Gradient-enhanced damage modelling of high-cycle fatigue

R. H. J. Peerlings1,*,†, W. A. M. Brekelmans1, R. de Borst2 and M. G. D. Geers1,3

1 Faculty of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
2 Koiter Institute Delft/Faculty of Aerospace Engineering, Delft University of Technology, P.O. Box 5058, 2600 GB Delft, The Netherlands
3 Department of Civil Engineering, Royal Military Academy, Avenue Renaissance 30, 1000 Brussels, Belgium

SUMMARY
Continuum damage mechanics can be used to model the initiation and growth of fatigue cracks. However, finite element analyses using standard fatigue damage formulations exhibit an extreme sensitivity to the spatial discretisation of the problem. The mesh sensitivity is caused by the fact that the underlying continuum model predicts instantaneous, perfectly brittle crack growth as soon as a crack has been initiated. The growth of damage localizes in a vanishing volume during this instantaneous growth. This localization is not so much due to loss of ellipticity of the problem, but is caused by the fact that the damage rate is singular at the crack tip. The damage rate singularity can be removed by the introduction of higher-order deformation gradients in the constitutive modelling. As a result, crack growth at a finite rate and with a positive amount of energy dissipation is predicted. Finite element analyses converge to this solution and are thus no longer pathologically dependent on the spatial discretization. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: fatigue; damage; fracture; localization; finite element method; mesh sensitivity

1. INTRODUCTION
Fatigue fracture processes can be divided into three stages: crack initiation, stable crack growth and final fracture. In the first stage repeated straining of the component results in an accumulation of microstructural damage, culminating in the formation of a macroscopic crack. In the stable growth stage, this crack propagates by a small amount per loading cycle, until the remaining cross section becomes too small to bear the load which is to be transferred by the component. The fracture process is then completed by dynamic growth of the crack.

Methods of design against fatigue traditionally focus on one of the first two stages of the fatigue damage process. Safe-life methods define failure as the initiation of a crack and try to predict the number of loading cycles that the component can resist before initiation occurs. Damage-tolerant approaches, on the other hand, assume that cracks and other defects will always exist and use

*Correspondence to: R. H. J. Peerlings, Faculty of Mechanical Engineering, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
†E-mail: ronp@wfw.wtb.tue.nl

Copyright © 2000 John Wiley & Sons, Ltd.

Received 3 December 1999
Revised 29 February 2000
analyses to predict the number of cycles resulting in catastrophic fracture. Existing computational approaches reflect this separation between initiation and propagation methods. In many cases, however, initiation and propagation analyses are overly conservative in the sense that a considerable part of the fatigue life is neglected.

Continuum damage mechanics can be used to describe fatigue crack initiation and propagation in a unified fashion [1, 2]. It introduces a macroscopic field variable which represents microstructural material damage in an average sense. This so-called damage variable appears in the constitutive relations which govern the deformation behaviour of the material, thus taking into account the effect of material degradation on the mechanical properties of the material. In the initiation stage, the damage variable gradually grows under the influence of mechanical loading. At a certain level of accumulated damage the material locally cannot transfer stresses anymore and a crack is initiated. The continued growth of damage then localizes near the crack (that is, the zone of completely damaged material) and the crack can propagate by a process of stress redistribution and damage growth. This type of fracture modelling is sometimes referred to as a local approach to fracture [3, 4].

Since damage models are formulated in terms of field variables, they can usually be cast into standard non-linear finite element formulations and implemented in finite element codes without much difficulty. However, finite element analyses using standard damage models often exhibit an extreme mesh sensitivity, i.e., numerical solutions 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. Accordingly, the mesh sensitivity of the analyses is not caused by the numerical methods which are used, but by the fact that the underlying continuum model does not properly describe the physical phenomena that takes place [6, 7].

The failure of standard continuum damage models in describing fracture processes can be understood if one realizes that the concept of a continuous damage variable presumes a certain local homogeneity—or at least smoothness—of the microstructural damage distribution. But the continuum models based on this concept allow for discontinuous solutions, in which the development of damage localizes in a surface, while the surrounding material remains unaffected. This localization of damage is in contradiction with the supposed smoothness of the damage field and thus affects the physical relevance of the damage modelling. A number of strategies have been suggested in the literature to prevent the so-called pathological localization of damage and deformation [7]. Among them, the most promising is perhaps the class of non-local and gradient models. Both approaches introduce spatial interaction terms in the constitutive model, either using integral (non-local) relations [6, 8] or gradients of some constitutive variable [9–11]. The additional terms have a smoothing effect on the deformation (and damage) fields, and thus preclude localization in a surface.

In this contribution the origins of damage localization in crack initiation and crack growth are examined in depth for an elasticity-based damage model of high-cycle fatigue. Classical localization criteria, which relate the localization of deformation to loss of ellipticity of the rate equilibrium problem, seem to be of limited use in this situation. Instead, the concentration of damage and the resulting non-physical behaviour are caused by the singularity of the damage rate at the crack tip. Based on this understanding, a gradient-enhanced damage formulation has been developed. In this enhanced model, damage rate singularities are avoided as a result of the non-locality introduced by the gradient terms. The gradient-enhanced model has been implemented in a finite element code and has been used to simulate fatigue crack initiation and propagation in a steel specimen.
2. LOCAL DAMAGE MODELLING

The restriction to high-cycle fatigue which is made here means that permanent deformations can be neglected and strains can be assumed to be small. Furthermore, the material behaviour is assumed to be linearly elastic in the absence of damage growth. The remaining class of models is often referred to as elasticity-based damage or damage coupled with elasticity. See for instance Lemaitre and Chaboche [1] for fatigue damage models which do allow for permanent deformations. As a further simplification it is assumed here that a single, scalar damage variable suffices to describe the local damage state and that the influence of this damage variable on the constitutive behaviour is isotropic.

