Mechanics of distributed cracking

Zdeněk P Bažant

Professor of Civil Engineering and Director, Center for Concrete and Geomaterials, Northwestern University, Evanston IL 60201

This paper reviews some interesting recent results on mesh sensitivity and strain localization instability, nonlocal and micromechanics models, and the size effect in failure due to distributed cracking.

As we know, failure of many materials involves propagation of a large system of densely distributed cracks rather than a single, precisely defined fracture. This is typical of concrete and rock, as well as stiff clays, ice, especially sea ice, filled asphalt and polymer concretes, and refractory concretes. The number of cracks or microcracks is extremely large, and their locations and orientations are random. Therefore, it is inevitable to treat the densely cracked material as a continuum. The constitutive relation must then exhibit strain-softening, a phenomenon where the matrix of tangential elastic moduli ceases to be positive definite. Unreasonable though such an approach might seem (84.40), it has nevertheless been proven useful provided that certain mathematical difficulties inherent to it are properly tackled.

Some of these difficulties were already pointed out in 1903 by Hadamard (03.1) and others were further analyzed by Thomas, Maier, Mróz, Bažant, Buri and Dougill, Sandler and Wright, Bažant and Belytschko, Darvall, Read and Hegemier, Wu and Freund, and others (61.4, 76.2, 77.4, 84.40, 84.46, 84.58). Recently, due to large errors in finite element predictions of complete failures of concrete structures or rock masses subjected to blast, impact, earthquake, thermal loading, creep or shrinkage, the modeling of distributed cracking came to the center of attention. The literature has become so extensive that a complete coverage in one paper is impossible. Therefore, I will focus my review mainly on the work done at our Center at Northwestern University which was carried out jointly with T B Belytschko, L Cedolin, T P Chang, P Gambarova, J K Kim, F B Lin, B H Oh, P Pfeiffer, P Prat, and A Zubelewicz, under the sponsorship of the Air Force Office of Scientific Research (Grant No 85-0009) and the National Science Foundation (Grants Nos CEE800-3148 and CEE821-1642). If I fail to mention some important contributions, their authors should realize that they will be in a good company.

Let me begin by showing a micrograph of the fracture tip region in Portland cement mortar, obtained by means of scanning electron microscopy by Mindess and Diamond (80.22). Instead of a continuous crack we see many discontinuous microcracks which are not arranged along one line but are also spaced laterally (Fig. 1). On a much larger scale, Fig. 2 shows typical cracking patterns in reinforced concrete structures, as observed on punching shear specimens tested at Northwestern University (85.6). On a still much larger scale, Fig. 3 shows a typical map of cracking caused by advancing a stope through a bushveld norite rock formation in a very deep mine, as determined by drill sampling techniques by Brumer and Rorke (84.14).

How should we model such distributed and microscopically chaotic crack systems in finite element analysis?

I. STRAIN-SOFTENING CRACK BAND AND LINE CRACK MODELS

While mathematicians may prefer to seek good solutions to simplified problems, we engineers must seek simplified solutions to good problems. So, rather than limiting attention to problems with sharp line cracks, for which a good solution could be readily found, we must look for a simplified solution to the actual problem which takes the disperse nature of cracking into account. A useful simplifying idea, introduced by Rashid (68.9) 18 years ago, is to imagine continuously distributed or smeared cracks and describe them as strain-softening (81.24). This is easily taken into account by reducing the elastic modulus in the direction normal to the cracks and introducing an orthotropic matrix of elastic moduli along with an inelastic stress decrement. The stress decrease due to cracking was initially considered as a sudden, i.e., vertical crack drop, but it is more correct to describe it as gradual strain-softening.

Strain-softening does not exist in the heterogeneous microstructure at sufficient resolution (84.40). It is merely an expedient macroscopic model accomplishing homogenization. Its mathematical formulation has recently been the subject of very interesting debates (84.40, 84.46). Some investigators, influenced by plasticity theory, have promulgated the view that strain-softening is a nonexistent and mathematically unacceptable property for a continuum. In my opinion they overstate, although they do describe certain difficulties correctly. Before I discuss them, I would like to point out some other branches of physics in which a similar problem arises. In the physics of gases, the liquid-vapor phase transition (83.1, 84.1) described by van der Waals’ equation of state (64.1, 80.18) involves a region where pressure decreases at decreasing volume. This behavior, which we might classify as strain-softening, characterizes metastable superheated liquids or supersaturated vapors. In astrophysics, a continuum equation of state of stellar matter which
exhibits what we would call volumetric strain-softening is a well-established concept, treated by field equations which differ from those for the classical local continuum (58.2, 64.3, 65.2, 73.11, 82.29). This concept explains the gravitational collapse of stars (e.g., the collapse of a white dwarf into a neutron star or the gravothermal collapse). It is, I think, more fruitful for us to accept the concept despite all its peculiar complications, and then attempt to deal with them.

One difficulty is that, in the strain-softening regime, the material can no longer propagate waves since, as Hadamard recognized (03.1), the wave speed ceases to be real and becomes imaginary (Fig. 4). However, this is not as problematic as thought until recently because the unloading tangential modulus remains positive even in the strain-softening range (76.2, 84.42, 84.41). By virtue of this fact, a solution to some wave propagation problems that involve strain-softening can be found, as I will show later.

First let me discuss the problems which arise in finite element analysis. Here the chief difficulty caused by strain-softening is the spurious mesh sensitivity. After large finite element programs with strain-softening models for distributed cracking had been written and widely applied, it was discovered that the convergence properties are incorrect and the calculation results are unobjective with regard to the analyst’s choice of the mesh (76.2, 79.1, 80.2, 83.7, 85.19, 85.45). In Fig. 5 we see some numerical results for a rectangular panel in tension obtained with three different meshes the sizes of which are in the ratio...
4:2:1. We assume that a starter notch exists in the center of the panel, and we plot the value of the load necessary to further extend the crack band, as a function of the crack band length. If the strength criterion is used to decide whether the crack band extends into the next finite element, the results of the three meshes are as different as the solid curves in Fig. 5.

Pathological behavior of this kind, by now confirmed by many investigators, can be remedied by adopting for crack band propagation some form of an energy criterion, just as in classical fracture mechanics of sharp line cracks. With the energy criterion, we obtain in Fig. 5 the dashed curves (80.2) which differ from each other only by a small numerical error. Similarly (Fig. 6), the mesh size is found to affect incorrectly the load-deflection diagrams, for example, those for a rectangular panel (Fig. 6) in which symmetric crack bands start from the sides. Such spurious mesh sensitivity can occur not only for the post-peak response but also for the value of the maximum load, as documented, e.g., for the finite element analysis of diagonal shear failure of beams (Fig. 7) (80.2, 83.7). The lack of objectivity or spurious mesh sensitivity is found both for the sudden (vertical) and the gradual stress drops (83.9).

The energy criterion, necessary to avoid the mathematical problem of inobjectivity or spurious mesh sensitivity, dictates that the energy dissipation due to localized cracking per unit length of fracture must be a material property which cannot depend on the chosen mesh size. The mesh-independence of energy dissipation can be achieved in finite element modeling in various ways, both for the sudden and the gradual stress drops. The essential physical aspect is that the distributed nature of cracking, which is in fact a consequence of material inhomogeneity (characterized by aggregate or grain size, reinforcement spacing, rock joint spacing, etc), blunts the front of localized fracture. This blunting, similar to that in ductile fracture of metals (Fig. 8), causes the fact that the stress decline along the line of fracture extension is gradual and occurs over a finite length which is approximately a certain constant multiple of the aggregate size. This length can approach the structure dimensions if the structure is not very large relative to the material inhomogeneity size.

One way to satisfy, for localized cracking, the mathematical requirement for a mesh-size independent energy dissipation, and capture at the same time the physical fact of fracture front blunting, is offered by the crack band model. As proposed in 1974, the cracking is assumed in this model to be uniformly smeared across a band having at its front a finite width \( w_e \) which is a material property (76.2). This is modeled by requiring the element size to be equal to \( w_e \) (Fig. 9). Fracture energy \( G_f \) is then given as \( w_e \) times the area under the uniaxial complete tensile stress-strain diagram, or more precisely the area under the loading and unloading diagrams emanating from the peak stress point (82.3). As the simplest form, the stress-strain diagram may be considered as bilinear, although a
diagram with a curved convex strain-softening branch and a very long tail is usually more realistic (85.19).

The triaxial form of the strain-softening constitutive relation is most easily obtained by a secant compliance formulation referred to the crack axis, such that the secant diagonal stiffness term corresponding to the direction normal to the crack is divided by the cracked area fraction $\mu$ (a damage-type parameter) which decreases to zero as the transverse normal strain increases (Fig. 10). This model (82.3, 84.12) can closely describe essentially all known fracture test data for concrete as well as rock (67.3, 68.3, 69.2, 70.6, 71.3, 71.6, 71.11, 72.1, 73.3, 73.4, 75.3, 76.7, 76.10, 77.8, 78.9, 78.13, 79.16, 80.22, 80.27, 80.28, 81.4, 81.11, 81.17, 81.23, 82.9, 82.28, 82.31, 82.32, 83.29, 83.42, 83.46, 84.30, 84.52, 85.21, 85.30, 85.31, 85.39, 85.44); see, e.g., Figs. 11 and 12 in which the optimum possible fits by linear elastic fracture mechanics are shown as the dashed lines and those for the crack band model as the solid lines. Figure 11 illustrates how the maximum load of fracture specimen depends on its size and notch length, while Fig. 12 shows how the specific energy required for crack growth depends on the length of the crack extension from the notch.

