(u) DTIC FILE COPY (2)

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

AD-A223 211	1b. RESTRICTIVE MARKINGS
	3. DISTRIBUTION / AVAILABILITY OF REPORT
4. PERFORMING ORGANIZATION REPORT NUMBER(S): No. 90-2/498-AR1	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR-90-0503
6a. NAME OF PERFORMING ORGANIZATION: Northwestern University; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR
6c. ADDRESS (City, State, and ZIP Code): 633 Clark Street, Evanston, IL 60208	7b. ADDRESS (City, State, and ZIP Code): AFOSR/NA, Bolling AFB DC 20332-6448
8a. NAME OF FUNDING / SPONSORING ORGANIZATION: Air Force Office of Scientific Research; 8b. OFFICE SYMBOL (If applicable): NA	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-87-0030
8c. ADDRESS (City, State, and ZIP Code): Building 410, Bolling Air Force Base, Washington, D.C. 20904-6448	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2302; TASK NO. C2; WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification): Micromechanics of Size Effect in Failure Due to Distributed Cracking (u)	
12. PERSONAL AUTHOR(S): Zdenek P. Bazant and Ted Belytschko	
13a. TYPE OF REPORT: Final Report; 13b. TIME COVERED: FROM 1/1/87 TO 12/31/89; 14. DATE OF REPORT (Year, Month, Day): February 26, 1990; 15. PAGE COUNT	
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

See reverse

DTIC ELECTE APR 30 1990 S D D

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☒ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: (u)
22a. NAME OF RESPONSIBLE INDIVIDUAL: Dr. Spencer T. Wu	22b. TELEPHONE (Include Area Code): (202) 767-6962; 22c. OFFICE SYMBOL: NA

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60208-3109

AFOSR-TR- 90-0503

THE TECHNOLOGICAL INSTITUTE
DEPARTMENT OF CIVIL ENGINEERING

February 27, 1990

distribution unlimited.

Dr. Spencer T. Wu
Air Force Office of Scientific Research
Directorate of Aerospace Sciences
Room A237, Building 410
Bolling Air Force Base, DC 20332-6448

Dear Spencer:

Enclosed are six copies of the Final Report on AFOSR Contract No. F49620-87-C-0030 with Northwestern University, entitled "Micromechanics of Size Effect in Failure Due to Distributed Cracking."

Sincerely yours,

Zdenek P. Bazant
Professor of Civil Engineering

Ted Belytschko
Professor of Civil and
Mechanical Engineering

ZPB/TB:gs
xc: Ms. Rebecca C. Applebaugh, Contract Administrator, AFOSR
Ms. Barbara E. Siegel, Research and Sponsored Programs, NU
Enclosures: 6 copies of annual scientific report

CONTENTS

Abstract 1

Introduction 2

1. SIZE EFFECTS 2

2. MICROMECHANICS STUDIES 6

3. NONLOCAL CONTINUUM APPROACH 10

4. LOCALIZATION INSTABILITIES 14

5. THERMODYNAMIC ANALYSIS OF STABLE PATH 16

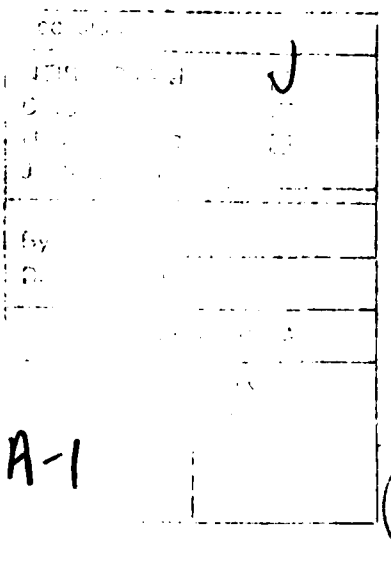
6. NUMERICAL AND ANALYTICAL STUDIES OF LOCALIZATION 17

7. CONCLUSION 18

PUBLICATIONS 20

FIGURES 25

APPENDICES 27





Abstract

An extensive study of the micromechanics aspects and size effects associated with strain softening damage due to distributed cracking in brittle heterogeneous materials such as concrete, rocks and ceramics has been carried out and presented in a sequence of publications. The results may be grouped in five categories: (1) Size Effects, (2) Micromechanics Aspects, (3) Nonlocal Continuum Models, (4) Localization Instabilities, (5) Thermodynamic Analysis of Stable Path, and (6) Numerical Implementation of Localization Limiters.

The law governing the size effect due to localization of strain softening into a finite size fracture process zone has been formulated, experimentally verified and calibrated. Knowledge of the size effect law has led to a method of determining nonlinear fracture properties on the basis of maximum loads of geometrically similar specimens of different sizes. The concept of a brittleness number characterizing the proximity of response of any structure to linear elastic fracture mechanics has been developed. Extensive experiments have been conducted on concrete, rock and ceramics. The size effect analysis has been extended to variable temperature and humidity conditions, as well as to fatigue fracture. Application to high strength concrete indicated that the fracture energy does not increases as much as strength and the brittleness of high strength concrete structures is higher than that of normal structures.

In the study of micromechanics, considerable effort has been devoted to random particle simulation of cracking and fracture in aggregate composites. A method to randomly simulate such microstructures has been formulated and the knowledge of the size effect has been used to determine fracture characteristics of random particle systems. Furthermore, analytical solutions of simplified situations with regular crack arrays have been carried out to gain insight into the nonlocal aspect of the macroscopic smoothing continuum and into the shape of the post-peak stress-displacement softening curve. For the description of the stress-strain relations for the fracture process zone, a microplane model in which the constitutive properties are described independently on planes of various orientations in the material has been developed and shown to be efficient and represent well available triaxial test data. Microcrack distributions in the fracture process zone have been observed microscopically, slicing specimens in which microcracks were propped open by a resin with fluorescent dye.

Nonlocal continuum models based on degradation of stiffness or degradation of the yield limit have been formulated and shown to allow effective solution of large finite element systems. The key property of these models is that nonlocal are only those aspects which are associated with strain softening while the elastic response is treated as local. Stability of the nonlocal models has also been analyzed. A method for measurement of the characteristic lengths of the nonlocal continuum model has been developed. As another aspect of micromechanics, softening instabilities have been solved analytically for localization into a planar band within a layer of finite thickness and localization into ellipsoidal regions. Conditions for instability as well as bifurcation on a stable path have been given. The conditions for stability and path bifurcation in a material with damage have been formulated on the basis of thermodynamic considerations.

INTRODUCTION

The salient property of failure of brittle aggregate materials such as concrete, rocks and ceramic composites is the existence of distributed cracking and damage. Fracture is obtained through the localization of cracking and is characterized by the presence of a microcracking zone in front of the fracture tip as well as bridging stresses behind it. The process of localization of cracking or damage leads to size effects which are intermediate between the size effects in the failures without localization, typified by plastic failures, and failures with localization of damage into a point, typified by linear elastic fracture mechanics. Due to the existence of distributed cracking, the classical continuum modeling of materials becomes inadequate, and some form of a nonlocal formulation is required in order to correctly capture the size effect and localization phenomena.

A three-year research program dealing with these problems has just been completed at Northwestern University. The research results have been reported in full detail in a number of papers (see the list in Appendix). The main results are briefly summarized in the present report, and the principal publications are attached as Appendices. The presentation of the results is grouped into six categories: (1) Size Effects, (2) Micromechanics Studies, (3) Nonlocal Continuum Approach, (4) Localization Instabilities, (5) Thermodynamic Analysis of Stable Path, and (6) Numerical and Analytical Studies of Localization.

1. SIZE EFFECTS

First, we will describe the results on the size effect which have been obtained on the basis of experiments, coupled with dimensional analysis, similitude arguments and asymptotic considerations. Discussion of the micromechanics aspects will be postponed to Section 2.

The size effect is the salient aspect of fracture mechanics. While all strength-based theories of failure, e.g., plasticity, exhibit no size effect, the classical linear elastic fracture mechanics exhibits a very strong size effect in which the nominal stress and failure (load

divided by dimension and thickness) is proportional to the inverse square root of the size. In brittle heterogeneous materials such as concrete, rocks or ceramics, the size effect is transitional between these two extremes. It has been shown that from the shape of the transition curve one can determine all the basic nonlinear fracture characteristics.

The principal results on the size effect are presented in Appendices A1 and A2. A large series of size effect tests of geometrically similar fracture specimens of various types has been conducted and analyzed. With reference to size effect, a new, unambiguous definition of the fracture energy of a material with a large fracture process zone has been conceived: The fracture energy is the specific energy required for crack growth in an infinitely large specimen.

Theoretically, this definition eliminates the effect of specimen size as well as shape and the type of loading on the fracture energy value since it can be demonstrated by asymptotic analysis that the stress field surrounding the fracture process zone in an infinitely large specimen is always the same, except for a multiplicative factor. The problem is, therefore, to identify the correct size effect law to be used for the extrapolation to infinite size. It has been shown that a rather simple law (Eq. 2 in A1; and 6.2), proposed by Bazant, is applicable for this purpose as an approximation. It has been experimentally verified that very different types of specimens including three-point bend fracture specimens, edge-notched tension specimens, and eccentric compression specimens indeed yield approximately the same fracture energy values.

A method to calculate the R-curves (resistance curves) has been formulated, and evaluation of the aformentioned tests showed that the R-curves for different specimen shapes are rather different although they have the same final asymptotic value (6.7 in A1).

It was further shown that, on the basis of the size-effect law, one can define a brittleness number (Eq. 8 and 10 in A1), which indicates how close the behavior of a specimen or structure of any geometry is to linear elastic fracture mechanics or to plastic limit analysis.

Appendix A2 describes determination of all the basic nonlinear fracture parameters on the basis of the size effect. Eq. 4 in A2 represents a size effect law formulated in terms of

the fracture energy and the effective length of the fracture process zone. What makes this formulation possible is the introduction of the intrinsic (shape-independent) dimension of the structure and the intrinsic nominal stress at maximum load (Eq. 6 in A2). Based on this more fundamental form of the size effect law, the brittleness number of a structure can be defined as a ratio of the intrinsic size to the effective length of the fracture process zone (Eq. 11 in A2).

Furthermore, the R-curve can be expressed in terms of the fracture parameters as given in Eq. 22 of A2. This expression for the R-curve has been used to calculate load-deflection curves of various types of specimens. Comparisons with experimental results have confirmed a previous theoretical indication that the master R-curve obtained from the size effect law is applicable only in the hardening regime, i.e., as long as the load is increasing. For the post-peak regime, it has been found that the R-curve must be assumed to be horizontal, with a value equal to that at the peak load. This behavior is due to the fact that in the post-peak regime the fracture process zone stops growing, separates from the notch tip and travels forward at roughly constant size.

Appendix A1 also gives experimental verification of the applicability of the size effect law to ceramics, in particular silicon carbide and slip-cast fused silica. In addition, Appendix A2 also contains a concise summary of the entire theory.

Appendix A3 exploits the size effect law for determining the dependence of the fracture properties of concrete on temperature and for water content of conerete specimens. It is found that the temperature dependence of fracture energy roughly follows a simple formula (Eq. 5 in A3) which has been derived on the basis of the activation energy theory for molecular bonds. The activation energy has been found to be strongly dependent on the moisture content. Experiments showed the fracture energy of dried concrete to be much less than that of wet concrete, especially at high temperatures.

Appendix A4 extends the size effect law to fatigue fracture. Crack growth in geometrically similar notched concrete specimens of various sizes, caused by load repetitions, has been measured by means of the compliance method. It was found that Paris law is valid only

for specimens of one size or asymptotically for very large specimens, but is invalid when specimens of different sizes are compared. However, a generalized form of Paris law, which is combined with a size effect law known from monotonic fracture, was shown to agree with test data quite well. The size-adjusted Paris law - a basic new result - gives the crack length increment per cycle as a power function of the amplitude of a size-adjusted stress intensity factor amplitude. The size adjustment is based on the brittleness number of the structure.

Appendix A5 presents application of the size effect method for the determination of fracture characteristics of high strength concrete. Tests of geometrically similar fracture specimens of different sizes revealed that the fracture properties of high strength concrete are rather different from those of normal concretes. Increase of the concrete strength by the factor of 2.6 was found to increase fracture toughness by only 25 percent, and at the same time to decrease the effective fracture process zone length by about 60 percent. This causes that the brittleness number of a given structure made of high-strength concrete is more than twice as large as that of the same structure made of normal concrete. This new result represents an adverse feature which will need to be coped with in design.

Appendix A6 presents an application of the size effect method to the measurement of fracture properties in mode III fracture (shear fracture). The test specimens were geometrically similar cylinders with a circumferential notch, loaded in torsion. The mode III fracture energy was found to be about three-times larger than the mode I (opening mode) fracture energy, provided the normal force resultant across the fracture plane is zero; otherwise, the fracture energy is very sensitive to the normal force value.

Finally, appendix A7 presents in simple terms the consequences of the size effect for the design codes. The results of this project have shown that the formulas for brittle failures of reinforced concrete structures in the existing design codes, which are based on plastic limit analysis, need to be revised by incorporation of the size effect, which is ignored at present. Experiments in which the diagonal shear failure loads of beams were measured on specimens of different sizes, (with the size range 1:16) have confirmed that the size effect is indeed significant from the designer's viewpoint.

2. MICROMECHANICS STUDIES

After discussing the results obtained on the basis of experiments, dimensional analysis, and asymptotic considerations, we will now describe the results obtained by micromechanics analysis.

Although some researchers have attempted to analyze the mechanics of fracture by subdividing the aggregate pieces as well as the mortar or cement matrix into many finite elements, in general this approach is only of limited usefulness due to preposterous demands for computer time and storage. Much more effective, and for most purposes adequate, is a simplified model of the microstructure in which the large aggregate pieces are considered as rigid, randomly arranged particles with axial interactions which allow for elastic response and interparticle cracking (softening). Such a model (which represents an adaptation of a similar earlier model of Cundall for sands) is presented in Appendix B1.

A method of random computer generation of a particle system meeting the prescribed particle size distribution has been developed; see the simulations in 6.2 of B1. The particles are assumed to be elastic, and the interparticle contact layers of the matrix are described by a softening stress-strain relation corresponding to a prescribed microscopic interparticle fracture energy. Both two and three-dimensional versions of the model are easy to program, but the latter poses at present extremely large demands for computer time. The model is shown to realistically simulate the spread of cracking and its localization.

With regard to the size effect, it has been shown that the random particle model correctly exhibits the transitional size effect described by the size effect law and agrees quite well with various experimental data. In this regard it should be noted that continuum models of the classical, local type are inherently incapable of describing the size effect (this is of course also true of finite element models based on a local continuum). Direct evaluation of fracture energy for the material represented by random particle models on the basis of interparticle forces is in general a difficult proposition since it is not clear which interparticle force work should be counted. However it is shown that the size effect method can be used to determine the fracture energy and the effective process zone lengths of the material represented by the

random particle system quite well, and unambiguously.

Another aspect of the distributed cracking in materials such as concrete is examined in Appendix B2, in which strain-softening damage due to distributed cracking is modeled by an elastic continuum with a quasi-periodic array of cracks of regular spacing but varying sizes. The strain due to cracking as well as the corresponding compliance are calculated analytically. The cracked material is homogenized in such a manner that the microscopic continuum strains satisfy exactly the condition of compatibility with the actual strains due to cracks, and the microscopic continuum stress satisfies exactly the condition of work equivalence with the actual stresses in the cracked material. The results show that, contrary to the existing theories, the damage variable used in continuum damage mechanics should be nonlocal, while the elastic part of the response should be local. The mathematical analysis also indicates that, for this particular crack system, the nonlocal continuum damage should be considered as a function of the spatial average of the cracking strain rather than its local value. This partially confirms the nonlocal damage concept.

The fracture energy represents the complete area and the stress-displacement curve characterizing the fracture process zone. This curve is very difficult to measure, only indirect inferences can be made. Therefore, determination of the shape of the stress displacement curve in the post-peak softening range needs to be aided by micromechanics analysis. Such an analysis is presented in Appendix B3, in which again the fracture process zone is assumed to consist of a two-dimensional array of circular cracks, in the initial stage, or circular ligaments, in the terminal stage of fracture. For both cases analytical solutions of the stress displacement curve are obtained. The solutions are approximate but asymptotically exact for very small cracks or very small ligaments. As an interesting result, the stress displacement curve in the post-peak range is found to exhibit snapback instability characterized by a maximum possible displacement value. However, it is observed that the snapback might be characteristic only of the elastic-fracture type of response, and might be suppressed by frictional phenomena or the effect of inhomogeneities, which could not be considered in the analytical solution.

The interactions between the sites of concentrated inelastic deformations in the microstructure are basically of two types:

(1) interactions at distance, and

(2) interactions between different orientations of the planes on which inelastic phenomena take place.

The first type of interactions can be approximately characterized, on the macro-level, by nonlocal continuum properties. If the continuum includes the nonlocal properties, the stress-strain relation for the fracture process zone needs to reflect only interactions between different orientations. This type of interactions is effectively described by the microplane model, presented in Appendix B4.

This model, which essentially represents generalization and modification of the classical slip theory of G.I. Taylor for plasticity of polycrystaline metals, can describe not only tensile cracking but also the general nonlinear triaxial response in compression and shear, as demonstrated by comparisons with experiments in B-4. In this approach, the constitutive properties are characterized separately on planes of various orientations within the material, called the microplanes (the original term in plasticity was "slip planes"). The advantage is that on the microplanes there are only two stress and strain components and no tensorial invariance requirements need to be observed. These requirements are satisfied automatically by integration over all spatial directions. The microplane strains are assumed to be the results components of the macroscopic strain tensor. The central assumption is that on the microplane level the stress-strain diagrams for monotonic loading are path independent and that all the path dependence on the macro level is due to unloading, which occurs selectively on microplanes of different orientations. The microplane theory yields the stiffness or compliance matrix of the material, which in general can be non-symmetric, due to frictional and dilatancy phenomena. Compared to previous macroscopic phenomenologic tensorial models, the microplane model developed involves fewer free material parameters and can be more easily adopted to tests triaxial test results. It has been shown that the model can represent the existing triaxial test data as well as tensile fracture data very well; see Appendix B5.

The microplane model requires a relatively large number of computations to obtain the stress tensor from the given strain tensor; one trades high computer time requirements for conceptual simplicity and versatility. However, with rapid progress in computer capabilities, these requirements are becoming quite manageable.

To obtain a more realistic picture of microcracking in the fracture process zone, and its evolution during the loading process, microscopic observations of cracks have been conducted. There have been two basic objectives:

(1) To observe open microcracks as they exist under the load in the post-peak softening regime, and

(2) To observe microcracks inside the specimen, not on its surface.

After the test, when the specimen is unloaded, most microcracks close. Consequently, they cannot be easily seen, and their previous width under load cannot be determined. These were serious limitations of all previous studies. Furthermore, microcracks have customarily been observed on the specimen surface. But it is likely that the crack pattern on the surface is quite different from that in the interior. Probably, the fracture process zone in the specimen interior (e.g., in the middle of specimen thickness) is wider than on the surface because of more restraints against localization. The surface layer of concrete inevitably contains a lower proportion of aggregate than the specimen interior, and a higher percentage of mortar; hence it is more homogeneous, which should promote localization. Another difference is due to the effect of Poisson ratio. In the specimen interior, there are transverse normal stresses, but not on the surface. This causes a reduced overall stiffness for the surface layer, and also produces a complicated three-dimensional stress field, which in linear elastic fracture mechanics is manifested by a three-dimensional singularity (whose exponent differs from 1/2 if the crack edge is normal to the surface).

To circumvent these limitations of previous studies, the following procedure has been conceived:

(1) Impregnate the fracture specimens under load, in a deformed post-peak state, using a resin with a fluorescent dye,

(2) Harden the resin by microwave heating while the specimen is under load (and thus the cracks remain open)

(3) Unload the specimen, cut it, polish the surface of the cut, and observe the cracks filled by hardened fluorescent resin under a microscope.

The experimental procedure, however, has turned out to be extremely difficult, tedious and therefore costly. Due to funding limitations, this part of experimental work could not yet be completed. Numerous types of specimens and methods of impregnation under load have been tried. During the last of year of investigation a method which appears feasible has finally been devised. The specimens are small compact tension specimens, of side 2 in., loaded by a wedge. All the wedge assembly must be made of a ceramic so as not to interfere with the microwaves. The wedge system makes it possible to block the wedge and thus keep the specimen deformed (and the microcracks open) while the specimen is moved to the impregnation chamber. In the impregnation chamber, which was built to this purpose, the specimen is first subjected to vacuum, and then to high pressure, while submerged in a bath of low viscosity polymer. In this manner it has already been possible to impregnate specimens successfully (an important point, difficult to achieve, is to impregnate not only those cracks that are connected to the surface, but also isolated microcracks, with no conduit to the surface - obviously the resin must pass through the fine pores of concrete). The method now appears to work but, due to its time-consuming nature and high cost, proceeds slowly. Fig. 1 shows the samples of some crack patterns observed in this manner, but no conclusions can yet be drawn from the limited data obtained so far.

3. NONLOCAL CONTINUUM APPROACH

It has previously been amply demonstrated that the classical, local continuum formulations, as well as their finite elemcnt implementations, cannot represent the transitional size effect observed in concrete-like materials. The reason is that, in local models, the strain softening always localizes to a region of zero volume, while in reality the size of this region,

representing the fracture process zone, is finite. The continuum model must include a mathematical device which limits localization to a zone of a certain finite, minimum size. Such a localization limiter is automatically achieved by nonlocal continuum.

Appendix C1 presents a rather effective and general nonlocal continuum model which can be implemented in large finite element programs. The model has been shown to avoid problems of convergence at mesh refinement and spurious mesh sensitivity, which are intractable with the classical continuum approach. The model is a generalization of plasticity, in which strain softening is characterized by degradation of the yield limit.

The key idea is to apply the nonlocal concept only to those parameters which cause the softening or degradation while keeping the total strains local. Compared to the previously advanced fully nonlocal (imbricate) continuum formulation, the new approach has advantage that the stresses are subjected to the standard differential equations of equilibrium and standard boundary or interface conditions, while in previous nonlocal formulations these equations were of higher order and involved additional terms. Two-dimensional finite element solutions of problems with several thousand degrees of freedom have been presented, documenting convergence and efficacy of the nonlocal numerical model.

Another type of nonlocal model in which strain softening is due to degradation of stiffness rather than degradation of the yield limit is presented in Appendix C2. This approach is more realistic for concrete, but does not have the advantage that the finite element program can be easily obtained by adapting a finite element program for plasticity, as in the previous case. The model in Appendix C2 generalizes the classical smeared cracking model for concrete, which has previously been shown to exhibit spurious mesh sensitivity. The crack band model for smeared cracking, developed previously, circumvents these deficiencies but is not completely general and can exhibit shear locking and directional mesh bias.

In Appendix C2 it is shown that all these problems can be avoided by a nonlocal generalization, in which the damage that characterizes strain softening is considered to be a function of the spatial average of the positive part or the maximum principal strain. Two alternatives of the model are presented: (1) Smeared cracking whose direction is fixed when

the cracks start to form, and (2) Smeared cracking whose orientation rotates with the maximum principal strain. Furthermore, fracture tests on specimens of various sizes are analyzed by finite elements. It is shown that the model correctly reproduces the experimentally observed size effect and agrees with the size effect law. Furthermore, orthogonal and slanted meshes are shown to yield, for the same specimen, approximately the same cracking zones and propagation directions. This confirms that mesh bias, which plagued previous models, is avoided.

A theoretically more consistent and fundamental approach to the degradation of stiffness of the material is the continuum damage theory, originated by the works of Kachanov. Again, this theory is incapable of representing the transitional size effect observed in concrete-like materials, and its finite element representations exhibit spurious mesh sensitivity due to unlimited localization and incorrect convergence when the mesh is refined. A nonlocal version of the continuum damage theory, which avoids all these problems, is presented in Appendix C3. The central idea is again to subject to nonlocal treatment only those variables that control strain softening, and to treat the elastic part of response as local. In the continuum damage theory, the only required modification is to replace the usual local damage energy release rate with its spatial average over the representative volume of the material whose size is a characteristic of the material (and roughly coincides with the width of the fracture process zone). Avoidance of spurious mesh sensitivity and proper convergence are demonstrated by numerical examples, including static strain softening in a bar, longitudinal wave propagation in a strain softening material, and static layered finite element analysis of a beam.

The nonlocal models presented in Appendices C1-C3 are relatively simple but not completely realistic for concrete and similar materials, especially in the case of nonlinear triaxial behavior. A more realistic model can be achieved by a nonlocal generalization of the microplane model, as presented in Appendix C4. The nonlocal concept is again applied only to those microplane strain components that are associated with strain softening, and not to those representing the elastic part of response. An effective numerical algorithm permitting large loading steps is developed, applying the idea of exonential algorithms previously formu-

lated for creep. Problems due to nonsymetry of the tangential stiffness matrix are avoided by using the initial elastic stiffness matrix in the incremental force displacement relations. Numerical finite element results with the nonlocal microplane model demonstrate a correct transitional size effect as observed in concrete and approximately described by the size effect law.

Introduction of the nonlocal damage formulation requires answering certain fundamental questions with regard to localization instabilities. These questions are addressed in Appendix C5. It is shown that the energy dissipation and damage cannot localize into regions of vanishing volume. The solution of the static strain localization instability in one dimension is reduced to an integral equation constrained by inequalities. It is shown from this equation that localization is controlled by the characteristic length of the material introduced in the spatial averaging rule. The calculated static stability limits are close to those obtained in the previous nonlocal studies, as well as to those obtained by the crack band model in which the continuum is treated as local but the minimum size of the strain softening region is prescribed as a localization limiter. Furthermore, the rate of convergence of static finite element solutions with nonlocal damage is studied and is found to be of a power type, almost quadratic. A smooth weight function of bell shape is found to yield an averaging operator that gives superior convergence.

The basic material property that characterizes the nonlocal aspects in models of all types is the characteristic length of the material. This length governs the minimum possible width of a zone of strain softening damage or, from another viewpoint, the minimum possible spacing of line cracks in discrete fracture models. A method of experimental determination of the characteristic length is presented in Appendix C6. The basic idea is to compare the response of two types of specimens, one in which the tensile softening damage remains distributed and one in which it localizes. The latter type of specimen is an edge-notched tensile fracture specimen, and the former type of specimen is of the same shape but without notches. Localization of strain softening damage is prevented by gluing to the specimen surface a layer of parallel thin steel bars and by using a cross-section of a minimum possible thickness that

can be cast with a given aggregate. The characteristic length is obtained as the ratio of the fracture energy (the energy dissipated per unit area) to the energy dissipated per unit volume. Evaluation of these energies from the tests of concrete indicated the characteristic length to be about 2.7 times the maximum aggregate size. However, this ratio will no doubt depend on the properties of the matrix and the interfaces, granulometry and other factors.

4. LOCALIZATION INSTABILITIES

Fracture of brittle heterogeneous materials that undergo cracking may be regarded as a process of localization of damage or cracking into a small zone, called the fracture process zone. The minimum possible size of the zone can be determined by stability analysis of localization. Illuminating in this regard are the solutions of some simple prototype situations such as localization of a field of homogeneous strain into a planar band of localized damage or into ellipsoidal regions. Analytical solutions to such problems are presented in Appendices D1-D3. These solutions describe vital aspects of the micromechanics of damage propagation.

In Appendix D1, the continuum is for the sake of simplicity considered as local but localization of strain softening into a region of vanishing value is precluded by requiring that the softening region, assumed to be in a state of homogeneous strain, must have a certain minimum thickness that is a material property, related to the characteristic length. Exact conditions of stability of an initially uniform strain field against localization are obtained for the case of an infinite layer in which the strain localizes into an infinite band. This solution generalizes a previous solution of Rudnicki and Rice, who treated localization into a band in an infinite space rather than in a layer of finite thickness. Consideration of a finite thickness is necessary for bringing to light the size effects. The localization problem is solved first for small strain. Then a linearized incremental solution is obtained taking into account the geometrical nonlinearity of strain. The stability condition (Eq. 13 in D1) is shown to depend on the ratio of the layer thickness to the softening band thickness; hence the size effect. It is found that if this ratio is not too large compared to 1, then the state of homogeneous strain may be stable well into the softening range.

Extending the study of localization in bands, Appendix D2 presents exact solutions for localization of strain into an ellipsoidal region contained within an infinite solid. The solution exploits Eshelby's theorem for eigenstrains in elliptical inclusions in an infinite elastic solid. The special cases of localization of strain into a spherical region in three dimensions and into a circular region in two dimensions are solved for finite bodies – spheres in three dimensions and circles in two dimensions. The solutions show that even if the body is infinite, the localization instability cannot occur at the start of strain softening, i.e., at the state corresponding to the peak of the stress strain diagram, but only later, when there is a finite negative strain softening slope. If the size of the body relative to the size of the softening region is decreased and the boundary is restrained, homogeneous strain softening remains stable into a larger strain. It is also found that, generally, localizations into elongated ellipsoidal region occur earlier, and the body in fact prefers to localize its strain into an infinite band. However, localization in a band is possible only in infinite bodies. For a finite body, it is not possible to satisfy the boundary conditions where the localization band runs into the boundary. In such a case ellipsoidal localizations with an ellipsoidal region that is much smaller than the dimension of the body represent a possible solution which approximately satisfies all the boundary conditions because stress disturbances decay exponentially with a distance from the ellipsoid.

The solutions to the localization instabilities can be used as checks for finite element programs for strain softening.

The localization solutions in Appendix D2 are limited to isotropic material models. Appendix D3 extends the analytical solutions to more complex constitutive laws, particularly to von Mises and Drucker-Prager plasticity with a degrading yield limit. It is shown how the conditions of onset of localization instability depend on the aspect ratios of the ellipsoid, and on the material properties, especially the friction angle and the dilatancy ratio. It is also shown that violation of normality promotes localization.

The solutions in Appendices D1-D3 deal only with localization instabilities but not with localization that happens as a bifurcation along a stable equilibrium path. That problem is

treated separately. It is shown that the bifurcation occurs as soon as the tangential moduli matrix becomes singular – a condition that coincides with Hill's classical bifurcation condition for localization into an infinite layer. It is shown that the bifurcation is normally of Shanley type, occurring in absence of neutral equilibrium while the controlled displacements at infinity increase. During a loading process in which the displacement increase is controlled at infinity, this type of bifurcation precedes the loss of stability of equilibrium due to an ellipsoidal localization mode. An exception, however, is the case when the tangential moduli change suddenly (which can happen, e.g., when the slope of the stress-strain diagram is discontinuous or when temperature is increased).

5. THERMODYNAMIC ANALYSIS OF STABLE PATH

In the analysis of static propagation of crack systems of damage zones in structures, one typically encounters equilibrium paths with multiple bifurcations and loss of uniqueness. Among the various possible post-bifurcating path, it is by no means guaranteed that the commonly used solution algorithms in finite element programs yield the correct path that actually occur. To obtain the correct solution, a thermodynamically based stability analysis of stable path is necessary. Such a theory is presented in Appendix E1. Stability of inelastic structures is analyzed in general on the basis of the second law of thermodynamics, and criteria for stable states as well as stable path are given. Novel results are the criteria in Eqs. 11-14 of E1, pertaining to the conditions or controlled displacements of controlled load at isothermal or isentropic conditions. These criteria reduce the stability analysis as well as the determination of the correct bifurcation path to the analysis of positive definiteness of the surface of the internally produced entropy increment (see Eq. 4 and 7 in Appendix E1).

The conditions for stable path have been applied to determination of the evolution of a system of interacting cracks. There are many problems where all the cracks in a crack system can grow simultaneously, or one grows preferentially as the others close. The latter behavior is an aspect of localization within the crack system. The conditions under which such localizations of crack growth into a single crack happen in some typical simple crack systems have been formulated in Appendix E1.

6. NUMERICAL AND ANALYTICAL STUDIES OF LOCALIZATION

Localization has been studied in the setting of viscoplastic materials by both analytical and numerical methods. The major findings of this study are:

1. The size of the localization band depends on the imperfection in viscoplastic materials (Appendix F1).

2. The size of the localization band in material models with localization limiters of gradient or nonlocal type depends on the parameter governing the higher-order term or the size of the nonlocal domain, respectively.

3. Finite element solutions with constant strain elements in which the imperfection is introduced by weakening one element provide unstable (and hence physically unattainable) solutions (Appendix F2).

4. A spectral overlay method has been developed wherein the region near a localization zone is treated by a spectral method, which provides the resolution needed to study the structure of the strain field in multi-dimensional localization (Appendix F3).

The fact that the localization band is scaled by the size of the imperfection is particularly interesting and significant since it indicates that even when a constitutive equation does not possess an inherent length scale, in application to physical problems, a length scale is introduced by the length scale of the imperfections. This also is in agreement with experiments which usually exhibit significant scatter in the size and structure of localization zones.

The new numerical method is based on the spectral interpolant, but in contrast to the spectral element method described, the finite element interpolants are not modified in this scheme. Instead, the spectral approximation is superimposed on the finite element approximation over a spectral patch which is placed over the region of large gradients. The boundaries of the spectral patch need not be coincident with the boundaries of the elements, so that high gradients inclined at an arbitrary angle to the mesh can be treated. With a

modest increase in the number of unknowns, it is possible to achieve striking resolution of the strain fields in problems such as viscoplastic shear banding.

The method originates from a desire to combine the generality of the finite element method with regard to complex shapes and boundary conditions with the infinite order convergence of the spectral method. While spectral interpolants can be introduced into finite elements in regions of high gradients, a tremendous number of unknowns is needed to achieve good resolution when the high gradient is skewed in relation to the mesh. The superposition technique described here enables one to orient the spectral resolution precisely along the high gradient region, or in other words, exactly where it is needed.

Furthermore, in most problems in nature with high gradients, only a single component becomes very large. We have taken advantage of this by using the spectral interpolant only in one direction in the spectral patch.

A variational method was used to obtain the discretized equations. As a consequence, it is only necessary that the spectral field vanish on its "long" boundaries. The continuity of tractions and any inhomogeneous natural boundary conditions are then handled by the finite element fields. Thus the method is easily applied to steep gradients which are skewed relative to the sides of the finite element mesh.

The method was applied to the problems of shear banding in a tensile specimen with several viscoplastic materials. The results show that the strain field is extremely localized in a band of an order of 0.01 of a typical element size. Thus constant strain elements would not be able to efficiently resolve the structure of this strain field. The shape of the localized strain field depends highly on the imperfection used to trigger localization.

7. CONCLUSION

Not too long ago, strain softening has been considered by mechanics theorists to be a meaningless concept. Yet experimental work kept providing clear indications of the existence of distributed damage and of various responses that could be explained only by strain softening. Therefore, in the past decade strain softening became the subject of serious studies.

one of which has been the present study supported by AFOSR. As a result of this as well as other studies, the problem of strain softening is at present understood quite well although not completely. It is known that strain softening as a material property must be associated with some characteristic length of the material, that the mathematical treatment must involve some type of localization limiter, that localization stabilities must be analyzed and stable post bifurcation paths determined, and that attention must be paid to the size effects - the principal consequence of strain softening from the practical viewpoint.

PUBLICATIONS

1. Bazant, Z. P., and Pijaudier-Cabot, G., "Measurement of Characteristic Length of Nonlocal Continuum," ASCE J. of Engineering Mechanics 115 (4), 1989, 755-767.

2. Pijaudier-Cabot, G. and Bazant, Z. P. "Dynamic Stability Analysis with Nonlocal Damage," Computers and Structures 29 (3), 1988, 503-507.

3. Bazant, Z. P., Sener, S., and Prat, P. C., " Size Effect in Torsional Failure of Plain and Reinforced Concrete Beams," Materials and Structures (RILEM), 21, 1988, 425-430.

4. Bazant, Z. P., Tabbara, M. R., and Kazemi, M. T., " Stable Path of Interacting Crack Systems and Micromechanics of Damage," Advances in Fracture Research, Proc. of 7th International Conference on Fracture, Houston, March 1989, K. Salama et al., eds., vol. 3, 2141-2152.

5. Bazant, Z. P., and Kazemi, M. T., "Brittleness and Size Effect in Concrete Structures," Advances in Cement Manufacture and Use, Engineering Foundation Conference held in Potosi, Missouri, 1988, E. Gartner, ed., pp. 23-30.

6. Bazant, Z. P., Kazemi, M. T., and Gettu, R., "Recent Studies of Size Effect in Concrete Structure," Transactions, Tenth International Conference on Structural Mechanics in Reactor Technology (SMIRT), Anaheim, CA, 1989, vol. H, pp. 85-93.

7. Lasry, D., and Belytschko, T., "Localization Limiters in Transient Problems," International Journal of Solids and Structures, 24 1988, pp. 581-597.

8. Lasry, D., and Belytschko, T., "Gradient Type Localization Limiters for Strain-Softening Materials," Advances in Inelastic Analysis, (s. Nakazawa, K. Willam, and N. Rebelo, eds), ASME-AMD Vol. 88, 1987, pp. 127-144.

9. Belytschko, T., and Lasry, D., "A Study of Localization Limiters for Strain-Softening in Statics and Dynamics," 1988, submitted to Computers and Structures.

10. Fish, J. and Belytschko, T., "A General Finite Element Procedure for Problems with High Gradients," in Computers and Structures.

11. Bazant, Z. P., and Tabbara, M. R., "Stable Propagation of Interacting Crack Systems and Modeling of Damage," ibid., vol. Q, pp. 1-8.

12. Bazant, Z. P., and Gettu, R., "Determination of Nonlinear Fracture Characteristics and Time Dependence from Size Effect," in Fracture of Concrete and Rock: Recent Developments. Proceedings of RILEM International Conference held at University of Wales, Cardiff, UK, S.P. Shah, S. E. Swartz, and Barr, B., eds., Elsevier, London, pp. 519-565.

13. Bazant, Z. P., Tabbara, M. R., Kazemi, M. T. and Pijaudier-Cabot, "Random Particle Model for Fracture of Aggregate or Fiber Composites," ASCE J. of Eng. Mech. 116, 1990, in press.

14. Bazant, Z. P., and Tabbara, M. R., "Random Particle Simulation of Damage and Fracture in Heterogeneous Materials," Proc. Int. Conf. on Num. Meth. in Engineering (NUMETA), held in Swansee, UK, Jan. 1990.

15. Bazant, Z. P., "Bifurcations and Thermodynamic Criteria of Stable Paths of Structures Exhibiting Plasticity and Damage Propagation," Computational Plasticity, Proc. of Second International Conference held in Barcelona, Spain, September 1989. D. R. J. Owen, E. Hinton, and E. Onate, eds., Pineridge Press, Swansee, UK, pp. 1-25.

16. Bazant, Z. P., "Identification of Strain-Softening Constitutive Relation from Uniaxial Tests by Series Coupling Model for Localization," Cement and Concrete Research 19, 1989, 973-977.

17. Bazant, Z. P., Sener, S. and Kim, J.-K. (1987). "Effect of cracking on drying permeability and diffusivity of concrete," ACI Materials Journal, 84 (Sept.-Oct), 351-357.

18. Bazant, Z. P. and Sener, S. (1987). "Size effect in torsional failure of concrete beams," J. of Struct. Engng. ASCE, 113, (10), 2125-2136.

19. Belytschko, T., Wang, X.-J., Bazant, Z. P., and Hyun, T. (1987). "Transient solutions for one-dimensional problems with strain-softening," Trans. ASME, J. of Applied. Mechanics ASME, 54, (3), 513-516.

20. Bazant, Z. P. and Pijaudier-Cabot, C. (1987). "Nonlocal damage theory," J. of Engng. Mechanics ASCE, 113, (10), 1512-1533.

21. Bazant, Z. P., and Pfeiffer, P. A. (1987). "Determination of fracture energy from size effect and brittleness number," ACI Materials Jour., 84, 463-480.

22. Bazant, Z. P. (1987). "Snapback instability at crack ligament tearing and its implication for fracture micromechanics," Cement and Concrete Research 17, 951-967.

23. Bazant, Z. P., and Prat, P. C. (1988). "Measurement of Mode III fracture energy of concrete," Nuclear Engineering and Design 106, 1-8.

24. Bazant, Z. P., and Pijaudier-Cabot, G. (1988). "Nonlocal continuum damage, localization instability and convergence," ASME J. of Applied Mechanics, 55, 287-293.

25. Bazant, Z. P., and Zubelewicz, A. (1988). "Strain-softening bar and beam: Exact nonlocal solution," International J. of Solids and Structures, 24 (7), 659-673.

26. Bazant, Z. P., and Prat, P. C. (1988). "Effect of temperature and humidity on fracture energy of concrete," ACI Materials Jour., 84 (July), 262-271.

27. Bazant, Z. P., and Lin, F.-B. (1988). "Nonlocal yield limit degradation," International J. for Numerical Methods in Engineering, 26, 1805-1823.

28. Pijaudier-Cabot, G., Bazant, Z. P., and Tabbara, M. (1988). "Comparison of various models for strain-softening," Engineering Computations, 5, (June), 141-150.

29. Bazant, Z. P. (1988). "Softening instability: Part I-Localization into a planar band," J. of Appl. Mech. ASME, 55, 523-529.

30. Bazant, Z. P. (1988). "Softening instability: Part II-Localization into ellipsoidal regions," J. of Appl. Mech. ASME, 55, 517-522.

31. Bazant, Z. P., and Prat, P. C. (1988). "Microplane model for brittle plastic material: I. Theory," J. of Eng. Mech. ASCE, 114, 1672-1688.

32. Bazant, Z. P., and Prat, P. C. (1988). "Microplane model for brittle plastic material: II. Verification," J. of Eng. Mech. Div. ASCE, 114, 1689-1702.

33. Bazant, Z. P., and Lin, F.-B. (1988), "Nonlocal smeared cracking model for concrete fracture," J. of Struct. Div. ASCE, 114, (11), 2493-2510.

34. Bazant, Z. P. (1988). "Stable states and paths of structures with plasticity or damage," J. of Eng. Mech. ASCE, 114 (12), 2013-2034.

35. Bazant, Z. P., and Pijaudier-Cabot, G. (1987). "Modeling of distributed damage by nonlocal continuum with local strains," Numerical Methods in Fracture Mech., (Proc. 4th Int. Conf. held in San Antonio, Texas), ed. by A. R. Luxmore et al., Pineridge Press, Swansea, U.K., 411-431.

36. Bazant, Z. P., "Recent Advances in Failure Localization and Nonlocal Models," Int. Conf on
Micromechanics of Failure of Quasibrittle Materials, held at Univ. of New Mexico. Albuquerque, in June 1990, ed. by M. L. Wang and S.-P. Shah, in press.

37. Bazant, Z. P., and Ozbolt, J., "Nonlocal microplane model for fracture, damage and size effect in structures," J. of Eng. Mech. ASCE, 116, 1990, in press.

38. Bazant, Z. P., and Ozbolt, J., "Nonlocal microplane model: tensile and compression fractures and partial damage," Proc.
Int. Conf. on Computer-Aided Design of R.C. Structures, H. Mang and N. Bicanic, held in Zell am See, Austria, Apr. 1990, Pineridge Press, Swansee, UK.

39. Bazant, Z. P., Lin, Feng-Bao, and Pijaudier-Cabot, G. (1987). "Yield limit degradation: nonlocal continuum with local strain," Computational Plasticity, (Proc., Int. Conf. held in Barcelona, Spain), ed. by D. R. J. Owen, E. Hinton and E. Onate, Pineridge Press, Swansea, U.K., 1757-1779.

40. Bazant, Z. P. (1987). "Fracture energy of heterogeneous material and similitude." Preprints, SEM-RILEM Int. Conf. on Fracture of Concrete and Rock (held in Houston, Texas, June 1987), ed. by S. P. Shah and S. E. Swartz, publ. by SEM (Soc. for Exper. Mech.) 390-402.

41. Bazant, Z. P., and Pijaudier-Cabot, G. (1988). "Nonlocal continuum damage and measurement of characteristic length," in Mechanics of Composite materials - 1988, AMD 92, ed. by G. J. Dvorak and N. Laws, Am. Soc. of Mech. Ingrs., N.Y. (Joint ASME/SES conference, Berkeley, CA), 79-85.

42. Bazant, Z. P., and Lin, F.-B. (1988). "Localization instability for softening in ellipsoidal regions and bands," ibid., 7-16.

43. Bazant, Z. P. (1989), "Stable states and stable paths of propagation of damage zones and interactive fractures," in Cracking and Damage, J. Mazars and Z. P. Bazant, eds., Elsevier, London, 183-207 (also Preprints, France-U.S. Workshop on Strain-Localization and Size Effect Due to Cracking and Damage, ed. J. Mazars and Z. P. Bazant held at ENS, Univ. Paris VI, Cachan).

44. Pijaudier-Cabot, G. and Bazant, Z. P. (1989). "Local and nonlocal models for strain-softening, and their comparison based on dynamic analysis," ibid., 379-390 (also Preprints, 1988).

45. Droz, P., and Bazant, Z. P. (1989). "Nonlocal analysis of stable states and stable paths of propagation of damage shear bands," ibid., 415-425 (also Preprints, 1988).

46. Bazant, Z. P., Prat, P. C., and Tabbara, M. R., "Antiplane Shear Fracture Tests (Mode III), ACI Materials Journal 87, (Jan.-Feb.), 1990, pp. 12-19.

47. Bazant, Z. P., and Carol, I., "Geometric damage tensor uncoupled from constitutive properties,"
Proc. Southeastern Conf. of Appl. Mech., Atlanta, Mar. 1990 (in press).

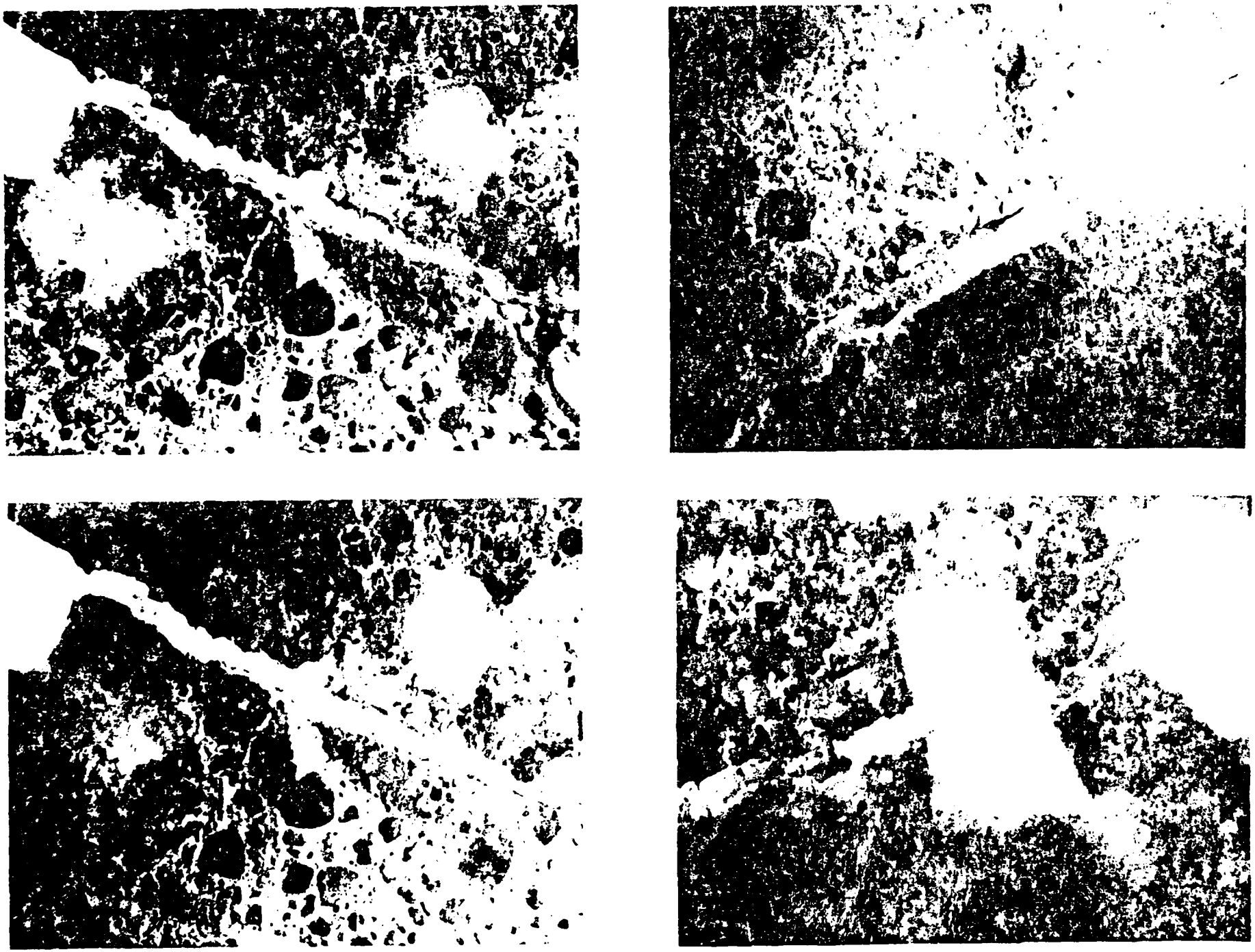


FIG. 1a

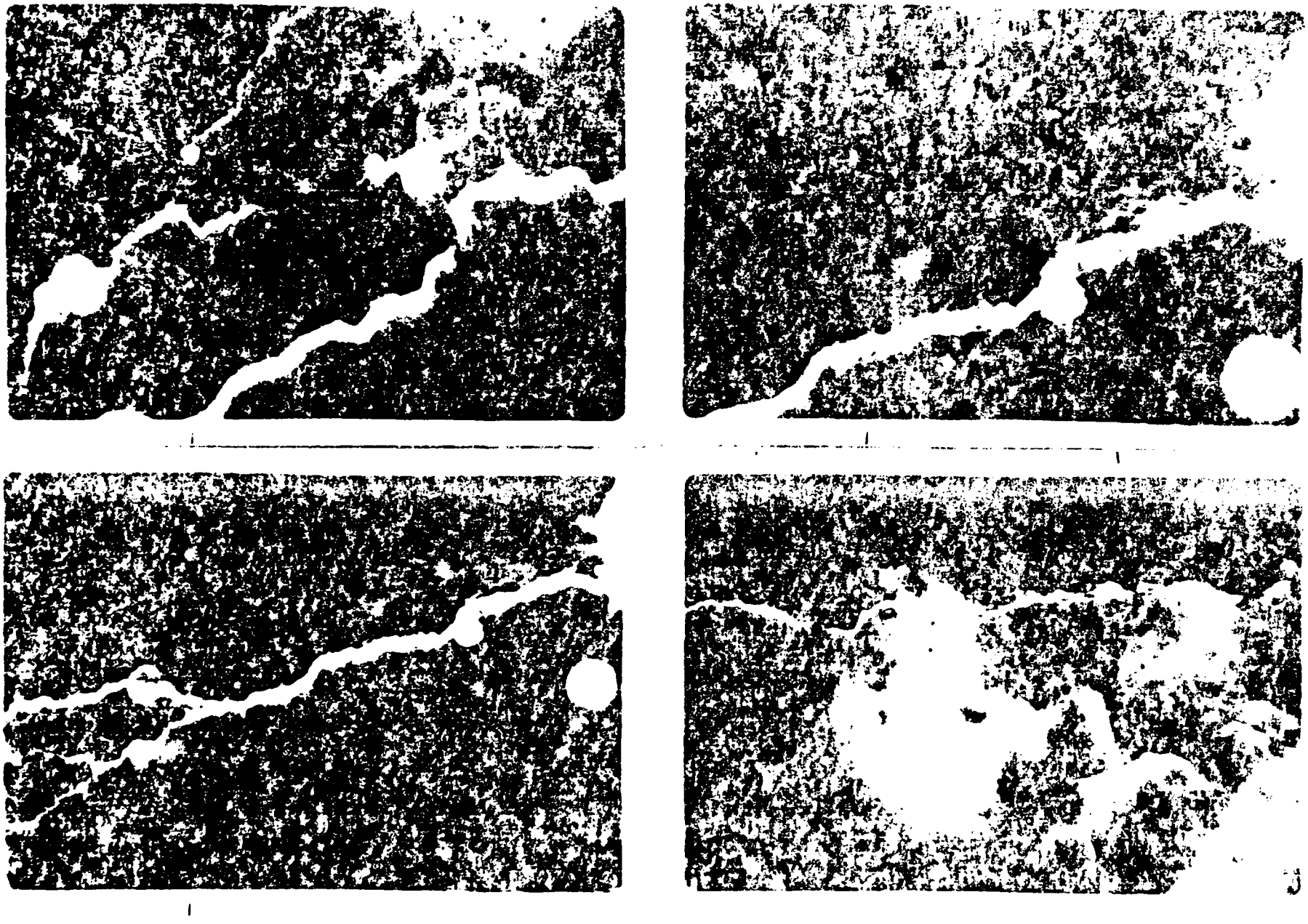


FIG. 1b

APPENDICES

ACI MATERIALS JOURNAL TECHNICAL PAPER

Title no. 84-M41

Determination of Fracture Energy from Size Effect and Brittleness Number



by Zdeněk P. Bažant and Phillip A. Pfeiffer

A series of tests on the size effect due to blunt fracture is reported and analyzed. It is proposed to define the fracture energy as the specific energy required for crack growth in an infinitely large specimen. Theoretically, this definition eliminates the effects of specimen size, shape, and the type of loading on the fracture energy values. The problem is to identify the correct size-effect law to be used for extrapolation to infinite size. It is shown that Bažant's recently proposed simple size-effect law is applicable for this purpose as an approximation. Indeed, very different types of specimens, including three-point bent, edge-notched tension, and eccentric compression specimens, are found to yield approximately the same fracture energy values. Furthermore, the R-curves calculated from the size effect measured for various types of specimens are found to have approximately the same final asymptotic values for very long crack lengths, although they differ very much for short crack lengths. The fracture energy values found from the size effect approximately agree with the values of fracture energy for the crack band model when the test results are fitted by finite elements. Applicability of Bažant's brittleness number, which indicates how close the behavior of specimen or structure of any geometry is to linear elastic fracture mechanics and to plastic limit analysis, is validated by test results. Comparisons with Mode II shear fracture tests are also reported.

Keywords: concretes; cracking (fracturing); crack propagation; dimensional analysis; energy; finite element method; measurement; specimens; tests.

The fracture energy of concrete is a basic material characteristic needed for a rational prediction of brittle failures of concrete structures. The method of experimental determination of concrete's fracture energy, and even its definition, has recently been the subject of intensive debate. Although in principle the fracture energy as a material property should be a constant, and its value should be independent of the method of measurement, various test methods, specimen shapes, and sizes yield very different results—sometimes differing even by several hundred percent.[1-7]

The crux of the matter is that the fracture of concrete, as well as brittle heterogeneous materials in general, is not adequately described by the classical idealization of a line crack with a sharp tip. In this idealization, which has been introduced for fracture with small-scale yielding in metals,[8] the fracture process is assumed to be concentrated in a zone that is so small compared to the body dimensions that it can be treated as a point. In concrete, the fracture process takes place over a relatively large fracture process zone whose size is, for the usual laboratory specimens, of the same order of magnitude as the size of the specimen itself. In other words, concrete is a material characterized by blunt fracture. The tip of the large visible crack is blunted by a zone of microcracking that lies ahead of the crack tip and is certainly rather long and possibly also quite wide relative to the size of material inhomogeneities.

The material behavior in the fracture process zone may be described by a strain-softening stress-strain relation or, to some extent equivalently, by a stress-displacement relation with softening that characterizes the fracture process zone over its full width. These mathematical models indicate that the fracture energy is not the only controlling parameter and that the size and shape of the fracture process zone as well as the shape of the softening stress-strain diagram have a significant influence. This is no doubt the source of difficulties and raises a basic question: can the fracture energy, the most important fracture characteristic of the material, be defined and measured in a way that is unaffected by the other influences?

The purpose of the paper is to show that this question can be answered in the affirmative. The key is the size effect—the simplest and most fundamental manifestation of the fracture mechanics aspect of failure. As we will see, if geometrically similar specimens are considered and the failure load is correctly extrapolated to a specimen of infinite size, the fracture energy obtained must be unique and independent of specimen type, size,

Received Oct. 1, 1986, and reviewed under Institute publication policies. Copyright © 1987, American Concrete Institute. All rights reserved, including the making of copies unless permission is obtained from the copyright proprietors. Pertinent discussion will be published in the September-October 1988 *ACI Materials Journal* if received by June 1, 1988.

Zdeněk P. Bazant, FACI, is a professor at Northwestern University, Evanston, Ill., where he recently served a five-year term as director of the Center for Concrete and Geomaterials. Dr. Bažant is a registered structural engineer and is on the editorial boards of a number of journals. He is Chairman of ACI Committee 446, Fracture Mechanics; a member of ACI Committees 209, Creep and Shrinkage in Concrete; and 348, Structural Safety; and a fellow of ASCE, RILEM, and the American Academy of Mechanics; and Chairman of RILEM's Creep Committee and of SMiRT's Concrete Structures Division.

Phillip A. Pfeiffer is a civil engineer at Argonne National Laboratory in the Reactor Analysis and Safety-Engineering Mechanics Program. Dr. Pfeiffer was formerly a graduate research assistant at Northwestern University, conducting both theoretical and experimental research in fracture mechanics applications and size effect in failure of concrete and aluminum.

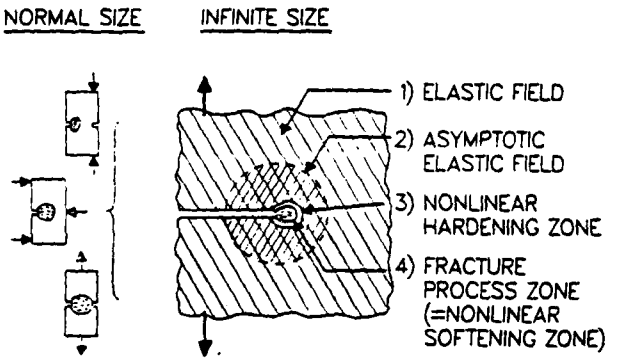


Fig. 1—Fracture process zone for normal size laboratory specimens and for extrapolation to infinity

and shape because the fracture process zone in the limit becomes vanishingly small compared to the specimen or structure dimensions.

With this asymptotic approach, the argument is reduced to finding and applying the correct size-effect law. Difficult as the precise answer to this question might seem, satisfactory results may nevertheless be achieved with an approximate size-effect law that was recently derived by Bažant on the basis of dimensional analysis and similitude arguments.[9-15] Bažant's law was approximately confirmed by comparisons with test data from fracture specimens,[11,16] as well as certain brittle failures of concrete structures.[17-22] Due to its approximate nature, the applicability range of Bažant's law is limited to a size range of perhaps 1:20, while for a much larger size range, e.g., 1:200, a more accurate and more complicated size-effect law would no doubt be required.

Since determination of fracture energy is of interest for finite element programs, we need to demonstrate also that approximately the same values of fracture energy are obtained when the fracture energy value is optimized to achieve the best fit of fracture test data with a finite element program.

DEFINITION OF FRACTURE ENERGY BY INFINITE SIZE EXTRAPOLATION

The size effect can be isolated from other influences if we consider geometrically similar specimens or structures. As mentioned, for normal-size fracture specimens, the fracture process zone is of the same order of magnitude as the specimen size (see Fig. 1). As already established theoretically as well as experimentally, the fracture process zone size is essentially determined by the size of material inhomogeneities, e.g., the maximum aggregate size. Therefore, the fracture process zone must become infinitely small compared to the specimen if an extrapolation to an infinite size is made, as shown in Fig. 1. (This in fact achieves conditions for which the small-scale yielding approximation used for metals[8,23] becomes valid.) Furthermore, due to the rather limited plasticity of concrete under tensile loadings, the hardening nonlinear zone surrounding the fracture process zone in concrete is rather small, and the boundary of the nonlinear zone lies very close to the boundary of the fracture process zone (see Fig. 1). Under this condition, the failure of an infinitely large specimen must follow linear elastic fracture mechanics. Based on this fact, it was shown[16] (see Appendix) that the fracture energy of concrete may be calculated as

$$G_f = \frac{g_f(\alpha_0)}{A E_c} \tag{1}$$

in which E_c = Young's elastic modulus of concrete; A = slope of the size effect regression plot for failure of geometrically similar specimens of very different sizes,[16] which will be explained later; and $g_f(\alpha_0)$ = nondimensional energy release rate calculated according to linear elastic fracture mechanics, which is found for typical specimen shapes in various handbooks and textbooks[8,23,24] and can always be easily determined by linear elastic finite element analysis; α_0 = relative notch length = a_0/d where a_0 = notch length and d = cross-section dimension.

Since the fracture energy is determined in this method from the size-effect law, its value is, by definition, size independent. This overcomes the chief obstacle of other methods, e.g., the RILEM work of fracture method,[4,25] which are plagued by a strong dependence of the measured values on the specimen size. However, another question arises in the size-effect approach: are the G_f values independent of the specimen type or geometry?

They must be independent. When the structure is infinitely large and the fracture process zone as well as the nonlinear zone are negligibly small compared to the specimen or structure size, nearly all the specimen is in an elastic state. Now it is well known from classical fracture mechanics that the asymptotic elastic stress-strain field near the crack tip is the same regardless of specimen geometry and the type of loading. This field is known to have the form $\sigma_{ij} = r^{-1/2} \phi_{ij}(\theta)$ in which σ_{ij} = stress components, r = radial coordinate from the crack tip, θ = polar angle, and ϕ_{ij} = certain functions listed in textbooks.[8,23] Therefore, the nonlinear zones in infinitely large specimens of all types are exposed on their entire boundary to exactly the same boundary stresses. It follows (assuming uniqueness of response) that the entire stress-strain field within the nonlinear zone, including the fracture process zone, must be the

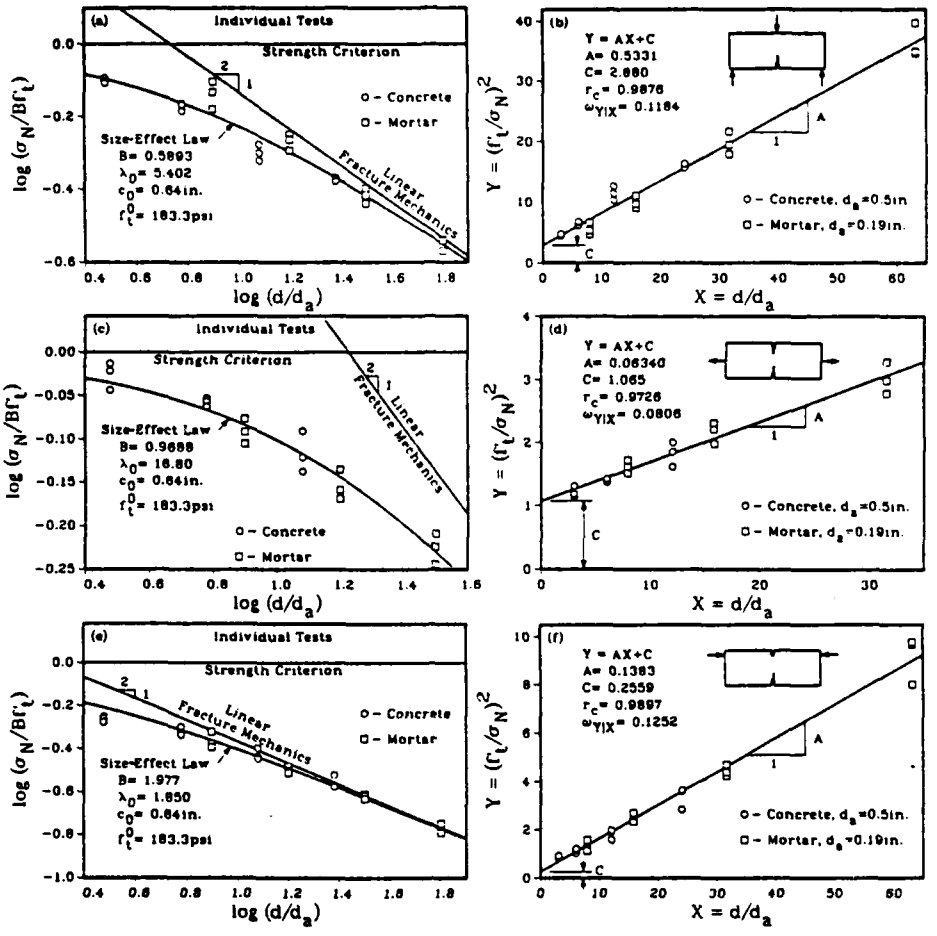


Fig. 2—Size-effect plot (a, c, e) and linear-regression plot (b, d, f) constructed from the maximum load values measured for three-point bent, edge-notched tension, and eccentric compression specimens of concrete and mortar for various specimen sizes

same for all specimen types, and so the energy that is dissipated in the fracture process zone per unit advance of the crack tip must be unaffected by the specimen or structure shape. Consequently, the following definition of fracture energy must give unique results independent of size as well as specimen type: *The fracture energy G_f of a microscopically heterogeneous brittle material is the specific energy required for crack growth in an infinitely large specimen.*

The extrapolation to infinite size and the asymptotic considerations for the stress field could of course be mathematically formulated by transformations of scale in the formulation of the nonlinear boundary value problem.

SIZE-EFFECT LAW

The fracture energy determination would be exact if we knew the exact form of the size-effect law to be used for extrapolation to infinite size. Unfortunately, we know this law only approximately, and the question is whether the approximate form is sufficiently accurate, and if so, over what size range? The simplest form of the size-effect law results from dimensional analysis and similitude arguments[11,14,15] if it is assumed that either the width or the length of the fracture process zone is a constant material property. This form is

$$\sigma_N = Bf_t' \left[1 + \left(\frac{d}{\lambda_0 d_a}\right)^r\right]^{-\frac{1}{r}} \qquad (2)$$

in which $r = 1$ according to the initial proposal,[11] σ_N = nominal strength at failure = P/bd where P = the maximum load, b = specimen thickness, and d = characteristic dimension of the specimen or structure (only geometrically similar specimens are considered); f_t' = strength parameter, which may be taken as the direct tensile strength; d_a = the maximum aggregate size; and B, λ_0 = two empirical constants to be determined by fitting test results for geometrically similar specimens of various sizes. Application of Eq. (2) to various types of brittle failures in concrete structures was demonstrated in References 17 through 22.

Note that for sufficiently small sizes d, the second term in the bracket is negligible compared to 1. This means that σ_N is proportional to the material strength or yield limit, and this represents the failure condition of plastic limit analysis, characterized by no size effect. In the plot of log σ_N versus log d, this failure condition is represented by a horizontal straight line [see Fig. 2(a)].

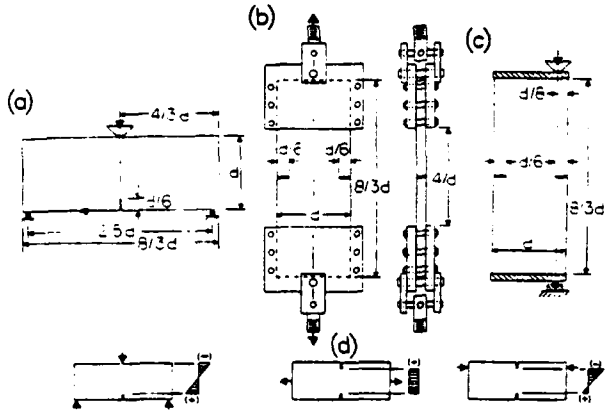


Fig. 3—Specimen geometry and loading detail for (a) three-point bent, (b) edge-notched tension, and (c) eccentric-compression specimens

As another limiting case when size d is extremely large, the number 1 in the bracket is negligible compared to the second term, and then σ_N is proportional to $d^{-1/2}$. This represents the strongest possible size effect and corresponds to the classical linear elastic fracture mechanics. In the plot of log σ_N versus log d, this limiting case projects itself as a straight line of downward slope $-1/2$. The plot of the size-effect law [Eq. (2)] consists of a gradual transition between these two limiting cases—plastic limit analysis for very small structures and linear elastic fracture mechanics for very large structures. The limiting cases are known exactly, and the question is the shape of the transition, which is the purpose of Eq. (2).

The size-effect law in Eq. (2) results by dimensional analysis from the hypothesis that the total energy release W of the structure caused by fracture is a function of: (1) the length a of the fracture, and (2) the area nd_a of the cracking zone, such that nd_a = width of the front of the cracking zone = constant (d_a = maximum aggregate size and $n \approx 1$ to 3 = empirical number). The original derivation[11] was simplified by truncation of the Taylor series expansion of a certain function. If this truncation is not made, a more general size-effect law is obtained[12-15]

$$\sigma_N = Bf_t' \left(C_0\xi^{-1} + 1 + C_1\xi + C_2\xi^2 + C_3\xi^3 + \ldots\right)^{-\frac{1}{r}} \quad \xi = \left(\frac{d_a}{d}\right)^r \tag{3}$$

in which B, C_0, C_1, C_2, . . ., r = empirical constants. In practice, however, no case where this more general form would be needed has yet been found. It appears, and our analysis of test results will confirm it, that Eq. (2) can fit quite well any existing data with $r = 1$. The size range of the existing test data is at most 1:10. Coefficient r might be needed for a broader size range, but this would be meaningful only if the statistical scatter were smaller than it normally is for concrete, since otherwise coefficient r cannot be determined unambiguously. Note that Eq. (1) is valid only for $r = 1$; for other r, see the Appendix.

Scatter-free values of maximum loads for geometrically similar structures of different sizes can be generated with a finite element program based on fracture mechanics. In Reference 13 it was shown that for a certain r-value, Eq. (2) can fit such results even if the size range is as broad as 1:500. However, this range is much broader than feasible to test in a laboratory and needed in practice. Alternatively, scatter-free results can be calculated for the line crack model of Hillerborg if the method of Green's function is used (private communication by Planas and Elices, June 1986). Such calculations, as well as similar calculations made by Rots,[27,28] Darwin, Hillerborg and others,[26] show that the response in general, and the shape of the size-effect curve in particular, are sensitive to the precise shape of the softening stress-strain diagram or stress-displacement diagram. Various possible shapes of this diagram yield different extrapolations to infinite size. From the practical viewpoint, though, this fact does not seem to pose a serious problem. We do not need to extrapolate to infinity in the mathematically true sense of the word. We need only to extrapolate to a specimen size that is sufficiently larger than the fracture process zone, and Eq. (1) seems sufficient for that.

Eq. (2), as well as Eq. (3), is valid for geometrically similar structures of specimens made of the same material. This implies the use of the same maximum aggregate size d_a. When d_a is also variable, one needs to introduce the following adjustment of strength[12,15,26]

$$f_t' = f_t^0 (1 + \sqrt{c_0/d_a}) \tag{4}$$

in which f_t^0 and c_0 are empirical constants.

TEST RESULTS

As already mentioned, the validity of fracture energy determination through the size-effect law [Eq. (1)] requires that different types of specimens must yield roughly the same results if made from the same concrete. In the previous study[16] where G_f was first determined from the size effect, test results for only one specimen type—the three-point bent specimen—were used. Therefore, the testing has been expanded to include also other specimen types such as the double-notched direct tension specimen sketched in Fig. 3(b) and the eccentric compression specimen shown in Fig. 3(c).

Fig. 3(d) shows the stress distribution across the ligament cross section drawn according to the bending theory. Although such stress distributions are unrealistic, they nevertheless reveal the great differences in the type of loading for the ligament cross section. For the notched tension specimen, the entire ligament is subjected to tension, which causes the fracture process zone to become very large. The opposite extreme is obtained for the eccentric compression specimen, for which the major part of the ligament is subjected to

compression and only a small part to tension [Fig. 3(d)]. In this case, the compression ahead of the crack that forms in the tensile zone prevents the fracture process zone from becoming very large. The bent specimen is a medium situation, in which roughly half of the cross section is subjected to tension and half to compression, and the fracture process zone is of medium size.

Thus we see that the choice of these three specimens covers the entire broad range of possibilities for the type of loading in the ligament cross section. Previously, rather different results for the fracture energy have been found for these specimens when the conventional methods were used.

Photographs of the test specimens are shown in Fig. 4(a), (b), and (c). Despite the very different types of loading for the ligament cross section, it was possible to use specimens of the same geometry except for the notches. All the specimens were of the same external shape as previously used in a shear fracture study.[39] The cross sections of the specimens were rectangular, and the length-to-depth ratio was 8:3 for all specimens (Fig. 3). The cross-sectional heights of the specimens were d = 1.5, 3, 6, and 12 in. (38.1, 76.2, 152.4, and 304.8 mm) (see Fig. 3). The thickness of the bending and compression specimens was b = 1.5 in. (38 mm) and that of the tension specimens b = 0.75 (19 mm), as shown in Fig. 3(b). For each specimen size and each type, three specimens were produced.

These three specimens were from different batches; however, from each batch of concrete or mortar one specimen of each size was cast. Notches of depth $d/6$ and thickness of 0.1 in. (2.5 mm) (same thickness for all specimen sizes) were cut with a diamond saw into the hardened specimens. The specimens were cast with the side of depth d in a vertical position. The concrete mix had a water-cement ratio of 0.6 and cement-sand-gravel ratio of 1:2:2 (all by weight). The maximum gravel size was d_a = 0.5 in. (12.7 mm), and the maximum sand grain size was 0.19 in. (4.83 mm). Mineralogically, the aggregate consisted of crushed limestone and siliceous river sand. Aggregate and sand were air-dried prior to mixing. Portland cement C 150, ASTM Type I, with no admixtures, was used.

To get information on the effect of aggregate size, a second series of specimens was made of mortar, whose water-cement ratio was 0.5 and cement-sand ratio was 1:2. The same sand as for the concrete specimens was used with the gravel omitted, and so the maximum aggregate size was d_a = 0.19 in. (4.83 mm). The water-cement ratio differed from that for concrete specimens to achieve approximately the same workability.

To determine the strength, companion cylinders of 3 in. diameter (76.2 mm) and 6 in. length (152.4 mm) were cast from each batch of concrete or mortar. After the standard 28-day moist curing, the mean compression strength was f'_c = 4865 psi (33.5 MPa), with a standard deviation (S.D.) = 550 psi (3.79 MPa) for the concrete specimens and 6910 psi (47.6 MPa) with S.D. = 207 psi (1.43 MPa) for the mortar specimens. Each

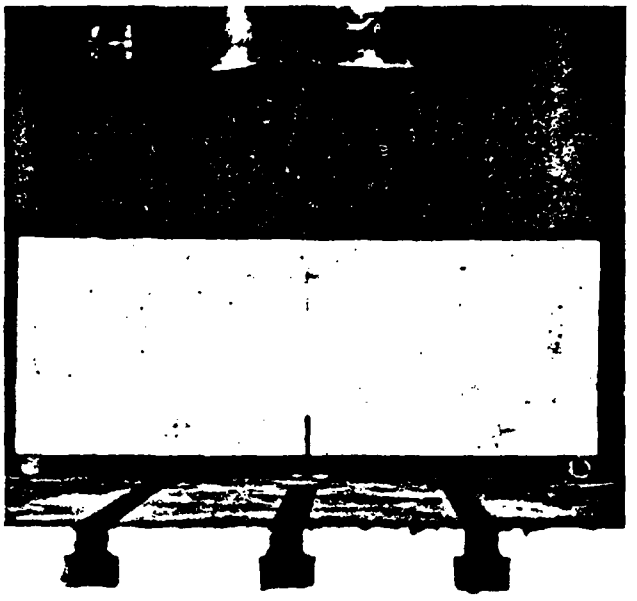


Fig. 4(a)—Three-point bent specimen of 6-in. (152-mm) depth

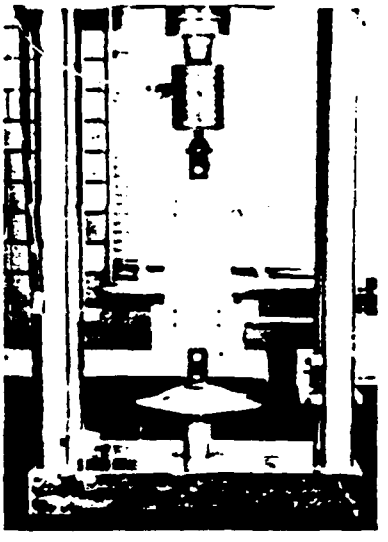


Fig. 4(b)—Edge-notched tension specimen of 6-in. (152-mm) depth

of these values was determined from three cylinders for each specimen type (see Table 1). The tensile strength was estimated as $f'_t = 6\sqrt{f'_c}$ psi and Young's modulus as $E_c = 57{,}000\sqrt{f'_c}$ psi, with f'_c in psi (1 psi = 6895 Pa).

The specimens were removed from their plywood forms one day after casting and were subsequently cured for 27 days (± 1 day) until the test in a moist room of 95 percent relative humidity and 78 F (25.6 C) temperature. All the specimens were tested in a 10-ton (89-kN) [Fig. 4(b)] or 60-ton (534-kN) [Fig. 4(d)] servo-controlled closed-loop MTS testing machine. The laboratory environment had a relative humidity of about

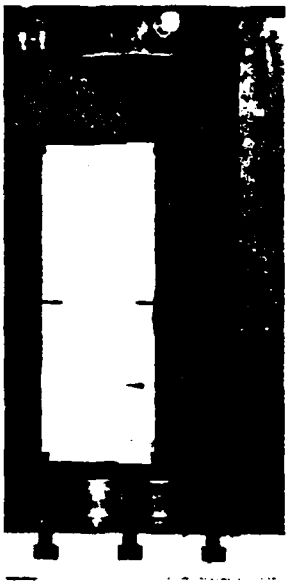


Fig. 4(c)—Eccentric-compression loadings specimen of 6-in. (152-mm) depth



Fig. 4(d)—Testing machine with 12-in. (305-mm) compression specimen

65 percent and temperature of about 78 F (25.6 C), and the specimens were exposed to this environment approximately three hours before the start of the test.

The loading for the three-point bent specimens, as shown in Fig. 4(a), applied three concentrated loads onto the specimen—one load through a hinge and two through rollers. The steel surfaces were carefully machined so as to minimize the friction of the rollers. The eccentric compression specimens were loaded in a vertical position by two hinges, as shown in Fig. 4(c). The loads were applied through steel plates 0.5 in. (12.7 mm) thick and glued to the specimen. For the eccentric compression specimens, the load was placed as close to the specimen corner as feasible but not so close that the specimen would fail by shearing off its corner rather than by cracking from the notch; by experimenting, the minimum possible distance was found to be $d/8$.

Table 1—Measured compressive strengths and estimated tensile strength and Young's elastic modulus for concrete and mortar, for all tests

Cylinder for loading	Individual tests, f'_c (psi)			Mean f'_c, psi	f'_t, psi	E_c, ksi
	1	2	3			
Concrete:						
Three-point bent	4798	4893	5160	4950	422	4010
Notched tension	4100	4190	4358	4216	390	3701
Eccentric compression	5218	5460	5610	5429	442	4200
Mortar:						
Three-point bent	6952	7000	7114	7022	503	4776
Notched tension	6604	6638	6952	6731	492	4676
Eccentric compression	6733	7003	7198	6978	501	4761

Note: 1 ksi = 1000 psi = 6.895 MPa.

A special loading grip was produced for the tensile specimens. It consisted [Fig. 4(b) and 3(b)] of a set of two aluminum plates compressed together by bolts. Sheets of hard rubber of 1-mm thickness were placed between the specimen surfaces and the plates to distribute the clamping evenly. The aluminum plates were designed so they could provide grips for the specimens of all the sizes. For the tension specimens the largest size was omitted, i.e., only specimens of d = 1.5, 3, and 6 in. (38.1, 76.2, and 152.4 mm) were used. A universal joint was provided at the top and bottom connections to the testing machine to minimize in-plane and out-of-plane bending effects.

The specimens were loaded at constant displacement rate. For each specimen size the displacement rate was selected to achieve the maximum load in about 5 min($\pm$ 30 sec).

ANALYSIS OF TEST RESULTS BY SIZE-EFFECT LAW

The test results are plotted as the square and circular data points in Fig. 2. The graphs of log σ_N versus log d, nondimensionalized with respect to f'_t and d_a, are on the left [Fig. 2(a), (c), and (e)]. At the right [Fig. 2(b), (d), and (f)], the same test results are shown in linear regression plots, based on the fact that Eq. (2) can be algebraically rearranged to a linear form. Instead of the plot $Y = AX + C$ with $X = d^r$, $Y = \sigma_N^{-2r}$, $A = C(\lambda_0 d_a)^{-r}$, $C = (B f'_t)^{-2r}$, it is more convenient to use the nondimensional plot of $Y' = A'X' + C'$ in which

$$X' = (d/d_a)^r,\ Y' = (f'_t/\sigma_N)^{2r}, \qquad (5)$$
$$C' = B^{-2r},\ A' = C'\lambda_0^{-r} = Af'^{2r}_t d_a^r$$

The value of r = 1 is used in Fig. 2. If Eq. (2) was followed exactly, the plot of Y' versus X' should be a straight line of slope A' and Y'-intercept C' (Fig. 2), which is why the deviations from this line represent statistical scatter. The plot of Y' versus X' has the advantage that one can apply linear statistical regression, yielding the size-effect law [Eq. (2) and (5)] as the regression line. The vertical deviations from the regression line may be characterized by the coefficient of

Table 2—Optimum values obtained by (a) linear regression and by (b) Levenberg-Marquardt algorithm

Loading	Material	C, psi^{-2} × 10^{-6}	A, in.$^{-1}$ psi^{2} × 10^{-6}	G_f, lb/in.	$\omega_{\gamma|X}$, %	$\omega'_{\gamma|X}$, %	ω_A, %	$\frac{\omega_{\gamma|X}}{\omega_\gamma}$, %	ω_G, %
a)									
Three-point bent	Concrete	22.34	7.253	0.219	10.90	5.56	6.81	14.60	7.07
Notched tension	Concrete	6.450	0.9134	0.205	7.89	3.42	14.84	13.91	14.92
Eccentric compression	Concrete	2.861	1.573	0.254	13.14	5.21	7.02	17.60	7.25
Three-point bent	Mortar	4.759	10.94	0.122	9.51	7.11	4.14	12.74	4.18
Notched tension	Mortar	4.439	1.133	0.131	7.74	3.67	10.22	13.64	10.44
Eccentric compression	Mortar	0.8600	2.726	0.129	11.28	6.60	4.81	15.10	5.09
b)									
Three-point bent	Concrete	17.75	8.269	0.192	13.15	4.04	7.21	17.61	7.45
Notched tension	Concrete	6.501	0.8888	0.211	7.92	3.41	15.32	13.97	15.40
Eccentric compression	Concrete	3.337	1.435	0.279	14.40	4.68	8.43	19.29	8.63
Three-point bent	Mortar	3.163	11.25	0.119	9.74	6.73	4.12	13.04	4.16
Notched tension	Mortar	4.291	1.174	0.126	7.80	3.62	9.95	13.76	10.17
Eccentric compression	Mortar	1.023	2.685	0.131	11.34	6.54	4.91	15.19	5.19

Note: 1 lb = 4.448 N, psi = 6.895 Pa.

Table 3—Results of least-square optimization when specimens with concrete and mortar are analyzed simultaneously

Loading	C'	A'	G_f Concrete, lb/in.	G_f Mortar, lb/in.	$\omega_{\gamma|X}$	$\omega'_{\gamma|X}$	ω_A	$\frac{\omega_{\gamma|X}}{\omega_\gamma}$	ω_G
Three-point bent	2.880	0.5331	0.229	0.129	0.1184	0.0982	0.0338	0.1283	0.0537
Notched tension	1.065	0.0634	0.210	0.118	0.0806	0.0375	0.0597	0.1060	0.0728
Eccentric compression	0.2559	0.1383	0.233	0.132	0.1252	0.0844	0.0309	0.1356	0.0519

Note: $G_f = g_f(\alpha_o) f_t'^2 d_a / A' E_c$.

variation $\omega_{\gamma|X}$ and the correlation coefficient r_c, both indicated in Fig. 2. The value of A in Eq. (1) represents the slope of the regression line in the plot of $Y = \sigma_N^{-2r}$ versus $X = d^r$, and from Eq. (5) it is readily found that $A = A' f_t'^{-2r} d_a^{-r}$.

By plotting the test results for concrete specimens and mortar specimens in the same diagram (Fig. 2), we are able to extend the size range of the data. Due to this fact, the plots are based on the generalized Eq. (2), which includes the effect of maximum aggregate size [Eq. (4)].

As for the value of exponent r, the overall optimum was found to be 0.954. However, for $r = 1$ the coefficient of variation of the deviations was only slightly larger. Therefore the value $r = 1$ is used for the sake of simplicity in all the plots as well as the numerical tables below.

Numerically, the test results are summarized in Tables 2 and 3. Tables 2(a) and (b) show the statistical results in which the concrete specimens and mortar specimens are treated separately, i.e., Eq. (2) is used without Eq. (4). Table 3 shows the statistical results when the data for all concrete as well as mortar specimens are treated collectively using both Eq. (2) and (4).

Table 2(a) shows the statistical results calculated separately for concrete specimens and mortar specimens and also separately for each specimen type. The optimization, based on linear regression, minimizes the sum $\Sigma\Delta(\sigma_N^{-2})$ where Δ stands for the difference between the measured value and the value according to the equation. The coefficient of variation of these deviations is $\omega_{\gamma|X}$. The table also lists the coefficient of variation $\omega'_{\gamma|X}$, which refers to the deviations in terms of σ_N.

Another possibility is to optimize the fit in terms of the deviations in σ_N, in which case the optimization problem is nonlinear. The fits, however, can be easily obtained with the standard computer library subroutine based on the Levenberg-Marquardt algorithm. The results are shown in Table 2(b).

The objective of our analysis is the value of G_f, which can be found by Eq. (1). Values of $g_f(\alpha_0)$ for the present three-point bent, notched tension, and eccentric compression geometries are 6.37, 0.693, and 1.68, respectively, as calculated by linear elastic finite element analysis. The E_c value is listed in Table 1 for each loading case for concrete and mortar. The G_f values in Table 2(a) for concrete, i.e., 0.219, 0.205, and 0.254 lb/in. (38.4, 35.9, and 44.5 N/m), show statistical scatter with a coefficient of variation of 11 percent. In Table 2(b), the resulting G_f values for concrete for the three specimen types are 0.192, 0.211, and 0.279 lb/in. (33.6, 37.0, and 48.9 N/m), and their coefficient of variation

Table 4—Measured maximum loads

Type of test	Depth d, in.	Maximum load P, lb			Mean P, lb
		1	2	3	
Three-point bent					
Concrete	1.5	405	408	417	410
	3.0	677	706	711	698
	6.0	990	1040	1096	1042
	12.0	1738	1739	1773	1750
Mortar	1.5	456	508	543	502
	3.0	703	752	777	744
	6.0	1005	1059	1104	1056
	12.0	1484	1582	1588	1551
Notched tension					
Concrete	1.5	385	405	413	401
	3.0	738	748	754	747
	6.0	1242	1290	1382	1305
Mortar	1.5	445	459	475	460
	3.0	768	786	830	795
	6.0	1292	1353	1401	1348
Eccentric compression					
Concrete	1.5	920	956	975	950
	3.0	1604	1645	1746	1665
	6.0	2500	2538	2791	2610
	12.0	3695	3711	4189	3865
Mortar	1.5	936	972	1104	1004
	3.0	1427	1458	1530	1472
	6.0	2156	2232	2272	2220
	12.0	2992	3010	3308	3103

Note: 1 lb = 4.448 N; 1 in. = 25.4 mm.

is 20 percent. We see that the optimization by linear regression gives somewhat less scattered results, and therefore we will prefer it from now on.

It is hard to decide whether the differences in the G_f values are purely random or whether they represent systematic differences between various specimen types. By extending the size range of the tests, random scatter can be minimized much more effectively than by increasing the number of specimens. Due to the cost of testing very large specimens, it has been decided to extend the size range in relative terms, i.e., in terms of d/d_a, by analyzing simultaneously the results for concrete and mortar specimens. The results of such statistical analysis, based on minimizing again the sum $\Sigma\Delta(\sigma_N^{-2})$, are shown in Table 3. In carrying out the statistical regression, parameter c_0, governing solely the effect of aggregate size, has been required to be the same for all three specimen types; its value came out to be c_0 = 0.64 in. (16.3 m). The coefficient of variation $\omega_{Y|X}$ in Table 3 characterizes the deviations in terms of the variable $Y' = [f_t^0(1 + \sqrt{c_0/d_a})/\sigma_N]^2$, and the coefficient of variation $\omega'_{Y|X}$ characterizes the deviations in terms of the variable $Y = \sigma_N$. In the absence of a direct tensile test, the values of f_t' for concrete and mortar were estimated from the ACI formula $f_t' \approx f_t^* = 6\sqrt{f_c'}$ where f_c' = average compressive strength = 4865 psi (33.5 MPa) and 6910 psi (1,120 MPa) for concrete and mortar, respectively. Then f_t^0 was determined so that the sum of squares of the difference of $f_t^0(1 + \sqrt{c_0/d_a}) - f_t^*$ for concrete and mortar be minimized; this yielded f_t^0 = 183.3 psi (1.26 MPa). Also, E_c was estimated for concrete and mortar from ACI formula $E_c = 57{,}000\sqrt{f_c'}$ as 3976 ksi (27.4 MPa) and 4738 ksi (32.7 MPa), respectively.

The principal results in Table 3 are the G_f values. For concrete they are 0.229, 0.210, and 0.233 lb/in. (40.1, 36.8, and 40.8 N/m), which gives the coefficient of variation of 5.5 percent. We see that these results are indeed less scattered than those obtained for the size range of concrete specimens alone. For mortar, the G_f values are 0.129, 0.118, and 0.132 lb/in. (22.6, 20.7, and 23.1 N/m), which give the coefficient of variation of 5.8 percent. Since the extension of the range of relative sizes reduces the relative scatter of results, it appears that the scatter is probably indeed of random origin, rather than due to omission of some unknown systematic influence, so the three types of specimens appear to give indeed approximately the same fracture energy value. This is of course to be expected theoretically, but only if the correct size-effect law is known. Thus we may conclude that Eq. (2) with Eq. (4), where $r = 1$, seems to be an acceptable approximate form of the size-effect law.

The measured maximum load values for all the individual specimens from which the present results were calculated are summarized in Table 4.

Our finding that the three fundamentally different specimen types yield roughly the same fracture energy is the principal result of the present study. Nevertheless, these results should eventually be subjected to a closer scrutiny using a much broader size range (much larger funds, of course, would also be needed).

It is interesting to determine the optimum fits of the data under the restriction that the G_f value be the same for all three specimen types. For this purpose, Eq. (1) with G_f from Eq. (2) may be algebraically rearranged to the linear plot

$$Y'' = X''/G_f + C'' \qquad (6)$$

in which $Y'' = \sigma_N^{-2}/g_f(\alpha_0)$, $X'' = d/E_c$, and $C'' = C/g_f(\alpha_0)$(with $r = 1$). Since for these variables the slope of the regression line is G_f^{-1} (Fig. 5) and the fracture energy is a material constant independent of the specimen type, the test results for the three specimen types should be fitted by regression lines with arbitrary vertical intercepts but with the same slope. The test results have been analyzed collectively for all the specimen types under this restriction, and they are shown for concrete and mortar in Fig. 5(a) and (b). Due to the scatter of the test data apparent in these figures, we cannot say that the test results prove that the slope (and thus the fracture energy) is the same for all the three specimen types; however, we must also admit that these plots do not reveal any systematic deviation from the constant common slope G_f^{-1}. To illustrate the scatter more clearly we may shift the data for each specimen type vertically so that the regression lines coincide (the shift being indicated by the vertical intercepts C_i for specimen types i = 1, 2, 3). The resulting plot is shown in Fig. 5(c) and (d) for concrete and mortar, and it is

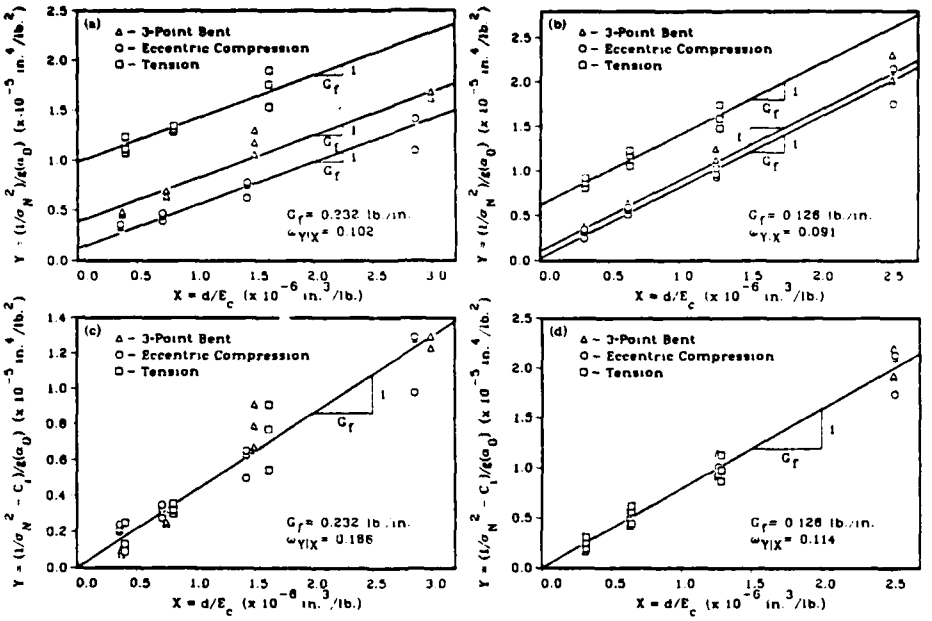


Fig. 5—Plots demonstrating independence of fracture energy from specimen type, (a, c) concrete and (b, d) mortar

seen again that although the large scatter prevents us from concluding that all three specimen types yield the same slope, no systematic deviations from a common slope are apparent.

Comparing the size-effect plots in Fig. 2(a), (c), and (e), we may note that in the case of the eccentric compression specimen the curve is quite close to the asymptote of slope $-\frac{1}{2}$ for linear elastic fracture mechanics, while for the case of tension specimen the curve is quite remote from this asymptote, and for the three-point bent specimen an intermediate situation occurs. From comparisons of these graphs, we may observe that the size range of the tension specimens would have to be increased about 20 times to approach the asymptote as closely as the compression specimen. Obviously, the tension specimen is by far the best for exploring, with relatively small specimen sizes, the behavior near the horizontal asymptote for the strength criterion, and the eccentric compression specimen is best for finding the linear fracture mechanics asymptote, and through it the value of G_f. What is the reason for these differences in behavior?

It is the difference in the size of the fracture process zone that causes these behavioral differences, as we will demonstrate by simplified analysis as well as finite elements. The difference can be understood intuitively by considering (according to the bending theory) linearized stress distributions across the ligament as shown in Fig. 3(d). Even though these stress distributions are no doubt far from the real ones, they make it clear that for the eccentric compression specimen the crack as it extends from the tension notch soon runs into a compression zone, and so the fracture process zone for the eccentric compression specimen can occupy only a small fraction of the ligament. The smaller the relative size of the fracture process zone, the closer the behavior should be to the limit of linear elastic fracture mechanics. For the tension specimen, on the other hand, the entire ligament is under tensile stress, and so nothing prevents the fracture process zone from being as long as the entire ligament itself. For such a large fracture process zone, the behavior should be close to the strength criterion limit, as corroborated by Fig. 2(c). For the three-point bent specimen, the size of the fracture process zone is also limited by the compression field of the top side of the ligament [Fig. 3(d)], but the tension part of the ligament is larger than it is for the eccentric compression specimen, and so the fracture process zone should be of medium length. This explains why the size-effect plots for this type of specimens are intermediate between the tension and eccentric compression specimens of the same exterior dimensions.

It must be emphasized that the size-effect method fails if the size range is insufficient compared to the width of the scatter band of the test results. The relevant data scatter is characterized by the standard deviation A of the regression line slope A and the corresponding coefficient of variation ω_A, which are defined as

$$\omega_A = \frac{s_A}{A}, \quad s_A = \frac{s_{Y|X}}{s_X \sqrt{n-1}} \tag{7}$$

where n is the number of all the data points in the linear regression, s_x is the coefficient of variation of the X values for all the points, and $s_{Y|X}$ is the standard deviation of the vertical deviations from the regression line. The values of ω_A are listed in Tables 2 and 3.

If the size range were so narrow that it would be approximately equal to the width of the scatter band, the data points would fill roughly a circular region (Fig. 6). One could still obtain a unique regression line, but its

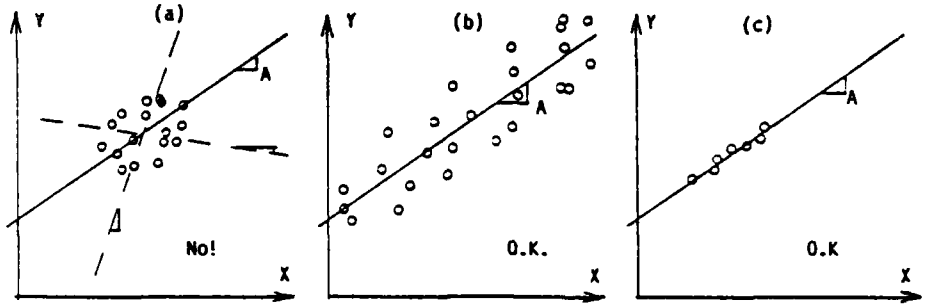


Fig. 6—(a) Size-effect method is inapplicable when the size range is too small compared to scatter; (b, c) size range is sufficient

slope would be highly uncertain and could even come out negative. To prevent this from happening, it is necessary that the value of ω_A would not exceed about 0.1. Even then, by testing very many specimens (large n) one could make ω_A sufficiently low even when the data points fill roughly a circular region rather than an elongated band. To prevent this from happening, one needs to require further that, for any n, the value of $\omega_{\gamma X}/\omega_X$ would not exceed about 0.15.

Obviously, if the test results are consistent, with a low scatter, one can do with a narrower size range than if the test results are highly scattered.

From ω_A one can approximately estimate the coefficient of variation ω_{G_f} of the fracture energy $G_f = g_f(\alpha_0)/AE_c$. If G_f is assumed to be perfectly correlated to E_c, then $\omega_{G_f} = \omega_A$. If G_f is assumed to be uncorrelated to E_c then, as a second-order approximation, $\omega_{G_f} = \omega_G$ where $\omega_G^2 = \omega_A^2 + \omega_E^2$ and ω_E is the coefficient of variation of E_c. In reality one may expect $\omega_A \leqslant \omega_{G_f} \leqslant \omega_G$.

It needs to be pointed out also that the size-effect approach can be destroyed by other size effects. They may arise from diffusion phenomena, e.g., the heating (and microcracking) caused by hydration heat, which may be significant for very large specimens, or the drying of the specimen, which is more severe for thinner or smaller specimens.

BRITTLENESS NUMBER

Our finding that the size-effect law yields approximately unique G_f values regardless of size and geometry makes it meaningful to base on this law a nondimensional characteristic[15,30] that indicates whether the behavior of a given specimen or structure is closer to limit analysis or to linear elastic fracture mechanics. The relative structure size $\lambda = d/d_a$ cannot serve as an objective indicator of this behavior. This is clear from the present tests. For $\lambda = 4$, e.g., the behavior of the tensile specimen is (according to Fig. 2) closer to plastic limit analysis based on the tensile strength, while the behavior of the eccentric compression specimen for $\lambda = 4$ is closer to linear elastic fracture mechanics. An objective indicator is the recently proposed Bažant's[15,30] brittleness number β. It is defined as

$$\beta = \frac{d}{\lambda_0 d_a} \tag{8}$$

and can be calculated after λ_0 has been determined either experimentally or by finite element analysis. The value of $\beta = 1$ indicates the relative size d/d_a at the point where the horizontal asymptote for the strength criterion intersects the inclined straight-line asymptote for the energy failure criterion of linear elastic fracture mechanics (Fig. 2). So $\beta = 1$ represents the center of the transition between these two elementary failure criteria.

For $\beta < 1$, the behavior is closer to plastic limit analysis, and for $\beta > 1$ it is closer to linear elastic fracture mechanics. For $\beta \leqslant 0.1$, the plastic limit analysis may be used as an approximation, and for $\beta \geqslant 10$, linear elastic fracture mechanics may be used as an approximation. For $0.1 < \beta < 10$, nonlinear fracture analysis must be used.

To find the brittleness number, one needs to calculate the coefficient λ_0, which represents the value of d/d_a at the point of intersection of the horizontal straight line and the inclined straight line in Fig. 2. The inclined straight line is given by the equation $\sigma_N = [G_f E/g_f(\alpha_0)d]^{1/2}$, which is valid according to linear elastic fracture mechanics for any two-dimensional structure. The horizontal line is given by the equation $\sigma_N = Bf_t'$ when coefficient B is obtained by plastic limit analysis. By equating both expressions for σ_N, Bažant[31] obtained the following expressions for the transition value d_0 of the characteristic dimension d of the structure and of the corresponding transition value of the relative structure size d/d_a

$$d_0 = \frac{G_f E_c}{f_t'^2 B^2 g_f(\alpha_0)}, \qquad \lambda_0 = \frac{d_0}{d_a} \tag{9}$$

Therefore[31]

$$\beta = B^2 g_f(\alpha_0)\frac{f_t'^2 d}{G_f E_c} \tag{10}$$

Practical calculations may generally proceed as follows: First solve the structure by plastic limit analysis,

which yields the value of B. Second, solve the structure by linear elastic fracture mechanics, which yields the value of g_f. Third, calculate β from Eq. (10). This method of calculating β requires no laboratory tests. Alternatively of course, the brittleness number of a certain type of structure of a given size can be determined by fitting the size-effect law either to test results or to finite element results for ultimate loads of structures of similar geometry but sizes that differ from the given size. (These finite element results must be based on a softening stress-strain or stress-displacement relation.)

It must be emphasized that the value of B must be calculated by plastic limit analysis rather than an allowable elastic stress formula. For example, for an unreinforced beam of span L, rectangular cross section of width b and net depth d in the middle, with concentrated load P at midspan, we have $M_u' = PL/4 = f_t' bd^2/4$ = plastic ultimate bending moment at midspan. From this we get $\sigma_N = P/bd = f_t' d/L$ or $\sigma_N = Bf_t'$ where $B = d/L$ = constant (for geometrically similar beams). It would be of course incorrect to use the elastic formula $M_u' = f_t' bd^2/6$ which would yield $B = 2d/3L$.

The brittleness number can serve as a basic qualitative indicator of the type of response. In this sense it is in fact analogous to the nondimensional characteristics used, e.g., in fluid mechanics, such as the Reynolds number.

Some researchers have tried to characterize the effects of structure size on the qualitative fracture behavior by means of some nondimensional combination of G_f, f_t', and E_c. This is, however, insufficient because these parameters cannot reflect differences in structure geometry. For example, Carpinteri[32] characterized the effect of structure size on its brittleness by the nondimensional ratio $s = G_f/b f_t'$ and Hillerborg[23] by the nondimensional ratio of some structural dimension to the characteristic length $l_{ch} = E_c G_f/f_t'^2$, defined in the same manner as the size of the small-scale yielding zone in metals.[8,23] However, these nondimensional ratios, i.e., the Carpinteri's and Hillerborg's brittleness numbers, are objective only for comparisons of different sizes of structures of the same geometry. For the same value of either Carpinteri's or Hillerborg's brittleness number, the failure of a structure of one geometry can be quite brittle, i.e., close to linear elastic fracture mechanics, while the failure of a structure of another geometry can be quite ductile, i.e., close to limit analysis.

The effect of structure size on the brittleness of its response is manifested not only in the maximum load but also in the post-peak shape of the load-deflection diagram. As graphically illustrated in Fig. 6 of Reference 15, the total deflection is a sum of the deflections due to the fracture process zone and to elastic strains. As the size is increased, the former remains about the same while the latter increases, causing the post-peak diagram to become steeper and steeper and eventually reverse to a snapback-type load-deflection diagram that is unstable under both load and displacement controls.

R-CURVES FOR DIFFERENT SPECIMEN TYPES

According to classical linear elastic fracture mechanics, the specific energy required for crack growth R is constant and equal to the fracture energy G_f. For materials that do not follow linear elastic fracture mechanics, R varies with the length of crack extension from the notch. This variation is described by the so-called resistance or R-curve. When the R-curves were first observed for ductile fracture of metals, it was proposed[33,34] that the shape of the R-curve may be considered to be approximately a material property. For concrete, this concept was introduced by Wecharatana and Shah,[35] and further refined by Bažant and Cedolin.[36] Analysis of extensive test data from the literature showed that, as a crude approximation, the R-curve may be considered unique for a certain limited range of specimen geometries,[36] but it can be very different for some very different specimen geometries.[16] Nevertheless, even though the R-curve is not a unique material property in general, it represents a simple, convenient way to reduce nonlinear fracture analysis to a linear one.

For a specified specimen geometry and type of loading, there is a one-to-one relationship between the size-effect law and the R-curve.[16] If one is known, the other can be easily calculated. The method described previously[16] has been used to calculate the corresponding R-curves from the size-effect curves in Fig. 2. They are shown in Fig. 7. Fig. 7(a) and (c) give the plots of the specific energy required for crack growth R as a function of the crack length c measured from the notch. Fig. 7(b) and (d) show (for the same crack lengths c) the relative values defined as R for the given specimen divided by R for the three-point bent specimen taken as a reference (by definition, for the three-point bent specimen the relative values are all 1).

From Fig. 7 we may observe that for the three-point bent specimens and the eccentric compression specimens the R-curves are not too far apart. Therefore, their mean could be used as a material property for a crude approximation. For the tension specimen, however, the R-curve is very different. This proves that in general the R-curve of concrete cannot be considered to be a unique material property. This conclusion, however, does not apply to the asymptotic values.

The asymptotic value of the R-curve ($c \to \infty$) corresponds to the limiting case of elastic fracture mechanics and is determined by the straight-line asymptote of slope $-\frac{1}{2}$ in Fig. 2. While the R-curves are very different for small crack lengths, the asymptotic values are nearly the same, up to a reasonable scatter range that is inevitable for a material such as concrete. This conclusion, which is particularly conspicuous from the relative curves in Fig. 7(b) and (d), is a basic result of this study. Since the final asymptotic value of the R-curve represents the fracture energy G_f as defined in this paper, this conclusion agrees with the previous conclusion that the fracture energies are approximately the same for various specimen types, with a reasonable scatter that appears to be random rather than systematic.

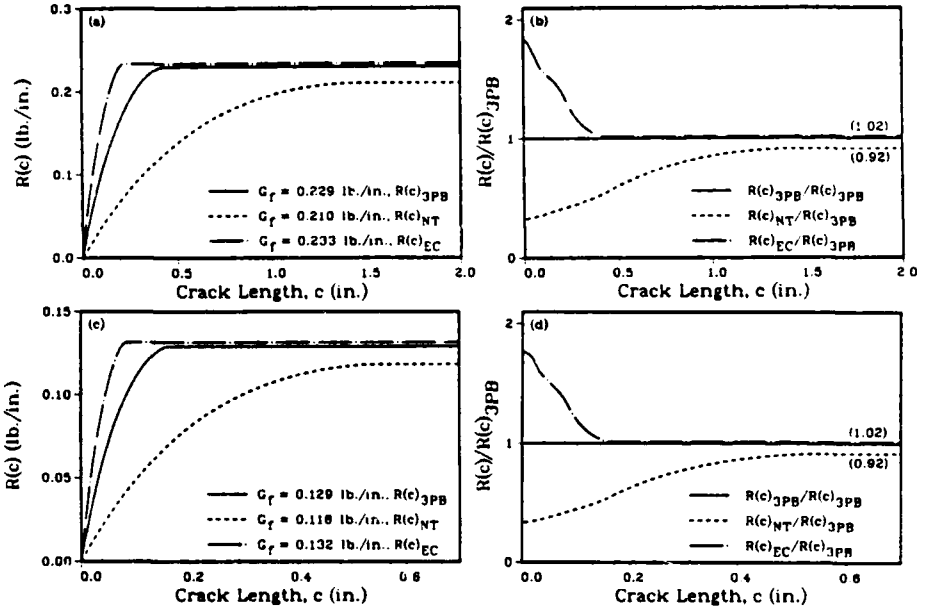


Fig. 7—R-curves for three-point bent, edge-notched tension, and eccentric-compression specimens of (a) concrete and (c) mortar, and (b, d) relative R-curves

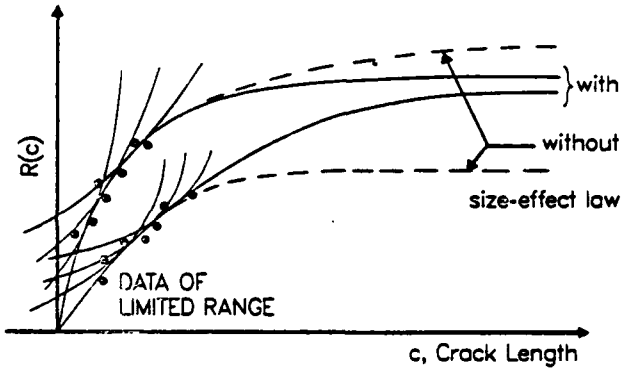


Fig. 8—Extrapolation of R-curves with and without size-effect law

With the help of the R-curves it may be explained why the methods currently used to define and measure the fracture energy do not lead to unique results and exhibit a large spurious dependence on the specimen size. These existing methods generally use the same specimen geometry and dimensions (size) and rely on measurements at various crack lengths or notch lengths. The work of fracture method recently adopted by RILEM[25] is of that type, as is the ASTM method for the measurement of R-curve. Because the specimens are of one size, the results correspond to a short segment of the size-effect curve, which in turn corresponds to a relatively short segment of the R-curve. This is so, even if the measurements are done at various crack lengths or various notch lengths, as demonstrated by finite element results in Fig. 7 of Reference 16. Thus one obtains for the R-curve a set of measured points occupying only a small part of the R-curve, as illustrated in Fig. 8. Obviously, extrapolation to $c \rightarrow \infty$ from such data for a limited size range is ambiguous. Due to inevitable statistical scatter, the extrapolated dashed or solid curves in Fig. 8 yield almost equally good fits of the measured data. It must be concluded that the existing methods that do not use specimens of very different sizes are inherently incapable of giving consistent results for the fracture energy as defined here. They yield fracture energy values illustrated by the asymptote of the dashed curves in Fig. 8, which are usable in the analysis of structures that are not much larger than the specimens tested, but are inapplicable to structures that are much larger.

These difficulties in determination of R-curve from scattered data for a limited size range are further compounded by the fact that the R-curve has to be determined as an envelope of a family of fracture equilibrium curves.[16,36] When these curves are scattered, an envelope simply cannot be constructed.[16]

FINITE ELEMENT ANALYSIS OF TEST RESULTS

The size effect in fracture can also be described by finite elements.[16,37] By optimizing the material fracture parameters so as to obtain the finite element fit of the present test data, it is possible to obtain the fracture energy. Is this result approximately the same as the fracture energy value obtained directly from the size-effect law?

To answer this question, the present test specimens were analyzed by finite elements in exactly the same manner as described on pages 302 to 303 of Reference 16. The analysis utilizes the crack band model with a square mesh of four-node quadrilateral elements in the fracture region. The fracture is simulated by a band of cracking elements of a single-element width. The cracking is described by gradual strain-softening in the elements of the crack band. To explore the effect of the shape of the stress-strain diagram for strain softening in

the crack band, the analyses are carried out both for linear softening [Fig. 9(a)], as in Reference 37, and for strain-softening given as an exponential passing through the same peak point [Fig. 9(b)]. The fracture energy in this approach is given by the area under the uniaxial tensile stress-strain diagram, multiplied by the width w_c of the crack band front. This width must be considered to be a material property independent of the element size and represents a certain small multiple of the aggregate size. The finite element solutions are obtained by step-by-step loading. They yield the maximum loads to which the size-effect law can be matched.

The results of the finite element calculations are compared with the direct analysis of test data by the size-effect law in Table 5 and Fig. 10. The size-effect law results are listed in the first column for G_f in Table 5. The finite element results obtained with linear and exponential strain-softening are listed in the last two columns of Table 5. These values were obtained by fitting the size-effect law to the finite element results for specimens of various sizes, and then matching the size-effect law to the curve [Eq. (2)] that optimally describes the test data. The fracture energy may then be obtained either from the asymptotic slope for the size-effect curve that fits the finite element results (Table 5), or directly from the area under the stress-strain curve (Fig. 9) considered in the finite element analysis, times width w_c (Table 6). For the three types of specimens used, the results are given in the last two columns of Table 5, separately for concrete and mortar. Since all the specimens were cast from the same concrete, the G_f values from finite elements with linear softening were forced to be the same for all the three types of tests (0.230 lb/in. or 40.3 N/m). The exponential softening is always introduced in such a manner that the areas under the stress-strain diagram, and thus also G_f, would be the same as for linear strain softening. The results for exponential strain softening (material parameters given in Table 6) are different, but only marginally so, and are close to the G_f values calculated directly from the size-effect law for each specimen type — for concrete as well as mortar. Compare the columns of Table 5.

Generally, the finite element calculations indicate relatively good agreement between the G_f values that give optimum fits of the measured maximum loads by the finite element program and those obtained by fitting the size-effect law to the measured maximum loads for various specimen sizes. The finite element results in Fig. 10 are compared directly with the test data and the size-effect law fits from Fig. 2 based on both the log (σ_N/f'_t) versus log (d/d_a) plot (a, c, and e) and the linear regression plot (b, d, and f). The finite element meshes used to obtain these results are drawn in Fig. 11. The coefficient of variation ω_{rsx} for the linear regression plots is given for each loading case and for both linear and exponential softening. Among these fits, the best that one could hope to obtain is the size-effect law, which has an average coefficient of variation of $\bar{\omega}_{rsx} = 0.11$ where $\bar{\omega}_{rsx} = \{[\omega^2_{rsx}(3PB) + \omega^2_{rsx}(NT) + \omega^2_{rsx}(EC)]/3\}^{1/2}$; $3PB$, NT, and EC stand for the three-point bent, notched tension, and eccentric compression specimens, respectively. For linear softening $\bar{\omega}_{rsx} = 0.14$ and for exponential softening $\bar{\omega}_{rsx} = 0.20$. Linear softening seems to match the test data somewhat better than exponential softening, for this particular set of data.

The finite element crack band model can also be used to obtain the R-curves.[16,37] The results of such calculations are shown in Fig. 12. The finite element R-curves obtained for both linear and exponential softening follow the general trend of the R-curves obtained by the size-effect law from the measured maximum loads (Fig. 7). The linear softening seems to give a slightly better agreement than the exponential softening when compared to the R-curves deduced directly from the test data.