The damage variable $D$ is defined such that $0 \leq D \leq 1$, where $D = 0$ represents the initial, undamaged material and $D = 1$ represents a state of complete loss of integrity. The effect of this damage variable on the stresses is given by the classical stress–strain relation of elasticity-based damage mechanics [1]

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

Einstein’s summation convention has been used in this relation; $\sigma_{ij}$ ($i, j = 1, 2, 3$) denote the Cauchy stress components and $\varepsilon_{kl}$ ($k, l = 1, 2, 3$) the linear strains. The elastic constants $C_{ijkl}$ are given by

$$C_{ijkl} = \frac{E}{(1 + \nu)(1 - 2\nu)} \delta_{ij} \delta_{kl} + \frac{E}{2(1 + \nu)}(\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk})$$

with $\delta$ the Kronecker delta and $E$ and $\nu$ the Young’s modulus and Poisson’s ratio, respectively. Relation (1) shows that the damage variable acts as a stiffness reduction factor. For increasing damage, the effective stiffness moduli $(1 - D)C_{ijkl}$ decrease, until they become zero for $D = 1$. After a certain amount of loading this critical state may have been reached in a finite region, in which the mechanical integrity and stiffness have been completely lost and which can therefore be regarded as the continuum damage representation of a crack.

The growth of damage under the influence of mechanical loading is described by relating the damage variable to the deformation [12, 13]. For this purpose a damage loading function is introduced:

$$f(\tilde{\varepsilon}, \kappa) = \tilde{\varepsilon} - \kappa$$

with $\tilde{\varepsilon}$ a positive equivalent measure of the strain state and $\kappa$ a threshold parameter. Since high-cycle fatigue damage is usually the result of plastic flow at the microscale, it seems natural to use the von Mises strain to define the equivalent strain measure

$$\tilde{\varepsilon} = \frac{1}{1 + \nu} \sqrt{-3J_2}$$

where

$$I_1 = \varepsilon_{kk}, \quad J_2 = \frac{1}{4} I_1^2 - \frac{1}{2} \varepsilon_{ij} \varepsilon_{ij}$$

The factor $1/(1 + \nu)$ in (4) scales the equivalent strain such that it equals the axial strain in the uniaxial, tensile stress case. The equation $f = 0$ defines a loading surface in strain space. For strain states within this loading surface ($f < 0$) there can be no growth of damage and the material behaviour is elastic. The damage variable can only increase when the equivalent strain reaches the
threshold value $\kappa$, i.e. when $f \geq 0$. The loading surface can therefore be related to the fatigue limit: if the strain state remains within the loading surface everywhere in a component, there will be no damage development and the component has an infinite fatigue life. In addition to the condition $f \geq 0$, it is assumed that the damage variable can only increase for continued loading, i.e., for $\dot{f} \geq 0$, and that it remains constant during unloading [13]. The rate of damage growth can then be written as

$$\dot{D} = \begin{cases} 
  g(D, \dot{e}) \dot{\varepsilon} & \text{if } f \geq 0 \text{ and } \dot{f} \geq 0 \text{ and } D < 1 \\
  0 & \text{else} 
\end{cases} \quad (6)$$

The dependence of $\dot{D}$ on the strain rate $\dot{\varepsilon}$ has been taken as linear in this relation in order to avoid rate (frequency) effects.

When dealing with loading histories composed of well-defined, discrete cycles, an evolution law in terms of the number of cycles and their amplitudes is often considered more practical. The number of cycles, $N$, is then regarded as a continuous, time-like variable and the growth of damage, which occurs during discrete time intervals within a cycle, is spread to a continuous evolution over the entire cycle. Such a cycle-based formulation can be obtained from (6) by integration over one loading cycle and approximating the growth of damage within this cycle, resulting in a relation of the form [1, 13, 14]

$$\frac{\partial D}{\partial N} = G(D, \varepsilon_a) \quad (7)$$

with $\varepsilon_a$ the amplitude of the equivalent strain cycle.

Fatigue damage growth relations are often formulated directly in a cycle-based format, but traditionally in terms of stresses rather than strains [4, 15–17]. Notice, however, that such a stress-based evolution can always be rewritten in a strain format by substitution of the stress-strain relations (1). Here (6) will be regarded as the primary definition. This relation does not require that the loading history consists of well-defined loading cycles. As a result, variable amplitude loading, overloads, etc., can be dealt with in a natural fashion, although the load interactions and retardation effects associated to these irregular loadings can be described only to a limited degree by the elasticity-based damage model. It will be shown in Section 5 that efficient numerical analyses of constant amplitude loading are nevertheless possible using definition (6) if the appropriate numerical techniques are used.

For the evolution function $g(D, \dot{\varepsilon})$ in Equation (6) an expression is used here which is slightly different from that proposed by Paas [14]:

$$g(D, \dot{\varepsilon}) = C e^{\kappa_D \varepsilon} \dot{\varepsilon}^\beta \quad (8)$$

with $C, \kappa$ and $\beta$ material parameters. The growth of damage and the fatigue life resulting from this evolution law can be solved in closed form for the situation of uniaxial, fully reversed loading with a constant strain amplitude $\varepsilon_a$. It is furthermore assumed that the equivalent strain equals the axial strain amplitude $\varepsilon_a$ at both extremes of the strain cycles (i.e. in tension and compression) and that there is no fatigue limit ($\kappa_0 = 0$). After substitution of (8), relation (6) can then be integrated over $N$ cycles, yielding

$$D = -\frac{1}{\kappa} \ln \left( 1 - \frac{2\kappa_C}{\beta + 1} \varepsilon_a^{\beta + 1} N \right) \quad (9)$$
The fatigue life \( N_f \) is obtained as a function of the strain amplitude from (9) by setting \( D = 1 \) and solving for \( N \), resulting in

\[
N_f = \frac{\beta + 1}{2\alpha C} (1 - e^{-\alpha}) \varepsilon_a^{-\frac{\beta}{1 + \beta}}
\]  