Aside from having to specify the shape of the strain-softening branch as well as the overall type of the strain-softening triaxial constitutive relation, we can characterize the crack band model by three material parameters: the tensile strength $f'_t$, the fracture energy $G_f$, and the cracking front width $w_c$. The characteristic slope $E_c (\leq 0)$ of the strain-softening diagram is a function of these parameters. $f'_t$ and $G_f$ being constant, the cracking front width $w_c$, which represents a characteristic length of the material, has an almost negligible influence on the capability to fit test data for localized fracture, and cannot be determined from them. It can be determined, however, if the strain-softening modulus $E_c$ is specified independently, for example, on the basis of tests in which cracking does not localize. Such situations of nonlocalized cracking prevail for direct tension or bending of a bar or beam with sufficient reinforcement, or for cooling or shrinkage cracking near the surface of a massive specimen. Absence of localization of cracking can perhaps be also assumed for unreinforced tension specimens that are sufficiently short. Based on $E_c$, the cracking front width $w_c$ is found to be approximately three times the maximum aggregate size for concrete (83.9). However, multiples from 1 to 10 make little practical difference if nonlocalized cracking is left out of consideration and only the data for localized cracking are analyzed.

For the majority of practical applications, an element size equal to several aggregate sizes is impracticably as well as unnecessarily small. Essentially identical results can be obtained with larger finite elements if the mesh is not too crude compared to the structure dimensions. For this purpose, it is necessary to adjust the average stress-strain relation for the

**FIG. 10.** Triaxial strain-softening law (83.9).

**FIG. 11.** Illustration of how the maximum load of fracture specimen depends on its size: ( ) nonlinear theory; (---) linear theory; (O) $d = 0$ in.; (■) $d = 10$ in. (test data by Walsh (72.9)).
finite element so as to ensure the same energy dissipation, ie, the same fracture energy $G_f$, for any element size $h$. We may imagine a band of width $w_0 = l$ to be embedded within the finite element (Fig. 13), and its relative normal displacement to be added to that corresponding to an elastic or elastoplastic (but nonsoftening) behavior. The average stress–displacement relations for the finite element then vary with the element size as shown in Fig. 13; the stress drop becomes progressively steeper until, for a certain characteristic element size $l_0$, it becomes vertical (85.3, 85.13). With larger element sizes, which are often needed in practice for the analysis of large structures, the average stress–displacement diagram is of the snap-back type, which is unstable under displacement control. Such a situation requires dynamic analysis; however, an approximate static analysis becomes possible by replacing the snap-back diagram by a diagram with a vertical stress drop which dissipates the same amount of energy (85.3, 85.13). Under this restriction, the dynamic snap-through (ie, the vertical stress drop in Fig. 13) begins and ends with states of equal kinetic energy, thus permitting static analysis. It is found that for finite elements larger than $l_0$, the strength limit for the sudden vertical stress drop must be decreased as $h^{-1/2}$ (79.1, 80.1, 83.7, 85.3, 85.4, 85.12, 85.13). The finite element results are then essentially independent of the mesh size, unless the mesh is too crude.

The constancy of the crack front width $w_0$ in the crack band model is certainly a simplifying assumption. While $w_0$ cannot depend on the mesh size, it may vary depending on the stress and strain fields surrounding the fracture process zone. The

FIG. 13. Fully localized softening or distributed cracking (85.3, 85.13).
effective crack front width probably increases as fracture extends from a sharp notch. This behavior can be described by the crack layer model of Chudnovsky (83.16, 82.11), in which the width of the crack layer (synonymous to crack band) depends on three path-independent integrals around the cracking front: the J-integral giving the energy influx, and the \( M \)- and \( L \)-integrals giving the energy changes due to a rotation of the cracking front and a lateral expansion of the crack layer. Some applications of this approach and its systematic thermodynamic formulation have already been made in the area of polymer crazing. A general analysis of continuum damage based on these integrals was made by G. Herrmann and J. R. Rice (84.44). No doubt, applications to concrete and geomaterials would be possible.

For the modeling of localized fracture, the essential property of the crack band model is not that its cracking front width \( w \) is finite, but that its finiteness forces the length of the fracture process zone in the direction of the crack axis (i.e., the length of the strain-softening zone) to be finite and large. This same essential property may be alternatively obtained with various types of line crack models (Fig. 14), patterned after the original ideas of Barrenblatt (59.1, 62.1) and Dugdale (60.1), which were extensively applied to ductile fracture of metals (73.6, 74.5, 74.9, 77.10) and were adapted for concrete in the pioneering works of Hillerborg et al (76.8, 81.17, 81.21, 83.31, 84.27, 85.31).

In Hillerborg’s model (as well as some similar preceding finite element models for metals), a sharp line crack is assumed to open up along the interelement boundary on which a stress–displacement relationship is specified as a material property. This is, nevertheless, approximately equivalent to the crack band model if the energy dissipation per unit length of fracture is the same, or if the transverse normal displacement on the line crack is equal to the accumulated transverse normal strain due to cracking over the width of the crack band. The equivalence of these two types of models can in fact be rigorously established by applying the \( J \)-integral, as Rice pointed out (85.41).

Another variant of the line crack approach, which was initially proposed for localization of strain-softening in shear (81.18), and is for concrete known as the composite damage (or composite fracture) model (84.53, 85.54), is to embed a line crack or a narrow crack strip inside the finite element (Fig. 14), superposing its relative displacement on the elastic deformation of the element. For the situations of localized fracture, e.g., the fracture specimens, this approach yields again essentially the same results as the fictitious crack model of Hillerborg et al, as well as the crack band model.

The models based on the line crack idea, however, are not equivalent to the crack band model for situations of nonlocalized distributed cracking (85.3, 85.13). This is due to the fact that they are characterized by only two material parameters, strength \( f' \) and fracture energy \( G_c \) (if the shape of the strain-softening diagram is given in advance), while the crack band model involves a third material parameter, \( w \), the characteristic length of the material. In this regard it is worthwhile to recall what has been learned in stability studies of crack systems.

### 2. STRAIN-LOCALIZATION INSTABILITY

An interesting parallel is provided by the propagation of a system of straight parallel equidistant cracks into a brittle elastic homogeneous half-space. Keer, Nemat-Nasser, and Ohshubo worked with me at Northwestern on stability of this system, and our work has been extended by Cleary at MIT (77.2, 78.11, 79.3, 79.5, 80.30, 83.17). The cracks are driven by cooling or drying, and the cooling stress profile can be either gradual, as obtained from heat conduction, or may have an abrupt temperature drop, as obtained when water circulates through the cracks (the motivation for this research was a dry rock geothermal energy scheme involving rock fracturing). The stability condition may be reduced to the condition that the second-order variation of the free energy of the system as a function of the lengths of the individual cracks must be positive definite with the exclusion of certain instability modes that are inadmissible, such as crack shortening for a nonzero stress intensity factor (Fig. 15). From this condition it is found that the parallel crack system with equal crack lengths may be stable or unstable; stable if the cooling temperature profile is uniform for a certain length and then has a sufficiently rapid drop, and unstable if the cooling profile is gradual, such as the profile obtained from heat conduction without convection in the cracks.

If the system is unstable, every other crack during the cooling process stops growing when its length becomes approximately equal to the spacing of the leading cracks (77.2, 78.11, 79.3). In this manner the fracture permanently localizes into the leading cracks (Fig. 15). On the other hand, if the cracks are stable, they can be driven into the half-space at constant spacing as deep as desired, and remain stable with a spacing as close as desired (79.5). Thus, from the macroscopic sense, i.e., over a volume involving many crack tips, the behavior at the cracking front is essentially the same as macroscopic
strain-softening. This softening, or distributed cracking, is stable over a cracking front which can be arbitrarily wide.

Similarly, and not surprisingly in view of the observations of reinforced concrete beams, a system of densely distributed parallel cracks in an elastic homogeneous brittle solid may be stable if there is a compression zone closely ahead of the crack tips, as found for instance for the cracking caused by tunnel excavation in a grouted soil (Fig. 16) or the cracking caused by bending in a beam, which contains sufficient reinforcement (80.8) (Fig. 17). (For bending cracks, the sufficient reinforcement percentage calculated from stability conditions is surprisingly small and roughly equal to the empirical minimum reinforcement prescribed in building codes.) The stabilizing influence of reinforcing bars provides an a posteriori justification to many of the finite element analyses of reinforced concrete structures in which the questions of possible localization instability of cracking and the associated spurious mesh sensitivity were ignored. It is the reason why many practical structures are fracture-insensitive and can be analyzed without paying attention to the problems we are now discussing.

