MODELS FOR QUASIBRITTLE FAILURE:
THEORETICAL AND COMPUTATIONAL ASPECTS

Milan Jirásek and Bořek Patzák
Laboratory of Structural and Continuum Mechanics,
Department of Civil Engineering,
Swiss Federal Institute of Technology at Lausanne,
CH-1015 Lausanne, Switzerland
e–mail: Milan.Jirasek@epfl.ch, Borek.Patzak@epfl.ch

Key words: quasibrittle materials, cracking, damage, localization, nonlocal continuum, mesh bias, adaptive methods, extended finite elements

Abstract. This paper deals with various aspects of the computational analysis of strain localization and failure in quasibrittle materials. It gives a general overview of three main classes of modeling approaches: models with propagating cohesive displacement discontinuities, softening continuum models with partial regularization by the fracture energy approach, and fully regularized softening continuum models. For the second class of models, the effect of mesh-induced directional bias is demonstrated and the importance of crack-induced anisotropy and of a clean resolution of a stress-free crack is discussed. For the third class of models, a mesh-adaptive technique based on an error indicator combined with a simple error estimator is briefly described and illustrated by an example. As an alternative to the usual h-refinement, an extended finite element method is outlined.
1 Introduction

The behavior of quasibrittle materials (such as concrete, rock, tough ceramics, or ice) subjected to increasing mechanical solicitations is characterized by diffuse microcracking that later localizes in relatively narrow zones, referred to as the fracture process zones. The localization of strain and damage eventually leads to a gradual development of macroscopic stress-free cracks. Despite a considerable progress in the past two decades, theoretical modeling and computational resolution of the localization process up to structural failure still remains a challenging issue of contemporary solid mechanics. The purpose of this paper is to summarize certain latest developments in this area and to outline the current trends and future prospects.

Narrow zones of highly concentrated evolving microdefects can be modeled in many different ways. One possible classification was proposed in [1]. Existing models were divided into three broad and partially overlapping classes, depending on the regularity of the kinematic description. The character of the displacement field and of the corresponding strain field for these three fundamental classes of models is illustrated in Fig. 1, which depicts the one-dimensional situation. The fracture process zone is represented either by a single point at which the displacement field has a jump (Fig. 1a), or by a finite interval. In the latter case, the strain field either has a jump at the boundary of the process zone (Fig. 1b), or remains continuous (Fig. 1c).

Figure 1: Representation of the process zone by a) a strong discontinuity, b) a band of localized strain separated by two weak discontinuities, c) a continuous profile of localized strain. Left column: displacement profile; right column: strain profile
In multiple dimensions, the foregoing characteristics of the kinematic description must be appropriately generalized. Models belonging to the first class admit the presence of a strong discontinuity, i.e., of a curve (in two dimensions) or surface (in three dimensions) across which the displacement field has a jump. In general, the discontinuity is assumed to transmit some cohesive tractions that are related to the opening and sliding component of the displacement jump by a special traction-separation law. Physically, such a discontinuity can be considered as a cohesive crack or slip line. The tractions transmitted by a cohesive crack are usually assumed to vanish when the crack opening exceeds a certain limit, and the crack then becomes stress-free.

Models belonging to the second class use a continuous description of the displacement field but admit the presence of weak discontinuities, i.e., of curves or surfaces across which certain components of the strain field have jumps. Typically, such weak discontinuities form at the internal boundaries that separate a band or layer of softening material from the surrounding material that undergoes unloading. The thickness of the softening band can be considered as a material property independent of the spatial discretization (e.g., finite element mesh), or as a measure of the minimum possible localization pattern that can be resolved on a given mesh. In the former case, the band thickness represents an intrinsic material length and the softening stress-strain law can be uniquely defined, while in the latter case the softening modulus must be adjusted according to the spatial discretization and the model can be interpreted as a regularized form of a cohesive crack model, with a strong discontinuity approached in the limit as the mesh is refined.

Finally, models belonging to the third class are characterized by continuity of both the displacement field and the strain field, and they represent the process zone as a band or layer of softening material, in which the strain gradually increases from the minimum value on the boundary of the band to the maximum value at the center. Such fully regularized models are obtained by proper enhancements of the standard continuum theory, e.g., by the incorporation of nonlocal averages or higher-order gradients of internal variables.

