. Brekelmans, R. de Borst and M.G.D. Geers, Gradientenhanced damage modelling of fatigue, *Int. J. Num. Meth. Eng.* (2000). Submitted.
- [20] Z.P. Bažant and G. Pijaudier-Cabot, Nonlocal continuum damage, localization instability and convergence, J. Appl. Mech., 55, 287–293 (1988).
- [21] R.H.J. Peerlings, R. de Borst, W.A.M. Brekelmans and J.H.P. de Vree, Computational modelling of gradient-enhanced damage for fracture and fatigue problems, D.R.J. Owen and E. Oñate, (eds.) *Computational Plasticity, Fundamentals and Applications. Proc. 4th Int. Conf.*, 975–986, Pineridge Press, Swansea, U.K. (1995).
- [22] R.H.J. Peerlings, R. de Borst, W.A.M. Brekelmans and J.H.P. de Vree, Gradientenhanced damage for quasi-brittle materials, *Int. J. Num. Meth. Eng.*, **39**, 3391–3403 (1996).
- [23] R.H.J. Peerlings, M.G.D. Geers, R. de Borst and W.A.M. Brekelmans, A critical comparison of nonlocal and gradient-enhanced softening continua, *Int. J. Solids Struct.* (2000). Submitted.
- [24] A. Huerta and G. Pijaudier-Cabot, Discretization influence on the regularization by two localization limiters, J. Eng. Mech., 120, 1198–1218 (1994).
- [25] M. Jirásek, Computational aspects of nonlocal models, Proc. European Conference on Computational Mechanics, TU München, München, Germany (1999).
- [26] M.G.D. Geers, R. de Borst, W.A.M. Brekelmans and R.H.J. Peerlings, Strain-based transient-gradient damage model for failure analyses, *Comp. Meth. Appl. Mech. Eng.*, 160, 133–153 (1998).
- [27] S. Suresh, Fatigue of Materials, Cambridge University Press, Cambridge, U.K. (1991).