In dynamics, another influence which tends to temporarily stabilize a nonlocalized cracking front is the inertia force. In dynamic finite element analysis, especially for loads of extremely short durations such as blast or impact, the times of response are not long enough for the cracking localization to take place, and nonlocalized distributed cracking is then the correct solution for many problems; see, eg, Fig. 18 which shows the cracking sequence for a prestressed concrete reactor vessel subjected to an explosive loading in the cavity, as calculated by Marchertas, Belytschko, and Bažant (76.13).

The crack band model has the advantage that it can deal with both localized and nonlocalized cracking, and it is no doubt for this reason that it is being widely applied in general finite element programs (Rots, de Borst, Kusters, Darwin, Pan, Marchertas, et al., 84.45, 84.18, 84.21, 84.20, 84.19, 84.41, 85.19, 85.45, 85.46, 85.20, 86.6, 85.44, 84.47, 84.15, 83.44, 83.30). In the case of nonlocalized cracking, a zone of many neighboring finite elements is simply allowed to exhibit smeared cracking, as practiced since the inception of this approach. However, if the models based on the line crack idea, eg, those of Hillerborg (76.8) or William and Sture (84.54), are used in a situation of nonlocalized cracking, the results are found to strongly depend on the element size, ie, are unobjective with regard to the analyst's choice of the mesh. This is illustrated by the example (84.15, 85.3, 85.13) of a concrete bar (Fig. 19) with 5% reinforcement, in which smeared cracking is stable without localizing. The finite element results obtained for three different mesh sizes differ from each other intolerably (Fig. 19) for the models based on a line crack (ie, the fictitious crack model, or the composite damage or composite fracture models), but are identical for the crack band model, which is in this case equivalent to the classical smeared cracking approach.
To sum up, the picture of objectivity at mesh refinement for localized and nonlocalized cracking situations is as shown in the table in Fig. 20. The crack band model appears to be objective in general, i.e., when one does not know in advance whether distributed cracking will or will not localize.

The limitation of the line crack models to localized cracking is due to the fact that they are characterized by only two material parameters, $f'$ and $G_s$, whereas the crack band model has three parameters, $f'$, $G_s$, and the characteristic length $l = w_c$. For parallel cracking, the line crack models can, of course, be easily modified so as to be applicable for both localized and nonlocalized cracking; in the fictitious crack model (i.e., interelement sharp cracks), it suffices to treat the element size $h$ as a fixed material property so as to enforce the required crack spacing, $h = w_c$, and in the composite damage or composite fracture models, to imagine that there is not just a single line crack or crack strip embedded within each finite element, but that there are as many line cracks or strips as needed, spaced at distance $w_c$, which is a material property (Fig. 17). Plausible though this generalization of the existing line crack models might be, it nevertheless runs into trouble for nonparallel cracking, simply because nonparallel straight-line cracks cannot have a constant spacing. On the other hand, if cracking is characterized by a multiaxial strain-softening constitutive law, there is no particular difficulty when the cracking direction varies over the space, or even when multidirectional cracking takes place, or when the direction of cracking rotates during the loading process.

Thus far we have seen how both the localized and the nonlocalized distributed cracking can be modeled. However, we have not yet addressed the question whether the cracking will or will not localize. This is a difficult problem whose handling in large finite elements programs has not yet been developed. Aside from the stability analysis of the regular parallel crack system mentioned before (Fig. 15), an exact analytical solution is possible for uniaxial tensile or compression strain-softening in a bar loaded under displacement control through a spring (76.2, 82.3) (Fig. 21). This simulates a compression or tensile specimen in a testing machine, whose flexibility is represented by the spring.

Stability requires that strain increments that are positive over some portion of the specimen length and negative over another portion (Fig. 21) lead to a positive second variation of the system's energy at fixed supports. If strain-softening is assumed to occur only over a negligible portion of the length, which would be correct for a perfectly homogeneous continuum (without inhomogeneous microstructure), it is found that instability always occurs at the peak stress point. Thus, the strain-softening could never be realized. The fact that it can be realized and measured proves that the existence of strain-softening within only a negligible portion of the length is an incorrect assumption. Correctly, the length of the strain-softening region must be assumed to be a material property, or else objectivity could not be achieved, and then the strain-localization instability occurs at a finite distance behind the peak rather than at the peak. The shorter the specimen, or the stiffer the machine, the farther is the instability point pushed beyond the peak stress point. The stability condition can be reduced to the requirement that, for an axial load applied at the boundary of the strain-softening region, the sum of the spring constant of the strain-softening region and the spring constant of the rest of the system including the unloading portion of the specimen and the supports must be positive. This appears to be a special case of the general requirement that the incremental stiffness matrix of the system be positive definite.

This type of stability condition must be met for all admissible (i.e., compatible) deformation increments of the system at which the prescribed boundary displacements are fixed and the prescribed loads are constant. Checking it, however, may be difficult. For a larger finite element system, the number of all possible deformation increments is extremely large. In theory, all possible combinations of loading and unloading in various elements that have so far entered the strain-softening regime must be checked for the positive definiteness of the associated

<table>
<thead>
<tr>
<th>Model</th>
<th>Localized cracking</th>
<th>Nonlocalized cracking</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Classical smeared cracking</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>2. Fictitious crack model</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>3. Composite damage model</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>4. Crack band model</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>
incremental stiffness matrix. Such checks are not normally carried out in the present finite element practice. If the instability for a certain mode is not checked, a possible instability may remain undetected. The iterations of the loading steps in the finite element program may converge and leave the unsuspecting analyst happy that he has obtained the solution, while in fact his solution cannot exist in reality since another solution releases more energy. The situation is not the same as in buckling of elastic systems in which the energy is compared only for infinitely close adjacent equilibrium configurations.

Due to loading-unloading transitions in which the tangential stiffness changes discontinuously, there is no continuity between various deformation increments with various combinations of loading and unloading, ie, the unstable path is not infinitely close to another stable path. For this reason, no trouble is encountered in computations unless the right instability mode is actually tried.

As an illustration, Fig. 22 shows various possible combinations of unloading and softening in various elements which ought to be checked. Figure 23 shows an example of a doubly symmetric rectangular panel in which two symmetrically located elements at each side have a slightly smaller strength and start softening simultaneously. One possible equilibrium extension of the cracking (or softening) zone is a band of two cracking elements next to each other. If one does not deliberately perturb the system, one remains unaware that another possible equilibrium solution is an asymmetric crack band of a single-element width. For each of these two loading paths the computer program works well and iterations in each loading step converge. However, only one of these two equilibrium solutions is correct; it is the asymmetric solution with the band of a single element width.

A somewhat simpler criterion for detecting localization instability may be stated by comparing all the possible equilibrium paths for the same loading step, characterized by the prescribed displacement increments and load increments. These paths are much less numerous than all the equilibrium and nonequilibrium kinematically admissible increments at fixed loading. It appears that among all possible equilibrium paths of the system, with various combinations of loading and unloading in the strain-softening elements, the actual path is that for which the second-order free energy release of the system, as indicated in Fig. 24, is maximized.

In general, stability checks with the foregoing criteria are too numerous to be feasible. It will no doubt be necessary to develop some practical ways to correctly select only a few equilibrium paths for which the stability checks are to be made.

A remark is in order on the recently popular theory of continuous damage mechanics, which originated with the work of Kachanov (58.3). Numerous papers have been written on stress-strain relations for damage resulting in failure of the material; yet the stability problems inherent to these formulations have been ignored, as if by calling strain-softening the damage the stability problems would go away. They do not. If the damage stress-strain relation results in an incremental material stiffness matrix that is not positive definite, the problems with unstable strain localization and spurious mesh sensitivity are again encountered. These problems went mostly undetected since little finite element computation has so far been done with the continuous damage mechanics formulation, and when it was done mesh sensitivity has not normally been checked. It has been also thought that stability problems are avoided by rate-dependent damage laws (84.46). However, if the matrix of the effective stiffnesses for the time interval under consideration is not positive definite, the stability problems must still be tackled (86.5). Valuable though many continuous damage theory developments were, the systematic lack of attention to the material instabilities due to damage—a more important problem—in my opinion—has been a flaw of this school of thought.
3. IMBRICATE NONLOCAL CONTINUUM

From the purely mathematical viewpoint, the fracture models based on a line crack permit that the mesh size be reduced to zero. From the physical viewpoint, though, it makes no sense to reduce the mesh to sizes that are smaller than a certain limit, say, the aggregate size. In the same sense, the mesh size can be arbitrarily reduced also for the crack band approach, provided that the mean downward slope of the strain-softening diagram is made appropriately milder to ensure the same fracture energy for fully localized crack band. From the physical viewpoint, though, the finite elements in the crack band approach should not be made smaller than a certain characteristic length, \( l \), which may be assumed to be about three times the aggregate size for concrete.

Nevertheless, mesh refinement is needed for three reasons: (1) to have a convergent discretization, the limit of which must be a continuum, (2) to accurately determine the energy dissipation due to cracking, and (3) to capture variations of the effective width of the crack band. It is obvious that the density of cracks or microcracks will not be uniform across the fracture process zone but will increase gradually towards its center. The crack band model assumes a uniform distribution of the cracking strain, with a sudden jump at the boundary of the band, but no doubt the correct picture is a smooth, nonuniform variation of the cracking strain across a crack band whose width depends on the problem. These objectives may be achieved by adopting the concept of nonlocal continuum.