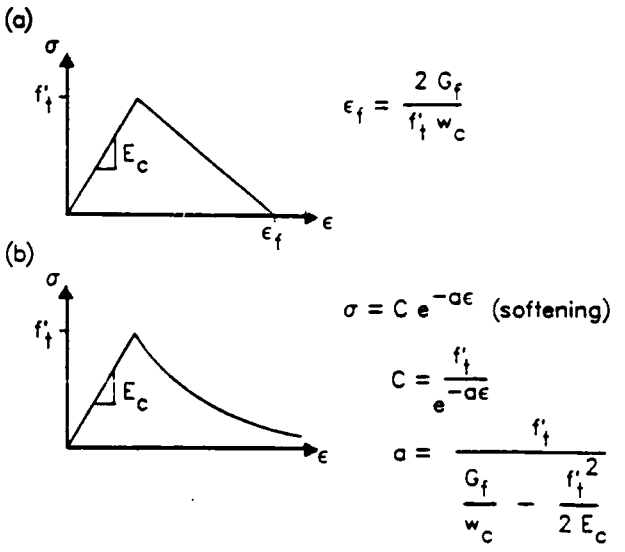


Fig. 9—Uniaxial stress-strain curves for (a) linear and (b) exponential softening

Table 5—Comparison of fracture energy from size-effect law and finite element results

Loading	Material	Size-effect law G_f, lb/in.	Linear softening G_f, lb/in.	Exponential softening G_f, lb/in.
Three-point bent	Concrete	0.229	0.230	0.249
Notched tension	Concrete	0.210	0.230	0.249
Eccentric compression	Concrete	0.233	0.230	0.249
Three-point bent	Mortar	0.129	0.130	0.141
Notched tension	Mortar	0.118	0.130	0.141
Eccentric compression	Mortar	0.132	0.130	0.141

Table 6—Material constants for strain softening used in finite element analysis

Softening Behavior	Material	f'_t, psi	E_c, ksi	w_c^*, in.	G_f, lb/in.
Linear	Concrete	390.7	3976	1.85	0.230
Linear	Mortar	519.7	4738	0.703	0.130
Exponential	Concrete	390.7	3976	2.85	0.249
Exponential	Mortar	519.7	4738	1.083	0.141

*w_c = crack front width.

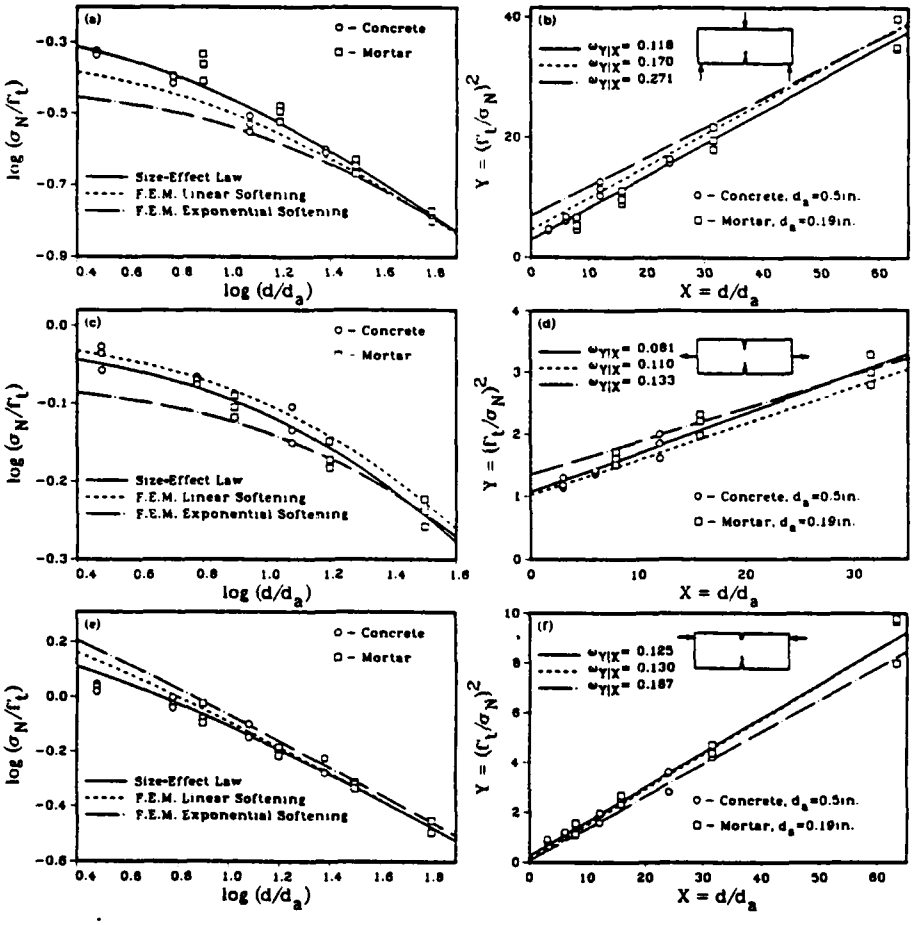


Fig. 10—Comparison of finite element results for three-point bent, edge-notched tension, and eccentric-compression specimens in (a, c, e) size-effect plot and (b, d, f) linear regression plot

SHEAR FRACTURE (MODE II)

All of our analysis thus far dealt with the opening fracture mode (Mode I). According to test results in a recent paper,[29] the size-effect law is also applicable to shear fracture (Mode II). Shear fracture can be produced on the same specimen as used here for tension and compression, although with a different type of loading, shown in Reference 29. For this loading the fracture energy for Mode II can be calculated from Eq. (1) using $g_f(\alpha_0) = 2.93$, as indicated by linear elastic fracture mechanics.[8,23,24] In finite element modeling, the crack band model is applicable to shear fracture as well, provided the smeared cracks within the finite elements are allowed to be inclined with regard to the crack band, letting them form with the orientation normal to the maximum principal stress. The compression stiffness and strain-softening of the concrete between the cracks, loaded in the direction parallel to the cracks, is taken into account.[29]

Tests for shear fracture, comparable with the preceding results, were conducted with the same type of concrete. Therefore, the results for shear fracture are also listed in Tables 7 and 8. The Mode II fracture energy G_f'' is found to be far larger than the Mode I fracture energy $G_f = G_f'$. This may be explained by compressional resistance to concrete between the cracks in the inclined direction parallel to the cracks.[29] The size-effect regression plots for Mode II are shown in Fig. 13(a) and (b) and the finite element results in Fig. 13(c) and (d).

Recently, Bažant and Prat[38] applied the size-effect law to Mode III fracture tests of cylindrical specimens with circumferential notches, subjected to torsion, and used again Eq. (1) to determine the Mode III fracture energy of concrete.

CONCLUSIONS

1. Fracture energy of a brittle heterogeneous material such as concrete may be defined as the specific energy required for fracture growth when the specimen or structure size tends to infinity. In this definition the fracture energy is a unique material property, independent of specimen size, shape, and type of loading.

2. The foregoing definition reduces the problem to the question as to which form of the size-effect law should be used for extrapolating the test results to infinite size. Although the exact size-effect law is not known, the present test results indicate that Bažant's

approximate size-effect law [Eq. (2)], with a correction for the maximum aggregate size [Eq. (4)], may be acceptable for practical purposes.

3. When the present method is used, different types of fracture specimens — such as the edge-notched tension specimen, three-point bent specimen, and notched eccentric compression specimen — yield approximately the same values of fracture energy. The observed scatter range is about the same as the usual range of inevitable scatter for concrete. Thus, the present method of defining and measuring the fracture energy appears to be approximately independent of both the specimen size and type — a goal not yet achieved with other methods.

4. Bažant's[15,30] brittleness number β [Eq. (8) and (10)], based on the size-effect law, may be used as a nondimensional characteristic of fracture similitude, which indicates how close the structure behavior is to linear elastic fracture mechanics or to limit analysis. Bažant's brittleness number, in contrast to Carpinteri's or Hillerborg's, is independent of the shape of specimen or structure, and so it can be used to compare the brittleness of structures of different shapes. For $\beta \leqslant 0.1$, the response of a structure of any shape is essentially ductile and plastic limit analysis applies, for $\beta \geqslant 10$ it is essentially brittle and obeys linear elastic fracture mechanics, and for $0.1 < \beta < 10$ the brittle and ductile responses mix and a nonlinear fracture analysis is required.

5. R-curves for various specimen types, calculated on the basis of the size effect from the maximum loads of specimens of different sizes, are very different for short crack lengths but approach a common asymptotic value for large crack lengths.

6. The present test results can be described by the finite element crack band model. The fracture energy values determined from the size effect approximately agree with the fracture energy value used in the finite element code, which represents the area under the tensile stress-strain diagram multiplied by the width of the cracking element at fracture front.

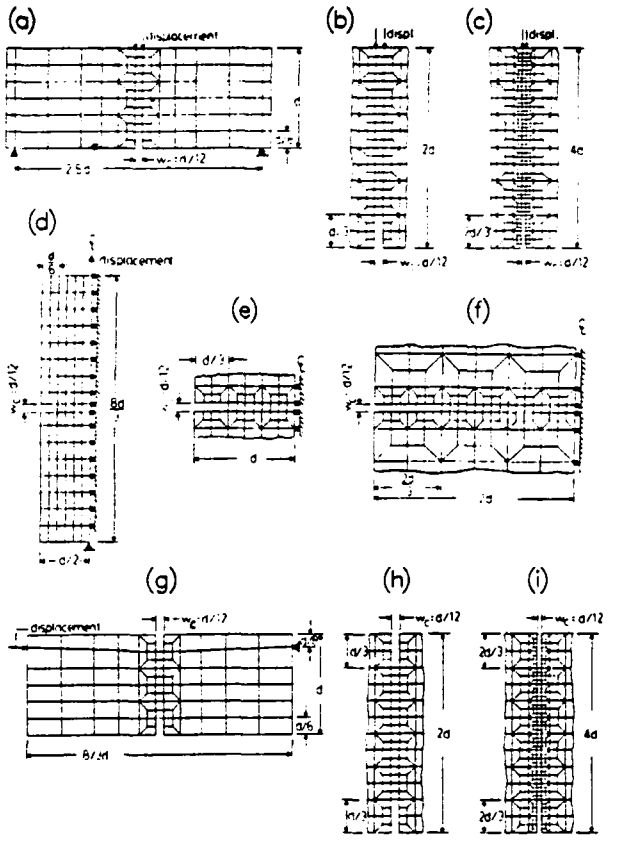


Fig. 11—Finite element meshes for (a, b, c) three-point bent, (d, e, f) edge-notched tension, and (g, h, i) eccentric-compression specimens

Table 7—Comparison of fracture energy from size-effect law and finite element results for shear fracture

Loading	Material	Size-effect law G_f, lb/in.	Linear softening G_f, lb/in.	Exponential softening G_f, lb/in.
Shear	Concrete	6.30	5.31	6.83
Shear	Mortar	3.34	2.82	3.62

Table 8—Material constants for strain softening used in finite element analysis of shear fracture

Softening behavior	Material	f'_c, psi	E_c, ksi	w_c, in.	G_f, lb/in.
Linear	Concrete	400.7	4000	1.85	0.246
Linear	Mortar	516.0	4750	0.703	0.130
Exponential	Concrete	400.7	4000	2.9	0.260
Exponential	Mortar	516.0	4750	1.1	0.138

ACKNOWLEDGMENTS

Partial financial support under Air Force Office of Scientific Research Grant No. 83-00092 and contract No. F49620-87-C-00300EF with Northwestern University is gratefully acknowledged.

Further financial support has been obtained under a cooperative research program with Universidad Politécnica de Madrid, funded under U.S.-Spanish Treaty (Grant CCA-8309071.)

REFERENCES

1. Shah, S. P., Editor, *Application of Fracture Mechanics to Cementitious Composites*, Martinus Nijhoff Publishers, Dordrecht, 1985, 714 pp.

2. Bažant, Z. P.; Kim, J. K.; and Pfeiffer, P., "Determination of Nonlinear Fracture Parameters from Size Effect Tests," NATO Advanced Research Workshop on Application of Fracture Mechanics to Cementitious Composites," Northwestern University, Evanston, Sept. 1984, pp. 143-169.

3. Wittmann, F. H., Editor, "Fracture Toughness and Fracture Energy of Concrete," *Proceedings*, RILEM International Conference on Fracture Mechanics of Concrete, EPFL, (held in Lausanne, 1985), Elsevier, Amsterdam, 1986, 699 pp.

4. RILEM Committee 50-FMC, "Determination of the Fracture Energy of Mortar and Concrete by Means of Three-Point Bend Tests on Notched Beams," RILEM Draft Recommendation, *Materials and Structures, Research and Testing* (RILEM, Paris), V. 18, No. 106, July-Aug. 1985, pp. 285-290.

5. Hilsdorf, H. K., and Brameshuber, W., "Size Effects in the Experimental Determination of Fracture Mechanics Parameters," *Application of Fracture Mechanics to Cementitious Composites*, Martinus Nijhoff Publishers, Dordrecht, 1985, pp. 361-397.

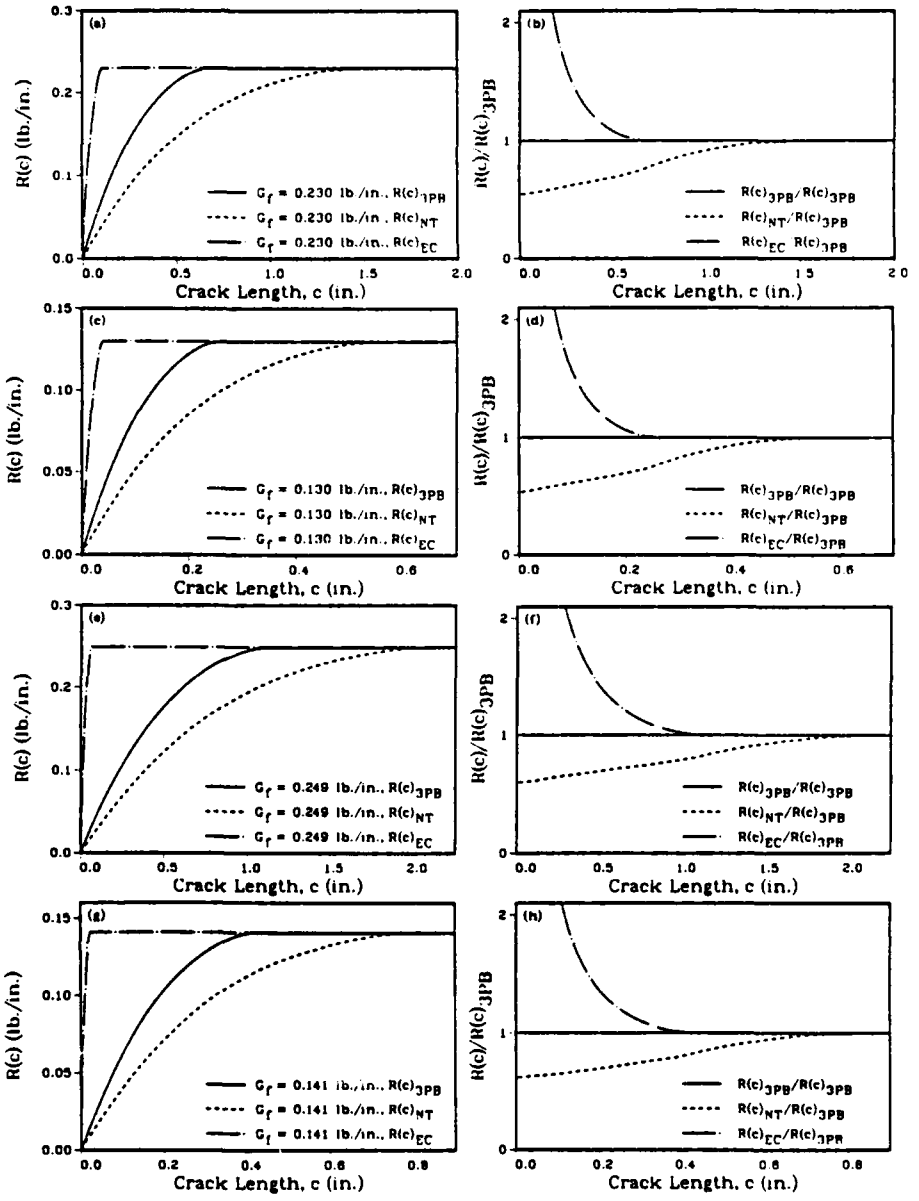


Fig. 12—Finite element results with linear softening for R-curves of three-point bent, edge-notched tension, and eccentric-compression specimens of (a) concrete and (c) mortar; relative R-curves for (b) concrete and (d) mortar; finite element results with exponential softening for three-point bent, edge-notched tension, and eccentric compression specimens of (e) concrete and (g) mortar; and relative R-curves for three-point bent specimens for (f) concrete and (h) mortar

6. Horwath, R., and Persson, T., "The Influence of the Size of the Specimen on the Fracture Energy of Concrete," *Report* No. TVBM-5005, Division of Building Materials, Lund Institute of Technology.

7. Hillerborg, Arne, "Results of Three Comparative Test Series for Determining the Fracture Energy G_f of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 18, No. 107, Sept.-Oct. 1985, pp 407-413.

8. Knott, J. F., *Fundamentals of Fracture Mechanics*, Butterworths, London, 1973, 273 pp.

9. Bažant, Z. P., "Fracture in Concrete and Reinforced Concrete," *Preprints*, IUTAM Prager Symposium on Mechanics of Geomaterials: Rocks, Concretes, Soils, Northwestern University, Evanston, Sept. 1983, pp. 281-316.

10. Bažant, Z. P., "Mechanics of Fracture and Progressive Cracking in Concrete Structures," *Fracture Mechanics of Concrete: Structural Application and Numerical Calculation*, Martinus Nijhoff Publishers, Dordrecht, 1985, pp. 1-93.

11. Bažant, Zdeněk P., "Size Effect in Blunt Fracture: Concrete, Rock, Metal," *Journal of Engineering Mechanics*, ASCE, V. 110, No. 4, Apr. 1984, pp. 518-535.

12. Bažant, Zdeněk P., and Kim, Jin-Keun, Authors' Closure to Discussion of "Size Effect in Shear Failure of Longitudinally Reinforced Beams," ACI JOURNAL, *Proceedings* V. 82, No. 4, July-Aug. 1985, pp. 579-583. Also, "Effects of Aggregate and Structure Sizes on Concrete Failure," correspondence to J. K. Kim, May 1985.

13. Bažant, Z. P., "Comment on Hillerborg's Comparison of Size Effect Law with Fictitious Crack Model," *dei Poli Anniversary Volume*, Politecnico di Milano, Oct. 1985, pp. 335-338.

14. Bažant, Zdeněk P., "Fracture Mechanics and Strain-Softening of Concrete," *Proceedings*, U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures (Tokyo, May, 1985), American Society of Civil E neers, New York, 1986, pp. 121-150.

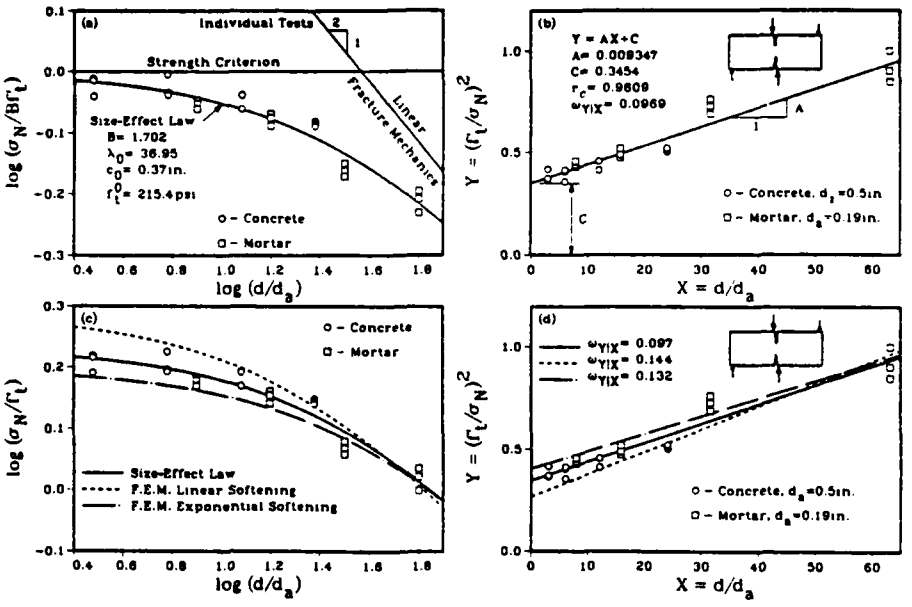


Fig. 13—(a) Size-effect plot, (b) linear-regression plot constructed from maximum load values measured for shear fracture specimens of concrete and mortar, for various specimen sizes, and (c) finite-element results in size-effect plot, (d) linear-regression plot

15. Bažant, Z. P., "Fracture Energy of Heterogeneous Material and Similitude," *Preprints*, RILEM-SEM International Conference on Fracture of Concrete and Rock (Houston, June 1987), Society for Experimental Mechanics, Bethel, pp. 390-402.

16. Bažant, Zdeněk P.; and Kim, Jin-Keun; and Pfeiffer, Phillip A., "Nonlinear Fracture Properties from Size Effect Tests," *Journal of Structural Engineering*, ASCE, V. 112, No. 2, Feb. 1986, pp. 289-307.

17. Bažant, Zdeněk P., and Kim, Jin-Keun, "Size Effect in Shear Failure of Longitudinally Reinforced Beams," ACI JOURNAL, *Proceedings* V. 81, No. 5, Sept.-Oct. 1984, pp. 456-468.

18. Bažant, Zdeněk P., and Cao, Zhiping, "Size Effect in Shear Failure of Prestressed Concrete Beams," ACI JOURNAL, *Proceedings* V. 83, No. 2, Mar.-Apr. 1986, pp. 260-268.

19. Bažant, Zdeněk P., and Cao, Zhiping, "Size Effect in Brittle Failure of Unreinforced Pipes," ACI JOURNAL, *Proceedings* V. 83, No. 3, May-June 1986, pp. 369-373.

20. Bažant, Zdeněk P., and Cao, Zhiping, "Size Effect in Punching Shear Failure of Slabs," *Report* No. 85-8/428s, Center for Concrete and Geomaterials, Northwestern University, Evanston, May 1985. Also, *ACI Structural Journal*, V. 84, No. 1, Jan.-Feb. 1987, pp. 44-53.

21. Bažant, Z. P., and Sener, S., "Size Effect in Torsional Failure of Longitudinally Reinforced Concrete Beams," *Journal of Structural Engineering*, ASCE, V. 113, No. 10, Oct. 1987, pp. 2125-2136.

22. Bažant, Z. P.; Sener, S.; and Prat, P. C., "Size Effect Tests of Torsional Failure of Concrete Beams," *Report* No. 86-12/428s, Center for Concrete and Geomaterials, Northwestern University, Evanston, Dec. 1986, 18 pp.

23. Broek, D., *Elementary Engineering Fracture Mechanics*, Sijthoff and Noordnoff International Publishers, Leyden, 1978, 408 pp.

24. Tada, H.; Paris, P. C.; and Irwin, G. R., *The Stress Analysis of Cracks Handbook*, 2nd Edition, Paris Productions, Inc., St. Louis, Mo. 1985.

25. Hillerborg, A., "The Theoretical Basis of a Method to Determine the Fracture Energy G_f of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 18, No. 106, July-Aug. 1985, pp. 291-296.

26. Bažant, Z. P., "Mechanics of Distributed Cracking," *Applied Mechanics Reviews*, V. 39, No. 5, May 1986, pp. 675-705.

27. Rots, J. G.; Nauta, P.; Kusters, G. M. A.; and Blaauwendraad, J., "Smeared Crack Approach and Fracture Localization in Concrete," *Heron* (Delft), V. 30, No. 1, 1985, 48 pp.

28. Rots, J. G., "Strain-Softening Analysis of Concrete Fracture Specimens," *Proceedings*, RILEM International Conference on Fracture Mechanics of Concrete, EPFL, Lausanne, 1985, pp. 115-126.

29. Bažant, Z. P., and Pfeiffer, P. A., "Shear Fracture Tests of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 19, No. 110, Mar.-Apr. 1986, pp. 111-121.

30. Bažant, Z. P., Seminar, Rensselaer Polytechnic Institute, Troy, Jan. 31, 1986.

31. Bažant, Z. P., "Material Modeling Principles," *Course* No. 720-D95, Northwestern University, Evanston, 1986-1987.

32. Carpinteri, A., "Notch Sensitivity in Fracture Testing of Aggregative Materials," *Engineering Fracture Mechanics*, V. 16, 1982, pp. 467-481.

33. Irwin, G. R., "Fracture Testing of High Strength Sheet Material," *ASTM Bulletin*, Jan. 1960, p. 29. Also, "Fracture Testing of High Strength Sheet Materials under Conditions Appropriate for Stress Analysis," *Report* No. 5486, Naval Research Laboratory, July 1960.

34. Krafft, J. M.; Sullivan, A. M.; and Boyle, R. W., "Effect of Dimensions on Fast Fracture Instability of Notched Sheets," *Cranfield Symposium 1961*, V. 1, pp. 8-28.

35. Wecharatana, Methi, and Shah, Surendra P., "Slow Crack Growth in Cement Composites," *Journal of Structural Engineering*, ASCE, V. 108, ST6, June 1982, pp. 1400-1413.

36. Bažant, Zdeněk P., and Cedolin, Luigi, "Approximate Linear Analysis of Concrete Fracture by R-Curves," *Journal of Structural Engineering*, ASCE, V. 110, No. 6, June 1984, pp. 1336-1355.

37. Bažant, Zdeněk P., and Oh, B. H., "Crack Band Theory for Fracture of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 16, No. 93, May-June 1983, pp. 155-177.

38. Bažant, Z. P., and Prat, P., "Measurement of Mode III Fracture Energy of Concrete," *Nuclear Engineering and Design*, in press.

39. François, Dominique, "Fracture and Damage Mechanics of Concrete," *Application of Fracture Mechanics to Cementitious Composites*, Martinus Nijhoff Publishers, Dordrecht, 1985, pp. 141-156.

APPENDIX — SIZE DEPENDENCE OF CRITICAL STRESS INTENSITY FACTOR AND DERIVATION OF EQ. (1)

The size-effect law [Eq. (2)] yields

$$\sigma_N = Bf'_t\ (\lambda_0\, d_a/d)^n \quad \text{for } d >> d_0 \tag{11}$$

where $d_0 = \lambda_0 d_a$ and $\sigma_N = P/bd$. This must be equivalent to linear elastic fracture mechanics, whose formulas, given for various specimen or structure geometries in textbook and handbooks,[4,23,24] may always be written in the form

$$\sigma_N = [G_f E_c / g_f(\alpha_0)]^n\, d^{-n} \tag{12}$$

for plane stress (for plane strain replace E with $E' = E/(1 - \nu^2)$ where ν = Poisson's ratio). Eq. (11) and (12) must be equivalent, and equating them one gets

$$G_f = \frac{g_f(\alpha_0)}{E_c} B^2 f'^2_t \lambda_0\, d_a \qquad \text{(any } r\text{)} \tag{13}$$

Then, noting that $B^2 f'^2_t \lambda_0\, d_a = A^{1/r}$ [Eq. (5)] one gets $G_f = g_f(\alpha_0)/(E_c A^{1/r})$, and for $r = 1$ one has Eq. (1). A slightly different derivation of Eq. (1) was originally given on p. 293 in Reference 16. Expressing Bf'_t from Eq. (13) and substituting for it in Eq. (2), one may write the size-effect law in the alternative form

$$\sigma_N = \left[\frac{G_f E_c}{g_f(\alpha_0)\, d_0}\right]^n \left[1 + \left(\frac{d}{d_a}\right)^r\right]^{-n/r} \tag{14}$$

which is based on fracture parameters G_f and $g_f(\alpha_0)$ instead of plastic limit analysis parameters f'_t and B.

Eq. (12) may further be used to determine the effect of size on the value of the critical stress intensity factor K_c, which is obtained when fracture tests are evaluated according to linear elastic fracture mechanics. From Eq. (12), $G_f = \sigma_N^2 g_f(\alpha_0) d/E_c$. From this, using the well-known[4] relation $K_c^2 = G_f E_c$ and substituting Eq. (2) for σ_N, Bažant[31] obtained

$$K_c = \left\{\frac{g_f(\alpha_0)\, d}{[1 + (d/d_0)^r]^{1/r}}\right\}^n Bf'_t = \left\{\frac{d/d_0}{[1 + (d/d_0)^r]^{1/r}}\right\} K_c^\infty \tag{15}$$

where $K_c^\infty = Bf'_t\ [g_f(\alpha_0) d_0]^n$ = limiting value of K_c for $d \to \infty$. This equation agrees well with the test data reported by François (Fig. 1 of Reference 39) and others.

APPENDIX A2

Journal of the Am. Ceram. Soc.
in press, 1990

SIZE EFFECT IN FRACTURE OF CERAMICS AND ITS USE TO DETERMINE FRACTURE ENERGY AND EFFECTIVE PROCESS ZONE LENGTH

by

Zdeněk P. Bažant and Mohammad T. Kazemi

Report No. 89-6/498s
(June 1989)

CENTER FOR ADVANCED CEMENT-BASED MATERIALS

ROBERT R. McCORMICK SCHOOL OF ENGINEERING AND
APPLIED SCIENCES IN THE TECHNOLOGICAL INSTITUTE
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS

Revised December 1989

SIZE EFFECT IN FRACTURE OF CERAMICS AND ITS USE TO DETERMINE FRACTURE ENERGY AND EFFECTIVE PROCESS ZONE LENGTH

By Zdeněk P. Bažant[1] and Mohammad T. Kazemi[2]

Center for Advanced Cement-Based Materials
Northwestern University
Evanston, IL 60208

ABSTRACT. - The paper shows that a previously proposed size effect law can be used to identify nonlinear fracture properties solely from measured maximum loads of geometrically similar ceramic fracture specimens of sufficiently different sizes. This law represents a first-order global approximation of the deviations from linear elastic fracture mechanics, independent of the type of the toughening mechanism in the fracture process zone. It provides a simple and unambiguous way to determine the size- and shape-independent values of the fracture energy, the effective length of the process zone, and the effective crack-tip opening displacement. It also yields the R-curve, which is geometry-(shape-) dependent. The proximity of response to linear elastic fracture mechanics is characterized by a brittleness number, which is shape-independent.

[Key words: size effect, fracture energy, fracture process zone, brittleness number, R-curve.]

---

[1] Professor of Civil Engineering;
Member, American Ceramic Society.

[2] Graduate Research Assistant

I. Introduction

Various ceramics and ceramic composites exhibit toughening mechanisms such as transformation [1-2], microcracking [2-4], crack deflection and crack bridging [2, 4-10]. These mechanisms greatly elevate the effective fracture toughness and endow the material with a pronounced R-curve behavior. Fracture behavior of ceramics has been studied intensely, however, one avenue of inquiry, which has met with great success in global modeling of nonlinear fracture of other brittle heterogeneous materials such as concrete, mortar and rock, has received little attention in the studies of ceramics. It is the size effect in fracture, which is most directly manifested in the values of the nominal strength of geometrically similar fracture specimens of different sizes. The size effect is the main practical consequence of fracture mechanics. It must be taken into account in design, and it can also be exploited for determining material fracture parameters merely from maximum load measurements, which are easy to carry out.

The principal purpose of this paper is to apply to ceramics the recent results on the size effect achieved in fracture studies of concrete and rock, and to demonstrate the usefulness of a size effect law [11] giving a first-order global approximation to the deviations from linear elastic fracture mechanics. For readers' convenience, the paper will also give a concise review of the basic results on the size effect in fracture which appeared in a number of recent papers [11-20]. Furthermore, the existing results will be extended by presenting statistical estimates of standard deviation and an effective method of regression of test data on apparent fracture toughness of similar

specimens of different sizes.

Although the effect of structure size as well as geometry on the strength and fracture toughness of ceramics has already been studied by a number of researchers [21-27], only one of the existing data, namely those of McKinney and Rice [23], appear to have a sufficient size range for determining fracture properties. These data, pertaining to one kind of silicon dioxide and two kinds of silicon carbide, will be utilized in the present study to illustrate the identification of fracture properties, and in particular, the determination of the R-curve. The fact that the specimen size affects the measured (apparent) fracture toughness values of ceramics has already been pointed out [23-25], but determination of fracture parameters and the R-curve from the size effect has apparently not yet been attempted in the study of ceramics.

It has often been thought that the R-curve behavior is important only for the ceramics in which the fracture process zone (including the frontal and crack bridging zones) extends over many millimeters. But everything is relative. What matters is the ratio of the structure size to the inhomogeneity size (or more precisely, to the process zone size, which is related to the texture, size and distribution of inhomogeneities [28]). Thus, even an R-curve whose range is much less than 1 mm can be important if the specimen is small enough.

A secondary purpose of this paper, intended to highlight the relativity of scale, is to call attention to the similarity of fracture between ceramics and concrete. In the literature on fracture of ceramics, only a few studies [e.g. 5-6, 10] have referred to the extensive and

exploding literature on the fracture of concrete. Yet many of the fracture problems of brittle heterogeneous materials with a pronounced R-curve behavior have already been successfully tackled in the literature on concrete. These advances have been driven principally by the needs of finite element analysis, which has been widely practiced for concrete because of the forbidding costs of real-size testing of certain structures.

The essential difference between fracture of concrete and ceramics is merely that of scale. In normal concrete, the typical maximum aggregate size is 20 mm, in high strength concrete 10 mm, and in dam concretes 80 mm or more (the corresponding fracture process zone lengths are typically 100 mm, 10 mm and 1000 mm, respectively). In ceramics, there are coarse-grained microstructures with grain size larger than 1 mm as well as fine-grained microstructures with grain size less than 5 μm (the corresponding process zone lengths are roughly between 20 mm and 100 μm). When the specimens of each material have the same cross section size, the degree of brittleness observed in fracture behavior will be vastly different. What matters, however, is the ratio of the structure size to the fracture process zone size. Thus, a concrete slab 400 mm in thickness and a fine-grained ceramic wafer 1 mm in thickness are likely to exhibit about the same type of fracture behavior, which is nonlinear and transitional between linear elastic fracture mechanics and plasticity. Practical applications necessitate extrapolations of scale – in concrete often to much larger structures, while in ceramics for electronic applications to much smaller structures. For both, knowledge of the law governing the size effect is essential, which is the principal

focus of this study.

The size effect that we are going to study is caused by the differences between the energy release rates of small and large structures. It is not a statistical size effect. The well-known Weibull-type weakest link statistics [22] governs the size effect in unnotched structures failing at the initiation of the first macrocrack from a microscopic flaw. But it plays only a minor role in notched fracture specimens. So it does in unnotched structures that fail only after a large stable crack or cracking zone has already formed [12]. Designs that ensure such a type of failure are enforced in concrete design codes by the requirement that the maximum load must be much larger than the crack initiation load. It may of course be appropriate to apply a similar requirement for ceramics. Such a requirement can be met, for example, by providing proper steel reinforcement in concrete or fiber reinforcement in ceramics.

II. Size Effect Law and Its Consequences

The most important consequence of fracture mechanics is the effect of size on the nominal stress at failure. To describe it, we consider geometrically similar structures or specimens of different sizes (with geometrically similar notches or initial cracks), and introduce for two-dimensional (2D) or three-dimensional (3D) problems the nominal stress at maximum load:

$$\sigma_N = c_n \frac{P_u}{bd} \quad \text{(for 2D)} \qquad \text{or} \qquad \sigma_N = c_n \frac{P_u}{d^2} \quad \text{(for 3D)} \tag{1}$$

where P_u = maximum load (ultimate load), b = thickness of

two-dimensionally similar structures or specimens, same for all the sizes, d = characteristic dimension of the structure or specimen, and c_n = coefficient introduced for convenience. For example, if d is the depth of a simply supported beam of span L, the elastic formula for the maximum bending stress is $\sigma_N = 1.5\ P_u L/bd^2 = c_n P_u/bd$ with $c_n = 1.5\ L/d$ (= constant if the structures are geometrically similar). Likewise, if the plastic bending formula, $\sigma_N = P_u L/bd^2$ is used, then $c_n = L/d$ (= constant).

As is well known, plastic limit analysis, as well as elastic analysis with an allowable stress criterion or any method of analysis with a failure criterion based on stress or strain, exhibits no size effect, i.e., geometrically similar structures of different sizes fail at the same σ_N. However, this is not true of fracture mechanics. Due to similarity of stress fields in similar two-dimensional elastic structures of different sizes, the total potential energy of the structure must have the form $U = (\sigma^2/2E')bd^2 f(\alpha)$ where $\sigma = c_n P/bd$, P = load, $f(\alpha)$ is a function of the relative crack length and depends on the shape of the structure, $\alpha = a/d$ (a is the crack length), $E' = E$ for plane stress, $E' = E/(1-\nu^2)$ for plane strain, E = Young's modulus of elasticity, and ν = Poisson's ratio. Therefore, the energy release rate is $G = -(\partial U/\partial a)/b = -(\partial U/\partial \alpha)/bd = -d(\sigma^2/2E')f'(\alpha)$, from which:

$$G = \frac{P^2 g(\alpha)}{E' b^2 d}, \qquad K_I = \sqrt{GE'} = \frac{P k(\alpha)}{b\sqrt{d}} \quad \text{(for 2D problems)} \tag{2}$$

where K_I is the stress intensity factor, $f'(\alpha) = \partial f(\alpha)/\partial \alpha$, $g(\alpha) = -f'(\alpha)c_n^2/2$, and $k(\alpha) = \sqrt{g(\alpha)}$. It may be checked that, for all structure shapes, the formulas for G or K_I have the form of Eq. (2). The values of

$k(\alpha)$ can be obtained by linear elastic analysis, such as finite element analysis, and for basic specimen geometries $k(\alpha)$ follows from the formulas for the stress intensity factors found in handbooks (e.g. Tada et al. [29]).

For three-dimensional similarity, $U = d^3(\sigma^2/2E')f(\alpha)$ where $\sigma = c_n P/d^2$, and $G\rho d = -\partial U/\partial a = -(\partial U/\partial\alpha)/d = -d^2(\sigma^2/2E')f'(\alpha)$, where ρd = length of the perimeter of the fracture front (ρ is a constant) and G = average energy release rate per unit length of the perimeter. Therefore,

$$G = \frac{P^2 g(\alpha)}{E'\rho d^3}, \qquad K_I = \sqrt{GE'} = \frac{Pk(\alpha)}{\sqrt{\rho d^3}} \qquad \text{(for 3D problems)} \tag{3}$$

In linear elastic fracture mechanics (LEFM), the size effect is very simple: $\sigma_N \propto d^{-1/2}$ (from Eqs. 1-3). A deviation from this law is caused by the existence of a finite, nonnegligible fracture process zone which surrounds the tip of the macroscopic continuous crack (and is associated with R-curve behavior).

The energy release rate, G, represents the energy flux into the nonlinear zone at the crack tip from the surrounding material which is in an elastic state. This definition of G coincides with Rice's J-integral but is less general than that. In the line crack models with bridging stresses, the J-integral, and thus also G, is equal to the area under the curve of bridging stress vs. crack opening displacement [e.g. 4, 30]. The J-integral gives the correct energy flux not only for integration contours lying outside the nonlinear zone but also for those lying in the hardening elasto-plastic part of the nonlinear zone, provided that 1) the contour does not pass through the fracture process zone, defined as the softening zone, and 2) that the unloading irreversibility is unimportant.

The reason that the contours passing through the fracture process zone must be barred from the J-integral is that microcracking and other damage in this zone dissipate energy. While in ductile fracture of metals the fracture process zone (in which voids and microcracks nucleate, grow and coalesce) is negligible in size compared to the nonlinear plastic-hardening zone, in ceramics, rock and concrete the fracture process zone is large while the plastic-hardening zone is negligible; see Fig. 1 taken from Ref. 13.

In this study we will consider only specimens or structures in which $g'(\alpha) > 0$. In such structures, said to be of positive geometry [31], the process zone starts to evolve from zero size and then grows as the load increases while remaining attached to the notch tip. If the structure is not large, the process zone length is significant compared to the length of the notch or initial crack, a_0, and the equivalent crack length $a = a_0+c$ at failure must be distinguished from a_0, where c = elastically equivalent length of the process zone. The crack length a is understood to be the equivalent crack length in the sense of linear elastic fracture mechanics (LEFM). The equivalent process zone length c (Fig. 2) is defined as the distance from the tip of the notch or critical crack to the tip of the equivalent crack which gives, according to LEFM, the same unloading compliance. The value of G is the energy release rate for the equivalent linearly elastic crack.

Let c_f be the value of c in an infinitely large structure ($d \to \infty$) at maximum load (which represents the limit of stability if the structure is under load control rather than displacement control). Further, let G_f be the the corresponding value of G required for crack growth at $d \to \infty$. In

the limit of $d \to \infty$, one has $c/d \to 0$ and $\alpha \to \alpha_0 = a_0/d$. Thus, in an infinitely large specimen, the fracture process zone occupies only an infinitesimal volume fraction of the body. Therefore, the entire body can be treated as elastic. Consequently, the stress and displacement fields surrounding the process zone are the asymptotic elastic fields. They are known to be the <u>same</u> for any specimen geometry, and so the shape and size of the fracture process zone must be the same. It follows that G_f and c_f, defined as the limits of G and c for an infinitely large specimen, are independent of the specimen shape. Therefore, unambiguous definitions of G_f and c_f as fundamental material properties, independent of specimen size and shape, can be given as follows [12,18]:

G_f and c_f are the energy required for crack growth and the elastically equivalent length of the fracture process zone, respectively, in an infinitely large specimen.

This definition of fracture energy can be mathematically stated as $G_f = \lim G_c = \lim (K_{Ic}^2/E')$ for $d \to \infty$, where G_c and K_{Ic} are the values of G and K_I [Eq. (2)], respectively, calculated from the measured peak load P_u and the initial crack or notch length a_0 using LEFM equations. More generally, G_f can be defined as the limit of J-integral for $d \to \infty$.

The value of G required for fracture growth is basically determined by the size of the process zone. Since the value of c is also determined by this size, the value of G for a growing crack may be assumed to be a function of the corresponding value of c (which serves as the basis of the R-curve approach). The value of c at $P = P_u$ determines the value of α and consequently the value of $g(\alpha)$ at failure, and so the ratio $G/g(\alpha)$ at maximum load should be approximately equal to $G_f/g(\alpha_f)$. Therefore, $G \simeq$

$G_f g(\alpha)/f(\alpha_f)$. Now we may substitute this expression into Eq. (2), make further the approximation $g(\alpha) \simeq g(\alpha_0)+g'(\alpha_0)(\alpha_f-\alpha_0)$ (from Taylor series expansion assuming that $g'(\alpha_0)>0$), denoting $\alpha_f = \alpha_0+c_f/d$ and $g'(\alpha_0) = dg(\alpha_0)/d\alpha$, and further set (from Eq. 1) $P_u^2 = (\sigma_N bd/c_n)^2$ and solve for σ_N. Thus we obtain the following approximate size effect law:

$$\sigma_N = c_n \left(\frac{E'G_f}{g'(\alpha_0)c_f + g(\alpha_0)d} \right)^{\frac{1}{2}} \tag{4}$$

This law may alternatively be written as [18]:

$$\tau_N = \left(\frac{E'G_f}{c_f+\bar{d}} \right)^{\frac{1}{2}} \tag{5}$$

where

$$\tau_N = \frac{P_u}{bd}\sqrt{g'(\alpha_0)} \quad , \qquad \bar{d} = \frac{g(\alpha_0)}{g'(\alpha_0)}\, d \tag{6}$$

τ_N = intrinsic (shape-independent) nominal stress at failure, and $\bar{d}$ = intrinsic (shape-independent) size of the structure [18], which was also introduced in a somewhat different context by Planas and Elices [31].

The quantity that makes $\bar{d}$ shape-independent is the ratio $g(\alpha_0)/g'(\alpha_0)$ [18], which has also been introduced for similar purposes by Planas and Elices [31] and for other purposes by Horii et al. [32]. Since the factor $g(\alpha)/g'(\alpha)$ in Eq. (6) takes the specimen shape into account only in an approximate way, the tests, whose results are used in regressions according to Eq. (5), should preferably involve only

geometrically similar specimens.

The size effect law has also been derived in two other ways: 1) by the method of energy release zones, and 2) by dimensional analysis and similitude arguments [11,12]; see Appendix. The latter derivation, which is more general, is based on the hypothesis that the total energy ΔU released due to crack formation is a function of both the initial crack length a_0 and the size of the process zone in an infinitely large specimen, which is a material property. This derivation yields the size effect law in the original form proposed in Ref. 11:

$$\sigma_N = \frac{Bf_u}{\sqrt{1+\beta}}, \qquad \beta = \frac{d}{d_0} \tag{7}$$

where f_u is an arbitrary measure of material strength, B and d_0 are two empirical constants, and β is called brittleness number [12, 17]. Eq. (7) is obtained both for two-dimensional and three-dimensional similarity [12]. Eq. (7) is applicable not only to notched specimens but also to unnotched structures provided that 1) a large stable crack develops before the maximum load is reached, and 2) the cracks for similar specimens of various sizes are geometrically similar (which has been shown to be often true for concrete structures).

The size effect law in Eq. (7), giving the approximate relation of σ_N to β, is plotted in Fig. 3. For large β such as $\beta>10$, Eq. (7) gives (with an error under 5%) the approximation $\sigma_N = Bf_u\beta^{-1/2}$, which is the size effect exhibited by LEFM. For small β such as $\beta<0.1$, Eq. (7) yields (again with an error under 5%) $\sigma_N = Bf_u$ = constant, that is, there is no size effect and the failure load is proportional to the strength of the

material. For $0.1<\beta<10$, the size effect is transitional between plastic limit analysis and LEFM. In this range, nonlinear fracture mechanics must be employed. Overall, the choice of the method of analysis and the type of failure are as follows [12]:

$$\left.\begin{array}{rl} \beta < 0.1 & \text{plastic limit analysis} \\ 0.1 \le \beta \le 10 & \text{nonlinear fracture mechanics} \\ \beta > 10 & \text{LEFM} \end{array}\right\} \qquad (8)$$

As we see, the brittleness number (Eq. (7)) is capable of characterizing the nature of failure regardless of structure geometry. Other definitions of the brittleness number were proposed by Hillerborg [33], Carpinteri [34], Gogotsi et. al [35], and Homeny et. al [36], but they do not represent absolute (shape-independent) measures of brittleness. Also it may be noted that the basic shape of the transitional size effect curve shown in Fig.3 was known prior to the formulation of Eq.(7) and was graphically sketched, without any formula, by various investigators, e.g. Walsh [37].

Eq. (7) has the same form as Eq. (4). But it should be noted that, in the form of Eq. (4), both parameters are calculated solely from the limiting case of infinite size. In general, this cannot be accurate for very small sizes. For most practical applications, the form of Eq. (4) can nevertheless be accepted. The values of G_f and c_f can be calculated from the size effect law parameters d_0 and B in Eq. (7). Taking the limit of Eq. (2) in which $P = P_u = bd\sigma_N/c_n$, expressing σ_N from Eq. (7), and noting that $\lim \alpha = \alpha_0$ for $d \to \infty$, one gets the formula [12]:

$$G_f = \frac{B^2 f_u^2}{c_n^2 E'} d_0 g(\alpha_0) \qquad (9)$$

Note that G_f can be determined without calculating $g'(\alpha_0)$. Furthermore, from Eqs. (4) and (7) [18]:

$$c_f = \frac{d_0 g(\alpha_0)}{g'(\alpha_0)} \tag{10}$$

The brittleness number, β, may now be expressed as [18]:

$$\beta = \frac{\bar{d}}{c_f} = \frac{g(\alpha_0)\, d}{g'(\alpha_0) c_f} \tag{11}$$

Note that the ratio d/c_f is geometry dependent. Geometry independence is achieved by the factor $g(\alpha_0)/g'(\alpha_0)$. Another formula can be obtained by expressing d_0 from Eq. (9) and substituting it into $\beta = d/d_0$ (Eq. 7). This yields [12, 17]:

$$\beta = \frac{B^2 g(\alpha_0)}{c_n^2} \frac{d}{\ell_u} = \frac{d}{d_0}, \qquad \ell_u = \frac{E' G_f}{f_u^2} \tag{12}$$

where ℓ_u is a material parameter with the dimension of length and d_0 represents the transitional size which corresponds to the intersection of the horizontal and inclined asymptotes in Fig. (3) and depends on structure geometry. The ratio d/ℓ_u is geometry dependent, and geometry independence of β is achieved by the factor $(B/c_n)^2 g(\alpha_0)$.

Since Eq. (12) involves the value of B that characterizes the limit load of a small structure, this equation is more accurate when β is small. On the other hand, Eq. (11) is more accurate when β is large since it ignores B (which characterizes small structures) and is based solely on c_f and the LEFM function $g(\alpha)$ (see Appendix).

In view of the approximate nature of the size effect law in Eq. (7), the infinite size needed to define G_f and c_f must not be interpreted literally. In practice, the infinite size should be assumed as a size

only one order of magnitude beyond the range for which the size effect law has been calibrated by tests or otherwise.

The size effect law in Eq. (7) has been shown to agree quite well with tests on concrete fracture specimens of different sizes and geometries, including three-point bend specimens, centric tension edge-notched specimens, eccentric compression specimens [17], and compact tension specimens [38]. Various specimen shapes were shown to yield about the same value of G_f according to Eq.9. A good agreement was also demonstrated for high strength concrete [20], various rocks [18,19] and aluminum alloys [16], and to some extent also ice [39]. The size effect law in Eq. (7) was shown to also apply for Mode II [40] and Mode III [41] fractures, and for double-punch compression fracture [42].

Some nonlinear fracture models utilize as a material parameter the crack-tip opening displacement at the maximum load, δ_c. Using the well-known LEFM expression for crack opening width, $\delta = (8K_I/E')(s/2\pi)^{1/2}$ where s = distance from the crack tip, and setting $s = c_f$, $K_I = K_{If}$, one gets, for an equivalent elastic crack in an infinitely large specimen [19, 31]:

$$\delta_c = \frac{8K_{If}}{E'\sqrt{2\pi}}\sqrt{c_f} \tag{13}$$

This value can also be determined from the size effect law, after obtaining from this law the values of c_f and K_{If}; $K_{If} = \sqrt{E'G_f}$ (= the value of apparent fracture toughness K_{Ic} for an infinitely large specimen).

In conventional testing, the apparent fracture toughness, K_{Ic}, is

usually determined by methods of linear elastic fracture mechanics without regard to the variations of the size of the fracture process zone, as if $\alpha = \alpha_0$ ($a = a_0$, $c = 0$) at failure. For that case, one gets from Eq. (2) $G_c = P_u^2 g(\alpha_0)/E'b^2 d$. Substituting $P_u^2 = (\sigma_N bd/c_n)^2 = (Bf_u bd/c_n)^2 d_0/(d+d_0)$ and expressing Bf_u by means of G_f from Eq. (9), one gets from Eqs. (5)- (7) and (11):

$$G_c = \tau_N^2 \bar{d} = G_f \frac{d}{d+d_0} = G_f \frac{\bar{d}}{\bar{d}+c_f} = G_f \frac{\beta}{\beta+1} \tag{14}$$

Since $K_{Ic} = \sqrt{EG_c}$, the apparent fracture toughness is found to vary as [17]:

$$K_{Ic} = \tau_N \sqrt{\bar{d}} = K_{If} \left(\frac{\bar{d}}{\bar{d}+c_f} \right)^{1/2} = K_{If} \left(\frac{\beta}{1+\beta} \right)^{1/2} \tag{15}$$

The size effect law in the form of Eqs. (5), (7) or (15) has the advantage that its parameters G_f (or K_{If}) and c_f can be determined from the measured peak loads P_u by linear regression [17-18], provided that a sufficient range of sizes is used, of course. Algebraic rearrangement of Eqs. (5) and (15) yields the linear plots:

$$Y = AX + C \qquad \text{or} \qquad Y' = A'X' + C' \tag{16}$$

in which

$$X = \bar{d} , \quad Y = 1/\tau_N^2 , \quad A = 1/K_{If}^2 , \quad C = c_f/K_{If}^2 \tag{17}$$

$$X' = 1/\bar{d} , \quad Y' = 1/K_{Ic}^2 , \quad A' = c_f/K_{If}^2 , \quad C' = 1/K_{If}^2 \tag{18}$$

Linear regression in each plot yields the slope A or A′ (and their coefficients of variation ω_A and $\omega_{A'}$) of the regression line and its vertical intercept C or C′ (and their coefficients of variations ω_C and

$\omega_{C'}$), from which the values of c_f and K_{If} (as well as their coefficients of variation ω_{cf} and ω_{KIf}) follow. The first plot [Eq. (17)], however, is preferable for calculation of K_{If} because its slope is independent of c_f, and normally the slope of a regression line is more accurate than the intercept.

The fact that the size effect law is amenable to linear regression makes it easy to obtain the statistics characterizing the uncertainty of G_f, c_f, K_{If} and δ_c. To calculate the coefficient of variation of K_{If} and c_f , one may consider the regression of $1/\tau_N^2$ vs. $\bar{d}$ [Eq. (17)]. Because $K_{If} = A^{-1/2}$ and $c_f = C/A$, we have the approximations:

$$\omega_{KIf} = \omega_A/2 \quad \text{and} \qquad \omega_{cf} = (\omega_A^2 + \omega_C^2)^{1/2} \tag{19}$$

where ω_A, ω_C, representing the coefficients of variation of slope A and intercept C, are obtained by standard formulas of linear regression statistics. These formulas are asymptotically exact if the distributions of A and C are normal and ω_A, ω_C are very small. Similarly, if we consider Eq. (18), we have the approximation:

$$\omega_{KIf} = \omega_{C'}/2 \quad \text{and} \qquad \omega_{cf} = (\omega_{A'}^2 + \omega_{C'}^2)^{1/2} \tag{20}$$

Also, since $G_f = K_{If}^2/E'$, the coefficient of variation of the fracture energy can be approximated as:

$$\omega_{Gf} = (4\omega_{KIf}^2 + \omega_{E'}^2)^{1/2} \tag{21}$$

These coefficients of variation characterize not only the material uncertainty but also the model uncertainty. It should be noted that they depend strongly on the range of sizes. By increasing the range, one can

decrease the coefficients of variation.

For certain specimen geometries, for example, the double cantilever specimen, the center-cracked panel loaded on the crack, the chevron-notched specimens [25], the indentation test [2, 26], and the double-punch compression test [42], $g'(\alpha)$ can be negative for a certain initial range of α-values. In this case, which we so far excluded from consideration, $g(\alpha)$ exhibits a minimum at a certain value, α_{min}, of the relative crack length. The peak load for ideally brittle materials (which have a zero-size fracture process zone) occurs when α equals the value of α_0 or α_{min}, whichever is larger. When $g'(\alpha)$ is negative, the crack propagation is more stable and easier to control than the tests in which $g'(\alpha)$ is positive. These observations have led various investigators to expect the toughness value obtained from the peak load of a test where $g'(\alpha_0)$ is initially negative to be size-independent, and therefore a true material property. Later, however, several investigators observed from this type of tests [25] that the calculated apparent fracture toughness values are size-dependent, and that the equivalent relative crack length at the peak load is also size-dependent and greater than both α_{min} and α_0.

An alternative, widely used approach to study the nonlinear fracture properties is to test specimens of constant size but with various crack or notch lengths [8, 24-27, 31, 36]. To get useful results, the size of the specimen should be large with respect to microstructural inhomogeneities as well as crack size. Some specimens with very small cracks (flaws) must also be included in the test series. Cook et al. [26-27] studied the effects of crack and grain sizes, using ceramic

specimens with different fine grain sizes and with flaws created by controlled indentation. Their study [26] shows in a different way the transition from strength plateau to the LEFM-type size effect as the flaw size increases. They emphasize that nonlinear behavior should be considered in extrapolating macroscopic crack test results to structures with microscopic flaws. This is in a way similar to using the present size effect law for extrapolating the results from test-size specimens to smaller structures.

III. Geometry-Dependent R-Curve

The size of the process zone initially grows as the load increases. This causes an increase of the resistance R(c) to fracture growth, representing the energy dissipated per unit specimen width and unit length of advance of the equivalent LEFM crack (whose tip lies roughly in the middle of the process zone). The fracture propagation condition may be written as G = R(c). In the early works [43-44], for the sake of simplicity, the R-curve was assumed to be a material property. Later, however, it was established that the R-curve depends on the specimen shape [e.g. Ref. 17]. As shown in Ref. 15, and refined in Ref. 18, the R-curve for a specimen of given geometry can be calculated from the material parameters obtained by the size effect method;

$$R(c) = G_f \frac{g'(\alpha_1)}{g'(\alpha_0)} \frac{c}{c_f} \tag{22}$$

in which

$$\frac{c}{c_f} = \frac{g'(\alpha_0)}{g(\alpha_0)}\left(\frac{g(\alpha_1)}{g'(\alpha_1)} - \alpha_1 + \alpha_0\right), \qquad \alpha = \alpha_0 + \frac{c}{d} \tag{23}$$

These equations define the R-curve parametrically. After determining G_f and c_f from the size effect law, a series of values of α_1 (which is here a dummy variable) may be chosen, and for each of them the length c of an elastically equivalent (traction-free) crack calculated from Eq. (23), and then R(c) determined from Eq. (22). If c is specified, α_1 may be solved from Eq. (23) by Newton iterations and subsequently R(c) computed. The dependence of the R-curve obtained according to Eq. (22) on the specimen geometry is introduced by the LEFM function $g(\alpha)$.

For readers' convenience, the derivation of Eqs. (22) and (23) is briefly as follows [18]. The energy balance at failure requires that $F(c,d) = G(\alpha,d)-R(c) = 0$ where $\alpha = a/d = \alpha_0+c/d$. If we change the size slightly from d to d+δd but keep the geometric shape (i.e. α_0 = constant), failure now occurs at c+δc, and since G = R or F = 0 must hold also for c+δc, we must have $\partial F/\partial d = 0$. Geometrically, the condition $\partial F/\partial d = 0$ together with $F(c,d) = 0$ means that the R-curve is the envelope of the family of fracture equilibrium curves $F(c,d) = 0$ for various sizes d. Because the R-curve is size-independent, we have $\partial R/\partial d = 0$ and so $\partial G/\partial d = 0$. Now we may substitute $P_u^2 = (\sigma_N bd/c_n)^2 = (Bf_u bd/c_n)^2/(1+d/d_0)$ where $(Bf_u)^2 = c_n^2 E' G_f/d_0 g(\alpha_0)$ (according to Eq. (9)) into $G = P_u^2 g(\alpha)/E' b^2 d$ (Eq. (2)). We thus obtain for the critical states:

$$G(\alpha,d) = G_f \frac{g(\alpha)}{g(\alpha_0)} \frac{d}{d+d_0} \tag{24}$$

Setting $\partial G/\partial d = 0$, and noting that $\partial\alpha/\partial d = \partial\alpha_0/\partial d + \partial(c/d)/\partial d = -c/d^2 = -(\alpha-\alpha_0)/d$ (because $\partial\alpha_0/\partial d = 0$ for geometrically similar structures), we

get:

$$\frac{d}{d_0} = \frac{g(\alpha)}{(\alpha-\alpha_0)g'(\alpha)} - 1 \tag{25}$$

Furthermore, substituting this, along with the relations $(\alpha-\alpha_0)d = c$ and $d_0 = c_f g'(\alpha_0)/g(\alpha_0)$ [from Eq. (10)] into Eq. (25), and setting $G(\alpha,d) = R(c)$, Eq. (22) is proven. Furthermore, elimination of d and G_f from Eqs. (25), (24) and (22), with $G = R$, yields Eq. (23).

The dependence of the fracture toughness on c (i.e. the R-curve of fracture toughness) can be determined from the relation $K_{IR} = \sqrt{E'R}$:

$$K_{IR} = K_{If} \left(\frac{k(\alpha)k'(\alpha)c}{k(\alpha_0)k'(\alpha_0)c_f}\right)^{1/2} \tag{26}$$

To get better insight, we can obtain from Eqs. (22) and (23) the initial and terminal slopes of the R-curve :

$$\left[\frac{\partial R}{\partial c}\right]_{c=0} = \frac{G_f}{c_f}\,\frac{g'(\alpha_m)}{g'(\alpha_0)}, \qquad \left[\frac{\partial R}{\partial c}\right]_{c=c_f} = 0 \tag{27}$$

Here α_m represents the limiting value of α as $d \to 0$, which gives the following condition, from which α_m may be solved:

$$\alpha_m = \alpha_0 + \frac{g(\alpha_m)}{g'(\alpha_m)} \tag{28}$$

The R-curve of Eq. (22) starts from zero, which means that the process zone forms right at the beginning of loading and that there is never any singularity at the crack tip. This type of R-curve also ensues from models in which crack bridging is the only toughening mechanism [32].

Some models for composite materials consider the R-curve to start

from some initial nonzero value [6, 30], interpreted as some small-scale value of the fracture energy (or toughness) of the material. This kind of R-curve implies that the crack tip can sustain, up to some value of K_I, a singular stress field without showing any damage, which is unreasonable. Nevertheless, for some composite materials (especially fiber composites) with fine-grained matrix, in which fibers or coarse particles create a relatively large bridging zone, a non-zero initial value of R-curve may be practical, and may be explained by assuming that the matrix itself will, on a smaller scale, exhibit a microscopic, relatively fast rising R-curve.

Eq. (22) or (26) applies only as long as the fracture process zone grows and remains attached to the notch tip, which is approximately up to the peak load. For the post-peak regime, in which the fracture process zone gets detached from the notch tip [19], the value of R does not follow the R-curve [Eq. (22)] but must be kept constant and equal to the value of R at the peak load. This behavior was first introduced as a hypothesis and was subsequently verified by comparisons of predicted load deflection curves with measurements [19-20]. Without keeping the postpeak value of R constant, close agreement could not be obtained.

Based on fracture tests, various investigators have found the R-curves to depend on the size and geometry of the specimen as well as the notch length [8-9, 16, 19-20]. Our preceding calculation of the R-curve indicates that this phenomenon has principally two sources: 1) The effect of $g'(\alpha)$ and $g(\alpha)$ which depend on specimen geometry [18], 2) The fact that R deviates from the Eq. (22) after the peak load, as just explained. Other minor effects, however, may contribute, too.

It may also be noted that, according to some tests on ceramic specimens [8-9], interaction of the process zone and the specimen boundary can even lead to a falling R-curve after a plateau is passed. The declining end of the R-curve may correspond to shrinking of the process zone as the end of ligament is approached; this would be significant for specimens with very small uncracked ligaments and would affect only the final part of the post-peak response.

IV. Application of the Size Effect Method to Ceramics

To demonstrate applicability of the size effect method, the test results of McKinney and Rice [23] on slip-cast fused silicon dioxide (SiO_2) and on two types of silicon carbide (SiC) refractories have been analyzed. The slip-cast fused silica used in these tests is a material consisting 99% of SiO_2, with a relatively large porosity (12%). It has the density 1900 kg/m^3, grain size 10-20 μm, and modulus of elasticity E = 57.9 GPa. The silicon carbides used in these tests had the porosity of 15%, density 2600 kg/m^3, and maximum grain size 2000 μm. So these ceramics, due to their grain structure, have a heterogeneity which may be significant for small sizes. One of the silicon carbides tested, of the type CN-137, consisted 78% of SiC, and 20% of Si_3N_4, with modulus of elasticity E =130 GPa. The other silicon carbide tested, of type CN-163, consisted 85% of SiC and 13% of Si_2ON_2, and had a modulus of elasticity E = 140 GPa.

All the test specimens were three-point-bend beams (see Fig.4). Their dimensions and failure loads are given in Table I, in which L = beam span, d = beam depth (chosen as the characteristic dimension), b =

width, and a_0 = length of the notch cut at midspan, and P_u = failure load. The specimens were not geometrically similar exactly, but the length of all the specimens was nearly four-times their depth.

As reported by the authors [23], the maximum loads of the slip-cast fused silica specimens of sizes d = 5.1 mm and 6.5 mm exhibited an unexplained drop, and therefore these results have been deleted from the data base. It should also be mentioned that the successive groups of smaller specimens were cut from the broken halves of the larger specimens after their fracture. Since the larger specimens had only small stresses at the crack location in the smaller, subsequently cut specimens, the results should not have been affected appreciably.

The three-point-bend beam results may be analyzed on the basis of Srawley's [29] approximation of function $k(\alpha)$:

$$k(\alpha) = \frac{1.5L\sqrt{\alpha}}{d} \frac{1.99-\alpha(1-\alpha)(2.15-3.93\alpha+2.7\alpha^2)}{(1+2\alpha)(1-\alpha)^{3/2}} \tag{29}$$

The linear regression plots based on Eq. (17) are shown in Fig. 5. From the linear regressions based on Eqs. (17) and (18), the values of G_f and c_f have been calculated and are listed in Table II. Aside from the linear regression, the values of G_f and c_f have also been obtained by direct nonlinear optimum fitting of Eqs. (5) and (15) to the test data, which was accomplished by a standard library subroutine using Levenberg-Marquardt algorithm for nonlinear optimization.

Table II also gives the coefficients of variation ω_τ and ω_K of the vertical deviations of the test data from the regression lines with the ordinates τ_N and K_{Ic}, respectively:

$$\omega_\tau = \left[\frac{\sum (\tau_{Ne} - \tau_{Np})^2}{n - 2} \right]^{1/2} \left[\frac{\sum \tau_{Ne}^2}{n} \right]^{-1/2} \tag{30}$$

$$\omega_K = \left[\frac{\sum (K_{Ice} - K_{Icp})^2}{n - 2} \right]^{1/2} \left[\frac{\sum K_{Ice}^2}{n} \right]^{-1/2} \tag{31}$$

where n = total number of specimens and subscripts e and p refer to the experimental values [Eqs. (6) and (2)] and the corresponding predicted values [Eqs. (5) and (15)], respectively. The coefficients of variation for the deviations (ω_τ or ω_K) in the corresponding nonlinear regression based on Eqs. (5) or (15) are, in all the cases, smaller. The reason is that the objective functions of the optimization are different.

The values of ω_{KIf} and ω_{cf} [Eqs. (19) and (20)] are also shown in Table II. We see that although the values of fracture toughness K_{If} for all optimization procedures are close, the results for the effective fracture process zone length c_f show much greater dispersion and are indicative only of the order of magnitude of c_f. Also it is seen that ω_{cf} is generally much larger than ω_{KIf}. So the effective length of the fracture process zone is much more uncertain than the fracture toughness. This is probably not just due to the modeling but is a natural property. Evidently, c_f is more sensitive to the range and scatter of the test results than is K_{If} or G_f, same as observed previously for concrete and rock. However, a broader range of specimen sizes could be used to determine c_f more accurately. For the rest of the present paper, the results of the nonlinear regression based on Eq. (5) will be used.

The ratio of c_f (Table II) to the grain size is found here to be a

relatively large number (30 to 60) for SiO_2. Swanson et al. [5] have measured the crack bridging zone length for alumina (with mean grain size of 20 μm, about the same as SiO_2), and found it to be about 100-times the grain size. The spacing of the bridging grains was about 2 to 5 grain size. Based on this observations, and considering that the fracture process zone length is roughly $2c_f$, the present value c_f = 0.6 mm seems reasonable for SiO_2. Also note that the slip-cast fused silicon dioxide used has a significant amount of porosity [23], which could help in toughening by shielding of crack tip.

The effective critical crack-tip opening displacements, δ_c, as calculated from Eq. (13) for SiO_2, SiC(CN-137), and SiC(CN-163), are 0.97μm, 34.54μm, and 53.12μm, respectively.

Fig. 6 gives the normalized values of the nominal stress at failure as a function of the brittleness number, β, in logarithmic scales. These plots indicate that the largest specimens of SiO_2 which were tested may be considered to be sufficiently large for their behavior to approach linear elastic fracture mechanics (this is also revealed by the brittleness number range β = 1.5 to 10). For SiC(CN-137) specimens, on the other hand, the response is far from linear elastic fracture mechanics and is actually closer to plastic limit analysis, as is clear from Fig.6b as well as the range of the brittleness number, which was β = 0.06 to 0.25. The brittleness number range for SiC (CN-163) specimens was β = 0.9 to 3.9 which is in nonlinear zone. In this regard it should be noted that the actual specimen sizes for both SiC materials CN-137 and CN-163 were almost the same, although the ranges of their brittleness numbers are very different. This illustrates that one cannot really

compare the apparent fracture toughness values measured on identical specimens of different materials. Such comparisons should be based only on the asymptotic values for the size approaching infinity, which can be obtained only if the size effect law is known.

Fig.7 shows the values of the apparent fracture toughness K_{Ic} of these ceramics as a function of the brittleness number, β, based on Eq.(15). A trend similar to Eq. (15) can be observed in various other experimental results [25, 26, 38]. From Fig.7 one can observe that, in contrast to the nominal strength values, the fracture toughness of the CN-137 type of SiC is more size dependent than it is for the other two ceramics tested, for the present size range. It means that, for this size range, the behavior of the CN-137 type of SiC is farther from linear elastic fracture mechanics than it is for the other two ceramics. By comparing Figs. 6 and 7 with Table II, it can be seen that, for the case when $\beta>1$ for all the specimens, linear regression based on Eq. (17) is not suitable (the case of SiO_2). The best test data are those which, similar to the case of SiC of the type CN-163, include values for both $\beta<1$ and $\beta>1$.

The foregoing results highlight the dependence of the conventional values of fracture energy, fracture toughness, and effective process zone length on the microstructure of materials, as already emphasized by some researchers [26, 45-46]. It should be noted that, in order to allow meaningful comparisons of fracture properties of ceramics with different grain size and different microstructures, the specimens used should have the same range of brittleness number rather than the same sizes. The size effect tests based on the size effect law make such comparisons possible.

Fig. 8 shows the effect of the notch length on the apparent fracture toughness of the three-point-bend specimens of different sizes, as calculated on the basis of Eq. (15) for the present geometry. It may be noticed from this figure that, for large sizes and for $0.15 < \alpha_0 < 0.80$, the apparent fracture toughness appears to be nearly independent of the notch length. In this kind of test, in which all the specimens have the same size but various notch lengths, it is important that the range of α_0 values include values below 0.15; otherwise it is hardly possible to detect the nonlinearity of fracture [26-27]. Fig. 8 also shows the importance of the choice of specimen size in this kind of experiments.

Fig. 9 shows the R-curves calculated from the size effect law according to Eqs. (22) and (23) for the present three-point-bend specimen geometry and for different relative notch lengths. For each different notch length, the R-curve is different, however the differences are insignificant for notch lengths exceeding 0.33 of the beam depth. It should be pointed out that for different types of fracture specimens, such as eccentric compression specimens, tensile edge-notched specimens, and compact tension specimens, the size effect method yields very different R-curves [18].

The equivalent length of the fracture process zone can be compared with the ligament length;

$$\alpha_r = \frac{c}{d - a_0} = \frac{\alpha - \alpha_0}{1 - \alpha_0} \tag{32}$$

Aside from α_m [Eq. (28)], the limiting values of α_r for $d \to 0$, denoted as α_{rm}, are plotted in Fig. 10 as a function of α_0. We see that, for $\alpha_0 > 0.15$, $\alpha_{rm} \simeq 0.24$, and that the plot of α_m vs. α_0 is a straight line.

V. Calculation of Load-Deflection Curve from Size Effect

Knowing the R-curve, one can easily calculate the load-point displacement, u [9, 19]. The complementary potential energy of the structure is $\Pi^* = C(a)P^2/2$ where C(a) = compliance at crack length a. The energy release rate at constant P is $G = (d\Pi^*/da)/b = P^2[dC(a)/da]/2b$. Substituting Eq. (2) for G, one gets a differential equation for C(a). Integration at constant P furnishes:

$$C(a) = C_0 + \frac{2}{bE'}\int_0^{\alpha} g(\alpha')d\alpha' \tag{33}$$

At the same time, from Eq. (2) for G = R [Eq. (22)]:

$$P = b\sqrt{\frac{E'd}{g(\alpha)}R(c)} \quad , \qquad u = C(a)P \tag{34}$$

Choosing various values of α, one can calculate the corresponding values of P, C(a) and u. (For an alternative derivation based on Castigliano' theorem, which is not as simple but is more instructive, see Ref. 47.)

Since there exist no pertinent size effect test data for ceramics, we illustrate the application of Eqs. (33) and (34) by showing in Fig. 11 the results for limestone [19], which behaves similar to ceramics. We see that the measured load-displacement curve in Fig. 11b has been predicted quite closely. The prediction was based solely on the measured maximum loads of geometrically similar three-point-bend specimens whose sizes ranged as 1:4. The size effect plot from which the material parameters were determined is shown in Fig. 11a. An equally close agreement has been achieved for high strength concrete with fine aggregate [20], which also

behaves similar to ceramics.

Note that, conversely, by fitting the measured P(u) curve according to Eqs. (33) and (34), one can obtain estimates of the values of E', G_f and c_f.

The work of fracture [21, 48], W_f, can be obtained from the area under the P(u) curve, as well as from the area under the R(c) curve [9]. It may be noted that the value of average fracture energy $\bar{G}_f = W_f/\ell$ where ℓ is the length of ligament [for three-point-bend specimens $\ell = (1-\alpha_0)d$] is in this manner found to be strongly size-dependent as well as shape-dependent [19-21, 38 and 49]. This shows that the work-of-fracture method does not yield unambiguous information on fracture energy, except if the results are extrapolated to infinite size, as proposed by Planas and Elices [49].

VI. Sensitivity of Fracture Properties to Micromechanics of Process Zone

As the foregoing theory illustrates, the size effect law yields, for a given structure geometry, a certain unique shape of the R-curve. Vice versa, from a measured R-curve, one can calculate the size effect curve $\sigma_N(d)$ [15]. Furthermore, from either the size effect law or from the R-curve, one can calculate the corresponding relation $\sigma_b(\delta)$ of bridging traction σ_b vs. crack opening δ in softening Dugdale-Barenblatt-type models. Vice versa, from $\sigma_b(\delta)$ one can calculate both the R-curve and the size effect curve for a given geometry [31-33]. Thus, the size effect law, the R-curve, and the $\sigma_b(\delta)$-curve, each of which can completely characterize the fracture properties, are uniquely related to each other.

All the aforementioned characterizations are sensitive to specimen or structure geometry. However, the sensitivity of the size effect law appears to be the least, and much less than that of the R-curve. This transpired from the experimental results for concrete and rock [17-19], which revealed that Eq. (7) works about equally well for very different specimen geometries.

From the viewpoints of micromechanics and development of new, tougher materials, there are of course great differences between various toughening mechanisms in ceramics, such as crack bridging, the dilatant transformation toughening, or crack-tip shielding by a microcracking zone induced by material heterogeneity. The details of such mechanisms and the microstructure properties which govern them influence, as is well known, the effective $\sigma_b(\delta)$ relation for crack bridging. Likewise, they do influence the shape of the size effect curve at least to some extent.

However, they are likely to influence the shape of the size effect curve only insignificantly because the size effect law has by now been shown to be applicable to materials of very different microstructures and toughening mechanisms. The present approximate, two parameter form of the size effect law in Eq. (7) suffices only for a global, first order approximation to deviations from linear elastic fracture mechanics, which is the sole object of interest here. The various possible mechanisms will probably influence only the additional parameters of a more complicated size effect law [see Appendix]. The only significant property which matters from the present viewpoint of macroscopic nonlinear fracture characterization is that there exists at the crack tip a process zone of a finite, nonnegligible size, and that the maximum possible effective

size c_f of this zone is, in fracture situations, approximately a material property. Knowledge of c_f cannot be used to distinguish between various toughening mechanisms such as crack bridging or microcracking; rather, c_f lumps their effects together.

The type of microstructure and the precise form of the toughening mechanism in the process zone decide, of course, the values of G_f (or K_{If}) and c_f. They will have to be considered if the values of G_f and c_f should be predicted from microstructure properties, rather than just determined empirically, by the fitting of test data, as already explained.

By various more detailed assumptions about the micromechanics of the process zone, further information can be extracted from size effect tests. Planas and Elices [31] computed the relation between c_f and certain other characteristics of the fracture process zone for crack bridging models with various softening $\sigma_b(\delta)$-relations. For example, for a linearly softening $\sigma_b(\delta)$ relation, which is widely used because of its simplicity and has the form of $\sigma_b = f_t(1 - \delta/\delta_f)$ where $\sigma_b(\delta_f) = 0$ and $G_f = f_t\delta_f/2$, Planas and Elices obtained $c_f = 0.420\ \ell_0$, where $\ell_0 = E'G_f/f_t^2$. For the same relation, Horii et al. [32] calculated the maximum length (for $d \to \infty$) of the fracture process zone to be $\ell_f = 0.732\ \ell_0$, from which, according to the linear softening model, $c_f = 0.573\ \ell_f$. For $\sigma_b(\delta)$-curves with a long tail, the ratio ℓ_f/c_f is larger, while the bridging stresses closer to the initial notch are smaller.

By estimating ℓ_0 from c_f for an assumed $\sigma_b(\delta)$ relation, one can calculate the local tensile strength as $f_t = (E'G_f/\ell_0)^{1/2}$. It should be noted that f_t could also be estimated by proper plastic analysis of the

structure for $d \to 0$, which does not depend on the form of the $\sigma_b(\delta)$ relation. These two values for f_t should be the same, which gives a constraint that can help in identifying a realistic function for $\sigma_b(\delta)$ from a generalized size effect law with three parameters [see Eqs. (A5) and (A6) in the Appendix].

VII. Conclusions

1. The size effect law in Eqs. (4), (5) and (7) gives a global characterization of nonlinear fracture properties of ceramics with toughening mechanics. It represents a first-order approximation to the deviations from linear elastic fracture mechanics. Its form is independent of the detailed toughening mechanics, but knowledge of this mechanism will be needed for predicting the size effect law parameters rather than measuring them.

2. The size effect law provides a simple and unambiguous way to determine from the maximum load data on geometrically similar specimens of different sizes the values of: (1) the fracture energy, G_f (or fracture toughness, K_{If}), and (2) the effective length of the fracture process zone, c_f. From these, one can also further obtain other parameters of the process zone such as the effective critical crack-tip opening displacement, δ_c, as well as the R-curve, which is geometry dependent. The material parameters can be identified by nonlinear optimization as well as by linear regression. Knowing c_f, one can determine for the given structural geometry the brittleness number, which characterizes the type of fracture behavior of any specimen or structure.

3. Silicon carbide [SiC(CN-137)] beam specimens of cross section

depths from 7 mm to 37 mm exhibit behavior that is closer to plastic limit analysis than to linear elastic fracture mechanics (LEFM). On the other hand, fracture behavior of the SiC(CN-163) specimens of the same sizes is transitional between plastic limit analysis and LEFM. The behavior of slip-cast fused silica beam specimens of depths from 5 mm to 32 mm is closer to LEFM.

Appendix

Derivation of Size Effect Law by Dimensional Analysis. - The size effect law for geometrically similar structures [Eq. (7)] can be derived most generally by dimensional analysis and similitude arguments [11] on the basis of the following hypotheses: (I) The energy release of the structure is a function of both: a) the length of fracture, a, and b) the characteristic size of the fracture process zone, c_f. (II) The length a at maximum load is not negligible compared to structure size d.

The total energy release due to fracture must be expressible in the form:

$$W = \frac{\sigma_N^2 bd^2}{2E'} f(\vartheta_1, \vartheta_2) , \qquad \vartheta_1 = \frac{a}{d} , \quad \vartheta_2 = \frac{c_f}{d} \tag{A1}$$

where $\sigma_N = c_n P_u/bd$; ϑ_1, and ϑ_2 are independent nondimensional parameters (their number follows from Buckingham's theorem of dimensional analysis), and f is a certain function which may be expected to be smooth. From the crack propagation condition $\partial W/\partial a = bG_f$ we get:

$$\sigma_N^2 = \frac{2E'G_f}{d} \left[\frac{\partial f(\vartheta_1, \vartheta_2)}{\partial \vartheta_1} \right]^{-1} \tag{A2}$$

We now choose the state $\vartheta_2 = 0$ (which corresponds to $d \to \infty$) as the reference state, and expand $\partial f/\partial\vartheta_1$ into Taylor series about this state; i.e. $\partial f/\partial\vartheta_1 = f_0 + f_1\vartheta_2 + f_2\vartheta_2^2 + f_3\vartheta_2^3 + \ldots$ where f_0, f_1, .. are constants if geometrically similar shapes (same ϑ_1) are considered. Substitution of this series into Eq. (A2) and truncation of the series after the linear term provides:

$$\sigma_N = \left(\frac{2E'G_f}{f_1 c_f + f_0 d} \right)^{1/2} \tag{A3}$$

This yields Eq. (7) if one denotes $d_0 = f_1 c_f / f_0$ and $Bf_u = (2E'G_f/f_1 c_f)^{1/2}$.

It may be noted that a more general size effect law, $\sigma_N = Bf_u(1+\beta^r)^{-1/2r}$, with an additional parameter r [14], can be derived under some less restrictive assumptions. However, fitting of test data for concrete specimens of various geometries indicates the optimum value to be $r \approx 1$.

Energy Explanation of Size Effect Law. - It is instructive to give also a simple energy explanation of the size effect. Consider the uniformly stressed specimen of width d in Fig. 12, in which a crack of length a_0, with a fracture process zone of width h, propagates from the left. The width h is a material property; $h = \kappa c_f$ where κ = constant. It may be imagined that the formation of a crack band of thickness h reduces to zero the strain energy density $\sigma_N^2/2E'$ in the cross hatched area. When the crack extends by Δa, the additional strain energy that is reduced comes from the densely cross-hatched strip of horizontal dimension Δa (Fig. 12). If the failure modes are geometrically similar, then the larger the panel, the larger is the crack band at failure. Consequently, the area of the densely cross-hatched strip is also larger; it equals $h\Delta a$ +

$2ka\Delta a$ where k = constant depending on the shape of the structure. This illustrates that, for the same σ_N, a larger structure releases more stored energy into the same extension Δa of the crack. But since the energy that can be dissipated by the crack extension Δa is independent of the structure size, the σ_N-value for a longer specimen must be smaller; hence the size effect. To quantify this argument, the energy released from the densely cross-hatched strip (per unit thickness) is :

$$G = \frac{1}{b}\frac{\partial W}{\partial a} = \frac{1}{\Delta a}\,(h\Delta a + 2ka_0\Delta a)\,\frac{\sigma_N^2}{2E'} = G_f \qquad (A4)$$

Solving for σ_N, one gets Eq. (7) with the notation $d_0 = h/2k\alpha_0$, $Bf_u = (2E'G_f/h)^{1/2}$ where $\alpha_0 = a_0/d$ = constant.

Generalization of Size Effect Law. - Some test results [19] indicate deviations from the size effect law for very small sizes. That there might be some discrepancies for very small sizes is suggested by the fact that there are two formulas for the brittleness number (Eqs. 11 and 12) and two ways to get size effect parameters. One uses both the LEFM characteristic G_f (pertinent to $d \to \infty$) and the plastic characteristic Bf_u ($d \to 0$). The other one only uses crack characteristics , G_f and c_f ($d \to \infty$). A simple empirical generalization, which makes an appreciable difference only for sufficiently small sizes, can be introduced by replacing Eq. (5) with:

$$\tau_N = \left(\frac{E'G_f}{\bar{c} + \bar{d}}\right)^{1/2}, \qquad \bar{c} = c_f\,\frac{\bar{d} + c_0}{\bar{d} + c_f} \qquad (A5)$$

in which τ_N and $\bar{d}$ are defined by Eq. (6), and its three parameters ($E'G_f$, c_f, c_0) can be obtained by nonlinear regression analysis of the test

data. The extra parameter, c_0, determines the location of the horizontal asymptote (Fig. 13), which is given by $\tau_N = (E'G_f/c_0)^{1/2}$. Considering this horizontal asymptote and Eq. (6), and writing the plastic limit load as $P_u = c_p bdf_t$, one obtains:

$$f_t = \left(\frac{E'G_f}{\ell_0} \right)^{1/2}, \qquad \ell_0 = c_p^2 g'(\alpha_0) c_0 \tag{A6}$$

where c_p depends on geometry and the material model used in plastic analysis. Eq. (A6) enables estimating ℓ_0 and f_t only from size effect test data. A broader range of test data would be needed to determine all three parameters by regression. Short of that, c_0 can be determined if the tensile strength, f_t, in the fracture process zone, representing the maximum value of bridging stresses, is known or determined separately from other types of tests (which is not an easy task).

Consider for example a very small three-point bend specimen (Fig. 4), to which plastic limit analysis can be applied. Since the compression strength (in the brittle materials) is much larger than f_t, the stress throughout the ligament is uniform and equal to f_t, except at the top face which can sustain a large compression force. Thus, as $d \to \infty$, the bending moment in the ligament is $M_u = f_t b(d-a_0)^2/2$, which yields $P_u = 4M_u/L = 2f_t(1-\alpha_0)^2 bd^2/L$; then $c_p = P_u/bdf_t = 2(1-\alpha_0)^2 d/L$. (Note that the value of c_p is twice [34] the value obtained by plastic analysis of a beam with equal yield stresses in tension and compression.)

By assuming a crack bridging model with a specific form for $\sigma_b(\delta)$, f_t can be related to c_f (see Section VI), which means the three parameters ($E'G_f$, c_f and c_0) in Eq. (5A) cannot be independent. By accepting $E'G_f$ as an independent parameter, c_0 and c_f can be related as

$c_0 = \mu c_f$ where μ will depend on the form of $\sigma_b(\delta)$ and the geometry of the structure. As an example, for the linear softening curve $\sigma_b(\delta)$ for which $c_f = 0.420\ell_0$ [31], we obtain $\mu = c_0/c_f = 2.384/c_p^2 g'(\alpha_0)$ using Eq. (A6). Fig. 13 is plotted for $\mu = 0.211$ which corresponds, for example, to a three-point bend specimen with a linear softening curve, and with $L/d = 4$ and $\alpha_0 = 1/3$ (from which $c_p = 2/9$ and $g'(\alpha_0) = 229$).

Acknowledgments: The basic concept of the size effect method was developed under AFOSR contract F49620-87-C-0030DEF with Northwestern University, and extensions to R-curve behavior were partially funded by the Center for Advanced Cement-Based Materials at Northwestern University (NSF grant DMR-8808432). Thanks are also due to Dr. K. R. McKinney of the Naval Research Laboratory, Washington, D.C., for kindly providing the details of his test results (Table I), supplementing Ref. 23. The second author wishes to express his thanks for a scholarship he received from Sharif University of Technology, Tehran, Iran.

References

1. B. Budiansky, J.W. Hutchinson, and J.C. Lambropoulos, "Continuum Theory of Dilatant Transformation Toughening in Ceramics," Int. J. Solids Structures, 19 [4] 337-55 (1983).

2. D.R. Clarke and K.T. Faber, "Fracture of Ceramics and Glasses," J. Phys. Chem. Solids, 48 [11] 1115-57 (1987).

3. A.G. Evans and K.T. Faber, "Crack-Growth Resistance of Microcracking Brittle Materials," J. Am. Ceram. Soc., 67 [4] 255-60 (1984).

4. M. Ortiz, "Microcrack Coalescence and Macroscopic Crack Growth Initiation in Brittle Solids," Int. J. Solids Structures, 24 [3] 231-50 (1988).

5. P.L. Swanson, C.J. Fairbanks, B.R. Lawn, Y.-W. Mai, and B.J. Hockey, "Crack-Interface Grain Bridging as a Fracture Resistance Mechanism in Ceramics: I, Experimental Study on Alumina," J. Am. Ceram. Soc., 70 [4] 279-89 (1987).

6. Y.-W. Mai and B.R. Lawn, "Crack-Interface Grain Bridging as a Fracture Resistance Mechanism in Ceramics: II, Theoretical Fracture Mechanics Model," J. Am. Ceram. Soc., 70 [4] 289-94 (1987).

7. S. Nemat-Nasser and M. Hori, "Toughening by Partial or Full Bridging of Cracks in Ceramics and Fiber Reinforced Composites," Mechanics of Materials, 6 [3] 245-69 (1987).

8. M. Sakai, J.-I. Yoshimura, Y. Goto, and M. Inagaki, "R-Curve Behavior of a Polycrystalline Graphite: Microcracking and Grain Bridging in the Wake Region," *J. Am. Ceram. Soc.*, 71 [8] 609-16 (1988).

9. M. Sakai, and M. Inagaki, "Dimensionless Load-Displacement Relation and Its Application to Crack Propagation Problems," J. Am. Ceram. Soc., 72 [3] 388-94 (1989).

10. F. Erdogan and P.F. Joseph, "Toughening of Ceramics through Crack Bridging by Ductile Particles," J. Am. Ceram. Soc., 72 [2] 262-70 (1989).

11. Z.P. Bažant, "Size Effect in Blunt Fracture: Concrete, Rock, Metal," *J. Engng. Mech., ASCE*, 110 [4] 518-35 1984.

12. Z.P. Bažant, "Fracture Energy of Heterogeneous Materials and Similitude"; pp. 229-41 in Fracture of Concrete and Rock, *SEM-RILEM Int. Conf.*, Houston 1987. Edited by S.P. Shah and S.E. Swartz. Springer-Verlag, NY, 1989.

13. Z.P. Bažant, "Mechanics of Fracture and Progressive Cracking in Concrete Structures"; pp. 1-94 in Fracture Mechanics of Concrete: Structural Application and Numerical Calculation. Edited by G.C. Sih and A. DiTommaso. Martinus Nijhoff Publishers, Dordrecht, 1985.

14. Z.P. Bažant, "Mechanics of Distributed Cracking," *Appl. Mech. Rev.*, 39 [5] 675-705 (1986).

15. Z.P. Bažant, J.-K. Kim, and P.A. Pfeiffer, "Nonlinear Fracture Properties from Size Effect Tests," J. Struct. Engng., 112 [2] 289-307 (1986).

16. Z.P. Bažant, S.-G. Lee, and P.A. Pfeiffer, "Size Effect Tests and Fracture Characteristics of Aluminum," *Engng. Fracture Mech.*, 26 [1] 45-57 (1987).

17. Z.P. Bažant and P.A. Pfeiffer, "Determination of Fracture Energy from Size Effect and Brittleness Number," *ACI Mater. J.*, 84 [6] 463-80

(1987).

18. Z.P. Bažant and M.T. Kazemi, "Determination of Fracture Energy, Process Zone Length and Brittleness Number from Size Effect, with Application to Rock and Concrete," Rep. No. 88-7/498d, Center for Concrete and Geomaterials, Northwestern Univ., Evanston, Ill. (1988); also *Int. J. Fracture*, in press.

19. Z.P. Bažant, R. Gettu, and M.T. Kazemi, "Identification of Nonlinear Fracture Properties from Size Effect Tests and Structural Analysis Based on Geometry-Dependent R-curve," Rep. No. 89-3/498p, Center for Advanced Cement-Based Materials, Northwestern Univ., Evanston, Ill. (1989); submitted for possible publication to Int. J. of Rock Mech. and Min. Sci.

20. R. Gettu, Z.P. Bažant, and M.E. Karr, "Fracture Properties and Brittleness of High Strength Concrete," Rep. No. 89-10/B627f, Center for Advanced Cement-Based Materials, Northwestern Univ., Evanston, Ill. (1989); submitted to ACI Material Journal.

21. H.G. Tattersall and G. Tappin, "The Work of Fracture and Its Measurement in Metals, Ceramics and other Materials," J. Materials Science 1 [3] 296-301 (1966).

22. G.K. Bansal, W.H. Duckworth, and D.E. Niesz, "Strength-Size Relation in Ceramic Materials: Investigation of an Alumina Ceramic," J. Am. Ceram. Soc., 59 [11-12] 472-78 (1976).

23. K.R. McKinney and R.W. Rice, "Specimen Size Effects in Fracture Toughness Testing of Heterogeneous Ceramics by the Notch Beam Method"; pp. 118-26 in Fracture Mechanics Methods for Ceramics, Rocks, and Concrete, ASTM STP 745. Edited by S.W. Freiman and E.R. Fuller,Jr. ASTM,

1981.

24. R. F. Pabst, "Determination of K_{Ic}-Factors with Diamond-Saw-Cuts in Ceramic Materials"; pp. 555-65 in Fracture Mechanics of Ceramics, Vol.2. Edited by R.C. Bradt et al. Plenum Press, New York, 1974.

25. J.L. Shannon, Jr., and D.G. Munz, "Specimen Size and Geometry Effects on Fracture Toughness of Aluminum Oxide Measured with Short-Rod and Short-Bar Chevron-Notched Specimens"; pp. 270-80 in Chevron-Notched Specimens: Testing and Stress Analysis, ASTM STP 855. Edited by J.H. Underwood et al. ASTM, Philadelphia, 1984.

26. R.F. Cook, B.R. Lawn, and C.J. Fairbanks, "Microstructure-Strength Properties in Ceramics: I, Effect of Crack Size On Toughness," J. Am. Ceram. Soc., 68 [11] 604-15 (1985).

27. R.F. Cook, C.J. Fairbanks, B.R. Lawn, and Y.-W. Mai, "Crack Resistance by Interfacial Bridging: Its Role in Determining Strength Characteristics," J. Mater. Res., 2 [3] 345-56 (1987).

28. F.E. Buresch, "A Structure Sensitive K_{Ic}-Value and Its Dependence on Grain Size Distribution, Density and Microcrack Interaction," ; pp. 835-47 in Fracture Mechanics of Ceramics, Vol. 4, Edited by R. C. Bradt et al. Plenum Press, New York, 1978.

29. H. Tada, P.C. Paris, and G.R. Irwin, The Stress Analysis of Cracks Handbook 2nd. ed., Paris Productions, St. Louis, 1985.

30. R.F. Cook, "Transient Fracture Resistance in the Weak Toughening Limit"; pp. 2747-55 in Advances in Fracture Research (ICF7, Houston, 1989), Vol. 4. Edited by K. Salama et. al Pergamon Press, Oxford, 1989.

31. J. Planas and M. Elices, "Size-Effect in Concrete Structures: Mathematical Approximations and Experimental Validation"; pp. 462-76 in

Cracking and Damage, Strain Localization and Size Effect, Proceedings of France-US Workshop, Cachan, France, 1988. Edited by J. Mazars and Z.P. Bažant. Elsevier, London, 1989.

32. H. Horii, Z. Shi, and S.-X. Gong, "Models of Fracture Process Zone in Concrete, Rock, and Ceramics"; pp. 104-15 in Cracking and Damage, Strain Localization and Size Effect, Proceedings of France-US Workshop, Cachan, France, 1988. Edited by J. Mazars, and Z.P. Bažant. Elsevier, London, 1989.

33. A. Hillerborg, "The Theoretical Basis of a Method to Determine the Fracture Energy G_F of Concrete," Materials and Structures, 18 [106] 291-96 (1985).

34. A. Carpinteri and S. Valente, "Size-scale Transition from Ductile to Brittle Failure: A Dimensional Analysis Approach"; pp. 477-90 in Cracking and Damage, Strain Localization and Size Effect, Proceedings of France-US Workshop, Cachan, France, 1988. Edited by J. Mazars and Z.P. Bažant. Elsevier, London, 1989.

35. G.A. Gogotsi, Ya.L. Groushevsky, and K.K. Strelov, "The Significance of Non-Elastic Deformation in the Fracture of Heterogeneous Ceramic materials," Ceramurgia (Ceramics) Int. 4 [3] 113-8 (1978).

36. J. Homeny, T. Darroudi, and R.C. Bradt, "J-Integral Measurements of the Fracture of 50% Alumina Refractories," J. Am. Ceram. Soc., 63 [5-6] 326-31 (1980).

37. P.F. Walsh, "Fracture of Plain Concrete," Indian Concrete J., 46 [11] 469-70 & 476 (1972).

38. F.H. Wittman, K.Rokugo, E. Brühwiler, H. Mihashi, and P. Simonin, "Fracture Energy and Strain Softening of Concrete as Determined

by Means of Compact Tension Specimens," Materials and Structures, 21 [121] 21-32 (1988); also E. Brühwiler, "Fracture Mechanics of Dam Concrete Subjected to Quasi-Static and Seismic Loading Conditions," Thesis No. 739, Laboratory for Building Materials, Swiss Federal Institute of Technology, Lausanne (1988) (in German).

39. Z.P. Bažant and J.-K. Kim, "Fracture Theory for Nonhomogeneous Brittle Materials with Application to Ice"; pp. 917-30 in *Civil Engng. in the Arctic Offshore*, Proc. Arctic'85, San Francisco. Edited by F.L. Bennett and J.L. Machemehl. ASCE, NY, 1985.

40. Z.P. Bažant and P.A. Pfeiffer, "Shear Fracture Tests of Concrete," *Mater. Struct.*, 19 [110] 111-21 (1986).

41. Z.P. Bažant and P.C. Prat, "Measurement of Mode III Fracture Energy of Concrete," Nuclear Engng. and Design, 106 [1] 1-8 (1988).

42. P. Marti, "Size Effect in Double-Punch Tests on Concrete Cylinders," ACI Material Journal, 86 [6] 597-601 (1989).

43. J.M. Krafft, A.M. Sullivan, and R.W. Boyle, "Effect of Dimensions on Fast Fracture Instability of Notched Sheets"; pp. 8-28 in Proc. Crack Propagation Symposium, Vol. 1. College of Aeronautics, Cranfield, U.K., 1961.

44. G. R. Irwin, "Fracture Testing of High-Strength Sheet Materials: A Report of a Special ASTM Committee"; pp. 29-40 in ASTM Bulletin. January 1960.

45. R.W. Rice, "Test-Microstructural Dependence of Fracture Energy Measurements in Ceramics"; pp.96-117 in Fracture Mechanics Methods for Ceramics, Rocks, and Concrete, ASTM STP 745. Edited by S.W. Freiman and E.R. Fuller,Jr ASTM, 1981.

46. B. Mussler, M.V. Swain, and N. Claussen, "Dependence of Fracture Toughness of Alumina on Grain Size and Test Technique," J. Am. Ceram. Soc., 65 [11] 566-72 (1982).

47. Z.P. Bažant, "Snapback Instability at Crack Ligament Tearing and Its Implication for Fracture Micromechanics," *Cem. Concr. Res.*, 17 [6] 951-67 (1987).

48. J. Nakayama, " Direct Measurement of Fracture Energies of Brittle Heterogeneous Materials," J. Am. Ceram. Soc., 48 [11] 583-87 (1965).

49. J. Planas, P. Maturana, G, Guinea, and M. Elices, " Fracture Energy of Water Saturated and Partially Dry Concrete at Room and at Cryogenic Temperatures"; pp. 1809-17 in Advances in Fracture Research (ICF7, Houston, 1989), Vol. 2. Edited by K. Salama et al. Pergamon Press, Oxford, 1989.

