A Précis of Some Recent Developments in Computational Failure Mechanics

René de Borst*, Harm Askes, Miguel A. Gutiérrez, Joris J.C. Remmers, Garth N. Wells

Koiter Institute Delft
Delft University of Technology, NL-2600 GB Delft, The Netherlands
e-mail: R.deBorst@LR.TUDelft.nl

Key words: Localisation, cohesive zones, discontinuities, meshless methods, random imperfections

Abstract
A concise overview is given of various numerical methods that can be used to analyse localisation and failure in engineering materials. The importance of the cohesive-zone approach is emphasised and various ways of incorporating the cohesive-zone methodology in discretisation methods are discussed. Next, a simple continuum damage (decohesion) model which preserves well-posedness of boundary-value problems via gradient enhancement is recalled. Using a meshless method the importance of the higher-order gradient terms is assessed. Finally, the model is used in finite element reliability analyses to quantify the probability of the emergence of various possible failure modes.
1. Introduction

Failure in most engineering materials is preceded by the emergence of narrow zones of intense strain ing. During this phase of so-called strain localisation, the deformation pattern in a body rather suddenly changes from relatively smooth into one in which thin zones of highly strained material dominate. In fact, these strain localisation zones act as a precursor to ultimate fracture and failure. Thus, in order to accurately analyse the failure behaviour of materials it is of pivotal importance that the strain localisation phase is modelled in a physically consistent and mathematically correct manner and that proper numerical tools are used.

Until the mid-1980s analyses of localisation phenomena in materials were commonly carried out using standard, rate-independent continuum models. This is reasonable when the principal aim is to determine the behaviour in the pre-localisation regime and some properties at incipient localisation. However, there is a major difficulty in the post-localisation regime, since localisation in standard, rate-independent solids is intimately related to a local change in the character of the governing set of partial differential equations. If this happens, the rate boundary value problem becomes ill-posed and numerical solutions suffer from spurious mesh sensitivity.

To remedy this problem, one must either introduce higher-order terms in the continuum representation that reflect the changes in the microstructure or take into account the inherent viscosity of most engineering materials. An alternative possibility is to by-pass the strain localisation phase and to directly incorporate the discontinuity that arises as an outcome of the strain localisation process. The latter possibility is pursued with so-called cohesive-zone models. We will start by describing them and discuss how they can be introduced in a numerical context. We will show that finite elements with ‘embedded’ localisation zones do not rigorously incorporate discontinuities in finite element models. Conversely, finite element formulations that exploit the partition-of-unity property of the shape functions [1] can, as will be discussed. Indeed, this concept even enables the modelling of a gradual transition from a (higher-order) continuum description to a genuine discontinuum in a numerical context. This is extremely powerful, since now the entire failure process, from small-scale yielding or the initiation of voids and micro-cracks up to the formation of a macroscopically observable crack, can be simulated in a consistent and natural fashion.

Like for finite element methods, the shape functions of meshless discretisations, e.g., the element-free Galerkin method [2], form partitions of unity. Meshless methods have a tremendous advantage for models in which higher-order terms are incorporated, since they inherently provide for the required higher-order continuity. Accordingly, enhanced continuum models can be implemented easily and the importance of the higher-order gradients can be assessed.

Finally, we will indicate how the heterogeneous character of materials at a macroscopic scale can be incorporated in numerical analyses of inelastic solids. In particular, the effect of stochastically distributed imperfections on the failure load in inelastic solids will be quantified in the framework of the finite element reliability method.
2 Cohesive-Zone Models

2.1 Formulation

An important issue when considering failure is the observation that most engineering materials are not perfectly brittle in the Griffith sense, but display some ductility after reaching the strength limit. In fact, there exists a small zone in front of the crack tip, in which small-scale yielding, micro-cracking or void growth and coalescence take place. If this fracture process zone is sufficiently small compared to the structural dimensions, linear-elastic fracture mechanics concepts apply. However, if this is not the case, the cohesive forces that exist in this fracture process zone must be taken into account, and cohesive-zone models must be utilised. In such models, the degrading mechanisms in front of the crack tip are lumped into a discrete line, and a stress-displacement ($\sigma - u$) relationship across this line represents the degrading mechanisms in the fracture process zone. Evidently, the shape of the stress-displacement relation is material dependent. The area under this curve represents the energy that is needed to create a unit area of fully developed crack. It is commonly named the fracture energy $G_f$ and has the dimensions of $J/m^2$. Formally, the definition of the fracture energy reads:

$$G_f = \int \sigma du$$

with $\sigma$ and $u$ the stress and the displacement across the fracture process zone. Cohesive-zone models were introduced by Dugdale [3] and Barenblatt [4] for elastic-plastic fracture in ductile metals, and for quasi-brittle materials Hillerborg [5] published his so-called Fictitious Crack Model, which ensured a discretisation-independent energy release upon crack propagation.

In the past years, cohesive-zone models have shown a strong revival and have been recognised to be an important tool for describing fracture and failure in engineering materials. Especially when the crack path is known in advance, either from experimental evidence, or because of the structure of the material, cohesive-zone models have been used with great success. Then, the mesh lay-out can be made such that the crack path coincides with element boundaries. By inserting interface elements between continuum elements along the potential crack path, a cohesive crack can be modelled exactly [6,7]. To allow for arbitrary crack propagation, Xu and Needleman [8] have inserted interface elements equipped with a cohesive-zone model between all continuum elements. Although analyses with this approach provide much insight, see also [9,10], they suffer from a certain mesh sensitivity, since crack propagation is not entirely free, but restricted to element boundaries, and, more importantly, the approach is not suitable for large-scale analyses.

Distributing the fracture energy $G_f$ over the full width of an element leads to so-called smeared formulations [11,12]. Since the fracture energy is now smeared out over the width of the area in which the crack localises, we obtain

$$G_f = \int \int \sigma d\epsilon(x) dx$$

with $x$ the coordinate orthogonal to the crack direction. When we assume that the strains are constant over a band width $w$ (an assumption commonly made in numerical analyses), we obtain

$$G_f = wg_f$$

with $g_f$ the work dissipated per unit volume of fully damaged material.
When doing so, the global load-displacement diagram can become almost insensitive to the discretisation. This is exemplified in Figures 1 and 2. Figure 1 shows the deformed mesh of a silicon-carbide specimen that is reinforced with carbon fibres (SiC/C composite). The dimensions of the specimen are 30 μm × 30 μm and a uniform loading is applied to the vertical sides. The fibres are assumed to remain elastic and the bond between the fibres and the matrix material is assumed to be perfect. A softening (decohesion) effect is solely adopted for the matrix material, for which a simple Von Mises plasticity model with linear isotropic softening has been used. After onset of softening a clear localisation zone develops, as shown in Figure 1 for a fine discretisation of 15568 triangular elements. Figure 2 shows that, at variance with results without the introduction of a fracture energy, cf. [13], the computed load-displacement curves are rather insensitive to the discretisation, since the coarse mesh (973 elements), the medium mesh (3892 elements) and the fine mesh practically coincide. Nevertheless, in spite of the reasonable results, the smearing out of the cohesive-zone model cannot prevent the local
change of character of the governing equations and the boundary-value problem still becomes ill-posed at a certain stage in the loading process. As a consequence, for another problem, or for slightly different material parameters, mesh-sensitive results may again be obtained.

### 2.2 Finite elements with embedded discontinuities

Finite element models with so-called embedded discontinuities provide a more elegant approach to implement cohesive-zone models in a smeared context [14-19]. There are two versions of these models, namely the strong discontinuity approach and the weak discontinuity approach. We will depart from the latter approach and define an element in which a band is defined within the element where the strains are different in magnitude than the strains in the remainder of the element:

\[
\varepsilon_{ij}^+ = \bar{\varepsilon}_{ij} + \frac{\alpha^+}{2} (n_i m_j + n_j m_i)
\]

and

\[
\varepsilon_{ij}^- = \bar{\varepsilon}_{ij} + \frac{\alpha^-}{2} (n_i m_j + n_j m_i)
\]

with \(n\) a vector normal to the band and \(m\) related to the deformation mode, e.g., \(m = n\) for mode-I behaviour and \(m\) orthogonal to \(n\) for mode-II behaviour. \(\alpha^+\) and \(\alpha^-\) are scalars indicating the magnitude of the strain inside and outside of the band, respectively, measured relative to the average, continuous strain \(\bar{\varepsilon}_{ij}\) in the element. The enhanced strain modes (second part of eqs (4)-(5)) are discontinuous across element boundaries. Consequently, they can be solved for at element level.