The nonlocal continuum is a classical concept introduced in the late 1960's by Kroner, Krumhansl, Kunin, Levin, Eringen, Edelen, and others (68.5, 68.6, 68.7, 71.8, 72.3, 83.24). The central hypothesis is that the stress at a point is not a function of the smoothed (homogenized) stress at the same point but a function of the (smoothed) strain distribution over a certain characteristic volume centered at that point (Fig. 25), the size of which is the characteristic length \( l \). Thus, one defines the mean strain, \( \bar{\varepsilon} \), as a certain averaging integral over the characteristic volume \( V \) with a suitable empirical weighting function \( \alpha(x) \) (Fig. 26). The constitutive relation then defines a broad-range stress \( \sigma \) as a function of the mean strain \( \bar{\varepsilon} \) (Fig. 26).

In the classical theory of nonlocal continuum, the continuum equation of motion (or of equilibrium) has been written in terms of the gradient of \( \sigma \). With regard to finite element analysis, however, this form of the continuum equation of motion is objectionable, because of two shortcomings: (1) The differential operator in the continuum equation of motion is not self-adjoint, with the result that the stiffness matrix of the finite element system is nonsymmetric even if the material stiffness matrix is symmetric (84.6, 85.5), and (2) for certain reasonable weighting functions \( \alpha \) in the averaging integral (and especially for the most useful uniform weighting function) the finite element formulation admits periodic zero energy solutions, with a wavelength equal to the characteristic length of the material (84.6, 85.12).

The latter shortcoming may be avoided by imagining an overlay of nonlocal and local continua, characterized by empirical coefficient \( c \) in Fig. 26. If \( c = 1 \), we have the usual or local continuum equation of motion, and, if \( c = 0 \), we have a purely nonlocal formulation. With \( c = 0 \), however, the aforementioned instability with a zero energy mode takes place. For a uniform weighting function this instability is avoided if \( c > 0 \). In theory, this would mean that an extremely small local component, e.g., \( c = 0.001 \), would suffice to avoid this instability; however, in practice one obtains excessive noise when \( c \) is close to the stability limit. Experience shows that \( c = 0.1 \) is usually the minimum needed to suppress this noise (84.6).

The problem with nonsymmetry of the operators and matrices may be avoided by applying the averaging operator \( H \) once more to the broad-range stress \( \bar{\varepsilon} \) before \( \sigma \) is substituted into the equation of motion (84.6, 85.5, 84.3) (see Fig. 26). It has been in fact rigorously demonstrated that the second application of the averaging operator \( H \) is required by a consistent use of the variational calculus (84.3).

For the purpose of some analytical solutions it is of interest to note that the integral averaging operator may be approximated by a differential operator involving a Laplacian \( \nabla^2 \) with a characteristic length \( \lambda \) (Fig. 26) related to the cross section \( l \) of the representative volume \( V \) (84.3).

When a uniform weighting function \( \alpha \) is used, a very simple discrete approximation is possible. This approximation consists of an imbricate (or regularly overlapping) finite element system as illustrated in Figs. 27 and 28 (84.3, 84.6, 85.12). When the finite elements are larger than the characteristic length \( l \), the imbricate and the usual finite element models are identical. For smaller finite elements, the element size must be kept constant, equal to \( l \), so that each finite element would automatically implement strain averaging over the length \( l \). Thus, the finite
elements must be regularly overlapping, or imbricated, and their thickness or cross section must be reduced so that the combined thickness of all overlapping finite elements remain 1. The discrete equations of motion for this finite element system represent difference equations, and it is easy to show that their continuum limit is exactly the continuum equation of motion in Fig. 26.

The imbricate finite element system is relatively easy to program. The only change that needs to be made in a usual, local finite element code is to redefine the integer matrix of nodal connectivity, which specifies the node numbers for each element number, and to program another integer matrix which gives the numbers of all finite elements that overlap each square of the mesh.

To validate a finite element program for strain softening, it is important to make a comparison with some exact solutions for strain-softening. An exact solution has been found to a wave propagation problem for a uniaxially stressed bar (85.5). The bar is subjected at its ends to sudden tensile stress jumps for strain-softening. An exact solution has been found to a wave propagation problem for a uniaxially stressed bar (85.5), the solution changes discontinuously when the limit of strain-softening is attained. Thus, the solution depends discontinuously on the boundary conditions. This means that the problem does not belong to the class of the so-called well-posed problems (as pointed out by Sandler, 84.46).

Nevertheless, a solution exists and is apparently unique (85.5). Moreover, a numerical finite element solution obtained with the standard explicit step-by-step algorithm is found to converge to the exact solution (Fig. 29), and to do so in fact quadratically (84.6). Clearly, the idea of strain-softening in a local continuum is not mathematically meaningless. However, it is physically unrealistic. According to the exact solution, the energy dissipation due to strain-softening which leads to breaking of the bar in the middle is zero (84.6, 76.2, 80.1). This cannot be true for any real material. We thus have another indication that a nonlocal continuum concept is required if the mesh is refined.

Figure 30 shows the profiles of the mean strain at various times obtained for progressively finer element subdivisions. The first subdivision has only five elements over the bar length, in which case the solution is local because the characteristic length is assumed to be 1/5 of the bar length. As the mesh is refined the width of the strain-softening region decreases, apparently converging to the Dirac delta function.

By contrast, for the imbricate finite element formulation (Fig. 31), the solution converges to strain distributions in which many elements next to each other undergo strain-softening (84.6). For the finest mesh, there are 195 elements over the bar, and about one-half of them are in the strain-softening regime. The energy consumed by failure of the bar is plotted in Fig. 32 as a function of the number of elements. While for the local finite element model the energy strongly depends on the mesh subdivision and apparently converges to zero, for the imbricate finite element solution, the energy consumed by failure of the bar remains approximately constant for various element subdivisions (Fig. 32). This behavior seems typical of all comparisons between local and imbricate finite element solutions for strain-softening.

In the foregoing solution, the strain softening is limited to an immobile point. However, as Belytschko et al just discovered, for problems with partial softening followed by hardening, strain-softening can develop in a finite region whose front can travel with a finite wave speed.

Another strain-softening problem, suggested by Belytschko as a one-dimensional problem which involves triaxial stresses, is
the imploding spherical wave (86.5). Figure 33 shows a hollow sphere whose volumetric stress-strain relation exhibits strain-softening. A stress jump applied at the sphere's surface produces an inward radial wave. An exact solution exists for elastic behavior, and it shows that the wave front grows. At a certain distance from the surface, depending on the magnitude of the suddenly applied stress, the strain-softening limit is inevitably reached. For elastic behavior, the explicit time-step finite element solution is found to converge to the exact solution. For strain-softening behavior, the local and imbricate nonlocal finite element solutions are shown in Fig. 33 for various element subdivisions, involving 10, 40, and 160 elements over the thickness. The characteristic length of the material is considered such that the local and nonlocal solutions are identical for a subdivision with 10 line elements of equal length. From the volumetric strain profiles at various times shown in the left column of Fig. 33, we see that the local solution does not converge. An interesting feature is that strain-softening does not arise merely at one point. The wave is able to penetrate beyond the point of first strain-softening and create further strain-softening points.

FIG. 30. Finite element solution of wave propagation in one-dimensional strain-softening bar, local continuum (84.6).

FIG. 31. Same as Fig. 30 but, imbricate continuum (84.6).
at apparently random locations which vary chaotically from one element subdivision to another. By contrast, the imbricate nonlocal solution, shown in the right column, appears to converge. Similar results are obtained for a cylindrical wave.

Figure 34 shows application of the imbricate analysis to a rectangular panel pulled vertically, with prescribed displacements of the boundary nodes on top and bottom (85.7). The properties of all finite elements are the same, except that in two elements at the centers of the sides the strength limit is assumed to be 2% smaller than in the other elements. Three meshes, with sizes in the ratio 1:1/3:1/5, are considered. The problem is static and an implicit step-by-step algorithm is used. Figure 35 shows the solutions obtained for various meshes, the local ones in the left column and the imbricate nonlocal ones in the right column. The load-displacement diagrams (top), as well as the transverse normal strain profiles across a strain-softening band growing horizontally across the panel, are rather different for the three meshes and the strain is seen to localize. By contrast, the imbricate nonlocal solutions are almost the same for the three meshes, and no strain localization in the middle of the panel occurs. The dependence of the energy consumed by strain-softening on the number of elements (middle) is similar as before; for the local solution, the dissipated energy value depends strongly on the number of elements and apparently converges to zero for an infinite mesh refinement, while for the imbricate nonlocal solution the dissipated energy is about the same for all mesh subdivisions.