Basic Notations

a, a_0	Effective and initial crack lengths
b	Width of specimen
c	Effective process zone length, $c = a - a_0$
c_f	Limit of c as $d \to \infty$, a material parameter
d	Characteristic size (dimension) of specimen
$\bar{d}$	Intrinsic size of structure, Eq. (6)
E'	Modulus of elasticity
f_t	Local tensile strength in softening models, a material parameter
f_u	An arbitrary measure of strength
G_c	Apparent fracture energy
G_f	limit of G_c as size approaches to infinity, a material parameter
K_{Ic}	Apparent fracture toughness
K_{If}	Limit of K_{Ic} as $d \to \infty$, a material parameter
L	Span of the beam
ℓ	Ligament size
ℓ_0	$= E'G_f/f_t^2$, a material length parameter
ℓ_f	Length of process zone
P_u	Maximum load
$R(c)$	Resistance to crack growth (critical energy release rate)
u	Load-point displacement
α, α_0	Relative effective and initial crack lengths ($= a/d$, a_0/d)
β	Brittleness number
σ_N	Nominal stress at maximum load, Eq. (1)
τ_N	Intrinsic nominal stress at maximum load, Eq. (6)

Table I. Test Data of McKinney and Rice (1981)

Material	d (mm)	b (mm)	L (mm)	a_0 (mm)	P_u (N)
SiO_2	31.7	25.2	127.0	7.2	489.3
	31.7	25.4	127.0	6.3	627.2
	25.6	25.5	101.6	4.0	609.4
	25.3	25.5	101.6	4.8	582.7
	14.9	12.0	59.4	3.5	193.5
	15.0	12.0	59.4	3.6	175.3
	15.0	12.4	59.4	3.5	196.2
	15.0	12.2	59.4	3.3	195.7
	11.6	11.7	46.7	1.9	200.2
	11.6	11.7	46.7	1.9	173.9
	11.7	11.7	46.7	2.2	177.9
	11.7	11.7	46.7	2.0	207.3
	4.8	4.8	19.3	1.3	31.6
	4.8	4.8	19.3	0.9	44.0
	4.8	4.9	19.3	1.4	28.5
	4.8	4.8	19.3	1.1	34.7
	4.8	4.8	19.3	1.4	36.0
	4.8	4.8	19.3	1.2	34.2
	4.8	4.8	19.3	1.8	26.2
	4.8	4.8	19.3	1.1	27.1
SiC CN-137	36.8	36.6	149.4	4.8	6316.2
	37.3	37.6	149.4	4.8	9429.8
	17.1	17.1	69.1	2.3	1610.2
	17.1	17.2	69.1	2.5	2090.6
	17.3	17.3	69.1	2.3	1934.9
	7.2	7.2	28.7	1.3	282.4
	7.2	7.2	28.7	1.5	342.5
	7.2	7.2	28.7	1.6	275.8
	7.2	7.2	28.7	1.5	291.3
SiC CN-167	36.6	37.3	149.4	4.8	3602.9
	37.3	37.1	149.4	4.8	3958.7
	17.4	17.5	67.6	2.4	1178.7
	17.0	16.5	67.6	2.4	1014.1
	16.3	17.5	67.6	3.4	902.9
	7.2	7.2	28.7	1.6	202.4
	6.7	6.7	28.7	1.4	160.1
	7.2	7.2	28.7	1.6	224.6
	7.2	7.2	28.7	1.6	197.9

Table II. Results for Data of McKinney and Rice (1981)

Material	Objective Function	K_{If} (MPa$\sqrt{m}$)	c_f (mm)	ω_τ %	ω_K %	ω_{KIf} %	ω_{cf} %
SiO_2	τ_N , Eq. 5	0.71	0.61	12.1	10.6	-	-
	K_{Ic}, Eq. 15	0.68	0.45	12.2	10.4	-	-
	$1/\tau_N^2$, Eq. 17	0.63	0.12	13.8	11.5	4.1	200
	$1/K_{Ic}^2$, Eq. 18	0.70	0.63	12.3	10.7	8.9	42.0
SiC CN-137	τ_N	9.64	21.33	13.9	16.1	-	-
	K_{Ic}	8.48	15.81	14.0	16.1	-	-
	$1/\tau_N^2$	7.61	13.04	14.3	16.5	49.6	102
	$1/K_{Ic}^2$	11.71	34.71	14.2	16.3	178	360
SiC iCN-163	τ_N	2.33	1.36	5.3	4.7	-	-
	K_{Ic}	2.29	1.26	5.3	4.7	-	-
	$1/\tau_N^2$	2.26	1.18	5.4	4.7	3.7	21.1
	$1/K_{Ic}^2$	2.33	1.39	5.3	4.7	7.0	23.0

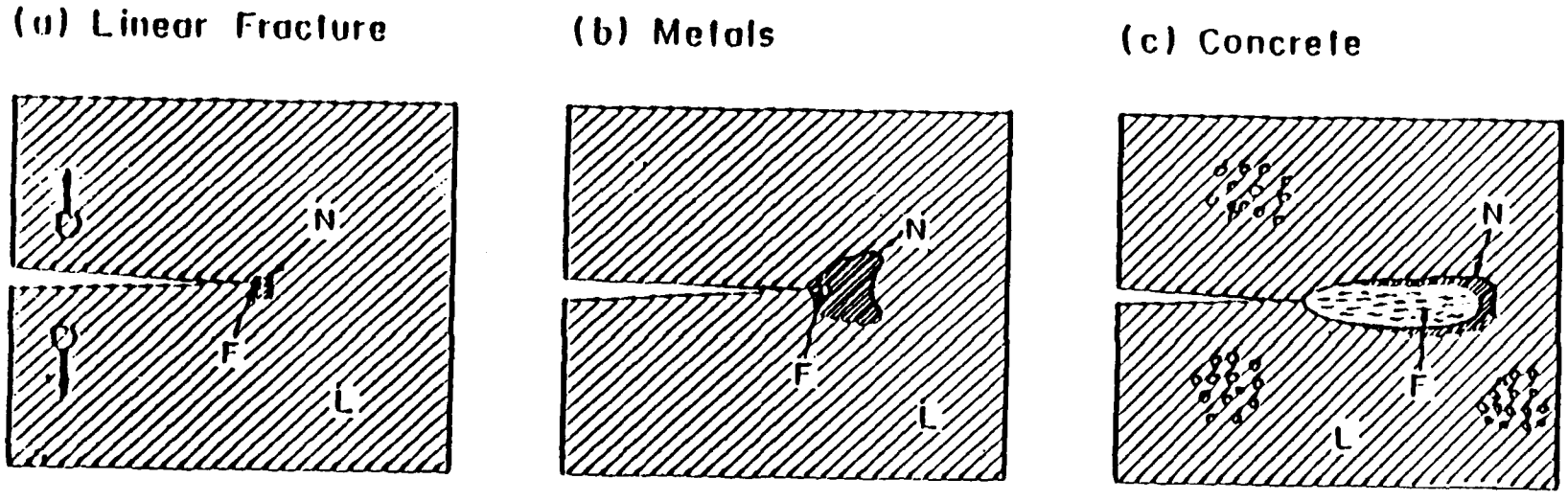


FIG. 1

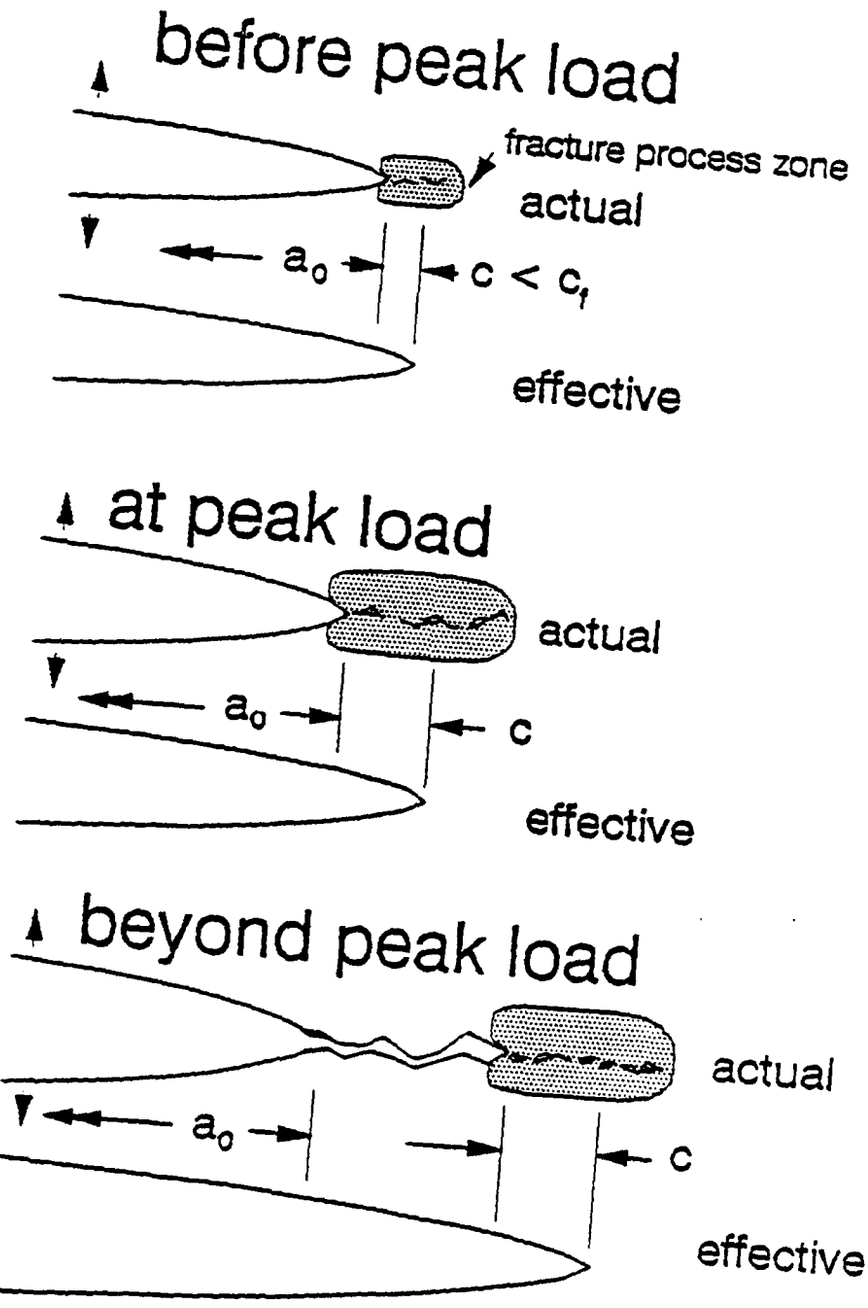

before peak load
fracture process zone
actual
a_0
$c < c_f$
effective
at peak load
actual
a_0
c
effective
beyond peak load
actual
a_0
c
effective

Fig.2

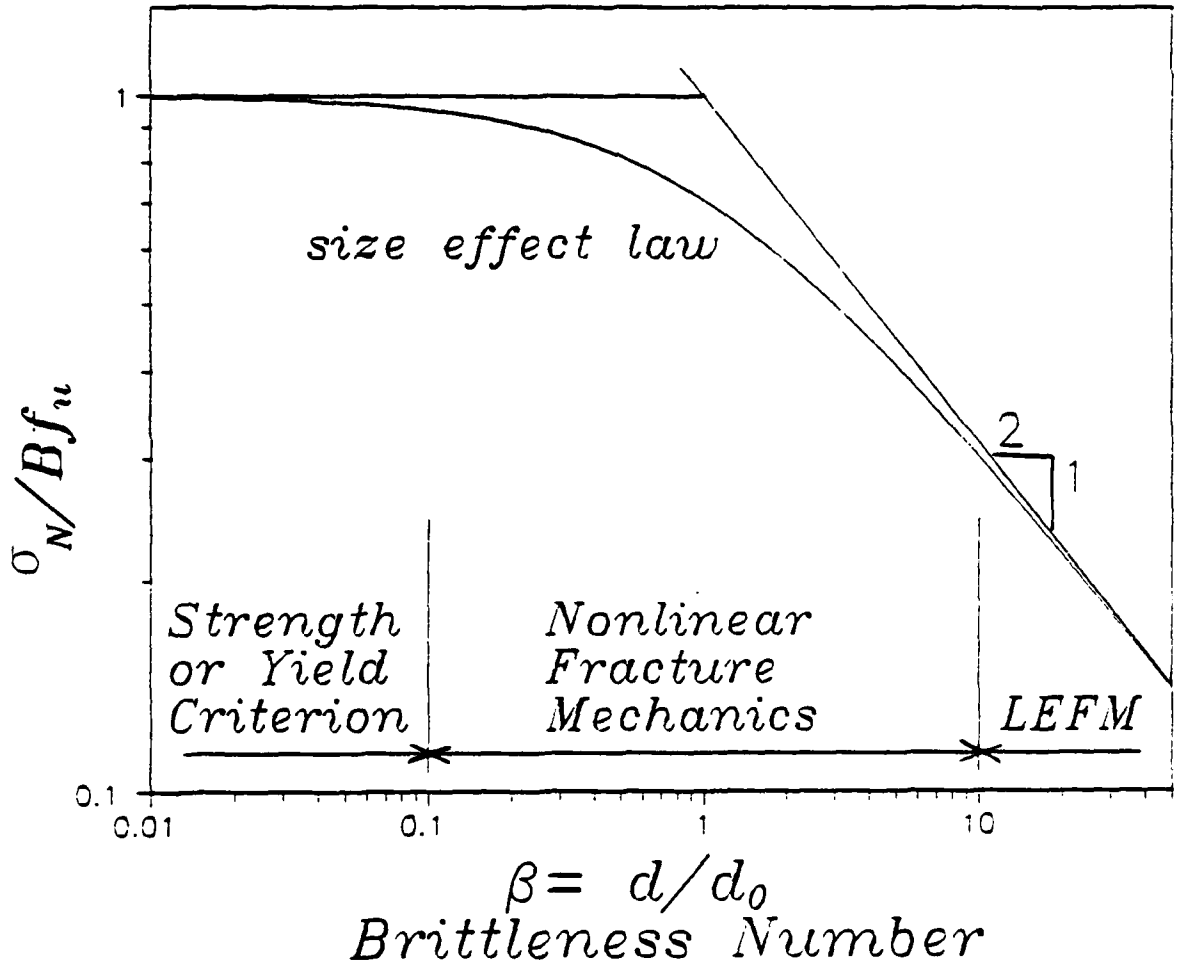

1
size effect law
2
1
σ_N/Bf_u
Strength or Yield Criterion
Nonlinear Fracture Mechanics
LEFM
0.1
0.01
0.1
1
10
$\beta = d/d_0$
Brittleness Number

Fig. 3

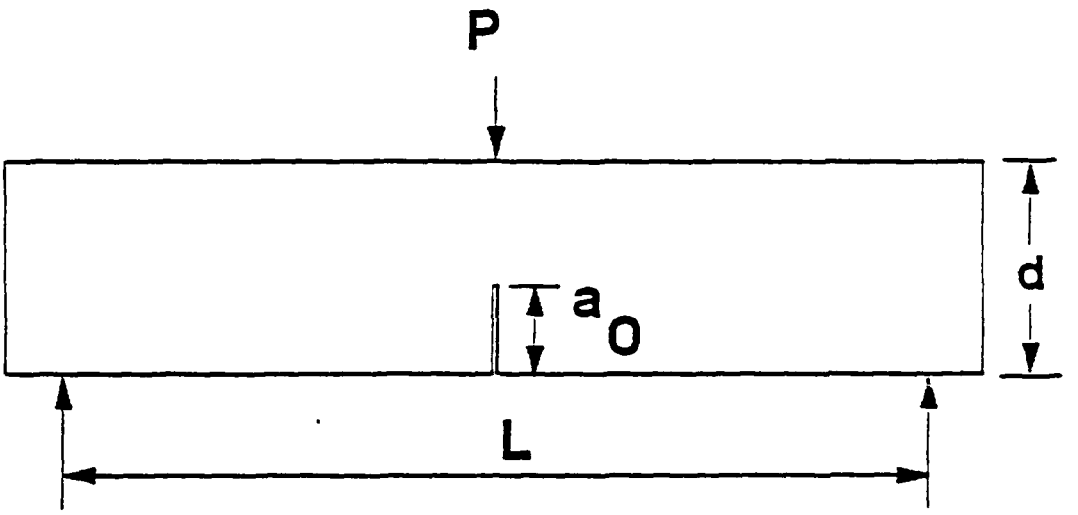


Fig. 4

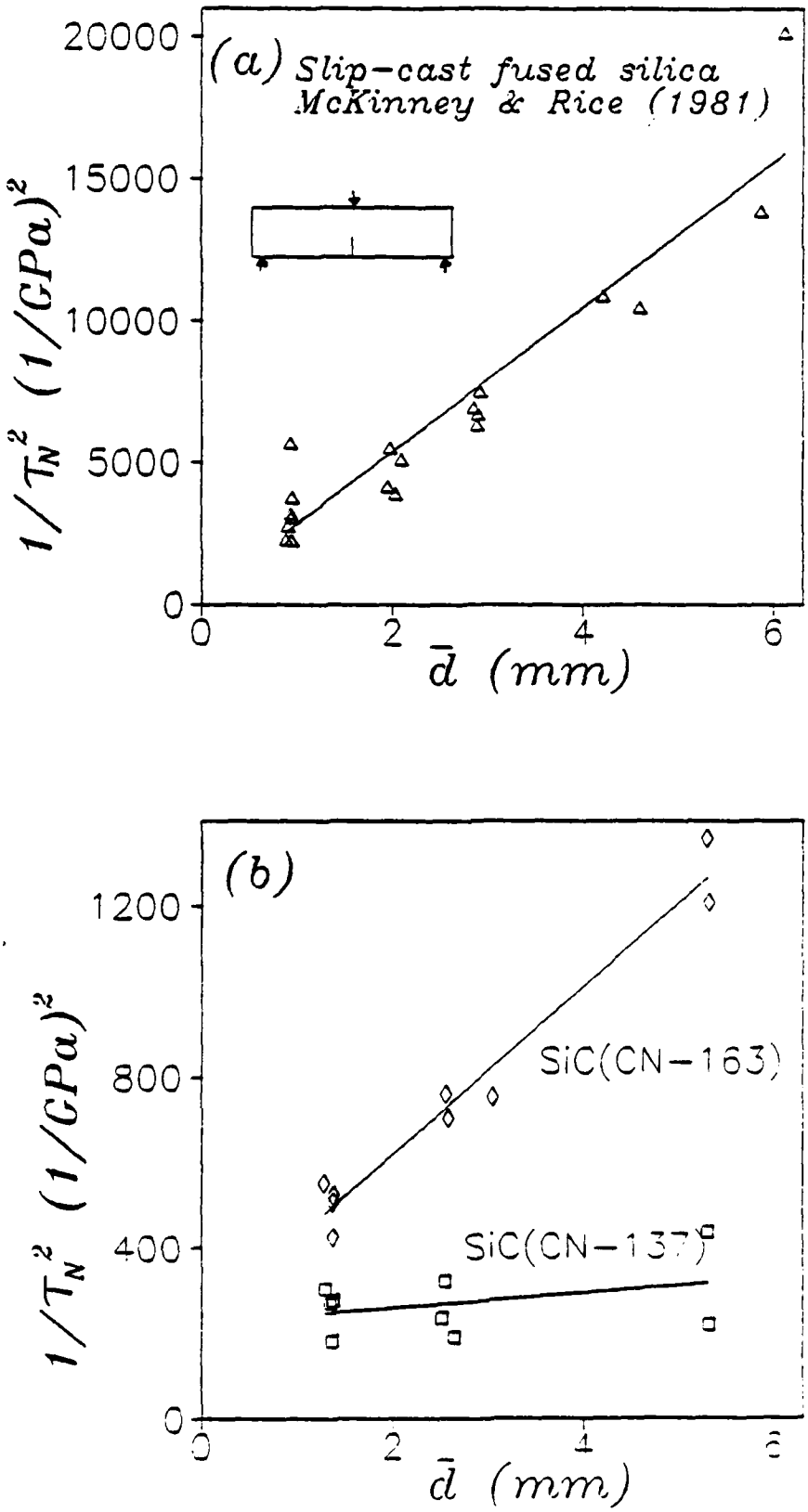
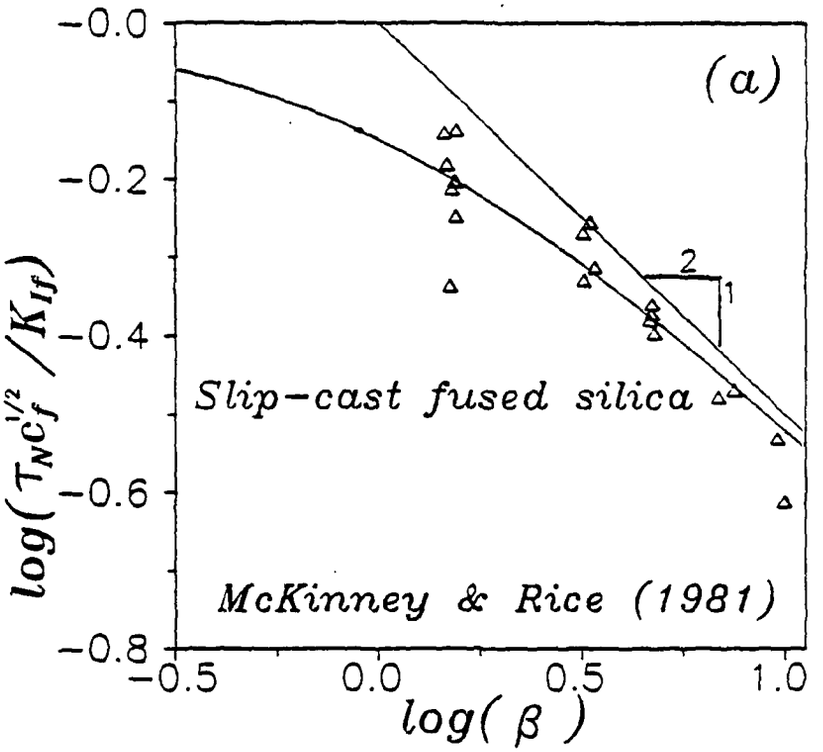

(a) Slip-cast fused silica
McKinney & Rice (1981)
20000
15000
10000
5000
0
0
2
4
6
$1/T_N^2$ (1/GPa)^2
$\bar{d}$ (mm)
(b)
1200
800
400
0
SiC(CN-163)
SiC(CN-137)
$1/T_N^2$ (1/GPa)^2
$\bar{d}$ (mm)

Fig. 5


(a)
Slip-cast fused silica
McKinney & Rice (1981)
2
1
−0.0
−0.2
−0.4
−0.6
−0.8
−0.5
0.0
0.5
1.0
log($\tau_N c_f^{1/2} / K_{If}$)
log(β)

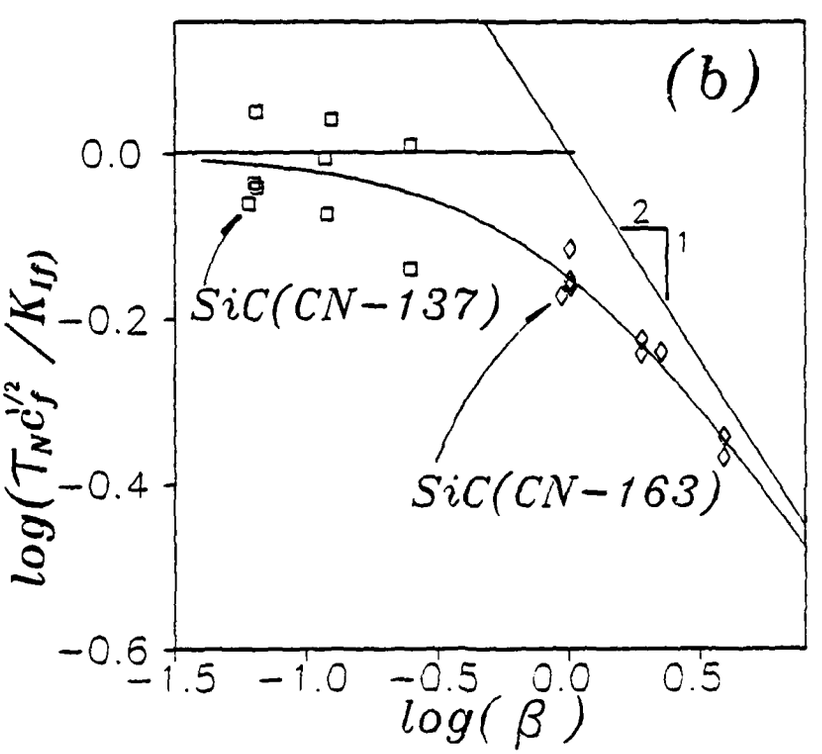

(b)
SiC(CN−137)
SiC(CN−163)
2
1
0.0
−0.2
−0.4
−0.6
−1.5
−1.0
−0.5
0.0
0.5
log($\tau_N c_f^{1/2} / K_{If}$)
log(β)

Fig. 6

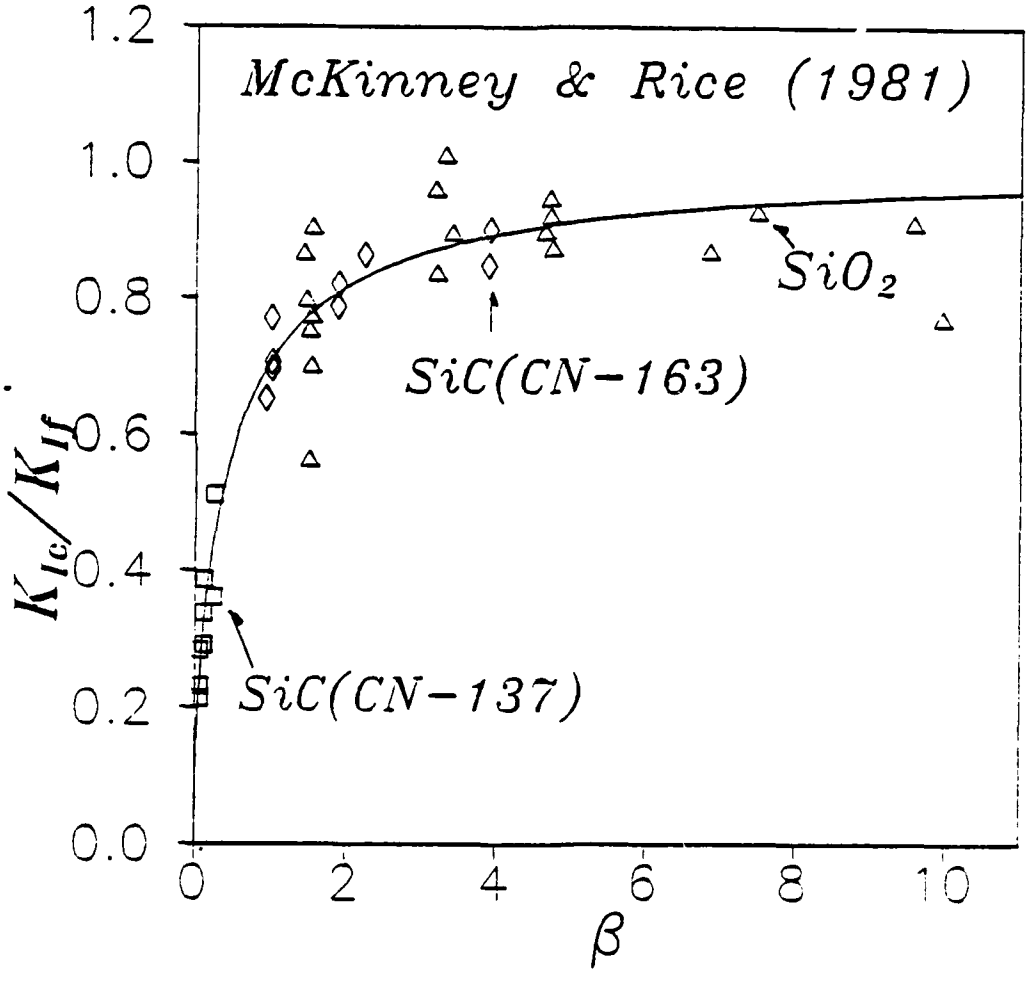

McKinney & Rice (1981)
SiO2
SiC(CN-163)
SiC(CN-137)
KIc/KIf
β
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10

Fig. 7

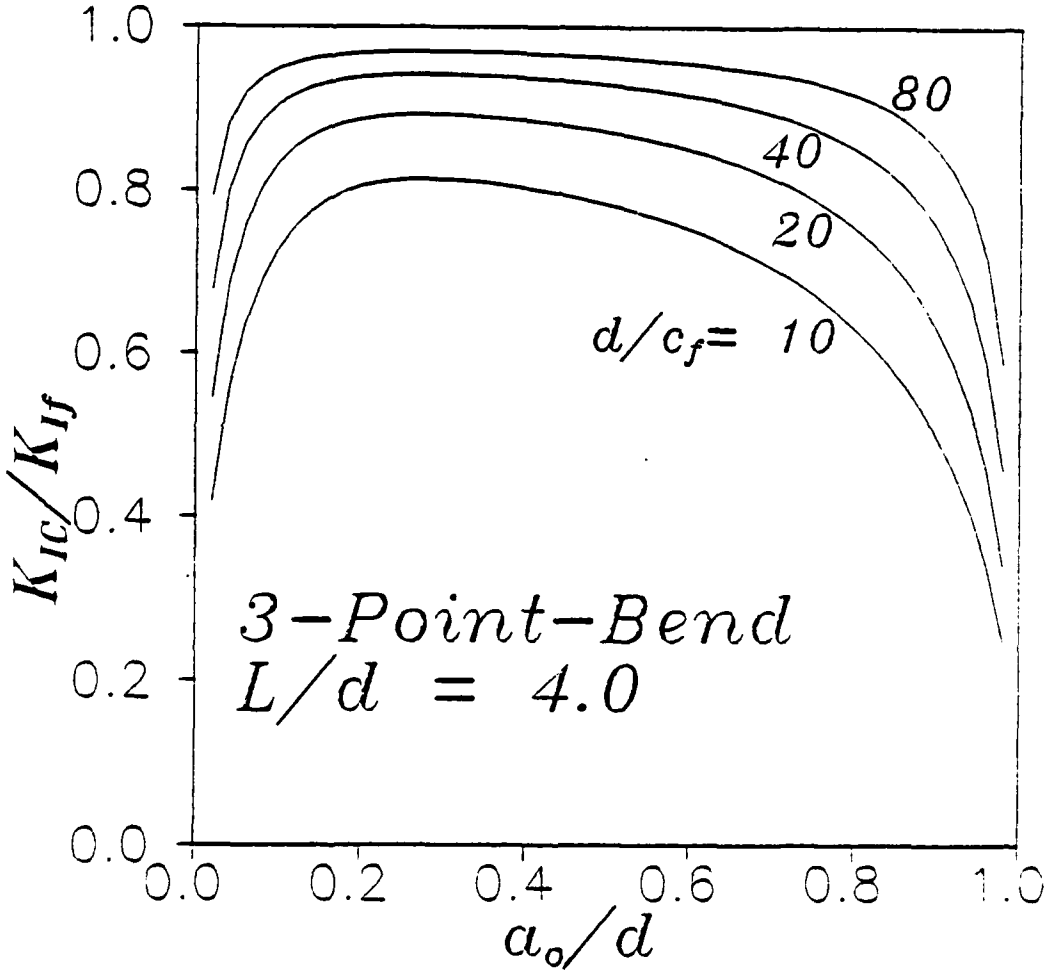

1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
K_{IC}/K_{If}
a_o/d
80
40
20
d/c_f= 10
3-Point-Bend
L/d = 4.0

Fig. 3

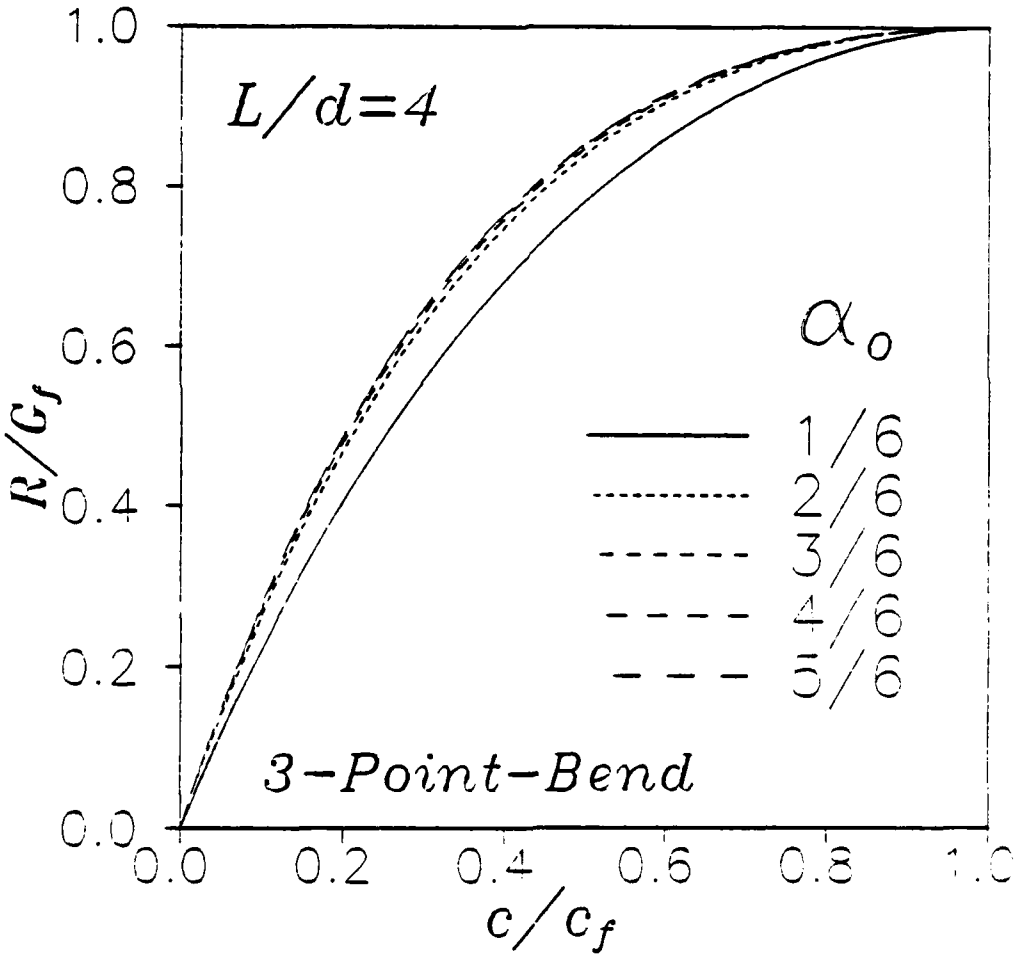

1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
L/d=4
α_o
1/6
2/6
3/6
4/6
5/6
3-Point-Bend
R/G_f
c/c_f

Fig. 9

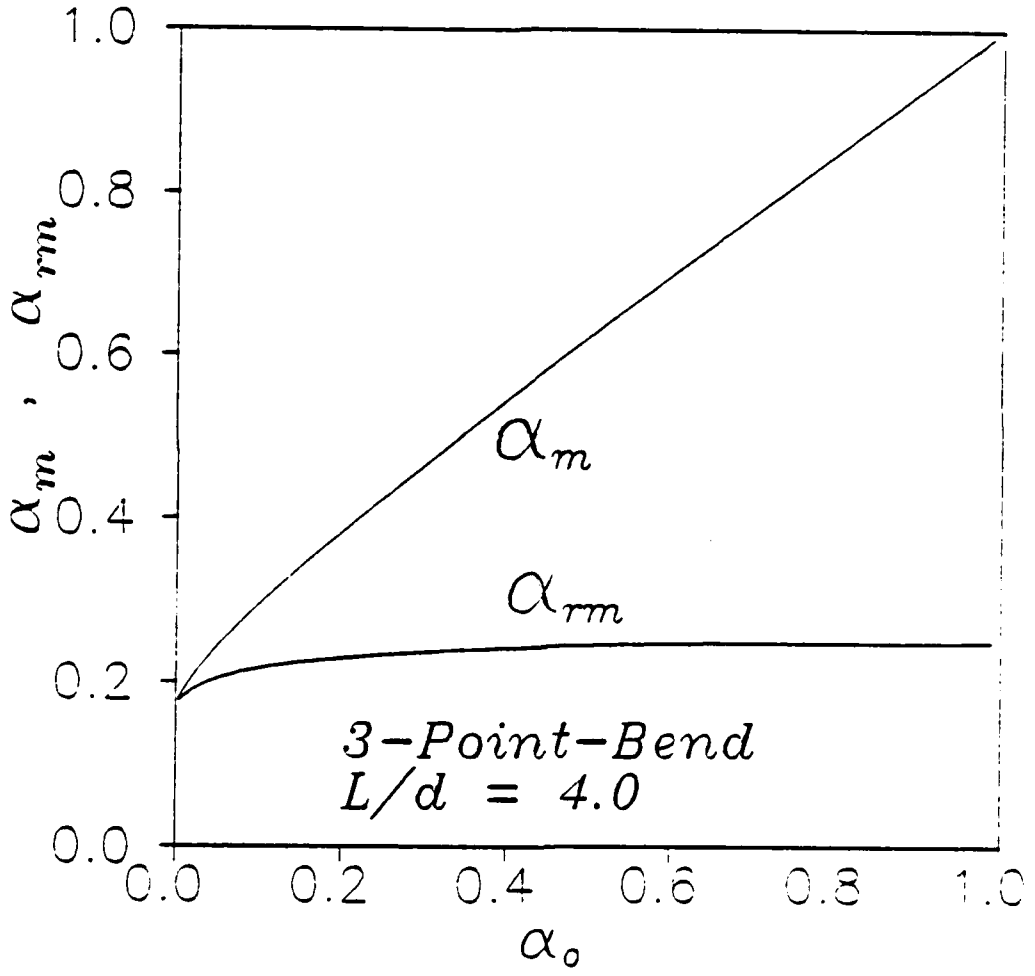
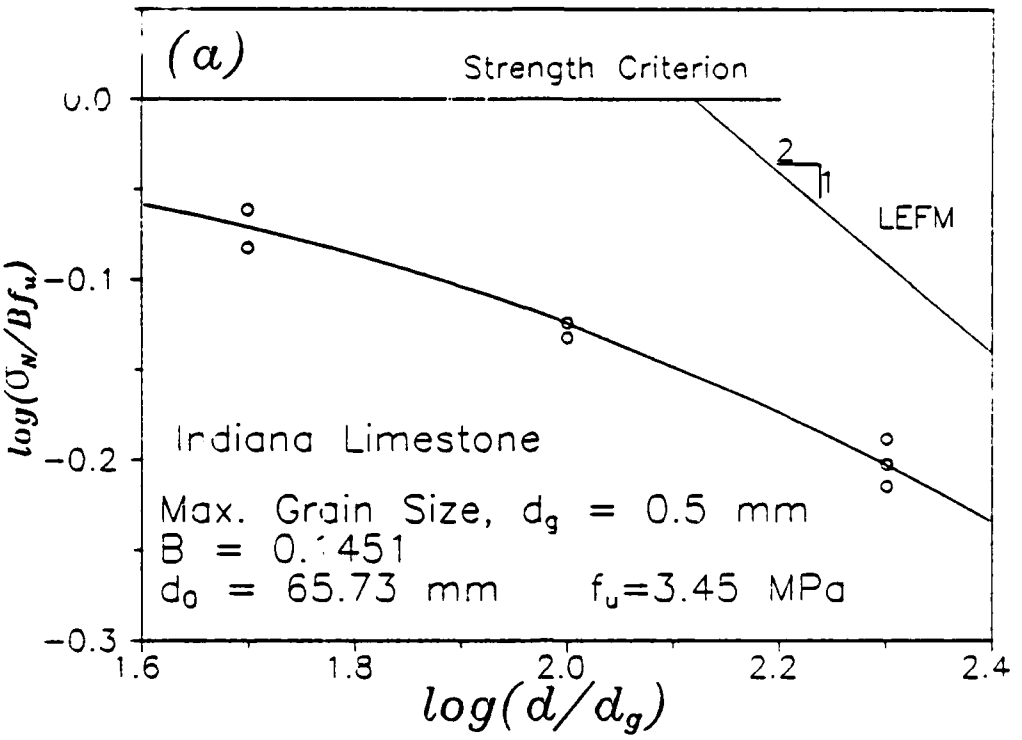

1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
α_m , α_{rm}
α_o
α_m
α_{rm}
3-Point-Bend
L/d = 4.0

Fig. 10


(a)
Strength Criterion
0.0
−0.1
−0.2
−0.3
2
1
LEFM
log(σN/Bfu)
Indiana Limestone
Max. Grain Size, dg = 0.5 mm
B = 0.1451
d0 = 65.73 mm
fu=3.45 MPa
1.6
1.8
2.0
2.2
2.4
log(d/dg)

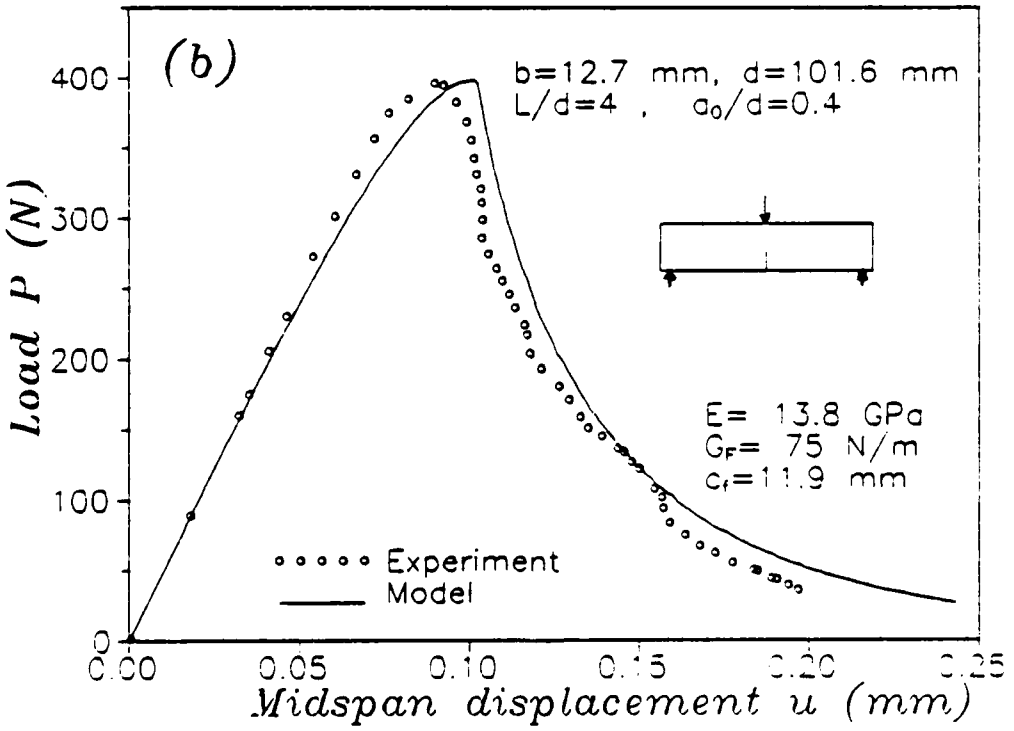

(b)
b=12.7 mm, d=101.6 mm
L/d=4 , a0/d=0.4
400
300
200
100
0
Load P (N)
E= 13.8 GPa
GF= 75 N/m
cf=11.9 mm
Experiment
Model
0.00
0.05
0.10
0.15
0.20
0.25
Midspan displacement u (mm)

Fig. 11

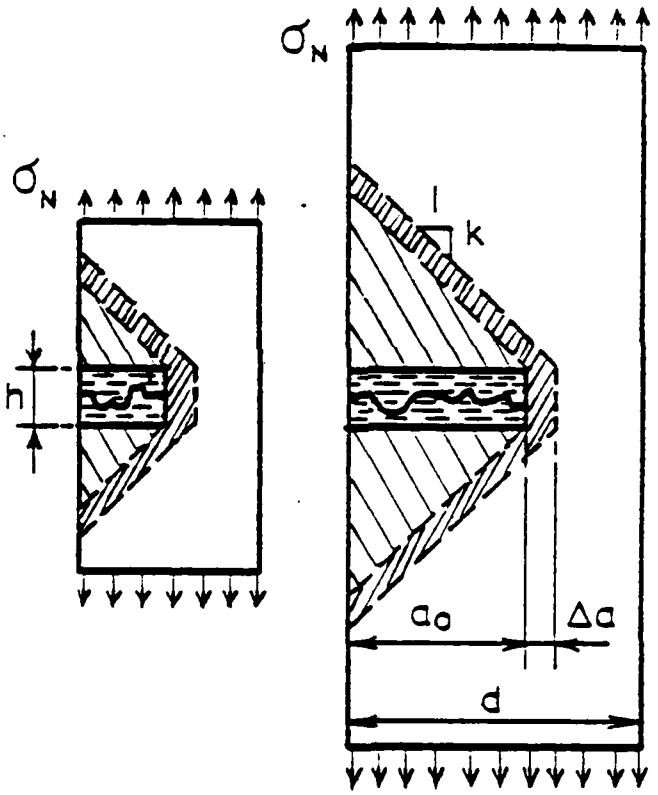


Fig. 12

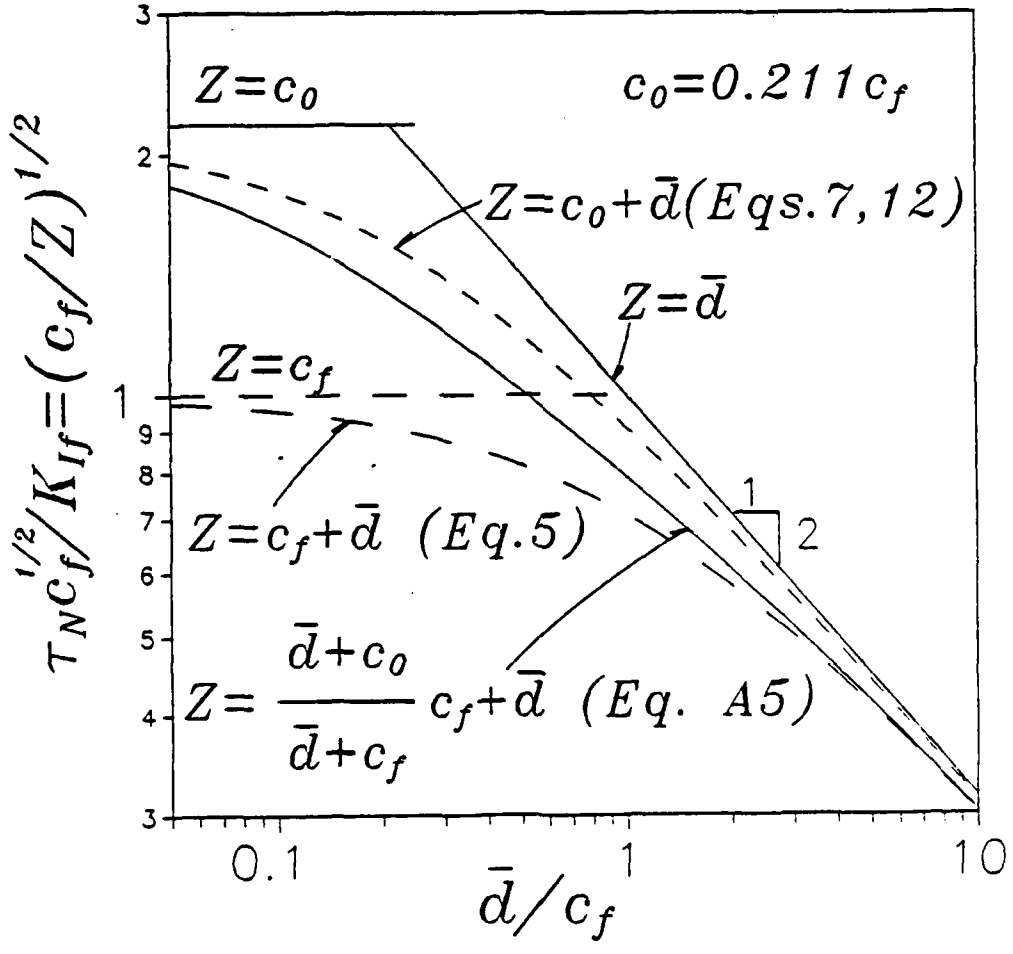

$Z=c_0$
$c_0=0.211c_f$
$Z=c_0+\bar{d}$ (Eqs. 7, 12)
$Z=\bar{d}$
$Z=c_f$
$Z=c_f+\bar{d}$ (Eq. 5)
$Z=\frac{\bar{d}+c_0}{\bar{d}+c_f}c_f+\bar{d}$ (Eq. A5)
1
2
$\tau_N c_f^{1/2}/K_{If}=(c_f/Z)^{1/2}$
$\bar{d}/c_f$
0.1
1
10
3
2
1
9
8
7
6
5
4
3

Fig. 13

ACI MATERIALS JOURNAL TECHNICAL PAPER

Title no. 85-M32

Effect of Temperature and Humidity on Fracture Energy of Concrete



by Zdeněk P. Bažant and Pere C. Prat

Fracture experiments were conducted at temperatures from 20 to 200 C (68 to 392 F) to determine the dependence of the Mode I fracture energy of concrete on temperature as well as the specific water content. The fracture energy values were determined by testing geometrically similar specimens of sizes in the ratio 1:2:4:8 and then applying Bažant's size effect law. Three-point bend specimens and eccentric compression specimens are found to yield approximately the same fracture energies, regardless of temperature. To describe the temperature dependence of fracture energy, a recently derived simple formula based on the activation energy theory (rate process theory) is used and verified by test results. The temperature effect is determined both for concrete predried in an oven and for wet (saturated) concrete. By interpolation, an approximate formula for the effect of moisture content on fracture energy is also obtained. This effect is found to be small at room temperature but large at temperatures close to 100 C (212 F).

Keywords: concretes; cracking (fracturing); energy; humidity; moisture content; temperature.

While heating in metals causes abrupt increases of fracture toughness, due to brittle-ductile transitions which change the fracture mechanism, the fracture toughness of purely brittle materials such as glass, ceramics, graphite, and rocks is known to smoothly decrease with increasing temperature[1-3]. For concrete, the determination of the influence of temperature on the fracture energy has been hampered by the fact that until recently an unambiguous definition and method of measurement of the fracture energy itself was unavailable. The basic classical measurement methods, such as the work of fracture method proposed for concrete by Hillerborg et al.[4,5] and adopted by The International Union of Testing and Research Laboratories for Materials and Structures (RILEM) or the methods based on compliance or crack opening displacement measurements, have not been found to yield the same fracture energy values when specimens of significantly different sizes or shapes were used.

Recently, however, a fracture energy definition that leads to unique values has been proposed.[6-10] The fracture energy is the specific energy required for fracture growth in an infinitely large specimen. This definition obviously requires knowledge of the size effect law which has to be used to extrapolate the test results to infinite size. Although the exact size effect law is unknown, the approximate Bažant's law[11] has been found to be sufficient in view of the inevitable statistical scatter of concrete.[6,7,10,12-14]

Measurements of concrete fracture energy based on extrapolation to infinite size were made by Bažant, Kim, and Pfeiffer.[9,10] Bažant and Pfeiffer[15] experimentally demonstrated that specimens of very different types, including the three-point bend specimens, edge-notched tensile specimens, and eccentric compression specimens, yield approximately the same fracture energy values, as far as the typical scatter exhibited by concrete permits one to discern. As for the effect of specimens size on fracture energy, none can be present due to the very nature of this fracture energy definition.

The availability of a measurement method that yields unique values of fracture energy makes it meaningful to study various factors that influence fracture energy. The objective of this study is to examine the effect of temperature as well as the related effect of humidity. The values of fracture energy need to be known to predict brittle failures of concrete structures on the basis of fracture mechanics, which promises to give more accurate results than the methods of limit analysis.

EXPERIMENTAL INVESTIGATION

The test specimens were beams of constant rectangular cross section loaded at three points [Fig. 1(a)]. To determine the size effect, geometrically similar specimens of various depth d were used. The depths were d

Received Feb. 4, 1987, and reviewed under Institute publication policies.
Copyright © 1988, American Concrete Institute. All rights reserved, including the making of copies unless permission is obtained from the copyright proprietors. Pertinent discussion will be published in the May-June 1989 *ACI Materials Journal* if received by Feb. 1, 1989.

Zdeněk P. Bažant, FACI, has been a professor at Northwestern University since 1973 and has also served there as director of the Center for Concrete and Geomaterials. He currently is Professor and Director, the Technological Institute, Department of Civil Engineering, Northwestern University. He is a registered structural engineer and a consultant to Argonne National Laboratory and several other firms and is on editorial boards of six journals. He is Chairman of ACI Commmittee 446, Fracture Mechanics; and is a member of ACI Committees 209, Creep and Shrinkage in Concrete; and 348, Structural Safety. Currently he is Chairman of RILEM Committee TC69 on creep and of IA-SMiRT Division H. In 1987, Professor Bažant visited the University of Tokyo as Kajima Foundation Fellow and during 1988 he is a NATO Senior Guest Scientist at E.N.S., Paris-Cachan.

ACI member Pere C. Prat is a research associate at the Materials Science Institute (CSIC), Barcelona, Spain. He obtained his civil engineering degree from the School of Civil Engineering of the Technical University of Catalonia, Barcelona, Spain, and his PhD from Northwestern University. His research interests include constitutive models for geomaterials, strain-softening behavior, fracture mechanics, nonlinear analysis of concrete structures, experimental methods, and finite element applications.

= 1.5, 3, 6, and 12 in. (38.1, 76.2, 152, and 305 mm), while the thickness was b = 1.5 in. (38.1 mm) for all the specimens. The length-to-depth ratio was L/d = 8:3 and the span-to-depth ratio was ℓ/d = 2.5, for all specimens. One notch of depth $d/6$ and thickness 0.1 in. (2.5 mm) (same for all specimens) was cut with a diamond saw three weeks after the specimens were stripped from the molds.

To check for a possible effect of specimen type, additional specimens of the same shape were tested in eccentric compression. These specimens, which had two symmetrical notches of depth $d/6$ and thickness 0.1 in. (2.5 mm) [Fig. 1(b)], included only specimen sizes d = 1.5, 3, and 6 in. (d = 12 in. would not fit the available oven).

From each batch of concrete, one specimen of each size as well as three control cylinders for compression strength, 3 in. (152 mm) in diameter and 6 in. (305 mm) in length, were cast. The means and standard deviations of the compression strengths after 28 days of moist curing are given in Table 1.

All specimens were cast with the side of depth d in the vertical position, using a water-cement ratio of 0.6 and a cement-sand-gravel ratio of 1:2:2 (all by weight). The maximum gravel size was d_a = 0.5 in. (12.7 mm) and the maximum sand grain size was 0.19 in. (4.8 mm). The aggregate consisted of crushed limestone and siliceous Illinois beach sand (Lake Michigan). Portland cement C 150 (ASTM Type I), with no admixtures and no air-entraining agents, was used.

The fracture specimens were removed from the plywood forms after one day and subsequently cured until about 2 hr before the test in a moist room with 95 percent humidity and a temperature of 25 C (77 F). Three identical specimens were cast simultaneously from each successive batch for each type of test. The age of the specimens at the time of the test was 28 days, except for the specimens tested in the hot water bath. For these, the tests had to be postponed to the age of 41 days due to certain scheduling problems.

The fracture tests were carried out at two types of humidity conditions — dry and wet. The dry tests included both the three-point loaded specimens and eccentric compression specimens. The temperatures in these tests were 20 C (68 F), 65 C (149 F), 120 C (248 F), and 200 C (392 F). Each of the tests started with preheating under no load in an oven. The temperature in the oven was raised gradually over a period of about 1 hr to the temperature of 120 C, which was then kept for 3 hr. The specime were not sealed and may be expected to have lost all their evaporable water. After this

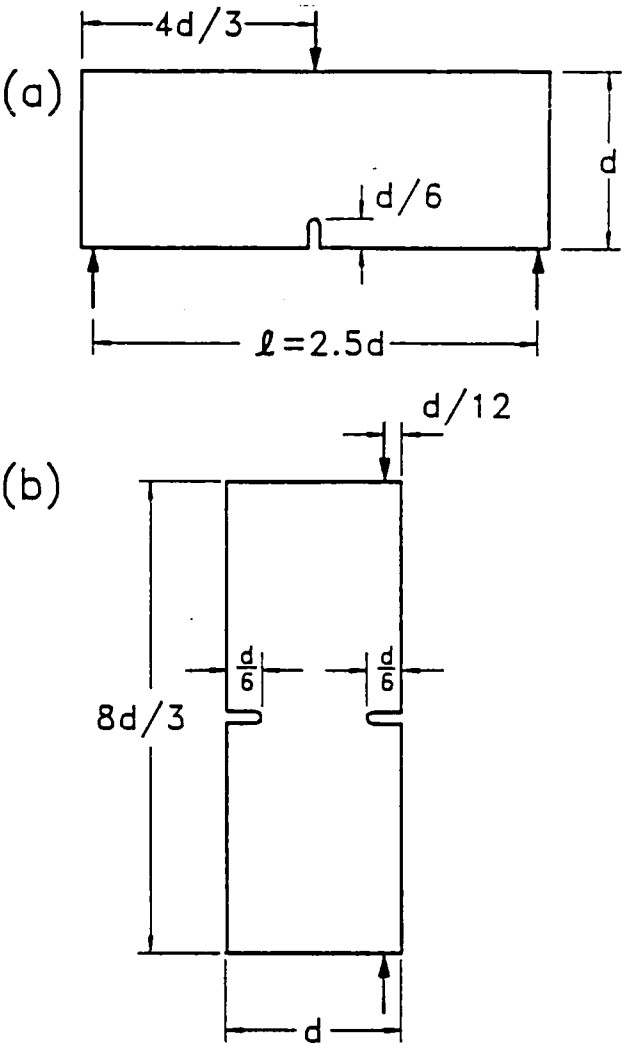


Fig. 1—Test specimens used: (a) three-point bend specimens; and (b) eccentric compression specimens

Table 1 — Compression strength of concrete (on 3 × 6-in. cylinders) in psi*

Test series	Mean f'_c	Standard deviation s
3-point bend, dry, 20 C	5550	36
3-point bend, dry, 65 C	5585	139
3-point bend, dry, 120 C	4975	90
3-point bend, dry, 200 C	5000	95
Eccentric compression, dry, 20 C	5410	78
Eccentric compression, dry, 65 C	5250	59
Eccentric compression, dry, 120 C	5075	117
Eccentric compression, dry, 200 C	5075	90
3-point bend, wet, 65 C†	5770	55
3-point bend, wet, 90 C‡	5725	84

* 1 psi = 6895 Pa.
† Corrected for age; measured was f'_c = 6035 psi, s = 58 psi at 41 days.
‡ Corrected for age; measured was f'_c = 5990 psi, s = 88 psi at 41 days.

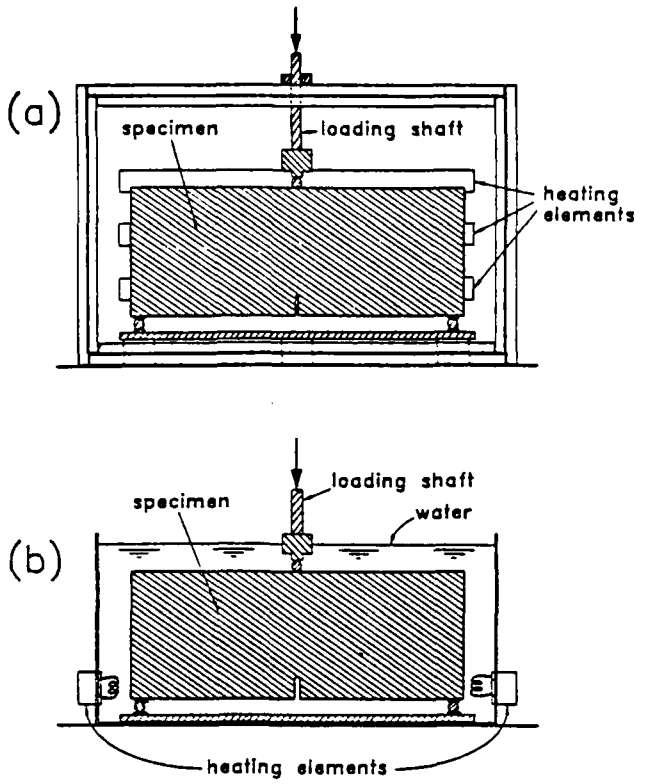


Fig. 2—Arrangement of tests: (a) in an oven; and (b) in a water bath

3 hr period, the temperature was gradually raised or lowered at the average rate of about 50 C/hr (90 F/hr), until the desired temperature of the fracture test was reached. The specimen was then loaded to failure while in the oven [Fig. 2(a)]. The loading shaft of the machine protruded through the wall of the oven, and the thermal insulation around the shaft caused no appreciable friction.

According to the linear heat conduction equation and the known thermal diffusivity of concrete, it was calculated that, after a sudden exposure to a different constant temperature, not more than 1 hr is needed for the temperature to become uniform within the specimen of thickness 1.5 in. (38.1 mm). Therefore, the heating times used must have been sufficient. The heating time in the oven must have also sufficed, since a uniformly dry state throughout the specimen thickness was obtained. This is not only well known empirically but is also justified by the measurements and theory in Reference 16, in which it was found that the diffusivity of moisture in concrete increases about 200 times when the temperature exceeds 100 C (212 F).

Although concrete drying by heating in an oven has been a standard practice in the investigations of creep and other properties, this procedure might, nevertheless, cause some damage such as microcracking and, thus, possibly affect the fracture test results. However, the fact that after such drying, the fracture energy of concrete becomes much higher than for a wet specimen at the same elevated temperature suggests that possible damage due to heating is probably not serious. The alternative to drying the specimens without heating is hardly feasible. Simplified calculations showed that the previously mentioned heating rates were slow enough to avoid development of high thermal stresses.

The wet tests included only three-point bend specimens. Each test started by placing the specimens into a water-filled tank [Fig. 2(b)] with the water at room temperature. The temperature of the water in the tank was raised gradually, at the rate of about 50 C/hr (90 F/hr), to the desired test temperature and was then kept constant for 1 hr before the fracture test in water was started.

The temperatures of the wet tests were 65 C (149 F) and 90 C (194 F). No wet tests were conducted at room temperature, since the response of wet specimens at room temperature may be assumed to be the same as that of previously tested unsealed and undried specimens, loaded right after curing. This assumption is justified by the fact that at room temperature the moisture diffusion is so slow that the environmental humidity cannot affect the state of concrete in the bulk of the specimen during the short time of exposure. At 120 C (248 F), by contrast, the diffusivity of moisture in concrete is about 200 times higher than it is at 90 C (194 F), and about 3000 times higher than at 20 C (68 F).[16] Therefore, the environmental relative humidity quickly spreads through the entire specimen.

No wet tests could be carried out at temperatures over 100 C (212 F), since the specimen could not be kept wet unless a pressure chamber was used. However, this would represent a different type of test.

All the tests were carried out in a closed-loop MTS machine. Only the maximum load values were needed to calculate the fracture energies. The loading rate was such that the time to failure was 3 to 5 min for all specimens.

All the dry tests were performed under stroke-control at a constant displacement rate. For certain extraneous reasons, the wet specimens had to be tested under load control; therefore, they failed right at the maximum load. Impossible though load-controlled testing is for other types of fracture tests (e. g., RILEM's work of fracture test), it is nevertheless acceptable for the present method since the post-peak response is not needed to determine the fracture energy. Theoretically, if the specimens exhibit high statistical heterogeneity, load-controlled tests can yield a lower maximum load than the displacement-controlled tests, but, according to the analysis of Bažant and Panula,[17] the differences that may reasonably be expected in the maximum load value are insignificant (under 3 percent). On the other hand, if the objective were the post-peak descending load-deflection curve, then load-controlled tests would of course be impossible.

The measured maximum load values P are given in Table 2 and the 28-day compression strengths f'_c for individual batches of concrete in Table 1. The 28-day f'_c values for the wet tests were not directly measured but were calculated from the compression strength f'_c (41) measured at the time of these tests, at which the age of concrete was 41 days. The adjustment for age was

based on the approximate formula[18]

$$f_c'(t) = f_c'(28)[1 + 0.277\log(t/28)]$$
$$\text{from which } f_c'(28) \approx 0.956 f_c'(41)$$

CALCULATION OF FRACTURE ENERGY FROM SIZE EFFECT

The size effect is the most important difference of fracture mechanics from the failure theories based on plastic limit analysis. The size effect is understood as the dependence of the nominal stress at failure σ_N on a characteristic dimension of the specimen d when geometrically similar specimens or structures are considered ($\sigma_N = P/bd$, where P = maximum load and b = specimen thickness). For stress-based failure theories such as the plastic limit analysis or elastic allowable strength design, there is no size effect, i.e., σ_N is constant. For classical linear elastic fracture mechanics, $\sigma_N \sim d^{-1/2}$. Due to the influence of a relatively large microcracking zone that blunts the crack front in concrete, the size effect is intermediate between plasticity and linear fracture mechanics and represents a transition from the former to the latter as the size is increased. This transition may be approximately described by Bažant's size effect law:[11] $\sigma_N = Bf_t'(1 + d/\lambda_0 d_a)^{-1/2}$ in which B, λ_0 = empirical constants and d_a = maximum aggregate size. According to this law, the plot of $Y = (f_t'/\sigma_N)^2$ versus $X = d/d_a$ (relative size) should be a straight line $Y = AX + C$, with $C = 1/B^2$ and the slope $A = C/\lambda_0$. Thus, a linear regression of the test results in the plot of Y versus X may be used to determine A and C, from which the size effect law parameters may then be calculated as $B = C^{-1/2}$ and $\lambda_0 = C/A$.

These regression plots also yield statistics of the errors, i.e., the deviations of the measured data points from the size effect law. The coefficient of variation, defined as $\omega_Y|_X = \{[\Sigma(Y_i^{test} - Y_i)^2]/(N-2)\}^{1/2}/\bar{Y}$ where $Y_i^{test} - Y_i$ are the vertical deviations of data points Y_i^{test} from the regression line, N is the number of all the data points, and $\bar{Y} = (\Sigma Y_i)/N$ = mean of all measured Y_i, is given in Fig. 3, 4, and 5, as is the correlation coefficient r.

From the size effect law, it follows[8-10] that the fracture energy G_f, defined as the energy release rate required for crack growth in an infinitely large specimen, may be calculated from the formula

$$G_f = \frac{g(\alpha_0)}{AE_c} f_t'^2 d_a \qquad (1)$$

in which $\alpha_0 = a_0/d$; a_0 = notch length; d = beam depth; E_c = modulus of elasticity of concrete; f_t' = direct tensile strength of concrete; $A = 1/(B^2\lambda_0)$ = slope of the regression line as already defined; $g(\alpha_0)$ = nondimensional energy release rate of the specimen according to the linear elastic fracture mechanics, which can be found for the basic specimen geometries in textbooks[19,20] and handbooks,[21] and can be, in general, determined by linear finite element analysis. For the three-point specimens used, $g(\alpha_0) = 6.37$, and for the eccentric compression specimens, $g(\alpha_0) = 1.68$.

Table 2 — Maximum loads at failure

Type of test	Depth, in.	Load at failure P, lb			Average P, lb
		1	2	3	
Three-point bend Dry, 20 C	1.5	390	420	375	395
	3.0	645	690	675	670
	6.0	990	1035	1080	1035
	12.0	1710	1689	1729	1709
Three-point bend Dry, 65 C	1.5	515	375	440	443
	3.0	715	565	670	650
	6.0	1090	955	1155	1066
	12.0	1595	1430	1600	1542
Three-point bend Dry, 120 C	1.5	348	384	366	366
	3.0	530	594	486	537
	6.0	954	840	864	886
	12.0	1296	1368	1332	1332
Three-point bend Dry, 200 C	1.5	304	322	320	315
	3.0	540	480	510	510
	6.0	792	744	816	784
	12.0	1180	1190	1200	1190
Eccentric compression Dry, 20 C	1.5	675	665	690	677
	3.0	1180	1160	1260	1200
	6.0	2040	2195	1945	2060
Eccentric compression Dry, 65 C	1.5	855	715	594	721
	3.0	1280	1223	1080	1194
	6.0	2100	2010	1740	1950
Eccentric compression Dry, 120 C	1.5	740	775	680	732
	3.0	1140	1240	1215	1198
	6.0	1860	1920	1720	1833
Eccentric compression Dry, 200 C	1.5	648	684	720	684
	3.0	960	960	900	940
	6.0	1545	1705	1625	1625
Three-point bend Wet, 65 C	1.5	342	390	408	380
	3.0	545	546	586	559
	6.0	811	780	732	774
	12.0	1250	1205	1210	1222
Three-point bend Wet, 90 C	1.5	290	350	290	310
	3.0	560	520	575	552
	6.0	640	690	745	692
	12.0	1055	1060	1010	1042

1 in. = 25.4 mm, 1 lb. = 4.448 N.

The tensile strength was not measured but was estimated from the formula $f_t' \approx 6(f_c')^{1/2}$ where f_c' and f_t' are in psi, f_c' = compression strength. A possible error in the f_t'-value, however, has no effect on G_f since the A-value obtained by regression is proportional to $f_t'^2$ if the ordinate is taken as $Y = f_t'^2/\sigma_N^2$.

Fig. 6 is a photograph of a typical set of specimens of different sizes. Fig. 7 shows typical specimens installed in the heating oven, and Fig. 8 shows typical specimens installed in the water bath.

The maximum loads measured are plotted in Fig. 3 through 5. The plots on the right demonstrate that the measured σ_N agrees quite well with the size effect law, which is shown as the solid curve. This agreement justifies the use of the size effect law. The plots on the left show the linear regressions that yield the optimum values of parameters B and λ_0. The fracture energy values calculated according to Eq. (1) from the Slopes A of the regression lines given on the left of Fig. 3 through 5 are tabulated in Table 3.

Since the wet specimens had to be tested at the age of 41 days rather than 28 days, a correction for the age difference became necessary. The fracture energy values were transformed to 28 days according to Bažant

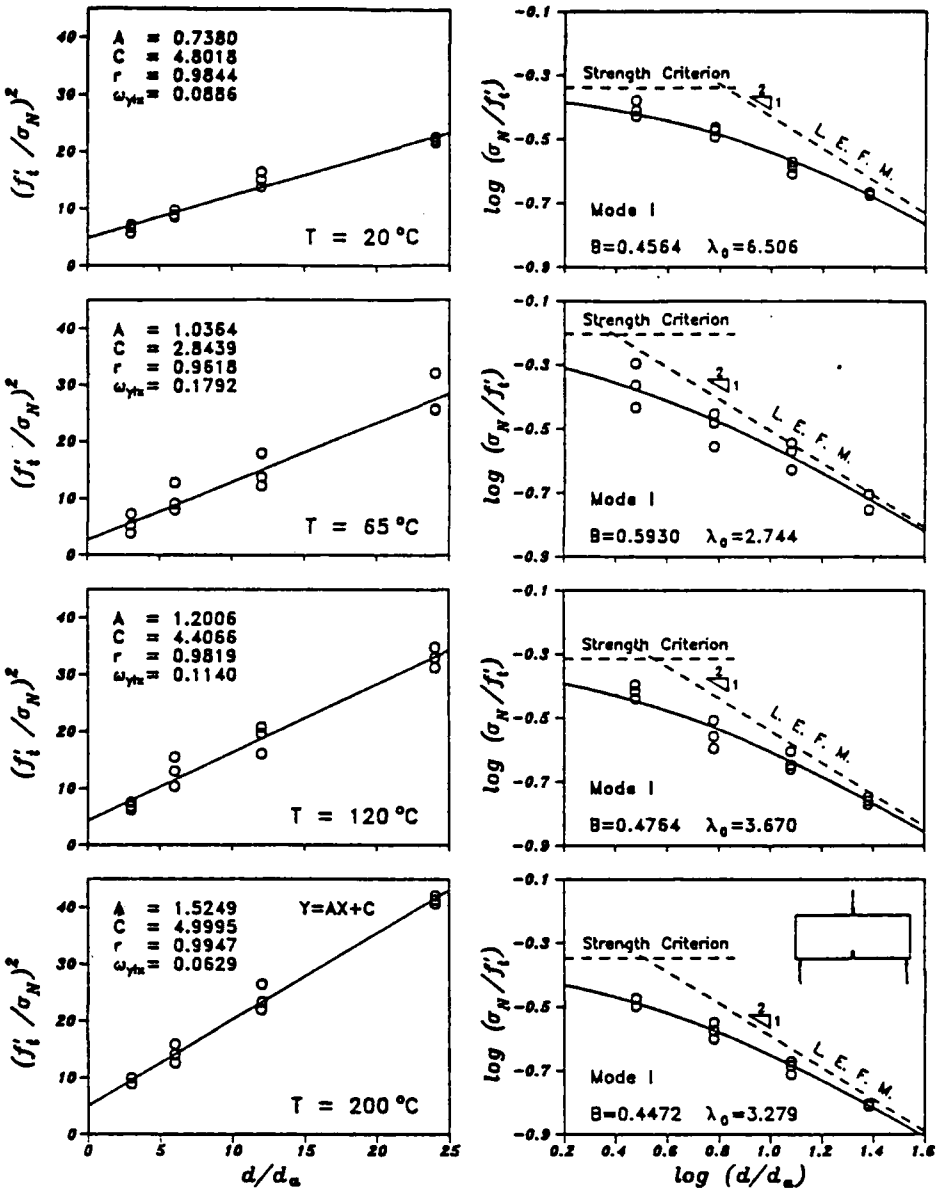


Fig. 3—One set of three-point bend specimens tested

and Oh's approximate formula[22] $G_f = (2.72 + 0.0214 f_t') f_t'^2 d_a/E$. The values in Tables 1 and 3 as well as those plotted in the figures are all transformed to 28 days.

ACTIVATION ENERGY FORMULATION

It is generally accepted that fracture is a thermally activated rate process. This means that the atomic bond ruptures that constitute the mechanism of fracture are provoked by the energies of thermal vibrations.[23] These energies are statistically distributed, as described by Maxwell distribution, and a rise in temperature increases the probability (or frequency) that the atom's energy would exceed the activation energy barrier of the bond. Therefore, a rise in temperature causes an increase in the rate of growth of fracture, which generally follows a formula of the type $\dot{a} = f(K) \exp(-U/RT)$,[3] where U = activation energy of bond rupture; R = universal gas constant; T = absolute temperature; K = stress intensity factor; and $f(K)$ = empirical monotonically increasing function. Recently, Evans[24] and Thouless et al.[25] verified for certain ceramics a special form of this formula

$$\dot{a} = v_c (K/K_c)^n e^{-U/RT} \quad (2)$$

in which v_c and n may be approximately considered as constants characterizing the given material and K_c = critical value of K (fracture toughness). Since concrete is a ceramic material, Eq. (2) may be expected to apply also. This equation is not exact but approximate only, for two reasons: 1) the proportionality of $\dot{a}$ to K^n is empirical; and 2) more than one mechanism of atomic bond rupture, with different activation energies, might be involved, and the type of this mechanism might change with temperature.

Due to the relations $K = (GE_c)^{1/2}$ and $K_c = (G_f'E_c)^{1/2}$, where G = rate of energy release from the structure

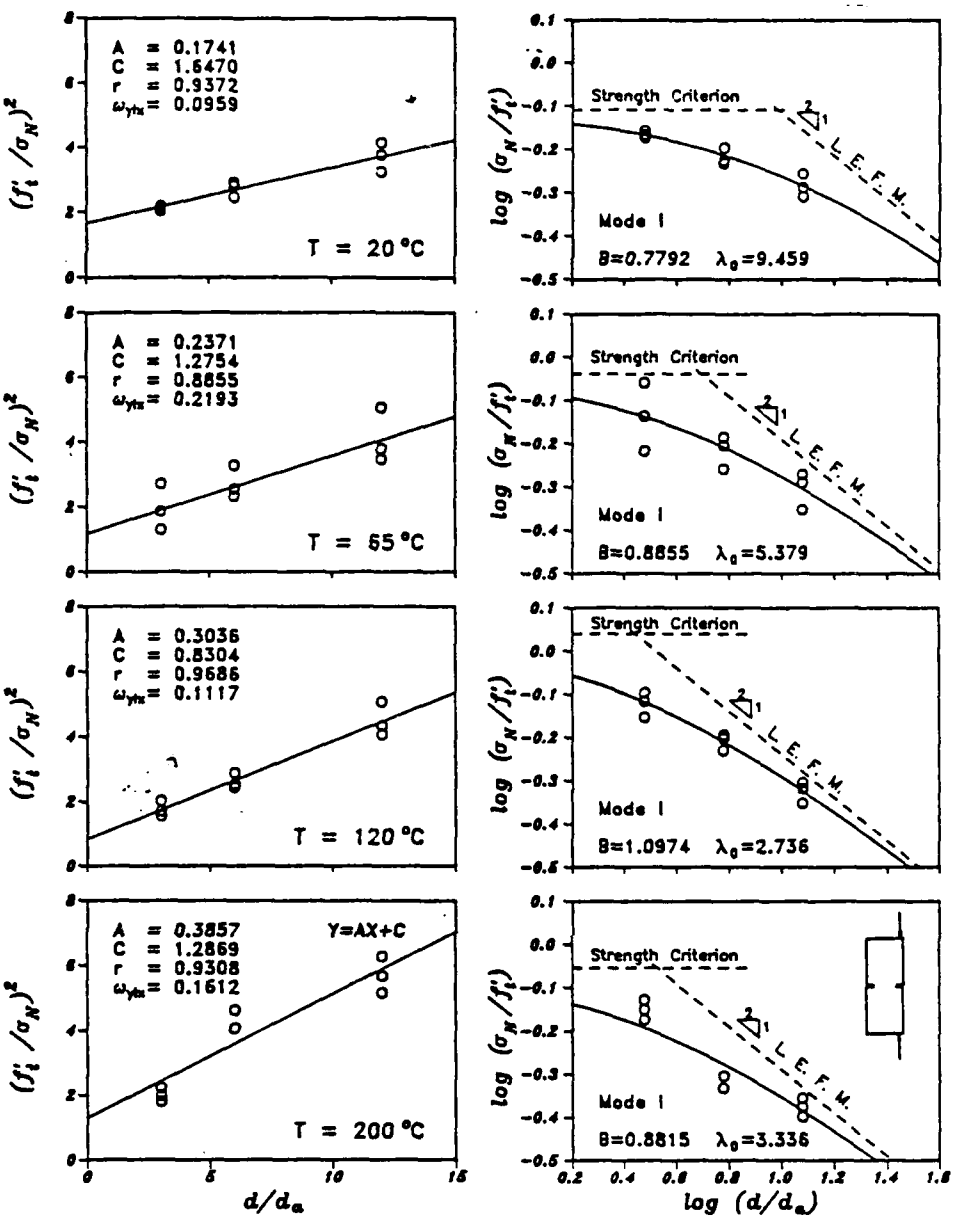


Fig. 4—Test results for dry eccentric compression specimens

into the fracture process zone,[19,20] Eq. (2) may be rewritten as

$$\dot{a} = v_c \left(\frac{G}{G_f'}\right)^{n/2} \exp\left[-\frac{U}{R}\left(\frac{1}{T} - \frac{1}{T_0}\right)\right] \qquad (3)$$

where T_0 = chosen reference temperature (normally $T_0 = 298$ K) and G_f' = value of fracture energy G_f at T_0. With regard to the finite size of fracture process zone in concrete, the relation $K = (GE_c)^{1/2}$ might at first be deemed inconsistent since it is based on linear elastic fracture mechanics. It is nevertheless consistent since, according to the size effect law, G is defined as the fracture energy of an infinitely large specimen, for which linear elastic fracture mechanics does apply.

Eq. (2) or (3) may serve as the basic relation for crack growth in time. Although the time-dependent fracture description in terms of the crack growth rate is no doubt physically more fundamental, the time-independent fracture description prevails in applications. In fact, what is known as fracture mechanics is a time-independent theory. So, deduction of the consequences of Eq. (2) for time-independent fracture description is needed. This was done in Reference 8.

The choice of the reference temperature T_0 in Eq. (3) is arbitrary. If temperature T is chosen as the reference temperature, then according to Eq. (3) the crack growth rate $\dot{a}$ at temperature T is simply expressed as

$$\dot{a} = v_c (G/G_f)^{n/2} \qquad (4)$$

(because $1/T_0 - 1/T = 0$ in this case). G_f represents the fracture energy at temperature T, while G_f' is the fracture energy at temperature T_0.

The crack growth rate expressions referred to T_0 or to T must be equivalent. Equating the expressions in Eq. (3) and (4), one obtains Bažant's[8] approximate formula

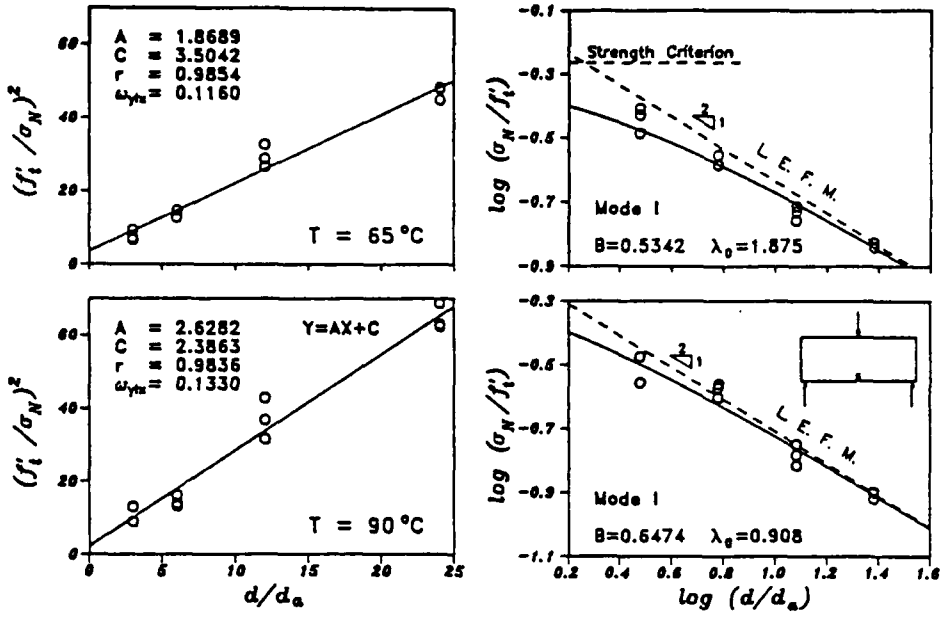


Fig. 5—Test results for wet three-point bend specimens

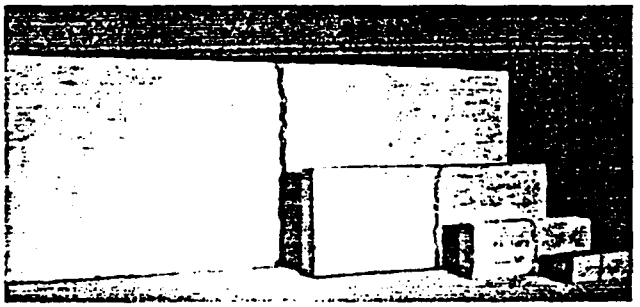


Fig. 6—One set of three-point bend specimens tested

$$G_f = G_f^\circ \exp(\gamma/T - \gamma/T_0) \qquad (5)$$

in which

$$\gamma = 2U/nR \qquad (6)$$

γ is a constant characterizing the given material.

Eq. (5) allows a simple determination of material parameters since it may be written in the form $Y = \gamma X + b$ where $Y = ln\ G_f^\circ$; $X = 1/T$; $b = ln\ G_f^\circ - \gamma/T_0$. So the values of γ and b may be found as the slope and the Y-intercept of the regression line in the plot of $ln\ G_f$ versus $1/T$, and $G_f^\circ = \exp(b + \gamma/T_0)$.

The values of activation energy U cannot be determined from the present types of experiments. According to Eq. (6), one would also need to determine exponent n. This would require measuring the rate of growth of the crack length at various load values.

The fracture energy values obtained experimentally at various temperatures are plotted in Fig. 9(b) as the data points. The linear regression plot, shown in Fig. 9(a), demonstrates that the present test results indeed fall quite close to a straight line in this plot. This confirms the validity of Eq. (5). The diagram on Fig. 9(b) shows the curves of Eq. (5) plotted for the optimum material parameter values of γ and G_f° obtained by linear regression. These values are (1 N/m = 0.005710 lb/in.):

Three-point bend, dry: $G_f^\circ = 34.48$ N/m, $\gamma = 581$ K
Eccentric compression,
dry: $G_f^\circ = 38.39$ N/m, $\gamma = 623$ K
Average of dry tests: $G_f^\circ = 36.38$ N/m, $\gamma = 602$ K
Three-point bend, wet: $G_f^\circ = 33.03$ N/m, $\gamma = 1875$ K

Note that the differences between the two specimen shapes (dry tests) are statistically insignificant, in view of the usual scatter for concrete. They are ± 5.5 percent for G_f° and ± 3.5 percent for γ, relative to the means. Since there is no significant difference between the eccentric compression and three-point bend dry tests, a combined line for both is plotted in Fig. 9(b), according to Eq. (5). The fracture energy $G_{f,dry}$ at reference temperature T_0 = 20 C (68 F) is obtained by linear regression using the combined data of both tests.

The portion of the wet specimen curve that lies above the temperature of 100 C (212 F) in Fig. 9(a) is strictly hypothetical because concrete cannot hold evaporable water above 100 C (unless the test was made in a pressurized chamber, but that would change G_f as well). Therefore, wet fracture energy for static loading above 100 C is physically meaningless.

For wet concrete at temperatures below about 20 C, Eq. (5) is probably inapplicable since the wet specimen curve in Fig. 9(a) would pass above the dry specimen curve. This would contradict the fact that the presence of moisture is known to weaken the bonds in hydrophylic porous solids.

Based on the previous results, the effect of the evaporable water content of concrete w per unit volume of concrete may be estimated. Let γ_0 and γ_1 be the values of γ at $w = 0$ (dry state) and at $w = w_1$ = saturation water content (wet state). As a crude estimate, we may assume a linear variation of γ as function of w. This

yields for any water content the approximation

$$\gamma = \gamma_0 + (\gamma_1 - \gamma_0)\frac{w}{w_1} \quad (7)$$

According to the results shown γ_0 = 602 K and γ_1 = 1875 K for the concrete tested.

The dependence of γ on w may be physically explained by a dependence of the activation energy U on w. Such a dependence is reasonable to expect since it is known that the activation energy of concrete creep depends on w. According to Eq. (7), we have

$$U = U_0 + (U_1 - U_0)\frac{w}{w_1} \quad (8)$$

with

$$U_0 = \tfrac{1}{2}(nR\gamma_0),\ U_1 = \tfrac{1}{2}(nR\gamma_1) \quad (9)$$

The linearity of Eq. (7) and (8) with respect to w is of course a hypothesis. To determine γ at intermediate water contents would be more difficult than the present tests. Providing control of the environmental relative humidity at arbitrary temperature would not be difficult; however, creating and maintaining a uniform value of some intermediate water content in the specimen would. It takes a very long time to dry a specimen to a uniform moisture content, and due to the coupling of temperature and humidity changes, it appears rather difficult to maintain a constant and uniform moisture content during the test. It would probably be necessary to make inferences from tests on specimens that are in a transient hygrothermal state. The advantage of the tests at either perfect saturation or perfect dryness is that a uniform moisture state is easily obtained and easily maintained at temperature changes.

That the activation energy should depend on the moisture content is not surprising. It is well established

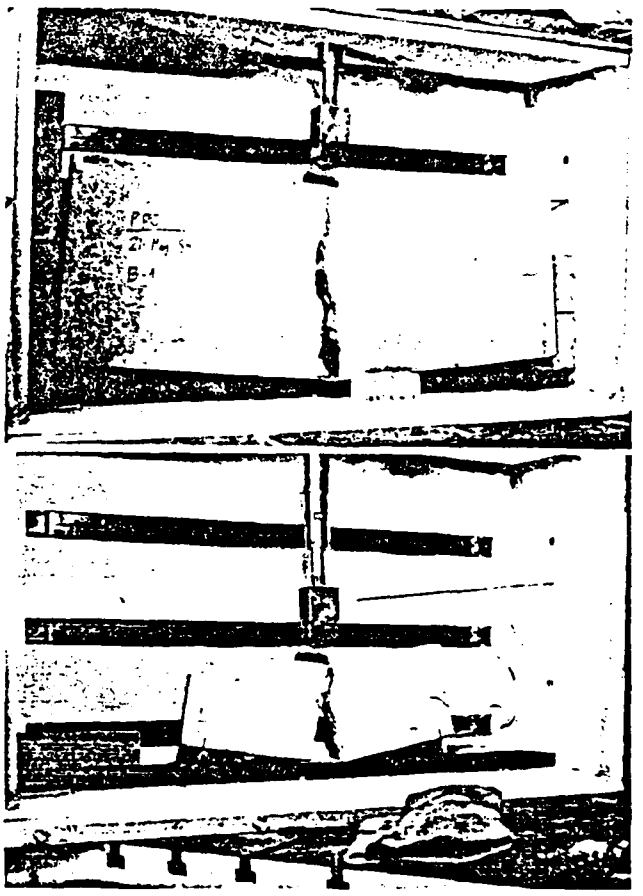


Fig. 7—Loaded specimens in an oven, installed in the testing machine

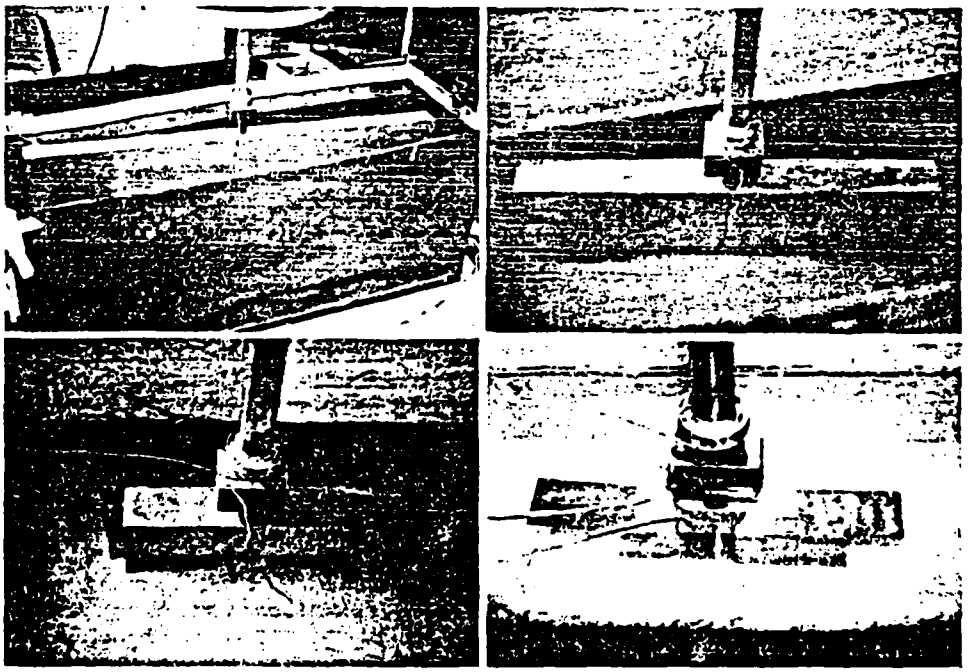


Fig. 8—Loaded specimens in the water tank

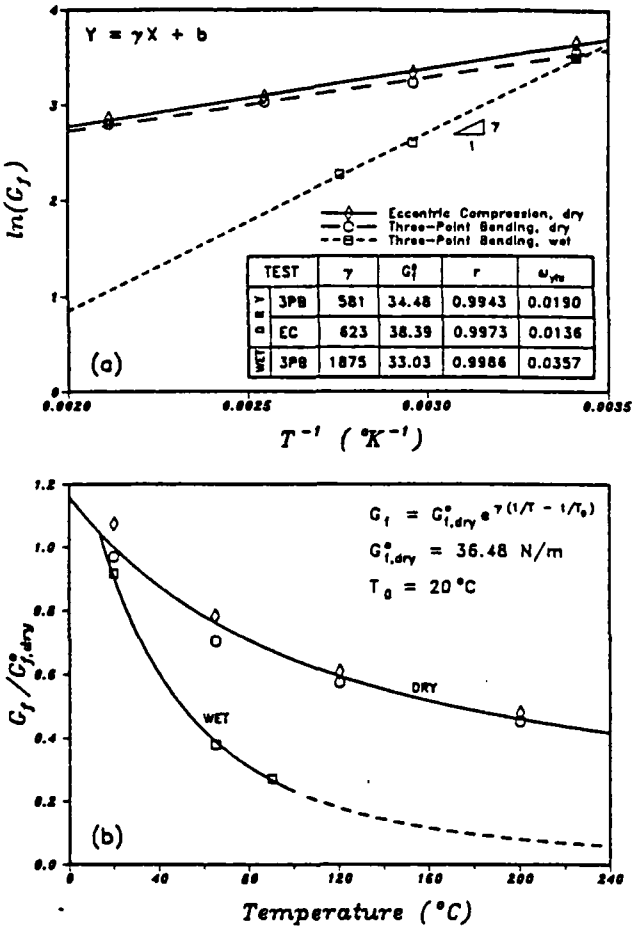


Fig. 9—Experimental results and theoretical formula for the effect of temperature on fracture energy for both dry and wet specimens: (a) linear regression lines; and (b) fracture energy versus temperature (γ = activation temperature in degrees K, G_f^0 = fracture energy at reference temperature T_0 = 20 C in N/m)

that the activation energy of concrete creep strongly varies with the moisture content, and the same is true of other porous hydrophylic materials such as wood. Also, the creep activation energy is known to depend on w nearly linearly.

The physical explanation of the dependence of U on w lies probably in the very high disjoining pressures caused by layers of hindered adsorbed water or interlayer water within the finest pores of the cement gel. These pressures are superimposed on the stresses due to load, thus lowering the energy barrier for bond ruptures. A mathematical model based on this mechanism would no doubt be of help for determining the dependence of U on w more accurately and for interpreting test results at intermediate values of w.

CONCLUSIONS

1. The size effect law is applicable not only at room temperature but also at elevated temperatures up to 200 C (392 F). Its parameters depend on temperature as well as specific water content of concrete.

2. The Mode I fracture energy of concrete significantly depends on temperature T. It decreases monotonically and smoothly as T increases. For concrete near saturation water content, this dependence is more pronounced than for predried concrete.

3. The effect of specific water content on the fracture energy is small at room temperature but grows as temperature increases and is very large at temperatures close to 100 C (212 F).

4. The measured dependence of Mode I fracture energy on temperature agrees with Bažant's[8] formula [Eq. (5)] based on the activation energy theory (rate-process theory).

5. Very different types of specimens such as the three-point bend and eccentric compression specimens yield approximately the same fracture energy values at various temperatures. This indicates that the present results for the fracture energy should represent true material properties.

Table 3 — Fracture energy values obtained from tests, N/m*

Test	Age	20 C (68 F)	65 C (149 F)	90 C (194 F)	120 C (248 F)	200 C (392 F)
3-point bend, dry	28	35.24	25.58	—	20.87	16.39
Eccentric compression, dry	28	39.01	28.42	—	22.06	17.36
3-point bend, wet	28†	33.30‡	13.72†	9.79†	—	—
3-point bend, wet	41	—	14.57	10.40	—	—

* 1 N/m = 0.005710 lb./in.
† Corrected from age 41 days.
‡ From Bažant and Pfeiffer, not in water bath but in air immediately after removal from moist room.

ACKNOWLEDGMENTS

Grateful acknowledgment is due to the U.S. Air Force Office of Scientific Research (Contract No. F49620-87-C-0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu) for partial support of the theoretical work and measurements, and for the preparation of specimens for the cooperative project between Universidad Politécnica de Madrid, Spain, and Northwestern University, funded under the U.S.-Spanish treaty (Grant CCA-8309071).

REFERENCES

1. Tetelman, A. S., and McEvily, A. J., *Fracture of Structural Materials*, John Wiley & Sons, New York, 1967, pp. 274-276.

2. Kanninen, M. F., and Popelar, C. H., *Advanced Fracture Mechanics*, Oxford University Press, New York, 1985, pp. 17-18, 26-30, and 82.

3. Cherepanov, G. P., *Mechanics of Brittle Fracture*, McGraw-Hill Book Co., New York, 1979 (translated from the Russian original published by Nauka, Moscow, 1974), pp. 174 and 455.

4. Hillerborg, A.; Modéer, M.; and Petersson, P.E., "Analysis of Crack Formation and Crack Growth in Concrete by Means of Fracture Mechanics and Finite Elements," *Cement and Concrete Research*, V. 6, No. 6, Nov. 1976, pp. 773-781.

5. Hillerborg, A., "The Theoretical Basis of a Method to Determine the Fracture Energy G_f of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 18, No. 106, July-Aug. 1985, pp. 291-296.

6. Bažant, Z. P., "Fracture Mechanics and Strain-Softening of Concrete," *Preprints*, U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures, Japan Society for the Promotion of Science, Tokyo, May 1985, pp. 71-92.

7. Bažant, Z. P., "Fracture in Concrete and Reinforced Concrete," *Mechanics of Geomaterials: Rocks, Concrete, Soils*, John Wiley & Sons, Chichester, 1985, pp. 259-304.

8. Bažant, Z. P., "Fracture Energy of Heterogeneous Materials

and Similitude," *Preprints*, International Conference on Fracture of Concrete and Rock (Houston, June 1987), Society for Experimental Mechanics, Bethel, 1987, pp. 390-402.

9. Bažant, Zdeněk P.; Kim, Jin-Keun; and Pfeiffer, Phillip A., "Nonlinear Fracture Properties from Size Effect Tests," *Journal of Structural Engineering*, ASCE, V. 112, No. 2, Feb. 1986, pp. 289-307.

10. Bažant, Z. P.; Kim, J.-K.; and Pfeiffer, P. A., "Determination of Nonlinear Fracture Parameters from Size Effect Tests," NATO Advanced Research Workshop on Application of Fracture Mechanics to Cementitious Composites, Northwestern University, Evanston, Sept. 1984, pp. 143-169.

11. Bažant, Zdeněk P., "Size Effect in Blunt Fracture: Concrete, Rock, and Metal," *Journal of Engineering Mechanics*, ASCE, V. 110, No. 4, Apr. 1984, pp. 518-535.

12. Bažant, Z. P.; Sener, S.; and Prat, P. C., "Size Effect Tests of Torsional Failure of Concrete Beams," *Report* No. 86-12/428s, Center for Concrete and Geomaterials, Northwestern University, Evanston, Dec. 1986, 18 pp.

13. Bažant, Zdeněk P.; and Kim, Jin-Keun, "Size Effect in Shear Failure of Longitudinally Reinforced Beams," ACI JOURNAL, *Proceedings* V. 81, No. 5, Sept.-Oct. 1984, pp. 456-468.

14. Bažant, Zdeněk P., and Cao, Zhiping, "Size Effect in Punching Shear Failure of Slabs," *ACI Structural Journal*, V. 84, No. 1, Jan.-Feb. 1987, pp. 44-53.

15. Bažant, Z. P., and Pfeiffer, P. A., "Shear Fracture Tests of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 19, No. 110, Mar.-Apr. 1986, pp. 111-121.

16. Bažant, Zdeněk P., and Thonguthai, Werapol, "Pore Pressure and Drying of Concrete at High Temperature," *Proceedings*, ASCE, V. 104, EM5, Oct. 1978, pp. 1059-1079.

17. Bažant, Zdeněk P., and Panula, Liisa, "Statistical Stability Effects in Concrete Failure," *Proceedinngs*, ASCE, V. 104, EM5, Oct. 1978, pp. 1195-1212.

18. Neville, A. M., *Properties of Concrete*, 3rd Edition, Pitman Publishing Limited, London, 1981, 779 pp.

19. Broek, D., *Elementary Engineering Fracture Mechanics*, Sijthoff and Noordhoff, International Publishers, Netherlands, 1978, 408 pp.

20. Knott, J. F., *Fundamentals of Fracture Mechanics*, Butterworths, London, 1973, 273 pp.

21. Tada, H.; Paris, P. C.; and Irwin, G. R., *The Stress Analysis of Cracks Handbook*, Del Research Corp., Hellertown, 1973.

22. Bažant, Zdeněk P., and Oh, B. H., "Crack Band Theory for Fracture of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 16, No. 93, May-June 1983, pp. 155-177.

23. Cottrell, A. H., *The Mechanical Properties of Matter*, John Wiley & Sons, New York, 1964, 430 pp.

24. Evans, A. G., and Fu, Y., "The Mechanical Behavior of Alumina," *Fracture in Engineering Materials*, Noyes Publications, Park Ridge, p. 71.

25. Thouless, M. D.; Hsueh, C. H.; and Evans, A. G., "A Damage Model of Creep Crack Growth in Polycrystals," *Acta Metallurgica*, V. 31, No. 10, 1983, pp. 1675-1687.

APPENDIX A4

Journal of the Am. Ceram. Soc.
submitted January 1990

SIZE EFFECT IN FATIGUE FRACTURE

by

Zdenek P. Bazant and Kangming Xu

Report No. 89-12/498f

CENTER FOR ADVANCED CEMENT-BASED MATERIALS

ROBERT R. McCORMICK SCHOOL OF ENGINEERING AND
APPLIED SCIENCES IN THE TECHNOLOGICAL INSTITUTE
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS

December 1989

SIZE EFFECT IN FATIGUE FRACTURE

By Zdeněk P. Bažant and Kangming Xu

Center for Advanced Cement-Based Materials
McCormick School of Engineering and Applied Science
Northwestern University, Evanston, Illinois 60208

ABSTRACT: Crack growth in geometrically similar notched concrete specimens of various sizes due to load repetitions is measured by means of the compliance method. It is found that Paris law, which states that the crack length increment per cycle is a power function of the stress intensity factor amplitude, is valid only (1) for one specimen size or (2) asymptotically for very large specimens. To obtain a general law, Paris law is combined with the size effect law for monotonic fracture proposed previously by Bažant. This leads to a size-adjusted Paris law which gives the crack length increment per cycle as a power function of the amplitude of a size-adjusted stress intensity factor. The size adjustment is based on the brittleness number of the structure, representing the ratio of the structure size to the transitional size. The latter size is indicated by the experiments to be much larger than that for monotonic loading, which means that the brittleness number for cyclic loading is much less than the brittleness number for monotonic loading. The crack growth is alternatively also characterized in terms of the nominal stress amplitude. In the latter form the size effect vanishes for small structures, while in terms of the stress intensity factor amplitude it vanishes for large structures. The curves of crack length vs. the number of cycles are also calculated and are found to agree with data.

INTRODUCTION

Repeated loading causes cracks to grow. This phenomenon, called fatigue fracture, has been studied extensively for metals and is now understood

relatively well (e.g. Kanninen and Popelar, 1985; Broek, 1986; and Rolfe and Barsom, 1977). For concrete, however, the knowledge of fatigue fracture is still rather limited. Aside from fatigue studies outside the context of fracture mechanics (e.g. Norby, 1958; Murdock and Kesler, 1958; Stelson and Cerrica, 1958; Hilsdorf and Kesler, 1966; Kaplan, 1959; and Avram et al.), fracture under cyclic loading was experimentally investigated by Buluch et al.(1987), Perdikaris and Calomino (1987), Suresh, Tschegg and Brockenbrough (1989), Suresh(1989), and Zhang et al. (1987). These investigations indicated that Paris law for crack growth under repeated loading (Paris, 1962; Paris, Gomez and Anderson, 1961; and Paris and Erdogan, 1963) can be at least to some extent transplanted from metals to concrete. However, one important aspect - the size effect - has apparently escaped attention so far, as far as fatigue is concerned. The objective of the present study is to investigate the size effect, both experimentally and theoretically.

FATIGUE TESTS OF SIMILAR NOTCHED BEAMS OF DIFFERENT SIZES

Two series of three concrete three-point-bend beam specimens of different sizes were tested in a closed-loop servo-controlled (MTS) testing machine; see Figs. 1 and 2. The ratio of cement : sand : aggregate : water in the concrete mix was 1 : 2 : 2 : 0.6. The aggregate was crushed limestone of maximum size 0.5 in. (12.7 mm). The sand was siliceous river sand passing through sieve No. 1 (size 5 mm). Type I portland cement, with no admixtures, was used. The specimens were cast in plywood molds with the beam side face in a horizontal position. From each batch of concrete, three beam specimens of different sizes (one of each size) were cast, along with companion cylinders of diameter 3 in. (76.2 mm) and length 6 in. (152 mm) for strength measurement. After the standard 28-day curing of the companion cylinders, their mean compression strength was f'_c = 4763 psi (32.8 MPa), with standard deviation 216 psi (1.49 MPa). The average direct tensile strength was estimated as $f'_t = 6\sqrt{f'_c}$ = 414 psi (2.86 MPa), average Young's modulus as $E = 57{,}000\sqrt{f'_c} = 3.93\times10^6$ psi (27,120 MPa) (all in psi). An additional series of nine notched beam specimens of three different sizes, three of each size, were cast from a different batch of concrete in order to determine fracture properties under monotonic loading by the size effect method.

The specimens of different sizes, both for cyclic tests and companion monotonic tests, were geometrically similar in two dimensions, including the lengths of their notches. The thickness of all the specimens was the same (b = 1.5 in. = 38.1 mm). The beam depths were d = 1.5, 3 and 6 in. (38.1, 76.2 and 127 mm), the span was $L = 2.5\ d$, and the notch depth was $a_0 = d/6$ (Fig. 2).

The specimens were cast vertically (i.e.with the loaded side on top) and demolded after 24 hours. Subsequently the specimens were cured in a

moist room of 95% relative humidity and 79 F (26.1 C) temperature until the time of the test. Just before the test, the notches were cut by a band saw. When tested, the specimens were days old. During the test, the specimens were positioned upside down, with the notch on top. The effect of specimen weight was negligible. During the cyclic tests (as well as the monotonic control tests), the laboratory environment had a relative humidity of about 65% and temperature about 78 F (25.6 C). The crack-mouth opening displacement (CMOD) was measured by an LVDT gage supported on metallic platelets glued to concrete.

In view of the difficulty of optical and other direct measurements, the crack length was measured indirectly, by the CMOD-compliance method. Although some doubts existed, the validity of this method for concrete has recently been confirmed by Swartz and Go (1984) and Perdikaris, Calomino and Chudnovsky (1986). The procedure was as follows: (1) Mark the notch lengths on the beam flank surface according to a logarithmic scale. (2) Cut the notch at midspan by a band saw. (3) Mount the LVDT gage and set up the specimen in the testing machine. (4) Apply the load (at the rate 0.0004 in./min.) and, after several load cycles between zero and maximum, record (with an X-Y recorder) the plot of load vs. CMOD, from which the specimen compliance is later calculated (the load is small enough so no crack would start). (5) Remove the specimen from the machine, extend the notch by sawing a cut up to the next mark, and repeat the procedure. An example of the relevant portion of the calibration curve of specimen compliance vs. the ratio of notch length to beam depth is seen in Fig. 3, which shows the measured data points as well as the curve of finite element results. Their deviation is within the range of normal experimental scatter in concrete testing.

In the cyclic tests, linear ramp load and frequency 0.033 to 0.040 Hz

was used. In all the tests, the load minima were zero and the load maxima were constant and equal to 80 % of the monotonic peak load for a specimen of the same size.

TEST RESULTS

A grapher was used to continuously plot the load history (Fig. 5a) and the CMOD history (Fig. 5b) in each cycle (only the terminal part of these histories near failure is shown in Fig. 5). The effective crack length, representing the coordinate of the tip of an equivalent elastic crack, has been determined from the compliance given by the slope of the unloading segment of the measured load-CMOD curve,according to the compliance calibration curve shown in Fig. 3. Based on the effective crack length, one can determine the equivalent stress intensity factor, K_I^{eq}. The maximum loads measured in the companion monotonic tests are given in Table 1. The table also gives the values obtained previously (Bažant and Pfeiffer, 1987) on the same type of concrete and the same specimens. The fact the these previous values were nearly identical confirms reproducibility. The measured curves of load P vs. CMOD are shown in Fig.5.

According to the fatigue fracture theory (e.g. Broek, 1986; Kanninen and Popelar, 1985), crack growth depends on the amplitude ΔK_I of the stress intensity factor K_I for the current effective (elastically equivalent) crack length, a. As is well known,

$$K_I = \frac{P\, f(\alpha)}{b\sqrt{d}} = \sigma_N \sqrt{d}\, \frac{f(\alpha)}{c_n} , \qquad \alpha = \frac{a}{d} \tag{1}$$

in which α = relative crack length, d = characteristic dimension of the specimen or structure, b = specimen thickness, P = load, $f(\alpha)$ = function depending on specimen geometry, which can be obtained by elastic finite element analysis and for typical specimen geometries is given in handbooks; for the present three-point bend specimens, $f(\alpha) = (1 - \alpha)^{-3/2}$

$(1 - 2.5\ \alpha + 4.49\ \alpha^2 - 3.98\ \alpha^3 + 1.33\ \alpha^4)$, as determined by curve-fitting of finite element results (Gettu, Bažant and Karr, 1989); and

$$\sigma_N = c_n\ P\ /\ bd \qquad (2)$$

where σ_N = nominal stress; and c_n = coefficient chosen for convenience, for example so that σ_n would represent the maximum stress according to the bending theory formula for the ligament cross section; in that case $c_n = 3L/2(d-a_0)$ = constant for geometrically similar specimens of different sizes. If the crack length is very small (microscopic), i.e. $a \ll d$, then $K_I = \sigma_y \sqrt{\pi a}$, where σ_Y is the local macroscopic stress calculated as if there were no crack; but in our case, due to the notch, the crack is not microscopic.

The fatigue test results are presented in Fig. 6, which shows the relative effective crack length, $\alpha = a/d$ (calculated from the compliance data), as a function of the number, N, of cycles, for each of the six specimens tested (two of each size). From this figure, the measured points in the plot of $\log(\Delta a/\Delta N)$ vs. $\log(\Delta K_I)$ have been constructed, as shown in Fig. 7a for a series of three specimens, one of each size, and in Fig. 7b for both series, two specimens of each size. Here ΔK_I is the amplitude of the equivalent stress intensity factor calculated on the basis of the current effective crack length a (of course $\Delta K_I = K_I$ because the lower load limit is 0 in the present tests); ΔN represents the increments between the consecutive cycles number N = 10, 110, 210,...710, 760, 810, 820, 830,...900, 905, 910,...965, 966, 967,....974, and Δa are the corresponding increments of a. The small, medium and large specimens failed at N = 974, 850 and 882, respectively, for the first group, and at N = 939, 1286 and 1083 for the second group. The units in Fig. 6 and 7 are in.(25.4 mm) for a and lb./in.$^{3/2}$ (1.099 N cm$^{-3/2}$) for K_I.

SIZE EFFECT IN PARIS LAW

As demonstrated extensively for metals, fatigue fracture follows approximately the empirical Paris law (Paris, Gomez and Anderson, 1961; Paris and Erdogan, 1963):

$$\frac{\Delta a}{\Delta N} = C \left(\frac{\Delta K_I}{K_{If}} \right)^n \qquad (3)$$

in which C and n are empirical material constants; n is nondimensional and, for reasons of physical dimensions, C has the dimension of length, and K_{If} = constant = fracture toughness. Δ in Eq.3 are the differences taken over one cycle or a sufficiently small number ΔN of cycles. When the range of ΔK_I is very broad, the test results for metals deviate below the plot of Eq. 3 for very low ΔK_I and above it for very high ΔK_I (Kanninen and Popelar, 1985), but for a broad enough range Eq.3 is good for metals.

According to Paris law, the plot in Fig. 7 would have to be a single inclined straight line of slope n. This is seen to be indeed true for each specimen of each size – see Fig. 7a; from the slope and intercept of the regression line of the data points for each specimen one gets; n = 11.78, log C = -40.2 (units in in. and lb.) for the smallest size, n = 9.97, log C = - 36.0 for the medium size, and n = 9.27, log C = -34.8 for the largest size, with the overall average n = 10.57. The optimum fits by the straight lines of slope equal to this average slope are shown in Fig. 8. Fig. 8b shows the data for all the six specimens (two of each size) compared to the same straight lines as in Fig. 8a. The fit is not exactly optimum, but is still satisfactory. As we see, however, there are gross discrepancies among various specimen sizes — not among the slopes n but among the C-values (Fig. 8a,b). Thus, the Paris law is invalid for concrete in general.

This is not really a surprising result, in view of the well-known deviations of concrete behavior from linear elastic fracture mechanics (LEFM). For monotonic loading, these deviations are quite well described

in terms of the size effect law proposed by Bažant (1984, 1986):

$$\sigma_N = B f_u (1 + \beta)^{-1/2}, \qquad \beta = d / d_0 \tag{4}$$

(Fig. 8a) in which B and d_0 are empirical constants characterizing both the material properties and specimen geometry; f_u is a measure of the tensile strength, e.g. $f_u = f'_t$ = direct tensile strength; and σ_N = nominal strength = nominal stress at maximum (ultimate) load P_u, $\sigma_N = c_n P_u / bd$ (Eq. 2). The meaning of d_0 is the transitional size (dimension) of the structure; for $d \gg d_0$ (i.e. $\beta \gg 1$) failure is governed by LEFM while for $d \ll d_0$ ($\beta \gg 1$) failure is governed by strength theory (or plasticity). For this reason, β is called the brittleness number (Bažant, 1987; Bažant and Pfeiffer, 1987; Bažant and Kazemi, 1989a,b). Eq. 4, describing a transition from the size effect of plasticity (i.e. no size effect in strength) to the size effect of LEFM (i.e. the maximum possible size effect in strength), is only approximate; but the accuracy appears to be sufficient for the size range up to 1:20.

Due to deviations from LEFM, the critical value K_{Ic} of stress intensity factor, K_I, at which the crack can propagate at monotonic loading, is not constant but depends on specimen size and geometry. An unambiguous definition of K_{Ic} (as well as of the corresponding fracture energy $G_f = K_{Ic}^2 / E$) can be given only in terms of extrapolation to a specimen of infinite size (Bažant, 1987; Bažant and Pfeiffer, 1987). At infinite size, the specimen geometry cannot matter because (1) the fracture process zone occupies an infinitely small fraction of the specimen volume so that the whole specimen is in an elastic state, and (2) the near-tip asymptotic elastic field to which the fracture process zone is exposed at its boundary is the same for any specimen geometry. Thus, $K_{If} = \lim K_{Ic}$ for $d \to \infty$ (more generally, K_{If} could be defined as the limit of Rice's J-integral).

To determine the size dependence of K_{Ic}, we first express K_{Ic} by substituting $P = P_u$ with $P_u = \sigma_N\, bd/c_n$ into Eq. 1. Then, substituting Eq. 4 for σ_N, we obtain (Fig.8b):

$$K_{Ic} = \frac{B\, f_u \sqrt{\beta d_0}}{\sqrt{1+\beta}} \frac{f(\alpha_0)}{c_n} = K_{If} \left(\frac{\beta}{1+\beta}\right)^{1/2} \tag{5}$$

in which K_{If} is a *constant* expressed as

$$K_{If} = B\, f_u \sqrt{d_0}\, \frac{f(\alpha_0)}{c_n} \tag{6}$$

where we recognized that K_{If} is the fracture toughness as defined above because it represents the limit of Eq. 4 for $d \to \infty$ or $\beta \to \infty$. Eq. 6, derived in Bažant (1987) (see also Bažant and Pfeiffer, 1987), means that the fracture toughness can be obtained from the size effect law parameters B and d_0. These parameters, in turn, can be obtained if the values of $(f_u/\sigma_N)^2$, calculated from the measured peak values P_u of the curves in Fig. 5, are plotted versus d. In this plot, the slope of the regression line is $1/Bd_0$ and the vertical axis intercept is $1/B$. Linear regression of the peak load values from the present monotonic fracture tests provided B = 0.520, d_0 = 2.86 in.(72.6 mm), from which K_{If} = 73.5 lb./in.$^{3/2}$ (1.099 N cm$^{-3/2}$).

Now, how should the size effect be manifested in fatigue fracture? The monotonic fracture represents the limiting case of fatigue fracture as $\Delta K_I \to K_{Ic}$ at $N \to 1$. If the value of ΔK_I is close to K_{Ic}, $\Delta a/\Delta n$ must become very large while K_{Ic} must at the same time follow the size effect law. This can be modeled by replacing in Paris law (Eq. 3) the constant fracture toughness K_{If} by the size-dependent equivalent fracture toughness K_{Ic} given by Eq. 5. So a generalized, size-adjusted Paris law may be written as

$$\frac{\Delta a}{\Delta N} = C \left(\frac{\Delta K_I}{K_{Ic}}\right)^n \tag{7}$$

The transition size d_0 in the size effect law (Eq. 3) applies only to

peak-load states at monotonic loading. For fatigue fracture, however, it is unclear whether the value of the transitional size d_0, needed to calculate K_{Ic} (Eq. 5), should be constant and the same as for monotonic fracture, or whether it should vary as a function of the ratio

$$\rho_K = \Delta K_I / K_{If} \tag{8}$$

Assuming d_0 = 2.86 in., same as for monotonic fracture, the optimum fit of the present test data by the size-adjusted Paris law, Eq. 7, is shown in Fig. 7c,d for the small, medium and large specimens by the dashed straight lines. As one can see, the fit has improved substantially compared to Fig. 7a,b, but not enough. This finding means that the transitional size d_0 cannot be constant but must depend on ρ_K.

Let us now assume that, for the present tests in which ρ_K = 0.8, the value of d_0 is 28.6 in., i.e. 10-times larger. Then the optimum fit of the present test data becomes as shown by the three straight lines in Fig. 7e,f. The closeness of fit is now satisfactory.

According to Eq. 7, if $\log(\Delta a/\Delta N)$ is plotted against $\log(\Delta K_{Is})$, the data points for the specimens of all sizes should have a common regression line. Fig. 9 confirms that for d_0 = 28.6 in. this is approximately indeed so.

As shown in Bažant and Kazemi (1989a,b), d_0 is proportional to the effective length of the fracture process zone. So the value of d_0 deduced from the present tests implies that the fracture process zone at load cycling with amplitude 80 % of the maximum monotonic load should be about 10-times larger than it is for monotonic loading. Although the factor 10 might be too large to still accept equivalent LEFM calculations for the present-size specimens, it nevertheless appears that the fracture process zone is greatly enlarged by load cycling. Intuitively, this is not an unexpected conclusion.

CRACK GROWTH DEPENDENCE ON NOMINAL STRESS

Expressing ΔK_I according to Eq. 1, Paris law may be written in terms of the amplitude $\Delta\sigma_N$ of the nominal stress:

$$\frac{\Delta a}{\Delta N} = \bar{C}\,(\sqrt{\beta}\ \Delta\sigma_N)^n, \qquad \bar{C} = C\left(\sqrt{d_0}\,\frac{f(\alpha)}{c_n K_{If}}\right)^n \tag{9}$$

where $\bar{C}$ is a constant if the specimens are geometrically similar, and $\beta = d/d_0$. Likewise, after expressing ΔK_I in Eq. 7 from Eq. 1, the size-adjusted Paris law may also be written in terms of the nominal stress amplitude:

$$\frac{\Delta a}{\Delta N} = \bar{C}\,(\sqrt{1+\beta}\ \Delta\sigma_N)^n \tag{10}$$

For very small specimen sizes, $d \ll d_0$, this law simplifies as

$$\frac{\Delta a}{\Delta N} = \bar{C}\,(\Delta\sigma_N)^n \tag{11}$$

i.e., the size effect in fatigue fracture disappears. On the other hand, for very large sizes, $d \gg d_0$, Eq. 10 asymptotically approaches Eq. 9 and the size effect becomes the strongest possible.

The present data for the first series of fatigue specimens of three sizes (one specimen per size), shown already in Fig. 7a,c,d, are now plotted in Fig. 10a with $\log \Delta\sigma_N$ as the abscissa, and in Fig. 10b with the logarithm of the size-adjusted nominal stress amplitude as the abscissa. The three straight regression lines predicted by Eq. 11 are also drawn.

Comparison of Fig. 10a with Fig. 7 reveals that, in terms of the nominal stress amplitude, the effect of size is much weaker than in terms of the stress intensity factor amplitude. This agrees with the conventional wisdom that concrete structures are relatively insensitive to fatigue. But for a sufficiently large size this ceases to be true.

The foregoing observation is of course incidental. It is due to the fact that all the specimens tested were relatively small. According to the size effect law, the effect of size on the nominal strength vanishes when

the size is very small, and is maximum when the size is very large. The opposite is true for the size dependence of the stress intensity factor required for fracture growth (Eqs. 4 and 7, Fig. 4). Thus, according to the present theory, the size effect in Fig. 10 should become very large, and in Fig. 7 very small, when the specimen or structure is very large. For large structures, fatigue is better characterized in terms of the stress intensity factor amplitude, while for small structures, fatigue is better characterized in terms of the nominal stress amplitude. (Note that a good absolute measure of structure size is its brittleness number — the structure is very large if $\beta \gg 1$ and very small if $\beta \ll 1$; see Bažant and Pfeiffer, 1987).

As we recall from Fig. 7, the larger the size, the smaller is the crack length increment per cycle for the same stress intensity factor amplitude. For the same nominal stress amplitude, though, the larger the size, the larger is the crack length increment per cycle; see Fig. 10. The last property is the same as in monotonic fracture — a size effect in the usual sense of a weakening of the structure with an increase of its size.

CRACK LENGTH VS. NUMBER OF CYCLES

It is also interesting to integrate Eq. 3. Replacing Δ with d, substituting $da = d_0 d\alpha$ and separating the variables, one has $N = \int (\Delta K_I)^{-n} d\alpha \; d_0/C$. Then, substituting Eq. 1 one gets

$$N = (d_0/C)(b\sqrt{d}/P)^n \int [f(\alpha)]^{-n} \, d\alpha \qquad (12)$$

Numerical integration yields the three curves of the relative crack length, α, versus the number of cycles plotted in Fig. 6 for the three specimen sizes. As we see, these curves agree with the measurements relatively well. This again confirms that fracture mechanics is applicable despite the fact that the size effect in terms of the nominal stress is in the present tests quite mild.

CONCLUSIONS

1. Paris law for fatigue fracture is applicable to concrete; but for structures with *macroscopic* (large) cracks this is so only (1) for one size of the specimen or structure or (2) asymptotically for very large sizes. Otherwise a modification of this law is required.

2. Paris law may be adjusted for the size effect by combining it with the size effect law proposed in Bažant (1984) for nominal strength in monotonic loading. This leads to a law in which the crack length increment per cycle is a power function of the amplitude of a size-adjusted stress intensity factor.

3. The size-adjusted Paris law is corroborated by the measurements of crack length growth due to fatigue loading of repeatedly loaded notched specimens of size range 1 : 4. The calculated growth of the crack length with the number of cycles also agrees with tests.

4. The fatigue crack growth can also be represented as a function of the nominal stress amplitude and the brittleness number. In the latter form, the size effect vanishes for very small structures, while in terms of the stress intensity factor amplitude it vanishes for very large structures. For sufficiently large structures, fatigue is better characterized in terms of the stress intensity factor amplitude, whereas for small structures fatigue is better characterized in terms of the nominal stress amplitude.

5. The transitional size, separating predominantly brittle failures from predominantly ductile (plastic or strength-governed) failures, appears to be much larger (about 10-times larger) for cyclic loading than it is for monotonic loading. Consequently, the brittleness number for a structure under cyclic loading is much smaller than it is for monotonic loading.

6. It must be emphasized that the scope of the present tests has been insufficient to check whether the foregoing conclusions are also valid when: (1) the stress intensity factor amplitude is varied; (2) the lower limit of the cyclic load is not zero; (3) the number of cycles per minute is varied; and (4) different specimen geometries and notch lengths are considered.

ACKNOWLEDGMENT.- Partial financial support has been received: (1) for development of the theory – from AFOSR contract F49620-87-C-0030DEF with Northwestern University, and (2) for experiments – from NSF Science and Technology Center for Advanced Cement-Based Materials at Northwestern University (NSF Grant DMR 880-8432).

References

Avram, C., et al., Concrete Strength and Strains, Elsevier Scientific Pub. Co., New York, 1981.

Barsom, J. M., and Rolfe, S. T., Fracture and Fatigue Control in Structures: Applications of Fracture Mechanics, N. J. Prentice-Hall, 1987.

Bažant, Z. P., "Size Effect in Blunt Fracture: Concrete, Rock, Metal", J. of Engrg. Mech., ASCE, Vol. 110, No. 4, Apr. 1984, pp. 518-535.

Bažant, Z. P., "Mechanics of Distributed Cracking", Applied Mech. Rev., Vol. 39,No. 5, May 1986, pp.675-705.

Bažant, Z. P., and Pfeiffer, P. A., "Determination of Fracture Energy from Size Effect and Brittleness Number", ACI Material J., Vol. 84, No. 6, Nov.-Dec., 1987, pp.463-480.

Bažant, Z. P., "Fracture Energy of Heterogeneous Material and Similitude", Fracture of Concrete and Rock, SEM/RILEM International Conference, Editors, Shah, S. P., and Swartz, S. E., Houston, Texas, June 1987, pp. 229-241.

Bažant, Z. P., and Kazemi, M. T., "Determination of Fracture Energy, Process Zone Length and Brittleness Number from Size Effect, with Application to Rock and Concrete", Report No. 88-7/498d, Center for Concrete and Geomaterials, Technological Institute, Northwestern University, July 1988.

Bažant, Z. P., and Kazemi, M. T.,"Size Effect in Fracture of Ceramics and Its Use to Determine Fracture Energy and Effective Process Zone Length", Report No. 89-6/498s, Center for Advanced Cement-based Materials, McCormick School of Engineering and Applied Science, Northwestern University, June, 1989.

Broek, D., The Practical Use of Fracture Mechanics, Boston, Kluwer Academic Publishers, 1988.

Baluch, M. H., Qureshy, A. B., and Azad, A. K., " Fatigue Crack Propagation in Plain Concrete", Fracture of Concrete and Rock, SEM/RILEM International Conference, Editors, Shah, S. P., and Swartz, S. E., Houston, Texas, June 1987, pp. 80-87.

Gettu, R., Bazant, Z. P., and Karr, M. E., "Fracture Properties and Brittleness of High Strength Concrete", Report No. 89-10/B627f, Center for Advanced Cement-based Materials, McCormick School of Engineering and Applied Science, Northwestern University, Oct., 1989.

Hilsdorf, H. K., and Kesler, C. E., " Fatigue Strength Concrete under Varying Flexural Stresses", ACI J., Proceedings, Vol. 63, No. 10, Oct. 1966, pp.1059-1075.

Kanninen, M. F., and Popelar, C. H., Advanced Fracture Mechanics, Oxford University Press, 1985

Kaplan, M. F., "Flexural and Compressive Strength of Concrete as Affected by the Properties of Concrete Aggregate", ACI J., Proceedings, Vol. 55, May 1959, pp.1193-1208.

Murdock, J. W.,and Kesler, C. E., "Effect of Stress on Plain Concrete Beams", ACI J., Proceedings, Vol. 55, No. 2 , Aug. 1959, pp.255-259.

Norby, G. M., "Fatigue of Concrete -- A Review of Research", ACI J., Proceedings, Vol. 55, No. 2, Aug. 1959, pp.191-219.

Paris, P. C., Gomez, M. P., and Anderson, W. E., " A Rational Analytic Theory of Fatigue", The Trend in Engineering, Vol. 13, No. 1, Jan. 1961.

Paris, P. C., and Erdogon, F., "A Critical Analysis of Propagation Laws", Transactions of ASME, J. of Basic Engrg., Vol. 85, 1963, pp.528-534.

Perdikaris, P. C. Calomino, A. M., and Chudnovsky, A., "Effect of Fracture on Fracture Toughness of Concrete", J. of Engrg. Mech., Vol. 112, No. 8, Aug. 1986, pp.776-791.

Perdikaris, P. C., and Calomino, A. M., "Kinetics of Crack Growth in Plain Concrete", SEM/RILEM International Conference on Fracture of Concrete and Rock, Houston, Texas, June 1987.

Stelson, T. E.,and Cerrica, J. N., " Fatigue Properties of Concrete Beams", ACI J., Proceedings, Vol. 55, No. 2, Aug. 1959, pp.255-259.

Suresh, S., Tschegg, E. K., and Blockenbrough, J. R., " Fatigue Crack Growth in Cementitious Materials under Cyclic Compressive Loads", Cement and Concrete Research, Vol. 19, No. 5, Sept. 1989, pp. 827-833.

Suresh, S., "Fatigue Crack Growth in Cementitious Solids under Cyclic Compression: Theory and Experiments", Fracture of Concrete and Rock: Recent Developments, Elsevier Applied Science, London and New York, Sept. 1989, pp.162-172.

Swartz, S. E., and Go, C. G, "Validity of Compliance Calibration to Cracked Beams in Bonding", J. of Experimental Mech., Vol. 24, No. 2, June 1984, pp.129-134.

Zhang, B., Zhu, Z. , and Wu, K., "Fatigue Rupture of Plain Concrete Analyted by Fracture Mechanics", Fracture of Concrete and Rock, SEM/RILEM International Conference, Editors, Shah, S. P., and Swartz, S. E., Houston, Texas, June 1987, pp. 58-63.

Table1 Measured Maximum Loads for Monotonic Loading

	Depth (in.)	Maximum Load P (lb.) 1	2	3	Mean P (lb.)
Present Tests.	1.5	386	413	428	408
	3.0	669	649	694	671
	6.0	1190	1152	1154	1165
Tests of Bazant and Pfeiffer, 1987.	1.5	405	408	410	408
	3.0	677	706	698	698
	6.0	990	1040	1042	1024
	12.0	1738	1739	1750	1742

Captions

Fig. 1 (a) Fatigue fracture specimens of different sizes, and (b) fracture specimens use in calibrating compliance.

Fig.2 Small (a) and medium (b) size fatigue fracture specimens, loading set-up, and LVDT gages for crack mouth opening and for deflection.

Fig. 3 Test specimen geometry.

Fig. 4 Example of calculated and measured compliance calibration curve.

Fig. 5 (a) Load-CMOD record for one specimen of medium size; (b) measured load history for the same specimen (CMOD = crack mouth opening displacement, in inches).

Fig. 6 Load-CMOD curves measured in monotonic loading of notched beam specimens of three sizes (1 in. = 25.4 mm, 1 lb. = 4.45 N).

Fig. 7 Measured and calculated growth of the relative crack length with the number of cycles, for all six specimens (two for each size).

Fig. 8 Logarithmic plots of crack length increment per cycle versus the stress intensity factor amplitude, with independent regression lines for specimens of three sizes (a in inches).

Fig. 9 Size effect law for (a) nominal strength and (b) critical stress intensity factor.