(10)

This equation can also be written as

\[
\varepsilon_a = \left(\frac{\beta + 1}{\alpha C} (1 - e^{-\alpha})\right)^{\frac{1}{\beta + 1}} \left(2N_f\right)^{-\frac{1}{\beta + 1}}
\]  

(11)

which is of the same form as the high-cycle part of the classical strain-based approach to fatigue [18], or Basquin’s law [19]

\[
\varepsilon_a = \frac{\sigma'_f}{E} (2N_f)^b
\]  

(12)

with the fatigue strength coefficient \( \sigma'_f \) and fatigue strength exponent \( b \) replaced by

\[
\sigma'_f = E \left(\frac{\beta + 1}{\alpha C} (1 - e^{-\alpha})\right)^{\frac{1}{\beta + 1}}, \quad b = -\frac{1}{\beta + 1}
\]  

(13)

Using expression (10) for the fatigue life, relation (9) can be rewritten in terms of the relative number of cycles \( N/N_f \):

\[
D = -\frac{1}{\varepsilon} \ln \left(1 - (1 - e^{-\alpha})\frac{N}{N_f}\right)
\]  

(14)

This relation has been plotted in Figure 1 for several values of the parameter \( \varepsilon \), along with experimental data taken from Hua and Socie [16]. The effect of variation of \( \varepsilon \) is quite clear: a higher value of \( \varepsilon \) results in an initially slower accumulation of damage, but a higher growth rate towards the end of the fatigue life. The value \( \varepsilon = 10 \) gives a reasonable fit of the experimental data. After determining this parameter, the other two parameters of the evolution law, \( C \) and \( \beta \), can then be solved from relations (13) for the fatigue strength coefficient and exponent, the values of which are available for many materials in fatigue handbooks.

3. LOCALIZATION AND MESH SENSITIVITY

Standard damage models and other continuum models of material degradation show a response in which all damage is concentrated in a surface, i.e. in a vanishing volume. This localization of damage is inconsistent with the definition of the damage variable as an average representation of microstructural damage. As a consequence, the behaviour of the model is physically unrealistic. Pathological localization of damage is often connected with the fact that the damage growth results in softening behaviour. This means that components of the stress rate, which can be written as

\[
\ddot{\varepsilon}_{ij} = \ddot{C}_{ijkl}\dot{\varepsilon}_{kl}
\]  

(15)

become negative for positive strain rates \( \dot{\varepsilon}_{kl} \). For the fatigue damage model which is considered here the tangential stiffness \( \ddot{C}_{ijkl} \) is given by

\[
\ddot{C}_{ijkl} = (1 - D)C_{ijkl} - g(D, \varepsilon)C_{ijmn}v_{mn} \frac{\ddot{\varepsilon}}{\dot{\varepsilon}_{kl}}
\]  

(16)
As the damage variable grows, the second term in this expression may begin to dominate the first term, so that negative stress rates may indeed appear. As a result of softening, the ellipticity of the rate equilibrium problem may be lost [20]. Loss of ellipticity occurs when the acoustic tensor $n_i\tilde{C}_{ijkl}n_l$ becomes singular for some unit vector $n$, i.e., when

$$\det(n_i\tilde{C}_{ijkl}n_l) = 0$$  \hspace{1cm} (17)$$

The vector $n$ is then the normal to a characteristic surface segment. Since solutions of linear partial differential equations with smooth coefficients can have discontinuities or discontinuous derivatives only across characteristic surfaces [21], this opens the possibility of jumps in the velocity solution. A discontinuity of the velocity across a characteristic surface results in a strain rate singularity on this surface, which in turn renders the damage rate singular (see Equation (6)). This means that for continued loading the damage variable immediately becomes critical on the characteristic segment and thus that instantaneous failure of this surface is predicted. In order to follow the stress drop resulting from this instantaneous loss of stiffness, material adjacent to the characteristic segment must unload elastically, so that the growth of damage indeed localizes in the surface segment.

Loss of ellipticity plays an important role in localization of deformation and damage in static fracture, e.g. in quasi-brittle damage [22]. In fatigue, however, the slow initial growth of damage implies that the ellipticity of the rate equilibrium problem is preserved until near the end of the fatigue life, when the growth of damage becomes faster. Indeed, there can be no loss of ellipticity at all for the cycle-based growth law (7). If $f_L$ denotes the frequency of the loading cycles, the rate of damage growth is given by (7) as

$$\dot{D} = f_L G(D, \varepsilon_a)$$  \hspace{1cm} (18)$$

Using this relation, the stress rate $\dot{\sigma}_{ij}$ can be written as

$$\dot{\sigma}_{ij} = (1 - D)C_{ijkl}\dot{\varepsilon}_{kl} - f_L G(D, \varepsilon_a)C_{ijkl}\dot{\varepsilon}_{kl}$$  \hspace{1cm} (19)$$

where $\dot{\varepsilon}_{kl}$ must be interpreted as the strain state for which $\tilde{\dot{\varepsilon}} = \varepsilon_a$. Likewise, $\sigma_{ij}$ represents the stress envelope rather than the stress variations within the loading cycles. The second term in (19), which represents the effect of damage growth on the stress rate, does not depend on the strain rate. As
a consequence, this term appears as a source term rather than a differential coefficient when (19) is substituted into the rate equilibrium equations $\frac{\partial \bar{\sigma}_{ij}}{\partial x_i} = 0$:

$$
(1 - D)C_{ijkl} \frac{\partial^2 \bar{v}_{k}}{\partial x_i \partial x_l} = f_L G(D, v_a) C_{ijkl} \frac{\partial \bar{e}_{kl}}{\partial x_i} + f_L \left( \frac{\partial G \partial D}{\partial x_i} + \frac{\partial G \partial v_a}{\partial x_i} \right) C_{ijkl} \frac{\partial \bar{e}_{kl}}{\partial x_i} 
$$

Ellipticity is lost when the characteristic determinant associated to this set of equations becomes zero, i.e. when

$$
\text{det}((1 - D)n_iC_{ijkl}n_j) = 0
$$

After substitution of the elastic constants according to (2) and some algebra, this condition can be rewritten as

$$
(1 - D)^3 \frac{E^3}{4(1 + v)^3} \frac{1 - \nu}{1 - 2\nu} = 0
$$

For $\nu < 0.5$ this equation cannot be satisfied when $D < 1$, so that it can be concluded that the rate equilibrium equations indeed remain elliptic until a crack is initiated.