2 Models with propagating cohesive discontinuities

The simplest example of a propagating displacement discontinuity is a stress-free crack with a sharp tip, at which the stress field has a singularity. If the material remains linear elastic, the process zone collapses to a single point and the model does not possess any characteristic length. Such models are suitable for the description of quasibrittle materials only on a very large scale, when the actual process zone is negligible with respect to the characteristic dimensions of the structure. Scaling of nominal strength is then given by a power law. On smaller scales, the model must be refined by introducing a cohesive traction-separation law, which removes the singularity and leads to the formation of a process zone with a finite length but still zero thickness. The length of the cohesive process zone is related to the characteristic length $l_c = E G_F / f_t^2$ that is set by the elastic modulus $E$, fracture energy $G_F$ (area under the traction-separation curve), and tensile strength $f_t$ (stress at which the discontinuity starts opening). Owing to the presence of a characteristic length, the model can reproduce a transitional size effect of a non-power type.
Computational resolution of discrete cracks was traditionally based on finite elements with displacement discontinuities at element interfaces, or on various forms of the boundary element method. If stress singularities are present around crack tips, the accuracy can be increased by using special elements, such as the quarter-node elements in which the singularity of the Jacobian at one of the nodes produces a singularity of the stress approximation [2].

If the displacement discontinuity is allowed to appear only at element interfaces, propagation of the crack requires frequent remeshing, otherwise the crack trajectory would be highly restrained. Even though efficient remeshing techniques are available for two- as well as three-dimensional problems [3, 4, 5, 6], increasing attention is being paid to alternative techniques that can handle displacement discontinuities in a more flexible manner, independently of the mesh. One fruitful research direction has been focused on discontinuities embedded inside finite elements at arbitrary locations and with arbitrary orientations [7, 8, 9, 10, 11, 12]. The standard finite element interpolations are enriched by terms that can reproduce a jump in the displacement field, either directly (by added discontinuous shape functions) or indirectly (by added strain modes that correspond to discontinuous displacements. Virtually all these models deal with nonconforming interpolations, i.e., compatibility is satisfied only in the weak sense. This makes it possible to treat the added degrees of freedom that correspond to discontinuous enrichments as internal ones and eliminate them on the element level. The main advantage is that only standard degrees of freedom (nodal displacements) are kept on the global, structural level, and the number of global equilibrium equations and structure of the stiffness matrix do not change when the crack propagates and enrichments are added to new elements. However, there is a price to pay for this convenience. A detailed analysis of the behavior of a single element with an embedded discontinuity reveals that, in order to guarantee uniqueness of the element response to any prescribed history of nodal displacements, severe restrictions must be placed not only on the element size (to prevent a non-unique response known as “local snapback”) but also on the element shape [13]. These restrictions become even more severe in the presence of multiple discontinuities in one element (needed to describe crack branching) and in three dimensions. When they are violated, the numerical algorithm evaluating the nodal forces cannot be expected to be robust and converge for all possible loading histories (even if the load is applied in very small incremental steps), and divergence on the element level occurs as soon as one of the discontinuities is introduced in an unfavorable position with respect to the basic element. Another inconvenience is that the tangent stiffness matrix of the embedded element is in general nonsymmetric, even if the material stiffnesses of the continuous material and of the cohesive discontinuity are symmetric.

Due to the high sensitivity of elements with embedded discontinuities to the position of the discontinuity, it is practically impossible to extend the technique to three dimensions without using special numerical tricks [14], and the beauty and simplicity of the original idea is lost. This has also been confirmed by the present authors in their unpublished work. The numerical robustness and versatility of the modeling approach dealing with displacement discontinuities can be substantially improved if the discontinuities are incorporated into the enriched interpolation using the partition-of-unity concept. The original idea of the partition-of-unity method [15, 16] was that the approximation space spanned by a standard basis (e.g., by the standard finite element shape functions) can be easily enriched by products of the standard basis functions with special
functions selected by the user and constructed, e.g., from the analytical solution of the problem under some simplifying assumptions. This permits the incorporation of a priori knowledge about the character of the problem and its solutions. Multiplication by the standard shape functions ensures that the enrichment functions have a limited support and that the corresponding degrees of freedom can be assigned to the nodes of the basic finite element mesh. This idea was adapted for linear elastic fracture mechanics in [17], with the enrichment constructed using the near-tip asymptotic solutions and simple Heaviside functions. The method was later called the eXtended Finite Element Method (X-FEM). It can efficiently handle three-dimensional cracks [18] and even branching and intersecting cracks [19].