The stress-strain relation in the band can be specified independently from that in the bulk of the element. Typically, a softening relation is prescribed which results in an energy dissipation per unit volume \(g_f\) upon complete loss of material coherence. For a band with a width \(w\), which is incorporated in the finite element formulation, we thus retrieve the fracture energy

\[
G_f = w \int \sigma \, d \varepsilon
\]

that is dissipated for the creation of a unit area of fully developed crack.

A problem resides in the determination of the length of the crack band, \(l_{\text{elem}}\) in a specific element. Obviously, for a given length \(l_{\text{elem}}\) the total energy dissipation in an element reads

\[
G_{f,\text{elem}} = l_{\text{elem}} G_f
\]

If the crack length in an element is estimated incorrectly, the energy that is dissipated in each element is also wrong, and so will be the total load-displacement diagram [19]. Different possibilities exist to calculate \(l_{\text{elem}}\), e.g. to relate \(l_{\text{elem}}\) to the area of the element \(A_{\text{elem}}\), to assume that the enhanced mode passes through the element midpoint and to calculate the band length accordingly, or to let the band connect at the element boundaries and to compute the band length in this fashion.

While the above considerations have been set up for the so-called weak discontinuity approach, in which the displacement is continuous, it is also possible to let the enhanced strain modes be unbounded. This so-called strong discontinuity approach can be conceived as a limiting case of the weak discontinuity approach for \(w \to 0\) [20]. The strain then locally attains the form of a Dirac function and the
displacement becomes discontinuous across a single discrete plane. Nevertheless, the integral over time of the product of the traction and the difference in velocities between both sides still equals the fracture energy.

The embedded discontinuity approaches enhance the deformational capabilities of the elements, especially when the standard Bubnov-Galerkin approach is replaced by a Petrov-Galerkin method, which properly incorporates the discontinuity kinematics [20]. As a consequence, the high local strain gradients inside localisation bands are better captured. However, a true discontinuity is not captured because the kinematics of eqs (4) and (5) are diffused over the element when the governing equations are cast in a weak format, either via a Bubnov-Galerkin or via a Petrov-Galerkin procedure.

2.3 Modelling of discontinuities by exploiting the partition-of-unity property

A method in which a discontinuity in the displacement field is captured rigorously has been developed recently on the basis of the partition-of-unity concept [21,22]. A collection of functions \( \phi_i \), associated with node \( i \), form a partition of unity if

\[
\sum_{i=1}^{n} \phi_i(x) = 1
\]

with \( n \) the number of discrete nodal points [1]. For a set of functions \( \phi_i \) that satisfy eq. (8), a field \( u \) can be interpolated as follows

\[
u(x) = \sum_{i=1}^{n} \phi_i(x) \left( a_i + \sum_{j=1}^{m} \psi_j(x) b_{ij} \right)
\]

with \( a_i \) the ‘regular’ nodal degrees-of-freedom, \( \psi_j(x) \) the enhanced basis terms, and \( b_{ij} \) the additional degrees-of-freedoms at node \( i \) which represent the amplitude of the \( j \)th enhanced basis term \( \psi_j(x) \).

A piecewise smooth displacement field \( u \) which incorporates a discontinuity with a unit normal vector \( \mathbf{n} \) pointing in an arbitrary, but fixed direction can be described by:

\[
u(x) = \tilde{u}(x) + H_{\Gamma_d}(x) \tilde{\mathbf{u}}(x)
\]

with \( \tilde{u}(x) \) the standard, continuous displacement field on which the discontinuity has been superimposed. The discontinuous field is represented by the smooth field \( \tilde{u}(x) \) and the Heaviside function \( H_{\Gamma_d} \), centered at the discontinuity plane \( \Gamma_d \). The displacement decomposition in eq. (10) has a structure similar to the interpolation of eq. (9). Accordingly, the partition-of-unity concept can be used in a straightforward fashion to incorporate discontinuities, and thus, cohesive-zone models in a manner that preserves the truly discontinuous character. Indeed, in conventional finite element notation, the displacement field of an element that contains a single discontinuity can be represented as:

\[
u = \mathbf{u} + H_{\Gamma_d} \tilde{\mathbf{u}} = \mathbf{N} \mathbf{a} + H_{\Gamma_d} \mathbf{N} \mathbf{b} = \mathbf{N} (\mathbf{a} + H_{\Gamma_d} \mathbf{b})
\]

where \( \mathbf{N} \) contains the standard shape functions, and \( \mathbf{a} \) and \( \mathbf{b} \) collect the conventional and the additional nodal degrees-of-freedom, respectively. The numerical development now follows standard lines by casting the balance of momentum in a weak format, and, in the spirit of a Bubnov-Galerkin approach, taking a decomposition as in eq. (11) also for the test function. For small displacement gradients a complete derivation can be found in [23], while the extension to large displacement gradients is given in [24].
It is emphasised that in this concept, the additional degrees-of-freedom cannot be condensed at element level, because, at variance with the ‘embedded’ displacement discontinuity approach, it is node-oriented and not element-oriented. It is this property which makes it possible to represent a discontinuity in a rigorous manner.

From eqs (9) and (11) we infer that the partition-of-unity concept can naturally be conceived as a multiscale approach. Decomposing $u(x)$ formally as

$$u(x) = u_C(x) + u_F(x)$$

with

$$u_C(x) = \sum_{i=1}^{n} \phi_i(x) a_i$$

representing the coarse-scale and

$$u_F(x) = \sum_{i=1}^{n} \phi_i(x) \sum_{j=1}^{m} \psi_j(x) b_{ij}$$

representing the fine scale. A more formal relation to the Variational Multiscale Formulation [25] has been given in the companion paper by Munts et al. [26].

As an example we consider the double cantilever beam of Figure 3 with an initial delamination length $a_0$. This case, in which failure is a consequence of a combination of delamination growth and structural instability, has been analysed using conventional interface elements by Allix and Corigliano [27]. The beam is subjected to an axial compressive force $2P$, while two small perturbing forces $P_0$ are applied to trigger the buckling mode. Two finite element discretisations have been employed, a fine mesh with three elements over the thickness and 250 elements along the length of the beam, and a coarse mesh with only one (!) element over the thickness and 100 elements along the length. Figure 4 shows that the calculation with the coarse mesh approaches the results for the fine mesh very well. For instance, the numerically calculated buckling load is in good agreement with the analytical solution. Steady-state delamination growth starts around a lateral displacement $u = 4$ mm. From this point onwards, delamination growth interacts with geometrical instability. Figure 5 presents the deformed beam for the coarse mesh at a tip displacement $u = 6$ mm. Note that the displacements are plotted at true scale, but that the difference in displacement between upper and lower part of the beam is for the major part due to the delamination and that the strains remain small. Another way of post-processing the results could have obviated this slightly misleading picture.

![Figure 3. Double cantilever beam with initial delamination under compression.](image-url)
Debonding (coarse mesh)  
Debonding (dense mesh)  
Perfect bond

Figure 4. Load-displacement curves for delamination buckling test.

Figure 5. Deformation of coarse mesh after buckling and delamination growth (true scale).

3 Continuum-discontinuum transition

The above approach enables the gradual and consistent transition from a continuum to a discontinuum description. In [28] tractionless discontinuities have been inserted in a softening, viscoplastic medium when the stress has become lower than a threshold level, say 1% of the initial yield strength. Accordingly, the viscous property of the continuum has been used to ensure a mesh-independent analysis in the softening continuum prior to local failure and the creation of a traction-free discontinuity.

For many materials the viscosity is so low that addition of it to the constitutive model is not sufficient to restore well-posedness of the boundary value problem during the strain localisation phase. Indeed, models that exploit the non-local interactions in the fracture process zone can be physically better motivated and numerically more effective. Among these models, the gradient-enhanced models have shown to be computationally the most efficient, either in a plasticity-based format [29-31], a damage-based format [32,33], or a combination of both [34]. Especially the gradient-enhanced damage model of Peerlings et al. [32,33] has proven to be very robust and effective, not only for damage evolution under monotonic loading conditions, but also for fatigue loading [35]. For this reason it is summarised below.