As a result of the fact that loss of ellipticity does not occur at all or is limited to the very last stages of the fatigue damage process, the damage growth remains stable and affects a finite volume throughout the crack initiation phase. Once a crack is initiated, however, the continued growth of damage nevertheless localizes in a surface and the predicted crack growth still becomes non-physical. This localization during crack growth is not due to loss of ellipticity, but is related to the strain singularity which is inevitably present at the tip of the crack. Once the damage variable becomes critical in a certain point and a crack is thus initiated, the displacement and velocity vectors must become discontinuous across this crack. This implies that the strain (rate) field at the crack tip becomes singular. Since the damage growth rate is directly related to the equivalent strain—or its amplitude in the cycle-based approach—the strain singularity at the tip results in an infinite damage rate. For continued deformation all stiffness is therefore lost instantaneously at the most critical point in front of the crack tip and the crack thus starts to propagate. Since the material adjacent to the crack must unload elastically in order to follow the resulting stress drop, the width of the crack remains zero. This implies that the strain and damage growth rate at the crack tip remain singular as the crack grows and consequently that the crack grows at an infinite rate. No work is needed in this fracture process, since it involves damage growth in a vanishing volume. It is emphasised that this mechanism of damage localization and instantaneous crack growth is activated even if the rate equilibrium equations remain elliptic. Loss of ellipticity may cause premature initiation of a crack and thus result in perfectly brittle crack growth, but the pathological propagation behaviour is nevertheless due to the singularity of the damage rate at the crack tip.

Finite element solutions try to follow the non-physical behaviour of the continuum model, but are limited in doing so by their finite spatial resolution. In standard finite element methods the displacement field must be continuous. The displacement jumps and singular strains of the actual solution can therefore only be approximated by high, but finite displacement gradients in the finite element solution. As a consequence, a finite volume is involved in the damage process, and a positive amount of energy is dissipated. Also, because the damage growth rate at the tip of the damage band remains finite, the crack propagates at a finite velocity. When the spatial discretization grid is refined, however, the finite element approximation becomes more accurate.
in the sense that the displacement gradients which describe the discontinuities become stronger. Consequently, the predicted fracture energy becomes smaller and the crack propagates faster. In the limit of infinitely small elements, the actual solution is retrieved, i.e. a vanishing fracture energy and an infinite crack growth rate. This convergence of the finite element approximation to the actual, non-physical solution of the problem is the origin of the apparent mesh sensitivity of damage models and other continuous descriptions of fracture.

An example of the apparent mesh sensitivity is given in Figure 2. The diagram shows the steady-state fatigue crack growth rate predicted by a finite element analysis vs the size of the elements which were used in the analysis. The problem geometry, loading conditions and modelling for which these results have been obtained will be detailed in Section 6. The dependence of the crack growth per cycle $\frac{da}{dN}$ on the element size $h$ is quite strong in this example: a decrease of the element size by roughly one decade leads to an increase of the crack growth rate by almost three decades.

4. GRADIENT ENHANCEMENT

An effective method to avoid pathological localization of damage and thus mesh sensitivity in finite element analyses is to add non-local terms to the constitutive model. This approach has been successfully applied to damage models for static and dynamic fracture and for creep [6, 8, 23, 24]. The spatial interactions resulting from the non-locality prevent the damage growth from localizing in a surface. Instead, the damage growth occupies a finite band, the width of which is related to the internal length scale provided by the non-locality. Since the volume affected by damage is in reality strongly related to the scale of the microstructure of a material, the internal length must be related to the scale of this microstructure.

In its traditional, integral form non-locality can be introduced in the fatigue model of Section 2 by defining a non-local equivalent strain $\tilde{\varepsilon}$ according to (cf. Pijaudier-Cabot and Bažant [8])

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

(23)

with $\Omega$ the problem domain and $\psi(y; x)$ a weight function defined by

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

(24)

where $\rho = |x - y|$ denotes the distance between point $y$ and $x$ and $l$ denotes the internal length of the non-local continuum. The factor $1/\Psi(x)$ in (23), with $\Psi(x)$ defined by

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

(25)

scales $\tilde{\varepsilon}$ such that it equals $\tilde{\varepsilon}$ for homogeneous strain states. The loading function (3) and the damage evolution law (6) are now rewritten in terms of the non-local strain:

$$f(\tilde{\varepsilon}, \kappa) = \tilde{\varepsilon} - \kappa$$

(26)
and

\[ \dot{D} = \begin{cases} 
q(D, \dot{\varepsilon}) \dot{\varepsilon} & \text{if } f \geq 0 \text{ and } \dot{f} \geq 0 \text{ and } D < 1, \\
0 & \text{else} 
\end{cases} \]  

(27)

Substitution of relations (23) and (27) into the equilibrium equations results in a set of integro-differential equations. This complication is avoided if the integral relation (23) is replaced by a differential approximation. An effective differential form can be obtained by substitution of a truncated Taylor expansion for the local equivalent strain \( \dot{\varepsilon} \) in (23), followed by some mathematical manipulation [22]. This leads to the following approximation of relation (23):

\[ \ddot{\varepsilon} - c \nabla^2 \ddot{\varepsilon} = \ddot{\varepsilon} \]  