4. MICROMECHANICS MODELING

Although a realistic stress–strain relation with strain-softening is not all that is required for finite element analysis of distributed cracking by the crack band approach, it is an essential ingredient. Numerous researchers, including Dougill, Bažant, Prat, Kim, Darwin, Pecknold, Murray, Darve, Zienkiewicz, Dafalias, Mroz, Hueckel, Maier, Dragon, Zubelewicz, and others have tackled this question from the macroscopic phenomenologic viewpoint and produced various sophisticated constitutive models with strain-softening (endochronic, plastic, fracturing, plastic-fracturing, bounding surface, etc) (71.10, 76.3, 76.5, 76.6, 76.9, 76.12, 77.1, 77.6, 77.11, 78.5, 78.7, 78.12, 79.8, 79.9, 79.12, 79.14, 79.18, 80.4, 80.6, 80.13, 80.14, 80.26, 81.10, 82.10, 82.16, 82.21, 82.22, 83.2, 83.18, 83.23, 83.25, 83.35, 84.35, 84.38, 84.40, 85.16, 85.47, 85.52). One promising approach is a damage constitutive law based on an idea of Kachanov (58.3),

**FIG. 32.** Energy consumed by failure of strain-softening bar (85.6).

**FIG. 33.** Spherical wave, in strain-softening material, strain profiles (86.5).
FIG. 34. Mesh refinement strain-softening zones for refined meshes (85.7).

FIG. 35. Convergence at mesh refinement for Fig. 34 (85.7).
at first used extensively for metals. In this approach the constitutive properties depend on a certain damage variable (scalar, or vector, or tensor) which is specified by an evolution law. For materials that exhibit distributed cracking, such damage formulations were developed by Mazars, Løland, Resende, and Martin, Krajcinovic, Fosnega, Popov, Ortiz, and others (77.9, 78.8, 80.19, 81.9, 81.12, 81.14, 82.21, 83.36, 84.33, 84.50, 82.17, 84.43, 84.34, 85.42, 83.39, 84.37). The latest Ortiz's formulation (84.37), in which the damage measure coincides with the fourth-order tensor of elastic moduli of the material, appears particularly realistic and promising.

Many of the damage constitutive laws involve some form of rate dependence, in which case the strain-softening may be disregarded as stress relaxation. It is unclear, though, whether the modeling of strain-softening as stress relaxation is more than a simplifying device suitable merely for a limited range of applications. As it seems, this type of formulation does not permit that strain-softening be obtained for various loading rates differing by orders of magnitude, as observed in testing of concrete. On the other hand, the strain rate in most materials affects the response as a parameter, and in this form the damage constitutive laws are no doubt useful for modeling strain-softening failures.

After less than complete satisfaction with macroscopic phenomenologic modeling, intensive efforts are now being made to deduce the constitutive relation from physics and micromechanics analysis. Such formulations, involving statistics of crack distributions, their sizes and openings, statistics of slips, thermodynamics of material cracking, etc., have been formulated by Kachanov, Vakulenko, Zaitsev, Wittmann, Roelfstra, Seaman, Shockey, Bui, Ehrlicher, Stout, Dienes, Horii, Nemat-Nasser, Davison, Chudnovsky, Ortiz, Margolin, Costin, and others (73.3, 75.8, 76.4, 78.6, 79.11, 79.17, 80.17, 80.21, 81.27, 82.15, 82.27, 83.19, 83.20, 83.21, 83.40, 83.49, 84.56, 85.28, 85.33, 85.48, 85.53, 86.8).

One simple and effective micromechanics approach to the stress-orientation relation is the microplane model. It is based on G I Taylor's basic idea (38.1), developed long ago in the slip theory of plasticity for polycrystalline metals by Batdorf and Budianski (49.1). To rocks and soils, these ideas were applied by Zienkiewicz and Pande (77.18, 82.16, 82.23, 82.24) in what they called the overlay model, and more recently to concrete by Băzant, Oh, and Gambarova (85.9, 84.8). The basic idea of these formulations is to characterize the inelastic behavior by a relation between the stress and strain components on a plane of an arbitrary inclination in the material, for which recently the general term “microplane” has been introduced. The stresses or strains from the microplanes of all orientations are suitably superimposed to obtain the macroscopic stress or strain tensor.

The macroscopic stiffness matrix is obtained from the condition that the rate of energy dissipation expressed in terms of the macroscopic and microcracking stresses and strains must be the same for any specified strain rate tensor. In formulations of this type it was normally assumed that the microstructure is statically constrained, which means that the stresses on each microplane are assumed to be the resolved components of the macroscopic stress tensor (49.1). Another equally simple possibility is to assume a kinematically constrained microstructure, in which the strains rather than the stresses on a microplane of any orientation are the resolved components of the macroscopic strain tensor. The latter, recent approach appears to be more realistic for concrete (85.9), for which the microplanes may be imagined to represent the thin contact layers between rigid aggregate particles. The deformation of these layers is essentially determined by the relative displacements of the aggregate particles, i.e., is kinematically constrained. Aside from that, it is found that the kinematic rather than static constraint is necessary to obtain a stable model for the case of strain-softening, and in the practical sense this is the overriding reason for introducing this approach.

The basic relations for a simple kinematically constrained microplane model in which only normal strains on the microplanes are permitted are written in Fig. 36; \( n_i \) are the direction cosines of the normal to each microplane, \( e_{ij} \) is the normal component, \( e_{ij} \) is the macroscopic strain tensor, and \( D \) is the macroscopic stiffness matrix obtained by integration over a unit hemisphere. Function \( f(\hat{n}) \) indicates the frequency of the microplanes for various orientations and is able to introduce anisotropic effects of the material. For isotropy, \( f(\hat{n}) = \text{const.} \) A more sophisticated model which takes into account not only the normal but also the shear components on each microplane has been formulated (84.13), but it is not clear at the present whether consideration of the shear strains is necessary since a shear strain on one plane is equivalent to normal strains on certain inclined planes. If the normal stress-strain relation for each microplane exhibits strain-softening, one obtains strain-softening also macroscopically, and it is easy to fit various tensile strain-softening data for concrete. It seems that the microplane model should be also capable of modeling compression splitting as well as the inelastic behavior of concrete under compression and shear in general, but work on this problem is still in progress (P C Pratt, Northwestern University).

Conceptually, the attractiveness of the microplane approach rests on the fact that we do not need to deal with relations between the tensors of stress and strain, worrying about their proper invariance. Rather, we need to describe only a simple relation between the components of stress and strain on one plane, and the tensorial invariance restrictions are then automatically satisfied by integration over all spatial directions.

With regard to application in finite element programs, it is important to integrate over a unit hemisphere (Fig. 36) as efficiently as possible, since this integration must be repeated at each integration point of each finite element, at each loading step. Various efficient Gaussian-type numerical integration formulas for the spherical surface are available in the literature (71.12, 85.9), and one of the most efficient ones is illustrated by the point arrangement on an icosahedron in Fig. 37, in which the points indicate the discrete spatial orientations used in approximating the integral over a unit hemisphere by a finite sum, with certain suitable weights for each direction (86.3).
error of the integration formula may be estimated by assuming a certain type of loading, eg, uniaxial tension, and applying it at various orientations with regard to the set of integration points. If the integration formula were exact, identical response curves of stress versus strain would be obtained. Due to the approximate nature of the numerical formula, the response curves have a certain spread which must be acceptable, both in the hardening and softening ranges (Fig. 37). Achieving a narrow spread of the response curves for the softening range is much harder than for the hardening range, and it necessitates in three dimensions the use of at least 21 integration points (85.9) (ie, discrete spatial directions). For two-dimensional problems, the number of discrete directions may be reduced to about eight (84.8).

In the microplane approach, we trade conceptual simplicity for a penalty in computer time. Compared to the phenomenologic macroscopic constitutive models, the basic relations are simpler and involve fewer material parameters. Aside from that, the microplane approach is more realistic, especially when complicated loading paths in the stress or strain space are considered (eg, shear and normal extension applied either proportionally or in sequence). The microplane model can capture the development of cracking in various directions at the same point of the structure during the loading sequence.

The macroscopic constitutive law does not provide complete information on the material properties. It is equally important to use micromechanics to determine material parameters that limit strain localization, such as the characteristic length \( l \) in the imbricate formulation or the cracking front width \( w_\text{e} \) in the crack band model. To obtain these parameters, the micromechanics model cannot be restricted to a single idealized point of the macroscopic homogenized (smoothed) continuum, but the spatial fields themselves must be directly modeled in terms of an idealized microstructure.

An example of this approach is the rigid particle contact model introduced for concrete by Zubelewicz (83.49) and subsequently developed in detail by Zubelewicz and Bažant (see Fig. 38). In this model, which is patterned after Cundall's distinct element method for frictional interaction of the grains of sand, gravel, and other bulk solids, one assumes that all inelastic deformation is concentrated into contact layers between rigid particles. In Fig. 39 this interaction is described simply as a sudden (vertical) stress drop when the specified strength limit is reached. The computer program generates a random system of rigid particles and their contact planes, and solves the problem in an incremental manner considering in each loading step the force and moment equilibrium conditions for each particle loaded by the interparticle contact forces. Figure 39 shows the calculated load-displacement curve and the curve of dissipated energy obtained for the example in Fig. 38, in which the top and bottom boundaries are subjected to prescribed uniform displacement. Evidently, the calculated tensile load displacement diagram seen in Fig. 39, with its very long tail of the strain-softening response, looks quite realistic.