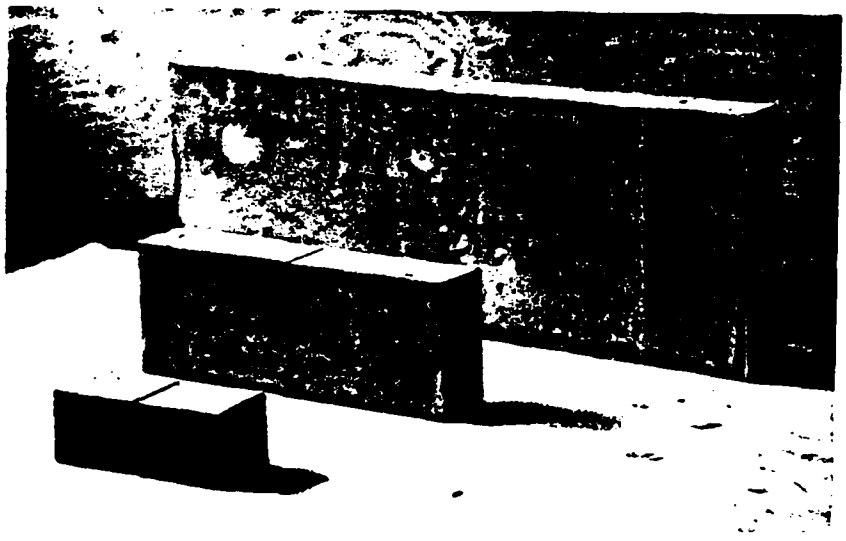
Fig. 10 Logarithmic plot of crack length increment per cycle versus the size-adjusted crack length amplitude, keeping the same d_0 as in monotonic tests (left - one, and right - two specimens of each size).

Fig. 11 Same as Fig. 10, but the transition size d_0 is ten-times larger.

Fig. 12 Logarithmic plots of crack length increment per cycle versus the amplitude of size-adjusted stress intensity factor, two specimens of each of three sizes, and optimum fit by size-adjusted Paris law (Eq. 7).

Fig. 13 Logarithmic plots of crack length increment per cycle versus nominal stress amplitude; left - without, and right - with size adjustment.

(a)



(b)

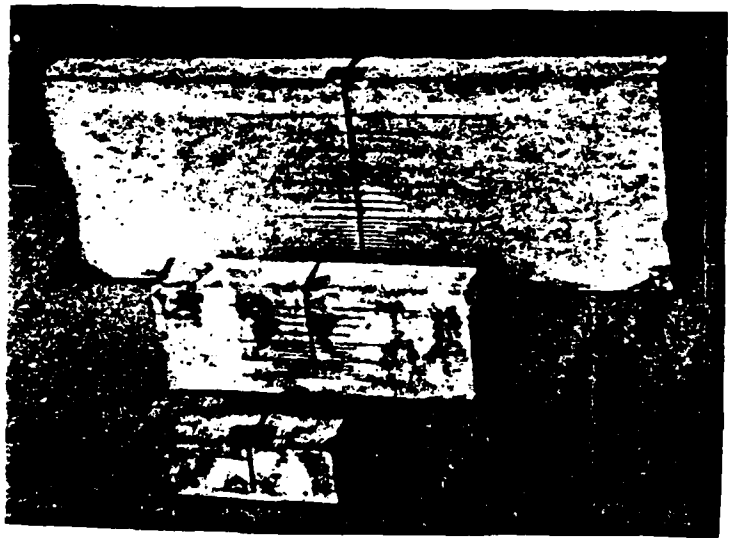
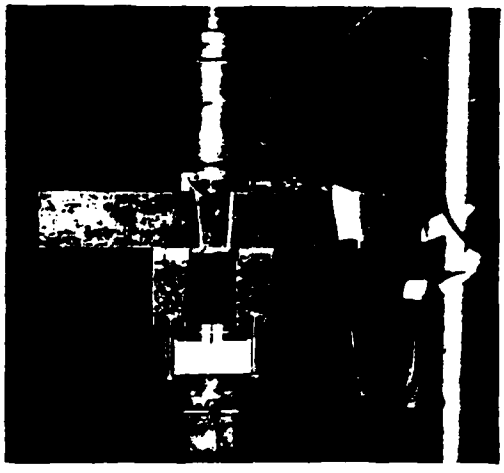


Fig. 1

(a)



(b)

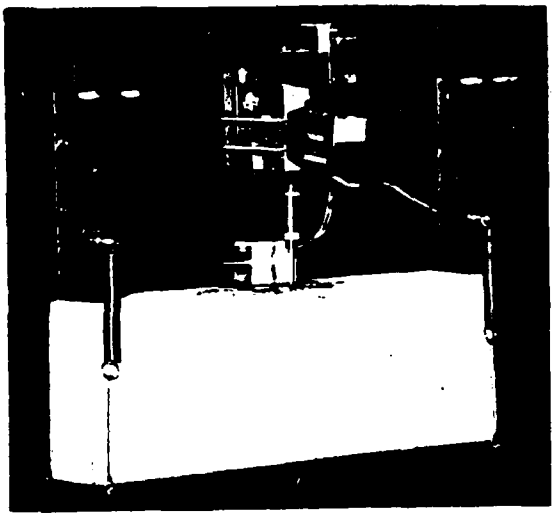
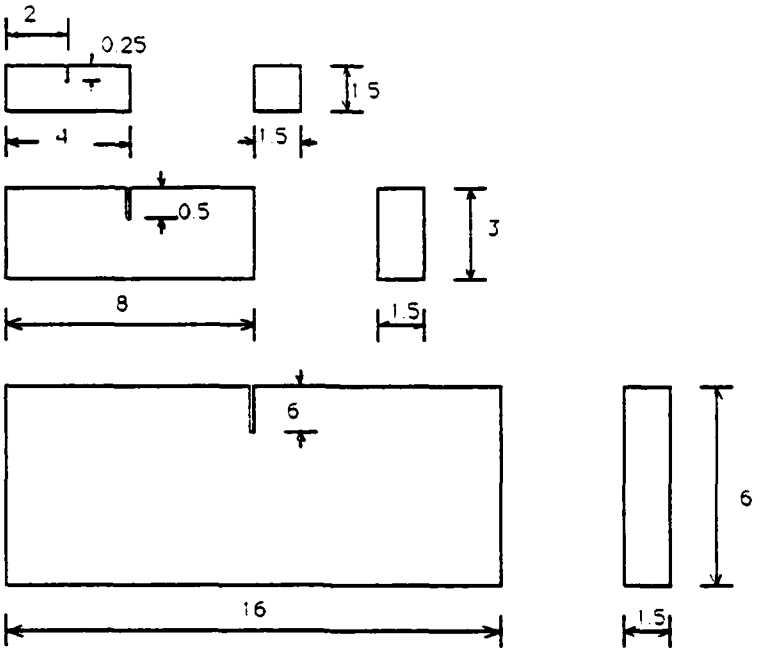
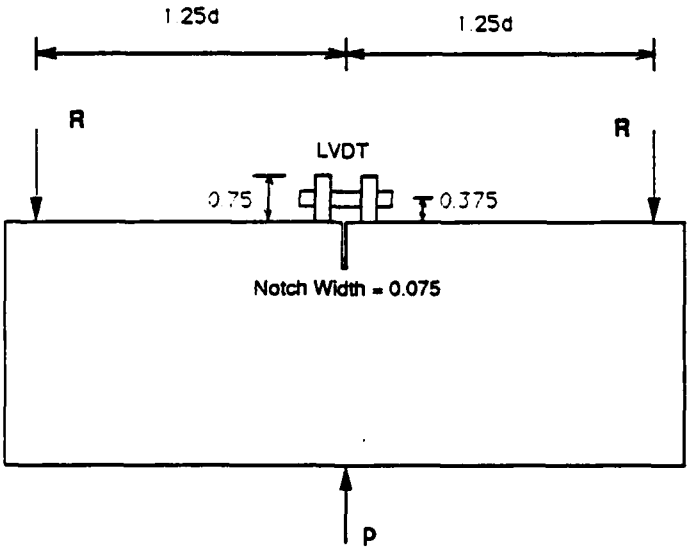


Fig. 2


2
0.25
1.5
4
1.5
0.5
3
8
1.5
6
6
16
1.5

(a)


1.25d
1.25d
R
R
LVDT
0.75
0.375
Notch Width = 0.075
P

(b)

Fig. 3

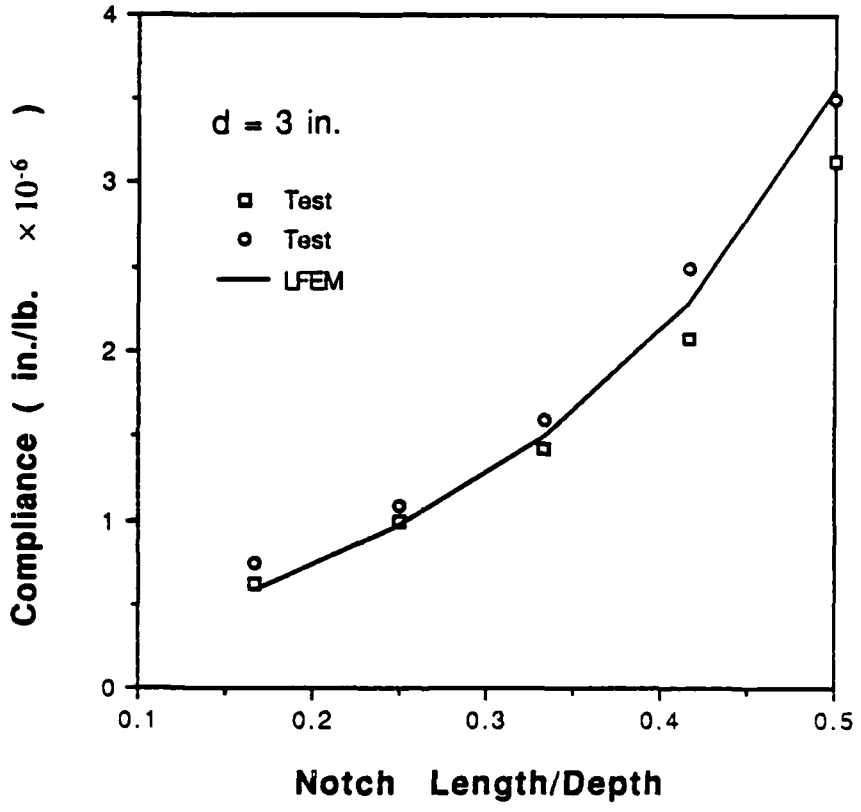
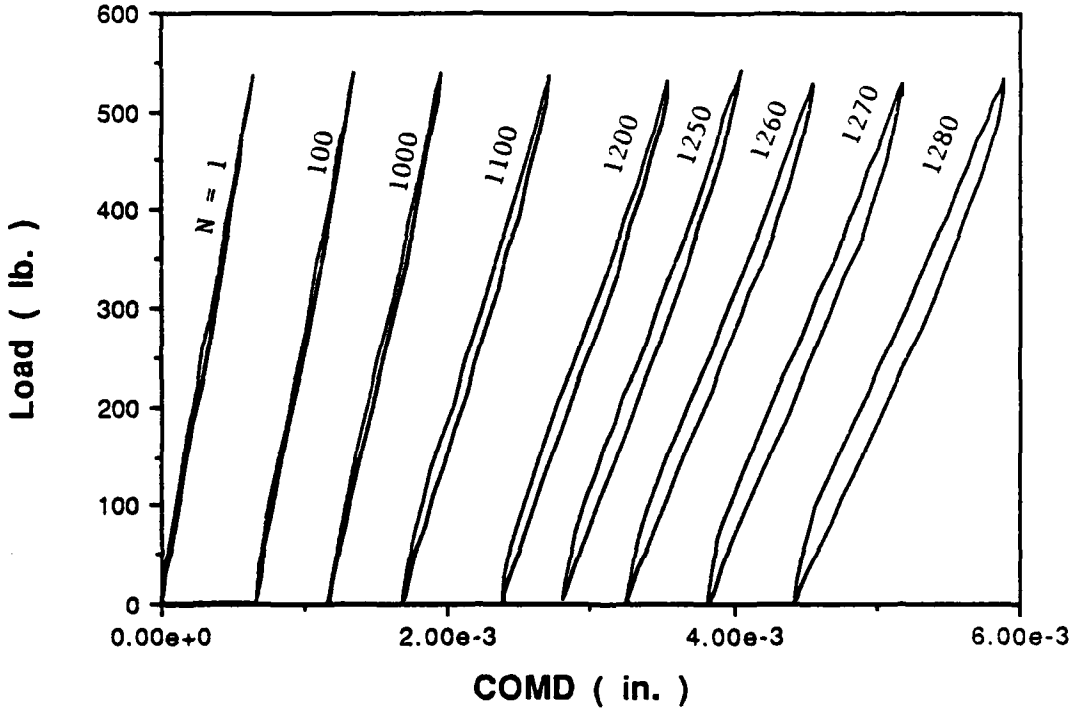
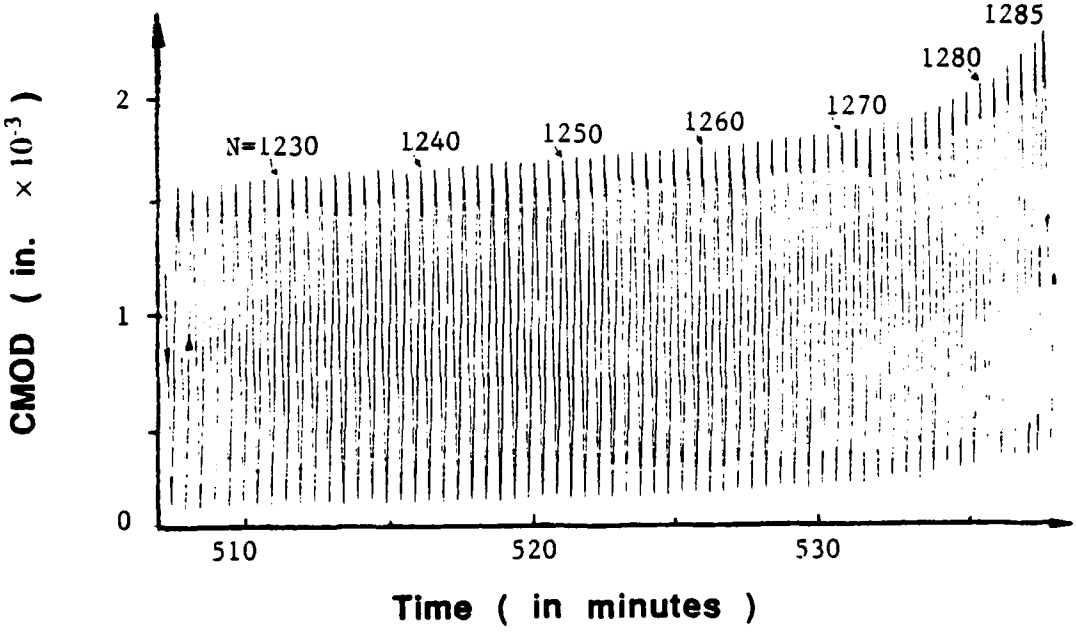

d = 3 in.
Test
Test
LFEM
Compliance (in./lb. × 10-6)
Notch Length/Depth
0
1
2
3
4
0.1
0.2
0.3
0.4
0.5

Fig 4


600
500
400
300
200
100
0
Load (lb.)
N = 1
100
1000
1100
1200
1250
1260
1270
1280
0.00e+0
2.00e-3
4.00e-3
6.00e-3
COMD (in.)

(a)


CMOD (in. × 10⁻³)
2
1
0
N=1230
1240
1250
1260
1270
1280
1285
510
520
530
Time (in minutes)

(b)

Fig. 5

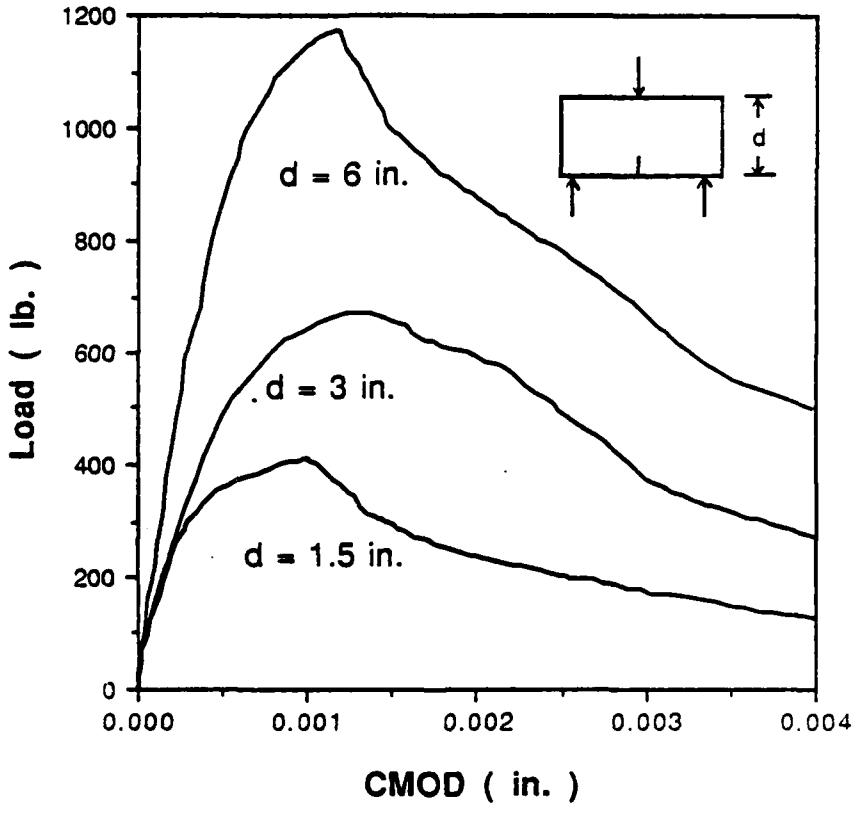

1200
1000
800
600
400
200
0
Load (lb.)
d = 6 in.
d = 3 in.
d = 1.5 in.
d
0.000
0.001
0.002
0.003
0.004
CMOD (in.)

Fig. 6

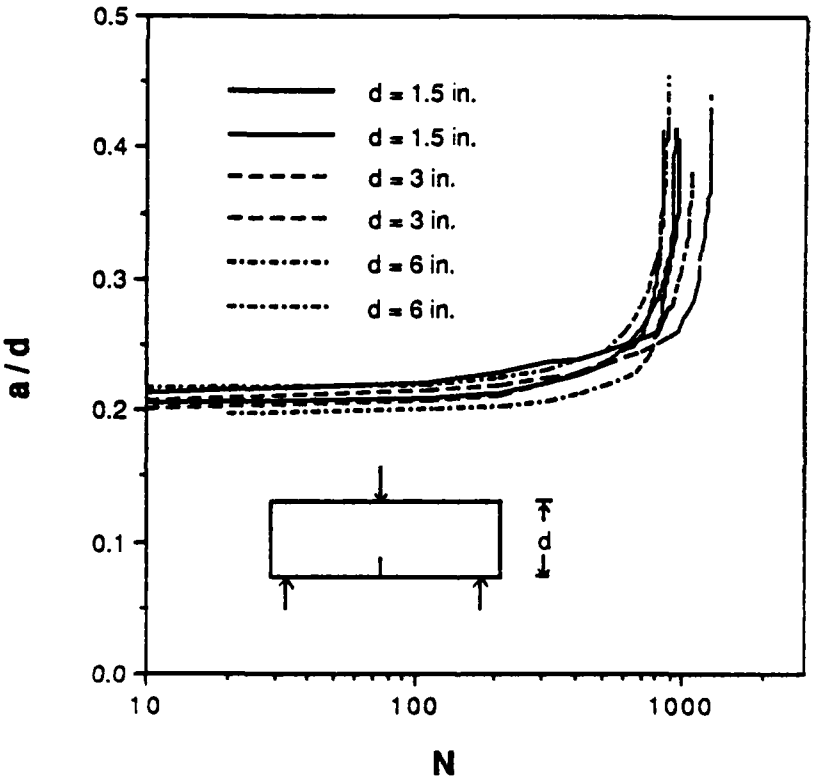
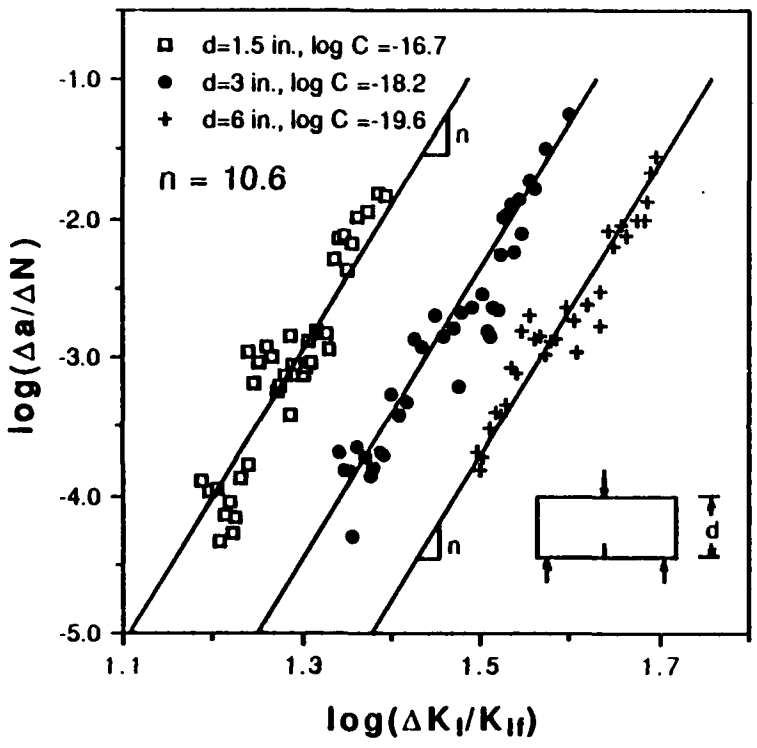
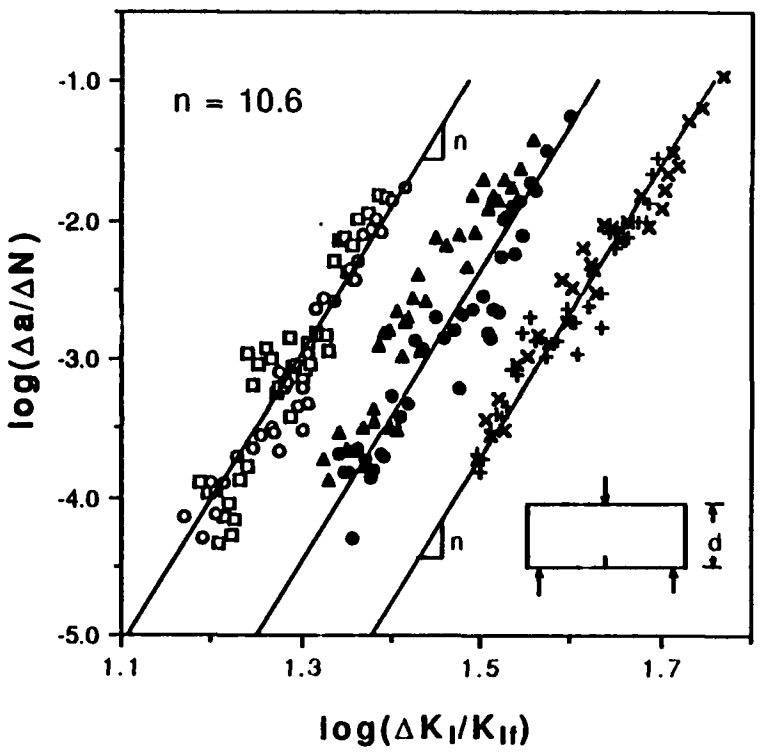

0.5
0.4
0.3
0.2
0.1
0.0
a/d
d = 1.5 in.
d = 1.5 in.
d = 3 in.
d = 3 in.
d = 6 in.
d = 6 in.
d
10
100
1000
N

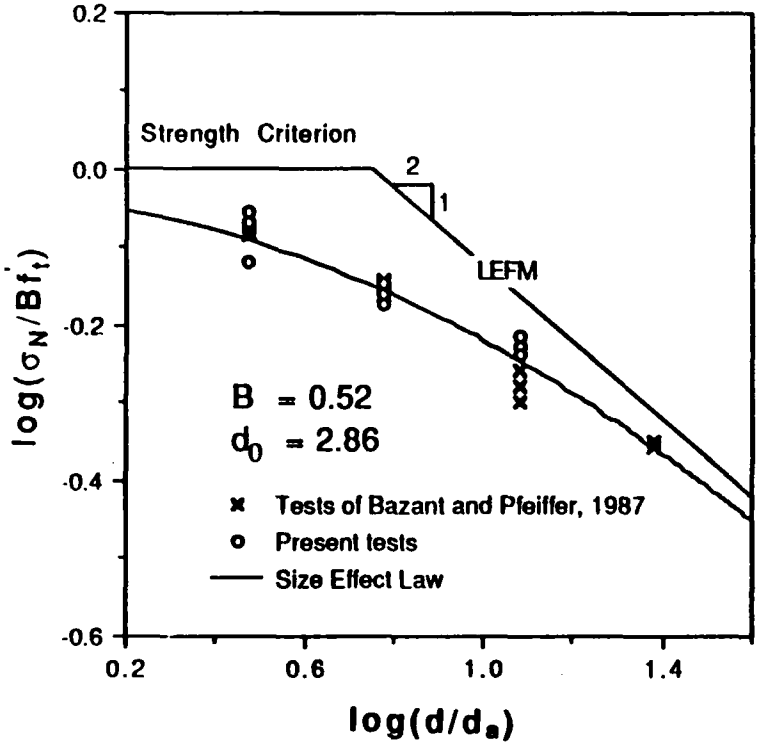
Fig. 7


d=1.5 in., log C =-16.7
d=3 in., log C =-18.2
d=6 in., log C =-19.6
n = 10.6
n
n
d
-1.0
-2.0
-3.0
-4.0
-5.0
1.1
1.3
1.5
1.7
log(Δa/ΔN)
log(ΔK_I/K_II)

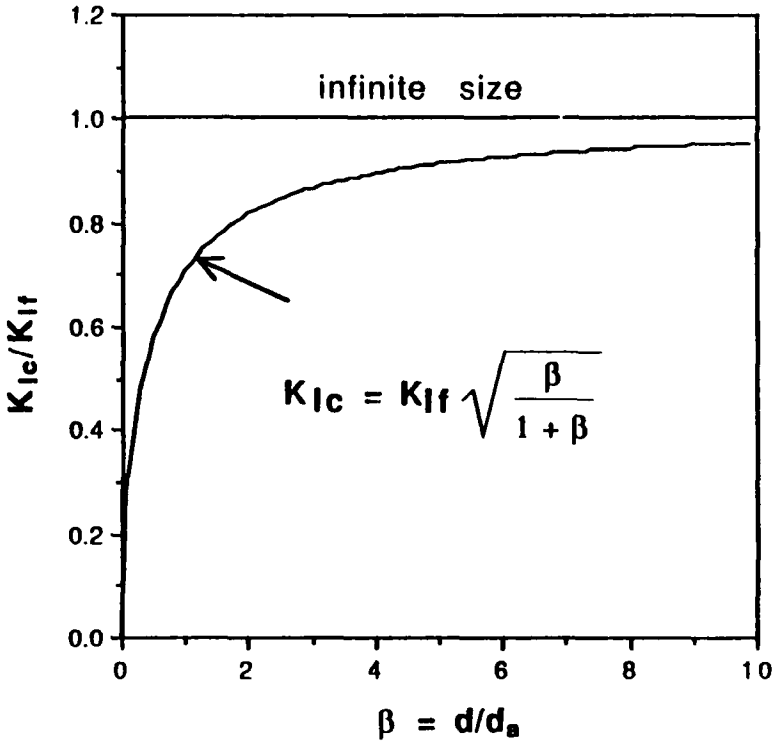
(a)


n = 10.6
n
n
d
-1.0
-2.0
-3.0
-4.0
-5.0
1.1
1.3
1.5
1.7
log(Δa/ΔN)
log(ΔK_I/K_II)

(b)

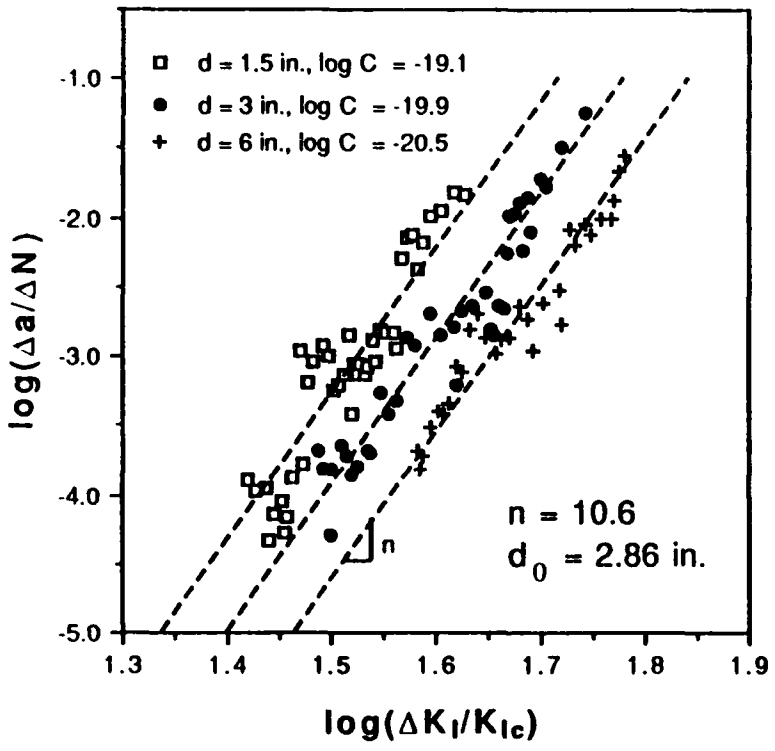

Strength Criterion
2
1
LEFM
B = 0.52
d_0 = 2.86
Tests of Bazant and Pfeiffer, 1987
Present tests
Size Effect Law
$\log(\sigma_N/Bf_t')$
$\log(d/d_a)$
0.2
0.0
-0.2
-0.4
-0.6
0.2
0.6
1.0
1.4

(a)

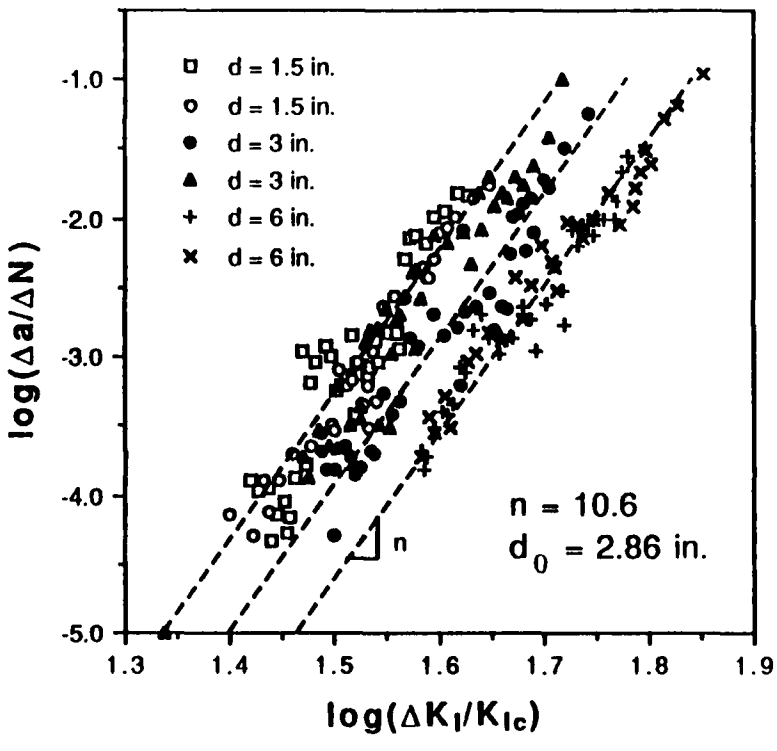

infinite size
$K_{Ic} = K_{If}\sqrt{\frac{\beta}{1+\beta}}$
K_{Ic}/K_{If}
$\beta = d/d_a$
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10

(b)

Fig 9


d = 1.5 in., log C = -19.1
d = 3 in., log C = -19.9
d = 6 in., log C = -20.5
n = 10.6
d0 = 2.86 in.
n
log(Δa/ΔN)
log(ΔKI/KIc)
-1.0
-2.0
-3.0
-4.0
-5.0
1.3
1.4
1.5
1.6
1.7
1.8
1.9

(a)


d = 1.5 in.
d = 1.5 in.
d = 3 in.
d = 3 in.
d = 6 in.
d = 6 in.
n = 10.6
d0 = 2.86 in.
n
log(Δa/ΔN)
log(ΔKI/KIc)
-1.0
-2.0
-3.0
-4.0
-5.0
1.3
1.4
1.5
1.6
1.7
1.8
1.9

(b)

Fig. 10

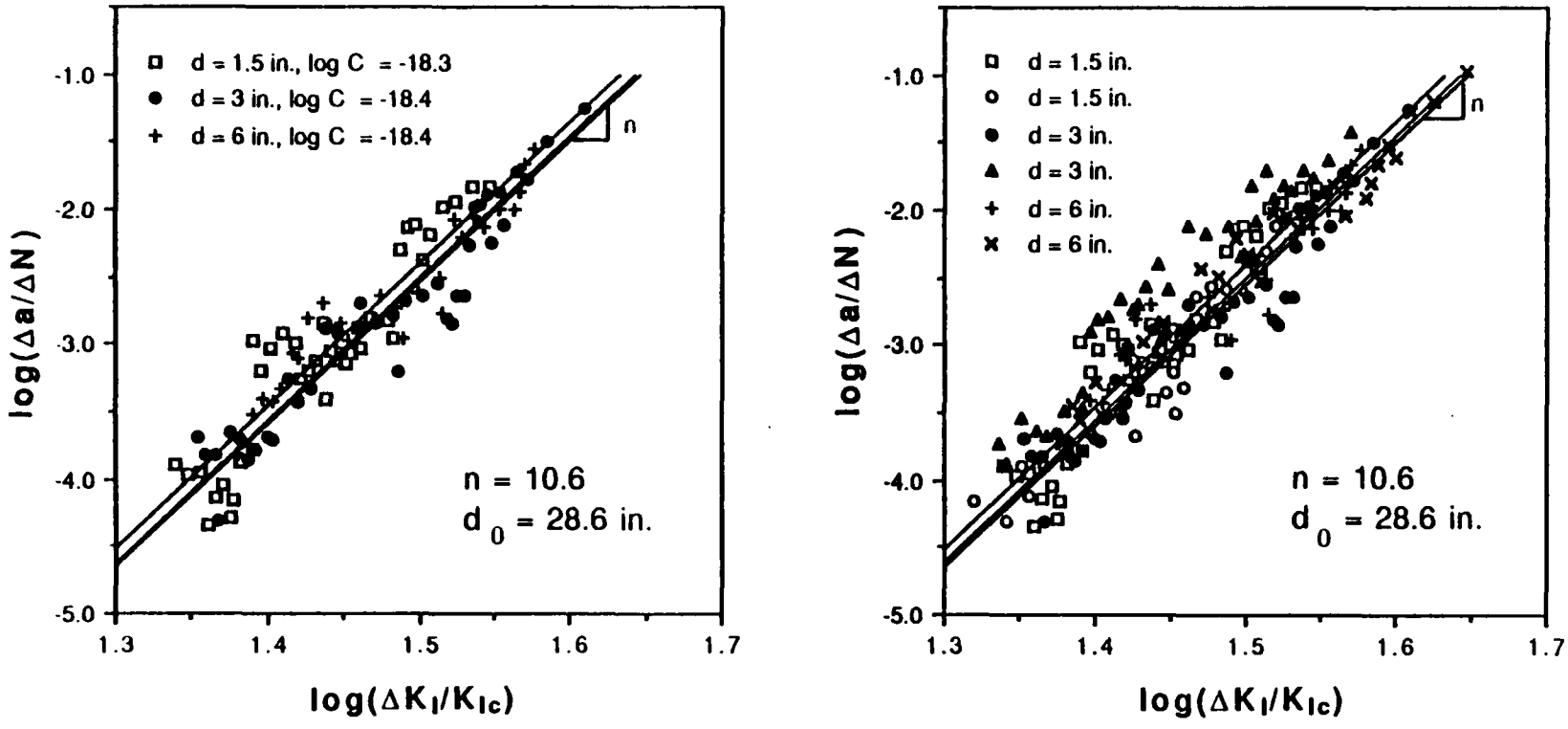

d = 1.5 in., log C = -18.3
d = 3 in., log C = -18.4
d = 6 in., log C = -18.4
n
n = 10.6
d 0 = 28.6 in.
log(Δa/ΔN)
log(ΔKI/KIc)
-1.0
-2.0
-3.0
-4.0
-5.0
1.3
1.4
1.5
1.6
1.7
(a)
d = 1.5 in.
d = 1.5 in.
d = 3 in.
d = 3 in.
d = 6 in.
d = 6 in.
n
n = 10.6
d 0 = 28.6 in.
log(Δa/ΔN)
log(ΔKI/KIc)
(b)

Fig 11

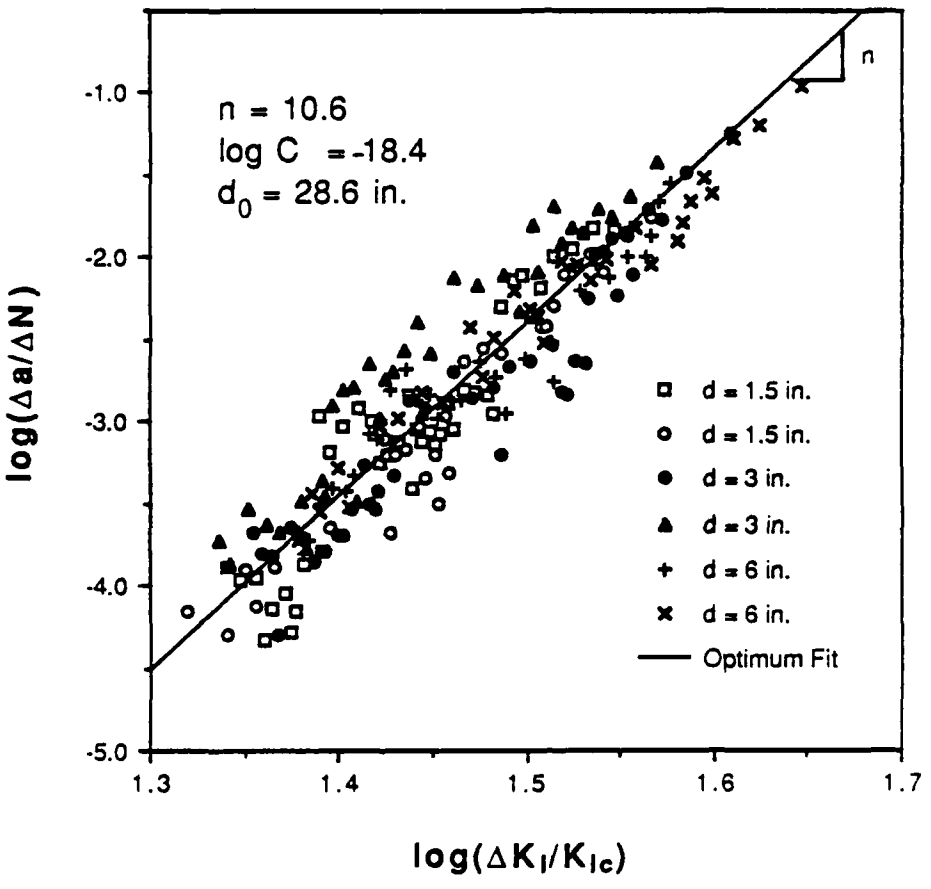
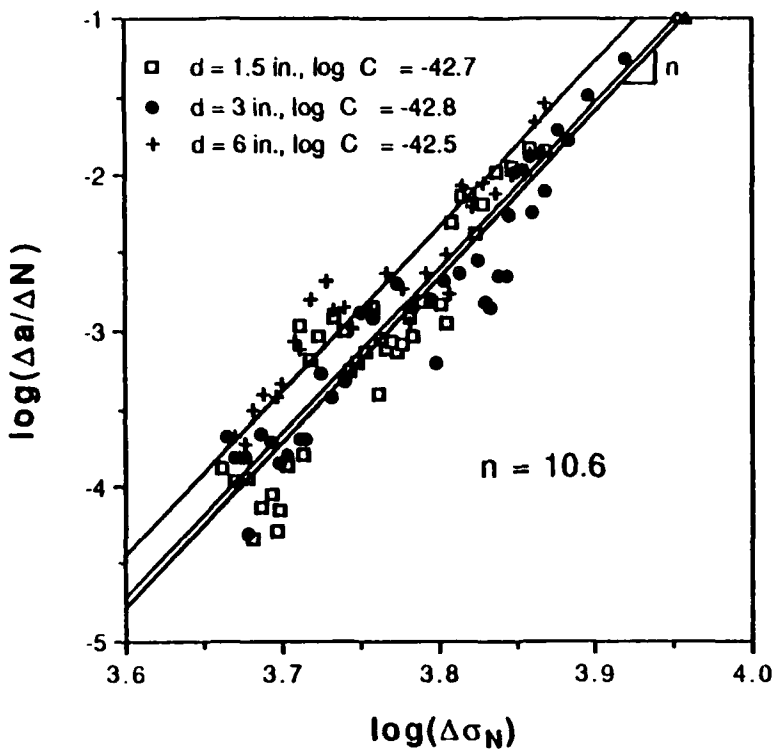
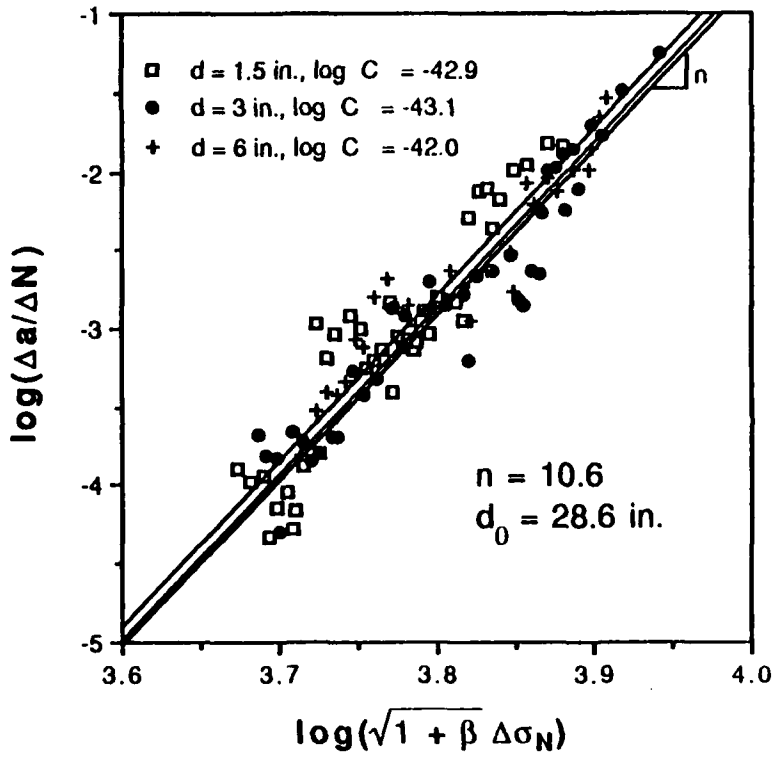

n = 10.6
log C = -18.4
d_0 = 28.6 in.
n
d = 1.5 in.
d = 1.5 in.
d = 3 in.
d = 3 in.
d = 6 in.
d = 6 in.
Optimum Fit
-1.0
-2.0
-3.0
-4.0
-5.0
1.3
1.4
1.5
1.6
1.7
log(Δa/ΔN)
log(ΔK_I/K_Ic)

Fig. 12


d = 1.5 in., log C = -42.7
d = 3 in., log C = -42.8
d = 6 in., log C = -42.5
n
n = 10.6
log(Δa/ΔN)
log(ΔσN)
-1
-2
-3
-4
-5
3.6
3.7
3.8
3.9
4.0

(a)


d = 1.5 in., log C = -42.9
d = 3 in., log C = -43.1
d = 6 in., log C = -42.0
n
n = 10.6
d0 = 28.6 in.
log(Δa/ΔN)
log(√1 + β ΔσN)
-1
-2
-3
-4
-5
3.6
3.7
3.8
3.9
4.0

(b)

Fig. 13

APPENDIX A5

Am. Concr. Inst. Materials Journal
accepted for publication, 1990

FRACTURE PROPERTIES AND BRITTLENESS OF HIGH STRENGTH CONCRETE

By

R.Gettu, Z.P.Bažant and M.E.Karr

Report No. 89-10/B627f

Center for Advanced Cement-Based Materials
McCormick School of Engineering and Applied Sciences
Northwestern University
Evanston, Illinois 60208

October 1989

MS 7056

FRACTURE PROPERTIES AND BRITTLENESS OF HIGH STRENGTH CONCRETE

By Ravindra Gettu[1], Zdeněk P. Bažant[2] and Martha E. Karr[3]

Center for Advanced Cement-Based Materials

Northwestern University, Evanston, Illinois 60208, USA

ABSTRACT

The size effect method for determining material fracture characteristics, which was previously proposed by Bažant and extensively verified for normal strength concrete, is applied to typical high strength concrete. Geometrically similar three-point bending specimens are tested and the measured peak load values are used to obtain the fracture energy, the fracture toughness, the effective length of the fracture process zone, and the effective critical crack-tip opening displacement. It is shown that the brittleness of the material can be objectively quantified through the size effect method. Comparison of the material fracture properties obtained with those of normal strength concrete shows that an increase of 160% in compressive strength causes: (1) an increase of fracture toughness by only about 25%, (2) a decrease of effective fracture process zone length by about 60%, and (3) more than doubling of the brittleness number, which may be an adverse feature that will need to be dealt with in design. The brittleness number, however, is still not high enough to permit the use of linear elastic fracture mechanics. It is also demonstrated that the R-curves derived according to the size effect law exclusively from the maximum loads of specimens of various sizes yield remarkably good predictions of the load-deflection curves.

keywords: high strength concrete, fracture, size effect, R-curve, fracture energy, fracture toughness, fracture process zone, nonlinear fracture model, brittleness

---

[1]Graduate Research Assistant

[2]Professor of Civil Engineering

[3]NSF Undergraduate Summer Research Fellow (1989)

INTRODUCTION

Concretes of strengths exceeding 80 MPa (12000 psi) are now commonly being used in the construction of high-rise buildings and offshore structures [1-5]. The utilization of concrete of such high strength has been spurred on by the superior mechanical properties of the material and the cost-effectiveness it provides [2,6]. The high strength concrete of today is a highly engineered material with several chemical and mineral admixtures. It is possible to obtain workable mixes with very low water contents by using super-plasticizers and retarders. Pozzolanic additives such as fly ash and silica fume are employed to alter the hydration reactions beneficially and also to fill the microscopic voids between cement particles. Small, round aggregates are used to achieve better mixing and attain higher surface areas for bonding. Typical high strength concrete has a very high strength matrix, is more compact and possesses well-bonded aggregate-mortar interfaces [7].

In the past, research on high strength concrete has primarily concentrated on increasing the "strength" of the material [8]. In the last decade, however, considerable effort has been spent in studying its mechanical properties and structural behavior [1-4,6,7,10-13]. Nevertheless, many aspects such as the fracture behavior need much more detailed investigation.

A complicating feature in the fracture analysis of brittle heterogeneous materials such as concrete is the nonlinear behavior. This is due to the fact that the fracture process is not concentrated at a point, the crack-tip, but is distributed over a zone whose size is not negligible when compared to the dimensions of the body. The existence of

the large fracture process zone is manifested in the size effect exhibited by concrete specimens and structures, which is considerable but not as strong as in linear elastic fracture mechanics (LEFM). The conditions for which LEFM is applicable to concrete are attained only for extremely large test specimens, whose testing would be very costly. However, material fracture properties which are unambiguous and, especially, size- and shape-independent can be defined only on the basis of a very large specimen or, more precisely, by means of extrapolation to a specimen of infinite size.

The simplest method to obtain size-independent material fracture properties is perhaps provided by extrapolation to infinite size on the basis of the size effect law proposed in Ref. 14 and 15 (see also Ref. 16). This law approximately describes the transition from the strength criterion, for which there is no size effect, to LEFM criterion, for which the size effect is the strongest possible. This law has been shown to agree well with concrete fracture tests in Mode I [17], as well as Mode II [18] and Mode III [19,20]. A good agreement was also demonstrated for certain ceramics [21], rocks [22,23], and aluminum alloys [24].

An important advantage of the size effect method is its simplicity. The method requires only the maximum load values of geometrically similar specimens which are sufficiently different in size. Such measurements can be carried out even with the most rudimentary equipment. Neither the post-peak softening response nor the true crack length need to be determined. Another advantage of the size effect method is that it yields not only the fracture energy (or fracture toughness) of the material but also the effective length of the fracture process zone, from which one can

further obtain the R-curve and the critical effective crack-tip opening displacement. Determination of the fracture process zone size is of particular interest for high strength concrete because, for various reasons, such concrete is suspected to be more brittle, and, therefore, to have a smaller fracture process zone than normal strength concrete.

In view of the aforementioned reasons, the size effect method has been adopted for the present experimental investigation of high strength concrete. The objectives will be to obtain the fracture energy and the process zone size. The third objective will be to investigate whether the R-curves derived solely from maximum load data yield load-displacement curves that agree sufficiently well with measurements.

REVIEW OF THE SIZE EFFECT LAW AND ITS IMPLICATIONS

Size Effect

Structures and test specimens of brittle heterogeneous materials, such as concrete, rock and ceramics, exhibit a pronounced size effect on their failure loads. This phenomenon, which is an important consequence of fracture mechanics, has been described by the size effect law proposed Ref. 14 and 15:

$$\sigma_N = \frac{Bf_u}{\sqrt{1+\beta}}, \qquad \beta = \frac{d}{d_0} \tag{1}$$

which applies to geometrically similar structures (or specimens) of different sizes; $\sigma_N = c_n P_u/bd$ = nominal stress at maximum load P_u (ultimate load), b = thickness of the structure or specimen, same for all the sizes (two-dimensional similarity), d = size or characteristic dimension of the structure (or specimen), c_n = coefficient introduced for

convenience, B and d_0 = parameters determined experimentally, f_u is a measure of material strength, and β is called the brittleness number. Eq. (1) also applies to structures that are similar in three-dimensions by taking $\sigma_N = c_n P_u/d^2$. In this paper, only two-dimensional similarity is considered.

Plastic limit analysis, as well as any analysis with failure criteria in terms of stresses and strains exhibits no size effect, that is, σ_N is independent of size if the structures are geometrically similar. According to Eq. (1), this behavior is obtained for very small bodies of concrete, $\beta<<1$. The application of fracture mechanics theory, however, results in σ_N being strongly size dependent. This is easily seen[25] from the LEFM relations for energy release rate G, and stress intensity factor K_I :

$$G = \frac{P^2\, g(\alpha)}{E' b^2 d}, \quad K_I = \sqrt{GE'} = \frac{P\, f(\alpha)}{b\sqrt{d}} \tag{2}$$

where P = applied load, $f(\alpha)$ is a geometry-dependent function of relative crack length $\alpha = a/d$ (a = crack length), $g(\alpha) = [f(\alpha)]^2$, $E' = E$ for plane stress, $E' = E/(1-\nu^2)$ for plane strain, E = Young's modulus of elasticity, and ν = Poisson's ratio. Values of $f(\alpha)$ can be obtained by elastic analysis techniques, such as the finite element method, and for basic specimen geometries, formulas for function $f(\alpha)$ can be found in fracture mechanics handbooks (e.g., Tada et al. [26], Murakami [27]).

When $g'(\alpha)>0$, which applies to most situations, LEFM indicates that the maximum load occurs at infinitesimal crack extensions. Therefore, α at maximum load is practically the same for bodies of different sizes. Setting $G = G_f$ (fracture energy) or $K_I = K_{Ic}$ (fracture toughness or critical stress intensity factor) along with $P = P_u$, Eq. (2) according to

LEFM yields the dependence of σ_N on size, which is σ_N = constant/$\sqrt{d}$. As is well known, LEFM criteria govern only the fracture behavior of very large concrete structures.

In usual-size concrete structures and specimens, the fracture process zone which forms in front of a propagating crack affects the behavior significantly. With increasing load this zone grows in size while remaining attached to the notch tip (provided that the specimen geometry is such that $g'(\alpha)>0$). The process zone shields the propagating crack tip and, thereby, increases the fracture resistance. The nonlinear fracture regime, where the influence of the process zone is dominant, lies between behavior governed by limit analysis and LEFM. As a result, there is considerable size effect on the failure of normal concrete structures and specimens but it is not as strong as that of LEFM. This transitional size effect is described by Eq. (1).

The size effect law (Eq. 1), giving the approximate relation of σ_N to β (the brittleness number), is plotted in Fig. 1. For large β, such as $\beta>10$, Eq. (1) gives (with an error under 5%) the approximation $\sigma_N \propto d^{-1/2}$, which is the size effect exhibited by LEFM. For small β, such as $\beta<0.1$, Eq. (1) yields (again with an error under 5%) $\sigma_N = Bf_u$ = constant, that is, failure loads are proportional to the strength of the material and there is no size effect. For $0.1<\beta<10$, the size effect is transitional between LEFM and plastic limit analysis. In this range, nonlinear fracture mechanics must be employed[17,25].

Fracture Parameters

Nonlinear fracture can be characterized by two material parameters, the fracture energy, G_f , and the effective length of the fracture process

zone, c_f . When determined from individual tests, however, these quantities are strongly size-dependent. Unambiguous definitions of G_f and c_f as fundamental material properties, independent of specimen size and shape [17,22,23,25], can be given on the basis of size effect: G_f and c_f are the energy required for crack growth and the effective (elastically equivalent) length of the fracture process zone, respectively, in an infinitely large specimen.

Mathematically, the definition of fracture energy can be stated as $G_f = \lim_{d\to\infty} G_m = \lim_{d\to\infty} (K_{Im}^2 / E')$, where G_m and K_{Im} are the apparent values of G and K_I , respectively, calculated according to LEFM (Eq. 2) for the measured peak load P_u and initial crack or notch length a_0 .

Eq. (1) can be used to determine G_f and c_f from experimental data using their foregoing definitions. A simple formula for G_f has been derived in Ref. 25:

$$G_f = \frac{B^2 f_u^2}{c_n^2 E'} d_0 g(\alpha_0) \tag{3}$$

For the effective length of the fracture process zone, the following expression has been derived in Ref. 22.

$$c_f = \frac{d_0 g(\alpha_0)}{g'(\alpha_0)} \tag{4}$$

The ratio of c_f to the true length of the process zone in concrete depends on the shape of the softening stress-displacement or stress-strain relations. For concrete, the ratio[21] seems to be about 2. However, the actual length of the process zone need not be known for the calculations.

Also, for an infinitely large specimen, the fracture toughness, K_{Ic} ,

can be obtained as:

$$K_{Ic} = \sqrt{E'G_f} = \frac{Bf_u}{c_n}\sqrt{d_0 g(\alpha_0)} \tag{5}$$

The practical applicability of Eq. (3) (or Eq. 5) has been verified in Ref. 18 through tests of normal concrete which showed that the three-point bend specimens, the double-notched tension specimens and the eccentric compression specimens yield about the same value of G_f. In this study we assume the same to be true of high strength concrete.

Due to the approximate nature of Eq. (1), the infinite size used in the definitions of G_f , c_f and K_{Ic} must not be interpreted literally. In practice, the infinite size should be assumed as a size just beyond the upper bound of the range for which the size effect law has been calibrated by tests or otherwise [25].

Using the relations $B = (E'G_f/g'(\alpha_0)c_f)^{1/2}c_n/f_u$ and $d_0 = c_f g'(\alpha_0)/g(\alpha_0)$ in Eq. (1), the size effect law can be reformulated in terms of the material fracture properties G_f and c_f :

$$\sigma_N = c_n\left(\frac{E'G_f}{g'(\alpha_0)c_f + g(\alpha_0)d}\right)^{\frac{1}{2}} \tag{6}$$

Alternatively, Eq. (6) can be put in a form convenient for the analysis of specimens which are not geometrically similar [22]:

$$\tau_N = \left(\frac{E'G_f}{c_f+\bar{d}}\right)^{\frac{1}{2}} \tag{7}$$

where $\tau_N = \sqrt{g'(\alpha_0)}\, P_u/bd$, $\bar{d} = d\, g(\alpha_0)/g'(\alpha_0)$; τ_N = intrinsic (shape-independent) nominal stress at failure, and $\bar{d}$ = intrinsic (shape-independent) size of the structure.

The brittleness number, β, (see Eq. 1) proposed by Bažant can also be expressed as [22,25]:

$$\beta = \frac{g(\alpha_0)\, d}{g'(\alpha_0) c_f} = \frac{\bar{d}}{c_f} \tag{8}$$

This number is capable of characterizing the type of failure (brittle or ductile) regardless of structure geometry[17]. It can quantify the proximity of the behavior of a structure to LEFM and is, therefore, a convenient and effective measure of the brittleness of a structure or specimen. It is also useful for comparing the behavior of specimens of different materials (with different compositions and strengths) such as high strength and regular concretes. Other brittleness numbers defined for concrete by Hillerborg [28] and Carpinteri [29,30] can compare only structures of similar geometry.

The size effect law in the form of Eq. (1), (6) or (7) has the advantage that its parameters B and d_0, or G_f and c_f, can be determined from the measured peak loads P_u by linear regression [14,15]. Eq. (1), applicable to geometrically similar specimens of different sizes, can be algebraically transformed to a linear plot:

$$Y = AX + C \tag{9}$$

in which

$$X = d,\ Y = (f_u/\sigma_N)^2,\ B = 1/\sqrt{C},\ \text{and}\ d_0 = C/A \tag{10}$$

The size range of the specimens used in the regression analysis must be sufficiently large in relation to the random scatter of material properties and test measurements. For the typical scatter of concrete, the minimum size range recommended is 1:4.

The size effect method can also be used to determine fracture

properties other than those already defined. Considering the infinitely large specimen again, material parameters such as the effective (elastically equivalent) crack-tip opening displacement at the peak load, δ_{ef} , can be related to c_f and K_{Ic} using LEFM relations [23]:

$$\delta_{ef} = \frac{8K_{Ic}}{E'}\sqrt{\frac{c_f}{2\pi}} \tag{11}$$

The value of δ_{ef} pertains to the effective opening when the stress at the notch (or initial crack) tip is just reduced to zero, and the fracture process zone is about to detach and advance away from this tip.

The size of the fracture process zone can also be characterized by the length parameter

$$\ell_0 = \left(\frac{K_{Ic}}{f_u}\right)^2 \tag{12}$$

first used by Irwin [31] for the size of the yielding zone in ductile fracture. For concrete, this length parameter was introduced by Hillerborg et al. [32]. The parameter ℓ_0 is a size-independent measure of the intensity of the toughening or crack-tip shielding mechanisms. It can be used to compare the brittleness of one material with another. A material with a relatively low ℓ_0 would be more brittle than a material with a higher ℓ_0 .

R-Curves

Another important consequence of the fracture process zone is the fracture resistance curve or R-curve. Since the fracture process zone evolves as it propagates, the resistance R(c) to fracture growth, representing energy dissipated per unit length and width of fracture extension, gradually increases. The function R(c), called the R-curve, was initially proposed to be used as a material property [33,34], independent of

the shape of the specimen or structure. It was, however, shown[35,36] that the R-curves for concrete strongly depend on the specimen geometry. A general derivation of the R-curve equations from the size effect law as presented in Ref. 22, yields the expression:

$$R(c) = G_f \frac{g'(\gamma)}{g'(\alpha_0)} \frac{c}{c_f} \tag{13}$$

in which γ is given by:

$$\frac{c}{c_f} = \frac{g'(\alpha_0)}{g(\alpha_0)} \left(\frac{g(\gamma)}{g'(\gamma)} - \gamma + \alpha_0 \right) \tag{14}$$

The foregoing relations (Eqs. 13 and 14) are valid as long as the fracture process zone remains attached to the tip of the initial crack or notch. This ceases to be true after the peak load[23]. Since the fracture process zone in the post-peak regime is detached from the tip (separated from the initial tip by a traction-free crack) and is advancing ahead with approximately constant size, it dissipates roughly the same amount of energy per unit crack extension. Consequently, the critical value of G after the peak load must be kept constant and equal to the value of R reached at the peak load. The effective R-curves given by Eqs. (13) and (14), are size-independent until the peak load but deviate afterwards into size-dependent horizontal branches. The size-dependence of R-curves in the post-peak regime has been previously observed for concrete [37] and other materials [23,24,38]. These effective R-curves tend to move toward complete coincidence with the actual R-curve as the size increases [23].

EXPERIMENTAL INVESTIGATION

Test Details

High strength concrete used in the construction industry varies widely in composition and strength. Even the demarcation between high and normal strength concretes is subjective and appears to be continually rising [2,8]. In order to study typical high strength concrete used in the industry, this investigation was conducted on material obtained directly from a batch mixed for the construction of a high-rise building in downtown Chicago. The concrete was provided by Material Services Corporation of Chicago, Illinois.

The concrete mix was designed to exceed a 28-day compressive strength of 83 MPa (12000 psi). Cement of much higher quality than standard (ASTM C 150) Portland cement was selected. Silica fume (micro-silica) and fly ash were used as mineral admixtures. The maximum aggregate size (d_a) in the mix was 9.5 mm (3/8 in.). The details of the mix composition are given in Table 1. The proportions of cement : sand : gravel : water : fly ash : micro-silica, by weight, were 1 : 1.35 : 2.42 : 0.35 : 0.25 : 0.04. The fresh concrete had a unit weight of 2410 kg/m^3 (151 lb/ft^3) and a slump of 235 mm (9.3 in.) as measured with a standard 305 mm (12 in.) cone.

Along with the fracture specimens, three 152 mm x 304 mm (6 in. x 12 in.) cylinders and three 102 mm x 204 mm (4 in. x 8 in.) cylinders were cast. These cylinders were capped with a sulphur compound and cured in a fog room. They were tested in compression after 14 days. The larger cylinders failed at an average maximum compressive stress of 85.5 MPa (12400 psi), with a standard deviation of 1.4%. This value, denoted as f'_{c14}, was taken to be the mean compressive strength of the

concrete. The smaller cylinders had an average compressive strength of 84 MPa (12200 psi), with a standard deviation of 0.4%. The 28-day compressive strength of the concrete, f'_c, as measured by testing two 152 mm × 304 mm (6 in. × 12 in.) cylinders, was 96 MPa (13950 psi).

Beam specimens of four different sizes, three in each size, were cast from the same batch of concrete. All specimens were compacted by rodding and stored in the molds for 24 hours. They were then demolded and cured under water until the tests. The specimens (Figs. 2 and 3) were 38 mm (1.5 in.) thick and their depths were 305, 152, 76, and 38 mm (12, 6, 3, and 1.5 in.). Their lengths were 8/3 times their depths and their spans 2.5 times the depths. Before testing, a notch 2 mm (0.08 in.) wide was cut at the midspan of each beam using a diamond band saw. The length of the notch was one-third the depth of the beam.

The three-point bending test was used due to the relatively simple test setup and the impossibility of crack bifurcation [22]. The peak loads of the largest specimens were measured by loading them under stroke control in a 534 kN (120 kip) load frame with an MTS control system. The other specimens were tested in a smaller load frame of 89 kN (20 kip) capacity, with a load cell operating in the 8.9 kN (2000 lb) range, and crack mouth opening displacement (CMOD) control with an MTS closed-loop control system was used. A Schaevitz linear variable differential transformer (LVDT) with a range of 0.25 mm (0.01 in.) was mounted across the notch of each beam to monitor the CMOD. The measured CMOD was used as feedback to control the tests at a constant CMOD rate. This type of loading stabilized crack propagation even after the peak load and thus made it possible to obtain complete load-deflection curves for the

specimens. The load-point displacements were measured between the tension faces of the beams and the cross-head of the loading ram with a Schaevitz LVDT of 1.27 mm (0.05 in.) range. The load-displacement and load-CMOD curves were continuously recorded for each beam. The test setup is shown in Figs. 4 and 5. All the beams were tested 14 days after casting and with CMOD rates such that they reached their peak loads in about 10 minutes. Typical curves for different specimen sizes are shown in Figs. 6 and 7, and the raw data are given in Table 2.

The test results from the largest beams were, however, excluded from the analysis which follows because their load values were inconsistent with the trend of the other tests. In retrospect, this inconsistency may be due to the difference in control parameters used for testing and in machine characteristics. Under stroke control, which was used only for the largest beams, it was not possible to obtain a stable response at the peak load. The instability of crack propagation may have caused a high crack propagation rate near the peak load. This might explain why the peak loads of these specimens were higher than expected.

The modulus of rupture of the material was used as the measure of its strength. Four specimens, each with a depth of 51 mm (2 in.) and a span of 152 mm (6 in.), were cut from two of the 152 mm (6 in.) deep fracture specimens which had already been tested. These unnotched beams were tested to failure in three-point bending under stroke control in the 89 kN (20 kip) load frame. From the peak loads the modulus of rupture of each specimen was calculated. The average value (f_u) was found to be 11.0 MPa (1600 psi), with a standard deviation of 5%. Carrasquillo, Nilson, and Slate [7] proposed an empirical formula for the modulus of rupture of high

strength concrete: $f_r = 0.94 \sqrt{f'_c}$ (stresses in MPa). Using the value of f'_{c14} in their relation, f_r was obtained as 8.7 MPa (1260 psi). The difference in the values of f_u and f_r may partly be due to errors in the predictions of the formula (known to be up to 15%) but it might also be due to the significant size effect which is known to exist in flexural tests on plain concrete beams [15,39].

As noted previously by several researchers[7,9,10,12], cracks in high strength concrete were seen to propagate through the coarse aggregates. This is an important difference from normal strength concrete in which cracks (in the vicinity of gravel particles) propagate mainly along the aggregate-mortar interfaces. The reduced crack deflection by the aggregates is due to the strong aggregate-mortar bond and the fact that the strength of the matrix approaches that of the aggregates. Obviously this near-homogeneous behavior decreases the width of the process zone as well as the shielding effect of the aggregates.

Analysis

Linear regression analysis (Fig. 8), using Eqs. (9) and (10), yields coefficients A = 1.58 mm^{-1} (40.1 $in.^{-1}$) and C = 51.7. The coefficient of variation of the deviations from the regression line is $\omega = s_{Y|X} / \bar{Y} = 15\%$, where $s_{Y|X}$ is the standard error of the Y estimate and $\bar{Y}$ is the mean of the experimental Y-values. From Eq. (10), the parameters of Eq. (1) are computed as B = 0.139 and d_0 = 32.8 mm (1.29 in.). Nonlinear regression analysis directly for B and d_0, using the Marquardt-Levenberg algorithm, results in B = 0.198 and d_0 = 13.3 mm (0.52 in.) with ω = 8.4%. Since the nonlinear regression provides a better fit, the latter values for the size effect parameters are used henceforth.

Fig. 9 shows the size effect curve (Eq. 1) based on the calculated parameters along with the experimental values of normalized nominal stress. It is seen that the trend of the data is fitted very well by the size effect law. For comparison, data from similar tests of normal strength concrete (particulars given later) are also shown. It is now important to observe that the data points of the high strength concrete beams lie closer to the LEFM criterion than those of usual concrete. This indicates that the same specimen made of high strength concrete behaves in a more brittle manner than that made of regular concrete. Yet the behavior of the present specimens cannot be described by LEFM. For LEFM relations to be applicable with errors less than 2%, the brittleness number β would have to be greater than 25, that is the beam depth would have to exceed 334 mm (13 in.).

For the specimen geometry used, the function $f(\alpha) = 6.647\sqrt{\alpha}(1-2.5\alpha+4.49\alpha^2-3.98\alpha^3+1.33\alpha^4)/(1-\alpha)^{3/2}$ was obtained by fitting the results of linear elastic finite element analysis. Using $g(\alpha_0) = 14.20$ and $g'(\alpha_0) = 72.71$ in Eq. (5), K_{Ic} is 30.0 MPa$\sqrt{mm}$ (864 psi $\sqrt{in.}$), and the value of c_f, from Eq. (4), is 2.6 mm (0.10 in.). This yields $c_f \simeq 0.3d_a$. From Eq. (12), ℓ_0 for the concrete tested is 7.4 mm (0.29 in.) which is about $0.8d_a$. Young's modulus of elasticity (E) may be computed from the formula of Carrasquillo et al. [7], $E = 3320\sqrt{f'_c} + 6900$ (in MPa), which yields $E \simeq 37.6$ GPa (5.45×10^6 psi). The fracture energy G_f ($= K_{Ic}^2/E$) is found to be 24.0 N/m (0.137 lbs/in.). The critical effective crack-tip opening can be computed by substituting the values of the basic fracture properties in Eq. (11); $\delta_{ef} = 4.11\times10^{-3}$ mm (1.60×10^{-4} in.).

For comparison, it is interesting to consider the results of similar

tests on normal strength concrete with d_a = 13 mm (0.5 in.). The relative proportions of cement : sand : gravel : water, by weight, were 1 : 2 : 2 : 0.6. The specimens and loading setup were identical to the present ones except that the notch lengths were one-sixth of the beam depth. The curing conditions were the same as those of the high strength specimens. The tests were conducted 28 days after casting and, therefore, the parameters corresponding to 14 days had to be estimated to this effect, the following relations were used: $f_r = 0.62\sqrt{f'_c}$ (ACI); $f'_t = 0.50\sqrt{f'_c}$ (ACI); $E = 4735\sqrt{f'_c}$ (ACI); $f'_c(t) = f'_c(28)\, t/(4+0.85t)$ (ACI); and $G_f \simeq (2.72+3.103f'_t)f'^2_t d_a/E$ where f'_c, f'_t, and E are in MPa, d_a is in mm, G_f is in N/mm and t is the age in days[40]. The material properties estimated for the age of 14 days are: f'_c = 32.5 MPa (4700 psi), E = 27.0 GPa (3.91×10^6 psi), f_r = 3.5 MPa (510 psi), K_{Ic} = 24.0 MPa$\sqrt{mm}$ (690.1 psi$\sqrt{in.}$), c_f = 6.6 mm (0.26 in.), G_f = 21.4 N/m (0.122 lbs/in), δ_{ef} = 7.29×10^{-3}mm (2.87×10^{-4}in.), and ℓ_0 = 46 mm (1.80 in.). The values of c_f and ℓ_0 are about $0.5d_a$ and $3.6d_a$, respectively. A graphical comparison of the properties of the high strength concrete and the regular concrete is presented in Fig. 10.

As concluded by recent investigators of high strength concrete fracture [11,12,41], G_f and K_{Ic} increase with strength but much less than the material strength itself. For an increase in strength (f'_c) of about 160%, G_f and K_{Ic} increase only by 12% and 25%, respectively. More significantly, the values of c_f and ℓ_0, especially the latter, decrease considerably with increase in strength. This implies that the size of the process zone must be smaller in high strength concrete than in regular concrete, and the crack-tip shielding by the fracture process zone must be

weaker. Consequently, the same structure made of high strength concrete is more brittle than that made of regular concrete. The brittleness number of any structure is more than doubled since it is inversely proportional to the value of c_f .

Conventional fracture analysis of the type applied to metals, based only on K_{Ic} or G_f , would yield the misleading conclusion that the ductility of concrete increases with strength; see e.g. de Larrard et al. [41]. This was also pointed out by John and Shah [11] on the basis of tests on high strength mortar.

According to RILEM recommendation [42], the fracture energy, denoted as G_f^R , is defined by the work-of-fracture method introduced for ceramics by Nakayama [43] and Tattersall and Tappin [44], and proposed for concrete by Hillerborg [28,45]. In this method, G_f^R can be determined from the area, W_0 , under the load-deflection curve of a fracture specimen. The total energy dissipated in the test is given by $W = W_0 + mg\, u_f$ where mg = weight of the specimen and u_f = deflection when the beam fractures completely. The RILEM fracture energy is assumed to be the average dissipated energy per unit cracked surface area; therefore, $G_f^R = W/(d-a_0)b$ where d = depth of the beam and a_0 = initial notch length. For practical reasons, u_f was taken to be the deflection when load had dropped to 10% of the measured peak load. The corresponding fracture toughness can be obtained as $K_{Ic}^R = \sqrt{G_f^R E_0}$, where E_0 is the Young's modulus corresponding to the initial compliance. The values of K_{Ic}^R/K_{Ic} for different sizes have been calculated from the present results and plotted in Fig. 11. It is obvious that the fracture toughness obtained by the RILEM method is strongly size-dependent and prone to considerable scattering. It also appears that for very large

sizes the values of K_{Ic}^{R} tend toward K_{Ic}, i.e., the fracture toughness from the size effect law. A similar result could be shown in terms of the fracture energy.

PREDICTION OF STRUCTURAL RESPONSE FROM R-CURVES

An effective, simple and reasonably accurate method of predicting the load-deflection curves is to identify the size effect law from the maximum load data, determine from it the geometry-dependent R-curve, and then utilize the R-curve along with LEFM relations[23]. This method is now used to predict the load-deflection curves of the high strength concrete beams. The predictions are then compared with the measured load-deflection curves.

The derivation of the equations used in the method is briefly outlined for convenience. Let u_c = load-point displacement of a specimen due to fracture, u_0 = displacement calculated as if there were no crack, and $u = u_c + u_0$ = the total displacement. W_P denotes the total energy that would be released if the fracture occurred at constant load P. Since $\partial W_P/\partial a = bG = P^2 g(\alpha)/E'bd$, we have:

$$W_P = b \int G(a)\, da = \frac{P^2}{E'b} \int_0^{\alpha} g(\alpha')\, d\alpha' \tag{15}$$

According to Castigliano's second theorem,

$$u_c = \frac{\partial W_P}{\partial P} = \frac{2P}{E'b} \int_0^{\alpha} g(\alpha')\, d\alpha' \tag{16}$$

At the same time, for G = R:

$$P = b \sqrt{\frac{E'd}{g(\alpha)} R(c)} = \frac{b\sqrt{d}}{f(\alpha)} K_{IR}(c) \tag{17}$$

where $c = (\alpha-\alpha_0)d$ = effective crack extension, and $K_{IR}(c) = \sqrt{E'R(c)}$ = effective fracture resistance. Thus, P and u_c can be obtained from Eqs. (16) and (17), for different chosen values of α or c, yielding the load-deflection curve of the structure or specimen with a propagating crack. Before the peak load, one uses the R-curve values derived from the size effect law (Eqs. 14 and 15), but after the peak load, as mentioned earlier, the R-curve, as well as $K_{IR}(c)$ is constant.

To obtain the total deflections, the elastic deflections u_b and u_s due to bending and shear must also be calculated. Assuming plane stress conditions and adopting the bending theory, the load-point displacement u_0 of the beam without crack is

$$u_0 = u_b + u_s , \quad u_b = \frac{PL^3}{4bd^3E} , \quad u_s = 0.6(1+\nu)\frac{PL}{bdE} \tag{18}$$

where L is the span of the specimen, and u_b and u_s are the contributions of bending and shear, respectively. The weight of the beam is approximately taken into account in the calculations by replacing it with a concentrated midspan load which produces an equivalent midspan moment.

Figures 12(a-c) compare the predictions of the foregoing method for the present high strength concrete beams with the measured load-deflection curves for specimens of three different sizes. The Young's modulus, in each case, was taken as that corresponding to the initial compliance of the specimen, and the Poisson's ratio was assumed to be 0.20. It can be seen that the R-curve method predicts the specimen response quite well. The R-curve derived from the size effect law is thus shown to be a useful tool for modeling structural behavior with ease and simplicity.

CONCLUSIONS

1. The size effect method provides reliable fracture properties for high strength concrete with a very simple experimental setup. These properties, namely, the fracture energy (or fracture toughness) and the effective length of the fracture process zone (as well as the effective critical crack-tip opening displacement derived from them) are size-independent and, therefore, true material properties.

2. As the strength increases, the fracture energy and the fracture toughness of concrete also increase, although much less than the strength. The effective size of the fracture process zone, however, diminishes. This results in decreased crack-tip shielding and, consequently, increase in the brittleness of specimen or structure. Increasing the compressive strength by 160% causes the brittleness number to more than double. This property can be disadvantageous and needs to be addressed in design.

3. Despite the aforementioned increase in brittleness, the behavior of contemporary high strength concrete, in the normal size range, is still governed by nonlinear fracture mechanics rather than linear elastic fracture mechanics.

4. The R-curves derived according to the size effect law solely from maximum loads can be utilized to provide reasonably accurate predictions of the load-deflection curves of concrete structures, even in the post-peak regime. The procedure requires only: (1) knowledge of the linear elastic fracture mechanics relations for the geometry of the structure and (2) the values of two basic material fracture parameters obtained from maximum loads through the size effect method.

ACKNOWLEDGEMENTS

This work was partially supported by the Center for Advanced Cement-Based Materials at Northwestern University (NSF Grant DMR-8808432). The underlying basic theory was developed under partial support from AFOSR contract 49620-87-C-0030DEF with Northwestern University. Summer support for M.E.Karr was provided under NSF Grant BCS-8818230. The authors are grateful to Jaime Moreno and Arthur King of Material Services Corporation, Chicago, Illinois, for donating the concrete used in the experiments.

NOTATION

a	effective crack length
b	specimen thickness
c	elastically equivalent crack growth
c_f	effective length of fracture process zone
c_n	coefficient for convenience (here, equal to 1)
d	characteristic specimen dimension (here, specimen depth)
d_a	maximum aggregate size
d_0	parameter determined experimentally
f'_c	28-day compressive strength
f'_{c14}	14-day compressive strength
f_u	some measure of material strength
$g(\alpha)$	function dependent on specimen shape
$g'(\alpha)$	derivative of $g(\alpha)$ with respect to α
ℓ_0	parameter related to the size of the process zone
A	slope of linear regression plot
B	parameter determined experimentally

C	intercept of linear regression plot
E	Young's modulus
E'	E, for plane stress; $E/(1-\nu^2)$, for plane strain
G	strain energy release rate
G_f	fracture energy
K_I	stress intensity factor
K_{Ic}	fracture toughness
P_u	maximum or peak load
α	relative crack length
α_0	initial relative crack or notch length
β	brittleness number
δ_{ef}	effective critical crack tip opening
ν	Poisson's ratio
σ_N	nominal stress at maximum load

REFERENCES

1. Shah, S.P., ed., *High Strength Concrete*, Proc. of Workshop, sponsored by National Science Foundation (Chicago), 1979, p. 226.

2. ACI Committee 363, "State-of-the-Art Report on High-Strength Concrete", *ACI J.*, July-Aug. 1984, pp. 364-411.

3. Russell, H.G., ed., *High-Strength Concrete*, SP-87, ACI, Detroit, 1985.

4. Holand, I., Helland, S., Jakobsen, B., and Lenschow, R., eds., *Utilization of High Strength Concrete*, Proc., Symp. in Stavenger,

Norway, Publ. by Tapir, Trondheim, 1987, p. 688.

5. Godfrey, K.A., Jr., "Concrete Strength Record Jumps 36%", *Civil Engng.*, ASCE, Vol. 57, No. 10, Oct. 1987, pp. 84-88.

6. Ahmad, S.H., and Shah, S.P, "Structural Properties of High Strength Concrete and Its Implications for Precast Prestressed Concrete", *PCI J.*, Vol. 30, No. 6, Nov.-Dec. 1985, pp. 92-119.

7. Carrasquillo, R.L., Nilson, A.H., and Slate, F.O., "Properties of High Strength Concrete Subject to Short-Term Loads", *ACI J.*, May-June 1981, pp. 171-178.

8. Carrasquillo, R.L., Nilson, A.H., and Slate. F.O., "High-Strength Concrete: An Annotated Bibliography 1930-1979", *Cem., Concr., and Aggregates,* ASTM, Vol. 2, No. 1, Summer 1980, pp. 3-19.

9. Bažant, Z.P., "Material Behavior Under Various Types of Loading", Discussion, see Shah, S.P., ed., 1979, pp. 79-92.

10. Carrasquillo, R.L., Slate, F.O., and Nilson, A.H., "Microcracking and Behavior of High Strength Concrete Subject to Short-Term Loading", *ACI J.*, May-June 1981, pp. 179-186.

11. John, R., and Shah, S.P., "Effect of High Strength Concrete and Rate of Loading on Fracture Parameters of Concrete", *Preprints, SEM/RILEM Int. Conf. on Fracture of Concrete and Rock,* (Houston) eds. Shah, S.P., and Swartz, S.E., Soc. for Exp. Mech., June 1987; also "Fracture Mechanics Analysis of High Strength Concrete", *J. Materials in Civil Engng.*, ASCE, Vol. 1, No. 4, 1989, pp. 185-198.

12. Tognon, G., and Cangiano, S., "Fracture Behaviour of High-Strength and Very High Strength Concretes", *Preprints of the Proc., Int. Workshop on Fracture Toughness and Fracture Energy - Test Methods for Concrete and Rock,* (Sendai, Japan) Oct. 1988.

13. Shah, S.P., "Fracture Toughness of Cement-Based Materials", *Mater. Struct.*, Vol. 21, 1988, pp. 145-150.

14. Bažant, Z.P., "Fracture in Concrete and Reinforced Concrete", *Preprints, IUTAM Prager Symp. on Mechanics of Geomaterials: Rocks, Concretes, Soils* (Evanston), ed. Bažant, Z.P., 1983; also in *Mechanics of Geomaterials, Rocks, Concrete, Soils,* ed. Bažant, Z.P., John Wiley, Chichester, 1985, pp. 259-304.

15. Bažant, Z.P., "Size Effect in Blunt Fracture: Concrete, Rock, Metal", *J. Engng. Mech.*, ASCE, Vol. 110, No. 4, Apr. 1984, pp. 518-535.

16. Bažant, Z.P., "Mechanics of Distributed Cracking", *Appl. Mech. Rev.*, Vol. 39, No. 5, 1986, pp. 675-705.

17. Bažant, Z.P., and Pfeiffer, P.A., "Determination of Fracture Energy from Size Effect and Brittleness Number", *ACI Mater. J.*, Vol. 84, No. 6, Nov.-Dec. 1987, pp. 463-480.

18. Bažant, Z.P., and Pfeiffer, P.A., "Shear Fracture Tests of Concrete", *Mater. Struct.*, RILEM, Vol. 19, No. 110, 1986, pp. 111-121.

19. Bažant, Z.P., and Prat, P.C., "Measurement of Mode III Fracture Energy of Concrete", *Nucl. Engng. and Des.*, Vol. 106, 1988, pp. 1-8.

20. Bažant, Z.P., Prat, P.C., and Tabbara, M.R., "Antiplane Shear Fracture Tests (Mode III)", *ACI Mater. J.*, Vol. 87, No. 1, Jan.-Feb. 1990, pp. 12-19.

21. Bažant, Z.P., and Kazemi, M.T., "Size Effect in Fracture of Ceramics and Determination of Fracture Properties", *Report*, Center for Advanced Cement-Based Materials, Northwestern Univ., Evanston, 1989; to be printed in *J. Am. Cer. Soc.*

22. Bažant, Z.P., and Kazemi, M.T., "Determination of Fracture Energy, Process Zone Length and Brittleness Number from Size Effect, with Application to Rock and Concrete", *Rep. 88-7/498d*, Center for Concrete and Geomaterials, Northwestern Univ., Evanston, 1988; also in *Int. J. Fracture*, 1990, in press.

23. Bažant, Z.P., Gettu, R., and Kazemi, M.T., "Identification of Nonlinear Fracture Properties from Size Effect Tests and Structural Analysis Based on Geometry-Dependent R-Curves", *Rep. 89-3/498p*, Center for Advanced Cement-Based Materials, Northwestern Univ., Evanston, 1989; submitted to *Int. J. Rock Mech. Min. Sci.*

24. Bažant, Z.P., Lee, S.-G., and Pfeiffer, P.A., "Size Effect Tests and Fracture Characteristics of Aluminum", *Eng. Fracture Mech.*, Vol. 26, No. 1, 1987, pp. 45-57.

25. Bažant, Z.P., "Fracture Energy of Heterogeneous Materials and Similitude", *Preprints, SEM/RILEM Int. Conf. on Fracture of Concrete and Rock* (Houston), eds. Shah, S.P., and Swartz, S.E., Soc. for Exp. Mech., 1987; also in *Fracture of Concrete and Rock SEM-RILEM Int.*

Conf., eds. Shah, S.P., and Swartz, S.E., Springer-Verlag, New York, 1989, pp. 229-241.

26. Tada, H., Paris, P.C., and Irwin, G.R., *The Stress Analysis of Cracks Handbook* 2nd. ed., Paris Productions, St. Louis, 1985.

27. Murakami, Y., ed.-in-chief, *Stress Intensity Factors Handbook*, Pergamon Press, New York, 1987.

28. Hillerborg, A., "The Theoretical Basis of a Method to Determine the Fracture Energy G_F of Concrete", *Mater. Struct.*, RILEM, Vol. 18, No. 106, 1985, pp. 291-296.

29. Carpinteri, A., "Notch Sensitivity in Fracture Testing of Aggregative Materials", *Eng. Frac. Mech.*, Vol. 16, No. 4, 1982, pp. 467-481.

30. Carpinteri, A., "Size Effect on Strength, Toughness, and Ductility", *J. Engng. Mech.*, ASCE, Vol. 115, No. 7, 1989, pp. 1375-92.

31. Irwin, G.R., *Structural Mechanics*, Pergamon Press, London, 1960; see also Irwin, G.R., and Wells, A.A., "A Continuum-Mechanics View of Crack Propagation", *Metallurgical Reviews*, Vol. 10, No. 38, 1965, pp. 223-270.

32. Hillerborg, A., Modéer, M., and Petersson, P.E., "Analysis of Crack Formation and Crack Growth in Concrete by Means of Fracture Mechanics and Finite Elements", *Cem. Concr. Res.*, Vol. 6, 1976, pp. 773-782.

33. Krafft, J.M., Sullivan, A.M., and Boyle, R.W., "Effect of Dimensions on Fast Fracture Instability of Notched Sheets", *Proc., Crack*

Propagation Symp., College of Aeronautics, Cranfield, 1961, pp. 8-26.

34. Heyer, R.H., "Crack Growth Resistance Curves (R-Curves) - Literature Review", *Fracture Toughness Evaluation by R-Curve Methods*, ASTM STP 527, 1973, pp. 3-16.

35. Wecharatana, M., and Shah. S.P., "Predictions of Nonlinear Fracture Process Zone in Concrete", *J. Engng. Mech.*, ASCE, Vol. 109, No. 5, Oct. 1983, pp. 1231-46.

36. Bažant, Z.P., Kim, J.-K., and Pfeiffer, P.A., "Nonlinear Fracture Properties from Size Effect Tests", *J. Struct. Div.*, ASCE, Vol. 112, No. 2, 1986, pp. 289-307.

37. Jenq, Y.S., and Shah, S.P., "A Fracture Toughness Criterion for Concrete", *Eng. Frac. Mech.*, Vol. 21, No. 5, 1985, pp. 1055-1069.

38. Sakai, M., Yoshimura, J.-I., Goto, Y., and Inagaki, M., "R-Curve Behavior of a Polycrystalline Graphite: Microcracking and Grain Bridging in the Wake Region", *J. Am. Cer. Soc.*, Vol. 71, No. 8, Aug. 1988, pp. 609-616.

39. Jenq, Y.S., and Shah, S.P., "A Two Parameter Fracture Model for Concrete", *J. Engng. Mech. Div.*, ASCE, Vol. 111, No. 4, Oct. 1985, pp. 1227-1241.

40. Bažant, Z.P., and Oh, B.H., "Crack Band Theory for Fracture of Concrete", *Mater. Struct.*, RILEM, Vol. 16, No. 93, 1983, pp. 155-177.

41. De Larrard, F., Boulay, C., and Rossi, P., "Fracture Toughness of

High-Strength Concretes", in Holand, I. et al. (1987).

42. RILEM Committee on Fracture Mechanics of Concrete, "Determination of the Fracture Energy of Mortar and Concrete by Means of Three-point Bend Tests on Notched Beams", *Mater. Struct.*, RILEM, Vol. 18, No. 106, 1985, pp. 285-290.

43. Nakayama, J., "Direct Measurement of Fracture Energies of Brittle Heterogeneous Materials", *J. Am. Cer. Soc.*, Vol. 48, No. 11, 1965, pp. 583-587.

44. Tattersall, H.G., and Tappin, G., "The Work of Fracture and its Measurement in Metals, Ceramics and Other Materials", *J. Mater. Sci.*, Vol. 1, 1966, pp. 296-301.

45. Hillerborg, A., "Existing Methods to Determine and Evaluate Fracture Toughness of Aggregative Materials - RILEM Recommendation on Concrete", *Preprints of the Proc., Int. Workshop of Fracture Toughness and Fracture Energy - Test Methods for Concrete and Rock*, (Sendai, Japan) Oct. 1988.

TABLE 1

Concrete Composition for about 0.8 m^3 (1 yd^3)		
Dundee Cement (ASTM Type I)	363 kg	(800 lb)
Sand (FA2)	490 kg	(1080 lb)
Gravel (CA15) (Crushed Limestone)	880 kg	(1940 lb)
Water	127 kg	(280 lb)
Fly Ash (Class C)	91 kg	(200 lb)
Micro-Silica (W.R.Grace Force 10,000)	16 kg	(35 lb)
Retarder (naphthalene-based) (W.R.Grace WRDA 19)	1 ℓ	(35 oz.)
High-Range Water Reducer (Pozzolith 100XR)	5.3 ℓ	(180 oz.)

TABLE 2: Summary of Test Results

Specimen Dimensions† (mm x mm)	Peak Load (N)	Deflection at Peak Load (mm)	Final Deflection* (mm)	Peak CMOD (mm)	E_0 (GPa)	W_0 (N-mm)
38.1 x 101.6	1668 1490 1735	n.r. 0.0610 0.0914	n.r. 0.1092 n.r.	0.0127 0.0114 0.0146	6.0 7.0 5.5	n.r. 74 n.r.
76.2 x 203.2	2357 2091 2535	0.0864 0.0686 0.0711	0.1727 0.1626 0.1422	0.0222 0.0152 0.0292	7.7 8.2 9.0	167 143 142
152.4 x 406.4	3825 3736 3825	0.1143 0.0940 0.1710	0.2184 0.3200 0.2388	0.0381 0.0437 0.0356	9.2 6.2 12.7	362 418 326
304.8 x 812.8‡	7072 7442 7695	n.m.	n.m.	n.m.	n.m.	n.m.

† All specimens 38.1 mm wide

‡ Stroke-controlled tests, not used in analysis

* Deflection when the load drops to 10% of peak load

E_0 - Young's modulus from initial load-deflection compliance

W_0 - Area under load-deflection curve

n.m. - values not measured

n.r. - values not recorded due to failure of instrumentation

List of Captions

1. The size effect law.
2. Geometry of the three-point bending specimens used.
3. Specimens of four different sizes.
4. Test setup.
5. Instrumentation used for the tests.
6. Typical load-displacement curves for different sizes.
7. Typical load-CMOD curves for different sizes.
8. Linear regression for size effect parameters.
9. Size effect curve.
10. Variation of properties with strength.
11. Fracture toughness from work-of-fracture method.
12. Predicted and experimental load-deflection curves for different sizes, (a) d=152.4 mm (b) d=76.2 mm (c) d=38.1 mm.

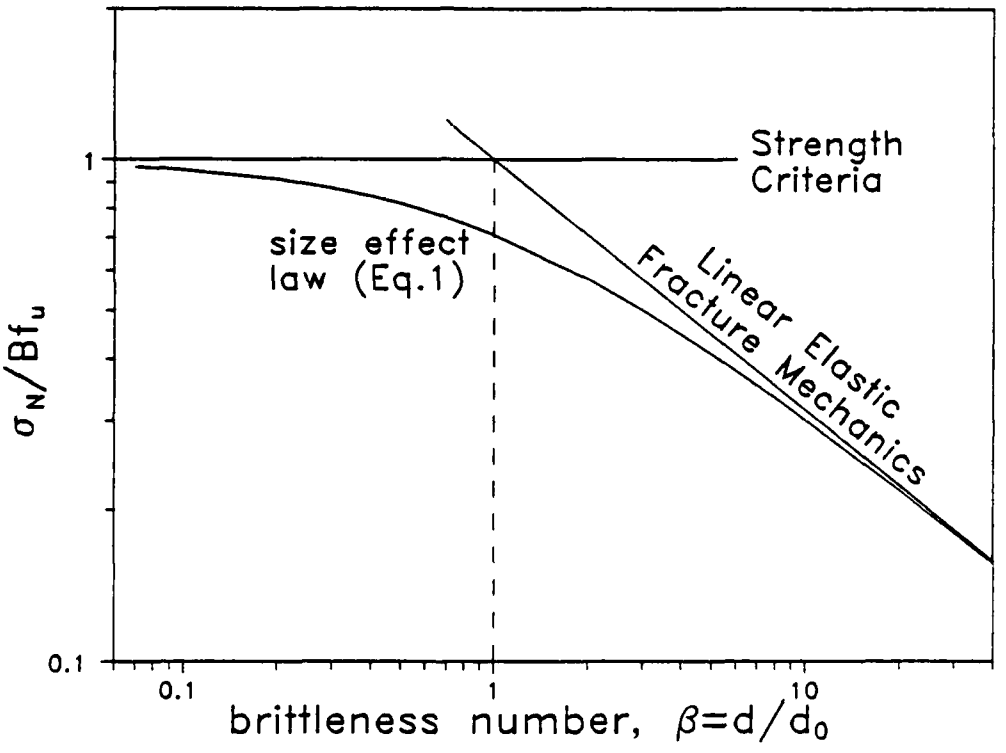

Strength
Criteria
size effect
law (Eq.1)
Linear Elastic
Fracture Mechanics
1
0.1
0.1
1
10
σ_N/Bf_u
brittleness number, $\beta=d/d_0$

FIG. 1

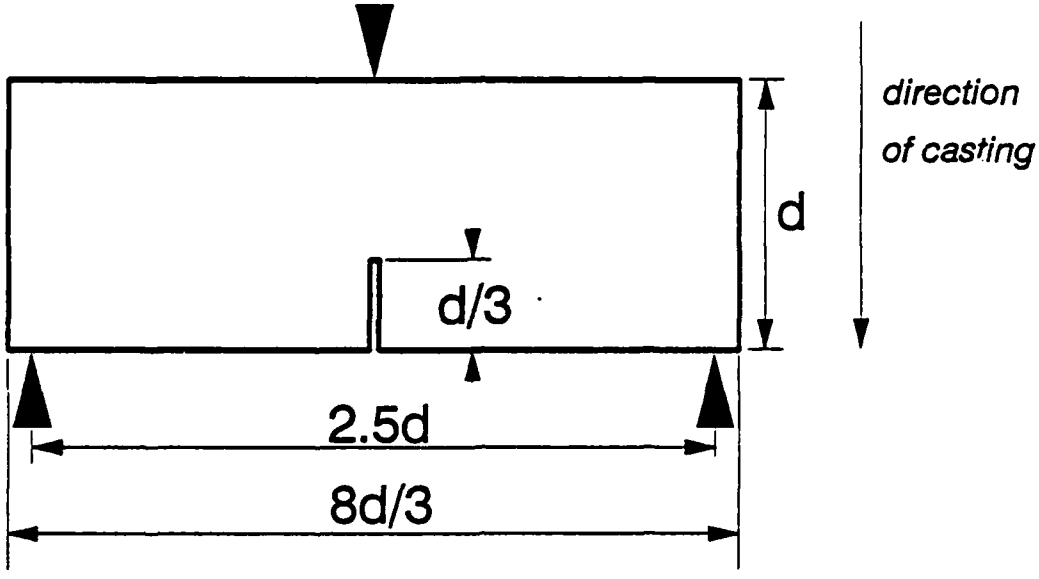


FIG. 2

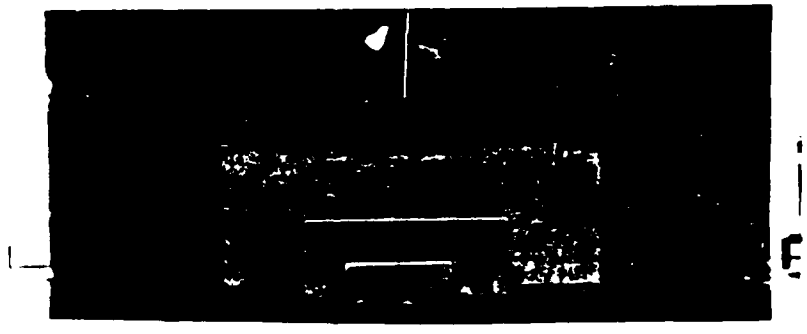


FIG. 3



FIG. 4



FIG. 5

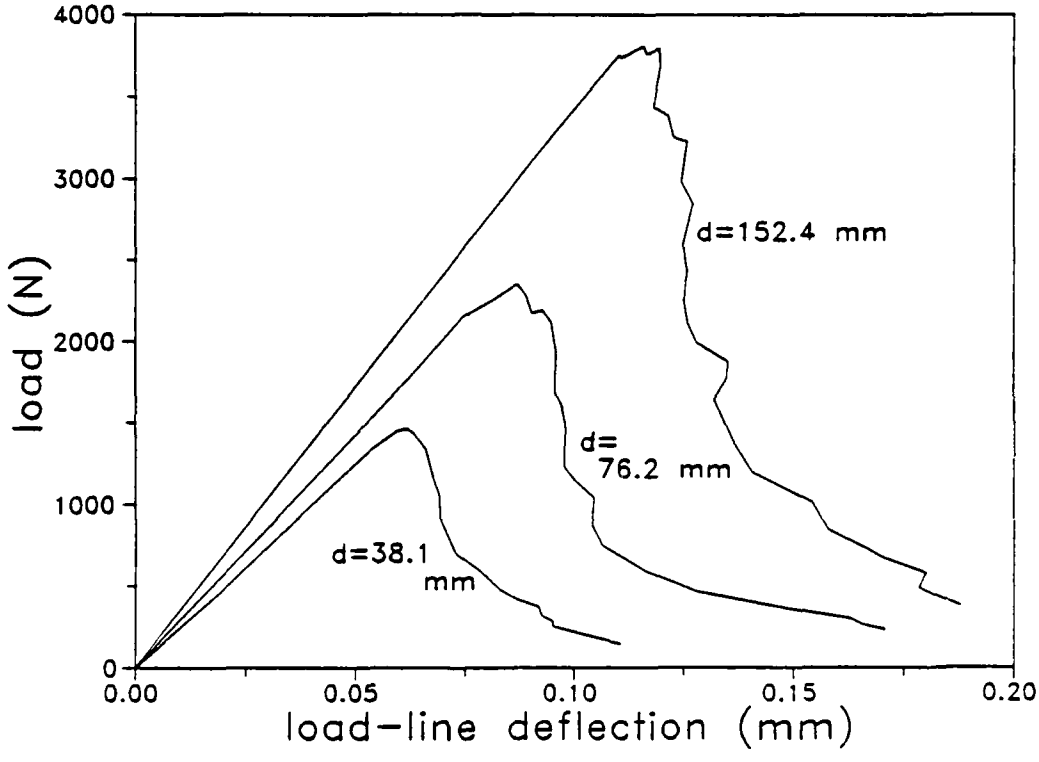

4000
3000
2000
1000
0
load (N)
0.00
0.05
0.10
0.15
0.20
load-line deflection (mm)
d=152.4 mm
d=
76.2 mm
d=38.1
mm

FIG. 6

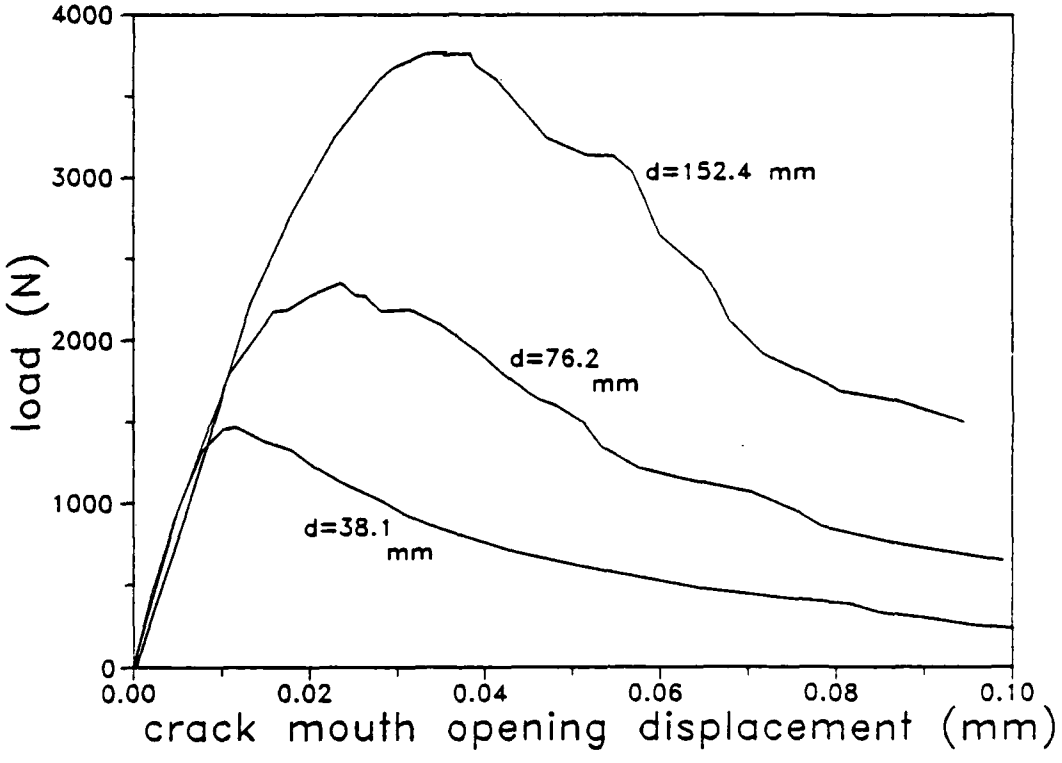

4000
3000
2000
1000
0
load (N)
d=152.4 mm
d=76.2 mm
d=38.1 mm
0.00
0.02
0.04
0.06
0.08
0.10
crack mouth opening displacement (mm)

FIG. 7

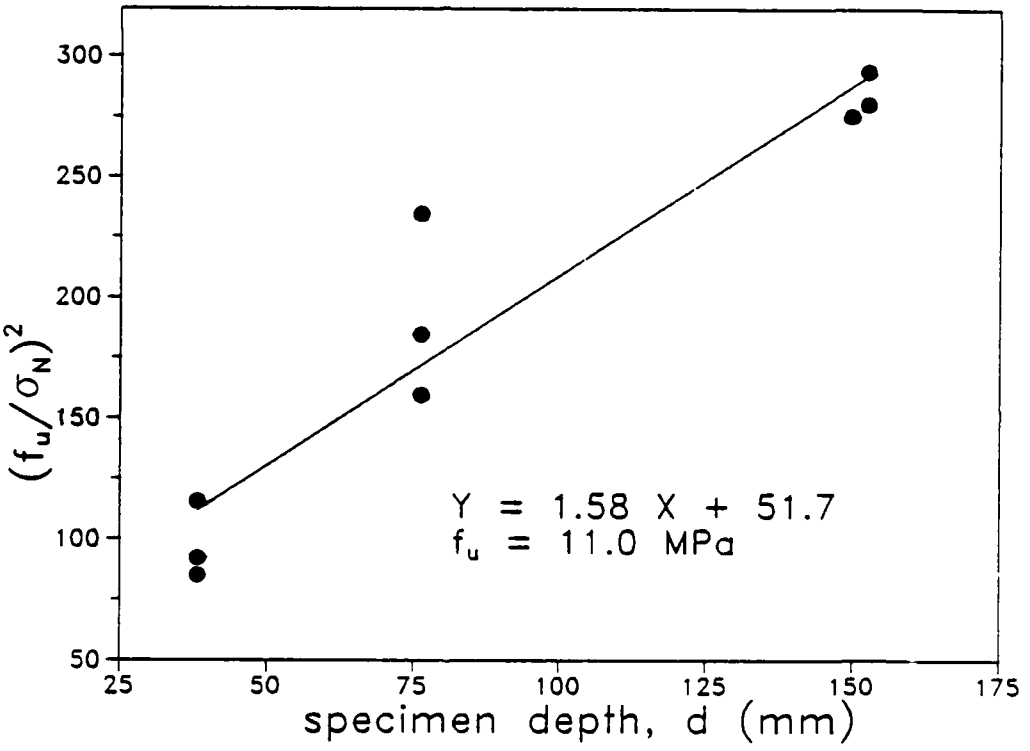

300
250
200
150
100
50
25
50
75
100
125
150
175
$(f_u/\sigma_N)^2$
specimen depth, d (mm)
Y = 1.58 X + 51.7
f_u = 11.0 MPa

FIG. 8

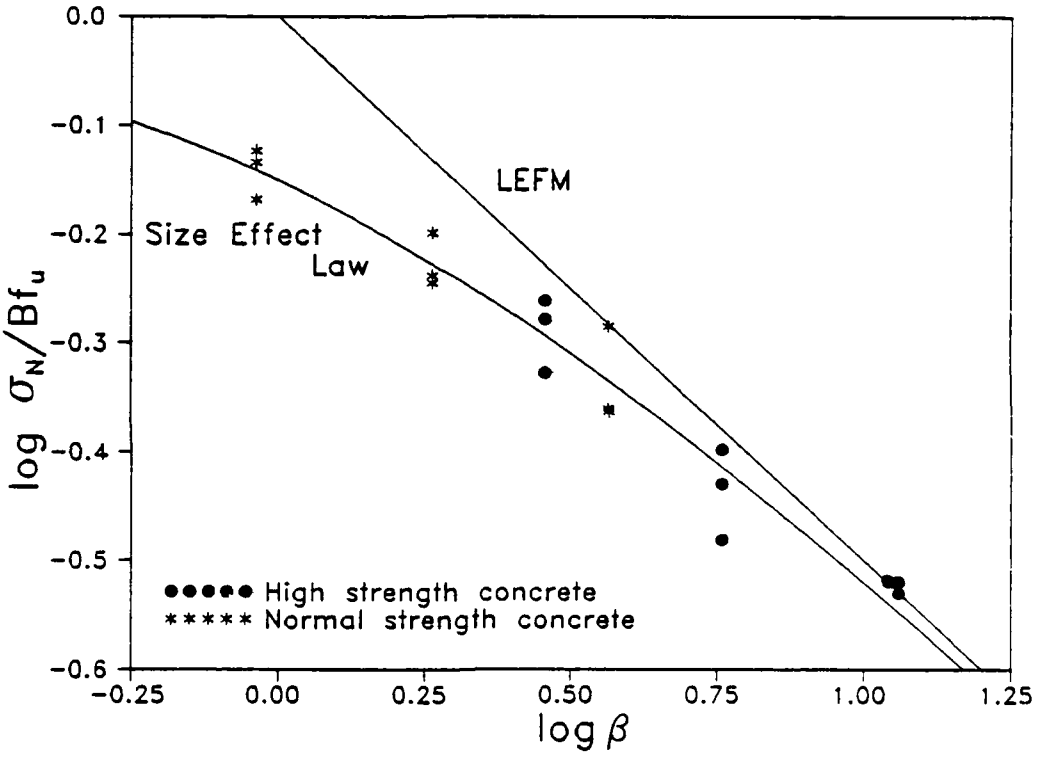

0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
log σ_N/Bf_u
LEFM
Size Effect
Law
●●●●● High strength concrete
***** Normal strength concrete
-0.25
0.00
0.25
0.50
0.75
1.00
1.25
log β

FIG. 9

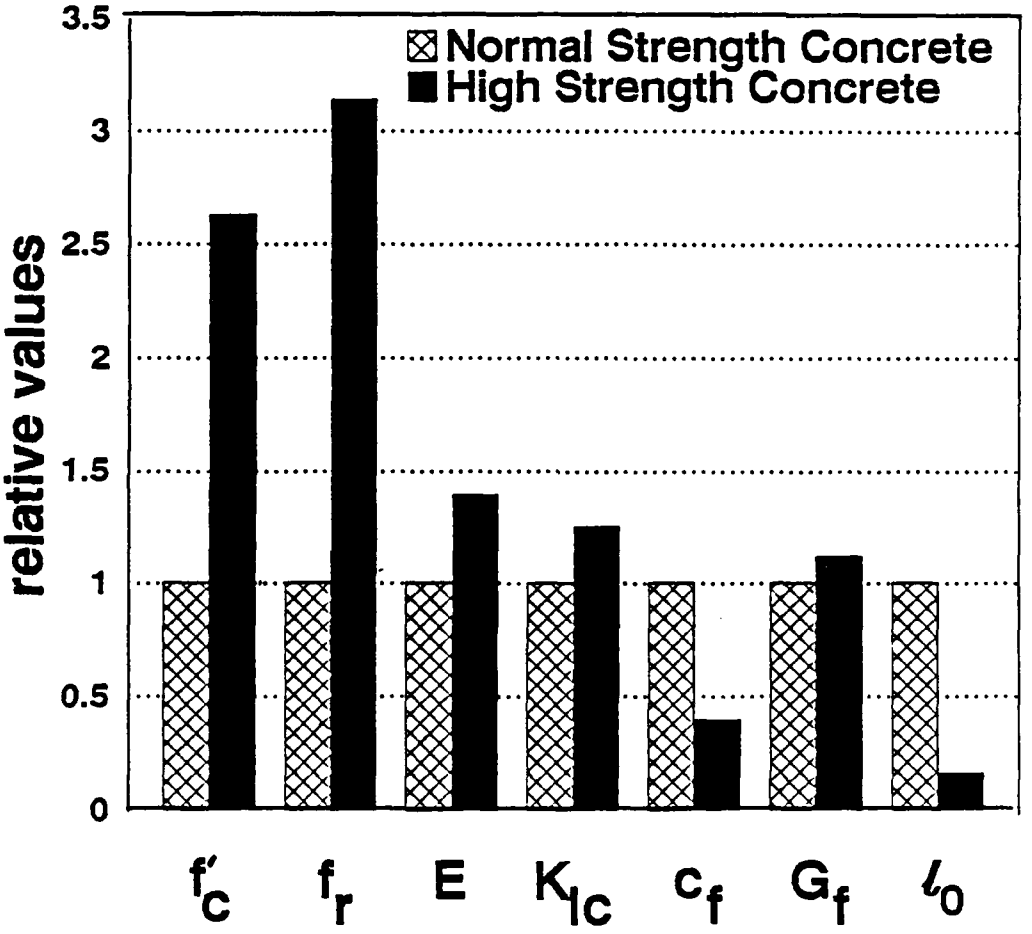

Normal Strength Concrete
High Strength Concrete
relative values
3.5
3
2.5
2
1.5
1
0.5
0
f'_c
f_r
E
K_{Ic}
c_f
G_f
ℓ_0

FIG. 10

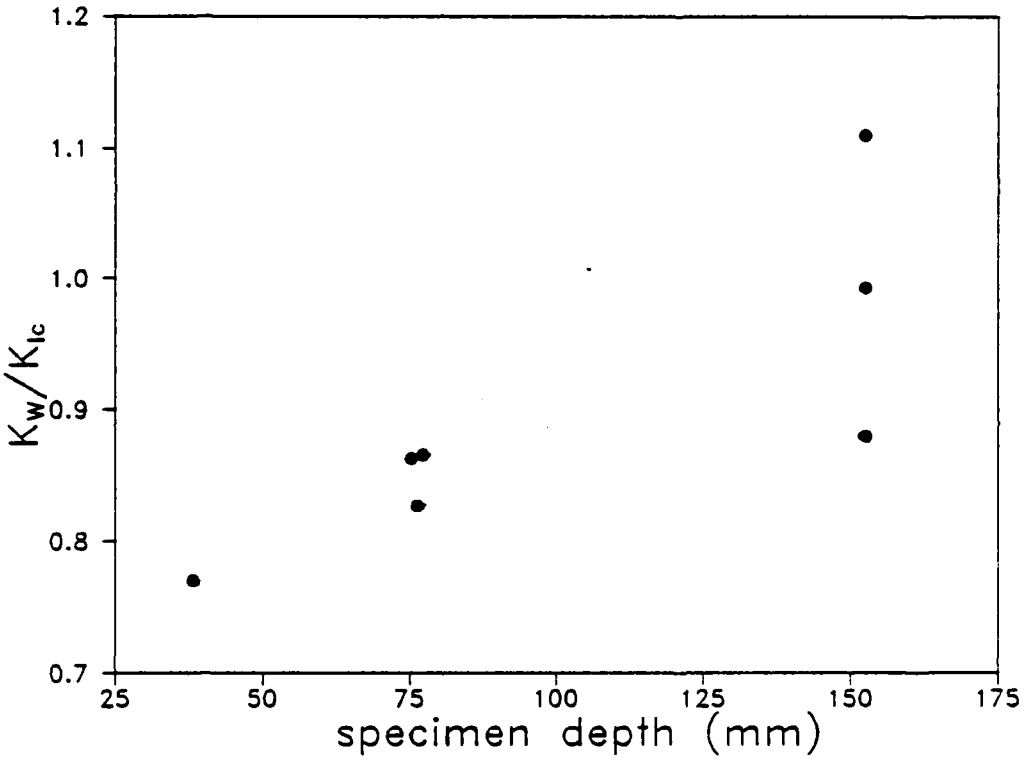

1.2
1.1
1.0
0.9
0.8
0.7
K_w/K_{Ic}
25
50
75
100
125
150
175
specimen depth (mm)

FIG. 11

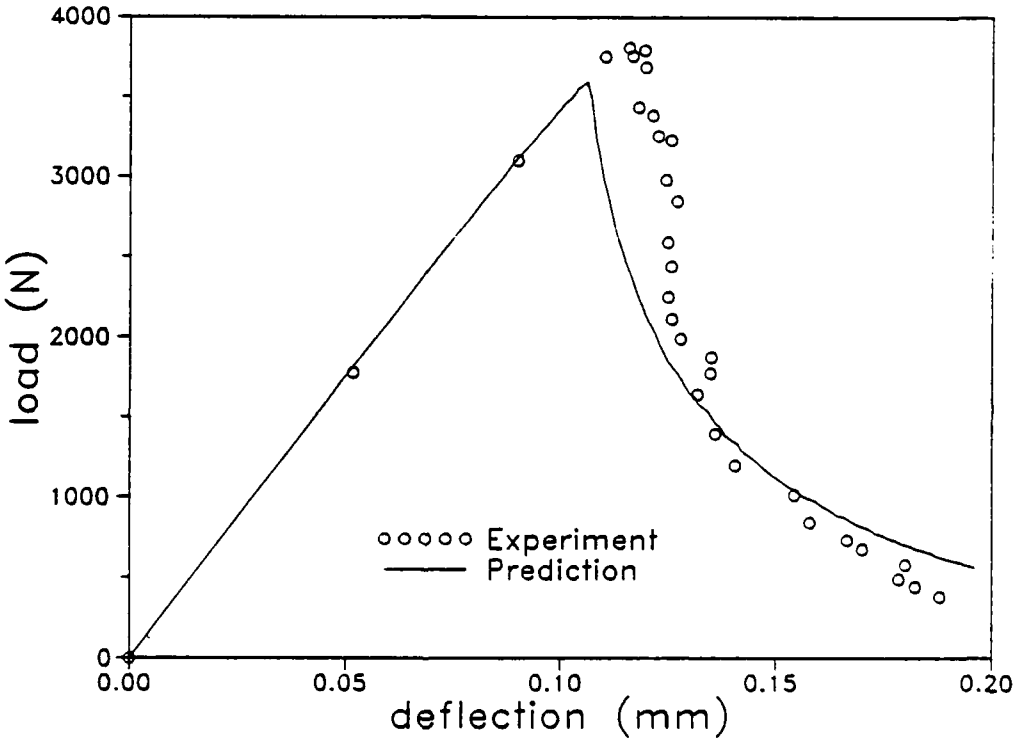

4000
3000
2000
1000
0
load (N)
0.00
0.05
0.10
0.15
0.20
deflection (mm)
Experiment
Prediction

FIG. 12(a)

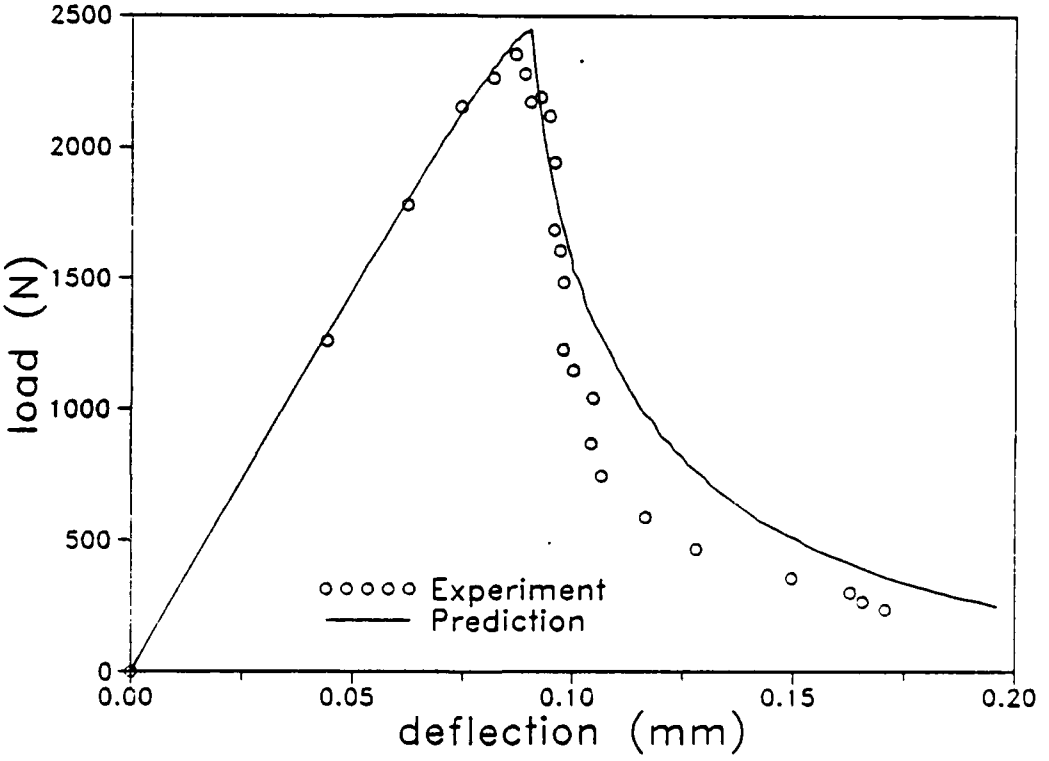

2500
2000
1500
1000
500
0
load (N)
0.00
0.05
0.10
0.15
0.20
deflection (mm)
Experiment
Prediction

FIG. 12(b)

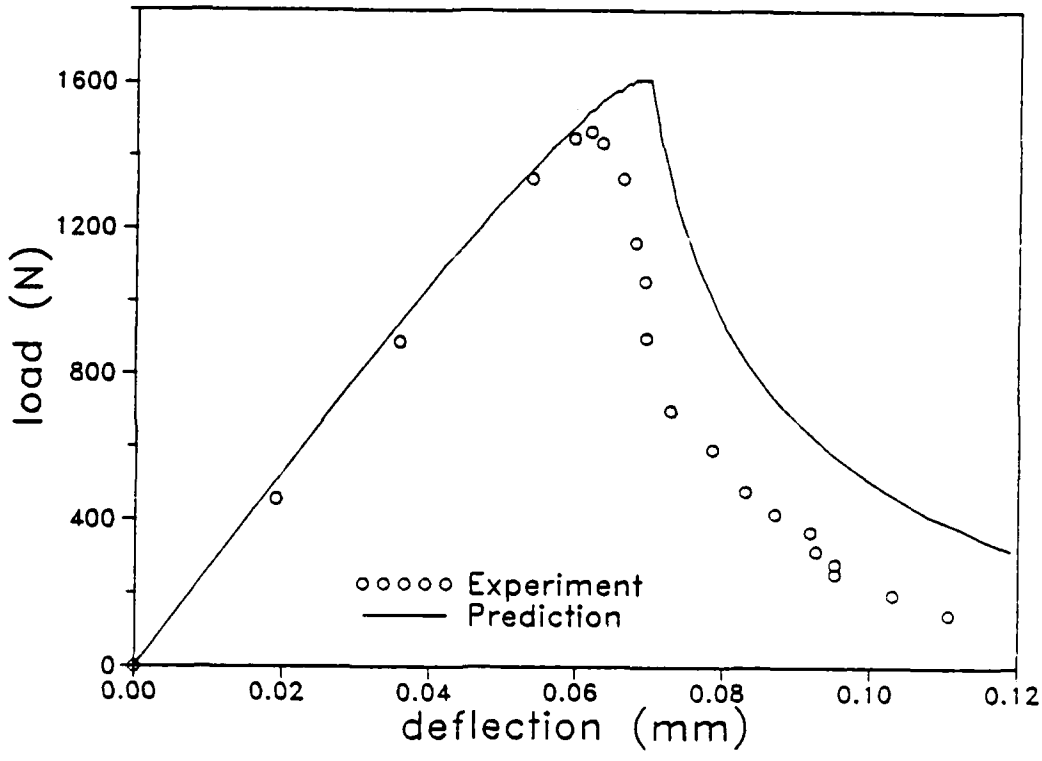

1600
1200
800
400
0
load (N)
0.00
0.02
0.04
0.06
0.08
0.10
0.12
deflection (mm)
Experiment
Prediction

FIG. 12 (c)

ACI MATERIALS JOURNAL TECHNICAL PAPER

Title no. 87-M2

Antiplane Shear Fracture Tests (Mode III)



By Zdeněk P. Bažant, Pere C. Prat, and Mazen R. Tabbara

Mode III (antiplane) shear-fracture energies of concrete and mortar were measured on the basis of the size-effect law. The specimens were cylinders with a circumferential notch at their midlengths, subjected to pure torsion at zero axial force. The specimens were geometrically similar and their sizes varied as 1:2:4. The mortar specimens were observed to behave in a manner closer to linear elastic fracture mechanics than did the concrete specimens. The ratio of Mode III to Mode I (tensile) fracture energies was found to be about 3 for concrete and about 8 for mortar. The results indirectly indicate that a volume expansion of the sheared fracture-process zone may be a significant mechanism causing transverse tensile stresses across the ligament. The size-effect measurements also yield an estimate of the size of the fracture-process zone, which is found to be nearly the same as for Mode I fracture.