(28)

where \( c = \frac{1}{2} I^2 \) and \( \nabla^2 \) denotes the Laplacian operator: \( \nabla^2 = \sum_i \partial^2 / \partial x_i^2 \). The non-local equivalent strain \( \ddot{\varepsilon} \) is now defined by (28) as the solution of a partial differential equation in which the local equivalent strain \( \dot{\varepsilon} \) appears as a source term. In order to uniquely define the relation between the local and non-local equivalent strains, a boundary condition must be added to partial differential equation:

\[ \frac{\partial \ddot{\varepsilon}}{\partial n} = 0 \]  

(29)

where \( \partial / \partial n \) denotes the derivative normal to the boundary of the problem domain. If definition (28) of the non-local equivalent strain is used in relations (26) and (27), a gradient-enhanced damage model is obtained, the response of which qualitatively agrees with that of the integral model [25]. Indeed, it can be shown that Equation (28) can also be written in the integral format (23) by redefining the weight function \( \psi \) [26, 27].

It can be verified that the gradient enhancement according to (26)–(29) removes the possibility of loss of ellipticity of the rate equilibrium problem [26, 28]. As a result, the deformation and damage cannot concentrate in a surface and the damage growth remains stable until crack initiation. However, when \( D = 1 \) somewhere in the component, and a crack is initiated, a strain singularity may be unavoidable at the crack tip. It is important that the damage growth rate remains finite, because the crack growth would otherwise be instantaneous (see Section 3). Since the damage growth rate depends on the non-local equivalent strain \( \ddot{\varepsilon} \) in the gradient formulation, this implies that \( \ddot{\varepsilon} \) must remain finite at the crack tip in order to have a finite crack growth rate.

Analytical expressions for the non-local strain can be obtained for a linear elastic crack in an infinite medium. One should realize that this situation is not representative for crack growth in a damaging material, since the development of damage in front of the crack may have an important effect on the strain singularity at the crack tip. Indeed, for a mode-III problem Liu and Murakami [29] have shown that the degree of singularity of the asymptotic strain field is larger for an assumed damage field than in the elastic case. For the cylindrical damage distribution considered by these authors the strain singularity varies from the elastic \( r^{-1/2} \) singularity for no damage to an \( r^{-1} \) singularity for widespread damage. The present analysis, which assumes the \( r^{-1/2} \) singularity predicted by linear fracture mechanics, can therefore only give an indication of the real crack growth behaviour. It is also relevant for situations where a singularity is a priori present as a consequence of the problem geometry, e.g. at sharp notches. The conventional, local damage theory then predicts immediate, complete fracture for each positive loading level. In order to avoid this non-physical behaviour, \( \ddot{\varepsilon} \) must also remain finite in case of such geometrical singularities.
Figure 3. Linear elastic crack problem.  

A two-dimensional, plane-stress configuration is considered here, see Figure 3. Cartesian co-ordinates \( x_1, x_2 \) and polar co-ordinates \( r, \vartheta \) will be used as convenient; the origin of both co-ordinate systems coincides with the crack tip. The crack is assumed to be loaded in mode-I. The asymptotic strains at the crack tip given by linear elastic fracture mechanics result in an asymptotic equivalent strain

\[
\tilde{\varepsilon} = \frac{K_1}{2E\sqrt{2\pi r}} \sqrt{(1 - \cos \vartheta)(5 - 3\cos \vartheta)}
\]  

(30)

with \( K_1 \) the stress intensity factor. The nonlocal equivalent strain is now obtained by solving partial differential equation (28) for this source term. If it is assumed that the solution is at least continuously differentiable across the \( x_1 \)-axis, the problem need only be solved for the half-plane \( \Omega^+ : x_2 > 0 \) using a Neumann boundary condition \( \partial \tilde{\varepsilon}/\partial x_2 = 0 \) at \( x_2 = 0 \). For \( x_1 < 0 \) this condition represents the free boundary of the crack surface and for \( x_1 > 0 \) it follows from the symmetry of the problem. This boundary value problem is solved by using Green’s function associated to it. The free-space Green’s function for Equation (28) in \( \mathbb{R}^2 \) reads [30]

\[
G_f(x; y) = \frac{1}{2\pi e} K_0\left(\frac{\rho}{\sqrt{e}}\right)
\]  

(31)

with \( \rho = |x - y| \) and \( K_0(z) \) the modified zeroth-order Bessel function of the second kind. It can easily be verified that Green’s function for the half-space problem with the homogeneous Neumann boundary condition is then given by

\[
G(x; y) = G_f(x; y) + G_f(x; y')
\]  

(32)

with \( x, y \in \Omega^+ \) and \( y' = [y_1, -y_2]^T \) the mirror point of \( y \) in the \( x_1 \)-axis. The solution of the half-space problem can now be written as

\[
\tilde{\varepsilon}(x) = \int_{\Omega^+} \tilde{\varepsilon}(y)G(x; y) \, d\Omega
\]  

(33)

or, using (32) and the symmetry \( \tilde{\varepsilon}(y') = \tilde{\varepsilon}(y) \) of the equivalent strain, as

\[
\tilde{\varepsilon}(x) = \int_{\Omega} \tilde{\varepsilon}(y)G_f(x; y) \, d\Omega
\]  

(34)
where $\Omega$ is defined as $\{x \in \mathbb{R}^2 | r > 0, \ -\pi < \vartheta < \pi \}$. In the limit $x \to 0$ the non-local strain approaches

$$\bar{\varepsilon}(0) = \int_{\Omega} \bar{\varepsilon}(y) G_t(0; y) \, d\Omega$$

(35)

Evaluation of the integral yields after some calculus [26]

$$\bar{\varepsilon}(0) = \frac{K_1 \Gamma^2(\frac{1}{3})}{\pi^{1/3} E \sqrt{\pi}} \left(1 + \frac{1}{2} \frac{1}{\sqrt{3}} \arctanh \left(\frac{1}{2} \sqrt{3}\right)\right)$$

(36)

with $\Gamma(z)$ the gamma function

$$\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} \, dt$$

(37)

Expression (36) is indeed finite for $c > 0$, so that the damage growth rate at the crack tip remains finite in the enhanced model. This in turn means that in this simplified situation a finite (initial) crack growth rate is obtained instead of the instantaneous fracture predicted by the standard, local damage model. It has already been mentioned that it is difficult to verify analytically whether the damage rate remains finite also in the coupled model, in which the damage influences the strain singularity. However, numerical results indicate that this is indeed the case, see Section 6. As a result, a finite crack growth rate is obtained.