A big advantage of this technique is that the displacement interpolation is conforming, with no incompatibilities between elements, and that the strain on both sides of a stress-free crack is fully decoupled, which was not the case for most of the previous elements with embedded discontinuities. The added degrees of freedom are global, but they can be assigned to the existing nodes of the basic finite element mesh, without any need for topology changes. Such degrees of freedom are easily inserted into the global set of equations and the resulting stiffness matrix preserves its banded character. Extension of the method to cohesive crack models is reported in [14]. Due to the absence of a stress singularity, no special enrichments around the crack tip are needed, and the enrichment functions are constructed as products of the Heaviside function with standard finite element shape functions that correspond to the nodes of those elements that are intersected by the crack. The cohesive zone model based on the partition of unity, so far studied only in two dimensions, seems to overcome the difficulties associated with the piecewise constant interpolation of the displacement jump used by many previous models, and it even restores the symmetry of the stiffness matrix.

3 Softening continuum models with partial regularization

As already alluded to in the Introduction, models that represent the fracture process zone by a finite band of localized strain can be interpreted in two different ways. One possibility is to consider the thickness of the localization band as a material parameter that has a precise physical meaning. In this case, the spatial discretization must be constructed such that the thickness of the numerically resolved band corresponds to the prescribed value. In a finite element simulation of a softening material, strain typically localizes into one layer of elements. The original idea proposed in [20] was to fix the element size so as to match the physical size of the process zone. This is of course a serious constraint on the mesh, because the trajectory of the softening band is in general not known in advance, so the mesh would need to have a constant density across the entire structure. Also, the effective thickness of the numerically resolved softening band depends on the orientation of the band with respect to the mesh lines. For example, in a square mesh the thickness of a zig-zag band in the diagonal direction is \( \sqrt{2} \) times larger than if the band propagates parallel to the element sides. A partial remedy is provided by the concept of softening bands embedded into finite elements [21, 22, 23]. Here, the elements can be larger (but not smaller) than the prescribed thickness of the process zone, and the results are objective with respect to the relative orientation of the softening band and the elements. Nevertheless,
there remains a constraint on the minimum element size, which can lead to problems when fine
details of the structural geometry need to be resolved. It is also disturbing from the theoretical
point of view that the notion of convergence upon mesh refinement cannot be introduced.

A refined technique, frequently used in practical applications, is based on the adjustment of the
softening modulus according to the element size [24, 25]. This is sometimes called the fracture
energy approach, because the aim is to properly reproduce the energy dissipation in the soften-
ing band. The area under the uniaxial stress-strain curve corresponds to the energy $g_F$ dissipated
per unit volume of the process zone under pure Mode-I failure. In a rigorous approach, only the
part of the energy dissipated after the peak (after the onset of localization) should be taken into
account. The product of this energy density per unit volume with the thickness of the localized
softening band $L_s$ gives the energy dissipation per unit area of the resulting stress-free crack,
i.e., the fracture energy $G_F$. Despite some continuing controversies regarding its objective ex-
perimental evaluation, this energy is usually considered as a fundamental material property.
Instead of treating $L_s$ as a fixed material parameter, one can interpret it as a mesh-related pa-
rameter giving the thickness of the numerically resolved localization band. If the area under the
stress-strain curve (after subtraction of the pre-peak energy dissipation) is adjusted according
to the formula $g_F = G_F / L_s$, energy dissipation in the softening band is described objectively
for meshes of an arbitrary density. Of course, if the mesh is too coarse, some discretization
errors can be introduced, but the important property is that the total dissipation and the global
load-displacement diagram converge to a physically meaningful limit as the mesh is refined. In
the simplest case of uniaxial tension, this limit exactly corresponds to the solution of a cohe-
sive crack model with the same fracture energy $G_F$ and an appropriate traction-separation law.
In fact, the softening continuum model can then be interpreted as a regularized version of a
displacement discontinuity model, with the displacement jump (crack opening) smeared over a
finite distance $L_s$ and replaced by an equivalent inelastic strain.