Figures 40–42 illustrate the interparticle contact forces and the development of cracks. The heavy solid, dashed, and light solid lines indicate the lines of interparticle forces exceeding 0.8, 0.6 and 0.4 of the strength limit. Note the concentration of interparticle forces in the vicinity of the notch tip and the general disappearance of heavily loaded contacts at subsequent stages of softening, as the material outside the fracture process zone is getting unloaded. The cracks obviously propagate discontinuously, and not along the line of the symmetry of the panel. This is due partly to the randomness of the microstructure, and probably also to instability of the straight crack direction in double cantilever fracture specimens (cf. Cotterell and Rice, 80.11).
The latest stage of crack evolution is shown in Fig. 43, in which the approximate range on the cracked area is shaded. With regard to the crack band model, it is of interest to note that the mean width of the fracture process zone in this simulation is roughly three times the particle size. Hence, this model reveals the length restriction on the extent of localization of cracking, a characteristic which is impossible to obtain with pointwise micromechanics models, such as the microplane model. There is no doubt much further potential in the rigid particle contact approach. However, it would be difficult to apply this approach to real large structures because an increase in particle size would change the fracture energy, same as an increase in mesh size.

Other very valuable studies of the propagation of microcracks in a random microstructure of concrete have been made by Wittmann and Zaitsev (85.50, 85.53, 83.47, 82.33, 81.27, 81.26, 77.17). Finite elements have been applied to model the microstructure in the concrete-mortar system by Darwin (80.20), and recently Roelfstra and Wittmann use them in detailed modeling of the microscopic process of fracture as well as creep, water diffusion, and shrinkage in concrete (82.33, 85.11, 85.50).

The most important aspect of the micromechanics approach I wish to emphasize is that it is not enough to use micromechanics to obtain the macroscopic stress-strain relation, which has been the exclusive objective of nearly all the studies up to now. As Belytschko and I have been emphasizing, it is equally important to use micromechanics to determine the macroscopic parameters and laws which prevent complete localization of damage into a region of zero volume. Examples are the minimum crack band width \( w_c \) in the blunt crack band model, the characteristic length, the weighting function, and the form of the averaging law in the imbricate nonlocal model. The stress-strain relation suffices to describe the macroscopic material properties of a statistically heterogeneous material only if the body is in the state of uniform stress and strain, but not if the state is nonuniform. Likewise, it is insufficient to test material properties on specimens which are designed to be as nearly as possible in a state of homogeneous strain, and incorrect to interpret the test results by assuming a homogeneous strain distribution.

If the properties of the particle system in Fig. 38 were described merely by a macroscopic stress-strain relation, the
cracking zone would localize into a line. The fact that it does not can be captured only by properties other than the stress–strain relation, properties that serve to restrict the localization. It is time to realize that the thrust of micromechanics studies should be redirected towards the localization restricting characteristics of damage models.

5. SIZE EFFECT

As Willard Gibbs once remarked, the principal objective of science is to find the viewpoint from which a given problem appears the simplest. What is such a viewpoint for distributed cracking? In my opinion, it is the size effect. It must be taken into account in extrapolating from reduced-scale laboratory tests to full-size real structures. We may distinguish size effects of different kinds and sources.

One kind of size effect is manifested in ductility, i.e., the ratio of the deformation at the failure load to the deformation at the maximum load. The simplest example is a uniaxially stressed bar loaded in a displacement controlled manner through a spring (Fig. 44). The strain along the bar remains uniform until a certain critical point at which two different equilibrium increments become possible (76.2, 82.3, 85.4). For one of them the strain remains uniform, and for the other the strain becomes nonuniform at uniform stress—it increases (and is uniform) over a segment \( h \) of specimen length \( L \) and decreases (while remaining uniform) over the remaining segment of length \( L - h \), i.e., segment \( h \) undergoes loading and segment \( L - h \) unloading. The latter nonuniform deformation increment can occur at fixed load-point displacement, and represents an instability of failure. If \( h \to 0 \), the failure occurs at the peak stress point, i.e., no strain-softenin can be realized. The fact that strain-softenin can be observed in tests means that the strain-localization length \( h \) must be finite. As proposed in 1974 (76.2), this length represents the band width in the crack band theory (\( h = \omega \)), which is assumed to be a material constant.

Assuming a certain realistic stress–strain diagram for the material (Fig. 44), we obtain the diagrams in Fig. 44 showing the ductility, i.e., the ratio of the failure strain to the strain at peak stress, as a function of the relative size \( L/h \). The diagrams are plotted for various ratios of the spring constant to the equivalent spring constant for unloading of the bar. The dashed curves represent the simplified original solution (76.2) based on the crack band theory and the usual local continuum, and the solid curves represent the exact solution obtained by Bazant and Zubelewicz (1985) according to the imbricate nonlocal formulation in which the averaging operator is approximated as a differential operator with the Laplacian, as mentioned before (Fig. 26). An exact solution is possible in terms of trigonometric and hyperbolic functions. We see that, for a certain ratio \( \alpha \) of the characteristic length of the imbricate continuum to the crack band width imposed for a local continuum discretization, the ductility diagrams are nearly identical.

Another kind of size effect, usually the most important one, is the effect of the structure size on the nominal stress at failure, \( \sigma_N \), when geometrically similar structures are compared. In a two-dimensional problem, \( \sigma_N \), defined as the maximum load \( P \) divided by structure thickness \( h \) and the characteristic structure dimension \( d \), is independent of the structure size if the solutions based on plastic limit design or elastic allowable stress design (as well as hardening plasticity, viscoplasticity, or viscoelasticity) are considered. For fracture mechanics, however, there is a cardinal difference: \( \sigma_N \) decreases as the structure size increases.

The structural size effect was observed in fracture testing by Walsh (72.9) and others (84.30, 84.31, 85.6, 85.10). It was found that the size effect of linear elastic fracture mechanics is too severe. This size effect is of the type \( \sigma_N = d^{-1/2} \) (82.9, 84.26, 81.4, 81.3, 81.16, 73.14, 74.3, 73.7) or \( \log \sigma_N = -\frac{1}{2} \log d + \text{const} \), which means that the plot of log \( \sigma_N \) versus log \( d \) for geometrically similar structures of different sizes is a straight line of downward inclination \( -\frac{1}{2} \) (Fig. 45). The test results, however, point to a milder size effect, somewhat like that shown by the solid curve.

An approximate formula for the size effect law, shown in Fig. 45, was obtained by means of simplified solutions of certain simple problems by crack band theory. The same size effect law was also shown to ensue from the approximate, but apparently quite realistic, hypothesis (Fig. 45) that the energy release rate at failure depends on: (1) the length \( a \) of the fracture or the crack band, and (2) the area of the cracking zone \( nd_a \), where \( nd_a \) is the cracking front width which is assumed to be a constant multiple of the aggregate size \( d_a \) (84.5).

If only the second part of the hypothesis were made, one would obtain the size effect law of limit analysis or elastic allowable stress design, according to which \( \sigma_N \) is independent of the structure size, as shown by the horizontal dashed line in Fig. 45. This is the classical strength (or yield) criterion of failure. If only the first part of the hypothesis were made, one would obtain the size effect of linear elastic fracture mechanics, given by the straight line of slope \( -\frac{1}{2} \), as already mentioned.

It has been shown by dimensional analysis and similarity arguments (84.5, 85.13) that the complete hypothesis leads, with
confirmed this formula (Fig. 45) by analyzing notch effect reduces to a certain reasonable approximation, to the size effect law given and plotted in Fig. 45; \( B \) and \( \lambda_0 \) are two empirical constants characterizing the shape of the structure. Zaitsev and Kovler (85.54) confirmed this formula (Fig. 45) by analyzing notch sensitivity with a force approach. Intuitively, the key point is that in a small structure the area of the crack band occupies a large portion of the whole structure, and in a large structure it occupies a relatively small portion. For a relatively small structure, for which \( d/d_\lambda \) is negligible compared to 1, the size effect reduces to \( \sigma_N = Bf_\lambda \), which is the strength criterion of plastic limit analysis or elastic allowable stress design. For a very large structure, for which 1 may be neglected as it is much smaller than \( d/d_\lambda \), we obtain \( \sigma_N = d^{-1/2} \). Thus, we see that the size effect law represents a gradual transition from the strength or yield criterion to the linear elastic fracture mechanics criterion (84.5, 85.13).

Most available laboratory tests of structures are conducted on reduced size models, which are made, according to the conventional wisdom, as small as possible for a given aggregate size. The results then lie on the left, i.e., near the strength-criterion end of the size effect curve (Fig. 45). Extrapolation of such test results into real structure sizes is at present done according to the strength criterion. This is obviously questionable. In fact, the existing codes for the design of concrete structures (77.3, 78.3) do not take fracture mechanics into account and their formulas exhibit no size effect, i.e., always correspond to a horizontal line in the size effect plot. Nevertheless, the size effect should be observed for all failures that are of a brittle type, for which there exists no yield plateau after the maximum load is reached but the load declines at increasing displacement. Such brittle failures must exhibit size effect. In the case of concrete, they include the diagonal shear failure and the torsional failure of beams, the punching shear failure of slabs, the torsional-punching shear failure, the bending failure of plain concrete, the failure of pavement under wheel load, the beam and ring failures of unreinforced pipes, etc.