5. FINITE ELEMENT IMPLEMENTATION

In mathematical terms the essential difference between the gradient-enhanced damage formulation introduced in Section 4 and the classical, local damage models consists of the additional partial differential equation (28). This equation must be solved simultaneously with the standard equilibrium equations. For finite element implementations this means that $\bar{\varepsilon}$ must be discretised in addition to the displacement components. Furthermore, a non-standard time discretization is used in order to limit the number of time increments needed for practical simulations.

5.1. Spatial discretization

The discrete form of the equilibrium equations follows from the standard transition to the weak form and a Galerkin discretization of the displacements by $u = Na$:

$$\int_{\Omega} B^T \sigma \, d\Omega = \int_{\Gamma} N^T t \, d\Gamma$$

(38)

where the matrices $N$ and $B$ contain the displacement interpolation functions and their derivatives, respectively, and the column matrices $\sigma$ and $t$ contain the Cauchy stresses and boundary tractions. Similar to the equilibrium equations, the weak form of Equation (28) can be discretised as [22]

$$\int_{\Omega} (cB^T \tilde{B} + \tilde{N}^T \tilde{N}) \, d\Omega = \int_{\Omega} \tilde{N}^T \tilde{e} \, d\Omega$$

(39)

with $\tilde{N}$ and $\tilde{B}$ containing the interpolation functions of the non-local equivalent strain and their derivatives and $\tilde{e}$ being the column with nodal values of $\tilde{e}$. The finite element interpolations of the
displacements and the non-local strain need to satisfy only the standard, $\mathcal{C}_0$-continuity requirements. The order of each of the interpolations can be selected independently, although some combinations may result in stress oscillations [26].

5.2. Time integration

Since damage growth is defined by relation (27) in a rate format, it must be integrated over each time increment of the numerical analysis in order to obtain the damage at the end of the increment. The damage variable at the end of a time increment $\Delta t$ can formally be written as

$$ D(t + \Delta t) = D(t) + \int_t^{t + \Delta t} \dot{D}(\tau) d\tau $$

(40)

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

$$ \int_t^{t + \Delta t} \dot{D}(\tau) d\tau \approx \frac{1}{2}(\dot{D} + \dot{D}^{t + \Delta t})\Delta t $$

(41)

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

The evolution of $\dot{\varepsilon}$ during an interval $(t, t + \Delta t)$ which comprises a number of loading cycles is shown schematically in Figure 4. It has been assumed in this figure that the loading is proportional and fully reversed. Each loading cycle results in two maxima of $\dot{\varepsilon}$, one corresponding to the tensile part of the cycle and one to the compressive part. Both maxima, which are denoted by $\dot{\varepsilon}_{a1}$ and $\dot{\varepsilon}_{a2}$, follow the envelope $\dot{\varepsilon}_a$. The integration of the damage rate according to (40) still holds for the situation of Figure 4. However, the integral in the right-hand side of (40) can no longer be directly approximated by (41) because $\varepsilon$ and $\dot{\varepsilon}$, 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 $\Delta t$:

$$ D^{t + \Delta t} = D^t + \sum_{n=N}^{N+\Delta N} \int_{t_n}^{t_n + \Delta t_n} \dot{D} d\tau $$

(42)

where $t_n$ and $t_n + \Delta t_n$ correspond to the beginning and end of cycle $n$, respectively, and $N$ and $N + \Delta N$ correspond to $t$ and $t + \Delta t$. Taking into account that damage growth is possible only if $\dot{\varepsilon} \geq \varepsilon$ and $\dot{f} \geq \varepsilon \geq 0$ and using evolution law (27), relation (42) can be rewritten as

$$ D^{N + \Delta N} = D^N + \sum_{n=N}^{N+\Delta N} \left( \int_{\varepsilon_{a1}}^{\varepsilon_{a1}} g(D, \varepsilon) d\varepsilon + \int_{\varepsilon_{a2}}^{\varepsilon_{a2}} g(D, \varepsilon) d\varepsilon \right) $$

(43)
where it has been assumed that $\bar{\varepsilon}_{a1}, \bar{\varepsilon}_{a2} \geq \kappa$. Obviously, cycles in which $\bar{\varepsilon} < \kappa$ (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 $\bar{\varepsilon}_{a}$, 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 (41), the sum over the cycle numbers in (43) 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, $\Delta N$:

$$D^{N+\Delta N} = D^N + \frac{1}{2} (G(D^N, \bar{\varepsilon}^N) + G(D^{N+\Delta N}, \bar{\varepsilon}^{N+\Delta N})) \Delta N$$

(44)

with $G(D, \bar{\varepsilon}_a)$ defined by

$$G(D, \bar{\varepsilon}_a) = 2 \int_{\kappa}^{\bar{\varepsilon}_a} g(D, \bar{\varepsilon}) d\bar{\varepsilon}$$

(45)

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

Relation (44) is a non-linear 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 $D^{N+\Delta N}$ on the right-hand side of (44) by a predictor value $D^p$ based on a forward Euler step:

$$D^p = D^N + G(D^N, \bar{\varepsilon}^N) \Delta N$$

(46)

$$D^{N+\Delta N} = D^N + \frac{1}{2} (G(D^N, \bar{\varepsilon}^N) + G(D^p, \bar{\varepsilon}^{N+\Delta N})) \Delta N$$

(47)

Since the incremental damage growth is given by (46) and (47) in terms of the non-local 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.