Scalar-based damage models are normally rooted in an injective relation between the stress tensor $\sigma$
and the strain tensor $\varepsilon$:

$$\sigma = (1 - \omega)D^e : \varepsilon \quad (15)$$

Herein $D^e$ is the elastic stiffness tensor with the virgin elastic constants $E$ (Young’s modulus) and $\nu$ (Poisson’s ratio). $\omega$ is a monotonically increasing damage parameter, with an initial value 0, for the intact material, and an ultimate value 1, at complete loss of material coherence. It is a function of a history parameter $\kappa$: $\omega = \omega(\kappa)$, with $\kappa$ linked to a non-local strain measure $\tilde{\varepsilon}$ via a loading function $f = \tilde{\varepsilon} - \kappa \quad (16)$

such that loading occurs if $f = 0$, $f' = 0$ and $\omega < 1$. The non-local strain measure is coupled to a local strain measure $\varepsilon = \varepsilon(\varepsilon)$ via:

$$\varepsilon - c_1 \nabla^2 \varepsilon - c_2 \nabla^4 \varepsilon = \tilde{\varepsilon} \quad (17)$$

with $c_1$ and $c_2$ material parameters with the dimension length squared, respectively of length to the power four.

4 Meshless methods for localisation and failure

A clear disadvantage of the use of higher-order continuum theories is the higher continuity that may be required for the shape functions that are used in the interpolation of the non-local variable. The second-order implicit gradient damage theory (eqs (15)-(17) with $c_2 = 0$) is such that after partial integration $C^0$-interpolation polynomials suffice for the interpolation of $\tilde{\varepsilon}$. However, this no longer holds when the fourth-order term is retained. Then, and also because of the moving elastic-plastic boundary in gradient plasticity models, $C^1$-continuous shape functions are required, with all computational inconveniences that come to it. Here, meshless methods, which can easily be constructed such that they incorporate $C^\infty$-continuous shape functions, have a clear advantage. Below we shall apply one such method, namely the element-free Galerkin method [2], to second and fourth-order implicit gradient damage models.

In the element-free Galerkin method, approximants $u^h$ are constructed as

$$u^h = p^T(x) a(x) \quad (18)$$

with $p$ a vector that contains monomials and $a$ a coefficient vector. The approximants are found by minimising the moving weighted least squares sum with respect to $a$:

$$J = \sum_{i=1}^{n} w_i(x) \left( p^T(x_i) a(x) - u_i \right) \quad \text{with} \quad p(x_i) \text{ the value of } p \text{ in node } i, \quad w_i \text{ the weight function attached to this node and } u_i \text{ the nodal displacement.}$$

The weight functions should be smooth and should contain a certain minimum number of other nodes within its domain of influence.

For a three-point bending beam, analyses have been carried out using the element-free Galerkin method both for the second and for the fourth-order gradient damage model. It appeared that the results for both formulations are virtually identical [36].

The higher-order continuity that is incorporated in meshless methods makes them well suited for local-
isation and failure analyses using higher-order continuum models. Moreover, the flexibility is also increased compared to conventional finite element methods, since there is no direct connectivity, which makes placing nodes in regions with high strain gradients particularly simple. This can also be achieved by finite element methods with spatial adaptivity. Originally, adaptivity for localisation analyses was applied using standard continuum models [37]. However, it soon became clear that the inherent loss of ellipticity prevented error estimators to work properly [38]. Contemporary approaches therefore apply mesh adaptivity techniques for failure analyses in conjunction with cohesive-zone models or with regularised continuum models [38-40]. Along the same line of reasoning, the discretisation itself cannot provide a regularisation, neither for finite element methods, nor for meshless methods. Indeed, for the latter class of methods the nodal spacing directly relates to the width of the localisation zone that is resolved if no regularisation is provided for the continuum model.

5 Stochastically Distributed Imperfections

So far, the discussion has concentrated on localisation and the ensuing failure in solids which have uniform strength and stiffness properties. In reality, strength and stiffness have a random distribution over any structure. The distribution and the size of imperfections may have a profound influence on the localisation pattern and, therefore, on the ultimate failure load, as was demonstrated more than half a century ago by Koiter in his landmark dissertation [41] on the influence of imperfections in elastic solids. We may expect that this observation holds a fortiori if material degradation plays a role.