Keywords: concretes; cracking (fracturing); dimensional analysis; energy; measurement; mortars (material); shear tests; specimens; tensile stresses; tests.

Until recently it had been generally accepted that fracture in concrete is always tensile and propagates in the direction normal to the maximum principal stress. Such behavior is indeed observed in the diagonal shear failure and torsional failure of beams, as well as in punching-shear failure of slabs. The existence of shear fracture in brittle materials has been denied, and even such claims as "shear fracture is sheer nonsense" have been heard.

Doubts arose, however, when dynamic tests of certain concrete box structures loaded by short pressure pulses caused the top slabs to fail by direct shear next to the supports. As a result of this and other experiences, tests of shear fracture were undertaken at Northwestern University[1] using edge-notched specimens of Iosipescu geometry loaded in Mode II (in-plane shear). The size-effect method was used in those tests to determine the Mode II fracture energy. Interpretation of those tests was not unambiguous, because a perfectly antisymmetric Mode II field at the crack tips is not achieved in this type of test, not only in the actual nonlinear behavior, but also according to the theory of elasticity. The elastic solution of the stress field near the crack tips in those specimens indicates the presence of compressive stresses across the crack-line extension, which represents deviations from the ideal Mode II conditions. In theory, a perfect Mode II shear loading could be achieved in an elastic specimen of that type, but this would require applying additional antisymmetric tensile forces, as shown in Reference 1, which would be difficult to carry out. Moreover, as pointed out by Ingraffea and Panthaki[2] in a dialogue with Bažant and Pfeiffer,[3] the failure of these specimens might be partly caused by axial compression splitting (similar to the Brazilian split-cylinder test), which again prevents attaining pure Mode II shear conditions.

To gain additional insight, Bažant and Prat[4] decided to study Mode III (antiplane shear) rather than Mode II (in-plane shear). Using cylindrical specimens with a circumferential notch subjected to torsion (Fig. 1 and 2), they carried out Mode III fracture tests of concrete, which apparently had not been conducted previously. According to the theory of elasticity, the applied torque produces a stress and deformation field that is perfectly antisymmetric with regard to the crack plane; thus, a planar crack represents a pure shear crack. The size-effect method[5,6] has been used to determine the Mode III fracture energy of concrete G_f^{III}. Surprisingly, the value of G_f^{III} obtained from these tests has been much smaller than the value of the Mode II fracture energy G_f^{II} obtained from the preceding tests of notched beams, although still larger than the Mode I (opening mode) fracture energy G_f^I. It apperared that the value of shear-fracture energy might not represent a material constant but could be a function of the normal stress across the crack plane, and that the value of this stress might in turn depend on the volume changes in the fracture process zone. This conclusion was supported by the finding that torsional loading in a triaxial torsional testing machine that produces axial confinement and prevents free axial expansion of the cylinder leads

ACI Materials Journal, V. 87, No. 1, January-February 1990
Received Aug. 18, 1988, and reviewed under Institute publication policies.
Copyright © 1990, American Concrete Institute. All rights reserved, including the making of copies unless permission is obtained from the copyright proprietors. Pertinent discussion will be published in the November-December 1990 *ACI Materials Journal* if received by Aug. 1, 1990.

Zdeněk P. Bažant, FACI, is a professor at Northwestern University, Evanston, Ill., where he recently served as Director of the Center for Concrete and Geomaterials. Dr. Bažant is a registered structural engineer, a consultant to Argonne National Laboratory, and is on the editorial boards of several journals. He is Chairman of ACI Committee 446, Fracture Mechanics; a member of ACI Committees 209, Creep and Shrinkage in Concrete; and 348, Structural Safety; and Chairman of RILEM's committee on creep and SMiRT's Division H. In 1987, Professor Bažant visited the University of Tokyo as Kajima Foundation Fellow and has been NATO senior guest scientist at E.N.S., Paris-Cachan.

ACI member Pere C. Prat is a scientific research associate at the Materials Science Institute (CSIC), Barcelona, Spain. He obtained his civil engineering degree from the School of Civil Engineering of the Technical University of Catalonia, Barcelona, Spain, in 1982, and his PhD from Northwestern University in 1987. His research interests include constitutive models for geomaterials, strain-softening behavior, fracture mechanics, nonlinear analysis of concrete structures, experimental methods, and finite element applications.

ACI member Mazen R. Tabbara is a graduate research assistant at Northwestern University, Evanston, Illinois. His research interests include constitutive models for concrete, fracture mechanics, and finite element applications.

to a nonplanar failure mode; see the conical failure surfaces in Fig. 3.

The purpose of this paper is to examine these questions more deeply, report the results of a larger series of Mode III failure tests that included mortar specimens as well as concrete specimens, determine the Mode III fracture energy, and examine the applicability of the size-effect law. An additional purpose is to extract from the test results an estimate for the length of the fracture process zone. The significance of this research is expected to be an improved ability to predict brittle failures of concrete structures.

The torsional fracture specimen conceived by Bažant and Prat[4] was independently also introduced by Suresh and Tschegg[9] to study mixed-mode fracture of fatigued ceramics. Using combined torsional-axial loading, Suresh and Tschegg determined a failure surface in terms of the Mode III and Mode I stress-intensity factors. Their conclusions on the effect of axial stresses are similar to those made by Bažant and Prat and are extended further here.

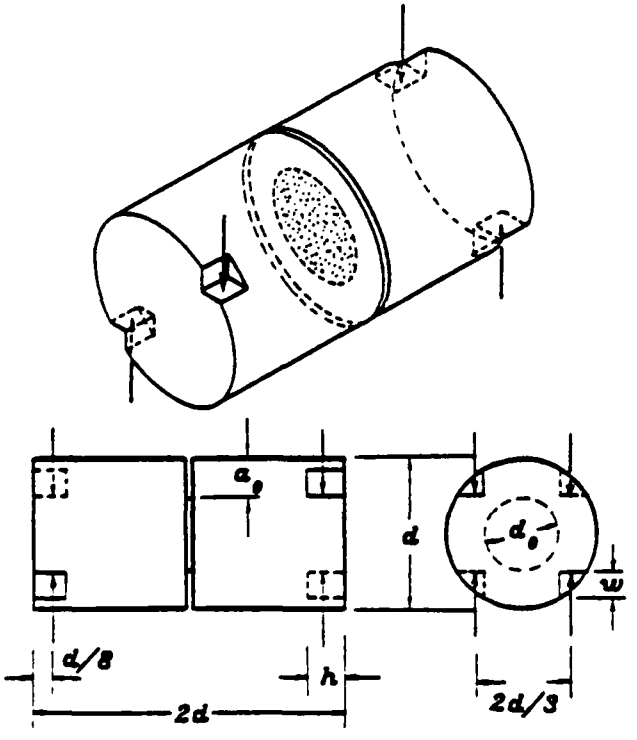


Fig. 1 — Torsional circumferentially notched fracture specimens

EXPERIMENTAL INVESTIGATION

To determine fracture energy by the size-effect method, geometrically similar specimens of significantly different sizes needed to be tested. Similar circumferentially notched cylinders of diameters d = 1.5, 3, and 6 in. (38, 76, and 152 mm) were used. The length-to-diameter ratio was ℓ/d = 2. Circular notches of thickness 1/16 in. (1.6 mm), located in the middle plane perpendicular to the axis of the cylinders (Fig. 1 and 2), were created by casting. The notch depth was a_0 = $d/4$.

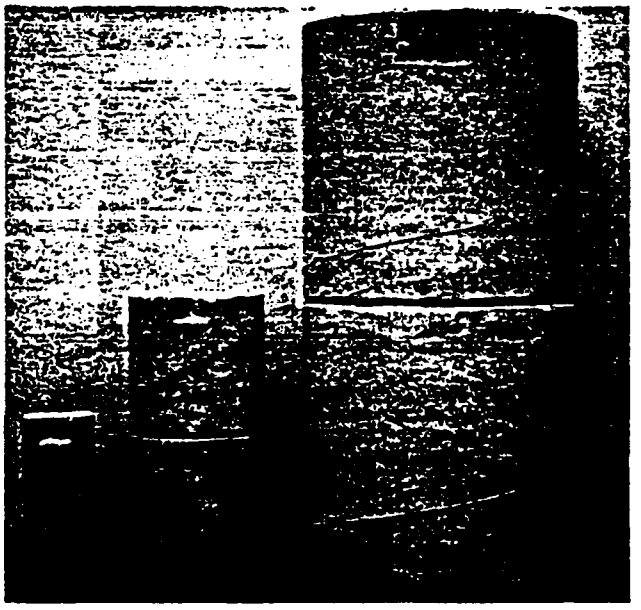


Fig. 2 — Geometrically similar notched specimens used in tests

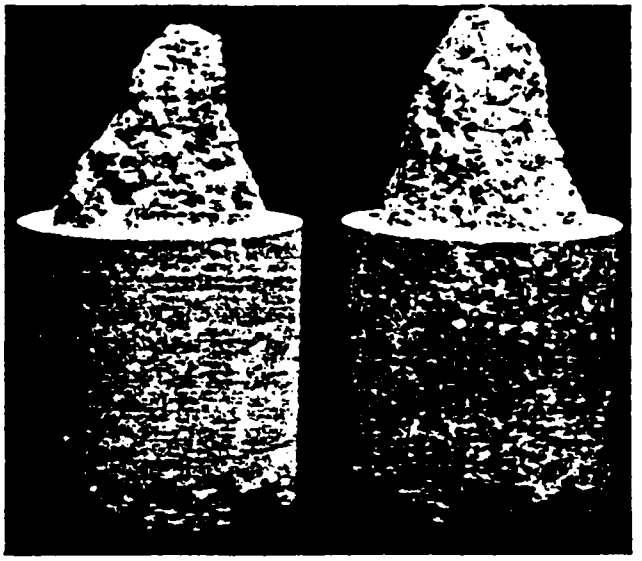


Fig. 3 — Conical failure surface observed in previous tests with restrained axial displacement

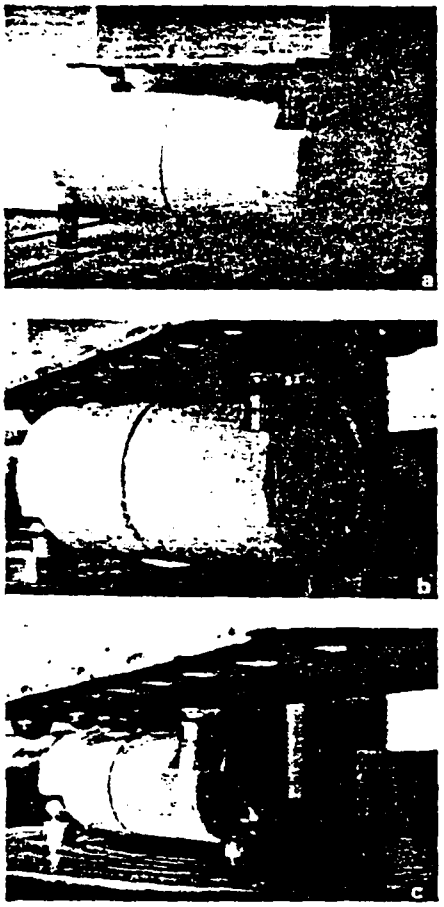


Fig. 4 — Specimens during testing in MTS machine

The torque T was applied at each end of the specimen as a couple, as shown in Fig. 1 and 4. The wedge-shaped cutouts on which the loads were applied were cast in a standard cylinder mold in which wooden inserts were placed. Each end-couple had an arm of $2d/3$ and was applied in a plane perpendicular to the axis of the specimen at a distance of $d/8$ from the ends.

Two test series were carried out, one for concrete and one for mortar. Each test series consisted of nine torsional specimens, three for each size (Fig. 2), plus three companion cylinders 3 in. (76 mm) in diameter and 6 in. (152 mm) long, to test for compression strength. All the specimens and control cylinders for each test series were cast from the same batch. The means and standard deviations of the compression strength f'_c after 28 days of moist curing are given in Table 1. The 28-day f'_c values for concrete were not measured directly, but calculated from the compression strength $f'_c(35)$, measured at the time of the tests at which the age of concrete was 35 days. The adjustment for age was based on the approximate formula[10] $f'_c(t) = f'_c(28)$ [1 + 0.277 log (t/ 28)] from which $f'_c(28) = 0.974\, f'_c(35)$.

All specimens were cast with the longitudinal axis in a vertical position, to insure statistical homogeneity and isotropy within the notch plane. The concrete mix had a water-cement ratio of 0.6, and a cement-sand-gravel ratio of 1:2:2 (all ratios by weight). The concrete had maximum gravel size d_a = 0.5 in. (12.7 mm) and maximum sand-grain size of 0.19 in. (4.8 mm). The aggregate consisted of crushed limestone and siliceous Lake Michigan beach sand from Illinois. ASTM C 150 Type I portland cement, with no admixtures and no air-entraining agents, was used. The mortar mix had a water-cement ratio of 0.5 and a cement-sand ratio of 1:2. The same sand as for the concrete specimens was used, i.e., the maximum aggregate size was d_a = 0.19 in. (5 mm). The aggregate composition was thus the same as for the concrete specimens except for the omission of gravel. The water-cement ratio differed from that of the concrete specimens in order to achieve approximately the same workability.

Table 1 — Compression strength f'_c, psi

Mix	Mean	Standard deviation
Concrete	5633	221
Mortar	5365	624

The specimens were cast in waxed cardboard molds from which they were removed after one day. Subsequently, they were cured in a moist room at 95 percent humidity and 80 F temperature until about 1 hr before the test. The tests of mortar specimens were made at the age of 28 days. The age of the concrete specimens at the time of test was 35 days, instead of the standard 28 days, because of a delay due to the repair of equipment. Fig. 4 shows the specimens of all sizes installed in the testing machine. The values of the strength and fracture energy for the age of 28 days were obtained by adjusting the 35-day values.

Fig. 5 exhibits the fractured concrete specimens after testing and demonstrates that the fracture surface was planar and did not propagate in the direction normal to the maximum principal stress, which is inclined. A similar fracture surface was observed for the mortar specimens. All the tests were carried out at room temperature in a closed-loop MTS machine under stroke-control conditions. The loading rates were chosen such that the time to maximum load was 3 to 5 min for specimens of all sizes. The size-effect method, used to calculate the fracture energy, required knowledge of only the maximum load, from which the maximum torque was then calculated.

SIZE-EFFECT ANALYSIS AND CALCULATION OF FRACTURE ENERGY

Size effect is understood as the dependence of the nominal stress at maximum load (failure) τ_N on a characteristic dimension of the specimen d, when geometrically similar specimens or structures are considered. In the case of three-dimensional similarity, τ_N is generally defined as $\tau_N = CP/d^2$, where P = maximum load, d = characteristic dimension of the body, and C = arbitrary nondimensional constant. For the present specimens $\tau_N = 16T/\pi d^3 = CP/d^2$, where T = maximum torque; $C = 16R/\pi d$, which is constant since R/d = constant for geometrically similar specimens; and $T = PR$ where R is the arm of the force couple applied at the ends. Note that the value of C has no effect on the

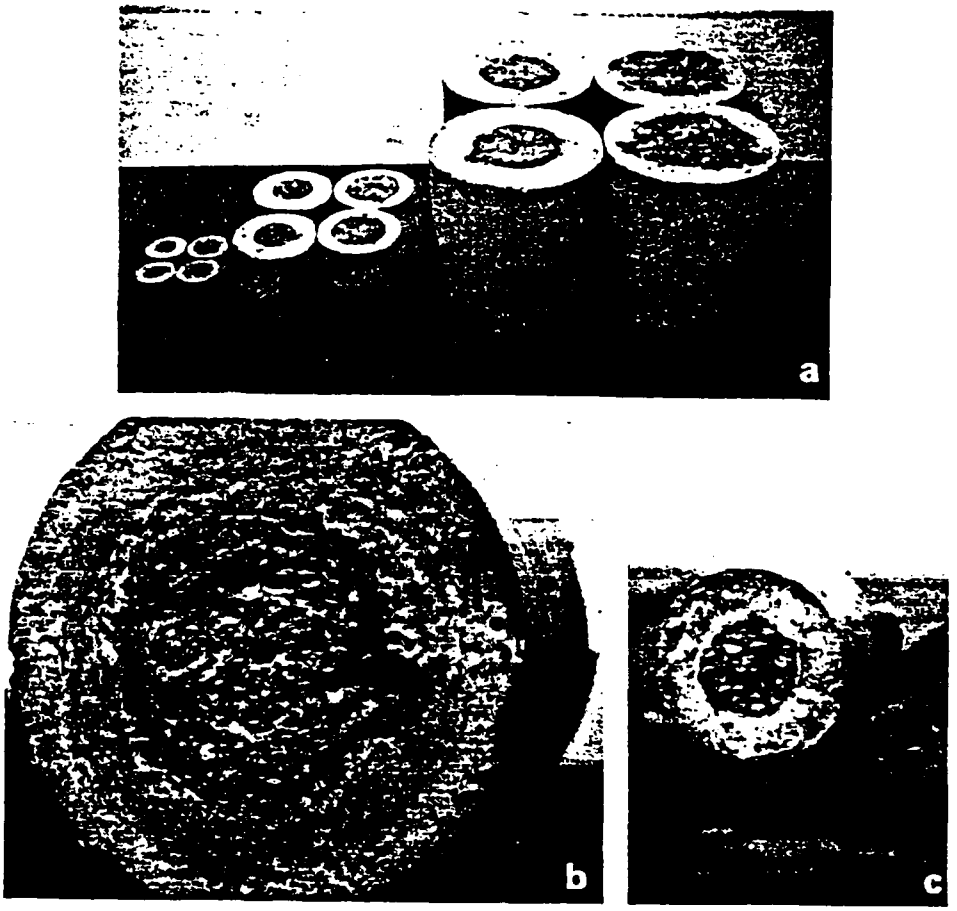


Fig. 5 — Fractured specimens after testing, showing the failure surfaces

values of fracture energy given by Eq. (3), which follows. The value of τ_N as defined represents the maximum elastic stress if there were no notch or crack.

Because the cracks in concrete propagate with a relatively large microcracking zone which blunts the fracture front, the size effect represents a transition between the plastic limit analysis for which there is no size effect (i.e., τ_N is constant), and the classical linear-elastic fracture mechanics for which the size effect is the strongest possible and which is of the type $\tau_N \sim d^{-1/2}$. The transition may be described by the approximate size-effect law[5]

$$\tau_N = Bf_t' \left(1 + \frac{d}{\lambda_0 d_a}\right)^{-1/2} \quad (1)$$

where B and λ_0 are empirical constants, d_a is the maximum size of aggregate, and f_t' is the tensile strength of concrete. Eq. (1) has been derived in References 5 and 11, and for three dimensions in Reference 12. It has been experimentally validated in Mode I (opening) fracture[6] and was also found to approximately apply to the quasi-Mode II (shear) fracture tests in Reference 1, and to diagonal shear.[13] Tests showed that Eq. (1) is also applicable to various brittle failures of concrete structures, such as failure of beams in torsion, punching-shear failure of slabs, failure of pipes, etc.[14-16]

To determine parameters B and λ_0, Eq. (1) may be transformed to the form

$$\left(\frac{f_t'}{\tau_N}\right)^2 = \frac{1}{B^2 \lambda_0}\left(\frac{d}{d_a}\right) + \frac{1}{B^2} \quad (2)$$

Eq. (2) represents a linear relation between $(f_t'/\tau_N)^2$ and d/d_a, and may be rewritten as $Y = AX + C$, where $Y = (f_t'/\tau_N)^2$, $X = d/d_a$, $C = 1/B^2$, and $A = C/\lambda_0$. A linear regression of the test results may be used to determine A and C, from which $B = C^{-1/2}$ and $\lambda_0 = C/A$. The regression analysis also yields the statistics of the errors. Fig. 6 and 7 give the coefficient of variation $\omega_{Y|X}$ of the vertical deviations from the regression line and the correlation coefficient ρ.

As proposed in References 11 and 12, the concrete fracture energy G_f may be uniquely defined as the energy release rate required for crack propagation in an infinitely large specimen. In theory, this definition must yield results that are independent of both the size and the shape of the specimen, provided that the correct size-effect law is known. The exact size-effect law is unknown, but the approximate size-effect law in Eq. (1) has been shown to be adequate for practical purposes, and valid for the size range 1:20 regardless of specimen shape. The following formula has been derived[5,11,12]

$$G_f = \frac{g(\alpha_0)}{AE_c} f_t'^2 d_a \quad (3)$$

in which $\alpha_0 = a_0/r$; a_0 = notch depth; r = radius of the

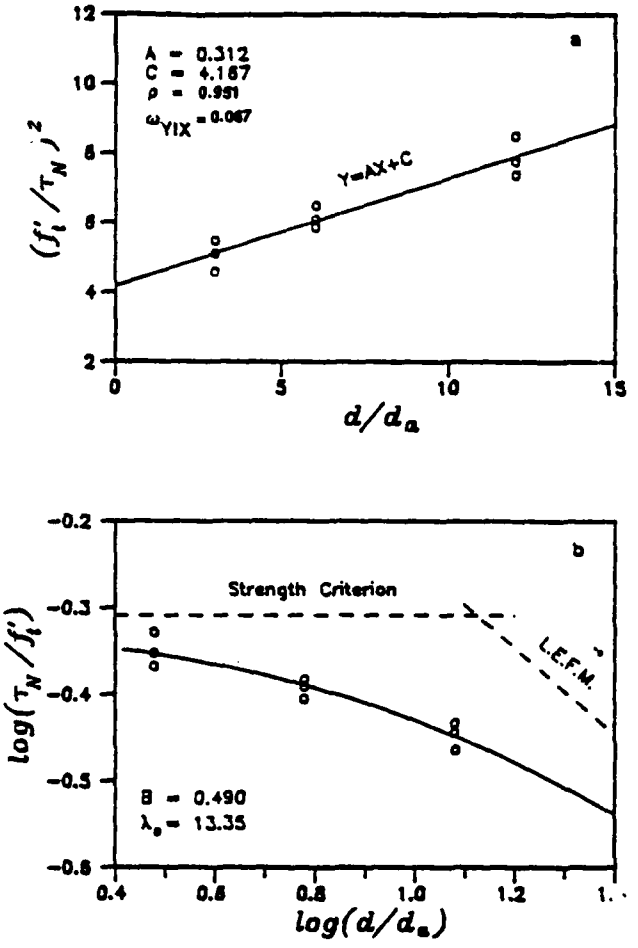


Fig. 6 — Results from tests of concrete specimens: (a) linear regression, and (b) size effect

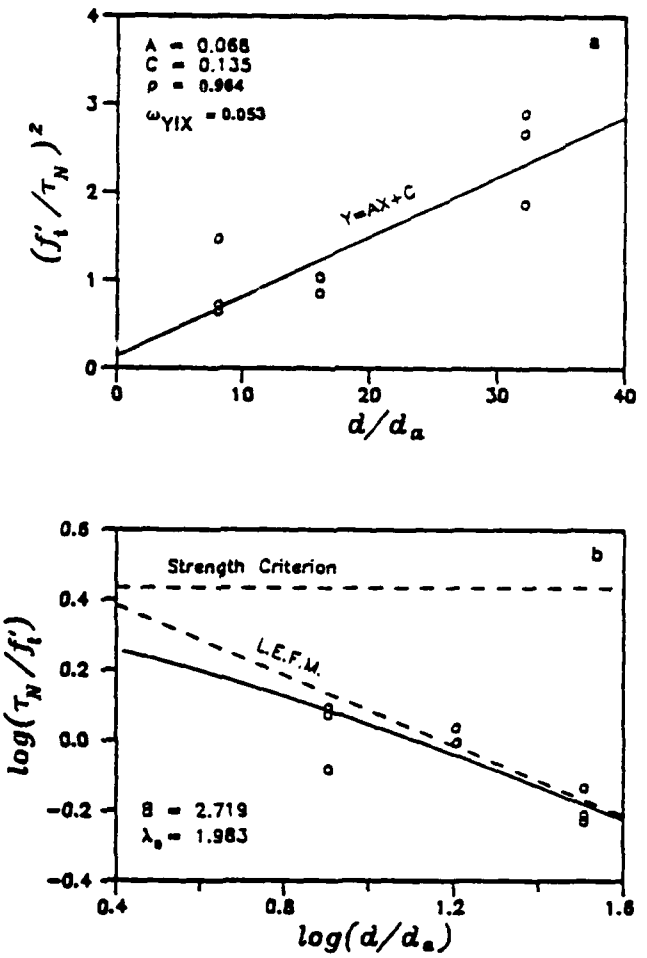


Fig. 7 — Results from tests of mortar specimens: (a) linear regression, and (b) size effect

Table 2 — Fracture energy G_f, lb/in. (N/m)

Mix	Mode I	Mode II	Mode III
Concrete	0.22 (39)	6.30 (1104)	0.61 (107)
Mortar	0.13 (23)	3.34 (585)	1.01 (177)

cylinder; E_c = modulus of elasticity of concrete; f_t' = direct tensile strength of concrete; $A = 1/(B^2\lambda_0)$ = slope of the regression line as already defined; $g(\alpha_0)$ = nondimensional energy release rate of the specimen according to the linear elastic fracture mechanics, which can be found for the basic specimen geometries in textbooks and handbooks,[17,18] and can in general be determined by linear finite element analysis. For the notch depth of $d/4$(i.e., $\alpha_0 = 0.5$), $\alpha\ (a_0) = 6.99\ (1 + \upsilon)$ where υ = Poisson's ratio. In the calculations it was assumed that $\upsilon = 0.18$.

The tensile strength was not measured but was estimated from the formula $f_t' = 6\sqrt{f_c'}$, where f_c' and f_t' are in psi (6895 Pa) and f_c' = standard cylindrical compression strength. Since the slope A of the linear regression plots is proportional to $f_t'^2$, an error in f_t' has no effect on the fracture-energy values.

The measured maximum torques are plotted in Fig. 6 and 7. The plots at the bottom demonstrate that the measured τ_N agrees with the size-effect law (solid curve). The plots at the top show the regression line, which yields the mean values of the parameters of the size-effect law.

The fracture-energy values obtained for concrete in this manner pertain to the age of 35 days. They were then transformed to 28 days, by using Bažant and Oh's empirical formula[19] $G_f = (2.72 + 0.0214\ f_t')\ f_t'^2\ d_a/E_c$, which yields $G_f(28) = 0.976G_f(35)$. The plots in Fig. 6 and 7 correspond to the 28-day values.

The values obtained for fracture energy are given in Table 2. For comparison, Table 2 also indicates the values for Modes I and II previously obtained from tests of essentially the same concrete and mortar.

Aside from the fracture energy, the size-effect law makes it possible to estimate the effective length c_f of the fracture process zone, using the formula[8]

$$c_f = \frac{d_0 g(\alpha_0)}{g'(\alpha_0)} \tag{4}$$

in which $g'(\alpha_0) = dg(\alpha)/d\alpha$ at $\alpha = \alpha_0$. The value of c_f represents the effective length of the fracture process zone for an infinitely large specimen, which is a material constant. The present test results yield the values

given in Table 3. For comparison, this table also shows the c_f values obtained from the previous test results[4] for Mode I. Notice the good agreement between Mode I and Mode III results, which suggests that c_f should indeed be a material property. Theoretically, the obtained value of c_f applies to an infinitely large specimen, but because Bažant's law is not sufficiently accurate for an inifinite size range, the value obtained for c_f applies to a size about ten times larger than that of the largest specimen tested.

ANALYSIS OF RESULTS

According to the size-effect law, the test results for mortar should be closer to linear elastic fracture mechanics than those for concrete, for the same specimens. This is confirmed by the present results; see Fig. 6 and 7 where the data points for mortar lie closer to the straight line asymptote of linear elastic fracture mechanics, whose slope is -0.5.

Comparison with the Mode II and Mode I fracture energies (Table 2) reveals a surprise. The values of G_f^{III} might have been expected to be about the same as G_f^{II}, and much larger than G_f^{I}. But this has not turned out to be the case; rather, the measured G_f^{III} is much smaller than G_f^{II}, but larger than G_f^{I}. The mechanism to explain this finding will require deeper study, but we shall attempt an approximate analysis.

A clue to the source of this discrepancy is provided by the experimetal observation that the results of shear fracture tests are very sensitive to the restraint of the specimen in the direction normal to the fracture plane (the axial direction of the cylinder for the present tests). In pilot tests, it was observed that very different results can be obtained if friction at the supports is not eliminated. The confinement of the shear fracture zone due to an axial friction force produced at the supports can no doubt raise the value of G_f significantly.

For the sake of comparison, another series of torsional tests of Mode III fracture was carried out with different equipment—a large triaxial torsional testing machine, which was very stiff because its 8.51 in. diameter test chamber is designed to resist chamber pressures up to 20,000 psi (138 MPa) plus axial forces up to 1,100.000 lb (4.89 mN). It was not surprising that the tests in this machine, which were made at essentially zero axial displacement at the ends of the cylindrical specimen, showed a different mode of failure, in which the failure surface was conical rather than planar (see Fig. 3). The axial restraint engenders an axial compressive force, which is known to be capable of altering the failure mode. The failure with a conical surface is nevertheless still of a shear type.

To explain the foregoing observations, it is logical to assume that G_f^{III} as well as G_f^{II} is not a material constant, but rather a material function $\phi(N_f)$ of the normal force N_f across the fracture process zone of length c_f, normal to the shear fracture plane (N_f is a force per unit length of the crack-front edge and has the dimension N/m). The function ϕ (N_f) should be the same for Mode II and Mode III shear fractures. For the same N_f,

Table 3 — Effective fracture process-zone length c_f, in. (mm)

Mix	Mode I	Mode III
Concrete	0.60 (15)	0.66 (17)
Mortar	0.06 (1.4)	0.04 (1.0)

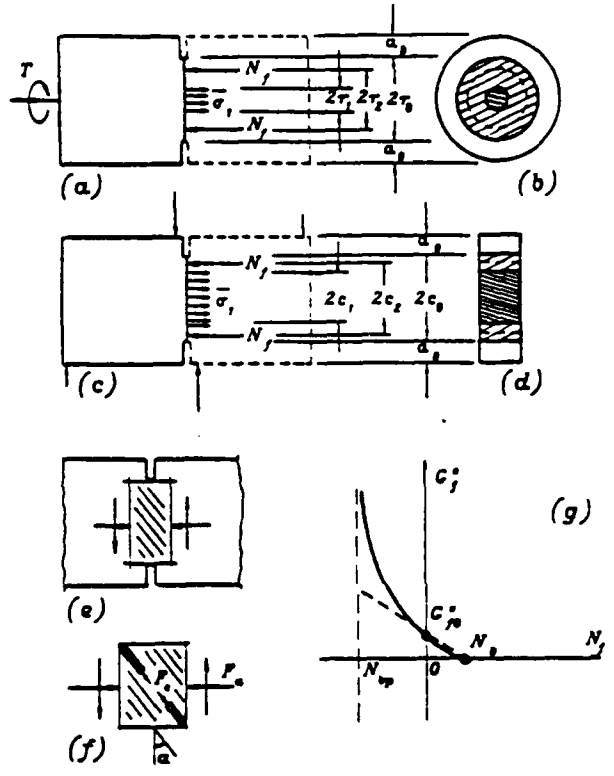


Fig. 8 — Transverse stresses and inclined microcracking induced in the fracture-process zone in shear-fracture tests

we should have $G_f^{II} = G_f^{III} = G_f^{I}$ = shear fracture energy (regardless of the mode) such that

$$G_f = \phi(N_f) \tag{5}$$

That G_f^{II} must in general strongly depend on N_f has already been corroborated in previous work[1] by finite element analysis of Mode II fracture tests. In fact, that study made the even more general assumption that both Mode I and II fractures can be modeled by the same stress-strain relation for the fracture process zone. The reason for different effective values of the fracture energy in Modes I and II can be found in the fact that, in shear, the microcracks in the fracture process zone are not parallel to the fracture plane, but are inclined to it by angle α [see Fig. 8(e) and (f)]. The column of intact material between adjacent inclined cracks carries compression force F_c. This force has an axial component F_s which must be resisted together by the tensile stresses in the undamaged portions of the ligament [σ_1 in Fig. 8(a)] and by the axial force provided by the support. If axial displacements are prevented, the inclined compressive forces F_c are high and can offer a large resistance to shear. Hence, the apparent values of frac-

ure energy must be large. If the axial displacements are free, the inclined compression forces must vanish and thus cannot contribute to transmit shear stresses across the fracture process zone, and then the apparent value of shear fracture energy is small.

The action just described was exhibited by a finite element model with a tensile softening stress-strain relation for the fracture process zone. Using such a model, Bažant and Pfeiffer[1] showed that the results of both Mode I and Mode II fracture tests can be matched with the finite element program using the *same* material properties. In this finite element analysis, one does not directly use the fracture energy. Rather, one uses a triaxial stress-strain relation such that the area under the implied uniaxial tensile stress-strain diagram equals G_f/w_c, where w_c = effective width of the fracture process zone.

As a crude approximate description of the role of the confining force N_f, we may write the longitudinal equilibrium condition assuming a zero axial resultant in the cross section. With the notation $\bar{\sigma}_1$ = average tensile stress in the tensile zone of the ligament and r_1 = radius of this zone, and the equilibrium condition is $\pi r_1^2 \bar{\sigma}_1 = -2\pi r_2 N_f$ or $N_f = -\bar{\sigma}_1 r_1^2/r_2$ where r_2 = radius up to the location of the compression force resultant N_f in the fracture process zone [Fig. 8(a) and (b)]. We may assume that $r_2 = k_2 r_1$ where k_2 = empirical constant greater than 1 but close to 1. Approximately, we may write $\bar{\sigma}_1 = k_1 f_t'$, where k_1 = empirical constant of the order of 1, perhaps $k_1 = 0.5$ to 1.0, and f_t' = direct tensile strength of concrete. Solving for N_f, we get

$$N_f = -\frac{k_1 f_t'}{2k_2} r_1 \tag{6}$$

For the previous Mode II fracture tests of Bažant and Pfeiffer,[1] a similar equilibrium condition for the fracture-plane cross section yields tensioned portion of the ligament [Fig. 8(c) and (d)]. Setting again $\bar{\sigma}_1 = k_1 f_t'$ we get

$$N_f = -c_1 k_1 f_t' \tag{7}$$

Let us assume that function $\phi(N_f)$ can be approximated linearly, i.e., $\phi(N_f) = G_f' - k_f N_f$, where G_f' = shear fracture energy at zero confining stress and k_f = nondimensional constant. For $N_f = c_f f_t'$ the fracture energy should become approximately zero, and so we have

$$\phi(N_f) = k_f (c_f f_t' - N_f) \tag{8}$$

Using Eq. (6) and (7) in Eq. (8), we obtain

$$\frac{G_f^{II}}{G_f^{III}} = \frac{c_f + k_1 c_1}{c_f + r_1 K_1/2K_2} \tag{9}$$

Since the fracture energy is defined for the limit case of an infinitely large specimen, we must assume that $c_f \ll c_1$ and $c_f \ll r_1$; therefore

$$\frac{G_f^{II}}{G_f^{III}} = 2k_2 \frac{c_1}{r_1} \tag{10}$$

According to the results of the tests made (Table 2), $G_f^{II}/G_f^{III} \approx 10$ for concrete and 3.3 for mortar. This ratio can be obtained from Eq. (10), given that $c_0 = r_0$, and if we assume intuitively that $k_2 = 1.5$, $c_1 = 0.7\, c_0$, and $r_1 = 0.21\, c_0$ for concrete; and $k_2 = 1.5$, $c_1 = 0.70\, c_0$, and $r_1 = 0.64\, c_0$ for mortar. These values do not appear to be out of the range of the reasonably expected behavior. So we may conclude that the difference between G_f^{II} and G_f^{III} is not all that surprising and might be explicable by a rational theory. But further tests as well as finite element studies will be required.

The previous analysis is based on the assumption of a linear dependence of G_f on N_f. In reality, this dependence must be expected to be nonlinear. Materials such as concrete exhibit a brittle-ductile transition at a certain confining pressure p_{bd}; for ordinary concrete, approximately, p_{bd} = 10,000 psi (69 MPa). Above this pressure value, there is apparently no strain softening and no fracture, and so $G_f \to \infty$. To satisfy this condition, the nonlinear dependence of G_f on N_f may be assumed to have the form

$$G_f = \phi(N_f) = \alpha_1 (N_f - N_{bd})^{-n} + \alpha_2 \quad (N_f < N_{bd}) \tag{11}$$

where α_1, α_2, and n = empirical positive constants and $N_{bd} = c_f P_{bd}$. Constants α_1 and α_2 may be determined from two conditions: (1) $G_f = 0$ for $N_f = c_f f_t'$, and (2) $G_f = G_{f_0}'$ for $N_f = 0$ where G_{f_0}' = shear fracture energy at zero normal force across the ligament. This yields

$$G_f = G_{f_0}' \frac{(N_f - N_{bd})^{-n} - (N_0 - N_{bd})^{-n}}{(-N_{bd})^{-n} - (N_0 - N_{bd})^{-n}} \tag{12}$$

where $N_0 = c_f f_t'$. This nonlinear form of the function $\phi(N_f)$ can considerably alter the values in Eq. (10) which correspond to the observed ratio G_f^{III}/G_f^{II}.

The foregoing analysis has one weakness in that the size-effect law in Eq. (1), which underlies the determination of fracture energy, might not be valid when the value of the normal force N_f across the fracture-process zone is different for various sizes. If this were so, a more sophisticated extrapolation to infinite size would be required to obtain the fracture energy value. It remains to be seen whether this aspect can significantly affect the present results for the practical size range.

CONCLUSIONS

1. The size-effect method can be used to determine the shear fracture energy of concrete. A cylinder with a

circumferential notch, loaded in torsion, is a suitable test specimen and yields consistent results.

2. As expected according to the size-effect law, the test results for the mortar specimens are much closer to the linear-elastic fracture mechanics than the test results for the concrete specimens.

3. Although according to the elasticity theory, the notched cylindrical specimen yields a perfect shear state (characterized by antiplane symmetry of the stresses near the crack-front edge), it does not yield such a perfect state in practice because of nonlinear effects. These effects result from the transverse confining normal stresses which are produced in the ligament cross section due to volume expansion from microcracking in the fracture process zone.

4. The value of the fracture energy of concrete obtained from Mode III tests is about three times larger than the Mode I fracture energy and about ten times less than the Mode II fracture energy obtained with the double-notched four-point-loaded specimens used in previous tests.[1] For mortar, it is about eight times larger than the Mode I fracture energy and three and one-half times less than the Mode II fracture energy.

5. The discrepancies between the Mode II and Mode III shear-fracture energy values seem to be explained by inclined microcracking in the fracture-process zone, the associated volume change, and the induced compressive force across the fracture front. Because of the influence of this force, the shear fracture energy for Modes II and III (unlike the Mode I fracture energy) is not a material constant but must be considered to be a material function of the confining normal force.

6. The fact that the relative difference between the Mode III and Mode I fracture energies is less for mortar than for concrete indicates that the confining force across the fracture process zone [Eq. (6)] is apparently less for mortar than for concrete. This might be explained by a smaller volume expansion of mortar.

ACKNOWLEDGMENTS

Development of the underlying theory was supported under AFOSR contract No. F49620-87-C-0030DEF with Northwestern University. Partial support for the tests of mortar specimens was received under a cooperative research program with Universidad Politecnica de Madrid funded under a U.S.-Spain Treaty (grant CCA-8309071). Interpretation of the results was partially supported from the Center for Advanced Cement-Based Materials at Northwestern University (NSF grant DMR-880 8432).

REFERENCES

1. Bažant, Z. P., and Pfeiffer, P. A., "Shear Fracture Tests of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 19, No. 110, Mar.-Apr. 1986, pp. 111-121.

2. Ingraffea, A. R., and Panthaki, M. T., "Analysis of Shear Fracture Tests of Concrete," *Proceedings*, U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures (Tokyo, May 1985), American Society of Civil Engineers, New York, 1987, pp. 151-173.

3. Bažant, Z. P., and Pfeiffer, P. A., "Comment on Ingraffea and Panthaki's Analysis of Shear Fracture Tests of Concrete," *Proceedings*, U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures (Tokyo, May 1985), American Society of Civil Engineers, New York, 1987, pp. 174-183.

4. Bažant, Z. P., and Prat, P. C., "Measurement of Mode III Fracture Energy of Concrete," *Nuclear Engineering and Design* (Lausanne), V. 106, 1988, pp. 1-8. Also, *Report* No. 87-12/428m, Center for Concrete and Geomaterials, Northwestern University, Evanston, 1987.

5. Bažant, Zdeněk P., "Size Effect in Blunt Fracture: Concrete, Rock, Metal," *Journal of Engineering Mechanics*, ASCE, V. 110, No. 4, Apr. 1984, pp. 518-535.

6. Bažant, Zdeněk P., and Pfeiffer, Phillip A., "Determination of Fracture Energy from Size Effect and Brittleness Number," *ACI Materials Journal*, V. 84, No. 6, Nov.-Dec. 1987, pp. 463-480.

7. Bažant, Zdeněk P.; Kim, Jin-Keun; and Pfeiffer, Phillip A., "Nonlinear Fracture Properties from Size Effect Tests," *Journal of Structural Engineering*, ASCE, V. 112, No. 2, Feb. 1986, pp. 289-307.

8. Bažant, Z. P., and Kazemi, M., "Determination of Fracture Energy, Process Zone Length and Brittleness Number from Size Effect, with Application to Rock and Concrete," *Report* No. 88-7/498d, Center for Concrete and Geomaterials, Northwestern University, Evanston, May 1988, 36 pp. Also, *International Journal of Fracture*, in press.

9. Suresh, S., and Tschegg, E. K., "Combined Mode I - Mode III Fracture of Fatigue - Precracked Alumina," *Journal of the American Ceramic Society*, V. 70, No. 10, Oct. 1987, pp. 726-733.

10. Neville, A. M., *Properties of Concrete*, 3rd Edition, Pitman Publishing, Inc., Marshfield, 1981, 779 pp.

11. Bažant, Zdeněk P., "Fracture Mechanics and Strain-Softening of Concrete," *Proceedings*, U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures (Tokyo, May 1985), American Society of Civil Engineers, New York, 1987, pp. 121-150.

12. Bažant, Z. P., "Fracture Energy of Heterogeneous Material and Similitude," *Preprints*, SEM-RILEM International Conference on Fracture of Concrete and Rock (Houston, June 1987), Society for Experimental Mechanics, Bethel, pp. 229-241.

13. Bažant, Zdeněk P., and Kim, Jin-Keun, "Size Effect in Shear Failure of Longitudinally Reinforced Beams," ACI JOURNAL, *Proceedings* V. 81, No. 5, Sept.-Oct. 1984, pp. 456-468.

14. Bažant, Z. P.; Sener, S.; and Prat, P. C., "Size Effect Tests of Torsional Failure of Concrete Beams," *Report* No. 86-12/428s, Center for Concrete and Geomaterials, Northwestern University, Evanston, Dec. 1986, 18 pp. Also, *Materials and Structures, Research and Testing* (RILEM, Paris), V. 21, No. 126, Nov. 1988, pp. 425-430.

15. Bažant, Zdeněk P., and Cao, Zhiping, "Size Effect in Punching Shear Failure of Slabs," *ACI Structural Journal*, V. 84, No. 1, Jan.-Feb. 1987, pp. 44-53.

16. Bažant, Zdeněk P., and Cao, Zhiping, "Size Effect in Brittle Failure of Unreinforced Pipes," ACI JOURNAL, *Proceedings* V. 83, No. 3, May-June 1986, pp. 369-373.

17. Tada, H.; Paris, P. C.; and Irwin, G. R., *The Stress Analysis of Cracks Handbook*, Del Research Corp., Hellertown, 1973.

18. Benthem, J. P., and Koiter, W. T., "Asymptotic Approximation to Crack Problems," *Methods of Analysis of Crack Problems*, G. C. Sih, Editor, Noordhoff International Publishers, 1972, Chapter 3.

19. Bažant, Zdeněk P., and Oh, B. H., "Crack Band Theory for Fracture of Concrete," *Materials and Structures, Research and Testing* (RILEM, Paris), V. 16, No. 93, May-June 1983, pp. 155-177.

SHOULD DESIGN CODES CONSIDER FRACTURE MECHANICS SIZE EFFECT?

By Zdeněk P. Bažant

Professor of Civil Engineering
Center for Advanced Cement Based Materials
Northwestern University-Tech 2410, Evanston, Illinois 60208-3109

Abstract.- The paper reviews recent theoretical results and analyzes experimental data on size effect in brittle failure, which causes release of potential energy in reinforced concrete structures. After summarizing the size effect law and explaining the novel concept of a brittleness number, the results of recent Northwestern University tests of diagonal shear failure, punching shear failure, torsional failure and pullout failure are discussed. These results, which were obtained on geometrically similar specimens with sizes differing over a broad range, are found to be in excellent agreement with the theoretical size effect law. The experimental evidence is much stronger than that which was previously obtained by analyzing a large amount of test results from literature, which were not obtained on geometrically similar specimens and were limited to a narrow size range. It is also found that the test data on diagonal shear disagree with Weibull-type theory of the size effect, thus strengthening the theoretical argument against the use of these statistical theories for concrete structures whose maximum load is much larger than the cracking initiation load. The test results indicate that the presently considered fracture mechanics size effect ought to be incorporated into the formulas for the contribution of concrete to the ultimate load capacity in brittle failure of concrete structures. It is shown that such formulas can be based on the brittleness number. However, its prediction in the absence of empirical data for the structure type of interest will require some further research.

Introduction

It has long been known that ultimate loads of concrete structures exhibit size effect. The classical explanation has been Weibull's weakest-link theory which takes into account the random nature of concrete strength (Weibull, 1939; Zaitsev and Wittmann, 1973; Mihashi and Zaitsev, 1981; Mihashi, 1983; and Carpinteri, 1986). However, for reasons briefly explained in the Appendix, it now appears that the statistical theory does not suffice to describe the essence of the size effect observed in brittle failures of reinforced concrete structures and plays only a secondary role. The main aspect of the size effect in this type of failures is det rather than statistical, and is due to the release of the stored potential energy of the structure into the front of the cracking zone or fracture. This phenomenon is properly described by fracture mechanics in its recently developed nonlinear formulation which takes into account the distributed nature of cracking at the fracture front.

The purpose of this paper is to analyze the existing evidence as well as review some recent results obtained at Northwestern University.

Mathematical Description of Size Effect

The size effect is defined by comparing the ultimate loads (maximum loads), P_u, of geometrically similar structures of different sizes. This is done in terms of the nominal stress σ_N at failure. When the σ_N-values for geometrically similar structures of different sizes are the same, one says that there is no size effect. The size effect represents a dependence of σ_N on the structure size (characteristic dimension), d. For two-dimensional similarity, $\sigma_N = P_u/bd$, and for three-dimensional similarity, $\sigma_N = P_u/d^2$; b = thickness of a two-dimensional structure. The characteristic dimension, d, may be defined as any dimension of the structure, e.g., the depth of a beam or its span, since only the relative values of σ_N matter.

According to plastic limit analysis, as well as elastic analysis with allowable stress or any theory that uses a failure criterion in terms or stresses or strains, σ_N is independent of the structure size. This can be illustrated, e.g., by the elastic and plastic formulas for the strength of beams in bending, shear or torsion (Bazant, 1984).

Another theory of failure, conceived by Griffith (1921) and introduced to concrete by Kaplan (1961), is fracture mechanics. It was Reinhardt (1981a,b) who proposed that fracture mechanics should be used to describe the size effect in diagonal shear failure. He also showed that the size effect of classical, linear elastic fracture mechanics agrees reasonably well with some tests, although later it was found that nonlinear fracture mechanics is necessary in general.

In the linear form of fracture mechanics, in which all the fracture process is assumed to be happening at a point -- the crack tip -- the size effect is the strongest possible. In the plot of log σ_N vs. log d, it is described by an inclined straight line of slope - 1/2 (Fig.1), provided that the cracks at the moment of failure of geometrically similar structures of different sizes are also similar. This is true regardless of the structure shape.

Concrete structures in reality exhibit a transitional behavior between the size effect of strength or yield criteria (i.e., no size effect), represented in Fig. 1 by a horizontal line, and the size effect of linear elastic fracture mechanics; see the curve in Fig. 1. This size effect is generally ignored by the current design codes, but it now appears very important.

The aforementioned transitional size effect can be most simply explained by considering uniformly stressed rectangular panels of different sizes, as shown in Fig. 2. Each panel is assumed to have a weak spot in the middle of the left side, from which fracture originates. For a brittle heterogeneous

material such as concrete, it is important to take into account a relatively large zone of distributed cracking at the fracture front. The size of this zone is not proportional to the structure size but is approximately related to the maximum aggregate size. In the simplest approximation, it may be assumed that the width, h, of the cracking band at the fracture front is approximately constant, independent of the structure size. One may also assume the length of the fracture, a, to be proportional to dimension d of the structure, i.e., a/d = constant.

Formation of a fracture with crack band of thickness h and length a may be imagined to release the strain energy of density $\sigma_N^2/2E$ from the cross-hatched area in Fig. 2 (E = elastic modulus of concrete). When the fracture extends by Δa, the additional strain energy that is released into the fracture front comes from the densely cross-hatched strip of horizontal dimension Δa. Obviously, the larger the structure, the larger is the area of the cross-hatched strip, which is given by $h\Delta a + 2ka\Delta a$ where k = empirical constant depending on the structure shape. Now it is important to note that in a larger structure the energy that is released into a certain small extension Δa of the fracture comes from a larger area. Since the energy dissipated by fracture per unit length (and unit thickness) is approximately equal to the fracture energy, G_f, which is a material property, the load for a larger structure must be less so that the total energy release from a larger area would be the same. Hence the size effect.

The strain energy released from the densely cross-hatched strip is $(h\Delta a + 2ba\Delta a)\ \sigma_N^2/2E$. Setting this equal to the dissipated energy $G_f b\Delta a$, one obtains the size effects law (Bazant 1984):

$$\sigma_N = \beta f_t'(1 + \beta)^{-\frac{1}{2}}, \qquad \beta = d/d_0 \tag{1}$$

in which β is called the brittleness number of the structure and f'_t represents the direct tensile strength. It must be kept in mind that Eq. 1 is only approximate, valid within a size range of up to about 1:20. For a broader size range, a more complicated formula would be required. Nevertheless, the accuracy and size range of Eq. 1 seems sufficient for most practical purposes.

Since from some viewpoints the length of the distributed cracking zone at fracture front is more important than the width, it is interesting to note that a similar derivation, with the same result (Bazant, 1984; Bazant and Kazemi 1989) can be made for a sharp line crack having a crack band of constant length at the front. For more complicated structural geometries, the foregoing type of reasoning gets difficult, however, Eq. 1 can be derived generally by dimensional analysis and similar arguments (Bazant, 1984). This general derivation rests on two basic hypotheses: (1) The propagation of a fracture or crack band requires an approximately constant energy supply per unit length and width of fracture, and (2) the potential energy released by fracture from the structure is a function of both the length of the fracture and the size of the cracking zone (fracture process zone) at the fracture front.

The transitional size effect curve in Fig. 1 is also obtained by numerical models of the microstructure, such as the random particle model. In this model, a system of aggregate particles is generated randomly and each large aggregate particle is considered as a finite element interacting with its neighbors through a contact element representing the contact zone (Bazant, Tabbara, Kazemi, and Pijaudier-Cabot, 1989). Furthermore, nonlocal finite element models in which localization of cracking is restricted to a zone of a certain minimum size, also exhibit the same transitional type of size effect, while ordinary finite element codes are incapable of representing it (Bazant and Lin, 1988; Bazant and Ozbolt, 1989).

Eq. 1 has been extensively verified experimentally for both fracture test specimens and reinforced concrete structures. In the case of test specimens, similarity of the fracture shape and length is enforced by providing geometrically similar notches in specimens of different sizes. In real concrete structures, from which notches are absent, Eq. 1 is applicable only under the following two additional hypotheses: (3) the failure modes (i.e., fracture shapes and length) of geometrically similar structures of different sizes are, at the moment of maximum load, also geometrically similar, and (4) the structure does not fail at crack initiation. But this is not a serious limitation; a good design practice of course requires the maximum load to be much higher than the cracking initiation load, and this is to some extent also enforced by design codes.

The characteristic of the failure process which gives rise to the size effect is the propagating nature of failure. In plastic limit analysis the failure is always nonpropagating, simultaneous, with all the parts of the structure moving simultaneously in proportion to one parameter. The typical characteristic of such failures is that the load-deflection diagram, after reaching the maximum load, exhibits a horizontal plateau. It can be shown in general that when the horizontal plateau is lacking, i.e., the load decreases after the peak with increasing deflection, the failure cannot be simultaneous but must be propagating. Propagating failures need to be generally described by fracture mechanics, and they always exhibit size effect (of the energy release type).

Brittleness Number

Distinctions between brittle and ductile failures have long been emphasized in concrete textbooks, however the meaning of brittleness has been left hazy, unquantified. The notion of brittleness is closely connected with the size effect. Brittleness increases with size. It can be generally shown

that in small structures the load decreases relatively slowly with deflection after the peak, while in a similar large structure the load drops rapidly after the peak, and in a very large structure the load-deflection curve even exhibits the so-called snapback instability in which the load-deflection curve becomes vertical after which the failure is dynamic. Recognizing this connection, various authors, including Gogotsi, et al. (1978), Homeny, et al. (1980) for ceramics in general, and Carpinteri (1982) and Hillerborg (1985) for concrete in particular, proposed brittleness to be quantified by some brittleness number depending on the structure size, d. Unfortunately, Gogotsi, Homeny, Hillerborg, Carpinteri's brittleness numbers are not independent of the structure geometry and thus cannot be used as universal, absolute characteristics of brittleness (e.g., a brittleness number equal to 2 could mean a very brittle behavior for one structure geometry and a very ductile behavior for another geometry). These numbers only allow comparing the brittleness of similar structures.

A universal, absolute measure of brittleness is offered by the size effect law (Bazant, 1984), although it is only approximate since the exact size effect law is not known. As already mentioned, the ratio $\beta = d/d_0$ may be taken as a brittleness number (Bazant, 1987; Bazant and Pfeiffer, 1987). The value $d=d_0$ (or $\beta = 1$) corresponds in the size effect plot of $\log \sigma_N$ vs. $\log d$ to the point where the horizontal asymptote for the strength or yield criterion intersects the inclined asymptote for the linear elastic fracture mechanics (Fig. 1). For $\beta \leq 1$, the behavior is closer to plastic limit analysis, and for $\beta \geq 1$ it is closer to linear elastic fracture mechanics. With a practically sufficient accuracy, the nature of structure response and the type of analysis may be characterized as follows (Bazant, 1987):

$$\left.\begin{array}{rcll} \beta & < & 0.1 & \text{plastic limit analysis} \\ 0.1 \leq \beta & \leq & 10 & \text{nonlinear fracture mechanics} \\ \beta & > & 10 & \text{linear fracture mechanics} \end{array}\right\} \quad (2)$$

For $\beta < 0.1$, the error of plastic limit analysis is < 4.7% of σ_N, and for $\beta >$ 10, the error of linear elastic fracture mechanics is < 4.7%, compared to the nonlinear fracture mechanics solution in which the finite size of the fracture process zone (cracking band) is taken into account. If an error under 2% is desired, then the nonlinear range must be expanded to $1/25 \leq \beta \leq 25$.

Let us now consider the determination of the brittleness number when size effect data are absent. Two formulas have been derived for this purpose, one based on matching the size effect law to the asymptote of plastic limit analysis, and the other one based on matching the size effect law to linear elastic fracture mechanics. They read

$$\beta = B^2 g(a_0) \frac{f_t'^2}{EG_f} \tag{4}$$

$$\beta = \frac{g(\alpha_0)}{g'(\alpha_0)} \frac{d}{c_f} \tag{5}$$

in which $g(\alpha_0)$ is the nondimensional energy release rate corresponding to the initial relative crack length $\alpha_0 = a_0/d$ according to linear elastic fracture mechanics, $g'(\alpha_0)$ is its derivative, and c_f is the effective length of the fracture process zone (defined for extrapolation to a specimen of infinite size), which is a material property.

The first formula, Eq. 4 (Bazant, 1987; Bazant and Pfeiffer, 1987), is more accurate for small sizes, and the second formula, Eq. 5 (Bazant and Kazemi, 1988), is more accurate for large sizes. With either formula, calculations of the brittleness number necessitates knowing the shape and length of the equivalent linear elastic crack at maximum load. This crack is precisely defined only by extrapolation to a similar structure of infinite size. For typical concrete structures it is not yet clear at present what

shape and length this equivalent crack should have. However, once this shape and length are known, $g(\alpha_0)$ and its derivative can easily be calculated with a linearly elastic finite element program, and could easily be tabulated for typical structures. For the calculation of B, the existing design formulas based on plastic limit analysis can of course be used.

It might also be that simple empirical formulas could be obtained for various typical structure shapes to determine the value of the transitional size d_0, e.g., in relation to the cross section dimension and the maximum aggregate size, d_a. Then the brittleness number immediately results as $\beta = d/d_0$.

In view of the universality of the size effect law in Fig. 1 and the brittleness number, it appears that a simple adjustment can introduce the size effect into the existing code formulas based on limit analysis. It suffices to take the existing formula for the nominal stress due to concrete at ultimate load, v_u, and replace it by the expression:

$$v_u(1 + \beta)^{-\frac{1}{2}} \qquad \left[\geq v_u^{\min} \right] \qquad (6)$$

Here we indicate that there might be a lower limit $v_u^{\min}$ on the nominal strength, since there can be a transition to some nonbrittle frictional failure mechanism (it seems, for example, that this might be the case in the Brazilian split-cylinder test).

Previous Tests of Brittle Structural Failures

After the size effect law has been formulated, much effort has been devoted to comparing and validating it on the basis of the test data in literature. The efforts were especially focused on the diagonal shear failure of longitudinally reinforced concrete beams (Bazant and Kim, 1984, 1985; Bazant and Cao, 1986; Bazant and Sun, 1987). The latter study included essentially all the experimental data that could be extracted from the

literature, consisting of 461 beam tests. After appropriately eliminating the effect of numerous other factors according to various known approximate formulas, the existence of a size effect has been clearly demonstrated. It was also shown that incorporation of the size effect law in Fig. 1 into the existing ACI or CEB-FIP design formulas for the contribution of concrete to the ultimate strength of beams in diagonal shear brings about a distinct improvement, reducing the coefficient of variation of the deviations of the test results from the design formula.

Unfortunately, however, the results of these studies have not allowed any strong conclusions. The reason has been that, even after filtering out the effects of other factors, the statistical scatter has been enormous. The tests have been done at various laboratories, on various concretes, and on beams of various geometries. Most tests did not include various sizes. Those few test series that did (Walraven, 1978; Bhal, 1968; Kani, 1967; Leonhardt and Walther, 1962; Taylor, 1972; Chana, 1981; Rüsch et al., 1962) did not include a sufficiently broad range of sizes and, most seriously, did not use geometrically similar specimens. The same is true of the latest and largest study of the size effect in diagonal shear presented by Iguro, Shioya, Nojiri and Akiyama (1984).

The lack of geometric similarity in the previous test series has been the most serious impediment against their exploitation for the present purposes. To extract information on the size effect, adjustments for all the other influencing factors had to be made first. But since the influences of various other factors are known only approximately, a considerable error is inevitably introduced by such adjustments. This causes enormous scatter, which obscures the underlying trend of the size effect (Bazant and Kim, 1984, Bazant and Cao, 1984, Bazant and Sun, 1984, Bazant and Sener, 1984).

Aside from unprestressed beams without stirrups, the previous studies of test data from the literature dealt also with prestressed beams without

stirrups and with unprestressed beams with stirrups. In the latter case, the portion of the carrying capacity due to stirrups is of course free of size effect as the stirrups fail in a ductile manner. However, despite considerable scatter, the analysis of the data showed that, in contrast to the current design approach, the carrying capacity due to stirrups (which exhibits no size effect) is not simply additive to that due to concrete (which does exhibit a size effect). Rather, the presence of stirrups appears to have a strengthening influence on the portion of the carrying capacity due to concrete.

Similar problems have been encountered in an attempt to evaluate the size effect from the existing data on punching shear failures and torsional failures although, for the latter, the test results of Hsu (1968), Humphrey (1957), and McMullen and Daniel (1975) provided at least a clear indication of the existence of a significant size effect.

To sum up, despite a clear revelation of size effect, the enormous scatter and narrow size range of the data from the literature has made it impossible to conclude which size effect theory is the correct one. For example, the present size effect law fits most data from the literature no better than a formula based on Weibull-type statistical theory. This state of affairs, for example, permits one to conclude on the basis of the test data of Iguro et al. (1984) that the Weibull-type theory should be acceptable for diagonal shear failures of beams, even though its use is in fact questionable on physical grounds as already mentioned.

Evidence from New Tests of Geometrically Similar Structures

In view of the aforementioned limitations of the existing experimental evidence, a systematic testing program of brittle failure of reinforced concrete structures has been initiated at Northwestern University. To keep the costs down, all the tests were on concrete with reduced-size aggregate

(maximum sizes 3/8 or 1/4 in.). The tests included the diagonal shear failure of beams without stirrups (Bazant and Kazemi, 1989), the punching shear failure of circular slabs reinforced at bottom surface (Bazant and Cao, 1987), the torsional failures of plain and longitudinally reinforced concrete beams (Bazant, Sener and Prat, 1988), and the pullout failure of reinforcing bars (Bazant and Sener, 1988). The results of these tests, whose details are given in the aforementioned articles, are shown as the data points in Fig. 3-6. The data for punching shear (Fig. 4) have the size range 1:4, and so have the data for torsion (Fig. 5a,b) and the data for pullout (Fig. 6).

The tests of diagonal shear consisted of two series. In the first series (Fig. 3a), in which the size range was 1:4 and in which the longitudinal bars were straight, it was found that the diagonal shear failure was accompanied by pullout failure of bars, marked especially for the smallest size. Therefore a second series (Fig. 3b) was conducted on beams in which the bars had right-angle hooks at the ends, which prevented the pullout. The second series had the size range of 1:16. The beam specimens were similar in two dimensions, i.e., they had the same thickness for all the sizes.

From Fig. 3-6 it is clear that there is a strong size effect. It is also noteworthy that in the logarithmic scales the size effect curve does not tend to level off at large sizes, as predicted by fracture mechanics. The optimal fits according to Bazant's law are shown as the solid curves, and it is seen that the agreement is quite good, especially in view of the inevitable statistical random scatter of concrete strength in brittle failures. While the aforementioned studies of test data from literature only confirm the existence of size effect but could not decide which formula for the size effect is the correct one, the comparisons in Fig 3-6 can be said to support Eq. 1. This is especially clear for the second series of the diagonal shear tests (Fig. 6b), thanks to its broad size range.

The second series appears to represent the first test results that show that Weibull-type statistical theory does not apply. According to this theory, the strongest size effect in two dimensions corresponds to a straight line of slope - 1/6, whose optimal fit to the present data is shown as the dashed line in Fig. 3b. One can see that this line clearly disagrees with the trend of the data and gives too weak a size effect for the large sizes. (This is true provided one takes the value of Weibull modulus as m = 12, in agreement with the results of uniaxial tests, and assumes the Weibull threshold strengths to be $\sigma_0 = 0$; if this threshold were larger than 0, the size effect would be even milder than shown by the dashed line in Fig. 3b, exhibiting an approach to a horizontal asymptote, as shown by the dotted curve.)

The essential parameter which determines the strength of the size effect is the transitional size d_0. This parameter represents a combination of a material property with the effect of structure shape. The tests in Figs. 3-6 show that its values vary greatly. From the present limited results, it does not seem possible to give a simple empirical formula for calculating d_0. Thus, parameter d_0 has to be predicted from Eq. 4 or 5 ($d_0 = d/\beta$). However, further research will be needed to determine the shape and length of of the equivalent linearly elastic crack at maximum load, which is needed to calculate $g(\alpha_0)$. Eventually, of course, geometrically similar tests will have to be made on real-size beams and slabs with full-size aggregate.

Other Uses of Size Effect Law and Ramifications

Knowledge of the size effect law is as useful for determining material fracture properties as it is for extrapolating laboratory data to real structures. The material fracture properties needed for finite element analysis of concrete fracture include the fracture energy and the effective length of the fracture process zone (alternatively, from these two properties

one can determine the critical cracktip opening displacement used as a fracture parameter in some models).

In laboratory fracture specimens of typical sizes, the problem is that the fracture process zone extends over a larger part of the cross section. Consequently, the size and shape of the fracture process zone is greatly influenced by the boundaries, i.e., the shape of the specimen. This makes it very difficult to get unambiguous results by the usual adaptations of linear elastic fracture mechanics. The problem, however, disappears when one extrapolates the results to a specimen of infinite size. The point is that, in a specimen of infinite size, the fracture process zone occupies a negligibly small part of the structure volume, and so most of the structure is elastic. As one basic result of linear elastic fracture mechanics, the asymptotic stress field around the crack tip is exactly the same for any structure geometry. Thus, the fracture process zone in an infinitely large specimen is exposed always to the same stresses at its boundary, and so it must always take the same size and shape, and develop the same stress and strain distributions. Consequently, the fracture energy as well as the effective length of the fracture process zone are constants, independent of structure shape. Thus, the definition of fracture energy and the effective length of the process zone are unambiguous, exact (Bazant, 1987).

The only problem is that the size effect law that we know (Eq. 1) is only approximate, which means that the fracture properties obtained on the basis of Eq. 1 are not exact. However, they should be sufficient for most practical purposes since Eq. 1 is applicable over a size range up to about 1:20. By matching the infinite-size extrapolations according to Eq. 1 to the formulas of linear elastic fracture mechanics, simple expressions for determining the fracture energy as well as the effective length of the process zone have been devised (Bazant, 1987; Bazant and Pfeiffer, 1987; Bazant and Kazemi, 1988). By testing fracture specimens of very different geometries, it was

demonstrated that the material parameter values obtained by the size effect method are indeed approximately independent of the structure shape.

The size effect method of determining material fracture properties is not only unambiguous, but also the simplest, since it requires one to measure only the maximum load values of specimens of various sizes, which can be obtained even with the most rudimentary equipment. Other methods require measurement of the post-peak softening response and unloading response, and necessitate a closed-loop displacement control which calls for more sophisticated equipment.

Another important question is the effect of aggregate size, and more generally the influence of concrete composition. In principle, the size effect law (eq. 1) is valid only for specimens of the same material, which also implies the same aggregate size distribution. While recognizing that it is strictly impossible to change aggregate size without changing other composition parameters of concrete, it was suggested that the value of d_0 should be roughly proportional to the maximum aggregate size d_a, and that the value of f'_t in Eq. 1 should be decreased as the aggregate increases. The formula $f'_t = f^0_t[1 + (c_0/d_a)^{1/2}]$ has been proposed for this purpose by Bazant (1985), in which f^0_t and c_0 = constants. But this formula does not take into account the effect of aggregate shape, which might be very important according to the recent tests by V. Saouma at the University of Colorado, Boulder (private communication, 1989). This aspect is especially important for concrete dams, for which introduction of fracture mechanics into the failure analysis is imperative.

The finite element codes for concrete structure which normally utilize the smeared cracking concept, generally exhibit no size effect, which is an unacceptable feature of these codes. The finite element codes based on linear elastic fracture mechanics represent too strong a size effect. The basic requirement of acceptability of a finite element code for concrete structures

is that the code should exhibit a transitional size effect. This is achieved either by codes based on nonlinear fracture process zones (e.g., Hillerborg's fictitious crack model, or the crack band model) or more generally by nonlocal finite element codes. To document it, Fig. 7 compares the results of a finite element code with nonlocal smeared cracking (Bazant and Lin, 1988) to test results.

Conclusions

1. While the hundreds of test data available in the literature exhibit so much scatter that they provide only a relatively weak and ambiguous evidence for the size effect, merely suggesting its existence, the new Northwestern University tests of geometrically similar specimens prove the existence of a strong size effect and also indicate its approximate law. These new tests include the diagonal shear failure of beams, punching shear failure of slabs, torsional failure of beams, and pullout failure of bars. In view of this evidence, the existence of size effect must be suspected for all other brittle failures of concrete structures, e.g., the failures of splices, anchors and anchorages, the shear failure of deep panels, the cryptodome failure of the top slab of a reactor vessel, the failure of pipes, and the failure of concrete dams.

2. The size effect observed in the Northwestern University tests agrees quite well with the size effect law proposed by Bazant (Eq. 1). This means that the size effect is caused by the fact that, the larger the structure, the greater is the release of stored potential energy of the structure into the front of a propagating fracture or crack band front (provided that the nominal stress due to load is kept he same), while about the same energy is required for a unit advance of fracture front, regardless of the size.

3. Although verification by tests of real-size beams with full-size aggregate will be necessary, it is already clear that the size effect due to

potential energy release ought to be introduced into all the ultimate load formulas for brittle failures in concrete design codes, as indicated in Eq. 6. This means that the existing formulas which are based on the crack initiation load (at which the present type of size effect is absent) should be either replaced or better supplemented by formulas based on ultimate load, at which the present size effect operates.

4. The size effect is inextricably linked to brittleness. The brittleness, formerly a hazy concept, may be quantitatively characterized by the brittleness number, β, which, in Bazant's definition, takes into account the effect of shape. Generally, the values of the ultimate strength contribution due to concrete (rather than steel reinforcement), as given by the existing codes, should be modified by dividing them with the factor $(1 + \beta)^{1/2}$. Further research, however, is needed to determine the transition size d_0 required for the calculation of β when test data for the structure geometry of interest are unavailable. For this purpose it will be necessary to know not only the final failure surface but also the shape and length of the principal crack (or cracking band) at the moment of maximum load.

5. The test results for diagonal shear of beams disagree with Weibull-type statistical theory, and thus support the theoretical arguments against the use of these series for concrete structures for which the ultimate load is much larger than the crack initiation load and a large fracture or cracking zone grows prior to attainment of the maximum load.

6. The basic criterion for acceptability of a finite element code for concrete structures ought to be that it must exhibit a transitional size effect between plastic limit analysis and linear elastic fracture mechanics.

Acknowledgement. - A grant from Reinforced Concrete Research Council to Northwestern University, which supported the preparation of the paper is gratefully acknowledged. The underlying basic research was supported initially under AFOSR contract F49620-87-C-0030 DEF with Northwestern

University, and later by the Center for Advanced Cement Based Materials at Northwestern University (NSF Grant DMR-8808432).

Appendix.-Is Weibull's Statistical Theory Applicable?

Traditionally, the size effect has been explained statistically, by randomness of strength. The failure load of a chain is determined by the the strength of the weakest link in the chain, and the size effect arises from the fact that the longer the chain, the smaller is the minimum strength that is likely to be encountered in the chain. This explanation, which is no doubt correct for the failure of long uniformly stressed bars, is described by Weibull's weakest link statistics. Weibull-type theories have later been extended to reinforced concrete structures in general. In these extensions, the probability of failure of a structure and the mean nominal stress at failure are:

$$\text{Prob}\ (P) = 1 - \exp\left\{ -\int_V \left[\frac{\sigma(P,\underset{\sim}{x})}{\sigma_0}\right]^m \frac{dV(\underset{\sim}{x})}{V_r} \right\} \tag{7}$$

$$\bar{\sigma}_N = \frac{1}{bd}\int_0^\infty [1-\text{Prob}(P)]dP \tag{8}$$

in which V = volume of the structure, V_r= constant, m and σ_0 = material constants representing the Weibull modulus and the threshold stress. The key to applications of Weibull theory is function $\sigma(P,x)$, representing the stress caused by load P at location x. Determination of this stress distribution is easy only for structures which fail at initiation of macroscopic crack growth (which is the case for most metallic structures, e.g. the failure of an engine pylon in an aircraft). In such a case, the stress distribution is calculated according to elasticity for a structure without any crack. Such a simplification has been widely used in various Weibull-type analyses of concrete structures, but it is invalid.

Due to reinforcement as well as the existence of strain softening, concrete structures do not fail at crack initiation. In fact, design codes require the failure to be sufficiently higher than the crack initiation load. Consequently, a concrete structure undergoes pronounced inelastic deformation and macroscopic crack growths prior to reaching the failure load. This causes severe stress redistributions such that the distribution $\sigma(P,x)$ at failure is very different from the elastic distribution at no crack. The aforementioned stress distributions are difficult to determine, but the near-tip asymptotic elastic stress field might be a good approximation at distances not too large and not too small from the tip of the macrocrack at the moment of failure. Now, due to singularity of this stress field, the stress values farther away from the tip of a microcrack are relatively small and make in Fig. 8 a negligible contribution to $\bar{\sigma}_N$ compared to the stresses in the volume of the fracture process zone around the macrocrack tip. Due to this singularity, the volume in which the strength values matter is quite small, usually a small part of the entire volume of the structure. This causes the statistical size effect to become weak.

In any event, the random statistical size effect comes only after the size effect caused by the stress redistributions, which is deterministic, would exist even if the strength showed no randomness, and is explained by the release of stored energy associated with the stress redistribution. If there is any statistical size effect, it can be manifested only by the remaining portion of measured size effect which is unaccounted for by the fracture mechanics size effect. But no such unaccounted portion is apparent from the existing tests.

To sum up, explaining the size effect in reinforced concrete structures only by Weibull-type statistical theory is incorrect. Nevertheless, a combination of this theory with the size effect law in Fig. 1 (Bazant 1984) might be an improvement.

REFERENCES

Mihashi, H. (1983), "A Stochastic Theory for Fracture of Concrete", in Fracture Mechanics of Concrete, ed. by F. W. Wittmann, Elsevier, Amsterdam - New York, pp. 301-340.

Hsu, T. T. C., "Torsion of Structural Concrete - Plain Concrete Rectangular Sections", Torsion of Structural Concrete (SP-18), American Concrete Institute, Detroit, Michigan, 1968, pp. 203-238.

Humphreys, R., "Torsional Properties of Prestressed Concrete", Structural Engineer, Vol. 35, No. 6, London, UK, 1957, pp. 213-224.

McMullen, A. E., and Daniel, H. R., "Torsional Strength of Longitudinally Reinforced Concrete Members of Rectangular Cross-Section", thesis presented to West Virginia University, at Morgantown, W. VA, in 1972, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Leonhardt, F., and Walther, R., "Beitrage zur Behandlung der Schubprobleme in Stahlbetonbau", Beton- und Stahlbetonbau (Berlin), March 1962, pp. 56-64 and June 1962, pp. 141.149.

Rusch, H., Haugli, F. R., and Mayer, H., "Schubversuche an Stahlbeton-Rechteckbalken mit gleichmassing verteilter Belastung", "Bulletin No. 145, Deutscher Ausschuss fur Stahlbeton, Berlin, 1962, pp. 4-30.

Kani, G. N. J., "How Safe Are Our Large Reinforced Concrete Beams?", ACI Journal, Proceedings V. 64, No. 3, March 1967, pp. 128-141.

Bhal, N. S., "Über den Einfluss der Balkenhohe auf Schubtragfahigkeit von einfeldrigen Stahlbetonbalken mit und ohne Schubbewehrung", Dissertation, Universitat Stuttgart, 1968, 124 pp.

Walraven, J. C., "The Influence of Depth on the Shear Strength of Lightweight Concrete Beams Without Shear Reinforcement", Stevin Laboratory Report No. 5-78-4, Delft University of Technology, 1978, 36 pp.

Bazant, Z. P., (1984), "Size Effect in Blunt Fracture: Concrete, Rock, Metal", Journal of Engineering Mechanics, ASCE, Vol. 110, No. 4, pp. 518-535.

Bazant, Z. P. (1987), "Fracture Energy of Heterogeneous Materials and Similitude", Proceeding, SEM-RILEM International Conference on Fracture of Concrete and Rock, eds. Shah, S. P., and Swartz, S. E., Houston, pp. 390-402.

Bazant, Z. P., and Cao, Z., (1987), "Size Effect in Punching Shear Failure of Slabs", ACI Structural Journal, Vol. 84, pp. 44-53.

Bazant, Z. P., and Sener, S., (1987), "Size Effect in Torsional Failure of Concrete Beams", Journal of Structural Engineering, ASCE, Vol. 113, No. 10, pp. 2125-2136.

Bazant, Z. P., and Kazemi, M. T. (1989), "Size Effect on Diagonal Shear Failure of Beams without Stirrups", Report No. 89-8/498S, Department of Civil Engineering, Northwestern University, Evanston; submitted to ACI Structural Journal.

Bazant, Z. P., and Kazemi, M. T., (1989), "Size Dependence of Concrete Fracture Energy Determined by RILEM Work-of-Fracture Method", Report No. 89-12/B623S, Center for Advanced Cement-Based Materials, Northwestern University, Evanston; submitted to International Journal of Fracture.

Bazant, Z. P., and Kim, J. K., (1984), "Size Effect in Shear Failure of Longitudinally Reinforced Beams", Journal of American Concrete Institute, Vol. 81, No. 5, pp. 456-468; discussion Vol. 82, No. 4, pp. 579-583.

Bazant, Z. P., and Lin, F.-B., (1988), "Nonlocal Smeared Cracking Model for Concrete Fracture", Journal of Structural Division, ASCE, Vol. 114, No. 11, pp. 2493-2510.

Bazant, Z. P., and Ozbolt, J. (1989), "Nonlocal Microplane Model for Fracture, Damage and Size Effect in Structures", Report No. 89-10/498N, Center for Advanced Cement-Based Materials, Northwestern University, Evanston.

Bazant, Z. P. and Pfeiffer, P. A. (1987), "Determination of Fracture Energy from Size Effect and Brittleness Number", ACI Materials Journal, Vol. 84, No. 6, pp. 463-480.

Bazant, Z. P., and Sener, S., (1988), "Size Effect in Pullout Tests", ACI Materials Journal, Vol. 85, pp. 347-351.

Bazant, Z. P., Sener, S., and Prat, P. C., (1988), "Size Effect Tests of Torsional Failure of Plain and Reinforced Concrete Beams", Materials and Structures, Vol. 21, pp. 425-430.

Bazant, Z. P., and Sun, H.-H., (1987), "Size Effect in Diagonal Shear Failure: Influence of Aggregate Size and Stirrups", ACI Materials Journal, Vol. 84, No. 4, pp. 259-272.

Carpinteri, A., (1982), "Notch-Sensitivity and Fracture Testing of Aggregate Materials", Engineering Fracture Mechanics, Vol. 16, No. 14, pp. 467-481.

Carpinteri, A., (1986), "Mechanical Damage and Crack Growth in Concrete", Martinus Nijhoff Publishers, Dordrecht.

Chana, P. S., (1981), "Some Aspects of Modeling the Behavior of Reinforced Concrete under Shear Loading", Technical Report No. 543, Cement and Concrete Association, Wexham Springs, pp. 22.

Gogotsi, G. A., Groushevsky, Y. L., and Strelov, K. K., (1978), "The Significance of Non-Elastic Deformation in the Fracture of Heterogeneous Ceramic Materials", Ceramugia International, Vol. 4, No. 3, pp. 113-118.

Griffith A., (1921), "The Phenomenon of Rupture and Flow in Solids", Philosophical Trans., Royal Society of London, Series, A, Vol. 221, pp. 163-198.

Hillerborg, A., (1985), "The Theoretical Basis of Method to Determine the Fracture Energy G_f of Concrete", Materials and Construction (RILEM, Paris), Vol. 18, No. 106, pp. 291-296.

Homeny, J., Darroudi, T., and Bradt, R. G., (1980), "J-Integral Measurements of the Fracture of 50% Alumina Refractories", Journal of American Ceramic Society, Vol. 63, No. 5-6, pp. 326-331.

Iguro, M., Shioya, T., Nojiri, Y., and Akiyama, H. (1985), "Experimental Studies on Shear Strength of Large Reinforced Concrete Beams under Uniformly Distributed Load", Concrete Library International, Japan Society of Civil Engineers, No. 5, pp. 137-154.

Kaplan, M. F., (1961), "Crack Propagation and the Fracture of Concrete", American Concrete Institute Journal, Vol. 58, No. 11.

Mihashi, H., and Zaitsev, J. W., (1981), "Statistical Nature of Crack Propagation", Section 4-2 in Report to RILEM TC 50 - FMC, ed. Wittmann, F. H.

Taylor, H. P. J., (1972), "The Shear Strength of Large Beams", Journal of Structural Engineering, ASCE, pp. 2473-2490.

Weibull, W., (1939), "Phenomenon of Rupture in Solids", Ingenioersvetenskapsakad. Handl., Vol. 153, pp. 1-55.

Zaitsev, J. W., and Wittmann, F. H., (1974), "A Statistical Approach to the Study of the Mechanical Behavior of Porous Materials under Multiaxial State of Stress", Proceedings of the 1973 Symposium on Mechanical Behavior on Materials, Kyoto, Japan. 105 p.

CAPTIONS

FIG.1 Size effect law.

FIG.2 Explanation of size effect due to stored energy release from the cross-hatched areas.

FIG.3 Test results of Bažant and Kazemi (1989) for size effect in diagonal shear failure of beams without stirrups; (a) reinforcement allowed, (b) reinforcement pullout prevented by hooks.

FIG.4 Test results of Bažant and Cao (1986) for size effect in punching shear failure of circular slabs.

FIG.5 Test results of Bažant, Sener and Prat (1989) on size effect in failure of unreinforced beams and longitudinal reinforced beams without stirrups.

FIG.6 Test results of Bažant and Sener (1988) on size effect in pullout failure.

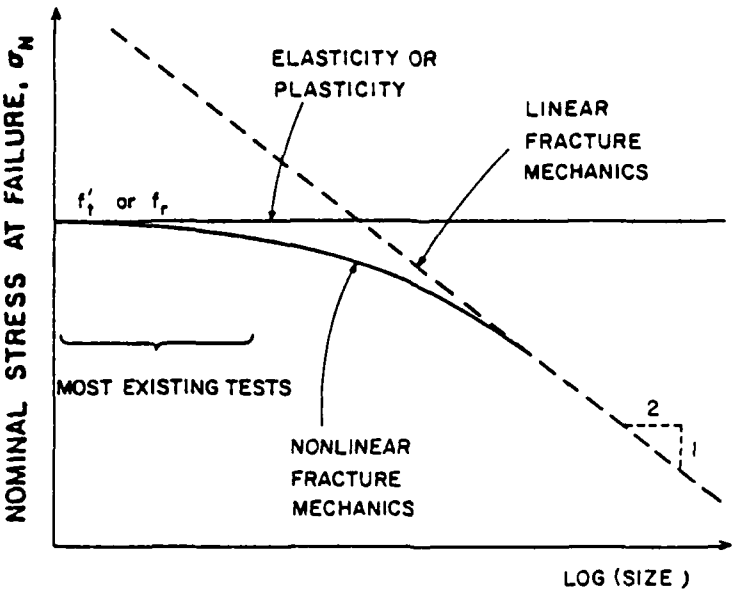

NOMINAL STRESS AT FAILURE, σ_N
ELASTICITY OR PLASTICITY
LINEAR FRACTURE MECHANICS
f'_t or f_r
MOST EXISTING TESTS
NONLINEAR FRACTURE MECHANICS
2
1
LOG (SIZE)

FIG. 1

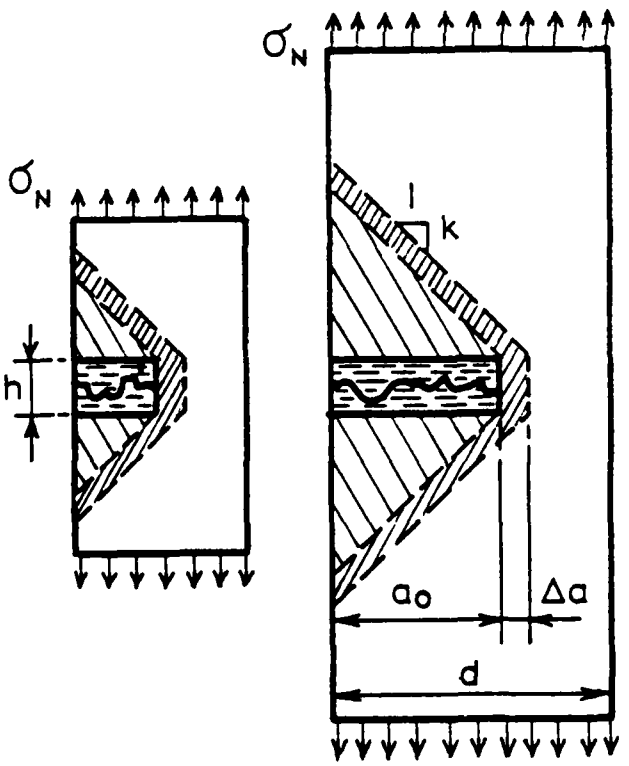

σ_N
σ_N
h
1
k
a_o
Δa
d

FIG. 2

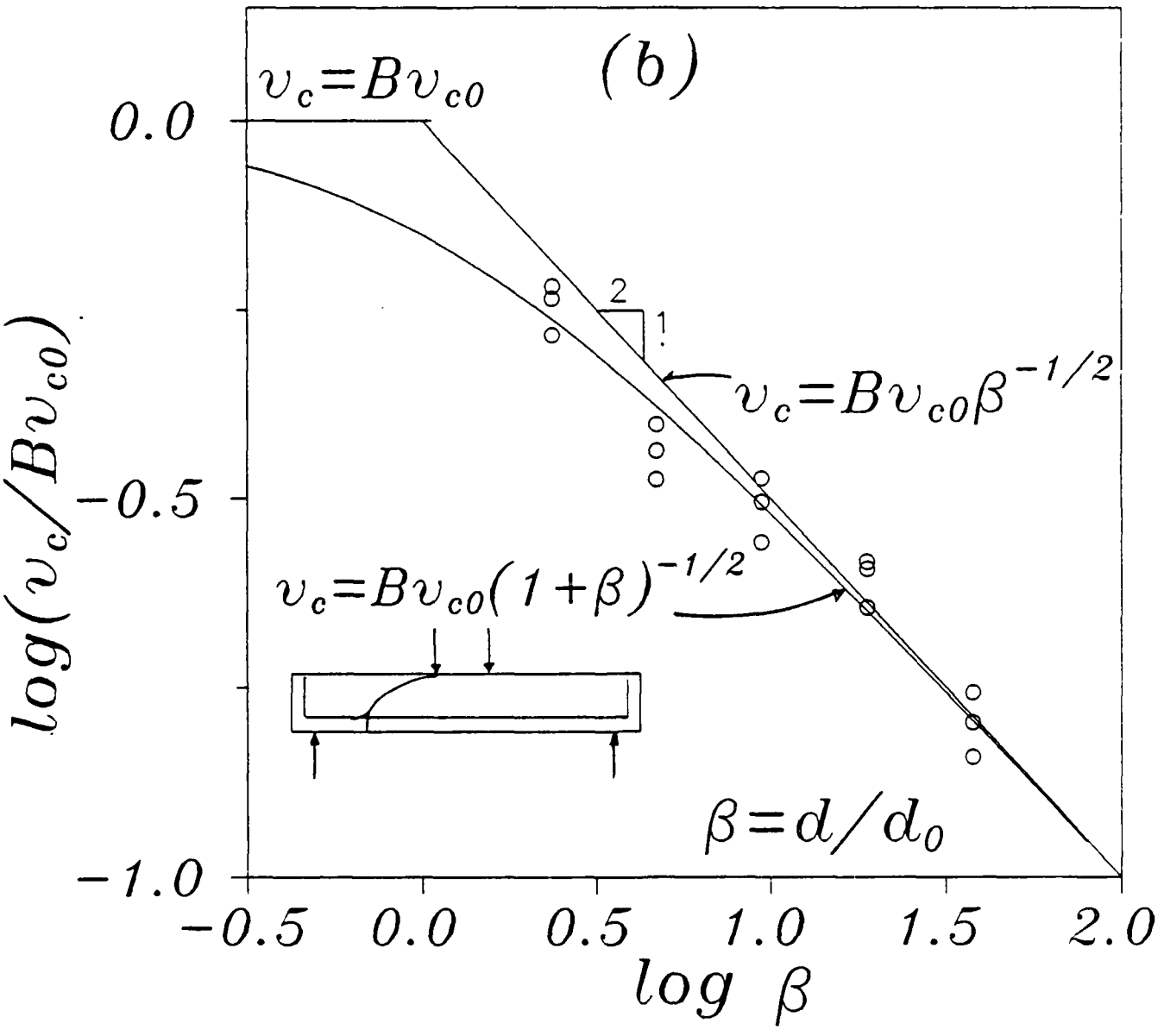

(b)
$\upsilon_c = B\upsilon_{c0}$
$\upsilon_c = B\upsilon_{c0}\beta^{-1/2}$
$\upsilon_c = B\upsilon_{c0}(1+\beta)^{-1/2}$
$\beta = d/d_0$
2
1
0.0
−0.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
$\log(\upsilon_c/B\upsilon_{c0})$
$\log\ \beta$

Fig.3 a

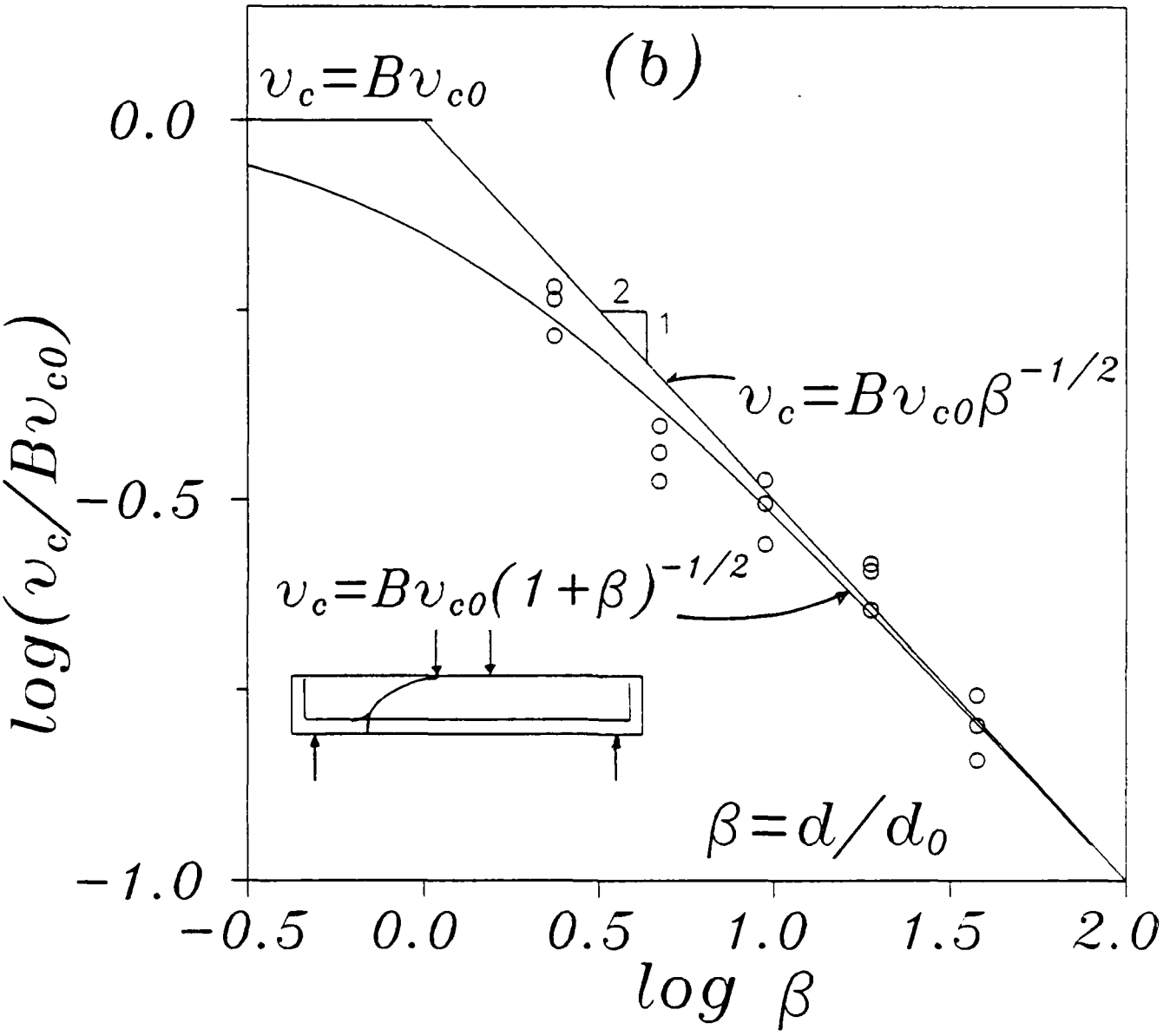

(b)
$v_c = Bv_{c0}$
$v_c = Bv_{c0}\beta^{-1/2}$
$v_c = Bv_{c0}(1+\beta)^{-1/2}$
$\beta = d/d_0$
2
1
0.0
−0.5
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
$\log(v_c/Bv_{c0})$
$\log\ \beta$

Fig.3 b

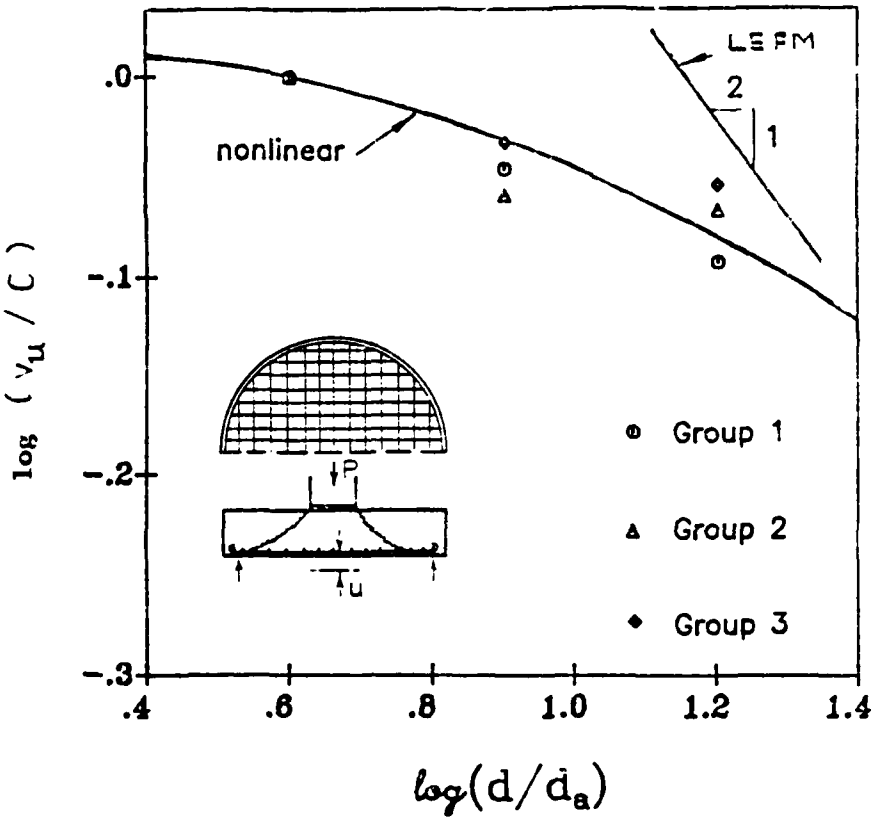

LEFM
2
1
nonlinear
.0
-.1
-.2
-.3
.4
.6
.8
1.0
1.2
1.4
log (v_u / C)
log(d/d_a)
P
u
Group 1
Group 2
Group 3

FIG. 4

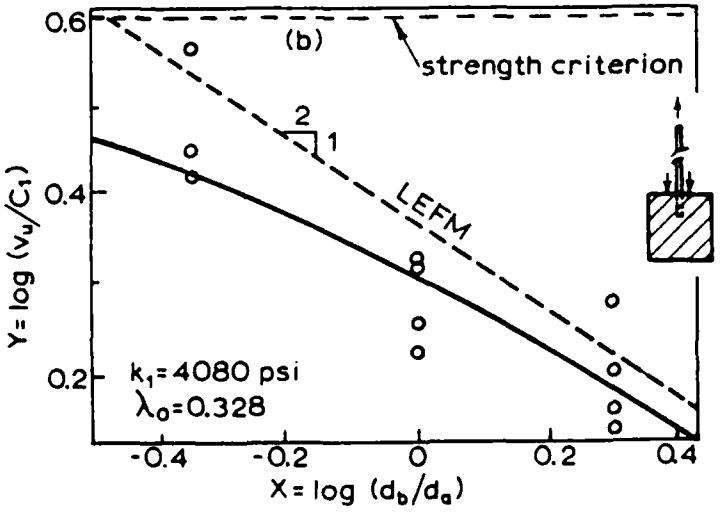

0.6
0.4
0.2
(b)
strength criterion
2
1
LEFM
Y= log (v_u/C_1)
k_1=4080 psi
λ_o=0.328
-0.4
-0.2
0
0.2
0.4
X= log (d_b/d_a)

FIG. 5

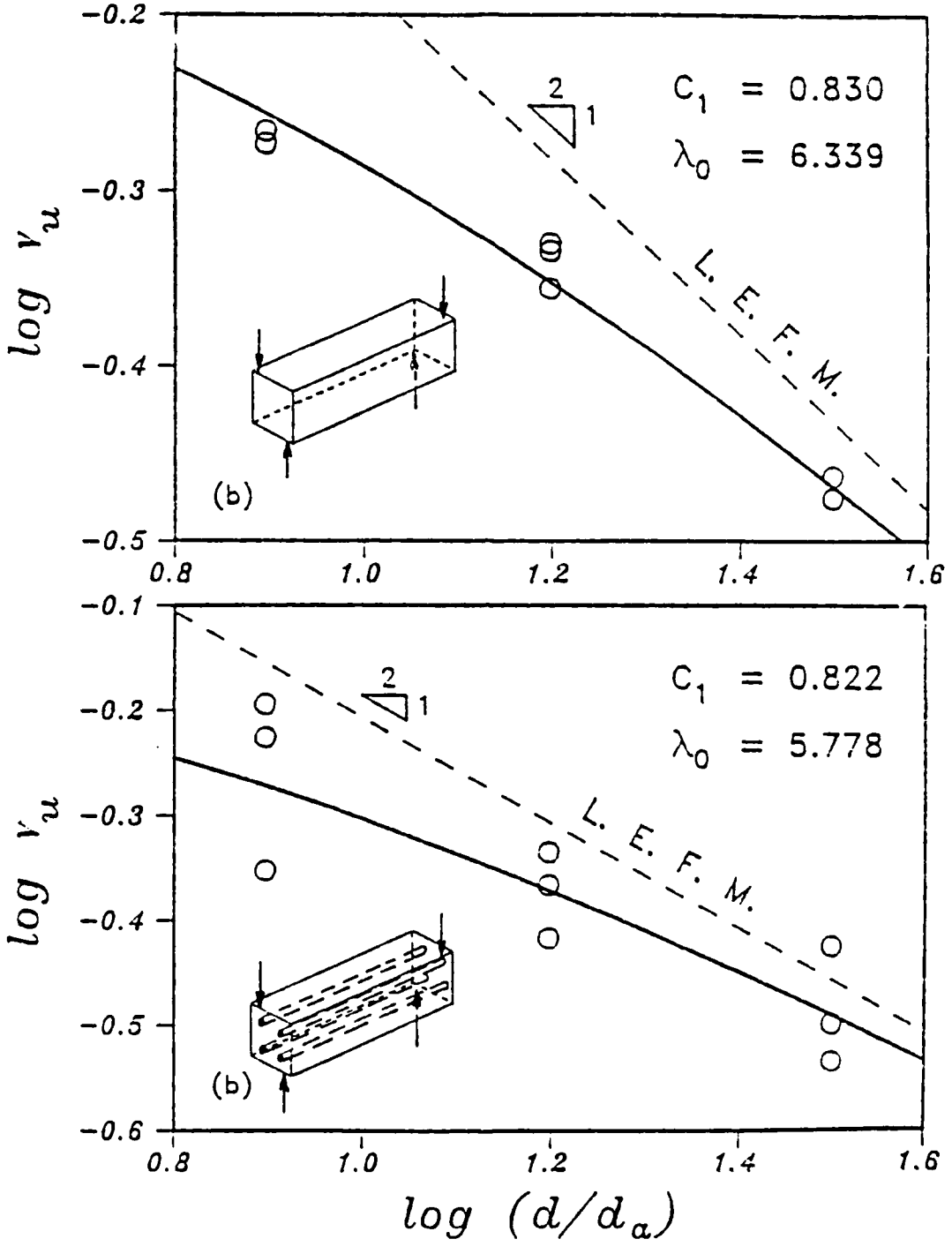

-0.2
-0.3
-0.4
-0.5
log v_u
2
1
C_1 = 0.830
λ_0 = 6.339
L. E. F. M.
(b)
0.8
1.0
1.2
1.4
1.6
-0.1
-0.2
-0.3
-0.4
-0.5
-0.6
log v_u
2
1
C_1 = 0.822
λ_0 = 5.778
L. E. F. M.
(b)
0.8
1.0
1.2
1.4
1.6
log (d/d_a)

FIG. 5

Journal of Engineering Mechanics
in press, 1990

RANDOM PARTICLE MODEL FOR FRACTURE OF AGGREGATE OR FIBER COMPOSITES

By Zdeněk P. Bažant[1], F. ASCE, Mazen R. Tabbara[2], S.M. ASCE, Mohammad T. Kazemi[2] and Gilles Pijaudier-Cabot[3], A.M. ASCE

ABSTRACT.- Presented is a particle model for brittle aggregate composites materials such as concretes, rocks or ceramics. The model is also applicable to the behavior of unidirectionally reinforced fiber composites in the transverse plane. A method of random computer generation of the particle system meeting the prescribed particle size distribution is presented. The particles are assumed to be elastic and have only axial interactions, as in a truss. The interparticle contact layers of the matrix are described by a softening stress-strain relation corresponding to a prescribed microscopic interparticle fracture energy. Both two- and three-dimensional versions of the model are easy to program, but the latter poses, at present, forbidding demand for computer time. The model is shown to realistically simulate the spread of cracking and its localization. Furthermore, the model exhibits a size effect on : (1) the nominal strength, agreeing with the previously proposed size effect law, and (2) the slope of the post-peak load-deflection diagrams of specimens of different sizes. For direct tensile specimens, the model predicts a development of asymmetric response after the peak load.

---

1 Professor of Civil Engineering, Center for Advanced Cement-Based Materials, Northwestern University, Evanston, Illinois 60208.
2 Graduate Research Assistant, Center for Advanced Cement-Based Materials, Northwestern University, Evanston, Illinois 60208.
3 Scientist, Laboratoire de Mécanique et Technologie, Ecole Normale Supérieure, Cachan, /CNRS/ Univ. Paris 6 France; formerly graduate Research Assistant at Northwestern University.

Introduction

It is now generally accepted that inelastic deformation and fracture of brittle heterogeneous materials such as concrete or fiber composites can be adequately modeled neither by classical (local) continuum material models nor by linear elastic fracture mechanics. Although nonlinear fracture models and nonlocal continuum models can no doubt go a long way towards a realistic description, they are inherently incapable of capturing those phenomena which are strongly affected by the randomness of material heterogeneity on the microscale and are localized to a small but non-negligible region. The continuum models can describe well the mean of the macroscopic material response but not its variance. Even when generalizations such as the stochastic finite element method are considered, the assumed spatial randomness can only be correct in the overall sense and cannot capture the effect of random local material inhomogeneities on the localization of damage and failure. It is for these reasons that a direct simulation of the random microstructure of these materials is useful. As demonstrated by Zubelewicz and Bažant (1987), a model simulating the microstructure can describe progressive distributed microcracking with gradual softening and with a large cracking zone (large relative to the aggregate size).

The principal objective of the present investigation is to study the effect of the size of the specimen or structure on the maximum load, the subsequent softening behavior, and the spread of the microcracking zone. These are the key questions of fracture modeling. We will model the material as a system of randomly configured circular particles. This model is intended to simulate the behavior of concrete, but can also be used for the behavior of unidirectional reinforced fiber composites in the transverse plane, with microcracks parallel to the fibers and

spreading in the transverse direction.

The model we will introduce will be a modification and refinement of that recently developed by Zubelewicz and Bažant (1987). The idea of particle simulation is an older one; it was proposed by Cundall (1971), Serrano and Rodriguez-Ortiz (1973), Rodriguez-Ortiz (1974) and Kawai (1980). These works dealt with rigid particles which interact by friction and simulate the behavior of granular solids such as sand. This approach was developed and extensively applied by Cundall (1978) and Cundall and Strack (1979), who called it the Distinct Element Method. An extension of Cundall's method to the study of microstructure and crack growth in geomaterials with finite interfacial tensile strength was introduced by Zubelewicz (1980, 1983), Zubelewicz and Mróz (1983), and Plesha and Aifantis (1983).

In this model we neglect the shear and bending interaction of neighboring particles throughout their contact layers. With such a simplification the present model becomes similar to the random truss model of Burt and Dougill (1977). These investigators verified that the truss model yields a realistic strain-softening curve but they did not study the fracture mechanics nor the size effect aspects. Furthermore, their method of system generation was not based on the idea of particles with a prescribed gradation.

Simplified Interaction Among Particles

Two adjacent circular or spherical particles i and j (Fig. 1a) are assumed to interact only in the axial direction, i.e., shear and moment interactions in the contact zone are neglected. As a generalization of the previous model by Zubelewicz and Bažant (1987), the particles are assumed to be elastic rather than rigid, characterized by elastic modulus E_a. The particles are embedded in a softer matrix representing the

portland cement paste in the case of concrete, or a polymer in the case of various fiber composites. The matrix is initially elastic, characterized by elastic modulus E_m. The stiffness of the connection between particles i and j is approximated as

$$S = (S_i^{-1} + S_m^{-1} + S_j^{-1})^{-1} \tag{1}$$

in which

$$S_i = E_a A_i / L_i, \quad S_m = E_m A_m / L_m, \quad S_j = E_a A_j / L_j \tag{2}$$

$$L_i = \gamma r_i, \quad L_j = \gamma r_j, \quad L_m = L - L_i - L_j \tag{3}$$

Here r_i and r_j are the radii of particles i and j, L is the distance of their centers, and the interaction is assumed to correspond to a strut (Fig. 1a) whose length is subdivided into segments L_i, L_j and L_m representing the particles and the interparticle contact layer of matrix. The lengths of these segments are determined by an empirical factor γ, for which the value $\gamma = 0.9$ has been used. A_i, A_j and A_m are the cross section areas of the strut segments describing the particles and the matrix; they are defined for the two-dimensional case of circles as

$$A_i = A_j = A_m = 2 \, \mathrm{Min}(r_i, r_j) \tag{4a}$$

and for the three-dimensional case of spheres as

$$A_i = A_j = A_m = \pi \, \mathrm{Min}(r_i^2, r_j^2) \tag{4b}$$

The strut segment L_m, representing the interparticle contact region of the matrix, is assumed to exhibit softening characterized by the triangular stress-strain relation in Fig. 1c. The softening behavior, which starts after reaching a certain specified strain ε_p for the peak stress, is determined by the fracture energy G_f^m of the interparticle layer (matrix and interface of matrix and particle) and is assumed to be a material constant. The fracture energy is equal to the area under the stress-strain diagram times length L_m (the fracture may, of course,

localize, but strain ε is defined as the average strain over length L_m, not as some local strain). Thus $G_f^m = L_m f_t^m \varepsilon_f/2$, with $f_t^m = E_m \varepsilon_p$ = strut strength. From this, the strain at the end of softening (at which complete interfacial crack forms) is

$$\varepsilon_f = \frac{2G_f^m}{L_m f_t^m} \tag{5}$$

In the numerical simulation which follows, the values used are such that $\varepsilon_f > \varepsilon_p$ even for the largest particle separations considered. However, for a small fracture energy or for interactions between particles that are sufficiently far apart, it could happen that $\varepsilon_f < \varepsilon_y$, which would mean that the response would exhibit a snapback instability. Although this would represent a complication for the numerical solution, it could be handled with minor modification.