It is interesting to note that relations (46) and (47) can also be obtained by applying Heun’s method to the non-local equivalent of the cycle-based evolution of damage, (7). If the corrector-step is omitted, the ‘jump-in-cycles’ procedure proposed by Lemaitre and Doghri [31] on the basis of this approach is retrieved. It has been found, however, that including the correction improves the accuracy considerably for the strongly progressive damage growth observed in fatigue, while the additional computational effort is limited [26].

The accuracy of the computed damage growth increases when a smaller step size $\Delta N$ 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 damage growth relations are strongly progressive (see for instance Figure 1). This implies that the solution fields associated to them diverge and thus the 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 (44) 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 \( \Delta N \), it is assumed that the computed value of the damage variable at \( N \), denoted as \( \tilde{D}_N \), contains a small error \( \delta \): \( \tilde{D}_N = D_N + \delta \). Substitution of this value in (44) gives the computed damage at the end of the increment as

\[
\tilde{D}_N^{N+\Delta N} = \tilde{D}_N + \frac{1}{2}(G(\tilde{D}_N^N, \varepsilon_a^N) + G(\tilde{D}_N^{N+\Delta N}, \varepsilon_a^{N+\Delta N}))\Delta N
\]

(48)

Linearizing this relation with respect to \( \delta \) and \( \Delta N \) results in a first-order approximation of the inherited error at \( N + \Delta N \): \( \tilde{D}_N^{N+\Delta N} - D_N^{N+\Delta N} \approx \left(1 + \frac{\partial G}{\partial D}\Delta N \right) \delta \)

(49)

The increment size is now selected such that the second term in the amplification factor equals a predefined constant \( \eta \):

\[
\Delta N = \frac{\eta}{\partial G/\partial D}
\]

(50)

Typically, values of \( \eta \) between 0.1 and 0.5 give a good balance between accuracy and efficiency. In principle, the derivative \( \partial G/\partial D \) in (50) 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 (50) in order to prevent excessively small or large cycle increments.

5.3. Iterative scheme

The spatial and time discretization reduce the equilibrium problem to a set of coupled, non-linear algebraic equations, which can be written as

\[
f_{ai} = f_{ae}
\]

(51)

\[
K_{ee}e - f_e = 0
\]

(52)

where

\[
f_{ai} = \int_\Omega B^T \sigma \, d\Omega, \quad f_{ae} = \int_\Gamma N^T t \, d\Gamma
\]

(53)

\[
K_{ee} = \int_\Omega (cB^T \dot{B} + \dot{N}^T \dot{N}) \, d\Omega, \quad f_e = \int_\Omega \dot{N}^T \dot{\varepsilon} \, d\Omega
\]

(54)

A full Newton–Raphson scheme is used to solve the set of Equations (51), (52) iteratively at the increment level. For this purpose, relations (51) and (52) are rewritten for iteration \( i + 1 \) as

\[
\delta f_{ai} = f_{ae} - f_{ai}^{i+1}
\]

(55)

\[
K_{ee} \delta e - \delta f_e = f_e^i - K_{ee} e^i
\]

(56)
where $e$, $f_{ai}$ and $f_e$ have been written as the sum of their value in the previous iteration $e', f_{ai}', f_e'$ and iterative corrections $\delta e$, $\delta f_{ai}$, $\delta f_e$. The variation of $f_e$ can be linearized using

$$\delta e = \left( \frac{\delta e}{\delta e} \right)^T B \delta a$$

(57)

where the derivative of the equivalent strain must be evaluated for $e = e'$. Similarly, changes of the internal nodal forces, $\delta f_{ai}$, are linearized by

$$\delta \sigma = (1 - D')CB\delta a - \delta D CB a'$$

(58)

where the matrix $C$ contains elastic constants (2). If the conditions for damage growth are satisfied, the iterative change of the damage variable follows from linearizing relation (47) as

$$\delta D = \frac{1}{2} \frac{\partial D}{\partial \tilde{e}} \Delta N \bar{\tilde{e}}$$

(59)

where the index $a$ of $\tilde{e}_a$ has been dropped and $D' = D_0$.

Using (57)–(59) in (55) and (56) results in the set of linear equations

$$\begin{bmatrix} K_{ai} & K_{ae} \\ K_{ea} & K_{ee} \end{bmatrix} \begin{bmatrix} \delta a \\ \delta e \end{bmatrix} = \begin{bmatrix} f_{ai} - f_{ai}' \\ f_e' - K_{ee} e' \end{bmatrix}$$

(60)

with

$$K_{ai} = \int_{\Omega} B^T (1 - D') CB \, d\Omega, \quad K_{ae} = - \frac{1}{2} \Delta N \int_{\Omega} B^T C \tilde{e} \Delta \tilde{N} \, d\Omega$$

(61)

$$K_{ea} = - \int_{\Omega} \tilde{N}^T \left( \frac{\tilde{e}}{\tilde{e}} \right)^T B \, d\Omega$$

(62)

5.4. Crack growth

A crack is represented in the damage model by a region of completely damaged material. It is important to realize 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

$$\frac{\partial \sigma_{ij}}{\partial \xi_i} = 0$$

(63)

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

$$(1 - D)C_{ijkl} \frac{\tilde{e}^2 u_k}{\tilde{e}_k \tilde{e}_l} - \frac{\partial D}{\partial \xi_i} C_{ijkl} \frac{\tilde{e} u_k}{\tilde{e}_l} = 0$$

(64)

For a given non-critical damage field $D(x) < 1$, 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 $D = 1$, 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 stress is imposed as a boundary condition. 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 discretization 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—non-physical—strains in the crack do not influence the surrounding material. But if this approach is followed for the gradient damage model, the non-local equivalent strain maps the (non-physical) 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 [32] and Geers et al. [33], 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 [26]. 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.