The technique adjusting the softening part of the stress-strain diagram according to the ele-
ment size provides only a partial regularization of the softening continuum. It removes the
pathological sensitivity to the refinement of the mesh and leads to objective global solution
characteristics. Nevertheless, the displacement and strain fields tend to a solution with a strong
discontinuity as the mesh is refined. Also, the orientation of the numerically resolved process
zone can be affected by mesh-induced directional bias.

To illustrate this point, consider the four-point shear test of a single-edge-notched beam, for
metals known as the Iosipescu beam. The test was adapted for concrete in [26]. The exper-
imentally observed crack trajectory is curved and has the shape shown by the dashed curve in
Fig. 2a. To compare the performance of several softening continuum models with partial regu-
larization by the fracture energy approach, the test is simulated on the quadrilateral mesh shown
in Fig. 2b. In the central part of the beam, where the crack is expected to propagate, the mesh is
quite fine and completely regular. The simulation is done using the following four constitutive
formulations:
1. Isotropic damage model with damage evolution driven by the equivalent strain related to the positive part of effective stress (Rankine-like failure envelope), described e.g. in [27, 28].

2. Standard rotating crack model with maximum principal stress criterion for crack initiation and with computational crack direction rotating with the principal strain axes [29, 30].

3. Rotating crack model with transition to a scalar damage model at the state when the crack opening exceeds a critical level or when the tangent shear modulus drops below a critical fraction of its elastic value [27].

4. Anisotropic damage model based on the principle of energy equivalence and on the microplane concept [31, 32].

The resulting crack trajectories (bands of localized strain) are compared in Fig. 3. For all the models, the crack starts propagating from the notch in the correct direction. The isotropic damage model (Fig. 3a) is sensitive to the mesh bias and the crack trajectory is soon attracted by the mesh lines—the crack propagates vertically and reaches the top surface of the specimen to the left of the loading platen, which does not agree with the experimental results. The rotating crack model (Fig. 3b) gives a somewhat better trajectory but the strain cannot fully localize due to stress locking [33]. Large stresses are present around the crack even when its opening is so

Figure 2: Four-point shear test: a) geometry and loading, b) finite element mesh
Figure 3: Crack trajectory in a four-point shear test: a) isotropic damage model, b) rotating crack model, c) rotating crack model with transition to scalar damage, d) anisotropic damage model
large that the crack should be stress-free, and the simulation fails to converge. A modification developed by the first author [27], based on the transition to a scalar damage formulation at late stages of the softening process, removes locking and leads to a fully localized softening band, which reaches the top surface of the specimen just right to the loading platen (Fig. 3c). This is an improvement over the isotropic damage model, but the trajectory is still too straight. The best results are obtained with the anisotropic damage model outlined in general terms in [31] and developed in [32]. This model is free of locking, and the numerical crack trajectory closely approximates the experimental one.

This example reveals the importance of anisotropy for the correct simulation of fracture propagating along general curved trajectories. The isotropic damage model is affected by the directional mesh bias and leads in this case to a shear-type failure. Models that capture the crack-induced anisotropy of the material are more successful in reproducing the correct failure mode, but they must be capable of cleanly reproducing a widely open stress-free crack within a continuum formulation. When the modified rotating crack model with transition to scalar damage is used, only a part of the process zone close to its tip is simulated in a truly anisotropic fashion; the remaining part of that zone is in the regime described by a fixed anisotropic stiffness tensor multiplied by a scalar parameter that tends to zero. This means that the ratios of the damaged stiffness coefficients are frozen at the moment of transition and the subsequent evolution of damage is isotropic. The microplane-based anisotropic damage model reflects the state of damage by a set of scalar parameters related to preselected spatial directions. The sum of contributions from all these directions provides the damage effect tensor, which is then used in the context of a damage formulation based on the principle of energy equivalence. This allows relatively general anisotropic damage evolutions and, at the same time, locking effects are excluded.