The size effect law agrees reasonably well with the results of laboratory fracture tests; see, e.g., Fig. 46 in which the classical test results of Walsh (72.9) for three-point loaded beams at three different beam sizes are plotted in a straight-line regression plot with the coordinate \( d/d_\lambda \) and ordinate \( (f_\lambda /\sigma_N)^2 \). In this plot, the size effect law appears as a straight line. Many subsequent results for different types of fracture tests provided similar results (85.6, 86.10). For brittle failures of real structures the data are much more limited; however, some information exists in the literature on the diagonal shear failure of beams (61.3, 62.2, 66.3, 68.1, 71.13, 72.8, 68.1, 75.5, 78.14, 81.19, 81.20); see the data points in Fig. 47, where the size effect plot is on the right and its linear regression on the left (84.9, 86.2).

The scatter of the data is large, due to the fact that test results from different laboratories on different concretes and different types of beams are combined in one diagram. Nevertheless, the downward trend with increasing beam depth is clearly apparent and contrasts with the present codes (77.3, 78.3) which imply no size effect (a horizontal line). It is also clearly apparent that the size effect is considerably milder than that for linear elastic fracture mechanics (the straight line of downward slope \(-1/2\)). Similar results, which are highly scattered but clearly confirm the downward trend, have been obtained at Northwestern University for other types of failure (85.6, 86.1).

The nonlinear fracture parameters have usually been determined on the basis of the so-called R-curves (resistance curves) which plot the energy required for crack growth as a function of the crack length \( c \) from the notch tip (60.3, 61.2, 73.14, 74.3, 82.32, 84.7, 84.26, 84.49). These curves are normally determined by measuring crack opening displacements and unloading compliance associated with various crack lengths in the same specimen. Such measurements are rather difficult,

![FIG. 46. Regression of Walsh's fracture data by size effect law (72.9).](image-url)
chiefly because of the ambiguity in the location of the crack tip. A considerably simpler measurement of nonlinear fracture parameters is made possible by the size effect law (84.10); it necessitates only the maximum load values for geometrically similar specimens of different sizes, which are easy to measure even with the most rudimentary equipment. The slope of the regression line as exemplified in Fig. 46 yields the fracture energy, and the Y-intercept of this line provides the remaining nonlinear fracture parameters. For each specimen size, it is possible to plot from these data one of the fracture equilibrium curves in Fig. 48. The R-curve then results as the envelope of these curves (84.10). A simple formula to obtain the approximate envelope from the parameters of the linear regression plot also exists. Furthermore, if the R-curve is given, it is possible to construct from it the size effect plot. So there is a one–two–one correspondence between these different characteristics of nonlinear fracture (84.10).

The size effect law in Fig. 45 is a simple approximation which is probably sufficient for most practical applications but is not likely to be valid for a very broad range of sizes. The most general size effect law can be expressed similarly to the formula in Fig. 45 if the expression in the denominator is replaced by a certain infinite series;

$$\sigma_W = B_f \left( a_0 \epsilon^{-1} + a_1 \epsilon + a_2 \epsilon^2 + a_3 \epsilon^3 + \cdots \right)^{1/2},$$

where $\xi = (d_0/d)'$ and $B_f, r, a_0, a_1, \ldots$ are constants (85.13). This yields a general asymptotic approximation which still has the same limits: the strength criterion and the linear elastic fracture mechanics. A special case, $\sigma_W = B_f [1 + (d'/d_0)^{1/2} - 1]^{1/2}$, is probably sufficient for most applications (85.2).

Another generalization of the size effect law in Fig. 45 is needed when concretes of greatly different aggregate sizes $d_0$ are considered. In that case the size effect law must be generalized as shown in Fig. 49, in which $c_0$ is an additional, third empirical material parameter (discussion closure of Ref. 84.9). This generalization is obvious if we realize that propagation of the cracking front cannot be continuous in the microscopic sense. The cracks must first form in the thin contact layers between the aggregate pieces, and only later they can join...
through the soft matrix of mortar. The formula in Fig. 49 results theoretically from the condition that a discontinuous microcrack must be produced in the next contact layer ahead of the microcracking front. Although fracture mechanics is used in this argument, the derivation is mathematically similar to that of the well-known Petch formula for the yield stress of polycrystalline metals, which was derived from dislocation theory.

The effect of the aggregate size is shown in the linear regression plots in Fig. 50 and the bottom of Fig. 51, in which the data of Chana, Taylor, and Iguro, Shioya et al, are used (85.13). Figure 51 shows the size effect regression plot of these data. The effect of aggregate size, presents an interesting question: given the structure size, should the maximum aggregate size be large or small? The answer depends on where the structure is located on the size effect curve (Fig. 52). While in the plot of \( \sigma_N \) vs the relative structure size \( d/d_s \), a smaller aggregate always yields a larger strength at equal \( d/d_s \), in the plots of \( \sigma_N \) vs the structure size \( d \), a smaller aggregate may yield either a larger or a smaller strength. If the structure is smaller than a certain critical size, it is better to use a small aggregate, and, if the structure is larger than this critical size, it is better to use a larger aggregate. This observation justifies, eg, the practice of using very large aggregate pieces in concrete dams.
Recognizing that material strength is a random variable, and introducing for it a certain suitable extreme value distribution, Weibull showed that the strength of a chain of elements decreases as the number of elements increases. It seems, however, that indiscriminate applications of this theory to all kinds of concrete structures have been misguided. For example, in the case of diagonal shear failure (Fig. 53), the range of locations at which the cracking front propagation can take place is very limited and is far smaller than the overall structure size. Furthermore, the condition of a series coupling of elements, implied in Weibull's theory, is obviously not met, even as an approximation. Thus, the statistical effect due to spatial random variation of strength should be relatively small. The major size effect is due to energy flow. It consists in the fact that in a larger structure there is much more energy available for being supplied to the cracking front. When test data of a limited size range are available, the size effect might seem to be in a reasonable agreement with the statistical theory; however, extrapolation to very large sizes is different. The reduction of structure strength is more severe for the fracture mechanics type of size effect. To be sure, a statistical aspect is always present in the failure of heterogeneous materials such as concrete or rock, but does not seem to be dominant for failures in which, by contrast to the failure of a chain, the failure occurs by propagation of a localized cracking band.

The size effect is not limited to tensional fracture. This effect, as well as the spurious mesh size sensitivity, is to be expected for all structures whose failure is due to strain-softening. This includes concrete failures due to compression or shear.

Tests of the shear-loaded specimens shown in Figs. 54 and 55, carried out at Northwestern University under displacement control in an MTS testing machine (Fig. 56), yielded the results shown in Fig. 57 (85.10). A strong size effect is evident. Furthermore, these tests verified that shear fracture is possible, despite previous claims that "shear fracture is shear nonsense," heard in some recent debates. If the cracks in Fig. 54 extended from the notch in a direction normal to the maximum principal tensile stress, they would not run vertically but deviate to the sides. Such cracks were observed by Arrea and Ingrafeia (82.2) for a different loading, in which the zone of high shear stress in the beam was relatively wide. In the present test, the zone of high shear stress is rather narrow. The fact that the cracks run vertically appears to be consistent with the criterion of the maximum energy release rate, the fundamental criterion for the fracture direction.

It must be also admitted that the specimen in Fig. 54 is not an ideal one for shear (Mode II) fracture, because, according to linear elastic fracture mechanics, the stress intensity factors $K_i$ and $K_ii$ are both nonzero (85.35). However, $K_i$ is small and has a negative value since the stresses are compressive, and so it should have a negligible effect. Pure Mode II situation could be generated by a slightly different loading system, which would, however, require additional tensile loads and would be much more difficult to realize (Bažant's discussion of Ref. 85.35).

![FIG. 54. Shear-loaded specimens (85.10).](image_url)

![FIG. 53. Specimens of sizes 1:2:4:8 after shear fracture (85.10, and Matériaux et constructions. RILEM, 1986, to appear).](image_url)
The regression plot for the shear fracture tests of geometrically similar specimens (Fig. 57, bottom) yields for the shear fracture energy a value which is much higher than that for Mode I (tensile) fracture—approximately 25 times higher. We can explain it (Fig. 58) by realizing that the shear stress is equivalent to tensions and compressions in diagonally inclined directions. The inclined tensile stresses produce a row of inclined microcracks; however, for the shear fracture to be complete, it is also necessary that the inclined compression stresses parallel to these microcracks would crush the material between the cracks, and the work of this crushing must be counted in calculating the fracture energy. Such an analysis yields shear fracture energy values which reasonably agree with those determined on the basis of the size effect regression plot.

The picture of the crack band with inclined microcracks in Fig. 59 further suggests a means of simulating shear friction and dilatancy on rough cracks in concrete and rock. If a shear stress is applied on a previously formed tensile crack, compression stresses are generated in the inclined directions across the rough crack modeled as a crack band. These inclined compression stresses have a horizontal component which represents the shear friction, as well as a normal stress component which either opposes dilatancy or leads to crack dilatancy if the normal stress component is not balanced. This type of mechanism of shear on cracks was analyzed using the microplane model (84.8), and it was found capable of closely representing various test data for aggregate interlock due to the shearing of cracks in concrete; see, eg, the comparisons of the results of the microplane model with the test data by Walraven and Reinhart (81.25).