6. APPLICATION TO METAL FATIGUE

The numerical implementation of the gradient damage model has been applied to the problem geometry of Figure 5. 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 [34]. The fatigue limit of this material, defined for $5 \times 10^8$ cycles, is 240 MPa. The damage threshold $\kappa_0$ is obtained by dividing this value by Young’s modulus $E = 210$ GPa, resulting in $\kappa_0 = 0.00114$. Poisson’s ratio has been set equal to
Figure 5. Geometry and loading conditions of the fatigue problem (dimensions in mm).

$v = 0.3$ and the von Mises equivalent strain definition (4) has been used. The parameter $z$ of the damage evolution law (6) has been set equal to 10. The damage growth curve which is obtained for this value is in reasonable agreement with experimental data for other steels (Figure 1). The remaining parameters $C$ and $\beta$ of the evolution law have been solved from expressions (13) for the fatigue strength coefficient ($\sigma _{f}^\prime = 827$ MPa) and the fatigue strength exponent ($b = -0.110$), yielding $C = 6.60 \times 10^2$ and $\beta = 8.09$. The internal length has been set to $\sqrt{c} = 0.1$ mm, which is a rough estimate of the maximum grain size in the material.

Because of symmetry, only the top half of the specimen has been modelled in the finite element analyses. The reference mesh contains a regular grid of elements with an edge length $h = 0.04$ mm in an area of approximately $0.65 \times 0.12$ mm$^2$ at the notch tip, see Figure 6(a) and 6(b). The discretization has been successively refined in this area to $h = 0.02$, 0.01 and 0.005 mm (Figure 6(c)–6(e)). The resulting discretizations consist of 612, 738, 1065 and 2046 plane–stress elements with bilinear displacement and non-local strain interpolations and a constant damage variable. A $2 \times 2$ Gauss scheme has been used for the spatial integration. The integration in time has been carried out with the explicit cycle-based scheme (46), (47) and the adaptive step size selection algorithm, with $\eta = 0.5$ and minimum and maximum increment sizes of 1 and $10^5$ cycles.

Figure 7 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 discretization on the crack shape is shown in Figure 8, in which the final crack pattern has been plotted for the four discretizations. The coarsest mesh (Figure 8(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 gives a good approximation of the crack shape in the finest discretization. The steady-state width of the crack does not vary between the three finest discretizations.

Figure 9 shows the length of the crack, $a$, vs the number of loading cycles $N$ for the four meshes. For an increasingly refined discretization 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 7). The steady-state crack growth rate has been plotted vs the element size $h$ in Figure 10. The corresponding values for the local damage model, which have already been given in Figure 2, have also been plotted in the diagram. Whereas the local model shows a drastic increase of the crack growth rate when the element size is reduced and goes to infinity in the limit $h \to 0$, the growth rate becomes practically constant at $1.42 \times 10^{-5}$ mm/cycle in the gradient model.

Figure 11 demonstrates the influence of the internal length $\sqrt{\varepsilon}$ on the crack growth behaviour. The reference value $\sqrt{\varepsilon} = 0.1$ mm has been reduced to 0.05 and 0.025 mm. The $h = 0.005$ mm mesh has been used in these analyses. For smaller values of the internal length the crack is initiated sooner (after 814 and 309 cycles, respectively) and it grows faster (approximately $2.51 \times 10^{-4}$ and $3.18 \times 10^{-2}$ mm/cycle). This trend is caused by the fact that the non-local equivalent strain follows the local equivalent strain more closely for smaller length scales. As a consequence, the
non-local equivalent strain at the notch tip and, after initiation of a crack, at the crack tip is higher, resulting in faster damage growth and thus in faster initiation and growth of the crack. The effect of the length scale on the final crack shape is shown in Figure 12. As expected, a smaller length parameter results in a smaller steady-state crack width. Furthermore, this constant crack width is reached after a smaller amount of crack growth as a result of the fact that a smaller area has been damaged before a crack is initiated.
7. 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 underlying, local damage modelling. The mesh sensitivity of the local model is caused by the fact that the damage process which represents the initiation and growth of cracks tends to localize in a vanishing volume. This pathological localization is a consequence of the singularity of the damage
rate at the crack tip, which is in turn caused by the strain singularity at the tip. As a result of the singular damage rate the material in front of the crack fails immediately and in a vanishing volume. The crack traverses the remaining cross-section at an infinite growth rate and the thickness of the corresponding damage band is zero. This mechanism is at least as important as the loss of ellipticity to which damage localization is usually attributed, not only in time-dependent fracture processes, as suggested by Murakami and Liu [35], but also in time-independent fracture [26, 28]. Loss of ellipticity may act as a premature initiator of singularities, but it is the instantaneous crack growth caused by the singular damage rate which renders the fracture process perfectly brittle.

The non-locality which is introduced by the gradient enhancement removes damage rate singularities. As a result, crack growth is no longer instantaneous and a finite, positive volume participates in the damage process. This also means that a positive amount of work is needed for the crack growth and that the fracture process is thus no longer perfectly brittle. The volume which is affected by damage and the width of the crack are governed by the intrinsic length scale introduced by the gradient enhancement. It has already been argued that this internal length must be related to the scale of the microstructure of the material, but the precise relations between microstructure and internal length parameters have yet to be developed and will also depend on the
Figure 12. Influence of the internal length scale $\sqrt{\varepsilon}$ on the final damage distribution and crack shape: (a) $\sqrt{\varepsilon} = 0.1$; (b) 0.05; and (c) 0.025 mm.

morphology of the microstructure and the behaviour of its constituents. This is also true for the relations which govern the development of damage, in particular the evolution law and the equivalent strain definition. It is believed, however, that the fact that these relations can and must be connected to the microstructural processes which are responsible for damage and fracture means that the resulting macroscopic models will be more accurate than the classical criteria, in which these processes are largely neglected.

REFERENCES