4 Regularized softening continua

4.1 General overview

Full regularization of the localization problem can be achieved by proper generalization of the underlying continuum theory. Generalized continua in the broad sense can be classified according to the following criteria:

1. Generalized kinematics.
   (a) Continua with microstructure.
   (b) Continua with nonlocal strain.

2. Generalized constitutive equations.
   (a) Material models with gradients of internal variables (or gradients of thermodynamic forces, or both).
   (b) Material models with nonlocal internal variables (or nonlocal thermodynamic forces, or both).
Here we focus on the second class of models, with enrichments on the level of the constitutive equations. Their advantage is that the kinematic and equilibrium equations remain standard, and the notions of stress and strain have their usual meaning. However, it is important to keep in mind that they represent the actual state of a heterogeneous material only on the macroscopic level, in the sense of mean values averaged over a certain representative volume. As shown in [34] for elastic composites with a random microstructure, the homogenization theory indicates that the usual local form of the stress-strain law on the macroscopic level is only the first approximation. When higher-order effects are taken into account, the constitutive equations become nonlocal. This is true already for an elastic composite, and the role of nonlocal interaction is further emphasized by inelastic processes, especially after the onset of localization. The reason is that the characteristic wave length of the deformation field decreases and becomes closer to the internal length of the material, related to the size and spacing of major inhomogeneities.

Even though the idea of a nonlocal continuum has a much longer history, nonlocal material models of the integral type were first exploited as localization limiters in the 1980s. After some preliminary formulations using the concept of an imbricate continuum [35], the nonlocal damage theory emerged [36]. Nonlocal formulations were then developed for a number of constitutive theories, including softening plasticity, smeared cracking, microplane models, etc. The basic concept seems to be sound, but the details of the formulation for a given material model are still to a large extent ambiguous. Formulations applying nonlocal averaging to different variables often give similar results at the onset of localization but their behaviors at later stages of the deformation process may be dramatically different and may exhibit some pathologies [37]. There are also controversies regarding the thermodynamic foundations of nonlocal modeling [38, 39]. From the practical point of view, the most serious deficiency is that no sufficiently general nonlocal model for complex materials such as concrete seems to be available. Existing formulations usually yield satisfactory results for a specific class of failure mechanisms (e.g., for tensile failure), but they are hard to extend to arbitrary loading scenarios. It is even questionable whether the general case can be covered using an isotropic nonlocal averaging scheme with a single characteristic length [40].

4.2 Adaptive analysis

Fully regularized models are in general computationally expensive, but they provide many improvements compared to partially regularized models based on the fracture energy concept. When a sufficiently fine mesh is used, the mesh-induced directional bias are either completely removed, or at least substantially alleviated. Owing to the continuous spatial distribution of the internal variables, they are ideally suited for mesh-adaptive simulations [41]. To illustrate that, the methodology developed by the present authors will be briefly described. It is based on the Zienkiewicz-Zhu error estimator [42] for regions that remain elastic, and a damage-based error indicator for regions of extensive cracking.

The proposed approach is truly adaptive, with mapping of displacements and internal variables, which allows to continue the analysis from the currently reached state, instead of restarting the analysis from the very beginning after the mesh refinement. The basic blocks of the adaptive procedure include the error estimator and indicator, remeshing criteria, primary unknown...
mapping algorithms, internal variable mapping algorithms, and mesh generator interface. The corresponding abstract interfaces have been designed, allowing to develop, use and combine different algorithms [43]. The error estimator/indicator is invoked at the end of each loading step to evaluate the quality of the solution. The remeshing criteria use the information about the error distribution and determine the further strategy. If an acceptable error level is not exceeded, the analysis continues on the current mesh. In the opposite case, the required mesh density is determined and the adaptive remeshing process is activated. The purpose of the mesh generator interface is to invoke the corresponding mesh generator, which produces the new discretization based on the problem geometry, boundary conditions and required mesh density. After a new discretization has been generated, the corresponding domain representation is read and the transfer of displacement and internal variables from the old to the new mesh is performed. The mapping of primary unknowns (displacements) is done first, usually using the shape function projections, and then the transfer of necessary internal variables is performed. After the internal history has been mapped, it is used together with the strain vector computed from the mapped displacements to update the internal state of each new integration point (to achieve local consistency). When the transfer is finished, the old discretization is deleted and an equilibrium iteration process brings the mapped configuration into global equilibrium. Afterwards, the solution continues with the next load increment.