The tangential softening modulus E_s (Fig. 1c) is defined by the relation $E_s(\varepsilon_f - \varepsilon_p) = -f_t^m$, from which

$$E_s = -\frac{f_t^m}{\varepsilon_f - \varepsilon_p} \tag{6}$$

If the length L_m is varied, the softening modulus E_s changes as shown by the dotted line in Fig. 1c, so as to preserve a constant value of G_f^m. In the case of unloading, the matrix material is assumed to follow a straight line parallel to the initial tangent. On reloading, this line is retraced up to the virgin stress-strain diagram, which is then followed at further increase of strain.

Let $\bar{u}_i$, $\bar{u}_j$ be the displacements at the ends of segment L_m in the direction ij. The strain in the matrix is

$$\varepsilon = (\bar{u}_j - \bar{u}_i)/L_m \tag{7}$$

Introducing the displacements u_i and u_j at particle centers i and j

(nodes) in the direction ij, we have

$$\bar{u}_i = u_i + \frac{P}{S_i}, \quad \bar{u}_j = u_j - \frac{P}{S_j} \tag{8}$$

substituting in Eq. 7, we get

$$\varepsilon = \frac{1}{L_m}\left(u_j - u_i - \frac{P}{S_j} - \frac{P}{S_i}\right) \tag{9}$$

in which P is the interparticle force. The strain given by Eq. 9 and its previous history makes it possible to calculate the stress in the strut according to the stress-strain relation for loading and unloading in the segment L_m.

Having neglected shear and moment interactions among particles, the particle system is equivalent to a pin-jointed truss and can be treated as a special case of a finite element system. The static response of the particle system to the prescribed loading is calculated in small displacement increments, in each of which the behavior of the connecting strut is described according to Eqs. 1-9. These equations are used to formulate the stiffness matrix of each connecting strut in the global coordinates, and these stiffnesses are assembled in each loading step into the structural stiffness matrix.

As an alternative, one can use the secant stiffness iteration (Owen and Hinton, 1980), which is based on the total force-displacement relations for the connecting struts. The solution algorithm proceeds in this case in the same manner as described for layered beam finite elements in Bažant, Pan and Pijaudier-Cabot (1987).

The formulation just outlined represents a simplification of the actual behavior on the microscale. The detailed growth of microcracks on a scale smaller than the aggregate sizes (i.e. the scale of grains in the matrix) is not modeled. Furthermore, the three-dimensional nonlinear

constitutive properties of the particle and matrix material are not taken into account. In this regard, however, we must realize that the physics of macroscopic phenomena generally depends only on the principal phenomena on the microscale. It is insensitive to various detailed phenomena and processes on the microscale, which tend to average out or cancel each other. Based on this experience, the description of the microscopic phenomena should always be greatly simplified if the purpose is a macroscopic model, but the simplification must be a realistic one. It seems that the present model might be adequate in this spirit.

Random Generation of Particle Configuration

An important aspect of the simulation is the generation of the random geometrical configurations of the particles which satisfies the basic statistical characteristics of the real material, such as concrete. The random configuration must be macroscopically homogeneous in space and macroscopically isotropic. The configuration must meet the required granulometric distribution of the particles of various sizes, as prescribed for the mix of concrete. The problem of random generation of the configurations with such constraints has various difficult and sophisticated aspects. These are for example discussed by Plesha and Aifantis (1983), who also review various pertinent mathematical theories.

After experimenting with various alternatives, the following simple procedure of random generation has been adopted:

1) The volume ratio for each particle size is specified in advance.

2) Using a random number generator from a standard computer library, generate the pairs of coordinates (x_i, y_i), $i = 1,2,\ldots$, of the particle centers (or coordinate triplets, in 3D), one pair after another, assuming a uniform probability distribution for these points throughout the specimen. This is particularly easy for specimens with rectangular

boundaries, for which the probability distribution of each coordinate is independent and uniform over a line segment, and equally easy for specimens which can be subdivided in a finite number of identical elements (e.g. squares or cubes).

3) For each generated coordinate pair (or triplet), check for possible overlaps of the new particle with the previously placed particles and with the specimen boundaries. Reject the coordinate pair if an overlap occurs.

4) Random generation of the coordinate pairs proceeds until the last particle of the largest size has been placed within the specimen. Then the entire random placement process of the particles is repeated for the particles of the next smaller size, and then again for the next smaller size, until the last particle of the smallest size has been placed within the specimen.

A different approach, which has been used, e.g., by Cundall (1978), is to actually simulate the production of the material. In the case of sand, this approach tries to simulate the pouring of the sand into a form. The particles are taken one by one and dropped vertically at random locations in the horizontal plane onto the heap of the sand already deposited. After coming in contact, each particle is allowed to roll or slide to a stable position. This approach produces a dense packing, which is no doubt quite realistic for sand. However, the random particle system obtained is not isotropic. Moreover, the programming is considerably more sophisticated than that for the presently used procedure. This simulation also ignores the fact that, in a material such as concrete or a fiber-reinforced polymer, each particle is wrapped in a layer of adhering matrix, which causes that perfect contacts are less likely than particle locations with some separation.

A salient feature of the present model is the existence of a contact layer of matrix separating the adjacent random particles. This makes it possible to model interaction between particle pairs that are not in contact. This feature has been handled in the manner previously introduced by Zubelewicz and Bažant (1987), in which each circular particle is assumed to be surrounded by a circular influence zone of a larger radius βr where r is the radius of the particle (Fig. 1a). After some numerical experimentation, the value $\beta = 1.65$ has been adopted. The particles are assumed to interact (i.e. the connecting strut to exist) only when their influence zones overlap. When these zones do not overlap, the particles are assumed to have no interaction (i.e., no connecting strut exists). It should be noted that the use of the influence zone greatly simplifies the random generation of the particle locations and makes the aforementioned procedure possible.

There is one limitation which is caused by the hypothesis that the particle interactions are only axial. The average Poisson's ratio ν for the initial elastic deformations cannot be specified as measured in experiments but is implied by the specified characteristics of the random particle system. For a random truss representing a system of particles of one size, the value of ν for a very large number of particles asymptotically approaches 1/3 for the two-dimensional case, and 1/4 for the three-dimensional case. This is a well-known property of large macro-homogeneous lattices. When the particle sizes are unequal the value of ν can be different, but for the present two-dimensional calculations it has been close to 1/3. This value of the Poisson's ratio is too high for concrete. To obtain a realistic value (approximately 0.18), one could include shear interactions, as in Zubelewicz and Bažant (1987) although other measures might be possible.

Numerical Studies

1) Response of Unnotched Specimens

The response of various unnotched concrete specimens has been calculated and compared to what is known from testing and other analytical studies. Examples of the randomly generated particle configurations are seen in Figs. 1b and 2. Fig. 1b also shows the truss to which the particle model is equivalent.

Geometrically similar specimens of various sizes, with depths = $3d_a$, $6d_a$, and $12d_a$ where d_a = maximum particle size (12mm), have been analyzed. The length of the specimens is always L = 8d/3. The diameters of the particles are 12, 8, 5, 3 and 1 mm. The volume ratio of particles for each diameter is selected to model concrete. Details of the number of particles for each size are given in Table 1.

1a) Size Effect. - A salient consequence of heterogeneity, which has fracture mechanics as well as probabilistic aspects, is the size effect on the failure load. If the failure is governed by criteria in terms of stress or strain (yield criteria), then geometrically similar specimens of different sizes must fail at the same value of the nominal stress at failure σ_N , defined as P_u/bd, where P_u = maximum load, d = depth of specimen, and b = thickness. To test this property, the specimens shown in Fig. 2 were subjected to prescribed uniform longitudinal displacement u at one end (Fig. 3a) and restrained against displacements at the opposite end. Displacement u has been incremented in small steps, and the axial force resultant P calculated for the following material properties: G_f^m = 24 N/m, f_t^m = 3 MPa, E_m = 30 GPa, and $E_a = 6E_m$. The diagrams of load P vs. load-point displacement u have been constructed; they are plotted in Figs. 3b, c, d for the small, medium and large specimens. Fig. 3e shows the load-deflection curves for specimens 1A, 2A and 3A scaled so

that the peak point would coincide with (1,1). The figure reveals that the post-peak declining slope gets steeper as the specimen size increases.

The data points for maximum loads are plotted in Fig. 4 in terms of stress σ_N as a function of the specimen depth d. The nominal stress is normalized with respect to f_u, an arbitrary measure of material strength, taken here as $f_u = f_t^m = 3$ MPa. The depth is normalized with respect to d_a, the diameter of the largest particle (12 mm). We see considerable scatter; therefore, the values of the nominal stress are averaged for each specimen size and are plotted as stars in Fig. 4. These plots clearly reveal that there is a size effect.

The size effect for materials that exhibit progressive cracking is known to approximately follow the size effect law proposed by Bažant (1984):

$$\sigma_N = \frac{Bf_u}{\sqrt{1 + d/d_0}} \tag{10}$$

where $d_0 = \lambda_0 d_a$ and B, λ_0 = empirical parameters. This law, which is approximate but usually applicable to size ranges up to 1:20, represents a gradual transition between the strength (or yield) criterion, for which there is no size effect on nominal strength, and linear elastic fracture mechanics, for which the size effect is the strongest possible ($\sigma_N \propto d^{-1/2}$). For data fitting it is convenient to transform Eq. 10 to the linear regression plot $Y = AX+C$, in which $X = d/d_a$, $Y = (f_u/\sigma_N)^2$, $C = 1/B^2$, and $A = C/\lambda_0$.

For comparison, the lines representing the size effect according to the strength (or yield) criterion and according to the linear elastic fracture mechanics are also shown in the Fig. 4b. From this it is evident that the size effect obtained is intermediate between the strength

criterion and the linear elastic fracture mechanics. This may be expected on the basis of the fact that the specimens do not fail at cracking initiation but only after a crack band has already developed, in which case nonlinear fracture mechanics should be applicable.

1b) Progressive Spread of Cracking. - Fig. 5 shows for the various specimens their microcrack patterns at the last calculated point on the load-deflection curve. Fig. 6c shows these patterns as they develop in specimen 1A at various stages of loading corresponding to points 1, 2, 3 and 4 shown in Fig. 6a. The fully formed and open cracks (for which the stress is reduced to 0, $\varepsilon > \varepsilon_f$) are shown by the solid lines, and the partially formed, developing (active) cracks which correspond to strain-softening states are shown by the dashed lines. Also shown are the previously formed cracks which are getting unloaded (Fig. 6b); these are represented by two stars. We see that the cracks first start at many random locations throughout the specimen, but later many of these partially formed cracks unload, and only some of them, lying in a narrow transverse band, open further and lead to the final fracture.

The unloading behavior is further documented in Fig. 7 which shows the evolution of the strain determined from the line segments 1 and 2. The first line segment always undergoes loading, crossing the developing crack band, while the second line segment undergoes partial softening but soon reverses to unloading (Fig. 7b).

1c) Scatter of Stress Profile and Symmetry Breakdown. - The distribution of the longitudinal component of the interparticle forces P_1 throughout the cross section at distance $x = 60$ mm from the fixed end is sketched, for specimen 1A, at various stages of loading in Fig. 8. Similar plots for specimen 3A at two locations $x = 160$ mm and $x = 240$ mm, are shown in Fig. 9. We see that the nonuniformity in the distribution is

getting more pronounced at later stages of loading. Furthermore, we see that the resultant of the longitudinal force tends to shift away from the center line, which indicates a tendency for these specimens to follow an asymmetric deformation pattern (a similar asymmetry is in fact indicated by analysis of the stable paths of fracture propagation). The eccentricity, e, of the force resultant, P, is plotted as a function of P for two specimens (1A and 1B) in Fig. 10, in which the tendency towards an increasing eccentricity is clearly seen. Development of such asymmetric response was observed in unnotched tensile specimens in the experiments of Bažant and Pijaudier-Cabot (1989). A similar asymmetry was seen to develop in the experiments of Rots, Hordijk, and de Borst (1987).

1d) Boundary Layer. - It is also apparent from Figs. 8 and 9 that the largest longitudinal forces tend to be transmitted through the interior of the specimen, and the boundary layer on the average transmits lower forces than the interior. This is a manifestation of what is known in concrete technology as the "wall effect". The boundary layer has on the average a smaller content of the stiff particles and a higher content of the soft matrix than does the interior of the specimen. This is inevitable when the material is cast into a form. On the other hand, if the specimen were cut out from a larger block of the material, then the boundary layer would have on the average exactly the same composition statistics as the interior. However, there would still be some wall effect, although probably a much smaller one, due to the fact that transverse stresses are transmitted across longitudinal straight lines in the interior, while transmission of the transverse stresses is interrupted at the surface.

1e) Macroscopic Poisson's Ratio. - To determine the Young's elastic modulus E and the Poisson's ratio of the equivalent (smeared, homogenized)

elastic continuum, the large specimen 3A was loaded in uniaxial tension by prescribing a very small uniform longitudinal displacement (u=0.004mm) at one end and restraining the specimen against displacement at the other end. Free sliding was allowed at the ends in the transverse direction. The displacements in the longitudinal direction, u_x, and the transverse direction, u_y, are shown in Fig. 11 for the nodes with the maximum size aggregates. Furthermore, in order to minimize the effect of boundary conditions, only nodes within an interior region ($0.8d \times 0.8L$) were considered.

Fig. 11a shows for the case of mean uniaxial stress $\sigma_x = 0.39$ MPa the linear regression line for u_x as a function of the x-coordinate; its slope represents the mean (macroscopic) strain in the x-direction, $\varepsilon_x = 0.1154 \times 10^{-4}$, from which E = 33.8 GPa. A similar regression line in Fig. 11b gives the strain in the y-direction, $\varepsilon_y = -0.4234 \times 10^{-5}$. The macroscopic Poisson's ratio is then obtained as $\nu = -\varepsilon_y/\varepsilon_x = 0.37$. This value is close to the value of $\nu = 1/3$, which is the theoretical value for a very large two-dimensional random lattice that is statistically uniform and isotropic. Neither the calculated value $\nu = 0.37$ nor the theoretical value $\nu = 1/3$ for very large specimens is realistic for most particulate composites, including concrete, for which, typically, $\nu = 0.18$. The correct Poisson's ratio of real particulate composites cannot be attained with the present model. The reason for this limitation is that only axial interactions between the particle centers are taken into account while shear stresses in the contact zones are neglected.

2) Response of Notched Specimens

As we have seen, cracking in unnotched specimens that are initially stressed uniformly develops quite randomly. This prevents it to bring to light fracture properties of the particle system. To see these

properties, notched specimens have also been simulated with the particle model. The notches were mathematically modeled by first generating the random particle configuration as if no notch existed, and then severing all the interparticle connections that cross the line of the mathematical notch depth, $a_0 = d/6$ ($\alpha_0 = a_0/d = 1/6$).

Fracture properties may be best determined by studying the size effect, since this effect represent the most important consequence of fracture mechanics. To this end, we analyze geometrically similar three-point-bend fracture specimens of three different sizes whose ratios are 1:2:4 . Thus, three new specimens were generated as shown in Fig. 12.

To save computer time, the truss modeling of the random particle system covered only the central region of the beam near the notch while the end portions of the specimen were modeled by a regular truss with linear elastic properties. This simplification has a negligible effect, on the values of maximum loads, as long as all the cracking occurs only outside the regular truss and far from its boundaries. Calculations verified that this was indeed the case. For the three-point-bend specimens the random particle region width was taken to be equal to $8d_a$(96 mm) which coincided with the span of the smallest size specimen. Thus, no truss was added to a specimen of this size .

In the calculations of maximum loads, three different sets of matrix properties have been considered: (1) $G_f^m = 24$ N/m, $f_t^m = 3$ MPa, (2) 24 N/m, 1.5 MPa and (3) 240 N/m, 3 MPa with $E_m = 30$ GPa and $E_a = 6E_m$ for all cases. For each set, the regression line (plotted in Fig. 13a) yields the values of slope A and intercept C, from which the values of B and λ_0 follow, see Fig. 13b. The size effect curve according to Eq. 10 is also plotted, along with all the calculated points for the different sets of matrix properties. The results show a good agreement with Eq. 10.

However, for the aforementioned values of material properties, the results are closer to linear elastic fracture mechanics than the test results obtained for similar specimens by Bažant and Pfeiffer (1987).

From the size effect on maximum loads, one can determine the fracture energy G_f of the idealized material represented by the random particle system, using the formula (Bažant and Pfeiffer, 1987) :

$$G_f = \frac{g(\alpha_0)}{AE} f_u^2 d_a \tag{11}$$

Here E = Young's modulus (macroscopic), and $g(\alpha_0)$ is the nondimensional energy release rate according to linear elastic fracture mechanics; $g(\alpha_0)$= 6.07 for the present specimens. On the basis of Eq. 10, one can also determine the size of the fracture process zone according to the formula (Bažant and Kazemi, 1988):

$$c_f = \frac{g(\alpha_o)}{g'(\alpha_o)} \lambda_o d_a \tag{12}$$

where $g'(\alpha_0)$ is the derivative of $g(\alpha)$ which is evaluated at α_0= 1/6, and is equal to 35.2 for the present three-point-bend specimens.

Table 2 gives the values of G_f and c_f calculated for the three sets of matrix properties. Taking property set (1) as reference, the results from (2) show that a decrease of interparticle strength f_t^m will decrease G_f and increase c_f. However, results from (3) show that an increase of G_f^m will increase both G_f and c_f. Table 2 also includes the values obtained by Bažant and Kazemi (1988) from experiments on mortar and concrete. These values are in the same range as calculated here and could be matched closer by adjusting the material properties of the model.

Numerical experience showed that the calculated maximum loads can be sensitive to the random arrangement near the notch tip and the way the notch is modeled. The length of the notch cut by a saw, which is simulated by severing the particle connection intersecting the notch, is

uncertain to the extent that the notch tip can be any where between the last severed connection and the first unsevered connection ahead of the tip. Furthermore, one should not delete a particle whose center falls into the width of the notch cut, since this can greatly distort the results. This means that the notch cut should be considered as a line rather than an actual cut of finite width.

Conclusions

1. Random generation of a particle system with a given particle size distribution can be accomplished quite easily, due to the fact that the particles are circular (or spherical) and are not in contact but are separated by contact layers whose thickness is unspecified except for a given maximum thickness value.

2. Modeling only the axial interactions of neighboring particles seems sufficient to obtain a realistic picture of the spread of cracking and fracture in concrete. The neglect of shear interactions might cause the fracture process zone to be shorter than the value reported from test on concrete specimens. However, the length of this zone can be increased by increasing the value of the interparticle fracture energy (G_f^m) or by decreasing the value of the interparticle strength (f_t^m).

3. The model can describe realistically the post-peak declining load-deflection diagram, which is predicted to get steeper as the specimen size increases.

4. In contrast to local continuum models, the present model is capable of describing the size effect on the nominal strength of unnotched specimens as well as notched specimens.

5. For direct tensile specimens, the model predicts development of asymmetric response after the peak load, such that the specimen bends and the axial force resultant becomes eccentric. This agrees with laboratory

tests.

6. Some improvements of the model might be appropriate in : (1) determining the cross section areas of the truss members so as to make them dependent on the distance between particles, (2) controlling the poisson's ratio without having to introduce shear interaction, (3) developing a method to connect the particle system in the cracking zone to a finite element mesh around this zone. This would significantly reduce the number of particles needed to analyze three-dimensional or large two-dimensional problems.

Acknowledgment

Partial financial support under AFSOR contract F49620-87-C-0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu, is gratefully acknowledged. Partial support for fracture studies was also received from the Center for Advanced Cement-Based Materials at Northwestern University (NSF Grant DMR-8808432). Supercomputer time on CRAY (X-MP/48) was received from the National Center for Supercomputing Applications (NCSA) under grant number MSM890006N.

References

Bažant, Z. P. (1984). "Size effect in blunt fracture: concrete, rock, metal." J. Engrg. Mech. Div., ASCE, 110 (4), 518-535.

Bažant, Z. P., and Pfeiffer, P. A., (1987) "Determination of fracture energy from size effect and brittleness number." ACI Materials Journal, 84 (6), 463-480

Bažant, Z. P., Pan, J., and Pijaudier-Cabot, G. (1987). "Softening in reinforced concrete beams and frames." J. Struct. Engrg., ASCE, 113 (12), 2333-2347.

Bažant, Z. P., and Kazemi, M.T. (1988). "Determination of fracture energy, process zone length and brittleness number from size effect, with application to rock and concrete." Report 88-7/498d Center for Concrete and Geomaterials, Northwestern University, Evanston, Illinois; also International J. of Fracture - in press.

Bažant, Z. P. and, Pijaudier-Cabot G. (1989). " Measurement of characteristic length of nonlocal continuum." J. Engrg. Mech. Div., ASCE, 115 (4), 755-767.

Burt, N. J. and, Dougill, J. W. (1977). " Progressive failure in a model Heterogeneous medium." J. Engrg. Mech., ASCE, 103 (3), 365-376

Cundall, P. A. (1971). "A computer model for simulating progressive large scale movements in blocky rock systems." Proc. Int. Symp. on Rock Fracture, ISRM, Nancy, France, 2-8.

Cundall, P. A. (1978). "BALL - A program to model granular media using the distinct element method." Technical Note, Advanced Technology Group, Dames and Moore, London, England.

Cundall, P. A., and Strack, O. D. L. (1979). "A discrete numerical model for granular assemblies." Geotechnique, 29, 47-65.

Kawai, T. (1980). "Some considerations on the finite element method." Int. J. Numer. Meth. Engrg., 16, 81-120.

Owen, D. R. J., and Hinton, E. (1980)." Finite elements in plasticity: theory and practice." Pineridge Press Ltd., Swansea, England.

Plesha, M. E., and Aifantis, E. C. (1983). "On the modeling of rocks with microstructure." Proc. 24th U. S. Symp. on Rock Mechanics, Texas A&M Univ., College Station, Texas, 27-39.

Rodriguez-Ortiz, J. M. (1974). "Study of behavior of granular heterogeneous media by means of analogical mathematical discontinuous models." Thesis presented to the Universidad Politecnico de Madrid, at Madrid, Spain, in partial fulfillment of the requirements for the degree of Doctor of Philosophy (in Spanish).

Rots, J. G., Hordijk, D. A., and de Borst, R. (1987). "Numerical simulation of concrete fracture in direct tension." Proc., 4th Int.

Conf. on Numerical Methods in Fracture Mechanics, San Antonio, Texas, A. R. Luxmoore et al., eds., Pineridge Press, Swansea, U.K., 457-471.

Serrano, A. A., and Rodriguez-Ortiz, J. M. (1973). "A contribution to the mechanics of heterogeneous granular media." Proc., Symp. on Plasticity and Soil Mechanics, Cambridge, U.K.

Zubelewicz, A. (1980). "Contact element method." Thesis presented to The Technocal University of Warsaw, at Warsaw, Poland, in partial fulfillment of the requirements for the degree of Doctor of Philosophy (in Polish).

Zubelewicz, A. (1983). "Proposal of a new structural model of concrete." Archiwum Inzynierii Ladowej, 29 (4), 417-429 (in Polish).

Zubelewicz, A., and Bažant, Z. P. (1987). "Interface element modeling of fracture in aggregate composites." J. Engrg. Mech. Div., ASCE, 113 (11), 1619-1630.

Zubelewicz, A., and Mróz, Z. (1983). "Numerical simulation of rock-burst processes treated as problems of dynamic instability." Rock Mech. and Engng., 16, 253-274.

Captions

Fig. 1 (a) Two adjacent circular particles with radii r_i and r_j and the corresponding truss member ij; (b) A typical randomly generated specimen and its corresponding mesh of truss elements; (c) Constitutive law for the matrix.

Fig. 2 Geometrically similar specimens of various sizes with randomly generated particles.

Fig. 3 (a) Direct tension specimens with d = 36,72 and 144 mm; (b),(c) and (d) load-displacement curves for specimens of different sizes; (e) normalized load-displacement curves for specimens 1A,2A and 3A.

Fig. 4 Linear regression plot (a) and size effect plot (b) constructed from the maximum load values calculated for direct tension specimens of various sizes.

Fig. 5 Cracking patterns at the last calculated point on the load-deflection curve for unnotched specimens in tension.

Fig. 6 (a) Load-displacement curve for specimen 1A; (b) Constitutive law for the matrix; (c) Evolution of cracking with load level.

Fig. 7 Load-strain curves for specimen 1A for (a) Full range of strain; (b) Small values of strain.

Fig. 8 Distribution of the longitudinal component of the interparticle forces P_i throughout the cross section of specimen 1A.

Fig. 9 Distribution of the longitudinal component of the interparticle forces P_i throughout the cross section of specimen 3A.

Fig. 10 The evolution of eccentricity as a function of load resultant.

Fig. 11 Linear regression for components of displacements of nodes with maximum size aggregates in (a) X - direction ; (b) Y - direction

Fig. 12 Three-point-bend specimens with d = 36,72 and 144 mm

Fig. 13 Linear regression plot (a) and size effect plot (b) constructed from the maximum load values calculated for three-point-bend specimens of various sizes.

Table 1. Number of Particles in Specimens

Specimen	d (mm)	L (mm)	Diameter (mm)				
			12	8	5	3	1
1A, B, C, D	36	96	2	16	14	68	35
2A, B	72	192	9	66	58	273	140
3A	144	384	39	264	235	1095	563

Table 2. Values of Fracture Energy G_f and Process Zone Size c_f

Three Point Bending	Particle Model			Experiment*	
	(1)	(2)	(3)	Mortar	Concrete
G_f (N/m)	16.5	6.8	31.2	20	37
c_f (mm)	3.5	4.2	5.9	1.9	13.5

* from Bažant and Kazemi (1988)

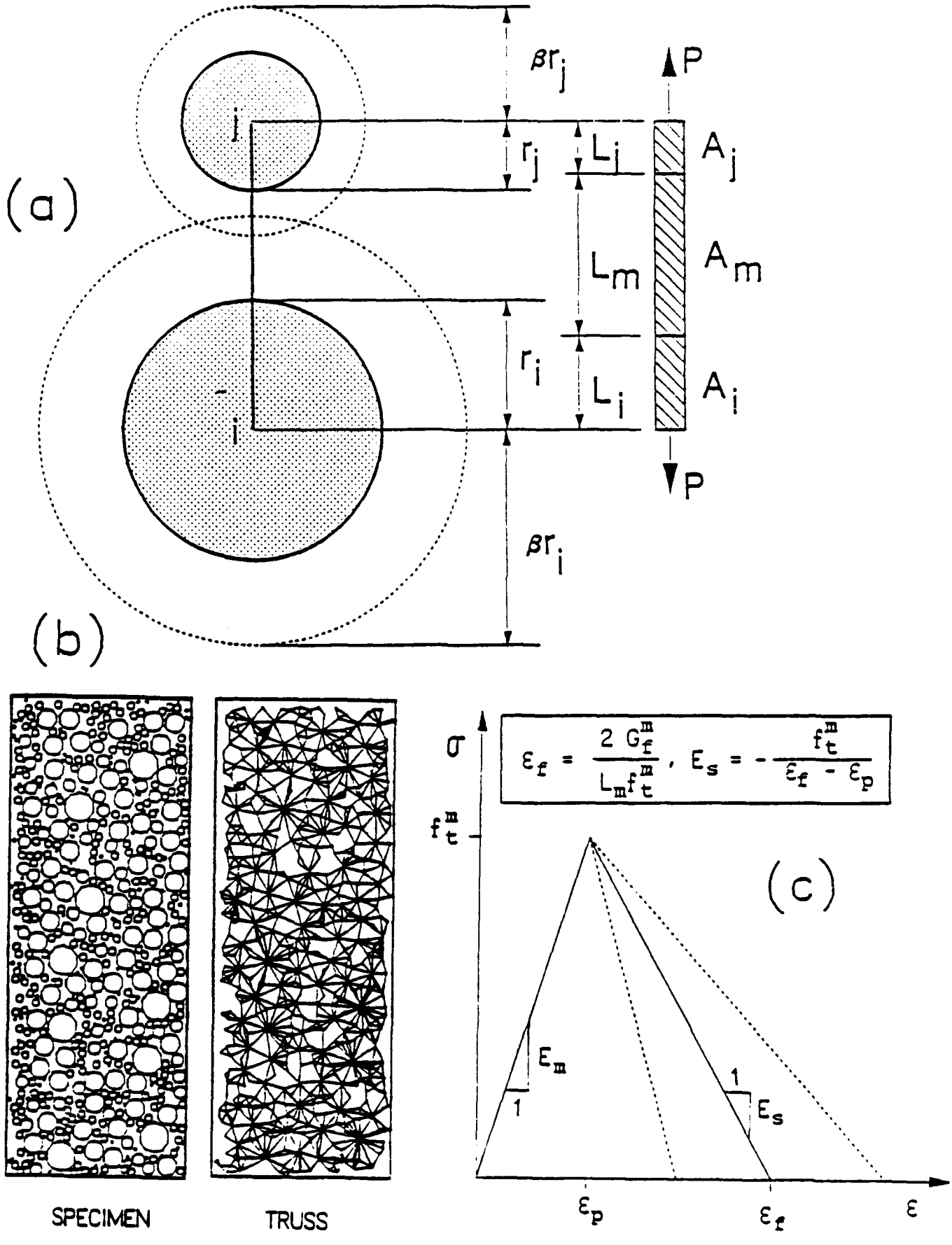

(a)
βr_j
P
j
r_j
L_j
A_j
L_m
A_m
r_i
L_i
A_i
i
P
βr_i
(b)
SPECIMEN
TRUSS
σ
ε_f = 2 G_f^m / (L_m f_t^m) , E_s = - f_t^m / (ε_f - ε_p)
f_t^m
(c)
E_m
1
1
E_s
ε_p
ε_f
ε

Fig. 1

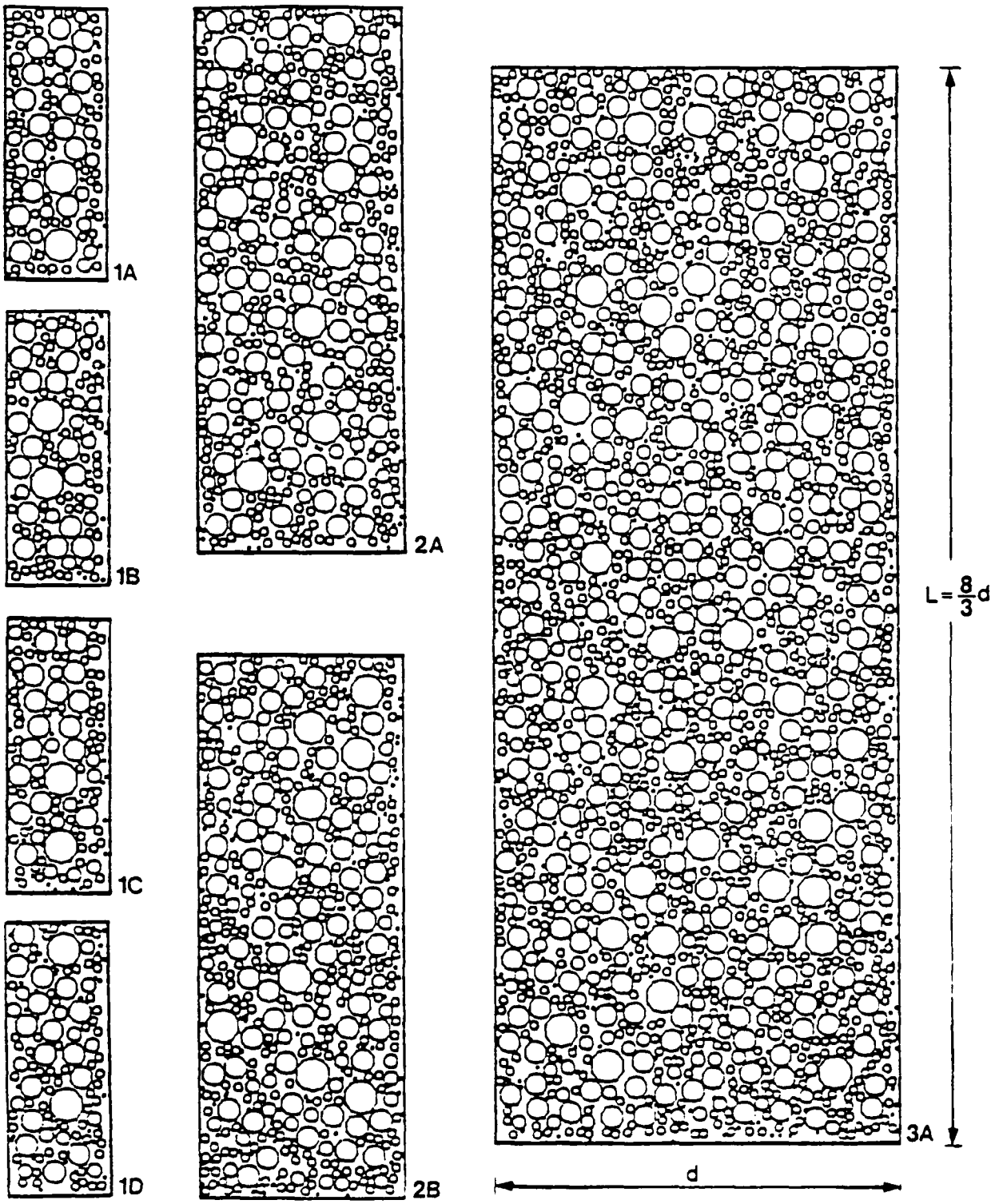


Fig. 2

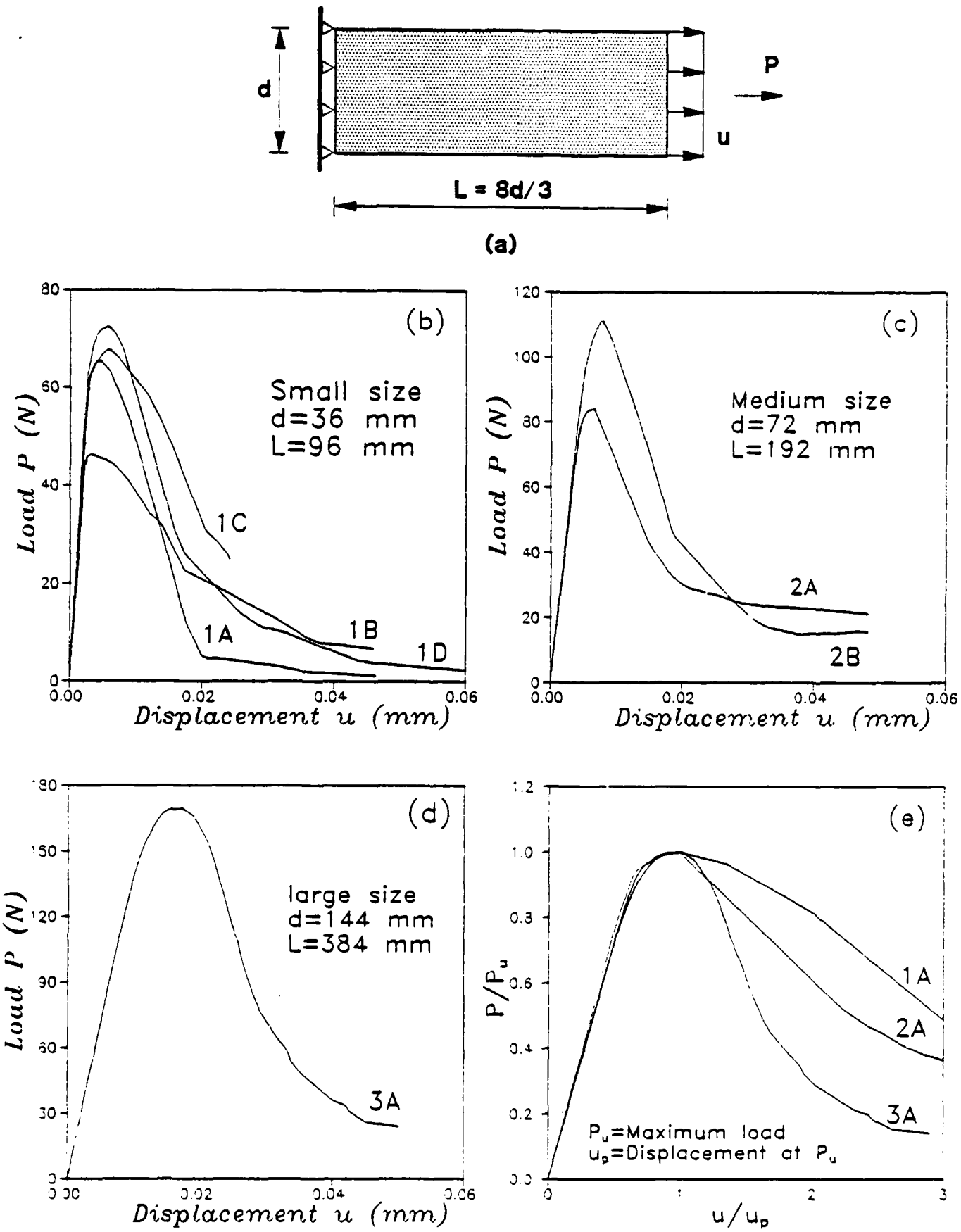
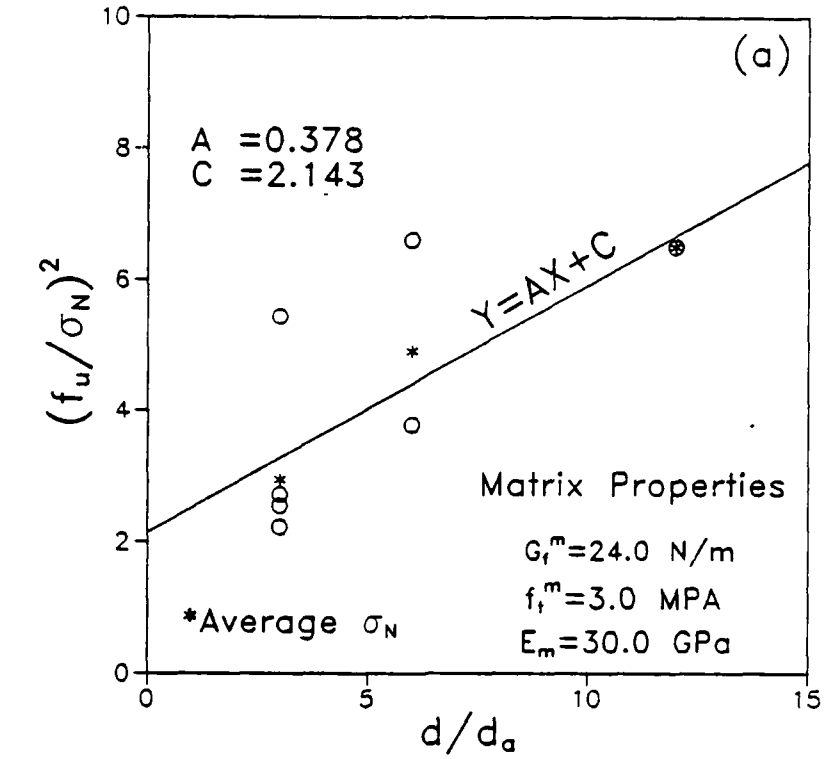

d
P
u
L = 8d/3
(a)
(b)
Small size
d=36 mm
L=96 mm
1C
1A
1B
1D
Load P (N)
Displacement u (mm)
(c)
Medium size
d=72 mm
L=192 mm
2A
2B
Load P (N)
Displacement u (mm)
(d)
large size
d=144 mm
L=384 mm
3A
Load P (N)
Displacement u (mm)
(e)
1A
2A
3A
P/Pu
u/up
Pu=Maximum load
up=Displacement at Pu

Fig. 3


(a)
A =0.378
C =2.143
Y=AX+C
Matrix Properties
G_f^m=24.0 N/m
f_t^m=3.0 MPA
E_m=30.0 GPa
*Average σ_N
$(f_u/\sigma_N)^2$
d/d_a
0
2
4
6
8
10
0
5
10
15

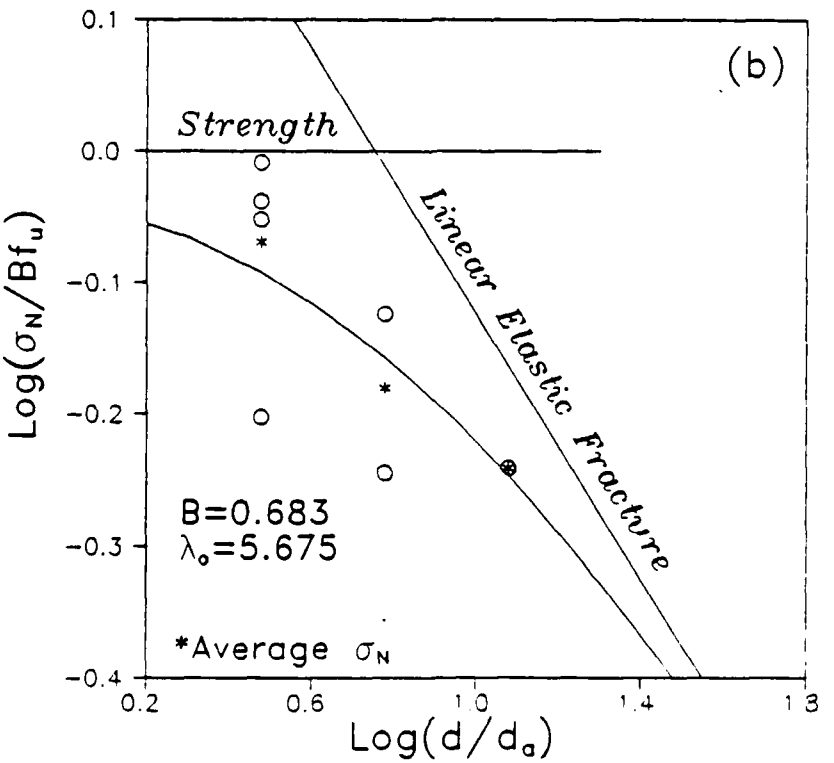

(b)
Strength
Linear Elastic Fracture
B=0.683
λ_o=5.675
*Average σ_N
$\text{Log}(\sigma_N/Bf_u)$
$\text{Log}(d/d_a)$
0.1
0.0
-0.1
-0.2
-0.3
-0.4
0.2
0.6
1.0
1.4
1.8

Fig. 4

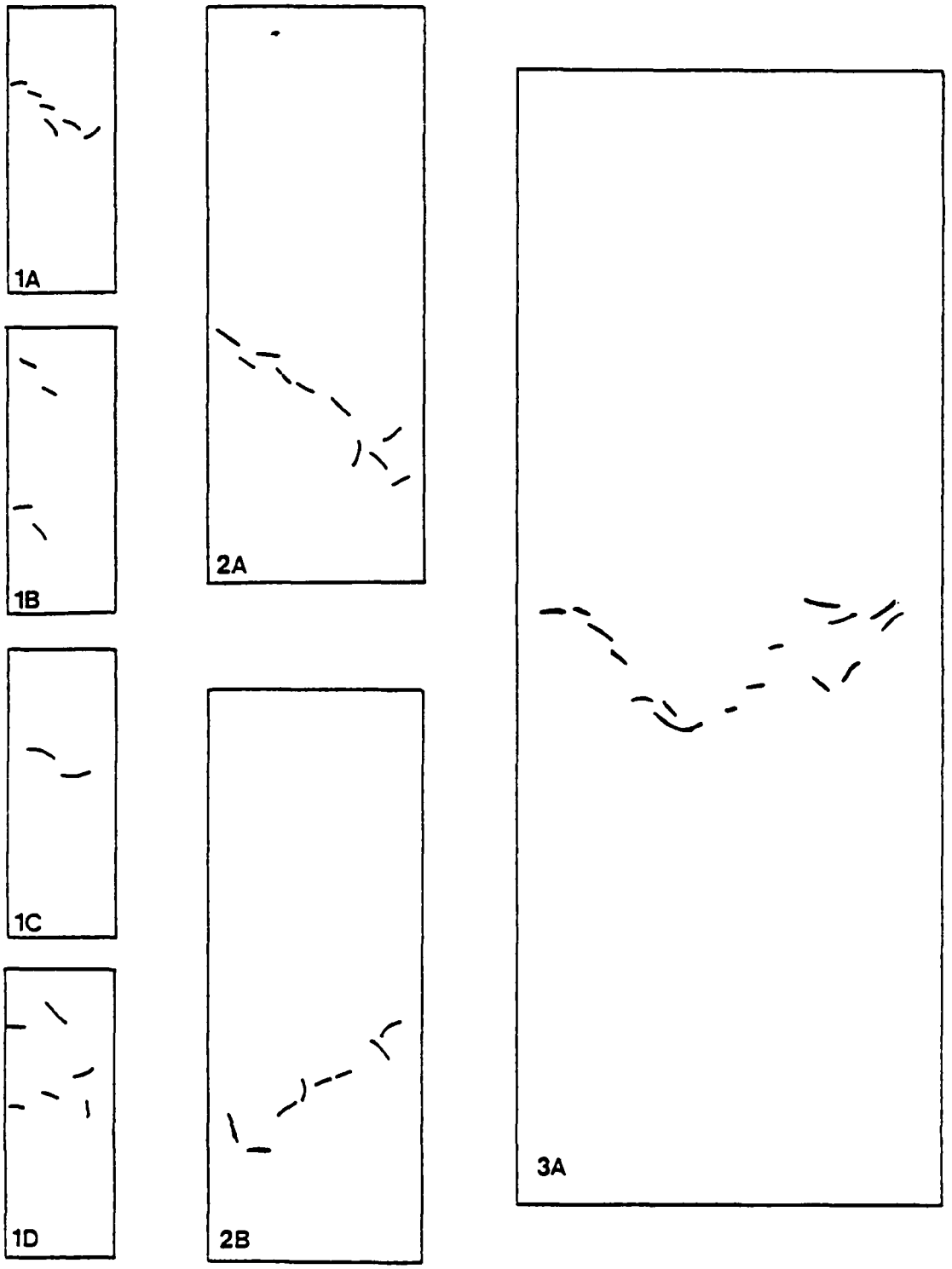
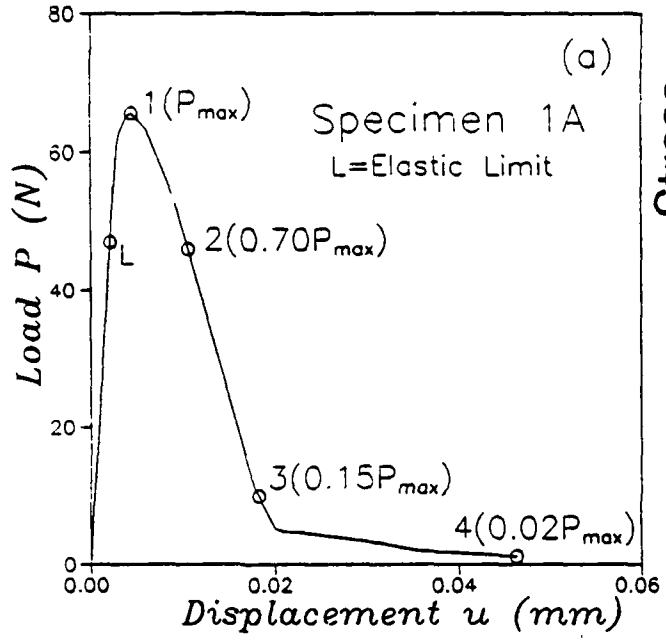
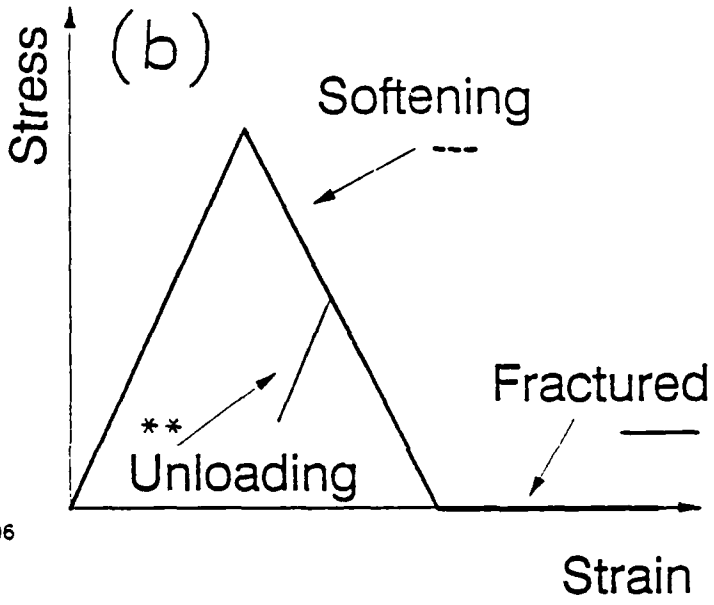
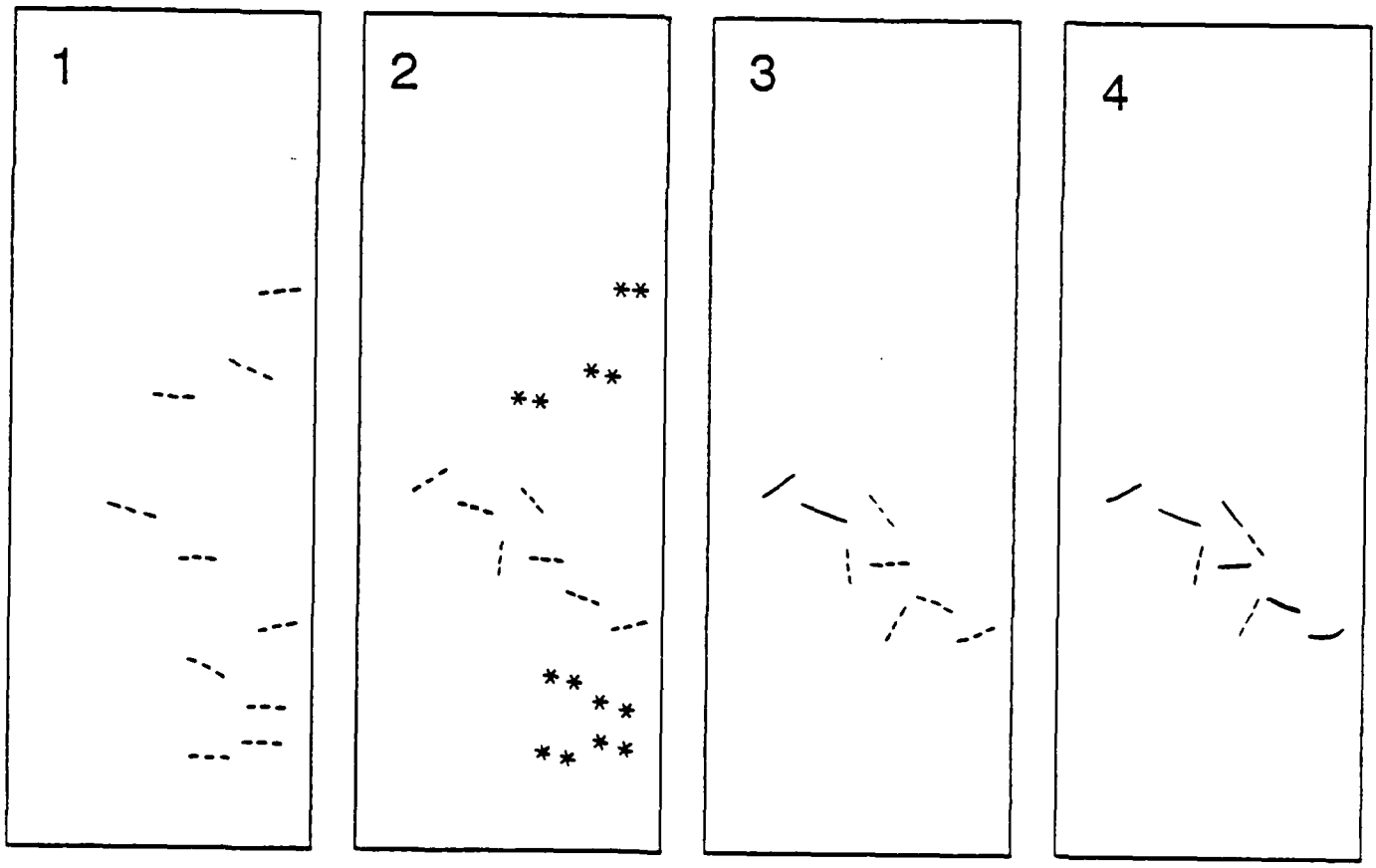


Fig. 5

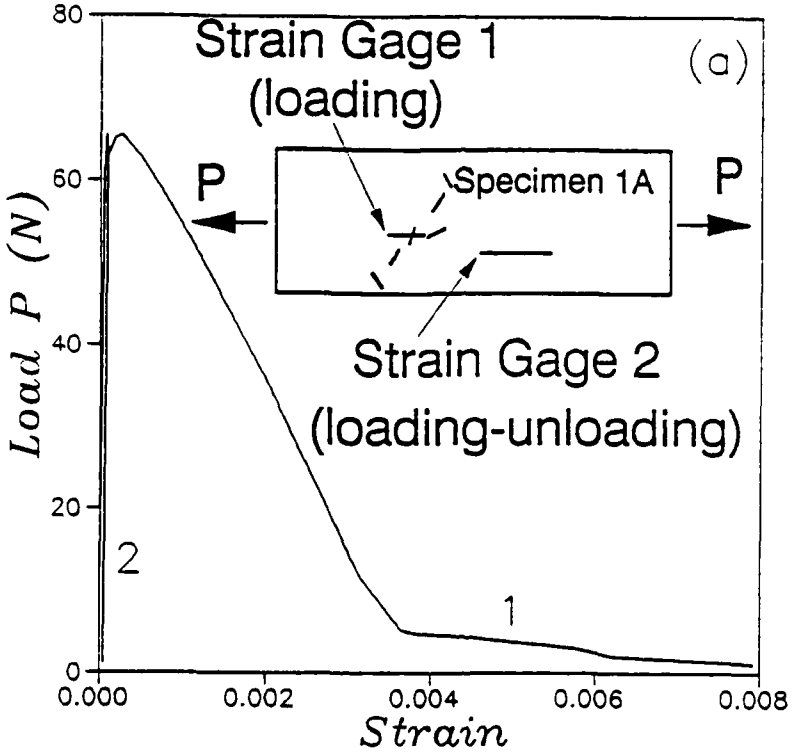

(a)
Specimen 1A
L=Elastic Limit
1(Pmax)
2(0.70Pmax)
3(0.15Pmax)
4(0.02Pmax)
L
Load P (N)
Displacement u (mm)
80
60
40
20
0
0.00
0.02
0.04
0.06


(b)
Stress
Softening
Unloading
Fractured
Strain


1
2
3
4

(c)

Fig. 6


80
60
40
20
0
Load P (N)
Strain Gage 1
(loading)
(a)
P
Specimen 1A
P
Strain Gage 2
(loading-unloading)
2
1
0.000
0.002
0.004
0.006
0.008
Strain

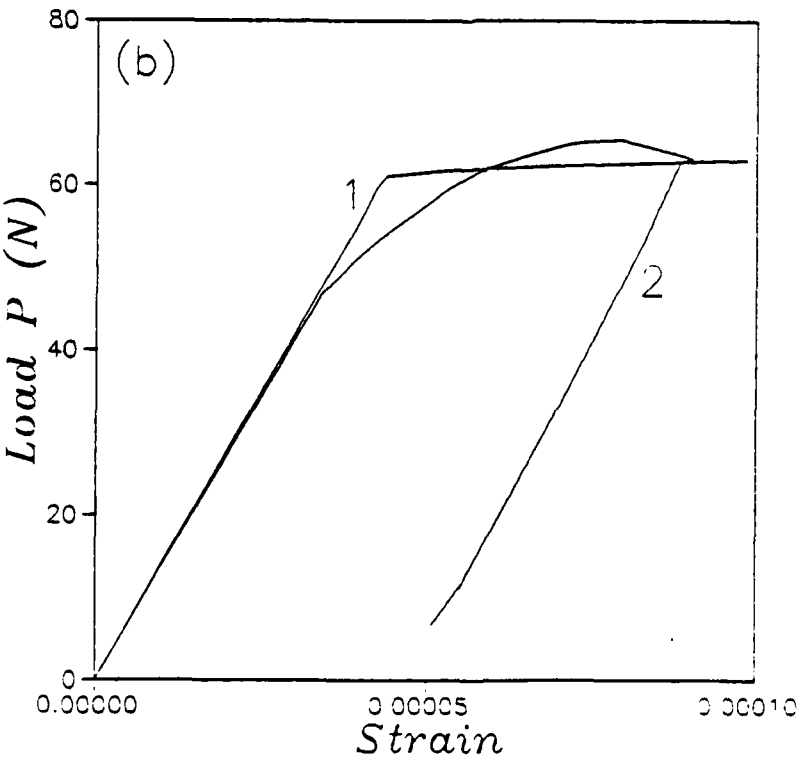

80
60
40
20
0
Load P (N)
(b)
1
2
0.00000
0.00005
0.00010
Strain

Fig. 7

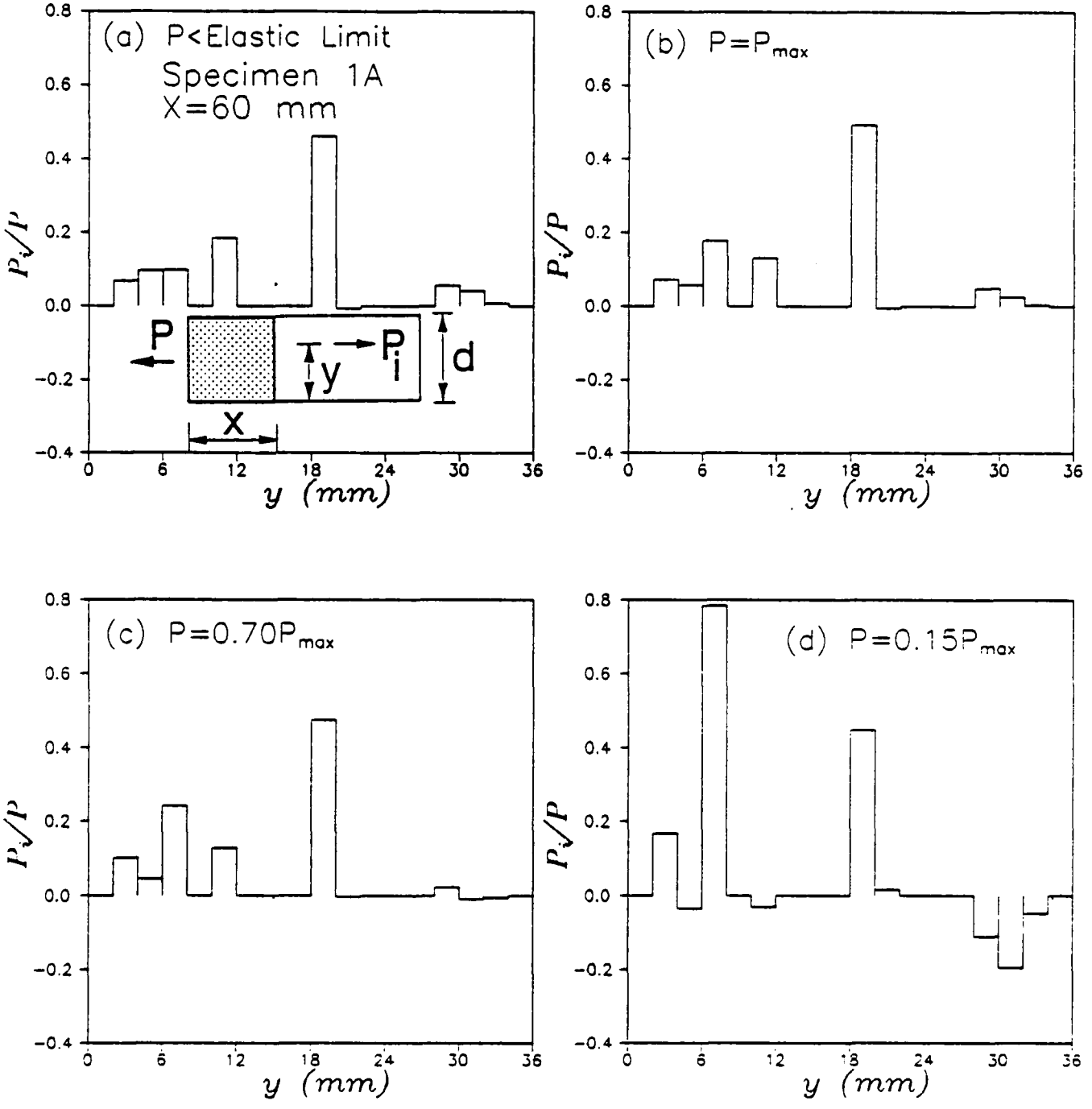

(a) P<Elastic Limit
Specimen 1A
X=60 mm
(b) P=Pmax
(c) P=0.70Pmax
(d) P=0.15Pmax
Pi/P
y (mm)
P
Pi
y
d
X
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
0
6
12
18
24
30
36

Fig. 8

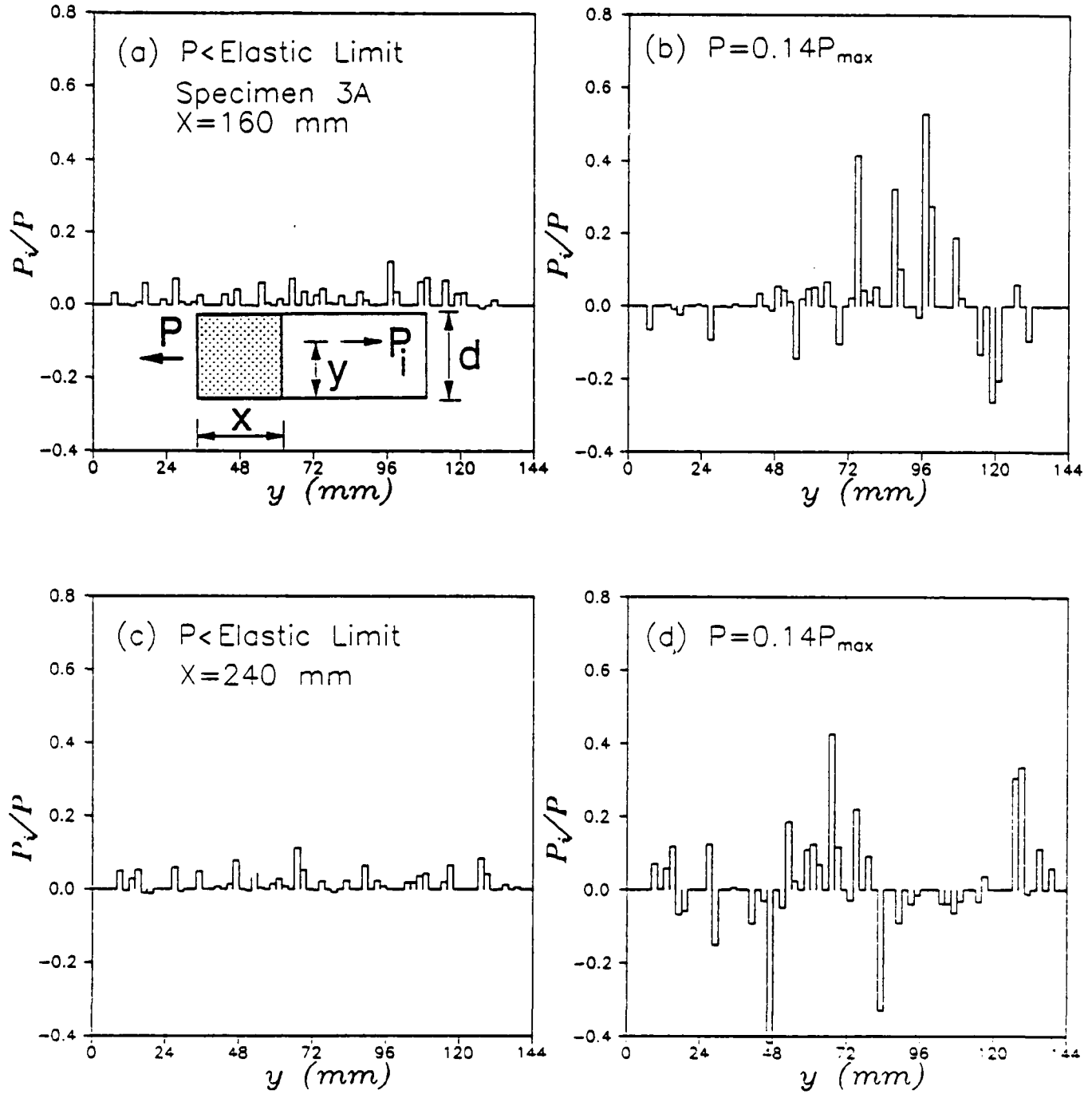

(a) P<Elastic Limit
Specimen 3A
X=160 mm
(b) P=0.14P_max
(c) P<Elastic Limit
X=240 mm
(d) P=0.14P_max
P_i/P
y (mm)
P
P_i
y
d
X
0.8
0.6
0.4
0.2
0.0
-0.2
-0.4
0
24
48
72
96
120
144

Fig. 9

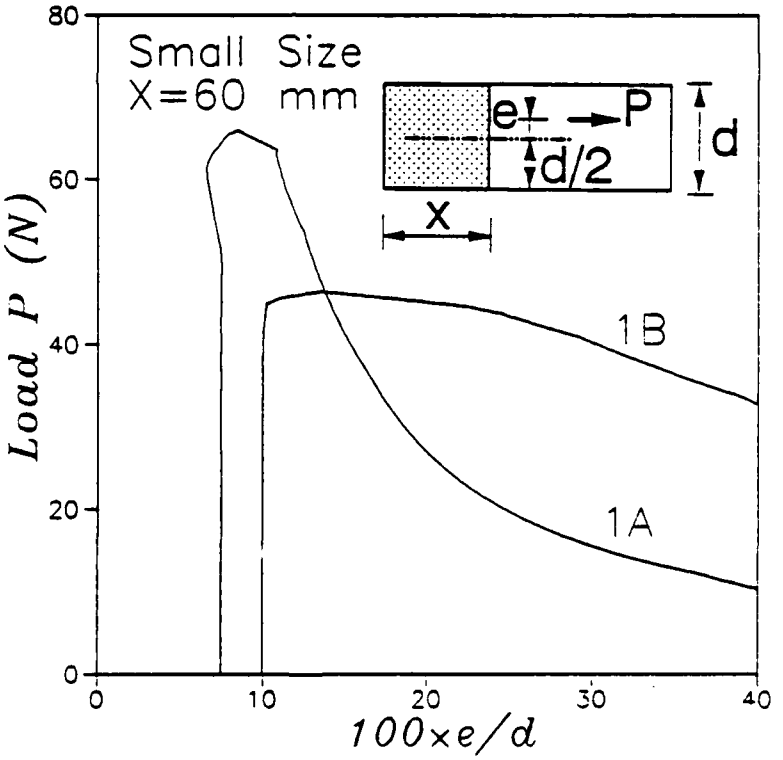
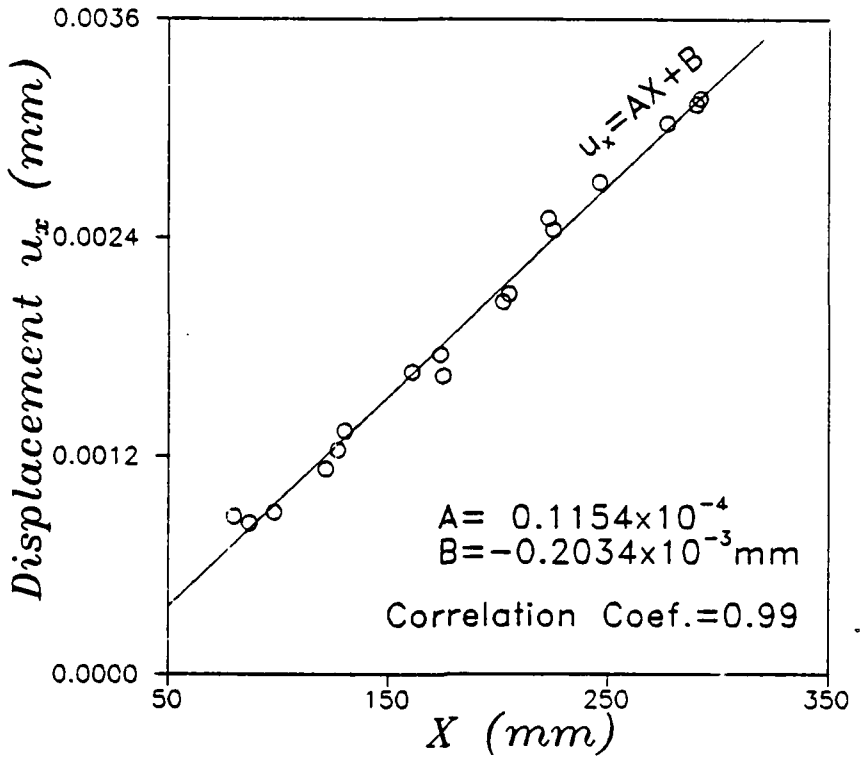

Small Size
X=60 mm
e
P
d/2
d
X
1B
1A
Load P (N)
100×e/d
0
20
40
60
80
0
10
20
30
40

Fig. 10


0.0036
0.0024
0.0012
0.0000
Displacement u_x (mm)
u_x=AX+B
A= 0.1154x10^{-4}
B=−0.2034x10^{-3}mm
Correlation Coef.=0.99
50
150
250
350
X (mm)

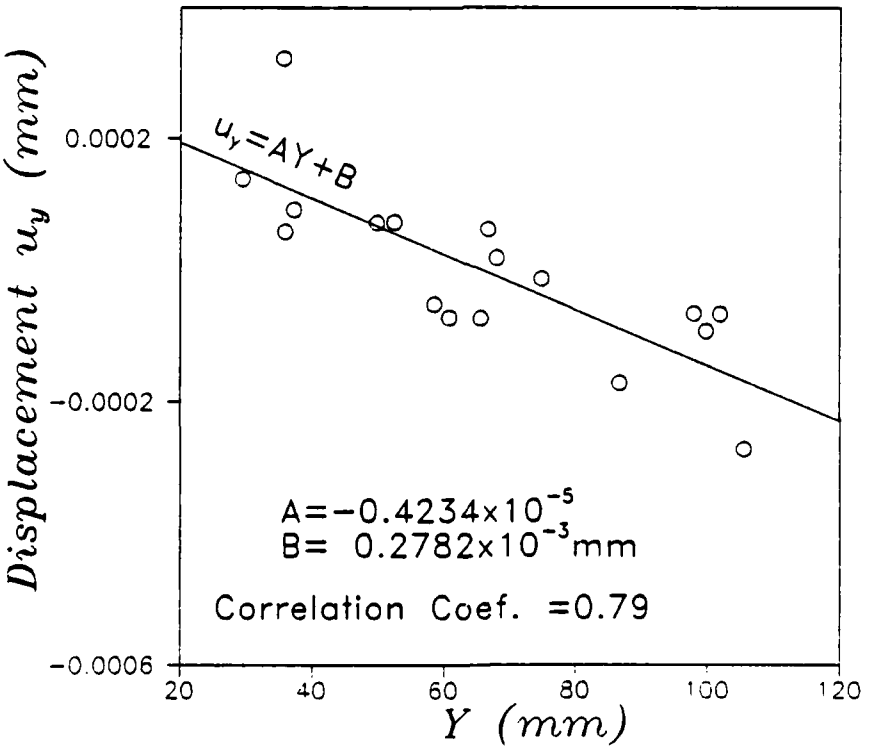

0.0002
−0.0002
−0.0006
Displacement u_y (mm)
u_y=AY+B
A=−0.4234x10^{-5}
B= 0.2782x10^{-3}mm
Correlation Coef. =0.79
20
40
60
80
100
120
Y (mm)

Fig. 11

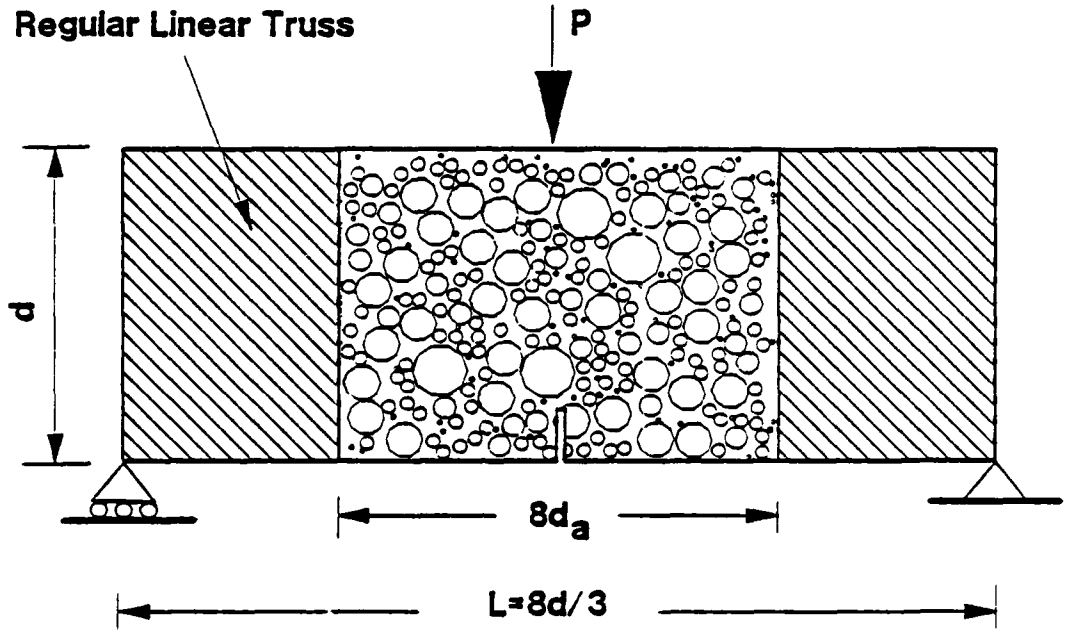


Fig. 12

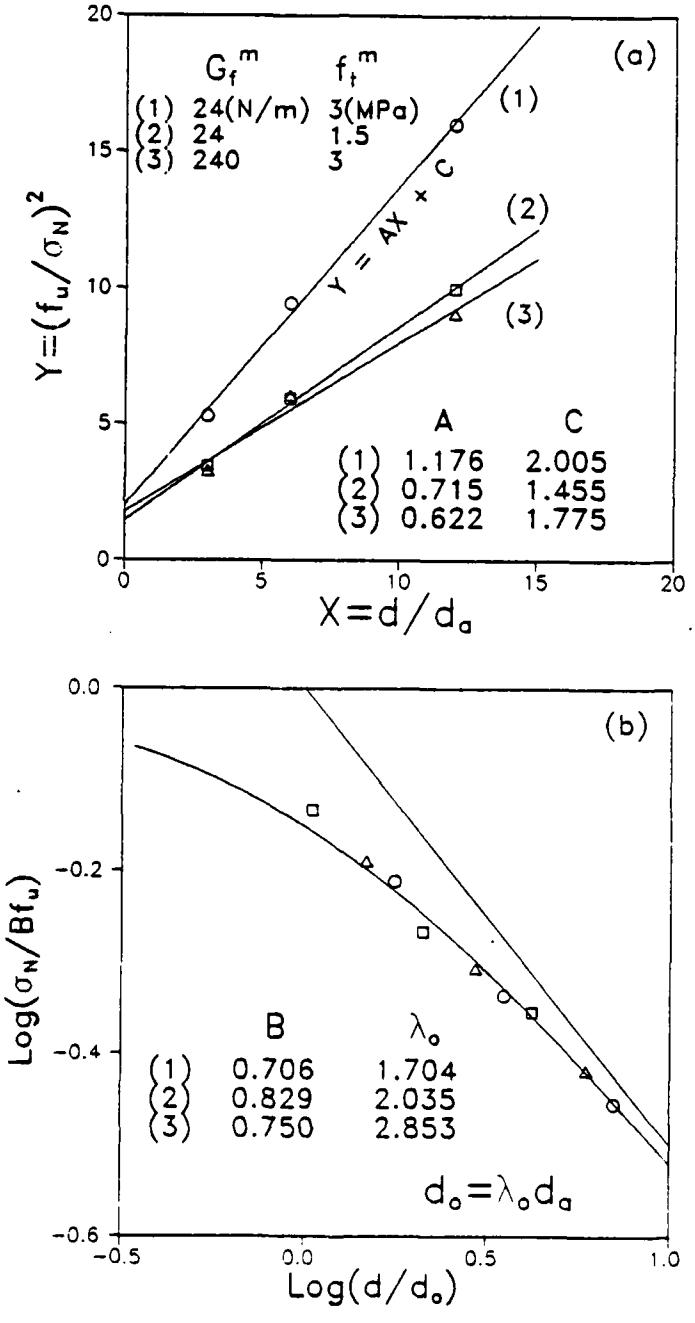

(a)
G_f^m f_t^m
(1) 24(N/m) 3(MPa)
(2) 24 1.5
(3) 240 3
(1)
(2)
(3)
$Y = AX + C$
A C
(1) 1.176 2.005
(2) 0.715 1.455
(3) 0.622 1.775
$Y=(f_u/\sigma_N)^2$
$X=d/d_a$
(b)
B λ_o
(1) 0.706 1.704
(2) 0.829 2.035
(3) 0.750 2.853
$d_o=\lambda_o d_a$
$Log(\sigma_N/Bf_u)$
$Log(d/d_o)$

Fig. 13

MECHANICS RESEARCH COMMUNICATIONS Vol. 14(5/6), 407-419, 1987. Printed in the USA.
0093-6413/87 $3.00 + .00 Copyright (c) 1987 Pergamon Journals Ltd.

WHY CONTINUUM DAMAGE IS NONLOCAL: JUSTIFICATION BY QUASIPERIODIC MICROCRACK ARRAY

Zdeněk P. Bažant
Professor of Civil Engineering
Northwestern University, Evanston, Illinois 60208, USA

(Received 24 August 1987; accepted for print 30 November 1987)

Abstract. - Strain-softening damage due to distributed cracking is modeled by an elastic continuum with a quasiperiodic array of cracks of regular spacing but varying sizes. As a model for the initial stage, the cracks are penny-shaped and small compared to their spacing, and as a model for the terminal stage the uncracked ligaments between the cracks are circular and small compared to their spacing. The strain due to cracks and the compliance per crack are calculated. The cracked material is homogenized in such a manner that the macroscopic continuum strains satisfy exactly the condition of compatibility with the actual strains due to cracks, and the macroscopic continuum stress satisfies exactly the condition of work equivalence with the actual stresses in the cracked material. The results show that, contrary to the existing theories, the damage variable used in continuum damage mechanics should be nonlocal, while the elastic part of the response should be local. In particular, the nonlocal continuum damage should be considered as a function of the spatial average of the cracking strain rather than its local value. The size of the averaging region is determined by the crack spacing.

Introduction

Introduction of nonlocal continuum concepts into the analysis of strain-softening structures [1] has recently met with considerable success. It has eliminated problems with spurious mesh sensitivity and incorrect convergence and has assured that refinements of the finite element mesh cannot lead to spurious localization of energy dissipation into a softening zone of a vanishing volume [1-10]. However, a physical justification of the nonlocal approach is still lacking. The objective of the present brief study is to show such a justification for a certain case of strain-softening that is caused by distributed cracking.

Array of Small Penny-Shaped Cracks

We consider an infinite elastic continuum containing an array of small circular (penny-shaped) cracks normal to axis x of cartesian coordinates x, y, z (Fig. 1a). The cracks lie in parallel planes $x = x_i = i\ell$ ($i = \ldots -2, 1, 0, 1, 2, \ldots$) and their centers lie at the nodes of a spatial cubic lattice whose nodal spacing for each of the lattice directions x, y and z is ℓ. To bring to light the nonlocal aspects, the crack array cannot be perfectly periodic. We make the array quasiperiodic by assuming that the radius a_i of the cracks on the planes $x = x_i$ slightly

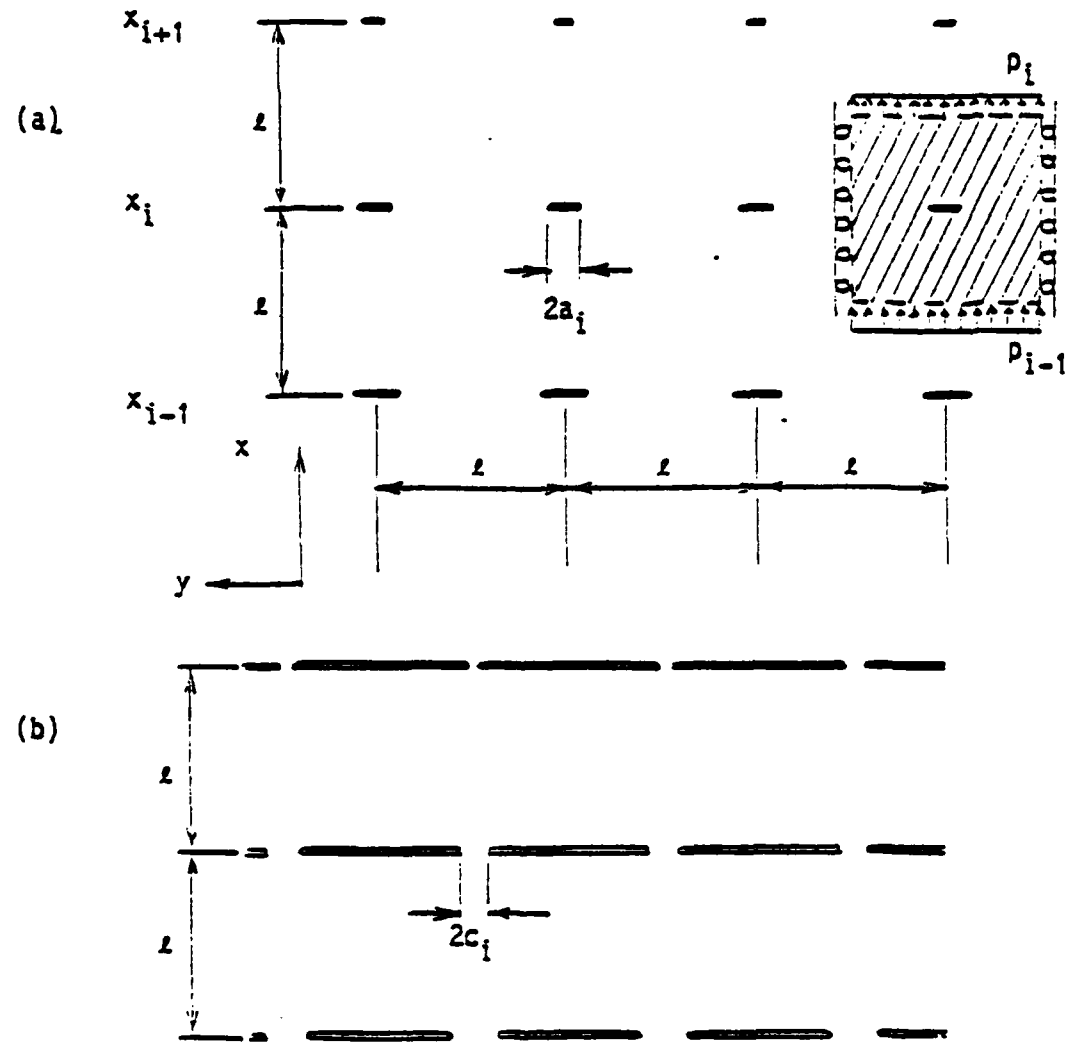


Fig. 1 - Quasiperiodic Crack Array in Infinite Elastic Continuum: (a) Small Penny-Shaped Cracks, (b) Small Circular Ligaments.

varies from plane to plane. All the cracks are assumed to be small compared to ℓ, i.e. $a_i \ll \ell$ for all i.

At $x \to -\infty$ we assume a fixed boundary, and at $x \to \infty$ a stress-free boundary. Small uniform normal distributed loads p_i are applied on the planes $x = x_i + \ell/2$. As the boundary conditions at $y \to \pm\infty$, $z \to \pm\infty$, we assume that the boundary points are sliding on the planes parallel to x. These boundaries develop additional lateral normal stresses σ'_z and σ'_y, which produce on the cracks stress intensity factors of zero values, and therefore they have no effect on the cracks. Consequently, these additional lateral stresses σ'_y and σ'_z, which are superimposed on the stress field due loads p_i, are uniform in each layer $-\ell/2 < x < x_i + \ell/2$.

Due to the loads in the x-direction, there are additional nonuniform stresses (of all components) near each crack. If a_i/ℓ is sufficiently small, the normal stress σ_x on each side of the plane $x_i + \ell/2$ is nearly uniform. Because of periodicity, each cell cross-hatched in Fig. 1a behaves as if its lateral boundaries were sliding. The average normal stress $\bar{\sigma}_x = \sigma_i$ on any plane x = const. is exactly the same for all x within the layer $x_i - \ell/2 < x < x_i + \ell/2$. Across the interfaces $x = x_i + \ell/2$ of these layers, σ_x varies discontinuously if p_i is nonzero. Note that the present boundary conditions and loading are introduced in such a manner that in absence of the cracks each layer is in a state of uniaxial strain, whose value varies from one layer to the next.

If the cracks are sufficiently small compared to ℓ (i.e., $a_i \ll \ell$), they do not interact and the formation of one crack does not release any appreciable amount of strain energy from the outside of the layer in which this crack is located. Therefore, the stress intensity factor K_i is the same as for a single crack in an infinite elastic solid subjected to remote stress σ_i, which is [11, 12]:

$$K_i = 2\sigma_i (a_i/\pi)^{1/2} \tag{1}$$

Accordingly, the energy release rate per crack circumference $2\pi a_i$ is [13, 14]:

$$\frac{\partial W_i}{\partial a_i} = 2\pi a_i \frac{K_i^2}{E'} = \frac{8}{E'} \sigma_i^2 a_i^2 \tag{2}$$

where $E' = E / (1 - \nu^2)$, E = Young's modulus, and ν = Poisson's ratio. By integration, the total energy release per crack is:

$$W_i = \frac{8}{3E'} a_i^3 \sigma_i^2 \tag{3}$$

By Castigliano's theorem, the displacement v_i due to the cracks lying in the plane $x = x_i$ is $v_i = \partial W_i / \partial P_i$ where P_i = force resultant per crack. Since $P_i = \ell^2 \sigma_i$, we have

$$v_i = \frac{1}{\ell^2} \frac{\partial W_i}{\partial \sigma_i} = \frac{16\, a_i^3}{3E'\ell^2} \sigma_i \tag{4}$$

This represents the difference of displacements between the planes $x = x_i - \ell/2$ and $x = x_i + \ell/2$. The average normal strain due to cracks in the interval $(x_i - \ell/2,\ x_i + \ell/2)$ may be defined as:

$$\bar{\gamma}_i = \frac{v_i}{\ell} = \left[\frac{16\ a_i}{3E'\ \ell^2}\right]^3 \sigma_i \tag{5}$$

We may call it the cracking strain. Furthermore, assuming the cracks to be propagating, we must have $K_i = K_c$ = critical stress intensity factor of the material. Thus, from Eq. 1,

$$a_i = \frac{\pi K_c^2}{4 \sigma_i^2} \tag{6}$$

Substituting this into Eq. 5, we obtain:

$$\sigma_i = \left[\frac{\pi^3 K_c^6}{12E'\ell^3 \bar{\gamma}_i}\right]^{1/5} = \frac{\bar{\gamma}_i}{C(\bar{\gamma}_i, \ell)} \tag{7}$$

in which function C represents the overall secant compliance due to cracks and is defined as:

$$C(\bar{\gamma}_i, \ell) = \frac{16}{3E'} \left[\frac{a_i(\bar{\gamma}_i)}{\ell}\right]^3 = \left[\frac{12\ E'\ \ell^3\ \bar{\gamma}_i^6}{\pi^3\ K_c^6}\right]^{1/5} \tag{8}$$

Eq. 7 yields a decreasing stress σ_i at increasing strain $\bar{\gamma}_i$, i.e. strain softening.

Note that compliance C depends on length ℓ, which may be regarded as the characteristic length of the cracked material. Dependence of the compliance or stiffness on some characteristic length is a typical property of nonlocal materials.

Homogenization by Macroscopic Continuum Approximation

The conditions of macroscopic equivalence of the actual cracked material and the homogenizing continuum have to be stated integrally for the basic, periodically repeated cell, in this case the layer $(x_i - \ell/2,$

$x_i + \ell/2$). The basic conditions we must impose are the compatibility of deformations over this cell and the equivalence of work. For our example, we can satisfy them exactly.

Compatibility of the macroscopic cracking strains $\gamma(x)$ with the displacement v_i due to cracks requires that

$$v_i = \int_{x_i-\ell/2}^{x_i+\ell/2} \gamma(x)\, dx \tag{9}$$

which implies that

$$\bar{\gamma}_i = \frac{1}{\ell} \int_{x_i-\ell/2}^{x_i+\ell/2} \gamma(x)\, dx = \langle \gamma(x_i) \rangle \tag{10}$$

The pointed brackets $\langle\ \rangle$ denote the averaging operator, and the superimposed bar is a label for the averaged (nonlocal) quantities.

Equivalence of work done by the stresses in the layer $(x_i - \ell/2,\ x_i + \ell/2)$ requires that

$$\sigma_i \delta v_i = \int_{x_i-\ell/2}^{x_i+\ell/2} \sigma(x)\ \delta\gamma(x)\, dx \tag{11}$$

where $\sigma(x)$ is the macroscopic continuum stress; and δv_i and $\delta\gamma(x)$ are any variations of the displacement due to cracks and the macroscopic cracking strains which are kinematically compatible, i.e. satisfy Eq. 9. Substituting for v_i from this equation, we obtain from Eq. 11:

$$\sigma_i \int_{x_i-\ell/2}^{x_i+\ell/2} \delta\gamma(x)\, dx = \int_{x_i-\ell/2}^{x_i+\ell/2} \sigma(x)\ \delta\gamma(x)\, dx \tag{12}$$

Since σ_i is constant within the layer, Eq. 12 may be rewritten as

$$\int_{x_i-\ell/2}^{x_i+\ell/2} [\sigma_i - \sigma(x)]\ \delta\gamma(x)\, dx = 0 \tag{13}$$

and because this must hold generally for all possible variations $\delta\gamma(x)$, we have

$$\sigma(x) = \sigma_i \text{ for } x_i - \ell/2 < x < x_i + \ell/2 \tag{14}$$

Note that this equation for stresses involves no averaging integral while Eq. 10 for strains does. For the type of loading we assumed (i.e. distributed loads p_i on planes $x = x_i + \ell/2$), $\sigma(x)$ is according to Eq. 14 a piecewise constant function. If the loads were not concentrated in these planes but distributed also over x, $\sigma(x)$ as well as $\sigma(x_i)$ would vary continuously and $\sigma(x)$ would not exactly equal $\sigma_i(x)$ (representing the

average of stresses σ_x over any plane x = const.). The reason we assumed the applied loads to be concentrated into discrete planes was to construct an example for which the continuum homogenization conditions can be satisfied exactly.

The presence of the spatial averaging integral in Eq. 10 is what ultimately impresses on the macroscopic continuum a nonlocal character, as we will see. At the same time, the absence of the averaging integral from Eq. 14 causes the nonlocal character to be restricted only to the cracking strains. In this regard it may be noted that in the theory of heterogeneous materials [15] the work equivalence condition is usually stipulated in a slightly different form, namely $l\sigma_i\epsilon^* = \int \sigma(x)\epsilon^* dx$ over an interval of length l, ϵ^* being an arbitrary uniform strain (virtual strain). This condition would imply that $\sigma_i = \int \sigma(x)dx/l$. However, it would not guarantee work equivalence for the actual displacements v_i and the actual cracking strains $\gamma(x)$. Thus Eq. 14 appears better justified for the present case.

According to Eqs. 10 and 14, the relation $\bar{\gamma}_i = C(\bar{\gamma}_i, l)\sigma_i$ (Eq. 7) has the following generalization:

$$\langle\gamma(x)\rangle = C(\langle\gamma(x)\rangle, l)\ \sigma(x) \tag{15}$$

We see that the continuum cracking strain in this stress-strain relation is nonlocal, i.e., is processed through a spatial averaging operator. However, this is not true of stress $\sigma(x)$.

The total displacement u in an elastic body with cracks represents a sum of the displacement obtained for the same body under the same loads if there are no cracks, and the additional displacement v due to the creation of the cracks while the loads are kept constant. Thus, the total relative displacement u_i of the macroscopic continuum between the planes $x = x_i - l/2$ and $x = x_i + l/2$ is

$$u_i = le(x_i) + v_i \tag{16}$$

where e is the elastic normal strain in the x-direction which is produced by stresses $\sigma(x)$ in the same elastic continuum if there are no cracks; and $\gamma(x)$ represents the increase of this elastic strain due to introduction of cracks of radii a_i. Strains e(x) are obtained by stress analysis of the continuum with no cracks, which is local.

Substituting $v_i = \ell\bar{\gamma}_i$ and defining the macroscopic strain at $x = x_i$ as $\epsilon(x_i) = u_i/\ell$, we get from Eq. 16 $\epsilon(x_i) = e(x_i) + \bar{\gamma}_i$, and generalizing this to any x, we obtain:

$$\epsilon(x) = e(x) + \langle\gamma(x)\rangle \tag{17}$$

So the total stress-strain relation for the normal x-components of stress and strain in the macroscopic continuum must have the form

$$\epsilon(x) = e(x) + C[\,\langle\gamma(x)\rangle,\ \ell\,]\,\sigma(x) \tag{18}$$

The elastic strain is determined as the strain in the continuum with no cracks. For the type of boundary conditions and loading that we introduced, the continuum with no cracks is in a state of uniaxial strain, i.e. all the strain components except ϵ_x are zero. Thus

$$e(x) = C_0\,\sigma(x) \tag{19}$$

where C_0 is the elastic compliance in uniaxial strain,

$$C_0 = \frac{(1+\nu)(1-2\nu)}{(1-\nu)\,E} \tag{20}$$

To sum up, the stress-strain relation for the macroscopic continuum should not be fully nonlocal. The elastic strains should be local, while the (macroscopically smoothed) strains due to cracking should be nonlocal. This property of the nonlocal formulation, introduced on the basis of numerical experience and intuition in Refs. 3-8, has been found to be essential for achieving well behaving finite element solutions.

It is possible to take an alternative approach to homogenization in which the relative total displacement $u_i = \ell C_0\sigma_i + \int \gamma(x)dx$ between $x = x_i \pm \ell/2$ is obtained before homogenization. Instead of Eq. 9 we may now impose the compatibility condition in the form $u_i = \int\epsilon(x)dx$ over the layer. Writing the work equivalence in the form $\sigma_i\delta u_i = \int \sigma(x)\delta\epsilon(x)\,dx$ instead of Eq. 11, we again recover Eq. 14 by the same procedure, however from the foregoing expressions for u_i we obtain $\langle\epsilon(x)\rangle = e(x) + C[\,\langle\gamma(x)\rangle, \ell]\sigma(x)$. The fact that $\langle\gamma(x)\rangle$ is nonlocal and $e(x)$, $\sigma(x)$ are local agrees with Eq. 18 but the appearance of nonlocal $\langle\epsilon(x)\rangle$ disagrees. It would make the use of this stress-strain relation more complicated, probably unnecessarily so. It should nevertheless be kept in mind that this aspect of homogenization is not without ambiguity.

The idea of nonlocal continuum was originally introduced without any reference to strain softening [16-22] and all strains and stresses were considered as nonlocal, i.e. the continuum was fully nonlocal. When this

idea was first proposed [2] to deal with strain-softening damage such as cracking, the stress-strain relation was also assumed to be fully nonlocal:

$$\langle \epsilon(x) \rangle = \langle e(x) \rangle + C[\, \langle \epsilon(x) \rangle - \langle e(x) \rangle,\ \ell\,]\, \langle \sigma(x) \rangle \tag{21}$$

This stress-strain relation, whose finite element implementation leads to the so-called imbricate medium and an imbricated finite element system [2], was found to cause various problems in finite element programming, due to difficulties in the implementation of the boundary conditions and interface conditions as well as the existence of certain spurious zero-energy instability modes. Thus, even though these difficulties have been overcome by certain artificial devices [23], this original version of the nonlocal formulation for strain-softening is complicated and less than satisfactory from the theoretical viewpoint. From the comparison of Eqs. 18 and 21 we now understand why.

To recapitulate, the elastic strain e is the strain of the continuum with no cracks, and so it must be local. On the other hand the macroscopic cracking strain must be defined by averaging (smoothing) of the effect of the discrete cracks in order to satisfy the strain compatibility requirement, and so it must be nonlocal. The basic reason is that the displacement due to a crack (Eq. 4) is defined with a unique value only at a sufficient distance from the crack.

Nonlocal Continuum Damage Mechanics

In continuum damage mechanics [24-27, 1], the stress-strain relation is written in the form:

$$\epsilon(x) = \frac{C_0}{1 - \Omega(x)}\, \sigma(x) \tag{22}$$

where $\Omega(x)$ is called damage and C_0 is the elastic compliance for the continuum with no damage ($\Omega = 0$), i.e. no cracks. Comparison of Eqs. 18 and 22 and substitution $e(x) = C_0 \sigma(x)$ yields for the damage the expression:

$$\Omega(x) = 1 - \frac{C_0}{C[\, \langle \gamma(x) \rangle,\ \ell\,] + C_0} = f(\, \langle \gamma(x) \rangle\,) \tag{23}$$

i.e. the damage is a function, f, of the average cracking strain rather than the local cracking strain.

From Eq. 23, it is clear that continuum damage Ω should be considered to be a nonlocal variable. Moreover, Eq. 23 shows that Ω should be a function of the average of the macroscopic cracking strain $\gamma(x)$. Such a definition was introduced for reasons of proper convergence at mesh refinements in Refs. 3, 4 and 6. Ref. 3 also considered an alternative in which Ω is obtained as the average of the local damage energy release rate Y (or local damage ω) that depends on local γ. This is equivalent only approximately, not exactly. The same is true when Ω is considered to be a function of $\langle\epsilon(x)\rangle$ rather than $\langle\gamma(x)\rangle$, as done in Ref. 8 for other reasons.

The definition of damage in Eq.23, along with the stress-strain relation in Eq. 18 in which ϵ, ϵ^e and σ are local, was found to yield excellent results in various finite element applications. The solutions were found to converge properly on mesh refinement [3, 4-6] and the numerical implementation proved to be quite easy even in large finite element programs [5, 7, 8]. Moreover, the nonlocal version of such programs provided faster convergence than the local version. As one application, the problem of cave-in induced by compressive strain softening at the sides of a subway tunnel excavated in a cement-grouted soil was solved using meshes with up to 3248 degrees of freedom [7].

If $\gamma(x)$ varies so slowly that the change of γ over distance ℓ would be negligible, then of course $\langle\gamma(x)\rangle$ can be replaced by $\gamma(x)$, and the damage is then local. However, since strain softening causes strain localization, such a slow variation of $\gamma(x)$ cannot be assumed to occur for all x. This is why a local treatment of strain-softening damage is always inadequate.

Other Crack Systems and Generalizations

Many other types of crack systems can be analysed similarly. As one further example consider the terminal stage of damage in which the planes $x = x_i$ are fully cracked except for small circular ligaments of diameters $2c_i$ (Fig. 1b). The centers of these ligaments coincide again with the nodes of a cubic lattice of step ℓ, and the ligaments are assumed to be very small compared to ℓ, i.e. $c_i \ll \ell$ for all i. This problem was analysed for a different purpose in Ref. 28. It was shown that for K_i =

K_c) the curve v_i (σ_i) attains a maximum of v_i after which it exhibits a snapback, whose final stage is described by the equation

$$\sigma_i = \frac{E'^3}{4\pi K_c^2 \ell^2} v_i^3 \tag{24}$$

Again v_i is the relative displacement between the planes $x = x_i - \ell/2$ and $x = x_i + \ell/2$.

The average normal strain due to cracks in the interval $(x_i - \ell/2, x_i + \ell/2)$ may again be defined as $\bar{\gamma}_i = v_i/\ell$. Eq. 24 may then be written as

$$\sigma_i = \frac{\bar{\gamma}_i}{C(\bar{\gamma}_i, \ell)} \tag{25}$$

in which

$$C(\bar{\gamma}_i, \ell) = \frac{E'^3}{4\pi K_c^2} \ell \bar{\gamma}_i^2 \tag{26}$$

To satisfy the requirement of compatibility of displacement v_i with the macroscopic cracking strain $\gamma(x)$, the value of $\bar{\gamma}_i$ may be replaced by $\langle\gamma(x)\rangle$, and further analysis is the same as before (Eqs. 9-20), with the same conclusions.

The detailed form of the compliance function $C(\bar{\gamma}_i, \ell)$, as defined by Eq. 8 or 26 is for our conclusions about the nonlocal aspects obviously unimportant. What is important is that C depends on the characteristic length ℓ, and even more that it depends on the average cracking strain $\bar{\gamma}_i$ which is macroscopically equivalent to $\langle\gamma(x)\rangle$ rather than $\gamma(x)$. Various other types of crack arrays should therefore lead to similar conclusions.

The preceding analysis neglected the microscopic heterogeneity of the material, such as the differences in elastic moduli between the aggregate and the matrix found in concrete. This heterogeneity no doubt has a large influence but it alone might not necessitate a nonlocal treatment. According to numerical experience, the nonlocal approach is required only if strain softening takes place, and strain softening is the consequence of microcracking, void growth or other damage.

The crack spacing in real materials is randomly irregular. To model this, one might consider a certain statistical distribution of lengths ℓ. Accordingly, one could introduce a certain weighting function in the averaging integral (operator $\langle\ \rangle$) in Eq. 18. This has already been done in Refs. 3-8 for reason of numerical efficiency. The analysis should also be

generalized to three dimensions, although this would bring about much greater complexity.

The number of cracks increases during the progress of damage, and this reduces their average spacing. This might require the characteristic length to be considered as a variable. The question would then arise as to what is the proper domain over which the averages should be taken. Also, it needs to be recognized that there are other sources of softening damage than cracking, e.g., void nucleation and growth in metals, or interface slips with softening. Further extensions would be needed for such problems.

Conclusion

1. Based on the analysis of an array of small penny-shaped cracks it appears that the damage variable in continuum damage mechanics should be treated as nonlocal while the elastic part of the response should be local.

2. The nonlocal damage should be formulated as a function of the spatial average of the cracking strain over a zone whose size, representing the characteristic length of the macroscopic continuum, coincides with the crack spacing.

Acknowledgments. - Financial support from the U.S. Air Force Office of Scientific Research under contract No. F49620-87-C-0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu, is gratefully acknowledged. Thanks are due to graduate research assistant Gilles Pijaudier-Cabot for some stimulating discussions, and to Nadine Pijaudier-Cabot for her expert and prompt secretarial services.

References

1. Bažant, Z. P., "Mechanics of Distributed Cracking", Applied Mechanics Reviews", ASME, Vol. 39, No. 5, May 1985, pp. 675-705.

2. Bažant, Z. P., Belytschko, T. B., and Chang, T. P., "Continuum Theory for Strain-Softening", J. of the Engineering Mechanics Division, ASCE, 1984, Vol. 110, No. 12, pp. 1666-1692.

3. Pijaudier-Cabot, G., and Bažant, Z. P., "Nonlocal Damage Theory", J. of Eng. Mech., ASCE , Vol. 113, No. 10, Oct. 1987, pp. 1512-1533.

4. Bażant, Z. P., and Pijaudier-Cabot, G., "Modeling of Distributed Damage by a Nonloca Continuum with Local Strain", Reprints, 4th Int. Conf. on Numerical Methods in Fracture Mechanics, held in San Antonio, March 1987, ed. by A. R. Luxmore, University College, Swansea, pp. 411-437.

5. Bażant, Z. P., Lin, F. B., and Pijaudier-Cabot, G., "Yield Limit Degradation: Nonlocal Continuum with Local Strains", Preprints, Int. Conf. on Computational Plasticity, held in Barcelona, ed. by E. Onate, R. Owen and E. Hinton, University of Wales, Swansea, 1987, pp. 1757-1780.

6. Bażant, Z. P., and Pijaudier-Cabot, G., "Nonlocal Continuum Damage: Localization Instability and Convergence", Report No. 87-2/428n-I, Center of Concrete and Geomaterials, Northwestern University, Evanston, Il, Feb. 1987; also J. of Appl. Mech. ASME, in press..

7. Bażant, Z. P., and Lin, F. B., "Nonlocal Yield Limit Degradation", Report No. 87-7/428ny, Center for Concrete et Geomaterials, Northwestern University, July 1987 (submitted to J. Eng. Mech. ASCE).

8. Bażant, Z. P., and Lin, F. B., "Nonlocal Smeared Cracking Model for Concrete Fracture", Report No. 87-7/428ns (submitted to J. of Struct. Eng. ASCE).

9. Schreyer, H. L., and Chen, Z., "The Effect of Localization on the Softening Behavior of Structural Members", Proceedings of Symposium on Constitutive Equations: Micro, Macro, and Computational Aspects, ASME Winter Annual Meeting, New Orleans, Dec., K. Willam (Ed.), ASME, New York, 1984, pp. 193-203; also ASME J. of Applied Mechanics, Vol. 53, Dec. 1986, pp. 791-797.

10. Willam, K. J., Hurlbut, B., and Sture, S., "Experimental, Constitutive and Computational Aspects of Concrete Failure", Preprints, U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concretes, Tokyo, May 1985, pp. 149-172.

11. Tada, H., Paris, P. C., and Irwin, G. R., "The Stress Analysis of Cracks Handbook", 2nd ed., Paris Productions Inc., St. Louis (226 Woodbourne Dr.), Mo., 1985.

12. Murakami, Y., Ed., "Stress Intensity Factors Handbook", Pergamon Press, Oxford - New York, 1987.

13. Knott, J. F., "Fundamentals of Fracture Mechanics", Butterworth, London, 1983.

14. Broek, D., "Elementary Engineering Fracture Mechanics", International Publishers, Netherlands, 1974.

15. Hashin, Z., "Analysis of Composite Materials", (a survey), ASME J. of Appl. Mech. Vol. 50, Sept. 1983.

16. Kröner, E., "Elasticity Theory of Materials with Long-range Cohesive Forces", Int. J. Solids Struct., Vol. 3, 1968, pp. 731-742.

17. Kröner, E., "Interrelations between Various Branches of Continuum Mechanics", Mechanics of Generalized Continua, E. Kroner, Ed., Springer-Verlag, W. Berlin, W. Germany, pp. 330-340, 1968.

18. Krumhansl, J. A., "Some Considerations of the Relations between Solid State Physics and Generalized Continuum Mechanics", Mechanics of Generalized Continua, E. Kroner (Ed.), Springer-Verlag, W. Berlin, W. Germany, 1968, pp. 298-331.

19. Kunin, I. A., "The Theory of Elastic Media with Microstructure and the Theory of Dislocations", Mechanics of Generalized Continua, E. Kroner (Ed.), Springer-Verlag, W. Berlin, W. Germany, 1968, pp. 321-328.

20. Levin, V. M., "The Relation between Mathematical Expectation of Stress and Strain Tensors in Elastic Microheterogeneous Media" (in Russian), Prikl. Mat. Mekh. 35, 694-701, 1971.

21. Eringen, A. C., and Edelen, D. G. B., "On Nonlocal Elasticity", Int. J. Eng. Science, Vol. 10, 1972, pp. 233-248.

22. Eringen, A. C., and Ari, N., "Nonlocal Stress Field at Griffith Crack", Crist. Latt. and Amorph. Materials, Vol. 10, pp. 33-38, 1983.

23. Bažant, Z. P., and Chang, T. P., "Nonlocal Finite Element Analysis of Strain-Softening Solids", J. of the Engineering Mechanics Division, ASCE, Vol. 113, No.1, 1987, pp. 84-105.

24. Lemaitre, J., and Chaboche, J.-L., "Aspect phénomenologique de la rupture par endommagement", Journal de Mécanique Appliquée (Paris), Vol. 2, 1978, pp. 317-365.

25. Lemaitre, J., and Chaboche, J.-L., "Mécanique des Matériaux Solides", Dunod-Bordas, Paris, 1985.

26. Krajcinovic, P., and Fonseka, G. U., "The Continuous Damage Theory of Brittle Materials", ASME J. of Applied Mech., Vol. 48, 1981, pp. 809-815.

27. Mazars, J., and Pijaudier-Cabot, G., "Continuum Damage Theory: Application to Concrete", Int. Report No. 71, Laboratoire de Mecanique et Technologie, Cachan, France, Dec. 1986.

28. Bažant, Z. P., "Snapback Instability at Crack Ligament Tearing and its Implication for Fracture Micromechanics", Cement and Concrete Research, Vol. 17, No. 6, Nov. 1987, pp. 951-967.

CEMENT and CONCRETE RESEARCH. Vol. 17, pp. 951-967, 1987. Printed in the USA.
0008-8846/87 $3.00+00. Copyright (c) 1987 Pergamon Journals, Ltd.

SNAPBACK INSTABILITY AT CRACK LIGAMENT TEARING AND ITS IMPLICATION FOR FRACTURE MICROMECHANICS

Zdeněk P. Bažant
Professor of Civil Engineering
Northwestern University, Evanston IL 60201, USA

(Communicated by F.H. Wittmann)
(Received July 13, 1987)

ABSTRACT

The microcracking in the fracture process zone ahead of a major crack is assumed to consist, in the initial stage, of a two-dimensional array of small circular (penny-shaped) cracks and, in the terminal stage, of a two-dimensional array of small circular ligaments, all located on the main crack plane. Both cases are solved in three-dimensions according to linear elastic fracture mechanics. The solution is approximate but asymptotically exact both for very small circular cracks and very small circular ligaments. The spacing of the cracks as well as the ligaments is governed by the spacing of the large aggregate pieces. The curve of the transverse displacement v due to cracks versus the remote applied normal stress is calculated and is found to exhibit snapback instability at which a negative slope changes to a positive slope and v reaches its maximum possible value. Since several other influencing physical mechanisms were neglected in the analysis, it still remains to be verified whether the snapback instability does actually occur in the concrete fracture process. The asymptotic behavior at ligament tearing is further analyzed, based on St.-Venant's principle, for arbitrary general three- and two-dimensional situations and it is shown that when the ligament transmits a force (mode I, II or III), its final tearing is always characterized by snapback instability, which determines maximum possible displacement due to crack. When, however, the ligaments transmit only a moment (bending or torsional), there is no snapback instability.

Nature of Problem

The softening law for the fracture process zone of a heterogeneous material such as concrete may be characterized either by the function $\sigma = \sigma(\delta_f)$ or the function $\sigma = \sigma(\gamma)$ where σ is the normal stress across the plane of the crack ahead of the crack tip, δ_f is the total displacement across the fracture process zone of width w_c, and $\gamma = \delta_f/w_c$ = mean normal strain across the fracture process zone (Fig. 1). The shape of the curve $\sigma(\delta_f)$ or $\sigma(\gamma)$ is the central question in the modeling of fracture in concrete as well as other materials such as rock, ceramics, composites, ice, etc. (1).

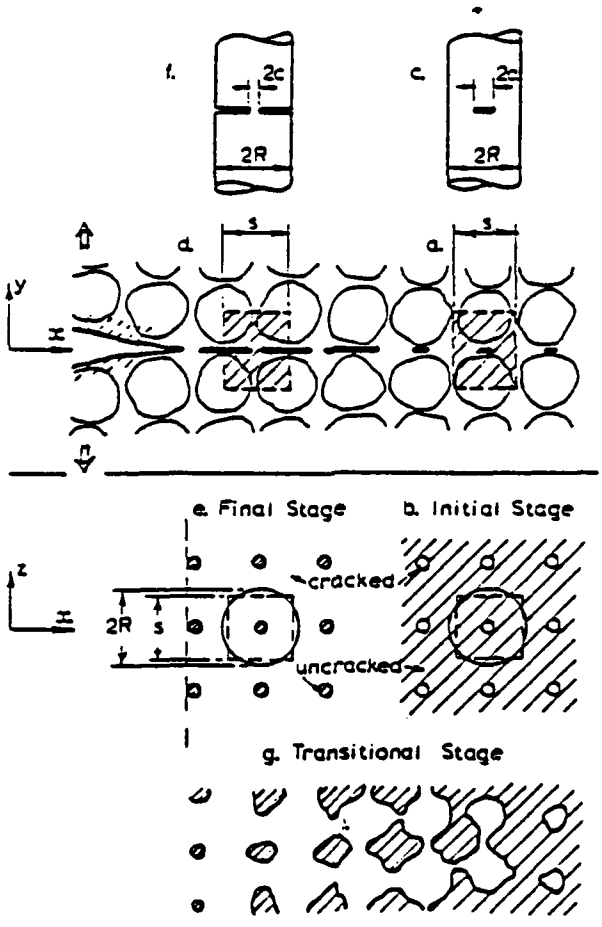


FIG. 1
Three-dimensional Idealization of the Microcracking Process Ahead of the Main Crack

The main microstructural influences on function $\sigma(\delta_f)$ or $\sigma(\gamma)$ include: 1) the nucleation, growth and coalescence of microcracks within the fracture process zone; 2) the heterogeneous microstructure, in which the difference of elastic moduli between the aggregate and the mortar matrix is the dominant feature; 3) the strength of the interfaces between the aggregate and the matrix, which is normally weaker than the strength (or fracture toughness) of either the aggregate or the matrix; and 4) various inelastic phenomena occurring away from microcrack tips, such as frictional slip, resistance to crack closure, etc. It is difficult to study all these effects simultaneously. This study* will focus on effect 1 and take into account effect 2 only insofar as the spacing of microcracks is concerned. The main objective is to study the $\sigma(\delta_f)$-curve from the stability viewpoint, especially in terminal asymptotic stage of ligament tearing, on which little information exists at present.

Idealization of Crack System Initiation, Growth and Coalescence

To facilitate analysis, we imagine the array of aggregate pieces to be

*based on Report (23)

regular, organized as a cubic lattice (Fig. 1). This idealization was used before to derive from the initiation condition of interaggregate cracks a formula for the decrease of concrete strength as a function of the maximum aggregate size (1,2), which was found to agree with test results. Now we use the same idealization to study crack initiation as well as terminal coalescence.