A particularly attractive aspect of the crack band simulation of shear fracture and aggregate interlock by the microplane model is that it makes it possible to analyze shear on tensile cracks which are only partially formed (84.8). Such discontinuous cracking is more typical of concrete structures than the presently used assumption of continuous cracks subjected to shear.

FIG. 56. Size effect fracture tests in MTS testing machine (at Northwestern Univ.).

FIG. 57. Results of shear-loaded specimen tests (Figs. 54 and 55) carried out at Northwestern Univ. under displacement control in an MTS testing machine (Fig. 56) [after Bažant and Pfeiffer (85.10)]. Top: (●) concrete; (■) mortar. Bottom: (●) concrete, $d_0 = 0.5$ in., $f'_s = 410$ psi; (●) mortar $d_0 = 0.15$ in., $f'_s = 550$ psi; $Y = A x^2 + C$, $A = 0.01002$, $C = 0.3688$, $r = 0.9522$, $\omega_{yl} = 0.0911$.

FIG. 58. Shear after total or partial tensile fracture and aggregate interlock tests.
CONCLUSION

To sum up, distributed cracking in brittle heterogeneous materials may be modeled as strain-softening, but strain-localization instabilities must then be carefully analyzed and a complete localization into a line prevented. As a device to prevent such localization, the concept of imbricate nonlocal continuum may be used. The nonlocal concept introduces a new challenge to micromechanics: micromechanics should be used to predict not only the constitutive relation between homogenized stress and strain, but also the properties that prevent localization, such as the characteristic length and the averaging operator.

Finally, what is the salient aspect of the mechanics of distributed cracking? In my opinion it is the size effect. As Einstein once remarked, a good theory should be as simple as possible but no simpler than that. In our case, the theory of distributed cracking must not be as simple as to miss the size effect. This is the principal issue.

ACKNOWLEDGMENTS

Partial financial support under AFOSR Grant No 83-0009 to Northwestern University is gratefully acknowledged. Mary Hill deserves thanks for her invaluable secretarial assistance.

BIBLIOGRAPHY

1903–1960
59.1 Batdorf, S B, and Budianski, B (1949). A mathematical theory of strain instabilities must then be carefully analyzed and a complete localization into a line prevented. As a device to prevent such localization, the concept of imbricate nonlocal continuum may be used. The nonlocal concept introduces a new challenge to micromechanics: micromechanics should be used to predict not only the constitutive relation between homogenized stress and strain, but also the properties that prevent localization, such as the characteristic length and the averaging operator.

Finally, what is the salient aspect of the mechanics of distributed cracking? In my opinion it is the size effect. As Einstein once remarked, a good theory should be as simple as possible but no simpler than that. In our case, the theory of distributed cracking must not be as simple as to miss the size effect. This is the principal issue.

ACKNOWLEDGMENTS

Partial financial support under AFOSR Grant No 83-0009 to Northwestern University is gratefully acknowledged. Mary Hill deserves thanks for her invaluable secretarial assistance.

BIBLIOGRAPHY

1903–1960
59.1 Batdorf, S B, and Budianski, B (1949). A mathematical theory of strain instabilities must then be carefully analyzed and a complete localization into a line prevented. As a device to prevent such localization, the concept of imbricate nonlocal continuum may be used. The nonlocal concept introduces a new challenge to micromechanics: micromechanics should be used to predict not only the constitutive relation between homogenized stress and strain, but also the properties that prevent localization, such as the characteristic length and the averaging operator.

Finally, what is the salient aspect of the mechanics of distributed cracking? In my opinion it is the size effect. As Einstein once remarked, a good theory should be as simple as possible but no simpler than that. In our case, the theory of distributed cracking must not be as simple as to miss the size effect. This is the principal issue.

ACKNOWLEDGMENTS

Partial financial support under AFOSR Grant No 83-0009 to Northwestern University is gratefully acknowledged. Mary Hill deserves thanks for her invaluable secretarial assistance.

BIBLIOGRAPHY

1903–1960
59.1 Batdorf, S B, and Budianski, B (1949). A mathematical theory of strain instabilities must then be carefully analyzed and a complete localization into a line prevented. As a device to prevent such localization, the concept of imbricate nonlocal continuum may be used. The nonlocal concept introduces a new challenge to micromechanics: micromechanics should be used to predict not only the constitutive relation between homogenized stress and strain, but also the properties that prevent localization, such as the characteristic length and the averaging operator.

Finally, what is the salient aspect of the mechanics of distributed cracking? In my opinion it is the size effect. As Einstein once remarked, a good theory should be as simple as possible but no simpler than that. In our case, the theory of distributed cracking must not be as simple as to miss the size effect. This is the principal issue.

ACKNOWLEDGMENTS

Partial financial support under AFOSR Grant No 83-0009 to Northwestern University is gratefully acknowledged. Mary Hill deserves thanks for her invaluable secretarial assistance.
700 Bažant: Mechanics of distributed cracking

mechanics, Proc 10th annual meeting, Soc of Eng Sci, Raleigh NC.
71.41 Rasmussen, S, and Herrmann, M A, and Herrmann, I. R. A numerical biaxial characterization for concrete, J Eng Mech ASCE 100(5), 935-948.
71.44 Balant, Z P. Mechanics of distributed cracking, Trans from Russian) 10(4), 87-95.
71.47 Janson, J, and Hult, I. Fracture mechanics and damage mechanics, a combined approach, J Mécanique Appli 1(1), 69-84.
71.51 Prevost, J H, and Hert, K. Soil mechanics and plasticity analysis of strain softening, Geotechnique 25, 279.
71.54 Secco, D C, and Dougall, J W. A quantitative approach of crack assessment and concrete design in concrete bridge design, Mag Concr Res 27(92), 151-206.
71.57 Balant, Z P. Mechanics of distributed cracking, Trans from Russian) 10(4), 87-95.
71.58 Balant, Z P. Mechanics of distributed cracking, Trans from Russian) 10(4), 87-95.
71.67 Kafé, A F, Nonlinear analysis of reinforced concrete slabs, and shells and for time dependent effects, PhD thesis, UC-SESM 76-6, Univ of California, Berkeley.
71.68 Minasian, S, and Nadeau, J S. Effect of notch width on KIC for mortar and concrete, Cement Concr Res 6, 529-534.
lightweight concrete beams without shear reinforcement, Stevin Lab Rep No 5-74-8, Delft Univ of Technology, Netherlands.

1979


79.5 Dragon, A, and Mrm, Z. A continuum model (or plastic-brittle behavior of rock and concrete, Int J Eng Sci 17, 137.


79.9 Keenan, J H. et al. Principles of thermodynamics, in Encyclopedia Britannica, 15th ed, vol 18, p 301 (Fig. 6).


Batant: Mechanics of distributed cracking

Eng Fracture Mech.


84.34 Mazars, J. Application de la mécanique de l’endommagement au comportement à la rupture du béton de structures, thèse de Doctorat d’Etat, Université de Paris 6.

84.35 Mier, J M van. Complete strain-strain behaviour and damaging stage of concrete under multiaxial conditions, in RILEM-CEB symposium on concrete under multiaxial conditions, INSA Toulouse, session 3, experimental results, vol 1, pp 75–85.


84.40 Read, H E, and Hegemier, G A. Strain softening of rock, soil, and concrete—a review article, Mech Mat.

84.41 Reinhart, H, W. Fracture mechanics of an elastic softening material for concrete, Heron 29 (2).


85.20 De Borst, R, and Nauta, P. Non-orthogonal cracks in a smeared finite element model, Eng Fracture Mech.

85.21 Determination of the fracture energy of mortar and concrete by means of three-point bend tests on notched beams, RILEM recommendations, Mat Construct (RILEM, Paris) 18 (106), 43–48.

85.22 Determination of the fracture energy of mortar and concrete by means of three-point bend tests on notched beams, RILEM recommendations, Mat Construct (RILEM, Paris) 18 (106), 43–48.


Zdeněk P Bažant was born and educated in Prague, obtaining a Civil Engineer degree from Czech Technical University in 1960 and a PhD in mechanics from the Czechoslovak Academy of Sciences in 1963. He moved to University of Toronto in 1967, then to UC Berkeley in 1968, and in 1969 joined the faculty of Northwestern University where he has been Professor of Civil Engineering since 1973, Structural Engineering Coordinator from 1974 to 1978, and Director of the Center for Concrete and Geomaterials since 1981. He is a Registered Structural Engineer in Illinois, a consultant to Arconne National Laboratory and other firms, and is on the editorial boards of a number of journals. Dr. Bažant is active as a fellow in ASCE, ACI, and RILEM, serving currently as chairman of the ACI Committee 446 on Fracture Mechanics, of the RILEM International Committee TC69 on Concrete Creep, and of IASAMRT Division H, Concrete Structures, and recently as chairman of the ASCE-EMD Committee on Properties of Materials. He is a member of the ASCE Committee on Probabilistic Methods, the ACI Safety Committee, and the ASTM Committee on Skiing. The distinguished nature of his research contributions has been recognized by the conferment of the RILEM Medal, the ASCE Huber Prize, the ASCE Lin Award, the 1980 Award Guggenheim and Ford Foundation Fellowships, the French Government ASTEF Fellowship, and his election as Fellow of the American Academy of Mechanics.