The following example illustrates the application of the developed strategy to the analysis of the four-point shear test described in Section 3. The specimen geometry is reproduced in Fig. 4a. The constitutive model used is a nonlocal version of the anisotropic damage model mentioned in Section 3. The mesh size in the elastic regions is controled by the Zienkiewicz-Zhu error estimator (error less than 15% required), and the maximum principal value of the second-order damage tensor is used as the error indicator for inelastic regions. A very simple remeshing criterion for the inelastic region is adopted, using a prescribed constant mesh density in the regions where the damage indicator exceeds 0.1. The mesh size prescribed in these regions is a fraction of the nonlocal interaction radius, to make sure that the nonlocal interaction between individual Gauss points is properly activated.

During the adaptive simulation, the initial mesh with 145 nodes and 234 elements (Fig. 4b) is gradually transformed into the final mesh with 4174 nodes and 8120 elements (Fig. 4c). The damage distribution at complete failure, shown in Fig. 4d, nicely corresponds to the experimental results. The damage evolution is shown on the adaptive discretization in Fig. 5.

Adaptive analysis of the four-point shear test has also been performed using a nonlocal formulation of the simple isotropic damage model with a Rankine-like equivalent strain measure. The damage distribution presented in Fig. 6 shows that the final process zone is closer to the experimentally observed crack trajectory than the crack band obtained with a local version of the model on a fixed mesh (Fig. 3). Nevertheless, the simulated trajectory is almost straight and the results are clearly worse than with the anisotropic model.
Figure 4: Adaptive analysis of a four-point shear test using the nonlocal anisotropic damage model: a) geometry and loading, b) initial mesh, c) final mesh, d) final damage
Figure 5: Adaptive analysis of a four-point shear test using the nonlocal anisotropic damage model: evolution of the damage pattern and of the mesh around the process zone.
4.3 Process zone resolution by extended finite elements

The previous subsection presented an adaptive approach based on $h$-refinement, i.e., on the adjustment of the element size, keeping the order of the elements constant (and usually low). However, low-order elements produce stress oscillations due to the mismatch between the interpolation of local and nonlocal quantities [44]. As shown in [1], $p$-adaptive methods based on an increase of the polynomial order of finite element shape functions lead to only a partial improvement. The undesired stress oscillations are usually reduced around the center of the process zone but they remain appreciable around the boundary of the process zone surrounded by material that experiences unloading. This is caused by the poor ability of polynomial approximations to capture the non-smooth transition between the unloading region with almost constant strain and the softening process zone with fast strain increase in the direction perpendicular to the boundary. Consequently, there is a strong need for an innovative adaptive procedure that goes beyond the commonly used $h$ or $p$-adaptivity. A tempting idea is to use the extended finite element method, already mentioned in Section 2 in the context of discontinuous displacement approximations.
For regularized models, the enrichments can be based on a smooth function gradually growing from 0 to 1, which represents a regularized counterpart of the Heaviside function. Some preliminary results are reported in [45]. To illustrate the potential of this promising approach, a three-point bending test is simulated using the nonlocal isotropic damage model with a Rankine-like equivalent strain measure. The geometry of the specimen is shown in Fig. 7a. The propagating crack geometry is assumed as a growing straight segment placed on the axis of symmetry of the specimen. The basic mesh used in this example is shown in Fig. 7b. It consists of 67 nodes and 92 constant-strain elements. During the adaptive analysis, only 6 additional generalized displacement degrees of freedom are introduced.

Fig. 8 compares the resulting load-displacement diagram with the diagram obtained using the standard finite element interpolation on a fine mesh, containing 976 nodes and 1868 constant-strain elements (Fig. 7c). The agreement is surprisingly good, given that the number of degrees of freedom used by the extended finite element model is very low. The main advantage of this technique is its ability to reasonably capture the evolution of the fracture process zone even on very coarse meshes. This can be illustrated using the obtained profiles of local strain (Fig. 9a) and damage (Fig. 9b).
Figure 8: Three-point bending test: load-displacement diagrams produced by extended finite elements on a coarse mesh and by standard finite elements on a fine mesh.

Figure 9: Three-point bending test: a) strain profile, b) damage profile.
Acknowledgment

This work has been supported by the Swiss Commission for Technology and Innovation (CTI) under project No. 4424.1.

References