If the aggregate is stiffer than the matrix (as is true of normal concretes), or if at least the aggregate-matrix interface is weaker, the microcracks are likely to initiate in the thin contact zones between aggregate pieces (Fig. 1a). The initiating microcracks which form in the fracture process zone ahead of the main crack (Fig. 1) may be imagined to be circular (penny-shaped), of diameter 2a. Their spacing in the plane (xz) of the main crack is the same as the spacing s of the large aggregate pieces, which may be given as $s = n_s d_a$ where d_a = maximum aggregate size and n_s = empirical factor greater than 1 but close to 1, perhaps $n_s = 1.1$.

The largest microcracks form on the main crack plane, and smaller microcracks form away from the crack plane. The reality is somewhere between the following two ideal limiting cases: 1) The microcracks arise only on the crack plane, thus forming a two-dimensional periodic array in the plane (xz); and 2) the microcracks form a three-dimensionally periodic array in space (xyz), being the same on the crack plane as well as the adjacent parallel planes.

Case 1 conforms to the classical Dugdale's and Barrenblatt's cohesive zone concept, in which all fracturing is imagined to be lumped into the crack plane. Case 2 corresponds to the more recent strain-softening models for cracking (1), which are now known to describe fracture realistically and in a theoretically consistent framework provided that the softening is treated in a nonlocal manner (3, 4). We analyze here only Case 1.

At the beginning, the cracks may be assumed to be circular (penny-shaped), of radius $a << s$ (Fig. 1a,b). Later the cracks develop complex irregular shapes in the crack plane as they gradually coalesce with each other. In the terminal stage of tearing, however, the uncracked area may be assumed to have again a simple shape, consisting of a periodic array of circular ligaments, of radius 2c, such that $c << s$ (Fig. 1 d,e).

For $a << s$, the situation may be idealized as a single penny-shaped crack within an infinite cylinder (Fig. 1c) of radius R. Force (and work) equivalence requires that the cylinder cross-section area of the cylinder be the same as one square of the lattice, i.e. $\pi R^2 = s^2$, from which $R = s/\sqrt{\pi}$. For $c << s$, the situation may be idealized as a single penny-shaped ligament of radius c within an infinite cylinder of radius $R = s/\sqrt{\pi}$ (Fig. 1d,e,f). While the linear elastic fracture mechanics does not apply to the macroscopic fracture, we may assume it to apply for the growths of the microcracks. The linear elastic solution for the penny-shaped crack (for any a(s)) was given by Bueckner (5) and Benthem (6); see Tada's handbook (7, p. 27.1). The solution for the penny-shaped ligament was given perhaps most accurately by Nisitani and Noda (8) (see Murakami's handbook (9, p. 643)), and previously by Bueckner (5), Benthem and Koiter (6), Harris (10), and others (cf. 7,9).

Circular (Penny-Shaped) Cracks: Initial Stage

For the penny-shaped crack, Benthem and Koiter (6) (see also p. 27.4 of Ref. 7 or p. 653 of Ref. 9) found that

$$K_I = \sigma\sqrt{R} f_1(\alpha), \qquad \alpha = a/R \tag{1}$$

where

$$f_1(\alpha) = \frac{2}{1-\alpha}\sqrt{\frac{\alpha}{\pi}(1-\alpha)}\left(1 + \frac{1}{2}\alpha - \frac{5}{8}\alpha^2 + 0.421\alpha^3\right). \tag{2}$$

K_I = Mode I stress intensity factor; $\sigma = P/\pi R^2 = P/s^2$ = average axial stress; P = resultant axial force on the cylinder (Fig. 1b). The error is under 1%. The necessary condition of propagation of the cracks is $K_I = K_{cm}$ = critical stress intensity factor of the microcracks (material constant). So the remote stress required for crack propagation is:

$$\sigma = \frac{K_{cm}}{f_1(\alpha)\sqrt{R}} . \tag{3}$$

To calculate the displacement v due to the cracks, we need to determine first the strain energy W_1 per crack that has been released due to the formation of the cracks. The energy release rate G is known to be (11, p. 108) $G = K_I^2/E'$ where $E' = E/(1-\nu^2)$, ν = Poisson ratio, E = Young's elastic modulus. By definition $\partial W_1/\partial a = 2\pi a G$, and so

$$\frac{\partial W_1}{\partial a} = 2\pi a G = 2\pi a \frac{K_I^2}{E'} = 2\pi a R \frac{\sigma^2}{E'} f_1^2(\alpha) . \tag{4}$$

Integrating, with the substitutions $a = R\alpha$ and $da = Rd\alpha$, we get

$$W_1 = 2\pi R^3 \frac{\sigma^2}{E'} G_1(\alpha) , \tag{5}$$

where

$$G_1(\alpha) = \int_0^\alpha \alpha' f_1^2(\alpha') d\alpha' . \tag{6}$$

The displacement at $y \to \infty$ due to the formation of the crack must, therefore, be

$$v = \frac{\partial W_1}{\partial P} = \frac{1}{\pi R^2} \frac{\partial W_1}{\partial \sigma} = 4R \frac{\sigma}{E'} G_1(\alpha) . \tag{7}$$

Eqs. 3 and 7 represent a parametric description of the stress-displacement curve. Choosing various values of α, and evaluating σ and v, one obtains the plot in Fig. 2 in which S = nondimensional remote applied stress and q = nondimensional displacement due to the crack;

$$S = \sigma \frac{\sqrt{R}}{K_{cm}} , \quad q = v \frac{E'}{K_{cm}\sqrt{R}} \qquad (R = s/\sqrt{\pi}) . \tag{8}$$

Small Circular Cracks: Asymptotic Approximation

An explicit, asymptotically exact solution can be obtained for the special case of small α ($\alpha \ll 1$ or $a \ll R$). Then $f_1(\alpha) = 2(\alpha/\pi)^{1/2}$, and so

$$K_I = 2\sigma\sqrt{a/\pi} . \tag{9}$$

Then $\partial W_1/\partial a = 2\pi a K_I^2/E' = 8\sigma^2 a^2/E'$. By integration, $W_1 = 8\sigma^2 a^3/3E'$. Note that in contrast to Eqs. 1, 4 and 6, these expressions are independent of R. By differentiation,

$$v = \frac{\partial W_1}{\partial P} = \frac{1}{\pi R^2} \frac{\partial W_1}{\partial \sigma} = \frac{16}{3\pi} \frac{\sigma}{E'} \frac{a^3}{R^2} = \frac{16}{3} \frac{\sigma}{E'} \frac{a^3}{s^2} , \tag{10}$$

where we used the relation $s^2 = \pi R^2$, which guarantees the force equivalence with a square grid of small circular cracks to be satisfied exactly. We see that the displacement v due to the cracks depends on R, and thus on spacing s. Setting $K_I = K_{cm}$, Eq. 9 yields $a = \pi K_{cm}^2/4\sigma^2$, and Eq. 10 becomes

$$v = \frac{\pi^2 K_{cm}^6}{12E'R^2} \frac{1}{\sigma^3} = \frac{\pi^3 K_{cm}^6}{12E's^2} \frac{1}{\sigma^3} \qquad (\sigma \le f_t') \tag{11}$$

Since the material can resist only finite stress, this equation is meaningful only for $\sigma < f_t'$ where f_t' = tensile strength. As argued already by Griffith, the creator of fracture mechanics, it follows from Eq. 9 that the material must initially contain flaws equivalent to circular cracks of radius $a = a_0$,

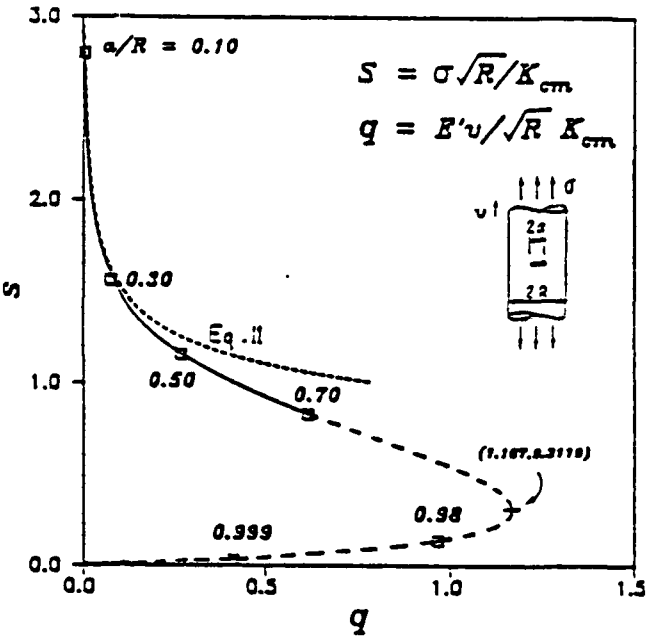


FIG. 2
Nondimensional Remote Stress vs. Nondimensional Remote Displacement Due to Cracks for the Initial State Idealization as a Circular (Penny-Shaped) Crack in a Cylinder (dashed segment, a/R > 0.70 is irrelevent since it is not initial stage).

$$a_0 = \frac{\pi}{4}\frac{K_{cm}^2}{f_t'^2} . \tag{12}$$

Although the present solution is valid only if all the initial flaws are located on one plane, the flaws are in reality distributed through the material uniformly. The initial displacement due to cracks when the cracks first become critical ($\sigma = f_t'$) then is, according to Eq. 11,

$$v_0 = \frac{\pi^3 K_{cm}^6}{12\ E's^2 f_t'^5} . \tag{13}$$

During the initial loading from $\sigma = 0$ to $\sigma = f_t'$, the cracks do not grow ($K_I < K_{cm}$), and so the σ-v relation is a straight line, $\sigma = C_0 v$, where $C_0 = f_t'/v_0$ = stiffness per unit area due to preexisting flaws;

$$C_0 = \frac{12}{\pi^3} E's^2 \left(\frac{f_t'}{K_{cm}}\right)^6 . \tag{14}$$

Assuming that the effective width w_c of the fracture process zone is such that the displacement at its boundary is nearly the same as v at $y \to \infty$, the total initial displacements δ_f across the fracture process zone before the cracks become critical is

$$\delta_f = \left(\frac{w_c}{E'} + \frac{1}{C_0}\right)\sigma \qquad (\text{for } v \le v_0) \tag{15}$$

where $E' = E/(1-\nu^2)$. E must obviously be interpreted as the theoretical elastic modulus of the material without any preexisting flaws. The elastic modulus measured in a tensile test is $E_{eff} = [E'^{-1} + (C_0 w_c)^{-1}]^{-1}$. After the cracks become critical

$$\delta_f = \frac{w_c\sigma}{E'} + \frac{\pi^3 K_{cm}^6}{6E's^2}\frac{1}{\sigma^3} \qquad (\text{for } v \ge v_0,\ a \ll s). \tag{16}$$

In analogy to Eq. 15, deformation $w_c\sigma/E'$ needs to be also added to the plot $v(\sigma)$ in Fig. 2 in order to get the plot of δ_f versus σ valid for any a.

Circular Ligaments: Terminal Stage of Crack Coalescence

Now consider the case of a penny-shaped ligament in an infinite circular cylinder of radius R. The result of elastic analysis (3,9) is:

$$K_I = \sigma\sqrt{R}\, f_2(\xi) \qquad (\xi = c/R) \tag{17}$$

where

$$f_2(\xi) = \tfrac{1}{2}\left[\frac{\pi}{\xi}(1-\xi)\right]^{\frac{1}{2}}\left(1 + \frac{\xi}{2} + \frac{3\xi^2}{8} - 0.363\xi^3 + 0.731\xi^4\right)\left(1 + 0.1\xi^2\sqrt{1-\xi}\right) \tag{18}$$

c = radius of ligament (Fig. 1e), $\sigma = P/\pi R^2$, and P = axial force. The remote applied stress required for crack propagation is:

$$\sigma = \frac{K_{cm}}{f_2(\xi)\sqrt{R}}\,. \tag{19}$$

The energy release rate is

$$-\frac{\partial W_2}{\partial c} = 2\pi c\,\frac{K_I^2}{E'} = 2\pi cR\,\frac{\sigma^2}{E'}\,f_2^2(\xi)\,. \tag{20}$$

By integration, with substitutions $c = R\xi$ and $dc = Rd\xi$,

$$W_2 = \int_1^{\xi}\frac{\partial W_2}{\partial c}\,dc = -\int_{\xi}^{1}\frac{\partial W_2}{\partial c}\,dc = 2\pi R^3\frac{\sigma^2}{E'}\,G_2(\xi) \tag{21}$$

where

$$G_2(\xi) = \int_{\xi}^{1}\xi' f_2^2(\xi')\,d\xi'\,. \tag{22}$$

The displacement v at $y \to \infty$, caused by the crack, may now be calculated as

$$v = \frac{\partial W_2}{\partial P} = \frac{1}{\pi R^2}\,\frac{\partial W_2}{\partial \sigma} = 4R\,\frac{\sigma}{E'}\,G_2(\xi)\,. \tag{23}$$

Eqs. 19 and 23 represent a parametric description of the stress-displacement curve. Choosing various values of ξ and evaluating σ and v, one obtains the plot in Fig. 3 in terms of nondimensional stress S and displacement q defined by Eq. 8.

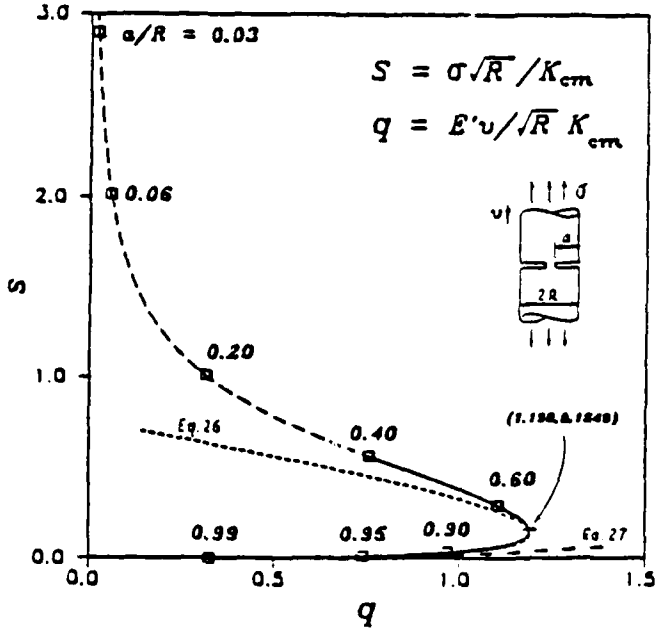


FIG. 3
Nondimensional Remote Stress vs. Nondimensional Remote Displacement Due to Crack for the Terminal State Idealization as a Circular Ligament in a Cylinder (dashed segment, a/R < 0.4, is irrelevant since it is not terminal stage).

Small Circular Ligaments: Asymptotic Approximation

An explicit asymptotically exact solution can be obtained for the special case of small ξ ($\xi \ll 1$ or $c \ll R$). Then $K_I = \frac{1}{2}\sigma(\pi R\xi^{-3})^{1/2}$, i.e.

$$K_I = \frac{P}{2\sqrt{\pi c^3}} . \tag{24}$$

Then $-\partial W_2/\partial c = 2\pi c K_I^2/E' = P^2/2E'c^2$. Note that, in contrast to Eqs. 17 and 20, these equations are independent of R, but only if they are written in terms of the applied force P rather than the applied stress σ (for very small cracks, Eqs. 9-10, this is true only if the equations are written in terms of σ rather than P). By integration, $W_2 = (P^2/2E')(c^{-1} - \kappa/s)$ where κ is a nondimensional integration constant. By differentiation

$$v = \frac{\partial W_2}{\partial P} = \frac{P}{E'}\left(\frac{1}{c} - \frac{\kappa}{s}\right). \tag{25}$$

Setting $K_I = K_{cm}$, Eq. 24 yields $c = (P/2K_{cm}\sqrt{\pi})^{2/3}$ and Eq. 25 then becomes

$$v = \frac{s}{E'}\left[\left(\frac{4\pi K_{cm}^2}{s\sigma^2}\right)^{1/3} - \kappa\right]\sigma. \tag{26}$$

For very small σ, we may neglect κ compared to the first term in the bracket, and we get

$$P = \sigma s^2 = \frac{E'^3}{4\pi K_{cm}^2} v^3 = \frac{E'^3}{4\pi K_{cm}^2} v^3 \quad \text{(small v)}. \tag{27}$$

For Eq. 26 which applies for larger (but still small) σ-values, we need to figure out the integration constant κ. If the solution were valid up to $c = R$, it would be proper to determine κ from the condition that $W_2 = 0$ for $c = R$. The solution, however, is valid only for $c \ll R$, and so we need some other condition to determine κ.

For any $\kappa > 0$, the curve $v(\sigma)$ according to Eq. 26 obviously has a point of maximum v. For arbitrary κ, this maximum point does not coincide with the exact maximum point in Fig. 3 based on Eqs. 19 and 23. Assuming that the point of v_{max} in Fig. 3 is still in the range of small enough c, we may determine κ from the condition that the graph of Eq. 26 would pass through the exact point of v_{max}, which is characterized by $S_{cr} = 1.198$ and $q_{max} = 0.1549$ (Fig. 3). Then, according to Eq. 8, $v_{max} = q_{max}K_{cm}\sqrt{R}/E'$ and $\sigma_{cr} = S_{cr}K_{cm}/\sqrt{R}$, and setting $v = v_{max}$, $\sigma = \sigma_{cr}$, Eq. 26 yields:

$$\kappa = \left(\frac{4\pi K_{cm}^2}{s\sigma_{cr}^2}\right)^{1/3} - \frac{E'}{s}\frac{v_{max}}{\sigma_{cr}} = \left(\frac{4\sqrt{\pi}}{S_{cr}^2}\right)^{1/3} - \frac{q_{max}}{S_{cr}\sqrt{\pi}} = 2.299 \tag{28}$$

Similar to the arguments related to Eqs. 14-16, the total displacement over the width w_c of the fracture process zone is, approximately,

$$\Delta_f = \left\{\frac{w_c}{E'} + \frac{s}{E'}\left[\left(\frac{4\pi K_{cm}^2}{s\sigma^2}\right)^{1/3} - \kappa\right]\right\}\sigma. \tag{29}$$

Interpretation of Mathematical Results

Although the displacement v due to cracks has been calculated for points infinitely far away from the crack plane, the displacement δ_{cr} due to cracking which arises over the finite width w_c of the fracture process zone must be nearly the same, i.e. $\delta_{cr} \approx v$. This is indicated by some solutions for interacting cracks on parallel planes at spacing s, which are available in the literature. E.g., the problem of interacting circular ligaments of diameter 2c and

spacing h in an infinite cylinder of radius R has been solved by Nisitani and Noda (8). Their results show that the interaction of cracks causes K_I to change by less than 1% if $h \geq R$ (see p. 644 of Ref. 9). For the planar problem of an infinite stack of parallel cracks of length 2a and spacing h between the crack planes, Yokobori and Ichikawa (cf. p. 202 of Ref. 9) found the crack interaction to cause a change of K_I by less than 1% if $2a < 0.08h$, and less than 10% if $2a < 0.21h$. Although Nisitani and Noda's solution is limited to $c/h \leq 0.5$, one can nevertheless invoke St.-Venant's principle to conclude that interaction of cracks on parallel planes of spacing s should be insignificant when approximately either $a \leq h/5$ or $c \leq h/5$. Since for concrete h must be about the same as s, the preceding solutions are asymptotically exact for crack initiation ($a \ll h$) as well as for terminal crack coalescence ($c \leq h$). So we may interpret v as the additional displacement across the fracture process zone due to the cracks, provided that all the cracks occur on one plane.

The most interesting feature of our results is the fact that the softening stress-displacement diagram $v(\sigma)$ exhibits a maximum displacement v_{max} after which both v and σ must decrease. This behavior is known in stability theory as the snapback instability. It means that, according to the present mathematical model, the ligament tearing cannot be stable even in a displacement-controled test after the critical state of v_{max} is attained. This conclusion, when first reached by the writer in March 1987, appeared to conflict with the current generally accepted softening models and thus to cast doubt on the assumptions of the present analysis. A stress-displacement diagram $v(\sigma)$ which declines with a negative slope all the way to a zero stress is widely used in the present practice of finite element fracture analysis of concrete, is embodied in Hillerborg's fictitious crack model (12,1) and is also implied indirectly in the crack band model (13,14,1). This present practice has apparently been also corroborated by extensive experimental evidence, particularly the measurements of the softening stress-displacement diagrams by Reinhardt and Cornelissen (15), Petersson et al. (16), Shah et al. (17), Willam et al. (18), Wecharatana (19) and others. Some of these tests included very large crack displacements, which exceeded 20-times the displacement v_p at peak stress and corresponded to a reduction of stress to less than 5% of its peak value. Yet the test specimens remained stable and no snapback instability has been reported.

During the discussion at the RILEM-SEM International Conference on Fracture of Concrete and Rock in Houston (June 1987), however, it has transpired that the experimentalists themselves have doubts about the interpretation of their test results, particularly the way the additional displacement due to cracking, accumulated over the fracture process zone, should be determined from measurements. The possibility of different interpretations of the test results then finally encouraged writing and publishing of this paper.

It has been very gratifying to the writer to learn in June 1987 that H. Horii et al. (20), studying concrete fracture, as well as M. Ortiz (21) studying ceramic composites, obtained a similar result independently. They discovered the snapback instability on the curve $\sigma(v)$ by modeling the fracture process zone ahead of the main crack as an infinite row of identical uniformly spaced line cracks on the extension line of the main crack. The plots of $v(\sigma)$ that they obtained look similar to that in Fig. 3. Their analyses, however, were planar (two-dimensional). This would be fully realistic only if the cracks ahead of the frontal edge of the main crack had the shape of infinite strips parallel to the frontal edge (in the z-direction, Fig. 1) and normal to the plane (xy) in which the problem is analyzed. Such crack strips, however, do not appear to be very realistic for concrete, especially for the initial

stage of small cracks as well as the final stage of small ligaments between the crack tips.

Due to the three-dimensional nature of the aggregate framework, the spacing of the initial small cracks, located in the contact layers between the adjacent aggregate pieces, is certain to be about the same in the direction x of fracture propagation and in the direction z normal to it (in the crack plane xz); see Fig. 1b. The same must be true for the final stage of very small ligaments, as is clear from Fig. 1e. So one may conclude that the three-dimensional model of small circular cracks should be more realistic. In this regard it may be also noted that, in a previous analysis (2,1) the assumption that the crack spacing is determined by the aggregate spacing gave a realistic and experimentally corroborated result for the dependence of concrete strength f_c' on the aggregate size.

The fracture energy G_f of concrete is equal to the cross-hatched area 023460 in Fig. 5a which is enclosed by the complete stress-displacement curve $\sigma(v)$. The softening curve $\sigma(v)$ begins at the tensile strength limit f_t'. Since infinitely small microcracks propagate only at infinitely large stress σ (Eq. 3), there must initially exist in the material certain microcracks of finite size a_0, as already mentioned (Eq. 12). The initial loading up to the strength limit occurs without any crack growth and is, therefore, elastic, described by the initial straight line (02 in Fig. 5a). In the line-crack models of the fracture process zone, such as Hillerborg's, the stress-displacement curve which describes fracture is assumed to start on the vertical axis at point f_t', and so the crack displacement δ_{cr} to be used in Hillerborg's model must be understood as $\delta_{cr} = v - \sigma/K$ where K is such that $v(f_t') - f_t'/K = 0$, i.e. $K = f_t'/v(f_t')$.

Since cracks neither extend nor shorten during unloading, the unloading path from any point on the $\sigma(v)$ curve (Fig. 5a) is a straight line toward the origin (Fig. 5a).

When stress σ is controlled, the crack becomes unstable at the peak-stress point f_t' (point 2 in Fig. 5a). When the displacement is controlled, the crack becomes unstable when the curve $\sigma(w)$ attains a vertical tangent ($\partial\sigma/\partial w = 0$); w is the load-point displacement, $w = v + \sigma/C$ where C is the elastic stiffness of the structure without the cracks, and v is the additional displacement due to the cracks. The higher the value of C, the smaller is the stress σ_{cr} at the snapback instability. The smallest σ_{cr} is obtained for $C \to \infty$, which coincides with the σ_{cr} value for the curve $\sigma(v)$ (Fig. 3). If the displacement is controlled, the stress after the snapback instability drops instantly to 0. This drop is called the snapdown. The snapdown path is dynamic and the motion is accelerated. The energy corresponding to the cross-hatched area 0454 to the left of the snapdown path in Fig. 5b is converted into kinetic energy, which is emitted as the energy of a sound wave (acoustic emission). In fact, the existence of sound emissions in a displacement-controlled fracture test implies the occurrence of snapback.

To conduct an approximate analysis in a static manner, the equilibrium snapback path (40 in Fig. 5a) must be replaced by a $\sigma(v)$-diagram that is equivalent in terms of energy. Thus, the equivalent $\sigma(v)$ diagram for static finite element analysis must preserve the correct area G_f enclosed by the $\sigma(v)$ curve, as indicated either by the vertical snapdown path 367 in Fig. 5c or the gradual equivalent softening path 89 in Fig. 5d. If static analysis is conducted on the basis of the actual snapdown path 45 in Fig. 5c, the results are incorrect in terms of energy, thus violating the most fundamental requirement of mechanics.

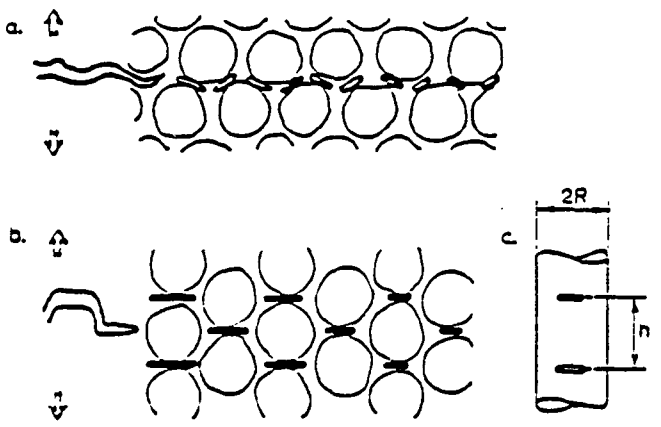


FIG. 4
Crack Systems Expected for Other Aggregate Configurations

The energy-equivalent snapdown along the vertical path 367 in Fig. 5 can actually occur only if a dynamic disturbance at point 3 imparts to the system a kinetic energy that is at least equal to the area 3463 (per unit length of fracture). Area 3463 is equal to the loss of kinetic energy along the path from point 3 to point 6 in Fig. 5c. Along the path 67 the kinetic energy is increased by an amount equal to the area 0760. The system ends at point 7 with the same kinetic energy as it had at point 3 if the areas 3463 and 0760 are equal, which is true if the area 02370 is equal to G_f.

It must be admitted that certain elastic fracture mechanisms which are neglected in the present analysis could considerably alter the shape of the $v(\sigma)$ curve including the snapback instability. These include: 1) inclination of the main crack with regard to a cubic lattice of aggregate; 2) other types of periodic lattices, e.g. of tetrahedral type (Fig. 4a,b); 3) random arrangement of aggregate pieces, which is of course the rule rather than an exception. Irregular random aggregate arrangements prevent the cracks to be all located on a single plane, contrary to the assumption of the present analysis. This has for consequence that the crack edges have not only a Mode I stress intensity factor, but also Mode II and Mode III (i.e. shear) stress intensity factors, and may therefore propagate in directions that are inclined with regard to the main crack plane (Fig. 4a). It remains to be seen whether such mechanisms can significantly alter the shape of the $v(\sigma)$ curve, especially the value of v_{max}.

The shape of the $v(\sigma)$-curve could be also significantly influenced by various inelastic phenomena. These include: 1) resistance to a crack closing due to fragmentation debris located in the crack space; 2) frictional slip on inclined cracks (Fig. 4a); and 3) irreversible deformations taking place elsewhere than microcrack tips. Since along the snapback path (40 in Fig. 5a) the portion of the crack that has formed previously is closing, the resistance to crack closing may significantly alter the snapback path and prevent complete recovery of displacement v as $\sigma \to 0$ (i.e., prevent return to point 0 in Fig. 5a). If crack closing is prevented completely, then there is no equilibrium path after the critical state of snapback (point 4, Fig. 5a) and the system can only follow the dynamic snapdown path (45 in Fig. 5b). But for reasons of stability it must follow this path anyway, and so a prevention of crack closing does not seem to alter the essential behavior.

It should also be recognized that the difference in elastic moduli between the matrix and the aggregate, as well as the fact that the microcrack fracture

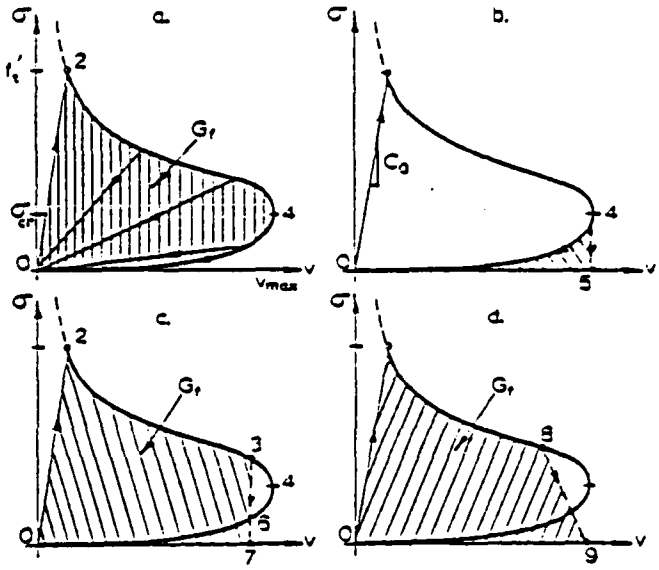


FIG. 5
Stress-Displacement Diagrams for Fracture Process Zone Modeling and Meaning of Fracture Energy.

toughness K_{cm} is probably lower at the aggregate-mortar interfaces than inside the matrix or the aggregate, has not been taken into account in the present analysis.

General Asymptotic Analysis for Small Ligament Transmitting Force or Moment

We will now show that snapback instability is a general characteristic of crack coalescence or terminal stage of crack ligament tearing in three as well as two dimensions. We consider the ligament size to be infinitely small compared to any cross section dimension of the structure. We assume the subsequent ligament shapes to be similar. Let P and M be the internal force of any direction and the internal moment about any axis transmitted across the ligament. (The special cases of P include a normal force or a shear force, and of M a bending moment or a twisting moment.) According to St.-Venant's principle, P or M can produce significant stresses and significant strain energy density only in a three-dimensional region whose size (L_1 and L_2 in Fig. 6) is of the same order of magnitude as the ligament size c. The strain energy produced in this region by P or M is

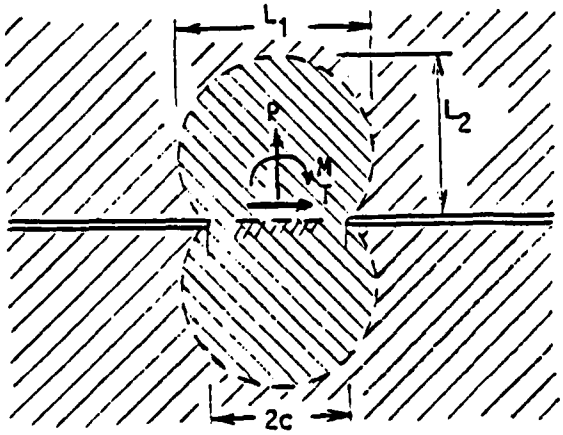


FIG. 6
Ligament Joining Two Halfspaces or Halfplanes

$$U_1 = \frac{P^2}{2EA} k_1 c = \frac{P^2}{2Ek_3 c}, \qquad U_2 = \frac{M^2}{2EI} k_2 c = \frac{M^2}{2Ek_4 c^3} \tag{30}$$

where $A = k_5 c^2$ = cross section area

of the ligament, $I = k_5 c^4$ = moment of inertia of the cross section of the ligament, and $k_1, k_2, \ldots, k_8$ = constants. The remote displacement v and rotation θ associated with P and M, respectively, are

$$v = \frac{\partial U_1}{\partial P} = \frac{P}{Ek_3 c}, \qquad \theta = \frac{\partial U_2}{\partial M} = \frac{M}{Ek_4 c^3}. \tag{31}$$

The energy release rates due to P and M per unit circumference of the ligament cross section are:

$$G_1 = -\frac{1}{k_7 c}\frac{\partial U_1}{\partial c} = \frac{P^2}{2Ek_3 k_7 c^3}, \qquad G_2 = -\frac{1}{k_8 c}\frac{\partial U_2}{\partial c} = \frac{3M^2}{2Ek_4 k_8 c^5}. \tag{32}$$

Setting $G_1 = G_f$ or $G_2 = G_f$ where $G_f = K_c^2/E$ = fracture energy of the material (a constant), we have:

$$\text{for P:}\quad c = \left(\frac{P^2}{2Ek_3 k_7 G_f}\right)^{1/3}; \qquad \text{for M:}\quad c = \left(\frac{3M^2}{2Ek_4 k_8 G_f}\right)^{1/5}. \tag{33}$$

Substituting this into Eq. 31 we get, for very small ligament size c, the asymptotic approximations:

$$\text{for P:}\quad v = c_1 P^{1/3}; \qquad \text{for M:}\quad \theta = c_2 M^{-1/5} \tag{34}$$

where c_1, c_2 = constants. Note that Eq. 34 for P agrees with Eq. 27 that we derived before in a different manner.

From Eq. 34 we conclude that for force loading of the ligament the curve v(P) must return to the origin (v = P = 0) as P → 0 (c → 0). This means that there must be a snapback at some finite P-value.

On the other hand, for moment loading of the ligament, the curve $\theta(M)$ tends to infinity as M → 0 (c → 0). So there can be no snapback.

For two-dimensional problems a similar asymptotic analysis is possible, but only for the moment loading. We have $I = k_6 bc^3$ where b = thickness of the body, and instead of Eqs. 30-32 we get

$$U_2 = \frac{M^2}{2EI} k_2 c = \frac{M^2}{2Ek_4 bc^2} \tag{35}$$

$$\theta = \frac{\partial U_2}{\partial M} = \frac{M}{Ek_4 bc^2} \tag{36}$$

$$G_2 = -\frac{\partial U_2}{\partial c} = \frac{M^2}{Ek_4 bc^3}. \tag{37}$$

Setting $G = G_f$, we have $c = (M^2/G_f E k_4 b)^{1/3}$, and substituting this into Eq. 36, we get, for small c:

$$\theta = c_2 M^{-1/3}. \tag{38}$$

So for moment loading in two dimensions there cannot be any snapback either.

For two-dimensional problems in which the ligament is loaded by a force, the foregoing approach fails because, as it turns out, the curve v(P) is not of a power type as P → 0. For a sufficiently short ligament, the stress field must be the same as that near a ligament joining two elastic halfplanes. For that problem it is known that $K_I = (P/b)(2/\pi c)^{1/2}$ where P = normal (centric) force and c = half-length of the ligament (Fig. 6). Therefore $-\partial W/\partial c = bG = bK_I^2/E' = 2P^2/\pi E' bc$, and by integration the total strain energy release is:

$$W = -\frac{2P^2}{\pi E' b} \ln \frac{c}{c_0} \tag{39}$$

where c_0 = integration constant. Furthermore,

$$v = -\frac{\partial W}{\partial P} = -\frac{4P}{\pi E' b} \ln \frac{c}{c_0} . \tag{40}$$

Setting $K_I = K_c$ = critical value of K_I, we also have $c = 2P^2/\pi b^2 K_c^2$ and substitution into Eq. 40 yields

$$v = -\frac{4P}{\pi E' b} \ln \frac{2P^2}{\pi b^2 K_c^2 c_0} \tag{41}$$

The curve $v(P)$ described by this equation is not of a power type, which explains why the type of approach used in Eqs. 30-38 would fail. The curve $v(P)$ obviously exhibits a snapback since, for $P \to 0$, $\lim v = 0$. The critical state is characterized by the condition $\partial v/\partial P = 0$, which yields the critical value $P_{cr} = (\pi c_0/2)^{\frac{1}{2}} K_c b/e$, from which $v_{max} = (2c_0/\pi)^{\frac{1}{2}} K_c/E'e$.

From the fact that ligament tearing in Mode I or II or III always exhibits snapback if the force across the ligament is nonzero it follows that the stress-displacement curve for the microcrack patterns in Fig. 4 must, too, terminate with a snapback.

Conclusions

1. To model the fracture process zone of a heterogeneous brittle material such as concrete, it may be assumed that the initial stage of cracking consists of a two-dimensional array of small circular (penny-shaped) cracks in the plane of the main crack, and the final stage consists of a two-dimensional array of small circular ligaments. Both cases may be approximately solved from the known value of the Mode I stress intensity factor for a penny-shaped centrally located transverse crack in an axially loaded long cylinder, and a small circular ligament in such a cylinder. These solutions may be considered to be asymptotically exact for very small cracks or for very small ligaments.

2. The spacing of the initial small circular cracks as well as the final circular ligaments may be considered to be determined by the spacing of the large aggregate pieces in concrete.

3. The solution indicates that the curve of the additional transverse displacement v due to the crack versus the remote normal stress σ exhibits a softening segment which does not descend all the way to zero stress, as assumed in the current fracture models, but terminates with a critical state of snapback instability at which the curve $\sigma(v)$ has a vertical tangent (Fig. 5). The displacement at this critical state, v_{max}, is the maximum possible displacement due to the cracks. After this critical state, the equilibrium path $\sigma(v)$ approaches the origin as a cubic parabola with a positive slope. This postcritical path is unstable and a dynamic snapdown occurs in reality (Fig. 5b). For approximate static analysis an energy-equivalent softening path without snapback (Fig. 5c,d) needs to be introduced.

4. The present calculation of the stress-displacement curve neglects: 1) the situation in which the cracks are inclined and are loaded in Modes II and III; 2) possible strong interactions between adjacent cracks (other than those meeting at the same ligament); 3) the effect of the differences of elastic moduli between the aggregate and the matrix; 4) the fact that the aggregate-matrix interface may be weaker than the adjacent solid material; 5) the resistance to crack closing due to fragments located in the crack space; and 6) the inelastic phenomena taking place elsewhere than microcrack tips. It remains to be determined whether the essence of the preceding theoretical conclusions remains valid if all these phenomena are taken into account.

5. General asymptotic analysis shows that when the crack ligament transmits a force, its final tearing is always characterized by snapback instability and

there exists a maximum displacement. When, however, the ligament transmits only a moment, there is no snapback instability and the rotation due to the applied moment grows beyond any bounds. These conclusions are valid for three-dimensional as well as two-dimensional situations, for normal as well as shear forces, and for moments about any axis, bending as well as torsional (i.e., for Modes I, II and III).

Acknowledgment. - Grateful appreciation is due to the Air Force Office of Scientific Research for financial support under Contract No. F49620-87-C-0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu. Thanks are also due to Graduate Research Assistants M. Tabbara and F.-B. Lin for help in computer plotting.

Appendix. - Stability of Three-Point Bent Specimen

The stability analysis of crack ligament tearing is not only of interest for micromechanics-based modeling of the fracture process zone, but also for various macroscopic problems. To illustrate and corroborate the general conclusions based on St.-Venant's principle, in particular Eq. 38, consider as an example the three-point bent beam specimen (Fig. 7), for which L = span, h = beam depth, a = length of crack (or of crack plus notch), c = h - a = length of ligament, b = beam thickness. In the literature there exist for this specimen various approximate K_I-formulas. However, those which are often used are invalid for small a/h. Nevertheless, a formula which is valid for the entire range $0 < a/h < 1$, is asymptotically exact for $c \to 0$, and generally has an error under 0.5%, has been derived by Srawley (22) for the case L = 4h:

$$K_I = \frac{3PL}{2bh^2}\sqrt{\pi a}\,F_1(\alpha), \quad F_1(\alpha) = \frac{1.99 - \alpha(1-\alpha)(2.15 - 3.93\alpha + 2.7\alpha^2)}{(1+2\alpha)(1-\alpha)^{3/2}} \tag{42}$$

where $\alpha = a/h$. The rate of release of the total strain energy W of the whole specimen is

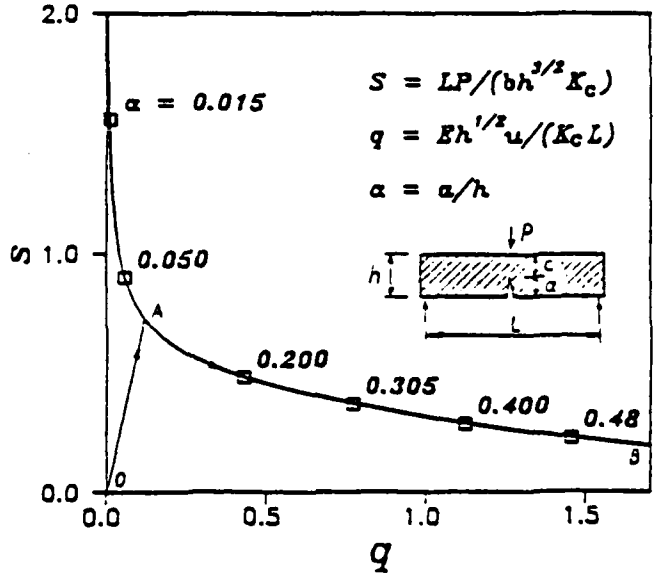


FIG. 7
Nondimensional Stress-Displacement Diagram
Calculated for the Three-Point Bent Beam Specimen.

$$\frac{\partial W}{\partial a} = b\,\frac{K_I^2}{E} = \frac{9\pi}{4bE}\,\frac{P^2L^2}{h^3}\,\alpha F_1^2(\alpha). \tag{43}$$

By integration

$$W = \int_0^a \frac{\partial W}{\partial a}\,da = \int_0^\alpha \frac{\partial W}{\partial \alpha}\,h\,d\alpha = \frac{9\pi}{4bE}\,\frac{P^2L^2}{h^2}\,F_2(\alpha) \tag{44}$$

in which

$$F_2(\alpha) = \int_0^\alpha \alpha' F_1^2(\alpha')d\alpha'. \tag{45}$$

The deflection v under the load may now be calculated as

$$v = \frac{\partial W}{\partial P} = \frac{9\pi}{2bE}\,\frac{PL^2}{h^2}\,F_2(\alpha). \tag{46}$$

If the crack is propagating, we must have $K_I = K_c$ = critical value of K_I (i.e., fracture toughness of the material). Thus, Eq. 1 yields:

$$P = \frac{2bK_c}{3\sqrt{\pi}}\,\frac{h^{3/2}}{L\sqrt{\alpha}\,F_1(\alpha)}\,. \tag{47}$$

Defining the nondimensional load S and the nondimensional displacement q,

$$S = \frac{L}{K_c bh^{3/2}}\,P, \qquad q = \frac{E\sqrt{h}}{K_c L}\,u \tag{48}$$

we have from Eqs. 7 and 6:

$$S = \frac{2}{3\sqrt{\pi\alpha}\,F1(\alpha)}\,, \qquad q = 3\sqrt{\frac{\pi}{\alpha}}\,\frac{F_2(\alpha)}{F_1(\alpha)}\,. \tag{49}$$

The last equations represent a parametric definition of the non-dimensional load-deflection diagram S(q). Evaluation of S and q for various α-values yields the diagram in Figs. 7 or 8. We see that indeed there is no snapback, as indicated already by Eq. 38. The extent to which the actual measured P(v) diagram in the three-point bent test deviates from the shape in Fig. 8 is an indication of inelastic behavior and existence of a large fracture process zone.

Consider now the asymptotic behavior for small ligament, c/h → 0. For that case Eq. 42 simplifies to the form:

$$K_I = 0.995(PL/b)(\pi/h^3)^{1/2}\xi^{-3/2} \qquad (\xi = c/h). \tag{50}$$

Using $\partial W/\partial\xi = -h\,\partial W/\partial a = -hbK_I^2/E$, and integrating, we get $W = 3.11\,P^2L^2\xi^{-2}/(2bh^2E)$, from which

$$v = \partial W/\partial P = 3.11\,PL^2\xi^{-2}/(bh^2E). \tag{51}$$

Setting $K_I = K_c$ and solving ξ from Eq. 50, Eq. 51 becomes

$$v = (1.45/E)(b^2K_c^2L/\pi)^{2/3}\,P^{-1/3}. \tag{52}$$

Exponent -1/3 agrees with Eq. 38 derived before for two-dimensional moment loading in general.

For very small α, $v \sim P^{-6}$, as can be verified from Eqs. 46-47.

References

1. Z.P. Bazant, "Mechanics of Distributed Cracking," Applied Mechanics Reviews 39 (5), 675 (1986).

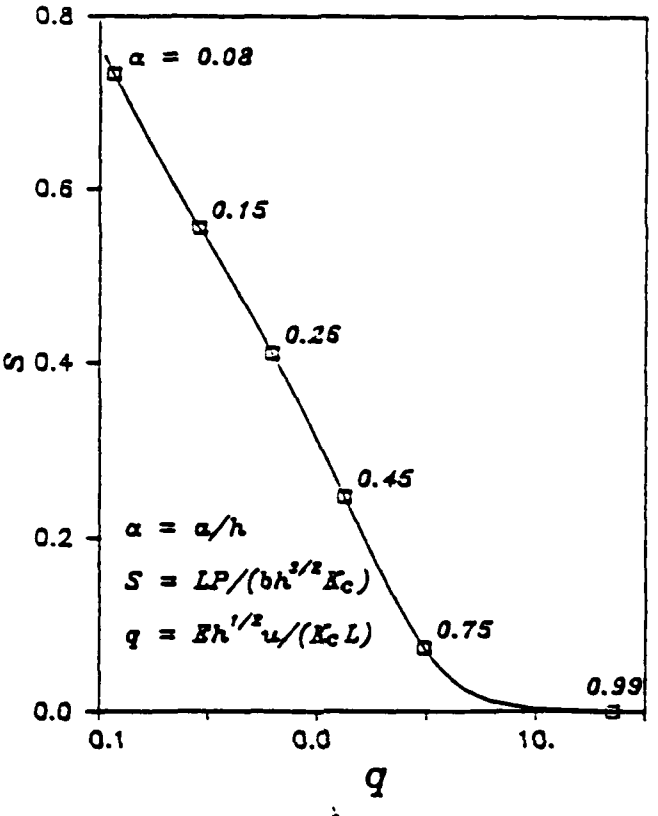


FIG. 8
Same as Fig. 7 but in Semilogarithmic Scale

2. Z.P. Bazant and J.-K. Kim, Discussion closure to the paper "Size Effect in Shear Failure of Longitudinally Reinforced Beams," ACI Journal 82 (4), 581 (1985).
3. G. Pijaudier-Cabot and Z.P. Bažant, "Nonlocal Damage Theory," Report No. 86-8/428n, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL; also J. of Engng. Mechanics ASCE (1986).
4. Z.P. Bažant and G. Pijaudier-Cabot, "Modeling of Distributed Damage by Nonlocal Continuum with Local Strain," Preprints, 4th Intern. Conf. on Numerical Methods in Fracture Mechanics," ed. by A.P. Luxmore, R.J. Owen (U. of Wales, Swansea, U.K.) and M.F. Kanninen (Southwest Research Inst., San Antonio), held in San Antonio, Texas (1987), pp. 411-432.
5. H.F. Bueckner, Discussion of the paper by P.C. Paris and G.C. Sih, "Stress Analysis of Cracks," Fracture Toughness Testing and Its Applications, ASTM-STP 381, 82 (1965).
6. J.P. Benthem and W.T. Koiter, Mechanics of Fracture 1, Method of Analysis and Solutions of Crack Problems, ed. by G.C. Sih, Noordhoff Intern. Publishing, Leyden, 131-178 (1973).
7. H. Tada, P.C. Paris and J.K. Irwin, "The Stress Analysis of Cracks Handbook," 2nd ed., Paris Productions Inc., 226 Woodbourne Dr., St. Louis, Mo. (1985).
8. H. Nisitani and N. Noda, "On the Tension of a Cylindrical Bar Having an Infinite Row of Circumferential Cracks," Trans. Japan Soc. Mech. Engrs. 50 (543), 847 (1984).

9. Y. Murakami, Ed., Stress Intensity Factors Handbook, Pergamon Press, Oxford-New York (1987).
10. D.O. Harris, "Stress Intensity Factors for Hollow Circumferentially Notched Round Bars," Trans. ASME, Ser. D, J. Basic Engng. 89, 49 (1967).
11. J.F. Knott, Fundamentals of Fracture Mechanics, Butterworth, London (1983).
12. A. Hillerborg, M. Modéer and P.E. Petersson, "Analysis of Crack Formation and Crack Growth in Concrete by Means of Fracture Mechanics and Finite Elements," Cement Concr. Res. 6 (6), 773 (1976).
13. Z.P. Bažant, "Instability, Ductility and Size Effect in Strain-Softening Concrete," J. Eng. Mech. ASCE 102, 331 (1976); discussions, 103, 357, 104, 501, based on Struct. Eng. Rep. No. 74-8/640, Northwestern University, Aug. 1974.
14. Z.P. Bažant and B.H. Oh, "Crack Band Theory for Fracture of Concrete," Mat. Construct. 16 (93), 155 (1983).
15. R. Ballarini, S.P. Shah and L.M. Keer, "Crack Growth in Cement Based Composites," Eng. Fract. Mech. 20 (3) (1984).
16. P.E. Petersson, Crack Growth and Development of Fracture Zones in Plain Concrete and Similar Materials, doctoral dissertation, Lund Inst. of Tech.
17. V.S. Gopalaratnam and S.P. Shah, "Softening Response of Plain Concrete in Direct Tension," ACI Journal 82 (3), 310 (1985).
18. K.J. Willam, B. Hurlbut and S. Sture, "Experimental, Constitutive and Computational Aspects of Concrete Failure" in Preprints, U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concretes, Tokyo, pp. 149-172 May (1985).
19. M. Wecharatana, Paper orally presented at SEM-RILEM Intern. Conf. on Fracture of Concrete and Rock, Houston, June (1987).
20. H. Horii, A. Hasegawa and F. Nishino, "Process Zone Model and Influencing Factors in Fracture of Concrete," Preprints, SEM-RILEM Intern. Conf. on Fracture of Concrete and Rock, Houston (1987), ed. by S.P. Shah and S.E. Swartz, publ. by Soc. for Exper. Mechanics, pp. 299-307 (1987).
21. M. Ortiz, "Microcrack Coalescence and Macroscopic Crack Growth Initiation in Brittle Solids," manuscript prepared at Brown University and privately communicated to Bazant in June 1987; also Int. J. of Solids and Structures, in press.
22. J.E. Srawley, "Wide Range Stress Intensity Factor Expression for ASTM E399 Standard Fracture Toughness Specimen," Int. J. of Fract. Mech. 12, 475 (1976).
23. Z.P. Bažant, "Snapback Instability at Crack Ligament Tearing and Its Implication for Fracture Micromechanics," Report No. 87-6/498s, Center for Concrete and Geomaterials, Northwestern University, June (1987).

APPENDIX B4

MICROPLANE MODEL FOR BRITTLE-PLASTIC MATERIAL: I. THEORY

By Zdeněk P. Bažant,[1] Fellow, ASCE, and Pere C. Prat,[2] Student Member, ASCE

ABSTRACT: A generalized microplane model for brittle-plastic heterogeneous materials such as concrete, which describes not only tensile cracking but also nonlinear triaxial response in compression and shear, is presented. The constitutive properties are characterized separately on planes of various orientations within the material, called the microplanes, on which there are only few stress and strain components and no tensorial requirements need to be observed. These requirements are satisfied automatically by integration over all spatial directions. The state of each microplane is characterized by normal deviatoric and volumetric strains and shear strain, which makes it possible to match any Poisson ratio. The microplane strains are assumed to be the resolved components of the macroscopic strain tensor. The central assumption is that on the microplane level the stress-strain diagrams for monotonic loading are path-independent and that all the path dependence on the macrolevel is due to unloading, which happens selectively on microplanes of some orientations. The response on the microplane is assumed to depend on the lateral normal strain, which does no work. In consequence, the incremental elastic moduli tensor is nonsymmetric, which is necessary to model friction and dilatancy. This tensor is also generally anisotropic and fully populated (i.e., none of its elements can be prescribed as zero). The model involves many fewer free material parameters than the existing comprehensive macroscopic phenomenologic constitutive models for concrete.

INTRODUCTION

During the last decade we have witnessed a tremendous progress in the development of constitutive models for brittle-plastic heterogeneous materials such as concrete, motivated chiefly by the needs of finite element analysis. The existing models are basically of two kinds: (1) Macroscopic phenomenologic models; and (2) micromechanics models. The macroscopic phenomenologic models include: (1) The classical plastic models (Chen and Chen 1975; Takahashi and Marchertas 1985), which include Drucker-Prager plasticity, the cap model, the critical state theory (originally developed for soils), models with rounded-triangle deviatoric section (Willam and Warnke 1974) and with slanted ellipse volumetric section (Lin et al. 1987), etc.; (2) the bounding surface model (Yang et al. 1985) which is really plasticity with a refined hardening rule; (3) the rotating active plane model (Zubelewicz and Bažant 1987); (4) the hypoelastic models

[1]Prof. of Civ. Engrg., Dept. of Civ. Engrg., Northwestern Univ., Evanston, IL 60208.

[2]Grad. Res. Asst.; Res. Assoc., Materials Sci. Inst. (CSIC), Barcelona, Spain.

Note. Discussion open until March 1, 1989. Separate discussions should be submitted for the individual papers in this symposium. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on August 7, 1987. This paper is part of the *Journal of Engineering Mechanics*, Vol. 114, No. 10, October, 1988. ©ASCE, ISSN 0733-9399/88/0010-1672/$1.00 + $.15 per page. Paper No. 22823.

(these include the orthotropic models, which, however, violate tensorial form invariance); (5) the total strain models (deformation theory of plasticity), also called nonlinear elastic models (Bažant and Tsubaki 1980; Cedolin et al. 1977; Gerstle et al. 1980; Gerstle 1981; Kotsovos and Newman 1978); (6) the continuum damage mechanics models (Janson and Hult 1977; Kachanov 1958; Krajcinovic and Fonseka 1981; Mazars 1981), which can be seen as a refinement of the total strain models; (7) the fracturing material of Dougill (1976); (8) the plastic-fracturing theory (Bažant and Kim 1979), which is a combination of the fracturing and plastic models; (9) the viscoplastic and endochronic models (Bažant 1974; Bažant and Bhat 1976; Bažant and Shieh 1980); and (10) the damage model of Ortiz (1985), in which all 21 anisotropic elastic moduli are treated as damage variables.

Despite significant initial success during the 1970s, however, the macroscopic phenomenologic approach seems to have run into a dead-end street. Probably only relatively minor further improvements can be expected. As the experimental data base in the literature is expanding, many aspects of the material response apparently cannot be modeled by a unified general theory in which the material parameters are the same for all types of loading.

A greater promise is offered by the second kind of models (i.e., those based on micromechanics of the inelastic phenomena in the material microstructure). These models include: (1) the classical slip theory of plasticity and its recent variants and extensions called the microplane models; and (2) particle simulation, which started as the distinct element method of Cundall for granular solids, and was recently developed for the fracturing of concrete as the interface element model (Zubelewicz and Bažant 1987). The latter models, in which a random simulation of the aggregate framework in the microstructure of concrete is used, pose enormous demands for computer time and storage when large structures or three-dimensional analysis is considered.

This study will focus on the microplane model. The original idea underlying this approach is due to G. I. Taylor (1938), who proposed in 1938 that the stress-strain relation be specified independently on various planes in the material, assuming that either the stresses on such a plane (presently called the microplane) are the resolved components of the macroscopic stress tensor, which is called the static constraint, or that the strain components on such plane are the resolved components of the macroscopic strain tensor, which is called the kinematic constraint. In practice, however, only the static constraint had been used until the recent development of the microplane model for tensile fracturing of concrete (Bažant and Gambarova 1984; Bažant 1984; Bažant and Oh 1985). The responses on the planes of various orientation were related to the macrolevel simply by superposition, while recently (Bažant and Gambarova 1984; Bažant 1984; Bažant and Oh 1985) the principle of virtual work has been used for that purpose, for both static and kinematic constraints. The static constraint formulation was developed initially for polycrystalline metals by Batdorf and Budianski in 1949 and is known as the slip theory of plasticity. This theory has long been viewed as the best available description of the plastic behavior of metals, especially for highly nonproportional loading paths with sharp corners. The static constraint approach was

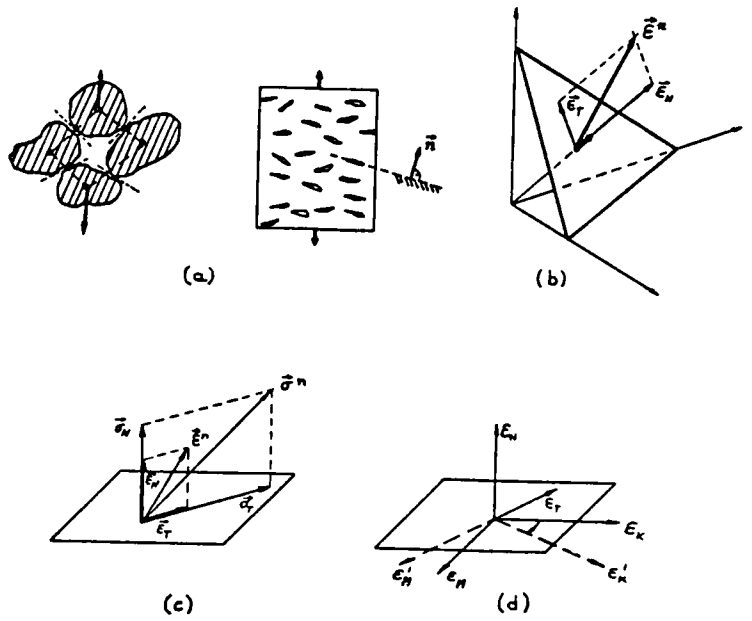


FIG. 1. (*a*) Assumed Microstructure; (*b*) Strain Components on Microplane; (*c*) Stress and Strain on Microplane; and (*d*) Lateral Strains

further developed for soils and rocks by Zienkiewicz and Pande (1977) and Xiong (1982), and more recently for creep of clay by Bažant and Kim (1986) and Bažant and Prat (1987).

In an effort to pursue this approach for tensile fracturing of concrete it was found (Bažant and Gambarova 1984; Bažant 1984; Bažant and Oh 1985) that, due to strain softening, the static constraint is impossible since it always leads to unstable response of snapback type. This problem has been overcome by adopting the kinematic constraint. Such a constraint seems to better describe the relation between the microscopic deformations and the macroscopic strain in an aggregate framework (Bažant and Gambarova 1984; Bažant 1984; Bažant and Oh 1985). The new term *microplane* had to be introduced (Bažant and Gambarova 1984; Bažant 1984; Bažant and Oh 1985) because the classical term *slip theory* gives the connotation of plasticity, which makes no sense as a description of concrete cracking. The prefix *micro* refers to the fact that the behavior is characterized on the weak planes of various orientations that are found in the microstructure, (e.g., the interaggregate contact planes) [Fig. 1(*a*)]. Keep in mind, though, that the microplane model does not represent a complete micromechanics formulation, which would have to reflect also spatial interactions between the weak planes, cracks and various defects on the microstructural level. The microplane model takes into account the microscopic interactions among various orientations but not those at distance. The latter ones are perhaps best described by nonlocal concepts.

From the mathematical viewpoint, the basic advantage of the mi-

croplane model is its conceptual simplicity. In contrast to the phenomenologic macroscopic models, one does not have to enforce any tensorial invariance requirements for the stress-strain relations, since the stress-strain relations refer to a single plane which cannot be rotated, by definition. The tensorial invariance requirements are automatically fulfilled postfacto (i.e., after the stress-strain relation is stated); this is done simply by integrating over the microplanes of all spatial orientations.

Due to this conceptual simplicity, one might hope to describe the material with fewer material parameters. This will be borne out by our analysis, which shows that only about four independent material parameters are needed to describe the general nonlinear triaxial behavior of any concrete with practically sufficient accuracy. By contrast, the existing phenomenologic macroscopic models for concrete, which are intended to describe a broad range of responses, all involve at least 15 unknown material parameters. When the endochronic theory, the first comprehensive nonlinear model for concrete, was presented in 1974 (Bažant 1974), the fact that it involved so many material parameters was unappealing and was justly criticized. Nevertheless, despite extensive efforts, neither this theory nor other general phenomenologic macroscopic models developed subsequently were able to succeed with fewer than about 15 material parameters. This attests to the presence of a fundamental limitation in the phenomenologic macroscopic approach.

The microplane model for concrete initially covered only tensile cracking (Bažant and Gambarova 1984; Bažant 1984; Bažant and Oh 1985), although it could also describe, by means of inclined tensile microcracks, the shear response of cracks or crack bands (Bažant and Gambarova 1984). The main objective of Part I of the present study is to develop a general microplane model that is applicable both to tension and compression, and both to brittle cracking and plastic response that characterize the damage in concrete. A further objective is to use as few material parameters as possible, and to cover a broader range of responses than the previous macroscopic phenomenologic models. Part II of this study will calibrate and verify this model by experimental data.

Basic Hypotheses

Hypothesis I

The strains on any microplane are the resolved components of the macroscopic strain tensor ε_{ij}, which represents a tensorial kinematic constraint [Fig. 1(*b*)]. (However, we will also introduce an additional static constraint of the nontensorial, scalar type.)

Hypothesis II

The microplane resists not only normal strains ε_N, but also shear strains, ε_{T_i}, and their vector has the same direction as the vector of shear stresses σ_{T_i} (i.e., $\sigma_{T_i} \sim \varepsilon_{T_i}$) [Fig. 1(*c*)].

Hypothesis III

The response on each microplane depends on the mean lateral strain ε_L, which is equivalent to a dependence on the volumetric strain ε_V ($\varepsilon_V = \varepsilon_{kk}/3$).

Hypothesis IV

The stress-strain curves of each microplane are path independent as long as this microplane suffers no unloading. During unloading and reloading, the curves of the stress and strain differences from the state at the start of unloading are also path independent.

Hypothesis V

The volumetric, deviatoric and shear responses on each microplane are mutually independent.

As already mentioned, hypothesis I makes it possible to obtain stable response during strain softening while the classical hypothesis of static constraint would not. At the same time one should be aware of the fact that the kinematic constraint in general yields the stiffest possible response for the given stiffness properties of the material components. This is suggested by the fact, known from the theory of elastic composites, that the Voigt estimate for the elastic modulus, which corresponds to a kinematic constraint, represents an upper bound, while the Reuss estimate, which corresponds to a static constraint, represents a lower bound. The reality would no doubt be better described by a mixed kinematic-static constraint. In fact, we will need to introduce an additional static constraint, but we will be able to keep it nontensorial (scalar), which preserves simplicity. Due to the mixed constraint, our model will not be the stiffest possible.

According to hypothesis I, the strain vector on a microplane whose direction cosines are n_i is $\varepsilon_j^n = \varepsilon_{jk} n_k$. The normal strain component and its vector then are

$$\varepsilon_N = n_j \varepsilon_j^n = n_j n_k \varepsilon_{jk} \quad \ldots\ldots (1a)$$

$$\varepsilon_{N_i} = n_i n_j n_k \varepsilon_{jk} \quad \ldots\ldots (1b)$$

The Latin lowercase subscripts refer to cartesian coordinates x_i ($i = 1, 2, 3$). Repeated Latin lower case subscripts indicate summation over 1, 2, 3. The magnitude of the strain vector on the microplane is $\|\varepsilon^n\| = \sqrt{\varepsilon_j^n \varepsilon_j^n} = \sqrt{n_i \varepsilon_{ji} n_k \varepsilon_{jk}}$. The vector of the shear (tangential) strain component is (Fig. 1*b*) $\varepsilon_T = \varepsilon^n - \varepsilon_N$. So the shear strain components and the shear strain magnitudes are:

$$\varepsilon_{T_i} = (\delta_{ij} - n_i n_j) n_k \varepsilon_{jk} = \frac{1}{2} (n_j \delta_{ik} + n_k \delta_{ij} - 2 n_i n_j n_k) \varepsilon_{jk} \quad \ldots\ldots (2a)$$

$$\varepsilon_T = \sqrt{\varepsilon_{T_i} \varepsilon_{T_i}} = \sqrt{n_k \varepsilon_{jm} n_m (\varepsilon_{jk} - n_i n_j \varepsilon_{ik})} \quad \ldots\ldots (2b)$$

in which δ_{ij} is the Kronecker's unit delta tensor. The purpose of the longer expression for ε_{Ti} is to make the tensor that multiplies ε_{jk} symmetric with respect to j and k. The reason for doing this is that the antisymmetric part of this tensor has no effect on the value of ε_{Ti} since ε_{jk} is symmetric.

In the original microplane model for tensile cracking of concrete, simplicity was achieved by neglecting the shear stiffness of the microplanes and dealing only with normal strains ε_N. On the macroscopic level, shear stiffness was provided by resistance against compression in inclined directions, which was exploited in the crack band microplane model (Bažant and Gambarova 1984). However, this approach proved

insufficient for the modeling of the response in compression, for various reasons. For example, simulation of the uniaxial compression tests with the original microplane model does yield a peak stress value; however, shortly into the strain-softening regime the system becomes unstable and exhibits snapback, which does not correspond to reality. After exploring a number of alternatives it has been concluded that the microplane model for general behavior must include not only the normal strain but also the shear strain on each microplane [see also Bažant (1984)]. It may be noted that the classical slip theory of plasticity involves only plastic shear response on each slip plane and ignores the normal response, which is the opposite of the original microplane model for tensile cracking.

Even with the inclusion of shear strains, however, a good description of the response in compression cannot be achieved, as learned in extensive numerical simulations. It appears that the response on a microplane is also affected by the mean normal strain ε_L, called the lateral strain, which influences the microplane from the planes that are normal to the microplane. This strain is defined as $\varepsilon_L = (\varepsilon_K + \varepsilon_M)/2$, where ε_K and ε_M are the normal strains in two orthogonal directions normal to that of ε_N; see Fig. 1(*d*). The value of ε_L is invariant with regard to rotations of the directions K and M around the normal to the microplane. It appears that the lateral strain ε_L considerably affects the brittleness of the response, although it does no work on the normal and shear stresses acting on the given microplane.

Instead of using the lateral strain ε_L, one may equivalently and more conveniently use the volumetric strain ε_V, which is expressed as $\varepsilon_V = (\varepsilon_N + \varepsilon_K + \varepsilon_M)/3$, and is a three-dimensional invariant. Defining the deviatoric normal strain as $\varepsilon_D = \varepsilon_N - \varepsilon_V$, one has $\varepsilon_D = \frac{2}{3}(\varepsilon_N - \varepsilon_L)$. Thus, hypothesis III is equivalent to the requirement that the volumetric and deviatoric normal strain components ε_V and ε_D should be used separately in the stress-strain relation for each microplane.

There is also another physical argument for introducing the shear strains on each microplane and for separating the volumetric and deviatoric normal strains. As is well known (Bažant and Oh 1985), the microplane model which exhibits only normal stiffness yields always the elastic Poisson's ratio $\nu = 0.25$ (in three dimensions; in two dimensions it is 1/3). Therefore, the artifice of a series coupling with a fictitious additional elastic element had to be used previously to adjust the Poisson's ratio to a physically correct value. If both the shear and normal strains are considered in a kinematically constrained microplane model but the deviatoric and volumetric components are not separated, the Poisson's ratio is restricted to the limits $-1 \leq \nu \leq 0.25$. This range could describe the Poisson's ratio observed for concrete, but the fact that Poisson's ratio does not exceed 0.25 is suspicious on physical grounds, because the microplane model should in principle also work for other materials for which the Poisson's ratio might exceed 0.25. It has been discovered by Bažant that an arbitrary Poisson's ratio, satisfying the thermodynamic requirements $-1 < \nu < 0.5$, can be obtained if, in addition to the consideration of both the shear and normal strains, the normal strains are further split into the volumetric and deviatoric components. It may be also noted that the additional scalar static constraint that we introduce later will have no effect on the elastic response, and therefore none on the Poisson's ratio.

As a reasoning behind hypothesis IV, one may recall that the nonlinear triaxial response of concrete in the hardening regime (prepeak states) is nearly path independent for monotonic loading paths. This is evident from the success of the total strain theory (deformation theory) of Cedolin et al. (1977, 1983) [later adapted and restated by Kotsovos and Newman (1978)], as well as the model of Gerstle et al. (1980) and Gerstle (1981) and the comprehensive model of Bažant and Tsubaki (1980). From these studies it transpired that the path dependence of response becomes significant only if there is unloading or strain softening. Strain softening, however, always implies unloading of the material on planes of some orientations at the microstructural level. This suggests that the main reason for the path dependence of concrete, as well as other similar materials, should be the unloading.

The typical response of the microplane model combines loading on the microplanes of some orientations and unloading on the microplanes of other orientations. There are numerous possible combinations of such loading and unloading on planes of various orientations, which cannot be captured by the existing phenomenologic macroscopic models. Thus it becomes clear that even if the microplane model assumes path independence on each microplane, the overall material response can and does exhibit significant path dependence. This has been the reason for introducing hypothesis IV.

The consequence of this hypothesis is that for monotonic loading, the secant moduli for each microplane can be considered to be functions of the current strains and stresses only. We will assume the stress-strain relation for monotonic loading to be of the form

$$\sigma_V = C_V(\varepsilon_V, \varepsilon_D, \varepsilon_T)\varepsilon_V \quad \ldots\ldots (3a)$$

$$\sigma_D = C_D(\varepsilon_V, \varepsilon_D, \varepsilon_T)\varepsilon_D \quad \ldots\ldots (3b)$$

$$\sigma_T = C_T(\varepsilon_V, \varepsilon_D, \varepsilon_T, \underline{\sigma})\varepsilon_T \quad \ldots\ldots (3c)$$

in which C_V, C_D, and C_T are the secant moduli, which depend on ε_V, ε_D, and ε_T, and C_T also on an invariant of stress tensor $\boldsymbol{\sigma}$. Modeling experience, however, indicated that the volumetric, deviatoric and shear responses on each microplane can be assumed to be mutually independent, as stated in hypothesis V. Thus, the stress-strain relations for monotonic loading are assumed to be decoupled as follows:

$$\sigma_V = C_V(\varepsilon_V)\varepsilon_V \quad \ldots\ldots (4a)$$

$$\sigma_D = C_D(\varepsilon_D)\varepsilon_D \quad \ldots\ldots (4b)$$

$$\sigma_T = C_T(\varepsilon_T, \sigma_C)\varepsilon_T \quad \ldots\ldots (4c)$$

in which C_V depends only on ε_V and C_D only on ε_D [Fig. 2(*c*)]. Among strains, C_T depends only on ε_T, but empirically it is found to also depend on the stress $\sigma_C = (\sigma_{II} + \sigma_{III})/2$ where σ_{II} and σ_{III} are the medium and minimum principal stresses. σ_C is a macroscopic stress invariant, which we will call the confining stress. Test data indicate that σ_C has a significant effect on the shear stiffness. The use of σ_C in Eq. 4 means that the

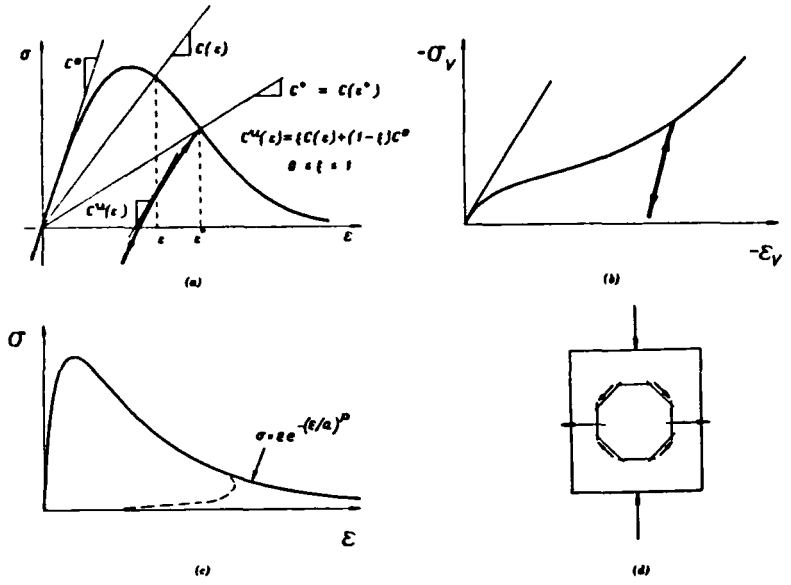


FIG. 2. (*a*) Unloading criterion; (*b*) Volumetric Stress-Strain and Unloading; (*c*) Strain Softening on Microplane and Snapback; and (*d*) Lateral Strains Produced by Uniaxial Compression

constraint of the microplanes is not solely kinematic but also static. The static constraint, however, is not tensorial but merely scalar. σ_C appears only as a parameter, not as a basic variable, and does not affect the elastic part of response.

The stress-strain relation for shear on the microplane, $\sigma_T = C_T\,\varepsilon_T$, neglects any possible effect of the direction of the shear strain vector ε_T on the components ε_{Tl} within the microplane. Whether the direction of this vector remains constant or rotates during loading, the secant modulus depends only on the final state. Although one might suspect this to be an oversimplification, good agreement with test data has been obtained. Note also that we do not need to specify the shear stiffness separately for each component ε_{Tl}, but do so only for the magnitude $\varepsilon_T = \|\boldsymbol{\varepsilon}_T\|$.

To handle unloading, one records in a step-by-step computation the maximum and minimum values of every strain component achieved so far (i.e., $\varepsilon_V^{\max}$, $\varepsilon_V^{\min}$, $\varepsilon_D^{\max}$, $\varepsilon_D^{\min}$, and $\varepsilon_T^{\max}$ ($\varepsilon_T^{\min} = 0$ always). These values are updated after each loading step. Hypothesis IV means that we need to specify for unloading or reloading a curve of a certain fixed shape to be followed by each strain component. After experimenting with various rules, the following simple rule, which is a slight modification of the rule introduced by Bažant and Chern (1985) and involves only one additional parameter, ξ, has been adopted:

$$d\sigma_V = C_V^u\, d\varepsilon_V \qquad \text{where} \quad C_V^u = \xi C_V(\varepsilon_V) + (1-\xi)C_V^0 \quad \ldots\ldots\ldots\ldots \quad (5a)$$

$$d\sigma_D = C_D^u\, d\varepsilon_D \qquad \text{where} \quad C_D^u = \xi C_D(\varepsilon_D) + (1-\xi)C_D^0 \quad \ldots\ldots\ldots\ldots \quad (5b)$$

$$d\boldsymbol{\sigma}_T = C_T^u\, d\boldsymbol{\xi}_T \qquad \text{where} \quad C_T^u = \xi C_T(\boldsymbol{\varepsilon}_T, \boldsymbol{\sigma}) + (1-\xi)C_T^0 \quad \ldots\ldots\ldots\ldots \quad (5c)$$

C_V^u, C_D^u, and C_T^u represent the tangential moduli during unloading or reloading (i.e., for unloaded states). Eq. 5*a* is valid for unloading of the volumetric component at $\varepsilon_V > 0$ (volumetric tension); if $\varepsilon_V < 0$ (volumetric compression), then the unloading modulus is $C_V^u = kC_V^0$ where k is the ratio of the unloading modulus to the initial elastic modulus of the macroscopic hydrostatic compression response [Fig. 2(*b*)]. In the data studies that follow, k is not obtained from curve fitting (and thus it is not a parameter of the model), but from direct measurement of the slope change at unloading. If such test results are unavailable, one may use $k = 1$, which has been found acceptable for all data sets used in this study.

The expressions in Eq. 5 graphically mean that the tangential modulus for unloading or reloading is a weighted average of the initial elastic modulus C_V^0, C_D^0 or C_T^0 and the secant modulus for loading at the same strain [Fig. 2(*a*)]; the case of ε_T is an exception to this rule if σ_C varies. Parameter ξ is assumed to be the same for all the strain components, although it could have different values for each of them.

The loading criterion is in a step-by-step computer program formulated as follows: If $\varepsilon_V\,\Delta\varepsilon_V \geq 0$ and $(\varepsilon_V - \varepsilon_V^{\max})(\varepsilon_V - \varepsilon_V^{\min}) \geq 0$: virgin loading (Eq. 4) is appropriate; otherwise, unloaded states occur (Eq. 5). The same criterion is introduced for ε_D and ε_T. The increment of shear strain $\Delta\varepsilon_T$ is defined as $\Delta\varepsilon_T = \|\boldsymbol{\varepsilon}_T + \Delta\boldsymbol{\varepsilon}_T\| - \|\boldsymbol{\varepsilon}_T\|$, because otherwise $\varepsilon_T\Delta\varepsilon_T \geq 0$ always (e.g., if $\Delta\varepsilon_T = \|\Delta\,\boldsymbol{\varepsilon}_T\|$).

The foregoing simple rule means that if unloading after previous compression continues into tensile states or if unloading after previous tension continues into compressive states, the tensile or compression strength limits for the microplanes, f_t^M and f_c^M, could be greatly exceeded. This could be grossly unrealistic. To prevent it from happening, one needs to impose for unloading of ε_V after previous compression the condition that, approximately, $\sigma_N \leq 0.75\,f_t^M$, and for unloading after previous tension of ε_D and ε_V the condition that $\sigma_N \geq -0.75\,f_c^M$. If this condition is violated, a more sophisticated unloading rule should be specified. If this is not done the computation must be aborted.

Lest it be thought that hypothesis V prevents the modeling of friction, several facts should be noted:

1. Even though the model has no shear friction on the microplane per se, there is a dependence of the normal strain σ_N on the lateral strain ε_L. This represents a frictional phenomenon as explained later. The dependence of C_T on $\boldsymbol{\sigma}$ is of course a frictional phenomenon too.
2. On the macroscopic level, the model exhibits friction due to the fact that a shear stress is manifested at inclinations $\pm 45°$ as compressive and tensile stresses that are not equal in magnitude and have unequal responses [Fig. 2(*d*)]. This was amply demonstrated for the crack band microplane model (Bažant and Gambarova 1984), in which the microplanes were assumed to possess no shear stiffness, yet the model was shown to describe friction on the macrolevel very well (e.g., upon applying a pure shear stress on the crack band, the inclined compressive stresses produce shear and normal components which are not canceled by the effect of the tensile stresses of opposite inclination since these tensile stresses are

smaller or vanish). The inclined normal strains caused by these inclined compressive and tensile stresses also produce dilatancy of the crack band subjected to shear.

3. Despite the neglect of shear stiffness on the microplanes, the crack band microplane model (Bažant and Gambarova 1984) was shown capable to describe all the available test results for friction and dilatancy on cracks in concrete. The reason is that a rough crack in shear behaves in the same manner as a crack band in shear. The contacts of the asperities are inclined and transmit inclined compressive stresses across the crack, which are the source of friction.

4. If any friction interaction were included for the microplane, it would have to model only the remaining part of friction which is in excess of that already exhibited by interactions among the microplanes of various orientations.

5. At a sufficient resolution of the microstructure, friction generally disappears. For example, there is no friction between atoms, only central forces. Although this analogy cannot be taken literally, it does illustrate that friction is a macroscopic phenomenon, caused by interaction of forces and deformations of various orientations in the material. There is no compelling requirement for attributing macroscopically observed friction to a coupling between normal and shear components on a single plane in the material.

Incremental Macroscopic Stress-Strain Relation

For incremental solutions with step-by-step loading, we need to differentiate Eq. 4, which yields (together with Eq. 5) the following incremental stress-strain relations for each microplane:

$$d\sigma_V = C'_V\, d\varepsilon_V - d\sigma''_V \quad \ldots\ldots (6a)$$

$$d\sigma_D = C'_D\, d\varepsilon_D - d\sigma''_D \quad \ldots\ldots (6b)$$

$$d\sigma_{T_r} = C'_T\, d\varepsilon_{T_r} - d\sigma''_{T_r} \quad \ldots\ldots (6c)$$

in which for virgin loading

$$d\sigma''_V = -\varepsilon_V\, dC_V, \qquad C'_V = C_V(\varepsilon_V) \quad \ldots\ldots (7a)$$

$$d\sigma''_D = -\varepsilon_D\, dC_D, \qquad C'_D = C_D(\varepsilon_D) \quad \ldots\ldots (7b)$$

$$d\sigma''_{T_r} = -\varepsilon_{T_r}\, dC_T, \qquad C'_T = C_T(\varepsilon_T, \boldsymbol{\sigma}) \quad \ldots\ldots (7c)$$

and for unloaded states:

$$d\sigma''_V = 0 \qquad C'_V = C''_V \quad \ldots\ldots (8a)$$

$$d\sigma''_D = 0 \qquad C'_D = C''_D \quad \ldots\ldots (8b)$$

$$d\sigma''_{T_r} = 0 \qquad C'_T = C''_T \quad \ldots\ldots (8c)$$

C'_V, C'_D, and C'_T represent the incremental elastic moduli for the current loading step for the microplane, and $d\sigma''_V$, $d\sigma''_D$, and $d\sigma''_{T_r}$ represent the

inelastic stress increments for the microplane, which occur only for virgin loading.

Following the same procedure as used in the original derivation of the microplane model (Bažant and Gambrova 1984; Bažant 1984; Bažant and Oh 1985; Bažant 1985), we use the principle of virtual work to approximately enforce the equivalence of forces on the microscale and macroscale:

$$\frac{4\pi}{3}\, d\sigma_{ij}\delta\varepsilon_{ij} = 2\int_S (d\sigma_N \delta\varepsilon_N + d\sigma_{T_r}\delta\varepsilon_{T_r}) f(\mathbf{n})\, dS \quad \ldots\ldots (9)$$

in which $\delta\varepsilon_{ij}$, $\delta\varepsilon_N$, and $\delta\varepsilon_{Tr}$ are small variations of the macroscopic strain and of the strain components on a microplane. The macroscopic work on the left-hand side is taken over the volume of a unit sphere, and the integral extends over the surface of a unit hemisphere, S. The factor 2 is used because the integrand values on diametrically opposite points of the sphere are equal. Function $f(\mathbf{n})$ is a weighting function of the normal direction $\mathbf{n}$ which in general can introduce anisotropy of the material in its initial state. For concrete, one can approximately assume isotropy, in which case one sets $f(\mathbf{n}) = 1$, but for some concretes (such as rolled dam concrete) consideration of initial anisotropy would be appropriate and would not add much to the complexity of analysis. (It is one advantage of the microplane model that analysis of inelastic anisotropic materials is not appreciably more difficult than isotropic ones.)

Substituting now $\delta\varepsilon_N = n_i n_j \delta\varepsilon_{ij}$, $\delta\varepsilon_{Tr} = \frac{1}{2}(n_i\delta_{rj} + n_j\delta_{ri} - 2n_i n_j n_r)\delta\varepsilon_{ij}$, according to Eqs. 1 and 2, and setting $d\sigma_N = d\sigma_V + d\sigma_D$, we get

$$\frac{2\pi}{3}\, d\sigma_{ij}\delta\varepsilon_{ij} = \int_S \left[n_i n_j (d\sigma_V + d\sigma_D) + \frac{1}{2}(n_i\delta_{rj} + n_j\delta_{ri} - 2n_i n_j n_r)\, d\sigma_{T_r} \right] f(\mathbf{n})\, dS\delta\varepsilon_{ij} \quad \ldots\ldots (10)$$

This variational equation must hold for any variations $\delta\varepsilon_{ij}$; consequently

$$d\sigma_{ij} = \frac{3}{2\pi}\int_S \left[n_i n_j (d\sigma_V + d\sigma_D) + \frac{1}{2}(n_i\delta_{rj} + n_j\delta_{ri} - 2n_i n_j n_r)\, d\sigma_{T_r} \right] f(\mathbf{n})\, dS$$

$$= \frac{3}{2\pi}\int_S \left[n_i n_j (C'_V\, d\varepsilon_V - d\sigma''_V + C'_D\, d\varepsilon_D - d\sigma''_D) + \frac{1}{2}(n_i\delta_{rj} + n_j\delta_{ri} \right.$$

$$\left. - 2n_i n_j n_r)(C'_T\, d\varepsilon_{T_r} - d\sigma''_{T_r}) \right] f(\mathbf{n})\, dS \quad \ldots\ldots (11)$$

in which we have substituted Eq. 6. According to Eqs. 1 and 2, we may further substitute $d\varepsilon_V = \delta_{km} d\varepsilon_{km}/3$, $d\varepsilon_D = n_k n_m d\varepsilon_{km} - d\varepsilon_V$, $d\varepsilon_{Tr} = \frac{1}{2}(n_m\delta_{rk} + n_k\delta_{rm} - 2n_r n_k n_m)d\varepsilon_{km}$. Thus, we finally obtain a macroscopic stress-strain relation of the form

$$d\sigma_{ij} = C_{ijkm}\, d\varepsilon_{km} - d\sigma''_{ij} \quad \ldots\ldots (12)$$

in which C_{ijkm} denotes the incremental stiffness tensor (elastic moduli tensor):

$$C_{ijkm} = \frac{3}{2\pi}\int_S \left[(C_D^t - C_T^t)n_i n_j n_k n_m + \frac{1}{3}(C_V^t - C_D^t)n_i n_j \delta_{km} + \frac{1}{4} C_T^t(n_i n_k \delta_{jm} + n_i n_m \delta_{jk} + n_j n_k \delta_{im} + n_j n_m \delta_{ik})\right] f(\mathbf{n})\, dS \quad \ldots\ldots\ldots\ldots (13)$$

and $d\sigma_{ij}''$ denotes the inelastic stress increments:

$$d\sigma_{ij}'' = \frac{3}{2\pi}\int_S \left[n_i n_j\, d\sigma_N'' + \frac{1}{2}(n_i\delta_{rj} + n_j\delta_{ri} - 2n_i n_j n_r)\, d\sigma_T''\right] f(\mathbf{n})\, dS \quad \ldots\ldots (14)$$

Nonsymmetry of Incremental Elastic Moduli Tensor and Isotropy

Upon carrying an interchange of indices i, j and k, m we notice a surprising fact: the incremental elastic moduli tensor is nonsymmetric unless $C_V^t = C_D^t$, which, however, cannot be expected to occur in general. This nonsymmetry, which represents a major difference from the previous constitutive models, can be understood from the following relations:

$$\sigma_N = \sigma_V + \sigma_D = C_V^t\varepsilon_V + C_D^t\varepsilon_D = C_V^t\varepsilon_V + C_D^t(\varepsilon_N - \varepsilon_V)$$

$$= C_D^t\varepsilon_N + (C_V^t - C_D^t)\varepsilon_V = C_V^t(\varepsilon_N - \varepsilon_D) + C_D^t\varepsilon_D$$

$$= C_V^t\varepsilon_N + (C_D^t - C_V^t)\varepsilon_D = \frac{1}{3}(C_V^t + 2C_D^t)\varepsilon_N + \frac{2}{3}(C_V^t - C_D^t)\varepsilon_L \quad \ldots\ldots\ldots (15)$$

Since $C_V^t = C_D^t$ is required for symmetry, we conclude that symmetry is obtained if σ_N depends only on the normal strain ε_N and does not separately depend on the volumetric strain ε_V and the deviatoric strain ε_D, and thus also the lateral strain ε_L. The lateral strain ε_L has the special property that none of the microplane stresses (neither σ_N nor σ_T) does any work on this strain. The phenomenon that a nonworking variable in this case ε_L, influences the response is in the generalized sense called friction (Bažant 1980). Frictional phenomena are undeniably real, they do occur. So the nonsymmetry of C_{ijkm} cannot be objected on physical grounds, although it violates Drucker's postulate and may cause instabilities and programming inconvenience.

Friction is of course impossible on physical grounds when the response is purely elastic, and in that case C_{ijkm} must be fully symmetric even if $C_V^t \neq C_D^t$. Moreover, C_{ijkm} must exhibit isotropy (if $f(n) = 1$). That these two properties are indeed true we demonstrate now. Using spherical coordinates θ and φ, we may substitute in Eq. 13:

$$n_1 = \cos\theta \quad \ldots\ldots (16a)$$

$$n_2 = \sin\theta\cos\varphi \quad \ldots\ldots (16b)$$

$$n_3 = \sin\theta\sin\varphi \quad \ldots\ldots (16c)$$

$$dS = \sin\theta\, d\theta\, d\varphi \quad \ldots\ldots (16d)$$

We also substitute the initial elastic moduli $C_V^t = C_V^0$, $C_D^t = C_D^0$, and $C_T^t = C_T^0$. Since C_V^0, C_D^0 and C_T^0 are independent of θ and φ, we may integrate

explicitly for all the possible combinations of subscripts i, j, k, m. For isotropy, we have by definition $f(\mathbf{n}) = 1$. In this manner we obtain the following elastic relation:

$$\begin{Bmatrix} d\sigma_{11} \\ d\sigma_{22} \\ d\sigma_{33} \\ d\sigma_{12} \\ d\sigma_{23} \\ d\sigma_{31} \end{Bmatrix} = \begin{bmatrix} A & B & B & 0 & 0 & 0 \\ B & A & B & 0 & 0 & 0 \\ B & B & A & 0 & 0 & 0 \\ 0 & 0 & 0 & C & 0 & 0 \\ 0 & 0 & 0 & 0 & C & 0 \\ 0 & 0 & 0 & 0 & 0 & C \end{bmatrix} \begin{Bmatrix} d\varepsilon_{11} \\ d\varepsilon_{22} \\ d\varepsilon_{33} \\ d\varepsilon_{12} \\ d\varepsilon_{23} \\ d\varepsilon_{31} \end{Bmatrix} \quad \ldots\ldots (17)$$

in which

$$A = \frac{1}{15}(5C_V^0 + 4C_D^0 + 6C_T^0) \quad \ldots\ldots (18a)$$

$$B = \frac{1}{15}(5C_V^0 - 2C_D^0 - 3C_T^0) \quad \ldots\ldots (18b)$$

$$C = \frac{1}{5}(2C_D^0 + 3C_T^0) \quad \ldots\ldots (18c)$$

The matrix in Eq. 17 must coincide with the well-known elastic moduli matrix for isotropic elastic materials, and so we must have

$$A = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} \quad \ldots\ldots (19a)$$

$$B = \frac{E\nu}{(1+\nu)(1-2\nu)} \quad \ldots\ldots (19b)$$

$$C = 2G = \frac{E}{1+\nu} \quad \ldots\ldots (19c)$$

in which E = elastic Young's modulus; G = elastic shear modulus, and ν = Poisson ratio. Equating the expressions for A and B in Eqs. 18 and 19, and solving for ν and E, we obtain the relations

$$\nu = \frac{5C_V^0 - 2C_D^0 - 3C_T^0}{10C_V^0 + 2C_D^0 + 3C_T^0} \quad \ldots\ldots (20a)$$

$$E = (1 - 2\nu)C_V^0 \quad \ldots\ldots (20b)$$

It may be verified that if one equates the expressions for A and C in Eqs. 18 and 19 and solves again for ν and E, one gets the same result. Note also that even if $C_V^0 \neq C_D^0$, the integration over θ and φ yields all the zeros in Eq. 17 that are required for material isotropy and symmetry.

The only thermodynamic restriction we have on C_V^0 and C_D^0 is that they must not be negative. From Eq. 20 one can then obtain any Poisson ratio in the range

$$-1 \leq \nu \leq 0.5 \quad \ldots\ldots (21)$$

that coincides with the well-known thermodynamic restriction on ν.

In fitting of the test data, it is convenient to consider E and ν as the basic parameters and choose the ratio $\eta_0 = C_D^0/C_V^0$. Then one obtains $C_V^0 = E/(1 - 2\nu)$ and, by solving Eq. 20,

$$C_I^0 = \frac{1}{3}\left[\frac{5(1-2\nu)}{1+\nu} - 2\eta_0\right]C_V^0 \quad \ldots\ldots (22a)$$

$$C_D^0 = \eta_0 C_V^0 \quad \ldots\ldots (22b)$$

Choosing various values of η_0, it was found that good fits are obtained approximately for $0.25 \leq \eta_0 \leq 1$.

For the extreme case, $\eta_0 = 1$ (i.e., $C_V^0 = C_D^0$). This does not imply that $C_V^t = C_D^t$ for further loading, but does imply that C_V^t is rather close to C_D^t for the initial inelastic response in the hardening range. Consequently, tensor C_{ijkm} is nearly symmetric for this initial range. Beyond the initial range, however, C_V^t and C_D^t may become significantly different (and tensor C_{ijkm} highly nonsymmetric) for two reasons: (1) C_V^t and C_D^t depend on ε_V and ε_D, respectively, whose values can be rather different; and (2) they change greatly in the case of unloading (either ε_V or ε_D may be loading and the other unloading on the same microplane). Note also that unloading occurs on some microplanes even during monotonic loading (e.g., for the uniaxial compression test).

It has been tried to determine whether replacement of tensor C_{ijkm} with its symmetric part could yield good fits of test data. It cannot (especially not the observed volume changes and lateral strains).

Numerical Implementation

To make computations efficient, the tensorial expressions $n_i n_j n_k n_m$, $\frac{1}{4}C_I^t(n_i n_k \delta_{jm} + n_i n_m \delta_{jk} + n_j n_k \delta_{im} + n_j n_m \delta_{ik})$, $n_i n_j \delta_{km}$, $n_i n_j$, and $\frac{1}{2}(n_i \delta_{rj} + n_j \delta_{ri} - 2n_i n_j n_r)$ should be calculated in advance of finite element analysis and stored in arrays in the computer's memory. The integrals over the unit hemisphere in Eqs. 13 and 14 need to be evaluated in finite element analysis in every loading step, every iteration of the step, and every integration point of every finite element. Therefore, evaluation of these integrals must be as efficient as possible. This evaluation is done by numerical integration using some suitable Gaussian-type integration formula for the surface of the hemisphere. Many such formulas are listed in Stroud's book (Stroud 1971), and some further formulas which are more efficient under certain circumstances have been derived by Bažant and Oh (1985, 1986). These formulas generally approximate the integrals in the form

$$\int_S F\, dS \approx 4\pi \sum_{\alpha=1}^{N} w_\alpha F_\alpha \quad \ldots\ldots (23)$$

in which subscript α refers to a certain discrete set of directions in space, and w_α are the weights (numerical integration coefficients) for these directions (i.e., for points on the unit hemisphere). In the present calculations, Stroud's formula with 28 points per hemisphere (Bažant and Gambarova 1984; Bažant and Oh 1985; Stroud 1971) has been used. For symmetric stress states the number of integration points can be reduced

because the values of $w_\alpha F_\alpha$ are equal for some points (e.g., for axisymmetric stress states, the summation in Eq. 23 involves only five terms). Stroud's 28-point formula is of the eleventh degree (i.e., it integrates exactly on the surface of a sphere all polynomials up to the eleventh degree). A lesser but practically still sufficient accuracy is obtained by Bažant and Oh's 21-point formula (Bažant and Gambarova 1984; Bažant and Oh 1985), and for relatively crude analysis one could use Collatz's formula with only 16 points. A smaller number of points gives unacceptable results, especially in the strain-softening range (Bažant and Gambarova; Bažant and Oh 1985).

Conclusions

1. The microplane model that was previously formulated for tensile cracking of materials such as concrete is extended to general nonlinear triaxial behavior including compression and shear loadings. This model represents a counterpart of the slip theory of plasticity in which the structure is constrained kinematically rather than statically and the normal inelastic strains are taken into account. As confirmed in Part II, which follows, the model is capable of a good description of a broad range of the existing test data for nonlinear triaxial behavior of concrete. It seems to describe realistically the dilatancy and friction; the brittle-ductile transition at increasing hydrostatic pressure; various degrees of path dependence; the stiffening response in hydrostatic pressure tests; the extended strain softening in tension as well as in compression; and other typical features.

2. The model achieves conceptual simplicity by specifying the constitutive properties independently on planes of various orientations in the material, which are called the microplanes and characterize principally the behavior of the weak planes within the microstructural framework. On each microplane there are only few stress and strain components and there are no tensorial invariance requirements to satisfy. The tensorial invariance requirements are satisfied automatically on the macroscopic level by integrating over all the spatial directions.

3. For monotonic loading, the stress-strain relation for each microplane is assumed to be path independent and all the path dependence that is observed macroscopically is a consequence of unloading which happens selectively on some microplanes.

4. On each microplane one distinguishes the volumetric and deviatoric normal strains and the shear strains, with different elastic constants for each. This makes it possible to match with the microplane system any Poisson ratio. The previous microplane models, by contrast, had been characterized by only a certain value of the Poisson ratio, and matching of the experimentally observed Poisson ratio required the artifice of coupling in series or in parallel with the microplane system an additional fictitious elastic element.

5. The response for each microplane is assumed to depend on the volumetric strain of the material, which implies dependence on the normal strain in the lateral direction. Since this strain does no work, a frictional aspect is introduced into the response, causing the incremental elastic moduli tensor in the nonlinear range to be nonsymmetric. This tensor is fully populated, (i.e., none of its components can in general be prescribed

as zero). The nonsymmetry of this tensor appears to be important for the modeling of friction, dilatancy, and pressure sensitivity.

6. To reflect internal friction, the inelastic part of shear stiffness on the microplane is considered to depend also on the macroscopic confining stress (defined as the average of the minimum and medium principal stresses). In consequence, the micro-macro constraint is not purely kinematic but mixed (kinematic-static). This causes the response to be less stiff than for a purely kinematic constraint. While the basic kinematic constraint is tensorial, the additional static constraint is scalar.

7. The model permits extensions to anisotropic materials without any additional complexity.

Acknowledgments

Grateful appreciation of financial support in the final stage of development of this model is due to the U.S. Air Force Office of Scientific Research, which provided funds under Contract No. F 49620-87-C-0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu.

Appendix. References

Batdorf, S. B., and Budianski, B. (1949). "A mathematical theory of plasticity based on the concept of slip." *Technical Note No. 1871*, Nat. Advisory Committee for Aeronautics, Washington, D.C.

Bažant, Z. P. (1974). "A new approach to inelasticity and failure of concrete, sand and rock: Endochronic theory." *Proc. Society of Engineering Science 11th Annual Meeting*, Duke Univ., Durham, N.C., 158–159.

Bažant, Z. P. (1980). "Work inequalities for plastic fracturing materials." *Int. J. Solids Struct.*, 16, 873–901.

Bažant, Z. P. (1984). "Microplane model for strain-controlled inelastic behavior." *Mechanics of engineering materials*, C. S. Desai and R. H. Gallagher, eds., John Wiley and Sons, Inc., New York, N.Y., 45–59.

Bažant, Z. P. (1985). *Mechanics of geomaterials: Rock, concrete, soils*. John Wiley and Sons, Inc., New York, N.Y.

Bažant, Z. P., and Bhat, P. D. (1976). "Endochronic theory for concrete." *J. Engrg. Mech.*, ASCE, 105(3), 407–428.

Bažant, Z. P., and Chern, J.-C. (1985). "Strain-softening with creep and exponential algorithm." *J. Engrg. Mech.*, ASCE, 111(3), 381–390.

Bažant, Z. P., and Gambarova, P. G. (1984). "Crack shear in concrete: Crack band microplane model." *J. Struct. Engrg.*, ASCE, 110(9), 2015–2035.

Bažant, Z. P., and Kim, S. (1979). "Plastic-fracturing theory for concrete." *J. Engrg. Mech.*, ASCE, 105(3), 407–428.

Bažant, Z. P., and Kim, J.-K. (1986). "Creep of anisotropic clay: Microplane model." *J. Geotech. Engrg.*, ASCE, 112(4), 458–475.

Bažant, Z. P., and Oh, B. H. (1983). "Crack band theory for fracture of concrete." *Matériaux et Constructions*, RILEM, Paris, 16(93), 155–177.

Bažant, Z. P., and Oh, B. H. (1985). "Microplane model for progressive fracture of concrete and rock." *J. Engrg. Mech.*, ASCE, 111(4), 559–582.

Bažant, Z. P., and Oh, B. H. (1986). "Efficient numerical integration on the surface of a sphere." *Zeitschrift für Angewandte Mathematik und Mechanik* 66(1), 37–49.

Bažant, Z. P., and Prat, P. C. (1987). "Creep of anisotropic clay: New microplane model." *J. Engrg. Mech.*, ASCE, 113(7), 1050–1064.

Bažant, Z. P., and Shieh, C.-L. (1980). "Hysteretic fracturing endochronic theory for concrete." *J. Engrg. Mech.*, ASCE, 106, 929–950.

Bažant, Z. P., and Tsubaki, T. (1980). "Total strain theory and path dependence of concrete." *J. Engrg. Mech.*, ASCE, 106(6), 1151–1173.

Cedolin, L., Crutzen, Y. R. J., and dei Poli, S. (1977). "Triaxial stress-strain relationship for concrete." *J. Engrg. Mech.*, ASCE, 103(4), 423–439.

Cedolin, L., dei Poli, S., and Iori, L. (1983). "Experimental determination of the stress-strain curve and fracture of concrete in tension." *Proc. Int. Conference on Constitutive Laws for Engrg. Materials*, Univ. of Arizona.

Chen, C. T., and Chen, W. F. (1975). "Concrete in biaxial cyclic compression." *J. Struct. Engrg.*, ASCE, 101(4), 461–476.

Dougill, J. W. (1976). "On stable progressively fracturing solids." *J. Appl. Math. Phys.* (ZAMP) 27, 423–436.

Gerstle, K. H. (1981). "Simple formulation of biaxial concrete behavior." *J. Am. Concr. Inst.*, 78(1), 62–68.

Gerstle, K. H., et al. (1980). "Behavior of concrete under multiaxial stress states." *J. Engrg. Mech.*, ASCE, 106(6), 1383–1403.

Gerstle, K. H. (1981). "Simple formulation of triaxial concrete behavior." *J. Am. Concr. Inst.* 78(9), 382–387.

Janson, J., and Hult, J. (1977). "Fracture mechanics and damage mechanics, a combined approach." *J. Mec. Theor. Appl.* 1(1), 69–84.

Kachanov, L. M. (1958). "Time of rupture process under creep conditions." Izv. Akad. Nauk SSR, Otd. Tkh. Nauk 8, 26–31.

Kotsovos, M. D., and Newman, J. B. (1978). "Generalized stress-strain relations for concrete." *J. Engrg. Mech.*, ASCE, 104(4), 845–856.

Krajcinovic, D., and Fonseka, G. U. (1981). "The continuous damage theory of brittle materials, part 1: General theory." *J. Appl. Mech. Trans.*, 48, 809–815.

Lin, F. B. et. al. (1987). "Concrete model with normality and sequential identification." *Comput. Struct.* 26(6), 1011–1026.

Mazars, J. (1981). "Mechanical damage and fracture of concrete structures." *Advances in Fracture Research, Proc. 5th International Conference on Fracture*, 4, 1499–1506.

Ortiz, M. A. (1985). "A constitutive theory for the inelastic behavior of concrete." *Mech. Mater.* 4, 67–93.

Pande, G. H., and Xiong, W. (1982). "An improved multi-laminate model of jointed rock masses." *Proc. 1st Int. Symposium on Numerical Models in Geomechanics* 218–226.

Stroud, A. H. (1971). Approximate calculation of multiple integrals. Prentice-Hall, Inc., Englewood Cliffs, N.J.

Sture, S., and Ko, H. Y. (1978). "Strain-softening of brittle geologic materials." *Int. J. Numer. Anal. Methods Geomech.*, 2, 237–253.

Takahashi, Y., and Marchertas, A. H. (1985). "A simple elasto-plastic constitutive model of concrete." *Proc. 4th Int. Conference on Struct. Safety and Reliability*,

Taylor, G. I. (1938). "Plastic strain in metals." *J. Inst. Metals* 62, 307–324.

Willam, K. J., and Warnke, E. P. (1974). "Constitutive model for the triaxial behavior of concrete." *IABSE Seminar on Concrete Structures Subjected to Triaxial Stresses*, ISMES, Bergamo, Italy, 1–30.

Yang, B.-L., Dafalias, Y. F., and Herrmann, L. R. (1985). "A bounding surface plasticity model for concrete." *J. Engrg. Mech.*, ASCE, 111(3), 359–380.

Zienkiewicz, O. C., and Pande, G. N. (1977). "Time-dependent multilaminate model of rocks—A numerical study of deformation and failure of rock masses." *Int. J. Numer. Anal. Methods Geomech.*, 1, 219–247.

Zubelewicz, A., and Bažant, Z. P. (1987). "Constitutive model with rotating active plane and true stress." *J. Eng. Mech.*, ASCE, 113(3), 398–416.

Zubelewicz, A., and Bažant, Z. P. (1987). "Interface element modeling of fracture in aggregate composites." *J. Engrg. Mech.*, ASCE, 113(11), 1619–1630.

MICROPLANE MODEL FOR BRITTLE-PLASTIC MATERIAL: II. VERIFICATION

By Zdeněk P. Bažant,[1] Fellow, ASCE, and Pere C. Prat,[2] Student Member, ASCE

ABSTRACT: The general microplane model formulated in Part I of this study is calibrated and verified by comparison with numerous nonlinear triaxial test data from the literature and good agreement is achieved. The model involves many fewer material parameters than the previous phenomenologic macroscopic models. Numerical implementation is also considered.

INTRODUCTION

In Part I, which immediately precedes this paper, a general microplane model for brittle-plastic materials such as concrete has been developed. The objective of Part II is to calibrate and verify this model by comparisons with the basic experimental data that exist in the literature and to discuss numerical implementation. All the definitions and notations from Part I of this study are retained.

The data from various basic material tests have been fitted directly by solving Eq. 13 of Part I, which represents a system of six algebraic equations relating the stress and strain increments. These equations have been rearranged so that the controlled strain or stress components appear on the right-hand side, and the measured ones appear on the left-hand side. Various standard rules required for stability of computations (Bažant and Shieh 1978) have been observed (e.g., in the strain-softening range one must prescribe the increments of the major axial strain rather than the axial stress; in proportional triaxial tests one must implement a prescribed ratio of the strain increments in various directions, etc.).

EMPIRICAL MATERIAL FUNCTIONS

For the functions defining the microplane secant moduli C_V, C_D, and C_T, the following expressions have been obtained intuitively and verified experimentally:

$$\text{for } \varepsilon_V \geq 0: \quad C_V = C_V^0 e^{-|\varepsilon_V/a_1|^{p_1}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

$$\text{for } \varepsilon_V < 0: \quad C_V = C_V^0 \left[\left(1 + \left| \frac{\varepsilon_V}{a} \right| \right)^{-p} + \left| \frac{\varepsilon_V}{b} \right|^q \right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

[1]Prof. of Civ. Engrg., Dept. of Civ. Engrg., Northwestern Univ., Evanston, IL 60208.

[2]Res. Assoc., Materials Sci. Inst. (CSIC), Barcelona, Spain; formerly, Grad. Res. Asst., Northwestern Univ.

Note. Discussion open until March 1, 1989. Separate discussions should be submitted for the individual papers in this symposium. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on August 7, 1987. This paper is part of the *Journal of Engineering Mechanics*, Vol. 114, No. 10, October, 1988. ©ASCE, ISSN 0733-9399/88/0010-1689/$1.00 + $.15 per page. Paper No. 22824.

for $\varepsilon_D \geq 0$: $C_D = D_D^0 e^{-|\varepsilon_D/a_1|^{p_1}}$ ..... (3)

for $\varepsilon_D < 0$:
$$\begin{cases} C_D = C_D^0 e^{-|\varepsilon_D/a_2|^{p_2}} & \text{..... (4)} \\ C_T = C_T^0 e^{-|\varepsilon_T/a_3|^{p_3}} & \text{..... (5)} \end{cases}$$

in which

$a_5 = a_3$ if $\sigma_C \geq 0$ ..... (6*a*)

$a_5 = a_3 - \frac{a_4}{f_c'}\sigma_C$ if $\sigma < 0$ ..... (6*b*)

in which a, b, p, q, a_1, a_2, a_3, a_4, p_1, p_2, and p_3 are positive empirical material constants. The dependence of a_5 on the confining stress σ_C reflects internal friction and represents an additional static constraint of nontensorial (scalar) type.

The expression for C_V in compression (Eq. 2) can be obtained independently of all others, by simple curve fitting of the test data for hydrostatic compression of the material. As is well known, the hydrostatic compression response exhibits neither maximum stress nor strain softening. Rather, it follows a curve of the shape sketched in Fig. 2(*b*) of Part I. Under hydrostatic loading ($\sigma_{11} = \sigma_{22} = \sigma_{33} = \sigma_V$, $\sigma_{12} = \sigma_{23} = \sigma_{31} = 0$), the stress-strain relation for the microplane system reduces to the simple equation $\sigma_V = C_V \varepsilon_V$ where σ_V now represents the volumetric (or hydrostatic, or mean) stress, $\sigma_V = \sigma_{kk}/3$.

Parameters a, b, p, and q may be identified independently of the others. This can be done by the fitting of one curve to the hydrostatic test data. The fitting can be easily accomplished by a trial-and-error approach. Alternatively, a good fit can be achieved by a nonlinear optimization algorithm such as the Levenberg-Marquardt is, which yields the values of the parameters a, b, p, and q. As still another alternative, one can make trial-and-error choices of only parameters a and p, and then identify parameters b and q by linear regression, plotting log Y vs. log ε_V where $Y = (C_V/C_V^0) - (1 + |\varepsilon_V/a|)^{-p}$; this plot is ideally linear (i.e., $\log Y = q \log \varepsilon_V - q \log b$, with slope q and vertical intercept $q \log b$).

The expressions for C_V, C_D, and C_T in Eqs. 1 and 3–5 are perhaps the simplest expressions that yield, for each stress-strain diagram for the microplane, a curve of the shape shown in Fig. 2(*c*) of Part I. According to these formulas, the strain-softening tail approaches zero stress asymptotically, as the strains approach infinity. Such a shape of the strain-softening curve has been assumed by many investigators, although test data for strain softening at very large strains, with stress less than about 15% of the peak stress, do not seem to exist.

In this regard it may be noted that in recent micromechanics analysis of the crack ligament tearing (Bažant 1987) it appeared that strain softening should terminate at a certain maximum strain after which the stress-strain curve exhibits snapback behavior (i.e., returns to the origin with positive slope) [see the dashed curve in Fig. 2(*c*) of Part I]. A modification of Eqs. 1 and 3–5 in this sense might have to be introduced in the future, but in the present fitting of test data the foregoing simple expressions were used for the sake of simplicity since a curve with a maximum strain would cause

numerical complications. This is of little practical consequence for the present analysis because the existing test data do not go into large enough strains for which the snapback after maximum strain [the dashed curve in Fig. 2(*c*) of Part I] would be reached.

Analysis of numerous test data reveals that the optimum values of exponents p_1, p_2, and p_3 are always nearly the same and can be fixed as follows:

$p_1 = 0.5$ ..... (7*a*)

$p_2 = p_3 = 1.5$ ..... (7*b*)

Parameters a, b, p, and q are also nearly the same for all concretes and can anyway be identified separately from the single hydrostatic test curve. The elastic properties can also be determined in advance separately. This leaves only five material parameters, a_1, a_2, a_3, a_4 and η_0, to be identified by simultaneous fitting of the test data for uniaxial and multiaxial loadings.

The values of the two elastic constants, Poisson's ratio ν and Young's modulus E, are fixed prior to the data fitting. The Poisson ratio is assumed to be $\nu = 0.18$ for all the cases, and the Young's modulus is chosen such that the initial slope of the curves obtained with the model fit the experimental data.

Verification with Experimental Results

Fits of various typical test data from the literature (Balmer 1949; Bažant et al. 1986; Bresler and Pister 1958; Goode and Helmy 1967; Green and Swanson 1973; Hognestad et al. 1955; Kotsovos and Newman 1978; Kupfer et al. 1969; Launay and Gachon 1971; Lin et al. 1987; Petersson 1981; Sinha et al. 1964; van Mier 1984) are exhibited in Figs. 1–10. The measurements are shown as the data points while the results of the present model are plotted as the solid curves. The failure envelopes in Figs. 8 and 9 have been constructed by simulating various standard triaxial and biaxial tests, respectively, and recording the maximum stress value for each test. Collecting these maximum stress values then provides the failure envelope. The basic information on most of these tests is summarized in Bažant and Shieh (1978) and Lin et al. (1987).

The optimum values of the material parameters corresponding to each of these fits are listed in Tables 1 and 2. The optimization by which the fits have been achieved was carried out simply by a trial-and-error approach. If all the data fits were optimized collectively, the scatter of the parameter values from one test series to another would probably be much less than that in Tables 1 and 2. More sophisticated computer optimization would be needed for that purpose.

The fits achieved with the present model are closer than those achieved previously with various phenomenologic macroscopic models. At the same time, the number of the free material parameters to be identified from the test data is greatly reduced. While the previous models that could reasonably approximate the present data required (in addition to two elastic constants) at least 15 material parameters to be identified from the measured data, the present model involves only five such material parameters, aside from further four parameters which are identified in advance

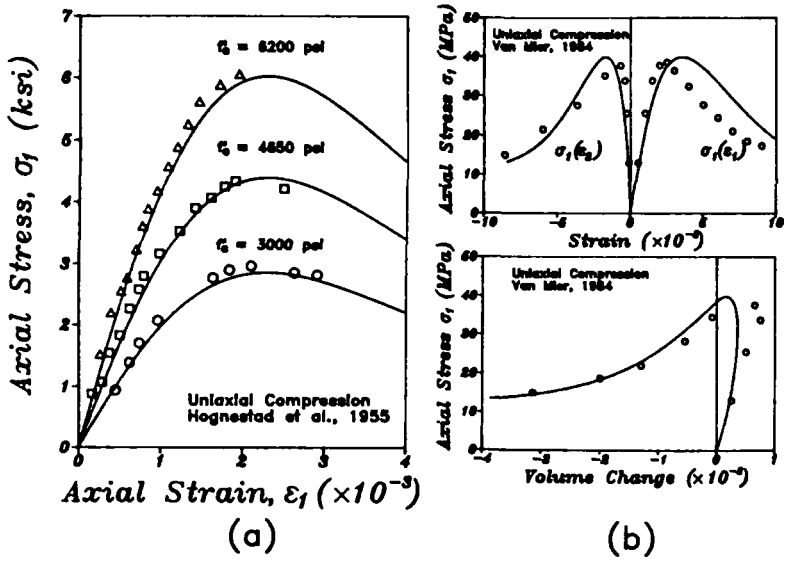


FIG. 1. Comparison with Uniaxial Compression Tests, Reported by: (*a*) Hognestad et al. (1955); (*b*) Van Mier (1984)

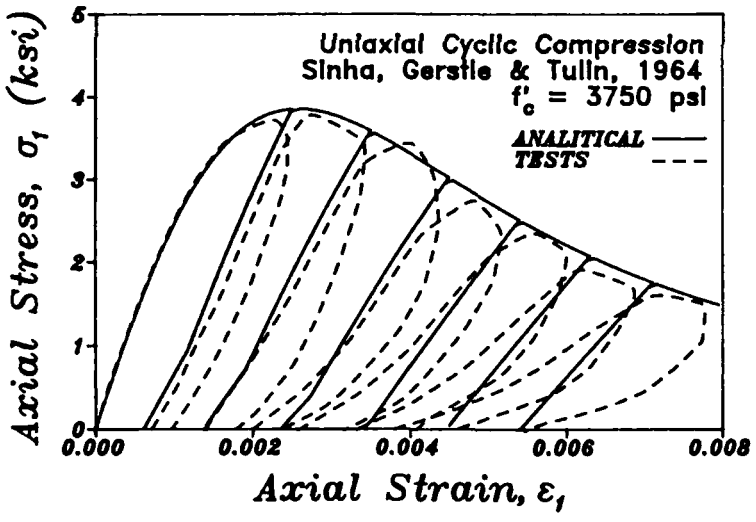


FIG. 2. Comparison with Uniaxial Cyclic Compression Tests, Reported by Sinha et al. (1964)

by simple curve fitting of the hydrostatic test and can be assumed in advance if the hydrostatic test reaching very high pressures is unavailable.