Thus, for realistic analyses of localisation and failure, material parameters like Young’s modulus, the tensile strength and the fracture energy should be considered as random fields, and the most probable realisation(s) should be sought which lead to failure or violate a certain serviceability criterion. Indeed, in such analyses, not only the scatter in material parameters, but also the uncertainty in the boundary conditions should be considered. The simplest, but also the most expensive method, would be to start a nonlinear analysis from different random distributions and to obtain the statistics of the response by carrying out a sufficient number of such Monte-Carlo simulations [42]. Evidently, this is very expensive and a more versatile approach is to utilise the finite element reliability method [43,44].

In the latter approach, the statistics of a certain measure for failure or loss of serviceability, say $Q$, are approximated as follows. First, the material parameters (or the boundary conditions) which are assumed to have a random distribution, are discretised and are assembled in a vector $V$, which is characterised by a joint probability density function $f_V$. Although the discretisation that is used to form $V$ can be different from the finite element discretisation that is used later, it simplifies the implementation if the random cells coincide with finite elements or patches of finite elements. To facilitate further computations, the vector $V$ of random variables is usually converted into a vector $Y$ which consists of uncorrelated variables with a standard normal distribution. A crucial step is then the (nonlinear) mechanical transformation, which, given a random distribution of the material parameters and/or boundary conditions, computes a random response, assembled in $Q$. If $q_0$ denotes the threshold value of the measure $Q$ that is used to assess failure or serviceability, the realisation $v$ of $V$ is sought which furnishes a local maximum of its own probability density through a suitable optimisation algorithm. A weakness of the approach is that it generates a local maximum, which implies that to obtain a good global estimate of the likelihood of failure or loss of serviceability the algorithm must be started from different initial conditions.
As an example we take the double-edge notched specimen of Figure 6. For detailed information regarding the parameters that have been used and the way in which the analysis has been carried out, the reader is referred to [43,44]. Tensile tests on specimens as depicted in Figure 6 tend to be sensitive to the boundary conditions, in particular when the brittleness of the material increases. When an imperfection is not imposed in the material, nor an asymmetry in the boundary conditions, the deformations will remain symmetric throughout the entire loading path. However, if either of these occurs, asymmetric crack propagation evolves from one of the notches at a generic stage in the loading process. The probability that either of these failure modes occurs can be simulated via the approach discussed above where the tensile strength is randomised, while starting from a symmetric as well as from an asymmetric realisation. In particular, the influence of the boundary conditions can be quantified. For instance, taking the upper loading platen fixed of the double-notched specimen, the probability of failure was found to be $P_s = 5.84 \times 10^{-2}$, irrespective whether the algorithm was started from a symmetric or from an asymmetric realisation. Indeed, the failure mode was purely symmetric. However, when the upper loading platen is allowed to rotate freely, an asymmetric mode was found with a probability of failure that is significantly higher than that of the symmetric model, namely $P_a = 0.41$. Again, this result was obtained irrespective of the starting realisation, which, for the symmetric realisation, is probably attributable to numerical round-off errors. Next, the analysis was repeated for a longer specimen ($L = 250\text{mm}$), while keeping the loading platens fixed. Not surprisingly, a symmetric failure mode was found with a probability of failure $P_s$ which is almost the same as for the shorter specimen. However, an asymmetric failure mode now also emerged, with a probability of failure $P_a = 0.13$, which is purely a consequence of the increased rotational freedom of the longer specimen.

It is emphasised that the inclusion of randomness of the material parameters in the analysis does not resolve the issue of loss of ellipticity at the onset of localisation when standard, rate-independent continuum models are considered [42]. Indeed, the above simulations were carried out using the implicit second-gradient continuum damage model of eqs (15)-(17). Here, the situation is similar to the use of mesh adaptivity techniques, which are also unable to remedy the loss of well-posedness which is the fundamental cause of the spurious results which are obtained when analysing localisation in standard, rate-independent continuum models.
References

[22] N. Moes, J. Dolbow, T. Belytschko, A finite element method for crack growth without remesh-


[42] J. Carmeliet, R. de Borst, Stochastic approaches for damage evolution in standard and non-