So we achieve a significant simplification of the numerical work needed to identify the material parameters. At the same time we achieve signifi-

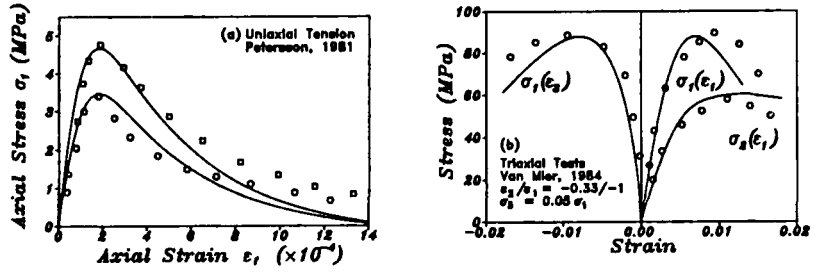


FIG. 3. Comparison with: (*a*) Uniaxial Tension Tests, Reported by Petersson (1981); (*b*) True Triaxial Tests, Reported by Van Mier (1984)

cant conceptual simplification. This is of course at the expense of extensive numerical work that has to be done by the computer when the present model is used in finite element analysis. However, with the recent advances in computing speed and power, this is no longer a serious obstacle [e.g., if the spherical integration formula with 21 microplanes is used, then one needs to store in the computer's memory 132 numbers for each integration point of each finite element (i.e., three stresses and three strain maxima per microplane, and six macrostrains); for the formula with 16 microplanes, this becomes 102 instead of 132].

From the fits shown in Figs. 1–10 it is obvious that the model can describe the major characteristics of the response of a brittle-plastic material such as concrete. This includes the progressive strain softening

TABLE 1. Optimum Values of Material Parameters: Basic Parameters

Test data (1)	η_0 (2)	a_1 ($\times 10^{-4}$) (3)	a_2 ($\times 10^{-3}$) (4)	a_3 (5)	a_4 (6)
Balmer	0.85	5.0	2.5	0.0018	0.0100
Bažant et al.	1.00	5.0	1.0	0.0010	0
Green et al.	1.00	4.0	2.0	0.0017	0
Hognestad et al.	0.50	5.0	2.5	0.0015	0
Kotsovos & Newman	0.85	5.0	5.0	0.0100	0.0160
Kupfer et al.	0.60	1.0	5.0	0.0011	0.0009
Petersson	0.25	0.3	0.5	0.0005	0
Sinha et al.	0.50	5.0	3.2	0.0016	0
van Mier	0.85	4.0	4.3	0.0018	0.0020

TABLE 2. Optimum Values of Material Parameters: Hydrostatic Curve Parameters

Test data (1)	a (2)	b (3)	p (4)	q (5)
Bažant et al.[a]	0.0050	0.0435	0.50	2.00
Bažant et al.[b]	0.0002	0.0850	0.36	3.00
Green & Swanson[c]	0.0008	0.1000	0.50	1.80
All Other Test Data	0.0050	0.2250	0.25	2.25

[a]Very high confining stress, small aggregate concrete.
[b]Very high confining stress, cement paste.
[c]High-strength concrete.

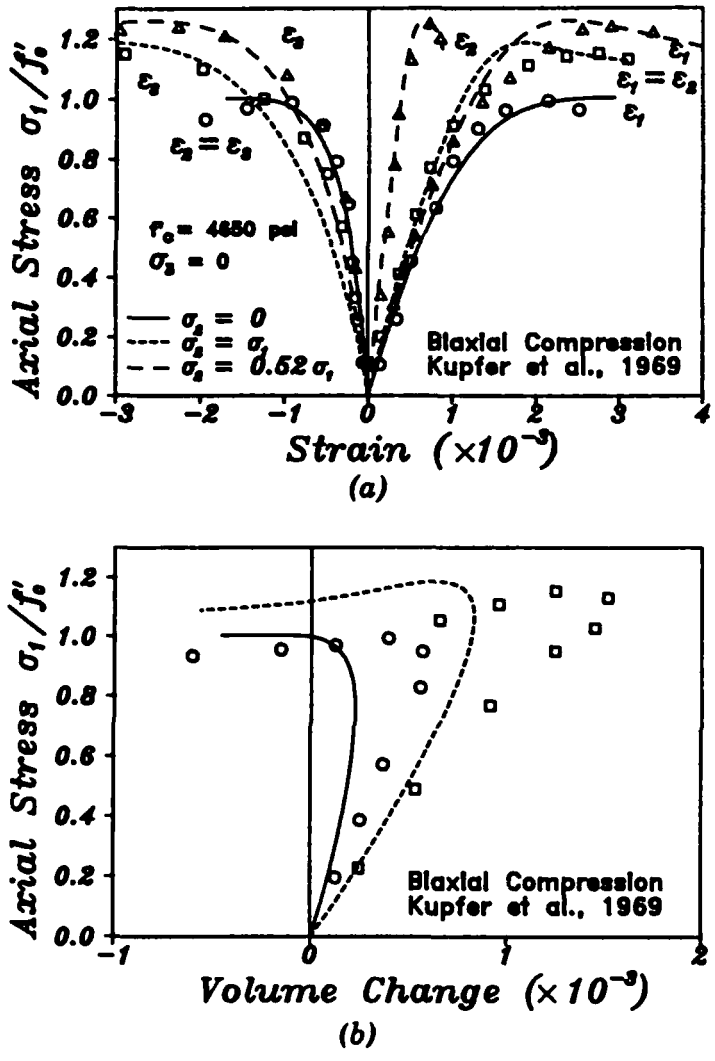


FIG. 4. Comparison with Biaxial Tests, Reported by Kupfer et al. (1969)

with a long tail, large-volume expansion due to strain softening in uniaxial as well as triaxial tests; transition from brittle response at small hydrostatic pressure to essentially plastic hardening response at high hydrostatic pressures; the absence of stress maximum in the hydrostatic pressure test; the difference in the shape of the compressive and tensile meridians in the Rendulić (volumetric) section of the failure envelope (Fig. 8); the rounded triangular shape in the deviatoric cross sections of the failure surface (Fig. 8); including a transition to a more circular shape as the hydrostatic shape is increased; the typical shape of the biaxial failure envelope (Fig. 9) and the compression-shear envelope (Fig. 10), etc.

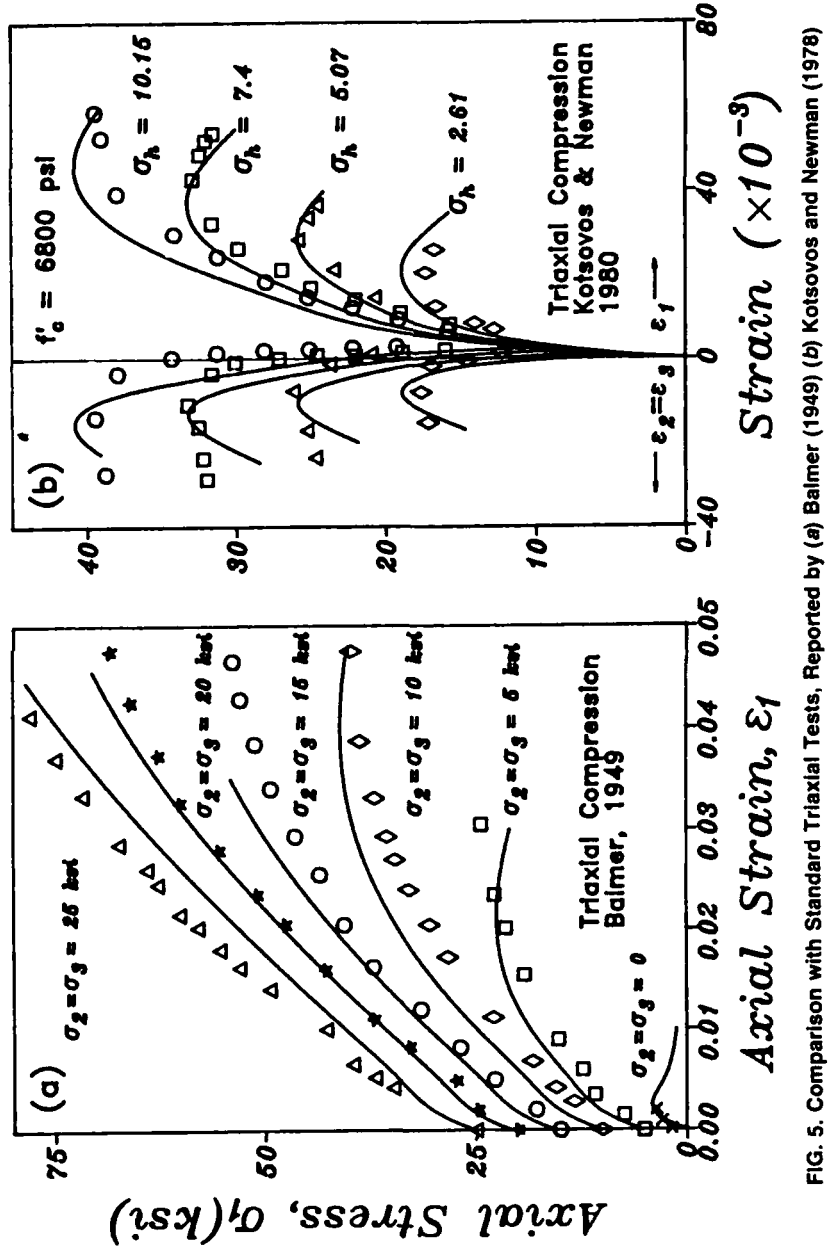


FIG. 5. Comparison with Standard Triaxial Tests, Reported by (a) Balmer (1949) (b) Kotsovos and Newman (1978)

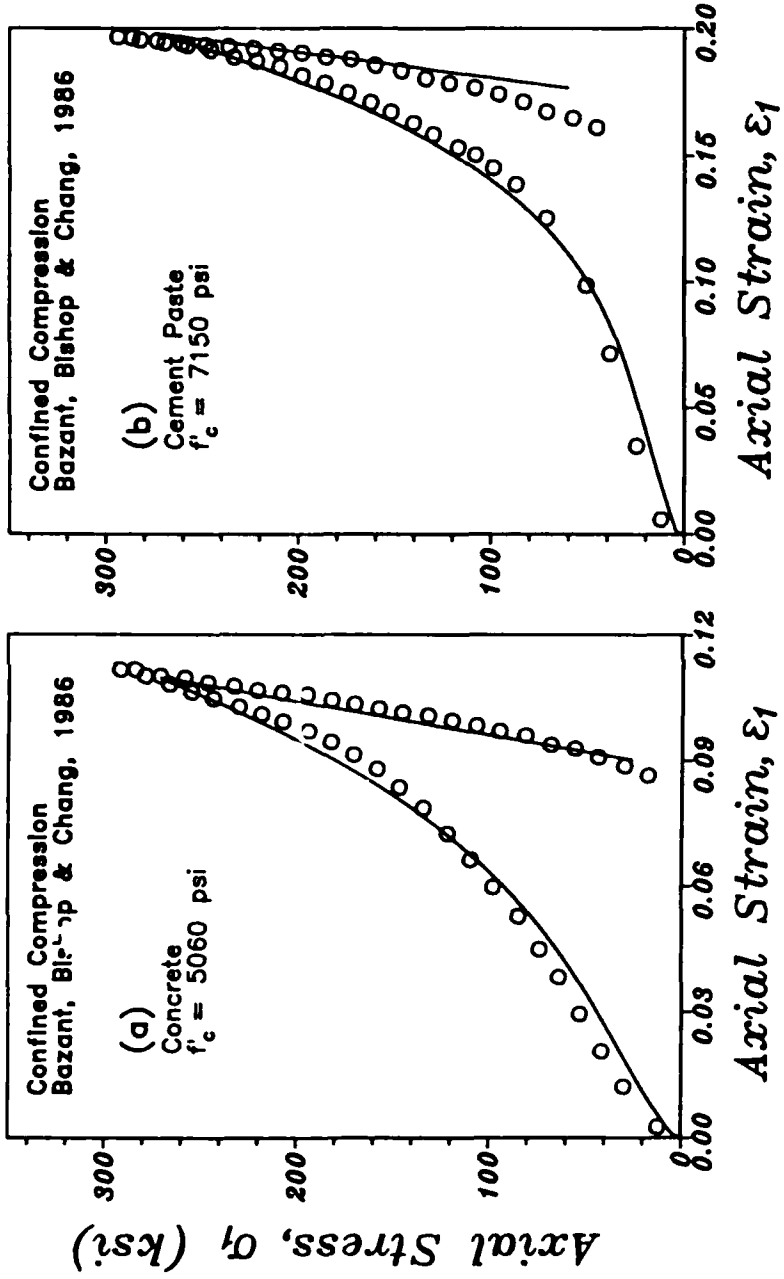


FIG. 6. Comparison with Confined Compression Tests, Reported by Bazant et al.

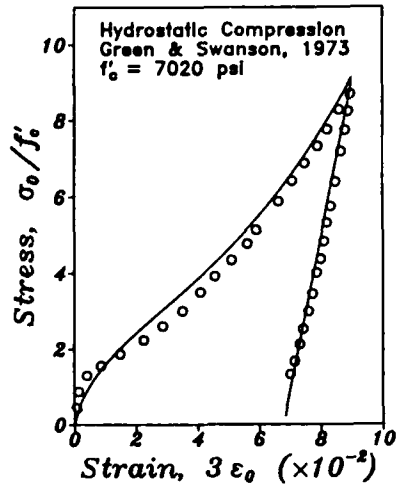


FIG. 7. Comparison with Hydrostatic Compression Tests, Reported by Green and Swanson (1973)

It is interesting to express the initial unloading slope C_u in Fig. 2 as a linear combination of the secant slope C_s and the initial elastic slope C_e [i.e., write $C_u = \xi' C_s + (1 - \xi')C_e$, and calculate $\xi' = (C_u - C_e)/(C_s - C_e)$]. Parameter ξ' is analogous to the material pressure ξ introduced for each microplane (Eq. 5 of Part I). From Fig. 2 we find that $\xi' = 0.57, 0.63, 0.65, 0.66, 0.66$, and 0.67 for the six unloading paths shown (the lowest value of ξ' corresponds to the first unloading). It is surprising that these ξ' values are smaller than the value $\xi = 0.75$ assumed for each microplane (i.e., the macroscopic initial unloading slopes are much closer to the initial elastic slope and much more remote from the secant slope than those for each microplane). The reason is that at the start of unloading only some microplanes start to unload, while some others still continue to load and strain-soften, thus contributing to make the overall macroscopic unloading slope relatively steeper.

Also of interest is to simulate with the present model the responses to various types of loading that are difficult to carry out in the laboratory and on which little or no information is available. Of particular interest is the path dependence of response, especially since assumptions about path independence on the microplane level are central to the present formulation. Fig. 11 (top) shows the responses to the loading paths defined in the figure, in which the final combined state of compressive stress and shear stress is obtained either proportionally or by various nonproportional paths such as uniaxial loading followed by shear. It is seen that the path dependence of the final state reached in these simulations is rather mild.

In the foregoing simulations, the principal stress directions rotate. Fig. 11 (bottom) shows further simulations of responses to loadings in which principal directions do not rotate and are similar to the loading in the cubical true triaxial devices. Shown in these figures are the responses for two principal compressive stresses applied proportionally or in sequence.

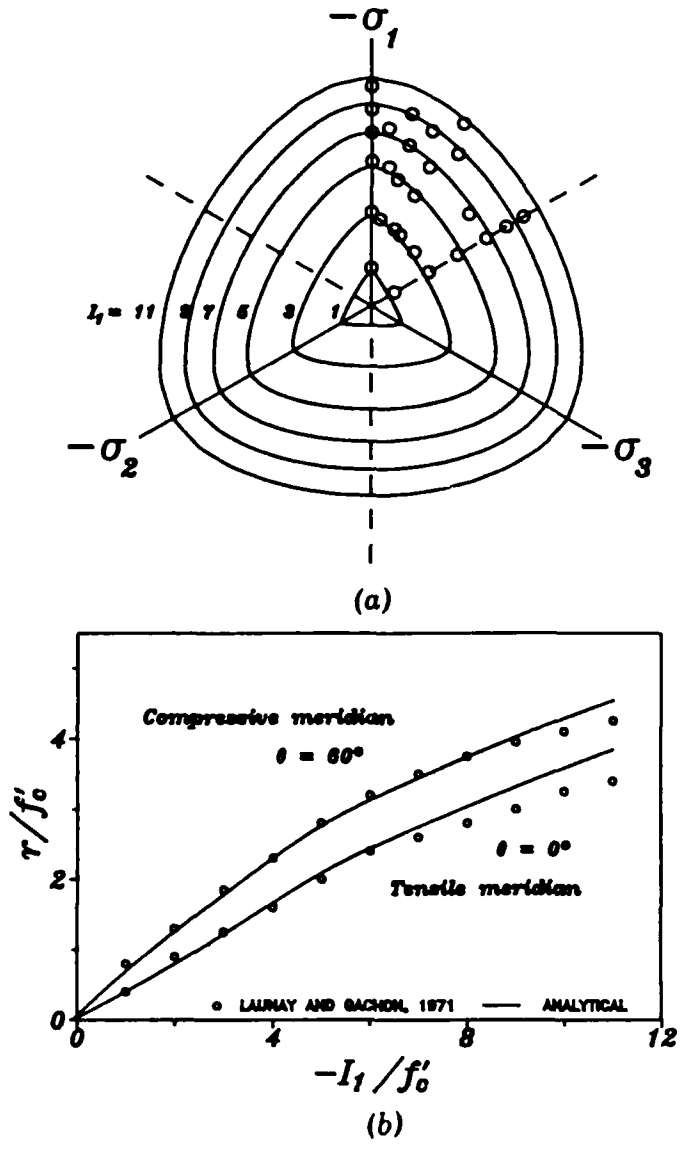


FIG. 8. Hydrostatic and Deviatoric Sections of the Failure Surface, Compared to Data Reported by Launay and Gachon (1971)

Interestingly, even though this loading does not reach the softening range, the path dependence appears to be more significant.

One advantage of the present model is its capability to describe compression failure, including the extended strain-softening range, without introducing any maximum stress and strain softening or crushing for

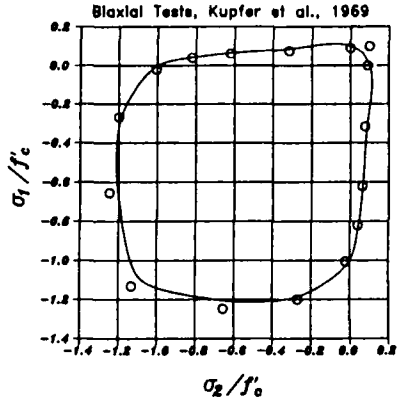


FIG. 9. Biaxial Failure Surface [Data from Kupfer et al. (1969)]

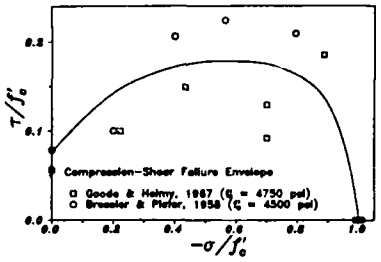


FIG. 10. Compression-Shear Failure Surface [Data from Bressler and Pister (1958) and Goode and Helmy (1967)]

compression on the microplane level. By simulating the uniaxial compression test and recording what happens on various microplanes during the loading process, one finds that the compression failure is obtained as a result of the following sequence: (1) The shear on the inclined planes on a uniaxial test causes slip (inelastic shear strain ε_T) on these planes; (2) the slip on the inclined planes results in a transverse extensional strain, manifested on the microplanes that are parallel to the uniaxial compression axis [Fig. 2(*d*) of Part I]; (3) the transverse extension causes softening in

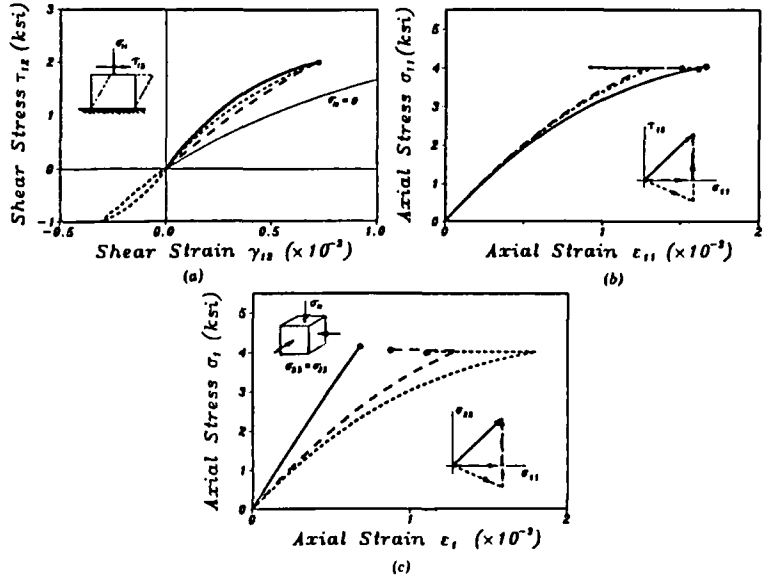


FIG. 11. Top: Path Dependence with Normal and Shear Stresses; Bottom: with only Normal Stresses (True Triaxial Tests)

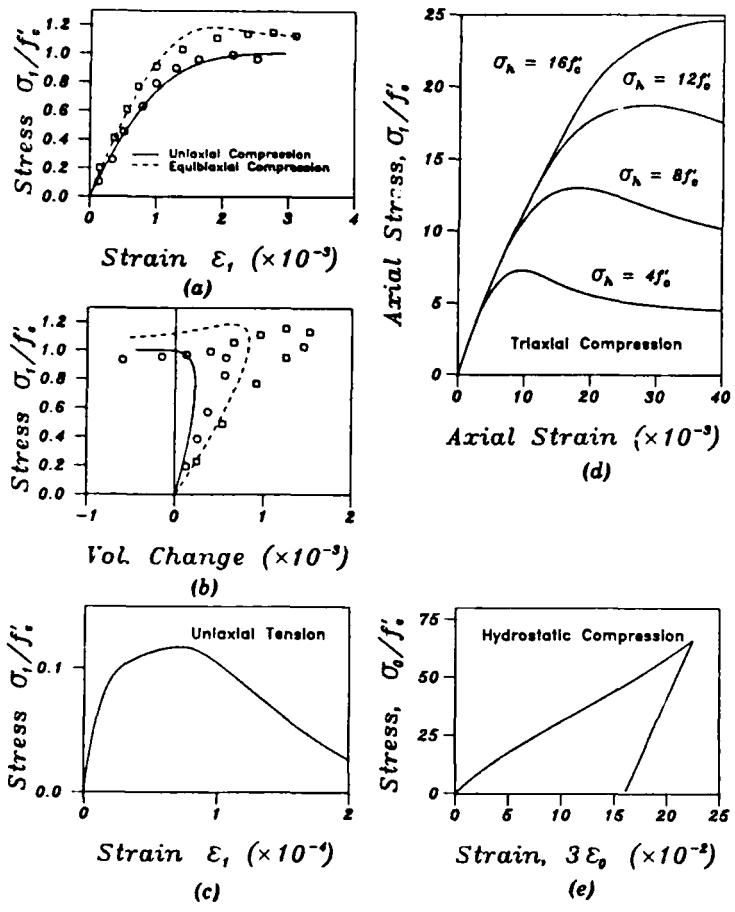


FIG. 12. Prediction of Concrete Response under Different Loading States, using the Parameters Corresponding to the Test Data Reported by Kupfer et al. (1969)

the transverse direction, manifested as splitting microcracks and cracks on planes parallel to the uniaxial compression axis; (4) the loss of resistance in the transverse direction transfers further stress on the inclined microplanes, thus leading to further softening in the transverse direction, etc. Note that if a maximum stress and strain softening were assumed to exist for the compression behavior on the microplane level, it would be impossible to obtain the correct response in the hydrostatic pressure test in which there is no stress peak and no strain softening.

The data sets that exist in the literature (Figs. 1–10) were obtained in different laboratories and therefore somewhat different parameters had to be used to fit each particular data set. Unfortunately, no consensus has been reached on the choice of some standard concrete which would be used in all the laboratories to make the results better comparable. To demonstrate that the response curves are reasonable for all the types of tests considered here (Tables 1 and 2) even if the same material parameters

are used for all the cases, the response curves corresponding to Figs. 1–10 have been run for the parameters of the concrete of Kupfer et al. (1969) (Tables 1 and 2). The curves are shown in Fig. 12.

The test data have been analyzed under the assumption that the deformation in the test specimens during the test was homogeneous. No doubt this was not so in many cases for the strain-softening regime, in which there is always a tendency for localization of strains. In view of this fact, the present formulation is applicable only for describing the behavior of the material elements whose size is roughly equal to the size of the test specimens used. Bodies larger than this have to be subdivided into finite elements of roughly this size in every region where strain softening is expected to occur, and finite elements larger or smaller than this are not allowed in such regions. However, if it is known that the deformation localizes, one can allow for approximate analysis the use of finite elements of a different size (larger or smaller) provided that the strain-softening portion of the stress-strain relation is adjusted so as to maintain the correct value of energy dissipation for the finite elements containing the localized strain-softening zone. This could be formulated in a manner similar to the crack band model for tensile fracture (Bažant and Gambarova 1984).

A general and theoretically sound method for handling strain localization instability due to strain softening is to adopt the recent nonlocal continuum formulation with local strain (Bažant and Pijaudier-Cabot 1987; Bažant et al. 1987; Pijaudier-Cabot and Bažant 1987). The present constitutive model can be implemented in such a nonlocal continuum formulation and this will be one objective of further study.

Conclusion

The conceptual simplicity of the present general microplane model is rewarded by ease of material identification. Several empirical material parameters can be fixed and considered the same for all concretes. Four of the material parameters can be identified independently of the others by simple fitting of a single stress-strain diagram for hydrostatic pressure loading. Finally, there remain only five material parameters that have to be identified by fitting simultaneously the responses measured in various types of triaxial loadings. By contrast, the existing phenomenologic macroscopic models that provide a comprehensive description of the material in many types of tests necessitate over 15 material parameters to be identified by data fitting. Good agreement with the existing basic test data is achieved.

Acknowledgment

The computer implementation and data fitting with the present model has been funded under U.S. National Science Foundation Grant No. MSM-8700830 to Northwestern University.

Appendix. References

Balmer, G. G. (1949). "Shearing strength of concrete under high triaxial stress—computation of Mohr's envelope as a curve." *Struct. Res. Lab. Report No. SP-23*, Denver, Colorado.

Bažant, Z. P. (1987). "Snapback instability at crack ligament tearing and its implication for fracture micromechanics." *Report No. 87-6/498s*, Center for Concrete and Geomaterials, Northwestern Univ., Evanston, Ill.
Bažant, Z. P., and Gambarova, P. G. (1984). "Crack shear in concrete: Crack band microplane model." *J. Struct. Engrg.*, ASCE, 110(9), 2015–2035.
Bažant, Z. P., and Pijaudier-Cabot, G. (1987). "Nonlocal continuum damage, localization instability and convergence." *Report No. 87-2/428n*, Center for Concrete and Geomaterials, Northwestern Univ., Evanston, Ill.
Bažant, Z. P., and Shieh, C.-L. (1978). "Endochronic model for nonlinear triaxial behavior of concrete." *Nucl. Eng. Des.*, 47, 305–325.
Bažant, Z. P., Bishop, F. C., and Chang, T.-P. (1986). "Confined compression tests of cement paste and concrete up to 300 ksi." *J. Am. Concr. Inst.* 33(4), 553–560.
Bažant, Z. P., Lin, F.-B., and Pijaudier-Cabot, G. (1987). "Yield limit degradation: Nonlocal continuum with local strains." *Proc. Int. Conference on Computational Plasticity*, held in Pineridge Press, Swansea, England, 1757–1779.
Bresler, B., and Pister, K. S. (1958). "Strength of concrete under combined stresses." *J. Am. Concr. Inst.*, 551(9), 321–345.
Goode, C. D., and Helmy, M. A. (1967). "The strength of concrete under combined shear and direct stress." *Mag. Concr. Res.* 19(59), 105–112.
Green, S. J., and Swanson, S. R. (1973). "Static constitutive relations for concrete." *Report No. AFWL-TR-72-2*, Air Force Weapons Lab., Kirtland Air Force Base.
Hognestad, E., Hanson, N. W., and McHenry, D. (1955). "Concrete stress distribution in ultimate strength design." *J. Am. Concr. Inst.* 52(4), 455–477.
Kotsovos, M. D., and Newman, J. B. (1978). "Generalized stress-strain relations for concrete." *J. Engrg. Mech.*, ASCE, 104(4), 845–856.
Kupfer, H., Hilsdorf, H. K., and Rüsch, H. (1969). "Behavior of concrete under biaxial stresses." *J. Am. Concr. Inst.* 66, 656–666.
Launay, P., and Gachon, H. (1971). "Strain and ultimate strength of concrete under triaxial stress." *Proc. 1st. Int. Conference on Struct. Mech. in Reactor Technology*, Berlin, West Germany.
Lin, F. B. et al. (1987). "Concrete model with normality and sequential identification." *Comput. Struct.* 26(6), 1011–1026.
Petersson, P. E. (1981). "Crack growth and development of fracture zones in plain concrete and similar materials." *Report No. TVBM 1006*, Lund Inst. of Tech., Sweden.
Pijaudier-Cabot, G., and Bažant, Z. P. (1987). "Nonlocal damage theory." *J. Eng. Mech.* ASCE, 113(10), 1512–1533.
Sinha, B. P., Gerstle, K. H., and Tulin, L. G. (1964). "Stress-strain relations for concrete under cyclic loading." *J. Am. Concr. Inst.* 62(2), 195–210.
van Mier, J. G. M. (1984). "Strain-softening of concrete under multiaxial loading conditions," thesis presented to University Eindhoven, the Netherlands, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING, VOL. 26, 1805–1823 (1988)

NON-LOCAL YIELD LIMIT DEGRADATION

ZDENĚK P. BAŽANT* AND FENG-BAO LIN†

Department of Civil Engineering, Northwestern University, Evanston, IL 60208, U.S.A.

SUMMARY

Presented is a new type of a non-local continuum model which avoids problems of convergence at mesh refinement and spurious mesh sensitivity in a softening continuum characterized by degradation of the yield limit. The key idea, which has recently been proposed in a general context and has already been applied to softening damage due to stiffness degradation, is to apply the non-local concept only to those parameters which cause the degradation while keeping the definition of the strains local. Compared to the previously advanced fully non-local continuum formulation, the new approach has the advantage that the stresses are subjected to the standard differential equations of equilibrium and standard boundary or interface conditions. The new formulation exhibits no zero-energy periodic modes, imbrication of finite elements is unnecessary and finite elements with standard continuity requirements are sufficient. Two-dimensional finite element solutions with up to 3248 degrees of freedom are presented to document convergence and efficacy. The formulation is applied to tunnel excavation in a soil stabilized by cement grouting, with the objective of preventing cave-in (burst) of the tunnel sides due to compression softening.

INTRODUCTION

As is now well documented, macroscopic description of distributed cracking or other types of damage in various materials requires constitutive models which exhibit strain-softening—a decrease of stress at increasing strain, or in general a loss of positive definiteness of the tangential moduli matrix of the material. This behaviour is observed macroscopically in many materials, including concrete, rocks, stiff soils, grauted soils, two-phase ceramics, various fibre composites, wood and particle board, paper, ice, especially sea ice, coal, filled elastomers, polymer and asphalt concretes, fibre-reinforced concretes, various refractories and metals exhibiting void nucleation and growth in large strains.

The loss of positive definiteness of the elastic moduli matrix was probably first considered by Hadamard[28] who pointed out that this property is inadmissible from the mathematical viewpoint since the wave speed ceases to be real, the hyperbolic differential equation of motion of the continuum changes its type to an elliptic differential equation, and consequently wave propagation becomes impossible. This argument was later restated and expanded by Hill, Thomas, Mandel, Truesdell and others.[5] About a decade ago, however, it became clear from experimental results that Hadamard's argument is not quite realistic since real materials in a strain-softening state possess not only a non-positive definite moduli matrix for further loading, but also another incremental moduli matrix for unloading which is always positive definite. Thus, Hadamard's argument about imaginary wave speed applies only to loading waves, while unloading waves can

* Professor of Civil Engineering

† Graduate Research Assistant; presently Assistant Professor of Civil Engineering, Polytechnic University, Brooklyn, N.Y. 11201, USA

0029–5981/88/081805–19$09.50

© 1988 by John Wiley & Sons, Ltd.

Received October 1987
Revised December 1987

still propagate in a strain-softening material. This fact makes a crucial difference. This is one reason which rendered strain-softening problems mathematically meaningful, even for a local continuum.

Solutions to certain wave propagation problems for a local continuum with strain-softening do exist, and some are unique.[2, 3, 7, 16, 21, 22] Explicit step-by-step finite element solutions converge to these continuum solutions, and appear to do so almost quadratically. It must be noted that, as Sandler[54] emphasized, problems of this type generally belong to the class of ill-posed initial-boundary-value problems. The ill-posedness *per se*, however, cannot be considered physically objectionable. Ill-posed problems, along with a change of type of the differential equation, are accepted as physically realistic in other branches of physics and engineering, e.g. fluid mechanics.[31] Thus, it appears that the problem with strain-softening models is physical rather than mathematical.

Noting that a positive stiffness must be considered for unloading in a softening material, Bažant[2, 3] showed in his stability analysis of static strain-softening in a bar that, according to a local continuum model, strain-softening must localize into a region of vanishing volume. The same type of behaviour is exhibited by dynamic wave propagation solutions of strain-softening structures.[6, 7, 12, 21, 22] The consequence of this behaviour is that the structure is predicted to fail with a zero energy dissipation. This is of course physically impossible and represents an unacceptable model.

To prevent such unacceptable behaviour, one must introduce mathematical models called localization limiters, which force the strain-softening region to have a certain minimum finite size and thus ensure the energy dissipation at failure to be finite. The simplest, albeit crudest, localization limiter, which was proposed in 1974[2] and was later adopted in the crack band model,[3, 6, 14] consists in limiting the finite element size by a certain minimum value, which represents a material property and is approximately related to the size of inhomogenities in the material. Although this approach appears to yield good results for problems of fracture mechanics type, in which the fracture is localized, it might nevertheless be too crude in that it makes it impossible to resolve the distribution of damage density throughout the failure region, thus precluding precise calculation of the energy dissipation, which is important especially for dynamic problems such as impact.

The most general form of a localization limiter appears to be the concept of non-local continuum. This concept, introduced on the basis of statistical analysis of heterogeneous materials by Kröner,[35] Krumhanzl,[36] Kunin, and Levin,[5] and widely applied by Eringen and Edelen[25] and others, was proposed for application to strain-softening models in References 4 and 6. At present it is clear that numerous particular forms of non-local formulations are possible. One new form which appears to be simple and effective for the modelling of strain-softening behaviour was recently introduced in References 19, 48 and 49 in the context of continuum damage mechanics. The basic idea was that only the softening damage should be treated as non-local while the elastic behaviour should be treated as local. Bažant and Pijaudier-Cabot[18, 19] extended this idea further, proposing that, most generally, an efficient model for strain-softening should subject to non-local description only those variables which cause strain-softening and should retain a local definition of strain. Thus the material model may be characterized as a non-local continuum with local strain. This concept may be applied to all possible constitutive models with strain softening. Its micromechanics justification has been presented.[58]

Ideally, two kinds of strain softening (or damage) may be distinguished.

1. Degradation of material stiffness, which is caused by microcracking, or by void nucleation and growth. In this case the unloading is ideally a straight line towards the origin (Figure 1(a)).

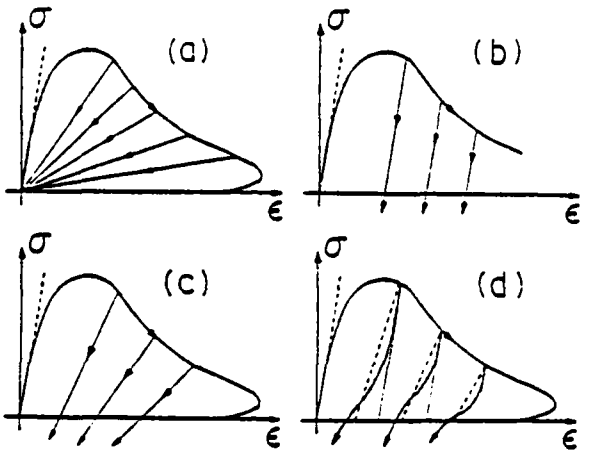


Figure 1. Various types of unloading characteristic of strain-softening materials

2. Degradation of the yield limit, which is associated with the reduction of the areas of cohesive or plastic–frictional connections within the material and does not cause a decrease of material stiffness. In this case the unloading slope is the same as the initial elastic slope (Figure 1(b)).

The real behaviour is better approximated as a combination of the foregoing two basic types (Figure 1(c)). The unloading slope in reality does not remain constant but varies during the progress of unloading (Figure 1(d)). The initial unloading slope is often very close to the initial elastic slope (Figure 1(d)). So, for the initial unloading behaviour, the yield limit degradation (Figure 1(b)) is a good simplification.

The objective of the present study, which was summarized at a recent conference,[13] is to formulate for yield limit degradation a non-local continuum model with local strain, verify that it exhibits proper convergence in displacements, stresses and energy, and apply it in finite element analysis of tunnel excavation. As for the degradation of stiffness, a non-local model with local strain has already been developed and demonstrated by finite element examples.[15, 18, 19]

Before embarking on our analysis, it is proper to mention that localization limiters may be given alternative forms. For example, the averaging integral in the non-local model may be expanded in Taylor series. This leads to higher-order gradients in the definition of strains and in the continuum equations of motion.[1-4] In particular, higher-order spatial gradients may be applied to the strength limit or yield limit, or to the yield function in general. This type of idea was apparently first proposed in 1952 by L'Hermite[42] as part of his study of shrinkage cracking. Later the idea of the dependence of strength limit on strain gradient was applied in the problems of bending of concrete beams with earthquake engineering applications.[32] Recently, the idea of introducing gradients of strength or of the yield condition was developed in the finite element context.[26, 43, 55, 56]

As another alternative, one may avoid the problem of localization by lumping all the softening deformations into a line. This leads to the line crack models with softening non-linear cohesive zone (line segment) ahead of the crack tip (cf. Reference 5), which were pioneered for concrete by Hillerborg with co-workers.[29, 30] This approach is applicable to cracks whose spacing is not less than a certain characteristic length that must be considered to be a material property.[5] Combined models, in which a line crack or strip is embedded in a larger finite element, were formulated by

Pietruszczak and Mróz.[47] Willam *et al.*[37] and others. Some authors have criticized the idea of strain-softening on various grounds:[51, 54] however, the criticized aspects have been overcome by the non-local models as well as other mathematical devices.

Much attention has recently been focused on the models of continuum damage mechanics.[33, 34, 37–41, 44] These models, however, were considered as local, and it has been overlooked that in such a form they are plagued by the same problems of localization (and the associated spurious mesh sensitivity) as the other strain-softening models.[8] It is nevertheless possible to rectify these problems by incorporating into the continuum damage mechanics formulation the concept of non-local continuum with local strain.[19, 49]

As another device to overcome the problems due to change of type of the differential equation, it has been proposed to introduce various forms of strain-rate sensitivity or artificial viscosity.[45] This seems, though, only a partial remedy, since strain and energy dissipation can localize to an arbitrarily small volume after a sufficient time period. Finally, it should be noted that the localization of strain and damage due to softening, analysed from the stability viewpoint already in 1974 (Bažant[2]), is not the same problem as the localization problem analysed by Rudnicki and Rice,[53] Rice,[52] Prevost[50] and others. In those problems, the localization was of primarily geometrically non-linear origin.

NON-LOCAL CONTINUUM WITH LOCAL STRAIN

The typical property of non-local continuum is that some of the variables in the constitutive equation are defined by spatial averaging. Thus, for example, the spatial average of the magnitude of plastic strain ε^p at location $\mathbf{x}$ may be defined by the equation

$$\langle \varepsilon^p(\mathbf{x}) \rangle = \frac{1}{V_r(\mathbf{x})} \int_V \alpha(\mathbf{s}-\mathbf{x}) \varepsilon^p(\mathbf{s})\, dV = \int_V \alpha'(\mathbf{x}, \mathbf{s}) \varepsilon^p(\mathbf{s})\, dV \tag{1}$$

in which

$$V_r(\mathbf{x}) = \int_V \alpha(\mathbf{s}-\mathbf{x})\, dV \tag{2}$$

$$\alpha'(\mathbf{x}, \mathbf{s}) = \alpha(\mathbf{s}-\mathbf{x})/V_r(\mathbf{x}) \tag{3}$$

The pointed brackets $\langle\ \rangle$ denote the averaging operator, V = volume of the body and $\alpha(\mathbf{x})$ = weighting function which defines the averaging; $\mathbf{s}$ is the general co-ordinate vector. V_r has approximately but not exactly the same meaning as the representative volume in the statistical theory of heterogeneous materials (Kröner;[35] Krumhansl[36]). The averaging may be specified by a uniform weighting function, $\alpha = 1$, which is non-zero only within a representative volume such as a circle in two dimensions or a sphere in three dimensions. However, numerical computations show much better convergence if the weighting function is smooth. Thus, a suitable form of the weighting function for numerical computation appears to be, for example, the normal (Gaussian) distribution function (error density function):

$$\alpha(\mathbf{x}) = e^{-(k|\mathbf{x}|/l)^2} \tag{4}$$

in which, for one, two and three dimensions

$$\text{1D:} \quad |\mathbf{x}|^2 = x^2, \qquad k = \sqrt{\pi} = 1{\cdot}772 \tag{5}$$

$$\text{2D:} \quad |\mathbf{x}|^2 = x^2 + y^2, \qquad k = 2 \tag{6}$$

$$\text{3D:} \quad |\mathbf{x}|^2 = x^2 + y^2 + z^2, \qquad k = (6\sqrt{\pi})^{1/3} = 2{\cdot}149 \tag{7}$$

l is the characteristic length, a material property which defines the diameter of the representative volume (a line segment, circle or sphere); l is determined from the condition that the representative volume has the same volume as the normal distribution function extending to infinity (x, y, z = Cartesian co-ordinates).

For a finite body, the normal distribution function extends beyond the boundary of the body. In that case the region outside the body is deleted from the integration domain V, both for the calculation of the average (equation (1)) and the calculation of the representative volume V_r (equation (3)). This fact causes V_r to depend on location x.

For numerical finite element computations, the integrals in equations (1)–(3) are approximated by finite sums over all the integration points of all the elements. Only the finite elements whose integration points are closer to point x than distance $2l$ need to be included in the sum since for a greater distance the value of α is negligible.

In the initial application of the non-local concept to strain-softening,[4] it was assumed that the strain was non-local. In such a case the variational derivation of the differential equations of motion or equilibrium and of the boundary conditions from the virtual work principle[4] yields field equations and boundary conditions of a non-standard form, with extra terms in the form of integrals or higher-order derivatives. This aspect causes considerable complications in numerical implementation. Recently[18, 19, 48] it was discovered that application of the averaging operator to strains is not necessary, and that a sufficient localization limiter may be obtained if the non-local averaging is applied only to those variables which control strain-softening. In that case the statement of the virtual work principle is the same as for the classical continuum, from which it has been mathematically proved that the differential equations of motion or equilibrium with their boundary conditions as well as interface conditions remain the same as in the classical theory (see Bažant and Pijaudier-Cabot[18, 19]). Consequently, the continuity requirements for the finite elements are the same as for a local finite element model (and no element imbrication, introduced in the initial non-local model for strain-softening, is needed). It has also been generally proved that the damage dissipation, which is controlled by the variables that govern softening, cannot localize to a region of vanishing volume.[19]

REVIEW OF CLASSICAL INCREMENTAL PLASTICITY

First we must define the local plasticity model which we want to generalized to non-local form. The yield surfaces are written as $F(\boldsymbol{\sigma}, \kappa)=0$ and the yield function is assumed to have the form $F(\boldsymbol{\sigma}, \kappa)=f(\boldsymbol{\sigma})-\kappa$, in which $\boldsymbol{\sigma}$ = stress tensor; function $f(\boldsymbol{\sigma})$ may be regarded as the effective stress σ^*, and κ is a hardening–softening parameter which may be interpreted as the yield limit; κ may in general not only increase (hardening) but also decrease (softening). We assume for the sake of simplicity that there is only one hardening–softening parameter. As usual, we require the increments $\Delta\boldsymbol{\varepsilon}^p$ of the plastic strain tensor $\boldsymbol{\varepsilon}^p$ to follow the normality rule

$$\Delta\boldsymbol{\varepsilon}^p = F_{,\sigma}\,\Delta\lambda = f_{,\sigma}\,\Delta\lambda \tag{8}$$

where a comma preceding a subscript denotes a derivative, i.e. $F_{,\sigma}=\partial F/\partial\boldsymbol{\sigma}$, and $\Delta\lambda$ represents a scalar proportionality coefficient. The loading–unloading criterion is defined by the conditions

$$F\Delta\lambda=0, \quad \Delta\lambda\geqslant 0, \quad F\leqslant 0 \tag{9}$$

which must all hold simultaneously.

We may now assume that κ depends on $\boldsymbol{\varepsilon}^p$ but not directly on $\boldsymbol{\sigma}$, and introduce Prager's continuity condition

$$F_{,\sigma}:\Delta\boldsymbol{\sigma} + F_{,\kappa}\,\kappa_{,\varepsilon^p}:F_{,\sigma}\Delta\lambda=0$$

where the colon (:) denotes a tensor product contracted on two inner indices. We may also denote

$$H = -F_{,\kappa}\,\kappa_{,\varepsilon^p}\,\varepsilon^p_{,\varepsilon^p} : F_{,\sigma}$$

where H may be called the plastic tangential modulus and κ is assumed to depend on ε^p. Thus the continuity condition becomes $F_{,\sigma} : \Delta\sigma - H\Delta\lambda = 0$, i.e. $\Delta\lambda = (F_{,\sigma} : \Delta\sigma)/H$. Substituting this and equation (8) into the incremental stress–strain relation

$$\Delta\sigma = \mathbf{D} : (\Delta\varepsilon - \Delta\varepsilon^p) \tag{10}$$

where $\mathbf{D}$ is the tensor of elastic moduli, we may solve for $\Delta\lambda$ and obtain the well-known relation[46]

$$\Delta\lambda = \frac{F_{,\sigma} : \mathbf{D} : \Delta\varepsilon}{F_{,\sigma} : \mathbf{D} : F_{,\sigma} + H} \tag{11}$$

Usually the effective plastic strain either is considered as the path length of plastic strain, i.e. $\Delta\varepsilon^p = (\Delta\varepsilon^p : \Delta\varepsilon^p)^{1/2}$, or defined by work equivalence $\sigma : \Delta\varepsilon^p = \sigma^* \Delta\varepsilon^p$. We choose the latter definition, and then

$$\Delta\varepsilon^p = \frac{1}{\sigma^*}\sigma : F_{,\sigma}\Delta\lambda \tag{12}$$

The yield function can be written in the form $F(\sigma, \kappa) = f(\sigma) - \kappa$; function $f(\sigma)$ may normally be assumed to be a homogeneous function of degree n, for which Euler's theorem states that $\sigma : f_{,\sigma} = nf$. Furthermore, the yield function may be written in such a form that $n = 1$ and $f(\sigma) = \sigma^*$. Also, the hardening–softening parameter κ may be assumed to be a function of the effective plastic strain ε^p rather than a general scalar function of ε^p, i.e. $k = k(\varepsilon^p)$. Then the aforementioned expression for H may be further simplified as follows:

$$\begin{aligned} H &= -F_{,\kappa}\,\kappa_{,\varepsilon^p}\,\varepsilon^p_{,\varepsilon^p} : F_{,\sigma} = \kappa_{,\varepsilon^p}\,\varepsilon^p_{,\varepsilon^p} : f_{,\sigma} \\ &= \kappa_{,\varepsilon^p}\frac{1}{\sigma^*}(\sigma : f_{,\sigma}) = \kappa_{,\varepsilon^p}\frac{nf}{\sigma^*} = \kappa_{,\varepsilon^p} \end{aligned} \tag{13}$$

NON-LOCAL GENERALIZATION

According to the concept of non-local continuum with local strain,[19, 56] equation (10) must be generalized as

$$\Delta\sigma = \mathbf{D} : (\Delta\varepsilon - \Delta\bar{\varepsilon}^p) \tag{14}$$

in which $\Delta\varepsilon$ remains to be local while $\Delta\bar{\varepsilon}^p$ is the non-local plastic strain increment. Note that we distinguish the superimposed bar, which is a label for non-local quantities, from the pointed brackets $\langle\ \rangle$, which represent an operator.

We now present the algorithm of the initial stiffness method for an implicit finite element solution with step-by-step loading. The computations in each loading step proceed as follows.

1. The calculation begins by carrying out an elastic finite element analysis of the structure using the applied load increments $\Delta\mathbf{f}$ (or prescribed displacement increments) and the nodal forces $\Delta\mathbf{f}''$. For the first iteration, $\Delta\mathbf{f}''$ are taken as zero. For the subsequent iterations, forces $\Delta\mathbf{f}''$ are determined from the previous iteration of this loading step; they represent the nodal forces equivalent (according to the principle of virtual work) to the unbalanced stresses determined in step 6 of this algorithm (defined below). This finite element analysis yields $\Delta\mathbf{u}$ and $\Delta\varepsilon$, where $\Delta\mathbf{u}$ is

the column matrix of all nodal displacement increments and $\Delta\varepsilon$ is the corresponding column matrix of strain increments (for all integration points of all finite elements). The elastic stress increments are obtained as

$$\Delta\sigma^e = \mathbf{D}:\Delta\varepsilon \tag{15}$$

and the final stress values are $\sigma_F = \sigma_I + \Delta\sigma^e$, where subscripts I and F refer to the initial and final values for the loading step.

2. Next we calculate for all the integration points of all the elements the values $F_F = F(\sigma_I + \Delta\sigma^e, \kappa_F)$, where κ_F is the final value of κ for the loading step and is taken for the first iteration as $\kappa_F = \kappa_I$, and for subsequent iteration as the final κ-value from the preceding iteration. If $F_F < 0$, the material at the integration point is still in the elastic range or unloaded.

3. Now $F_F > 0$. This means that the end value of stress is outside the loading surface and must therefore be adjusted to the loading surface. First we need to determine the ratio r $(0 < r < 1)$ such that $F(\sigma_I + r\Delta\sigma^e, \kappa_F) = 0$. Approximately, by linear interpolation, $F_I + r(F_F - F_I) = 0$, from which we calculate $r = F_I/(F_F - F_I)$ for all integration points of all elements.

4. To take into account the plastic strain increments, we calculate (for all the integration points of all elements) the value of H according to equation (13). For $F_{,\sigma}$, $\kappa_{,\varepsilon^p}$ and $\varepsilon^p_{,\varepsilon^p}$ we use the values corresponding to σ_I and ε^p_I for the first iteration, and to $\sigma_I + \Delta\sigma$ and $\varepsilon^p_I + \Delta\varepsilon^p$ for the subsequent iterations, with $\Delta\sigma$ and $\Delta\varepsilon^p$ taken from the preceding iteration. The strain increments corresponding to the previously determined stress increments $(1-r)\Delta\sigma^e$ are $(1-r)\Delta\varepsilon$. Thus, according to equation (11), we calculate

$$\Delta\lambda = \frac{F_{,\sigma}:\mathbf{D}:(1-r)\Delta\varepsilon}{F_{,\sigma}:\mathbf{D}:F_{,\sigma} + H} \tag{16}$$

for all the integration points of all elements. These are still the local $\Delta\lambda$ values. $F_{,\sigma}$ are here evaluated for σ_I for the first iteration and for $\sigma_I + \Delta\sigma$ for the subsequent iterations using always $\Delta\sigma$ from the preceding iteration. The corresponding local plastic strain increments are then calculated as $\Delta\varepsilon^p = F_{,\sigma}\Delta\lambda$ (equation (8)), for all the integration points of all elements.

5. Now comes the non-local operation which yields for each integration point of each element the average non-local plastic strain increment:

$$\langle\Delta\varepsilon^p(\mathbf{x})\rangle = \int_V \alpha'(\mathbf{x}, \mathbf{s})\Delta\varepsilon^p(\mathbf{s})\,dV \approx \sum_j \alpha'_{(i,j)}\Delta\varepsilon^p_{(j)} \tag{17}$$

The last expression is a numerical approximation of the integral. The subscripts in parentheses refer to the integration points of all finite elements in the system, and $\alpha'_{(i,j)}$ denotes the values of the function $\alpha'(\mathbf{x}, \mathbf{s})$ for the integration point $\mathbf{x}_{(j)}$ (of this or another element) relative to point $\mathbf{x}_{(i)}$, i.e. $\alpha'_{(i,j)} = \alpha(\mathbf{x}_{(j)} - \mathbf{x}_{(i)})/V_{r(i)}$. Note that, for sufficiently large $|\mathbf{x}_{(j)} - \mathbf{x}_{(i)}|$, i.e. for sufficiently remote integration points, the value of $\alpha'_{(i,j)}$ can be neglected because the normal probability function decays with distance very rapidly. For numerical implementation, it is convenient to evaluate all the coefficients $\alpha'_{(i,j)}$ in advance of the finite element solution and store them as a two-dimensional array in the computer's memory.

6. The definition of the non-local equivalent plastic strain increment $\Delta\bar{\varepsilon}^p$ may be considered similar to equation (12), i.e. $\sigma^*\Delta\bar{\varepsilon}^p = \sigma:\langle\Delta\varepsilon^p\rangle$. Accordingly, we evaluate for each integration point of each element

$$\Delta\bar{\varepsilon}^p = \frac{1}{\sigma^*}\sigma:\langle\Delta\varepsilon^p\rangle \tag{18}$$

and define non-local $\Delta\bar{\lambda}$ by the relation $\boldsymbol{\sigma}:F_{,\sigma}\Delta\bar{\lambda}=\boldsymbol{\sigma}:\langle\Delta\boldsymbol{\varepsilon}^p\rangle$. Then we may evaluate for each integration point of each element

$$\Delta\bar{\lambda}=\frac{\boldsymbol{\sigma}:\langle\Delta\boldsymbol{\varepsilon}^p\rangle}{\boldsymbol{\sigma}:F_{,\sigma}}=\frac{\sigma^*\Delta\bar{\varepsilon}^p}{\boldsymbol{\sigma}:F_{,\sigma}} \tag{19}$$

These non-local values differ from the local $\Delta\lambda$ given by equation (16) and are not equal to $\langle\Delta\lambda\rangle$. Also, note that equation (19) does not imply $\langle\Delta\boldsymbol{\varepsilon}^p\rangle$ to be equal to $F_{,\sigma}\Delta\lambda$. Therefore, the non-local plastic strain increments $\Delta\bar{\boldsymbol{\varepsilon}}^p$ which are normal to the loading surface may now be obtained for all the integration points of all elements as

$$\Delta\bar{\boldsymbol{\varepsilon}}^p=F_{,\sigma}\,\Delta\bar{\lambda} \tag{20}$$

The final stresses $\boldsymbol{\sigma}'_F$ for the loading step are first estimated as

$$\boldsymbol{\sigma}'_F=\boldsymbol{\sigma}_1+r\,\Delta\boldsymbol{\sigma}^e+\mathbf{D}:[(1-r)\Delta\boldsymbol{\varepsilon}-\Delta\bar{\boldsymbol{\varepsilon}}^p] \tag{21}$$

in which $(1-r)\Delta\boldsymbol{\varepsilon}$ corresponds to the stresses outside the yield surface. Next we obtain the final value $\kappa_F=\kappa(\bar{\varepsilon}^p_F)$, where $\bar{\varepsilon}^p_F=\bar{\varepsilon}^p_1+\Delta\bar{\varepsilon}^p=$ final value of the equivalent plastic strain. The $\boldsymbol{\sigma}'_F$ from equation (21) does not lie on the yield surface. We apply the radial return method (Figure 2) to adjust it to the yield surface and get the correct final stresses $\boldsymbol{\sigma}_F$. The differences between $\boldsymbol{\sigma}_1+\Delta\boldsymbol{\sigma}^e$ and $\boldsymbol{\sigma}_F$ represent the unbalanced parts of the elastic stress increments, and must therefore be balanced by applying additional residual forces. This is done later in step 1 of the next iteration. The objective of the iterations of the loading steps is to make the residual forces approximately vanish.

7. This is approximately achieved by imposing the condition $\sum(\text{residual force})^2<\varepsilon_1$, where the sum is over every nodal point and ε_1 is a specified small number representing the tolerance. If the tolerance is not satisfied, then we return to step 1 and start the next iteration of the same loading step. If it is satisfied we also return to step 1 but start the first iteration of the next loading step.

The loading–unloading criterion which is implied by steps 2, 3 and 6 of the foregoing non-local algorithm may be concisely stated as

$$F\Delta\bar{\lambda}=0,\quad \Delta\bar{\lambda}\geqslant 0,\quad F\leqslant 0 \tag{22}$$

As an alternative to the foregoing algorithm, we may apply the averaging operation directly to $\Delta\lambda$ instead of $\Delta\boldsymbol{\varepsilon}^p$. This is computationally more efficient since at each point there are many components of $\Delta\boldsymbol{\varepsilon}^p$ but only one $\Delta\lambda$-value. In this alternative, instead of equation (19) we calculate for each integration point of each element the non-local value as

$$\Delta\bar{\lambda}=\langle\Delta\lambda\rangle=\left\langle\frac{\boldsymbol{\sigma}:\Delta\boldsymbol{\varepsilon}^p}{\boldsymbol{\sigma}:F_{,\sigma}}\right\rangle=\sum_j\alpha'_{(i,j)}\left(\frac{\boldsymbol{\sigma}:\Delta\boldsymbol{\varepsilon}^p}{\boldsymbol{\sigma}:F_{,\sigma}}\right)_{(j)} \tag{23}$$

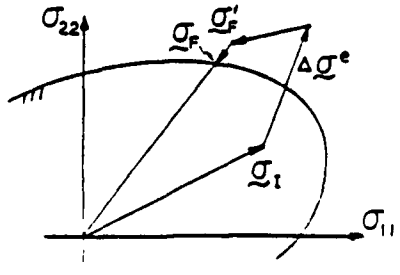


Figure 2. Radial return method for plastic analysis

The non-local plastic strain increments then are $\Delta\bar{\boldsymbol{\varepsilon}}^p = F_\sigma \Delta\bar{\lambda}$ (equation (20)) and the non-local effective plastic strain increment is $\Delta\bar{\varepsilon}^p = (\boldsymbol{\sigma}:\Delta\bar{\boldsymbol{\varepsilon}}^p)/\sigma^*$, from which we get the value of $\kappa(\bar{\varepsilon}^p_n)$. The remaining equations and the computational algorithm are the same as before. This alternative has been found to yield about the same numerical results as the originally described formulation with the averaging of $\Delta\boldsymbol{\varepsilon}^p$, and appeared to converge equally well.

EXAMPLE I. DAMAGE BANDS IN TENSIONED RECTANGULAR PANEL

Figure 3 shows the example of a rectangular panel subjected to uniform prescribed displacement on top boundary and fixed at the bottom boundary, with the lateral boundaries free. The problem is solved with three rectangular meshes using four-node quadrilateral finite elements. In order to initiate a softening band symmetrically, it is assumed that one element at the middle of each side has the initial yield limit smaller by 5 per cent than the remaining finite elements. The side of the finite element in the crudest mesh is assumed to be equal to the characteristic length l which figures in equation (4). The material is assumed to obey Tresca yield condition (Figure 4(b)), and the hardening–softening rule is specified in such a manner that after the attainment of the initial yield limit κ_0 the value of κ decreases linearly as a function of ε^p, with a slope H, until a zero value of κ is reached (Figure 4(b)). After that κ remains zero. Unloading and reloading are elastic.

Figures 4(a) and (b) show the curves of the reaction force P on top as a function of displacement u enforced at the top boundary. The results in Figure 4(a) are local, and those in Figure 4(b) are non-local, based on the present model. It is clear that the local response shows strong spurious mesh sensitivity in the softening range. Were the mesh refined further, the softening response would result in a snapback instability, i.e. the softening response after the peak point would revert to a positive slope. For the present non-local formulation, the results for the three meshes are approximately the same, and especially those for meshes 2 and 3 are nearly identical. This documents that the presently proposed non-local formulation eliminates the phenomenon of spurious mesh sensitivity which has plagued previous uses of strain-softening in finite element computation.

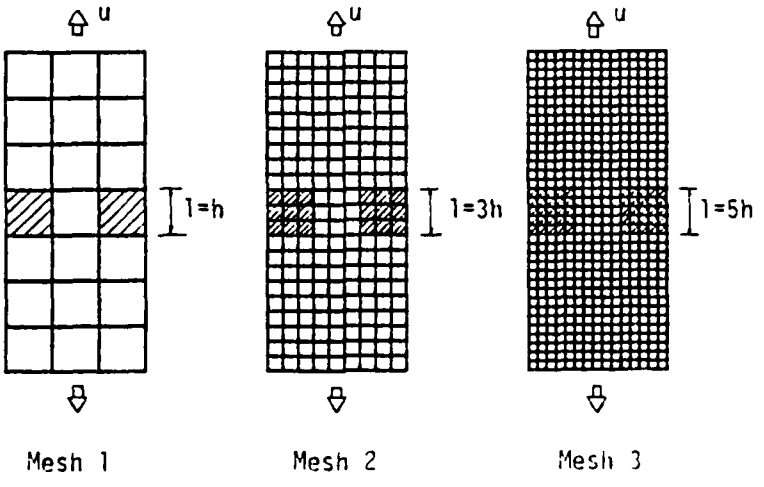


Figure 3. Rectangular panel with various mesh subdivisions

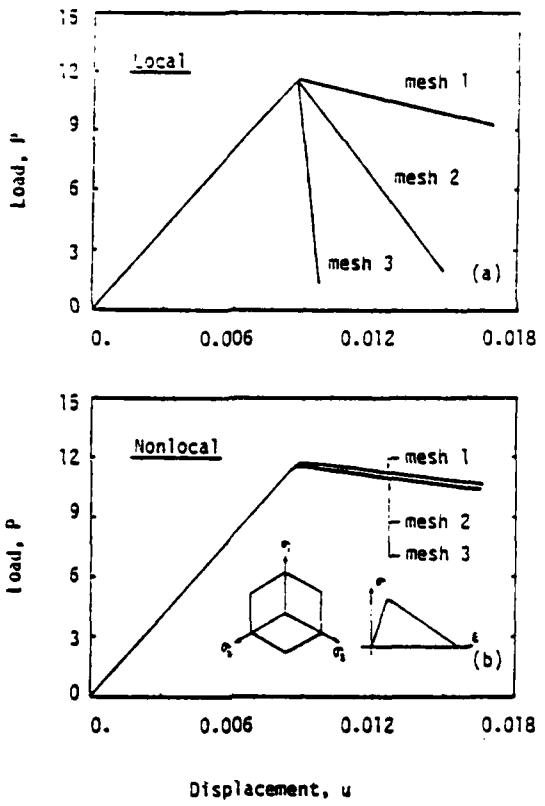


Figure 4. Load deflection diagrams calculated with various meshes according to the local and non-local plasticity model

The strain profiles across the cracking zone are shown in Figure 5. For the local case, the zones are seen to localize into sharper spikes when the mesh is refined. The most important aspect of strain-softening is the energy dissipation. As shown previously for the local solution, the energy dissipated due to strain-softening exhibits strong spurious mesh sensitivity due to localization of the softening front, and generally tends to zero as the mesh is refined to zero. This is documented for the local computation in Figure 6(a). With the present non-local model, by contrast, the energy dissipation appears to converge; see Figure 6(b) (N = number of elements in the mesh, and W = energy dissipated up to the instant at which the displacement on top is 0·025).

The softening zones at various stages of loading (for $u = 0{\cdot}00854$, 0·00868 and 0·00889) are plotted in Figure 7. Note that for the refined meshes the width of the softening zone is not a constant but varies. As a consequence of this fact, the energy dissipation rate per unit length of the strain-softening band is also not a constant but varies during the progress of loading. This has an interesting implication for non-linear line crack fracture models with a cohesive zone, such as Hillerborg's model for concrete: the fracture energy which would match the present results could not be considered as constant but would have to be variable. Moreover, due to the variation of the width of the softening band, the non-local solution cannot be equivalent to a line crack solution with a unique stress–displacement diagram. Rather, a different stress–displacement diagram would have to be considered for various stages of computation in order to match the present results.

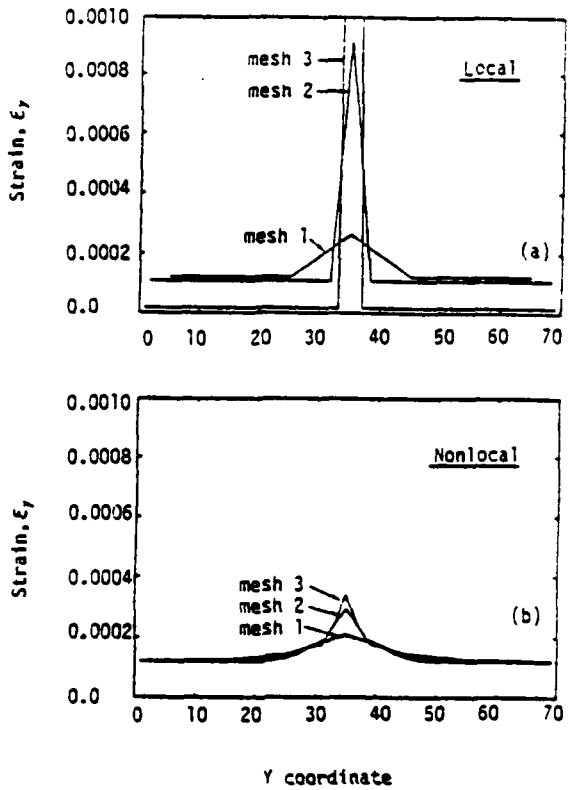


Figure 5. Profiles of vertical normal strain at vertical line of symmetry across strain-softening zone obtained for various meshes with local and non-local plasticity model

EXAMPLE II: COMPRESSION SOFTENING DUE TO TUNNEL EXCAVATION IN GROUTED SOIL

The present non-local model was applied to evaluate the safety of tunnel excavation in a softening soil, with particular reference to the construction of a subway in a non-cohesive granular soil stabilized by cement grouting. To make it possible to exacavate the tunnel by machines, the soil in the region of the tunnel is stabilized by injecting cement grout into bores drilled from the surface. During the excavation the surface of the excavation can be left unlined, which makes this construction method highly economical and susceptible to mechanization. Shortly after excavation, after the machine moves away, the tunnel surface is sprayed with fine concrete and later a relatively light concrete lining is cast *in situ*. The bores from the surface are spaced so densely that the grouted regions around the adjacent bores overlay, so that a grouted monolith is created by grouting. The grouted soil resembles concrete of very low strength and is highly inhomogeneous, with ungrouted pockets at a spacing of about 50 to 100 cm, and lumps of highly cemented stiff material.

The grouted soil mass can be treated as a continuum only on a scale sufficiently larger than the prevalent spacing of the weak pockets and highly cemented lumps. The characteristic length l of the homogenizing continuum should be roughly equal to this spacing, certainly in the order of magnitude, and therefore it is quite large, much larger than, e.g. for ordinary concrete. So this is an application where the modelling by non-local continuum seems particularly appropriate.

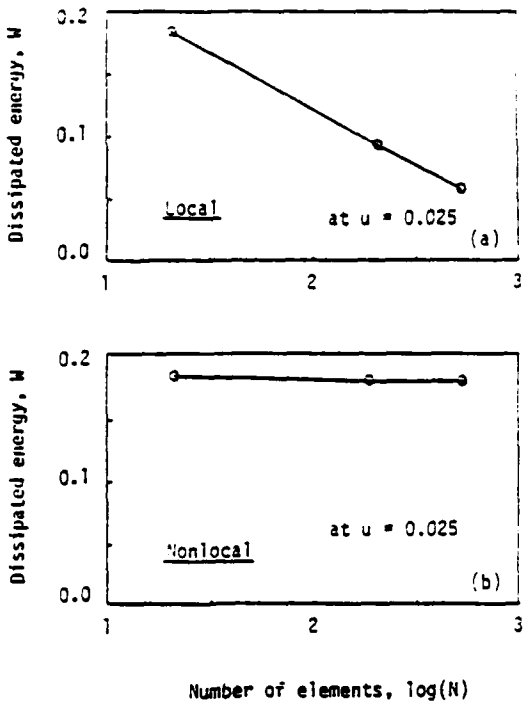


Figure 6. Convergence of energy dissipation at $u = 0{\cdot}025$ with increasing number of elements N in the mesh, for local formulation (top) and non-local formulation (bottom)

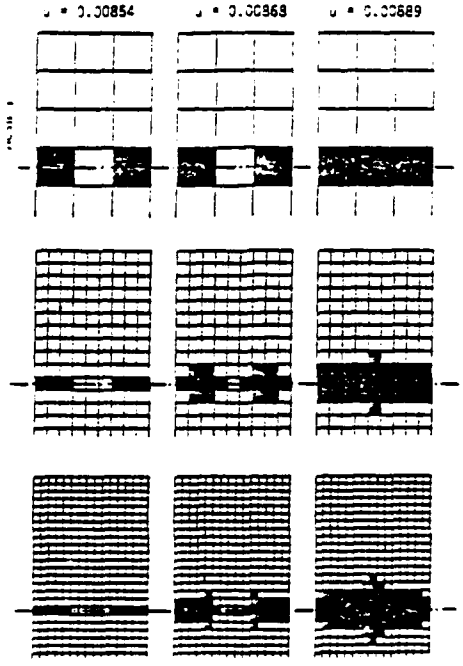
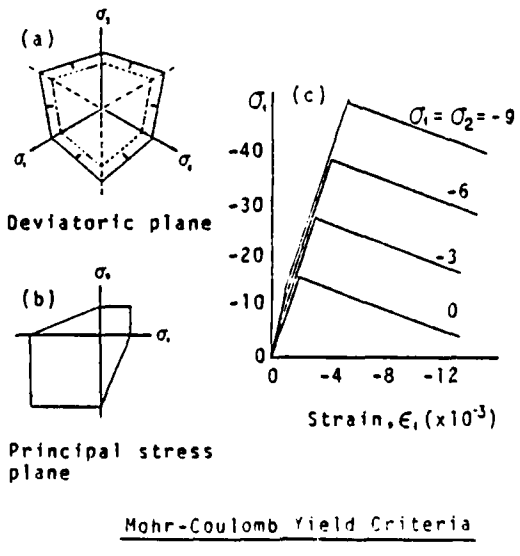


Figure 7. Softening zones at various stages of loading, for the three meshes in Figure 3

The tunnel and the calculation results are shown in Figures 8–12. The soil is characterized by a Mohr–Coulomb yield condition (Figure 8) with a yield limit that decreases with increasing plastic strain. Prior to excavation, the vertical and horizontal normal stresses in the soil are caused solely by gravity and they increase linearly with the depth. The values of the surface tractions on the tunnel contour are calculated from the initial stress values due to gravity. Excavation is simulated by progressively decreasing the values of these surface tractions to zero. The decrease of these surface tractions causes a decrease of the lateral confining stresses as well as an increase of the vertical compression stress near the tunnel sides. This results in strain-softening behaviour in compression near the sides of the tunnel and may lead to collapse, manifested as a cave-in (or burst) of the tunnel sides.

The geometry of the tunnel is defined in Figure 9. To assess convergence of the finite element solution with compressive strain-softening, the problem was solved (on a Cray X-MP/48 supercomputer) with four two-dimensional meshes shown in Figure 10, for which the number of degrees of freedom (unknown nodal displacement components) was $N = 218$, 608, 1840 and 3248. The characteristic length l of the grouted soil, shown graphically in Figure 10, has been assumed to be $l = 1{\cdot}8$ m, which is roughly equal to the typical spacing of the lumps of highly grouted soil and of weakly grouted pockets as observed at the construction site. The material properties are fully defined in Figure 8, in which $E = 11{\cdot}4$ ksi, $\nu = 0{\cdot}33$, $c = 5{\cdot}7$ psi, $\phi = 35°$, $H = -57{\cdot}8$ psi.

The problem was solved both conventionally (i.e. using the usual local form of the present plasticity model with strain-softening) and by the presently proposed non-local formulation. Figure 11 shows the plots (for the four meshes used) of the horizontal inward displacement u_1 at the tunnel side surface as a function of the degree D_e of excavation (defined as the ratio of the applied surface tractions to the surface tractions that existed initially, before the excavation started). Figure 12 shows for the completed excavation ($D_e = 1$) the profiles of the vertical normal stress σ_y on a horizontal line passing roughly at the midheight of the tunnel. The minimum point on these plots corresponds to the strength limit (the peak point of the stress–strain diagram). To the right of the minimum point, the soil is in a hardening state (elastic), and to the left it is in a



Mohr-Coulomb Yield Criteria

Figure 8. Mohr–Coulomb yield criterion with strain-softening due to yield limit degradation

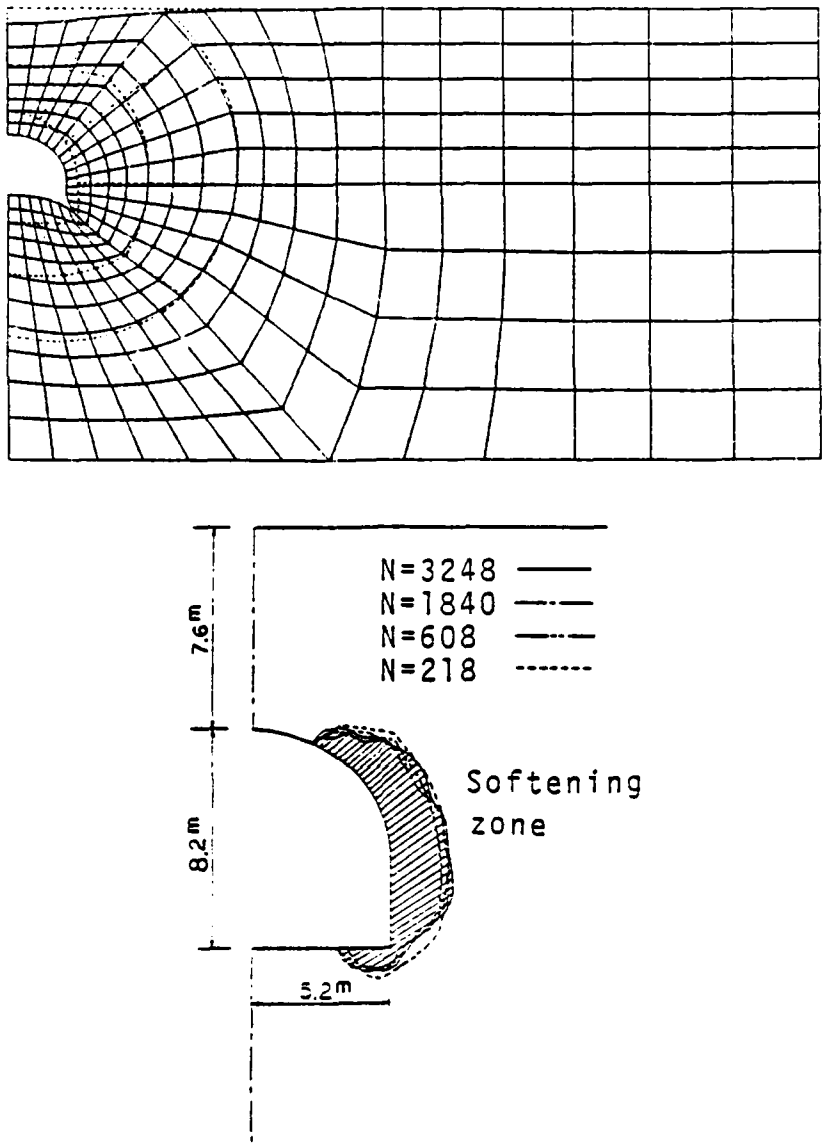


Figure 9. Boundaries of the softening zone at full tunnel excavation obtained for the four meshes shown in Figure 10, and the exaggerated deformation

state of plastic strain softening, in which the compression stress at tunnel surface is reduced to about 1/4 of the strength. The lateral expansion of the compressed softening soil is quite marked, as is evident from the deformed finite element mesh shown in Figure 9(a), in which the displacements are exaggerated 70 times.

The local analysis results (Figures 11, 12 top) for the four meshes reveal a poor convergence in the softening zone at the tunnel sides, i.e. strong spurious mesh sensitivity. These results are subjective since they strongly depend on the analyst's choice of the mesh. By contrast, the nonlocal analysis results (Figures 11, 12 bottom) for the four meshes exhibit a far better convergence.

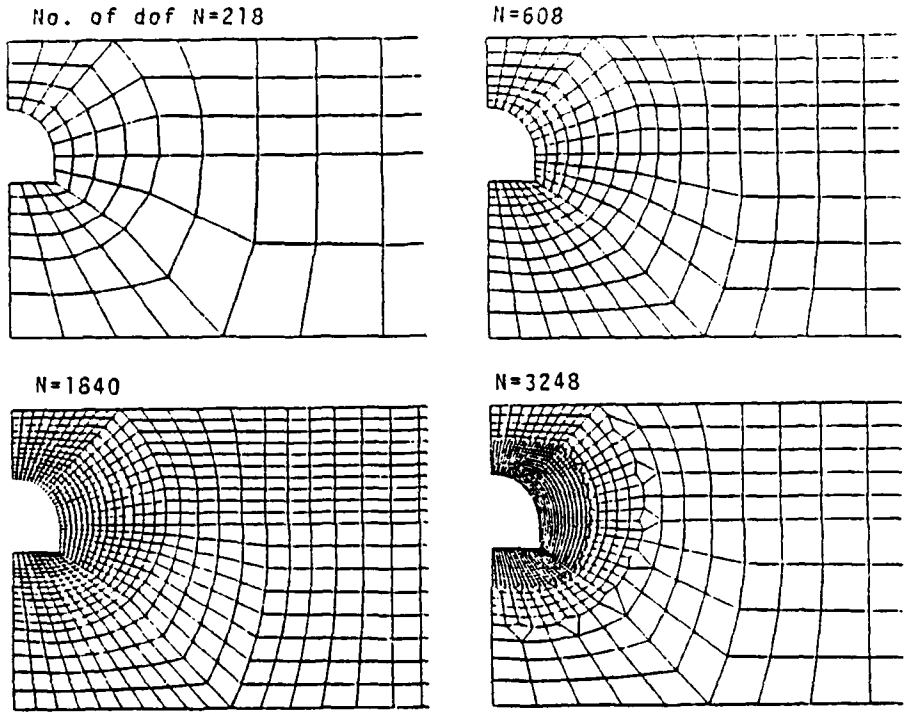


Figure 10. Finite element meshes in the analysis of a tunnel excavation in a grouted soil with a degrading yield limit

Especially the results for the meshes 3 and 4 are sufficiently close to each other for all practical purposes. So the non-local results appear to be objective.

The analysis indicates that a cave-in (burst) of the tunnel sides would not occur, i.e. the excavation without supporting the sides should be safe. In finite element calculations for various other values of material parameters, other dimensions and a different depth of the tunnel below the surface, it was often experienced that the displacements as a function of the degree of excavation would begin to increase rapidly and eventually convergence of the iterations of the loading step could not be achieved, even with the non-local formulation. Such response indicated that the tunnel would be unstable and could not be excavated safely.

Figure 9 also shows the boundaries of the softening zone obtained by non-local analysis with the four meshes. Note that all these boundaries are nearly identical, which confirms good convergence. The softening boundary is defined as the locus of the points at which the material is at the peak stress state and represents for the present constitutive model also the boundary of the elastic zone. The softening zone is in a state of a high vertical compressive strain and a high lateral tensile strain.

The numerical results confirm that the present type of non-local continuum exhibits no spurious zero-energy periodic instability modes which were encountered for other non-local models.[6, 11] As for the efficiency of computation, it has been experienced that for the large meshes the non-local finite element runs were usually running faster and converged more rapidly than the corresponding local ones. Apparently the fact that the non-local operator stabilizes the response improves the convergence to such an extent that it more than offsets the additional time the computer needs to evaluate the averages of plastic strains over the finite element mesh.

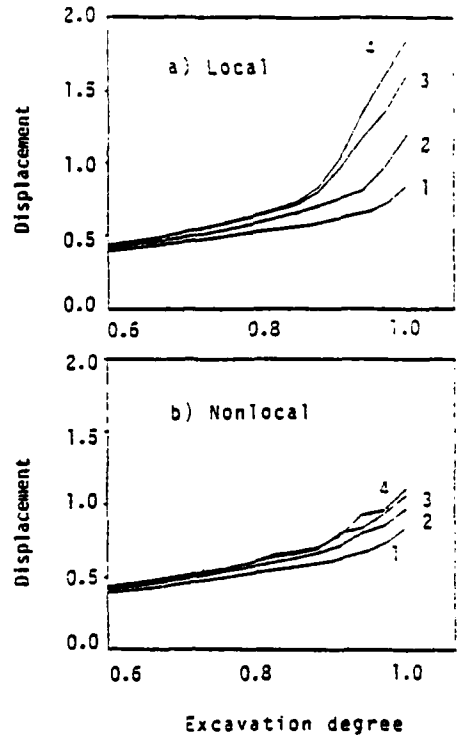


Figure 11. Response curves in the analysis of a tunnel excavation in a grouted soil with a degrading yield limit (displacement in m)

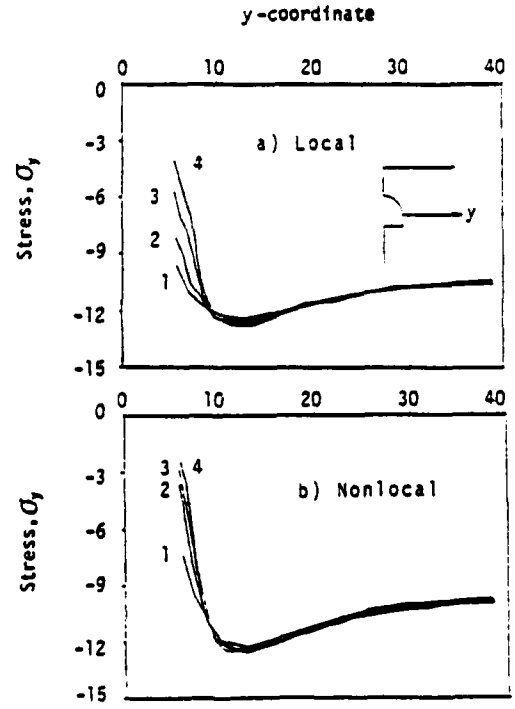


Figure 12. Stress profiles in the analysis of a tunnel excavation in a grouted soil with a degrading yield limit (stress in ton m^{-2} = 9807 Pa, y in m)

CONCLUSIONS

1. For strain-softening plastic materials with a degrading yield limit, localization of energy dissipation into a region of vanishing volume may be prevented by treating the plastic strains in the constitutive equation as non-local while the elastic strains remain local. This is achieved by processing the plastic strains through a spatial averaging operator, in which the characteristic length reflects the size of material inhomogeneities.

2. In contrast to the usual local approach, the finite element formulation based on the concept of non-local plastic strains exhibits correct and rapid convergence at mesh refinements and avoids spurious mesh sensitivity. The computer running times appear to be no longer than those for the local approach.

3. The new formulation has been verified for tensile as well as compressive strain softening due to yield limit degradation. It makes it possible to assess the safety against cave-in of tunnel excavation in softening soils.

ACKNOWLEDGEMENT

Partial financial support for the formulation of the basic concepts in this study was provided by U. S. Air Force Office of Scientific Research under Contract No. F49620–87–C–0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu. NSF provided funds for non-local finite element programming under Grant No. MSM–8700830 to Northwestern University, and for the use of the Cray X-MP/48 supercomputer under Grant No. 10131/MSM8700830.

APPENDIX—SOME DETAILS OF PLASTICITY MODEL

The meaning of tangential plastic modulus may be clarified by considering uniaxial loading with non-zero principal stress σ_1 and the corresponding strain ε_1. In this case **D** reduces to Young's modulus E and, from equation (11), $\Delta\lambda = \Delta\varepsilon_1 E/(E+H)$. The stress–strain relation becomes $\Delta\sigma_1 = E\Delta\varepsilon_1[1 - E/(E+H)]$, and the expression for H is

$$H = \frac{d\kappa}{d\varepsilon_1^p} = \frac{d\sigma_1}{d\varepsilon_1^p} = \frac{d\sigma_1}{d\varepsilon_1 - d\varepsilon_1^e} = \frac{1}{E_t^{-1} - E^{-1}} \tag{24}$$

Obviously, $H > 0$ signifies hardening ($E_t > 0$) and $H < 0$ softening ($E_t < 0$).

The envelope of Mohr circles for Mohr–Coulomb material is characterized by the equation: $\tau = c - \sigma \tan\phi$, where c = cohesion, ϕ = integral friction angle and σ, τ = normal and shear stress axes for Mohr's circle. This may be shown to be equivalent to the condition

$$\sigma_1 - \sigma_3 = 2c\cos\phi - (\sigma_1 + \sigma_3)\sin\phi \tag{25}$$

where σ_1, σ_3 = maximum and minimum principal stresses. The principal stresses are calculated as

$$\begin{Bmatrix}\sigma_1\\ \sigma_2\\ \sigma_3\end{Bmatrix} = \frac{I_1}{3}\begin{Bmatrix}1\\ 1\\ 1\end{Bmatrix} + 2\sqrt{\frac{J_2}{3}}\begin{Bmatrix}\sin(\theta+60°)\\ \sin\theta\\ \sin(\theta-60°)\end{Bmatrix} \tag{26}$$

in which $I_1 = \sigma_{kk}$, $J_2 = s_{ij}s_{ji}/2$, $J_3 = s_{ij}s_{jk}s_{ki}/3$, $\theta = \sin^{-1}(-1{\cdot}5\sqrt{3}\,J_3 J_2^{-3/2})/3$, $-30° \leqslant \theta \leqslant 30°$; s_{ij} = stress deviator. The yield surface thus becomes

$$F(\sigma, \kappa) = \frac{1}{3}I_1\sin\phi - \sqrt{J_2}\left(\cos\theta - \frac{1}{\sqrt{3}}\sin\theta\sin\phi\right) - \kappa \tag{27}$$

where $\kappa = c\cos\phi$. To model strain-softening, the cohesion c is gradually decreased.

The numerical values used in the tunnel analysis were $E = 11{\cdot}4$ ksi, $\nu = 0{\cdot}33$, $c = 5{\cdot}7$ psi, $\phi = 35^\circ$, $H = -57{\cdot}8$ psi.

The Tresca criterion is a special case of the Mohr–Coulomb criterion for $\phi = 0$, and the material parameters for Example I were $E = 3122$ ksi, $\nu = 0{\cdot}18$, $\sigma_y = 2c = 0{\cdot}4$ ksi, $H = -80$ ksi.

REFERENCES

1. E. C. Aifantis, 'On the microstructural origin of certain inelastic models', *J. Eng. Mat. Tech.*, No. 106, 326–330 (1984).
2. Z. P. Bažant, 'Instability, ductility and size effect in strain-softening concrete', *J. Eng. Mech. Div. ASCE*, **102**, 331–344; Discussions, **103**, 357–358, 775–777, **104**, 501–502 (1976) (based on *Struct. Eng. Report No. 74-8/640*, Northwestern University, Aug. 1974).
3. Z. P. Bažant, 'Crack band model for fracture of geomaterials', in Z. Eisenstein (ed.), *Proc. 4th Int. Conf. on Numerical Methods in Geomechanics, Vol. 3*. Edmonton, Canada, 1982, pp. 1137–1152.
4. Z. P. Bažant, 'Imbricate continuum and its variational derivation', *J. Eng. Mech. Div. ASCE*, **110**, 1693–1712 (1984).
5. Z. P. Bažant, 'Mechanics of distributed cracking', *Appl. Mech. Rev. ASME*, **39**, 675–705 (1985).
6. Z. P. Bažant, T. B. Belytschko and T.-P. Chang, 'Continuum theory for strain-softening', *J. Eng. Mech. Div. ASCE*, **110**, 1666–1692 (1984).
7. Z. P. Bažant and T. B. Belytschko, 'Wave propagation in a strain-softening bar: exact solution', *J. Eng. Mech. Div. ASCE*, **111**, 381–389 (1985).
8. Z. P. Bažant and T. B. Belytschko, 'Localization and size effect', in C. S. Desai *et al.* (eds.), *Int. Conf. on Constitutive Laws for Engineering Materials, Vol. 1*, Tucson, Ar., 1987, Preprints, pp. 11–33.
9. Z. P. Bažant and L. Cedolin, 'Blunt crack band propagation in finite element analysis', *J. Eng. Mech. Div. ASCE*, **105**, 297–315 (1979).
10. Z. P. Bažant and L. Cedolin, 'Fracture mechanics of reinforced concrete', *J. Eng. Mech. Div. ASCE*, **106**, 1287–1306 (1980); Discussion and closure, **108**, 464–471 (1982); see also Blunt crack band propagation in finite element analysis', *J. Eng. Mech. Div. ASCE*, **105**, 297–315 (1979).
11. Z. P. Bažant and T.-P. Chang, 'Instability of nonlocal continuum and strain-averaging', *J. Eng. Mech. Div. ASCE*, **110**, 1441–1450 (1984).
12. Z. P. Bažant and T.-P. Chang, 'Nonlocal finite element analysis of strain-softening solids', *J. Eng. Mech. Div. ASCE*, **113**, 84–105 (1987).
13. Z. P. Bažant, F.-B. Lin and G. Pijaudier-Cabot, 'Yield limit degradation: nonlocal continuum with local strains', in E. Onate *et al.* (eds.), *Int. Conf. on Computational Plasticity*, Barcelona, 1987, Preprints, pp. 1757–1780.
14. Z. P. Bažant and B. H. Oh (1983), 'Crack band theory for fracture of concrete', *Materials and Structures*, RILEM, Paris, France, Vol. 16, pp. 155–177.
15. Z. P. Bažant, J. Pan and G. Pijaudier-Cabot, 'Softening in reinforced concrete beams and frames', *Report No. 86-7/428s*, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL, 1986.
16. Z. P. Bažant and L. Panula, 'Statistical stability effects in concrete failure', *J. Eng. Mech. Div. ASCE*, **104**, 1195–1212 (1978).
17. Z. P. Bažant, G. Pijaudier-Cabot and J. Pan, 'Ductility, snapback size effect and redistribution in softening beams and frames', *J. Struct. Eng. ASCE*, **113**, 2348–2364 (1987).
18. Z. P. Bažant and G. Pijaudier-Cabot, 'Modeling of distributed damage by nonlocal continuum with local strain', in A. R. Luxmore *et al.* (eds.), *4th Int. Conf. on Numerical Methods in Fracture Mechanics*, San Antonio, Texas 1987, Preprints, pp. 411–432.
19. Z. P. Bažant and G. Pijaudier-Cabot, 'Nonlocal damage: continuum model, localization instability and convergence', *Report No 87-2/428n-1*, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL, 1987; also *J. Appl. Mech. ASME*, in press.
20. Z. P. Bažant and A. Zubelewicz, 'Strain-softening in bar and beam: exact nonlocal solution', Report, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL, 1986.
21. T. B. Belytschko, Z. P. Bažant, Y. W. Hyun and T.-P. Chang, 'Strain-softening materials and finite element solutions', *Comp. Struct.*, **23**, 163–180 (1986).
22. T. B. Beltyschko, X. J. Wang, Z. P. Bažant and H. Hyun, 'Transient solutions for one-dimensional problems with strain-softening', *J. Appl. Mech. ASME*, **54**, 513–518 (1987).
23. L. Cedolin and Z. P. Bažant, 'Effect of finite element choice in blunt crack band analysis', *Comp. Methods Appl. Mech. Eng.*, **24**, 305–316 (1980).
24. A. C. Eringen and N. Ari, 'Nonlocal stress field at Griffith crack', *Cryst. Latt. Amorph. Mater.*, **10**, 33–38 (1983).
25. A. C. Eringen and D. G. B. Edelen, 'On nonlocal elasticity', *Int. J. Eng. Sci.*, No. 10, 233–248 (1972).
26. H. Floegl and H. A. Mang, 'On tension stiffening in cracked reinforced concrete slabs and shells considering geometric and physical nonlinearity', *Ing. Arch.*, **51**, 215–242 (1981).
27. A. L. Gurson, 'Plastic flow and fracture behavior of ductile materials incorporating void nucleation, growth and interaction', *Ph.D. Thesis*, Brown University; see also *J. Eng. Mat. Tech.*, **99**, 2 (1975).
28. J. Hadamard, *Lecons sur la Propagation des Ondes*, Chapt. VI, H. Hermann et Cie Paris, France, 1903.

29. A. Hillerborg, (1984), 'Numerical methods to simulate softening and fracture of concrete', in G. C. Sih (ed.), *Fracture Mechanics Applied to Concrete Structures*, Martinus Nijhoff, The Hague.
30. A. Hillerborg, M. Modéer and Petersson, P. E. 'Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements', *Cement Concr. Res.*, No. 6(6), 773–782 (1976).
31. D. D. Joseph *et al.*, 'Hyperbolicity and change of type in the flow of viscoelastic fluids', *Arch. Rational Mech. Anal.* **87**, 213–242 (1986).
32. I. D. Karsan and J. O. Jirsa, 'Behavior of concrete under compression loadings', *J. Struct. Div. ASCE*, **95**, 2543–2563 (1969).
33. D. Krajcinovic, 'Constitutive equations for damaging materials, *J. Appl. Mech. ASME*, **50**, 355–360 (1983).
34. D. Krajcinovic and G. V. Fonseka, 'The continuous damage theory of brittle materials', *J. Appl. Mech. ASME*, **48**, 809–815 (1981).
35. E. Kröner, 'Elasticity theory of materials with long-range cohesive forces', *Int. J. Solids Struct.*, No. 3, 731–742 (1967).
36. J. A. Krumhansl, 'Some considerations of the relations between solid state physics and generalized continuum mechanics', in E. Kröner (ed.), *Mechanics of Generalized Continua*, Springer-Verlag, Berlin, 1968, pp. 298–331.
37. P. Ladevèze, 'On an anisotropic damage theory' (in French), *Internal Report No. 34*, Laboratoire de Mécanique et Technologie, Cachan, France, 1983.
38. F. A. Leckie, 'Constitutive equations of continuum creep damage mechanics', *Philos. Trans. Roy. Soc. London, Ser. A*, **288**, 27–47 (1978).
39. J. Lemaitre, 'How to use damage mechanics' (Division lecture presented at SMiRT7, Chicago, 1983), *Internal Report No. 40*, Laboratorire de Mècanique et Technologie, Ecole Normale Supérieure de l'Enseignement Technique, Paris-Cachan, France, 1983.
40. J. Lemaitre and J.-L. Chaboche, 'Aspect phenomenologique de la rupture par endommagement', *J. Mécan. Appl. (Paris)*, **2**, 317–365 (1978).
41. J. Lemaitre and J. L. Chaboche, *Mécanique des Matériaux Solides*, Dunod-Bordas, Paris, 1985.
42. R. L'Hermite and J. J. Grieu, 'Etude expérimentales récentes sur le retrait des ciments et des bétons', *Annal. I.T.T.P. (Paris)*, **5**, 494–514 (1952).
43. H. Mang and J. Eberhardsteiner, 'Collapse analysis for thin R. C. shells on the basis of a new fracture criterion', *U.S.–Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures*, Preprints 1985, pp. 217–238.
44. J. Mazars and G. Pijaudier-Cabot, 'Continuum damage theory: application to concrete', Internal report, Laboratoire de Mecanique et Technologie, Cachan, France, Dec. 1986.
45. A. Needleman and V. Tvergaard, *J. Mech. Phys. Solids*, **32**, 461 (1984).
46. D. R. J. Owen and E. Hinton, *Finite Elements in Plasticity: Theory and Practice*, Pineridge Press, Swansea, 1980.
47. S. T. Pietruszczak and Z. Mróz, 'Finite element analysis of deformation of strain-softening materials', *Int. j. numer. methods eng.*, **17**, 327–334 (1981).
48. G. Pijaudier-Cabot and Z. P. Bažant, 'Nonlocal damage theory', *Report No. 86-8/428n*, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL; 1986: also *J. Eng. Mech. ASCE*, **113**, 1512–1533 (1987).
49. G. Pijaudier-Cabot and Z. P. Bažant, 'Nonlocal damage: dynamic stability and convergence', *Report No 87-2/428n-II*, Center of Concrete and Geomaterials, Northwestern University, Evanston, IL, 1987; also *Comp. Struct.* (in press).
50. J. H. Prevost, 'Localization of deformation in elastic–plastic solids', *Int. j. numer. anal. methods geomech.*, **8**, 187–196 (1984).
51. H. E. Read and G. P. Hegemier, 'Strain-softening of rock, soil and concrete'—a review article, *Mech. Mat.*, **3**, 271–294 (1984).
52. J. R. Rice, 'The localization of plastic deformation', *Proc. 14th Int. Congr. of Theoretical and Applied Mechanics Vol. 1*, Delft, The Netherlands, 1976, pp. 207–220.
53. J. W. Rudnicki and J. R. Rice, 'Conditions for the localization of deformation in pressure-sensitive dilatant materials', *J. Mech. Phys. Solids*, **23**, 371–394 (1985).
54. I. S. Sandler, 'Strain-softening for static and dynamic problems', in K. J. Willam (ed.), *Proc. Symp. on Constitutive Equations: Micro, Macro and Computational Aspects*, ASME Winter Annual Meeting, New Orleans, ASME, New York, 1984, pp. 217–231.
55. H. L. Schreyer and Z. Chen, (1984), 'The effect of localization on the softening behavior of structural members', in K. J. Willam (ed.), *Proc. Symp. on Constitutive Equations: Micro, Macro, and Computational Aspects*, ASME Winter Annual Meeting, New Orleans, ASME, New York, 1984, pp. 193–203.
56. H. L. Schreyer and Z. Chen, 'One-dimensional softening with localization', *J. Appl. Mech. ASME*, **53**, 791–797 (1986).
57. K. J. Willam, B. Hurlbut and S. Sture, 'Experimental, constitutive and computational aspects of concrete failure', *U.S.–Japan Seminar on Finite Element Analysis of Reinforced Concretes*, Tokyo, 1985, Preprints, pp. 149–172.
58. Z. P. Bažant, 'Why continuum damage is nonlocal: justification by quasiperiodic microcrack array', *Mech. Res. Comm.*, **14**, 407–419 (1987).

APPENDIX C2

Nonlocal Smeared Cracking Model for Concrete Fracture

By Zdeněk P. Bažant,[1] Fellow ASCE, and Feng-Bao Lin,[2] Associate Member, ASCE

Abstract: The classical smeared cracking model widely used in finite-element analysis of concrete and rock cannot describe the size effect experimentally observed in brittle failures and exhibits spurious mesh sensitivity with incorrect convergence to zero energy dissipation at failure. The crack band model circumvents these deficiencies but has limitations with respect to mesh refinement, shear locking on zig-zag crack bands, and directional bias of the mesh. It is shown that all of these problems can be avoided by a nonlocal generalization, in which the damage that characterizes strain softening is considered to be a function of the spatial average of the positive part of the maximum principal strain. Two alternatives are presented. (1) Smeared cracking whose direction is fixed when cracks start to form; and (2) smeared cracking whose orientation rotates with the maximum principal strain. Furthermore, fracture tests on specimens of various sizes are analyzed by finite elements. It is shown that the model correctly reproduces the experimentally observed size effect and agrees with Bažant's size effect law. Orthogonal and slanted meshes are shown to yield approximately the same cracking zones and propagation directions. The model is easily programmed and computationally more efficient than the corresponding local version.

Introduction

The cracking in concrete structures has widely been modeled in finite-element analysis by adjustments of material stiffness. This approach, called smeared cracking, was proposed by Rashid (1968), refined by many others, and has provided good results in many practical applications (Bažant 1986; Be Borst 1984; De Borst and Nauta 1984, 1985; "Finite element analysis of reinforced concrete" 1982; Hand et al. 1973; Lin and Scordelis 1975; Suidan and Schnobrich 1973; Yuzugullu and Schnobrich 1973). Nevertheless, problems were also encountered. In 1974, it was shown by stability analysis of strain softening (Bažant 1976) and demonstrated numerically (Bažant and Cedolin 1979, 1980; Bažant and Oh 1983; Cedolin and Bažant 1980; Darwin 1985; De Borst 1984) that the classical smeared cracking model is unobjective, exhibiting spurious mesh sensitivity and convergence to an incorrect failure mode with zero energy dissipation. As a related deficiency, the numerical results obtained with geometrically similar meshes exhibit no size effect, while test results for brittle failures of concrete structures as well as fracture specimens show a pronounced size effect (Bažant 1984, 1986; Bažant and Pfeiffer 1987). As a remedy, Bažant proposed in 1974 a model (Bažant 1976), later called the crack band model (Bažant and Oh 1983, 1984), whose distinguishing characteristic is that no finite element is allowed to become smaller than a certain characteristic length w_c which is a material property and is related to the size of inhomogeneity in the material. Extended to de-

[1]Prof., Dept. of Civ. Engrg., Northwestern Univ., Evanston, Ill. 60208.
[2]Asst. Prof., Dept. of Civ. and Envir. Engrg., Polytechnic Univ., Brooklyn, NY 11201.

Note. Discussion open until April 1, 1989. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on September 2, 1987. This paper is part of the *Journal of Structural Engineering*, Vol. 114, No. 11, November, 1988. ©ASCE, ISSN 0733-9445/88/0011-2493/$1.00 + $.15 per page. Paper No. 22936.

scribe progressive cracking characterized by gradual strain softening (Bažant and Oh 1983), the crack band model was shown to agree with all of the basic experimental data from concrete or rock fracture testing.

An alternative is to model concrete fractures as interelement line cracks, as proposed by Ngo and Scordelis (1967). Adapting to concrete Dugdale and Barenblatt's ductile fracture model (Bažant 1986), Hillerborg et al. (Hillerborg 1985; Hillerborg et al. 1976) recognized that the line crack must be endowed with a softening cohesion zone near the crack tip, in order to obtain agreement with fracture test results. This approach was later found to be essentially equivalent to the crack band model, except for the case of many closely spaced parallel fractures. Although Ingraffea et al. (Ingraffea 1985; Ingraffea et al. 1985; Ngo and Scordelis 1967) succeeded in implementing line cracks (interelement cracks) in large finite-element programs, this approach seems generally more difficult to program than the smeared cracking. It might also be less realistic since no condition on minimum admissible crack spacing is used in this model but appears to be required for objectivity. Furthermore, it has not been demonstrated that remeshing with extreme mesh refinements at crack tip, as used in this approach, introduces no bias for the direction of crack progagation.

The existing crack band model nevertheless suffers with several disadvantages: (1) Refinement of the mesh in which the cracking zone width is subdivided into more than one element is not permitted, for reasons of objectivity; (2) in the case of zig-zag crack band propagation through the mesh, one needs to introduce some mathematical artifices to prevent spurious locking of the rugged opposite sides of the crack band (Droz 1987) (the same problem exists for shear bands); (3) the mesh, typically a square mesh, inevitably introduces a certain degree of directional bias for the crack propagation, favoring propagation along the mesh lines or along the diagonals of the mesh, and suppressing propagation directions of small inclination with regard to the mesh line (Marchertas et al. 1982; Pan et al. 1983; 1984); and (4) possible variations of the cracking zone width, which would cause variations in the specific energy required for propagation, cannot be taken into account (neither can this be done for the discrete crack models).

The foregoing disadvantages can be eliminated by adopting a nonlocal continuum approach (Eringen and Edelen 1972) whose application to strain softening was proposed by Bažant et al. (1984). The objective of the present study is to introduce a new, simple version of the nonlocal approach whose general idea was discovered recently (Bažant et al. 1987; Bažant and Pijaudier-Cabot 1987, 1988; Pijaudier-Cabot and Bažant 1986, 1987). The idea is that the nonlocal treatment should be applied only to those variables which cause strain softening, while the other variables, especially the elastic strain, should be local. In this form of nonlocal approach, the differential equations of equilibrium with the boundary conditions retain their standard form, no extra boundary conditions need to be introduced, the continuity requirements for finite elements remain the same as for the local formulation, and spurious zero-energy periodic instability modes, which caused problems for the original nonlocal formulation (Bažant et al. 1984), do not exist.

Alternative I. Cracking of Fixed Direction

In two dimensions, the crack band model may be characterized by the stress–strain relation:

$$\begin{Bmatrix}\epsilon_{11}\\ \epsilon_{22}\\ \gamma_{12}\end{Bmatrix}=\begin{bmatrix}C_{1111} & C_{1122} & 0\\ C_{2211} & C_{2222} & 0\\ 0 & 0 & C_{1212}/\beta\end{bmatrix}\begin{Bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{12}\end{Bmatrix}+\begin{Bmatrix}\epsilon'\\ 0\\ 0\end{Bmatrix} \quad \ldots\ldots (1)$$

in which ϵ' is the fracturing strain; ϵ_{11}, ϵ_{22}, and γ_{12} are the normal strains and the shear angle ($\gamma_{12} = 2\epsilon_{12}$); σ_{11}, σ_{22}, and σ_{12} are the normal and shear stresses; C_{1111}, C_{2222}, C_{1212}, and C_{1122} are the initial elastic compliances; and the numerical subscripts refer to the cartesian coordinates x_1 and x_2. Assuming isotropy, we have $C_{1111} = C_{2222} = 1/E'$, $C_{1122} = C_{2211} = \nu'/E'$, $C_{1212} = 2(1 + \nu)/E$, in which for plane stress $E' = E$, $\nu' = \nu$, and for plane strain $E' = E/(1 - \nu^2)$, $\nu' = \nu/(1 - \nu)$.

The cracks are assumed to be all parallel and normal to the axis x_1. As the first alternative of the smeared cracking model, we introduce the classical hypothesis that, at the time the cracks start to form, the direction of the cracks (i.e., of the axis x_1) is fixed as normal to the maximum principal stress σ_t at that instant and remains constant afterwards. If σ_t does not rotate, σ_{12} remains zero; however in general σ_t may rotate, and in that case one must take into account the shear stiffness due to aggregate interlock on rough crack surfaces. This is done by multiplying the shear compliance C_{1212} with an empirical coefficient β, called the shear retention factor, as proposed (for sudden cracking) by Schnobrich et al. (Hand et al. 1973; Suidan and Schnobrich 1973; Yuzugullu and Schnobrich 1973) and others (Lin and Scordelis 1975). More realistically, the coefficient β could be assumed to vary according to the rough crack model (Bažant and Gambarova 1981).

It has been shown (Bažant and Oh 1983) that Eq. 1 is equivalent to the equation:

$$\begin{Bmatrix}\epsilon_{11}\\ \epsilon_{22}\\ \gamma_{12}\end{Bmatrix}=\begin{bmatrix}C_{1111}/(1-\omega) & C_{1122} & 0\\ C_{2211} & C_{2222} & 0\\ 0 & 0 & C_{1212}/\beta\end{bmatrix}\begin{Bmatrix}\sigma_{11}\\ \sigma_{22}\\ \sigma_{12}\end{Bmatrix} \quad \ldots\ldots (2)$$

in which the normal stiffness C_{1111} in the direction normal to the cracks is reduced by the coefficient $1 - \omega$, where ω represents what is known in continuum damage mechanics as damage and may be regarded as the cracked area fraction; $\omega = 0$ corresponds to the initial state of no cracking (zero damage); and $\omega = 1$ corresponds to the final state of complete cracking (complete damage). Comparing Eqs. 1 and 2, we note that $\epsilon' = [\omega C_{1111}/(1 - \omega)]\sigma_{11}$.

Although in general the evolution of damage may be described in the form $\dot{\omega} = F(\omega, \epsilon)$, one may simply assume that, for monotonic loading, $(1 - \omega)^{-1} = \phi(\epsilon_{11})$ = function that characterizes the cracking in the material. Assuming unloading and reloading to follow the secant modulus (which is the basic hypothesis of continuum damage mechanics), we may write:

$$\text{If } \epsilon_{11} \geq \epsilon_{max} \text{ and } \Delta\epsilon_{11} \geq 0: \quad \frac{1}{1-\omega} = \phi(\epsilon_{11}); \quad \text{otherwise } \Delta\epsilon = 0 \ldots\ldots (3)$$

in which Δ refers to a small increment over the loading step; and ϵ_{max} denotes the maximum strain ϵ_{11} attained in all of the preceding steps. The assumption of secant unloading is of course a simplification, and various refinements

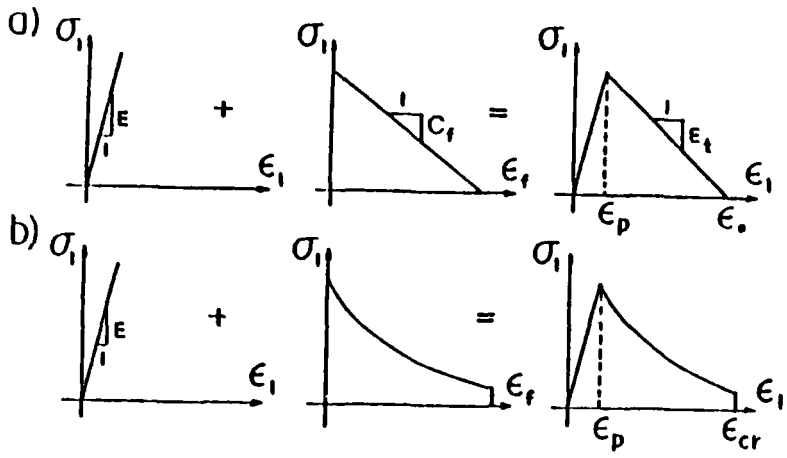


FIG. 1. *(a) Linear and (b) Exponential Strain Softening*

with more complex unloading rules may be introduced.

A widely used assumption has been a triangular stress–strain diagram for uniaxial loading [Fig. 1(*a*)]. In this case,

$$\begin{aligned} &\text{for } \epsilon_{11} \le \epsilon_p: && \phi(\epsilon_{11}) = 1 \\ &\text{for } \epsilon_p < \epsilon_{11} < \epsilon_0: && \phi(\epsilon_{11}) = \frac{E}{E_t}\frac{\epsilon_{11}}{\epsilon_0 - \epsilon_{11}} \qquad (4) \\ &\text{for } \epsilon_{11} \ge \epsilon_0: && \phi(\epsilon_{11}) = 0 \end{aligned}$$

in which ϵ_p is the strain at peak stress (cracking initiation); and ϵ_0 is the strain when the stress is reduced to zero (full cracking).

Recently various experimental evidence indicates that it is more realistic to assume a strain-softening curve with a steep initial decline followed by an extended tail. In that case, one may use

$$\begin{aligned} &\text{for } \epsilon_{11} \le \epsilon_p: && \phi(\epsilon_{11}) = 1 \\ &\text{for } \epsilon_p < \epsilon_{11} < \epsilon_{cr}: && \phi(\epsilon_{11}) = \frac{E}{f_t'}\epsilon_{11}e^{\alpha(\epsilon_{11}-\epsilon_p)} \qquad (5) \\ &\text{for } \epsilon_{11} \ge \epsilon_{cr}: && \phi(\epsilon_{11}) = 0 \end{aligned}$$

in which we introduce a certain critical (or maximum) strain ϵ_{cr} [Fig. 1(*b*)] that may not be exceeded in strain softening. Our introduction of ϵ_{cr} differs from the existing practice in which the strain-softening curve has been assumed to approach the strain axis in a continuous descent. Recent micromechanics analysis of crack ligament tearing (Bažant 1987), however, indicated that, for strain softening, there should exist a certain maximum strain after which the stress–strain diagram should exhibit a snapback, i.e., revert to a positive slope. The value of ϵ_{cr} has been considered to correspond to a point where $\sigma_{11} = 0.05\,\sigma_p$. This ϵ_{cr} value is so large that it has never been reached in the present computations.

Nonlocal Generalization with Local Elastic Strain

The nonlocal continuum is a continuum for which at least some variables are defined by spatial averaging. As discovered in Bažant and Pijaudier-Cabot (1987, 1988) and Pijaudier-Cabot and Bažant (1987), only those variables that cause strain softening should be considered as nonlocal. In the present context, this means that the damage ω should be made nonlocal. This is accomplished by specifying it as a function of the average (nonlocal) strain, which is defined by spatial averaging as follows:

$$\bar{\epsilon}_{11}(\mathbf{x}) = \frac{1}{V_r(\mathbf{x})}\int_V \alpha(\mathbf{s}-\mathbf{x})\langle\epsilon_{11}(\mathbf{s})\rangle dV = \int_V \alpha'(\mathbf{x},\mathbf{s})\langle\epsilon_{11}(\mathbf{s})\rangle dV \qquad (6)$$

in which

$$V_r(\mathbf{x}) = \int_V \alpha(\mathbf{s}-\mathbf{x})dV, \quad \alpha'(\mathbf{x},\mathbf{s}) = \frac{\alpha(\mathbf{s}-\mathbf{x})}{V_r(\mathbf{x})} \qquad (7)$$

The overbar is a label for the nonlocal variables, V = volume of the entire body, $\alpha(\mathbf{x})$ = given weighting function which is treated as a material property, and $\mathbf{s}$ = general coordinate vector [Fig. 2(*b*)]. V_r has approximately but not exactly the same meaning as the representative volume in the statistical theory of heterogeneous materials. The pointed brackets $\langle\ \rangle$ denote the positive part of the variable; i.e. $\langle\epsilon\rangle = \epsilon$ if $\epsilon > 0$ and 0 if $\epsilon \le 0$.

The weighting function could be defined as uniform ($\alpha = 1$) over volume V_r represented as a circle in two dimensions, and zero outside V_r. However, it has been experienced that the calculations converge better if the weighting function is smooth. A suitable choice is the normal (Gaussian) distribution function:

$$\alpha(\mathbf{x}) = e^{-(k|\mathbf{x}|/l)^2} \qquad (8)$$

in which k is a constant whose value is $k = 2$ for the case of two dimensions; $|\mathbf{x}|^2 = x^2 + y^2$; and l is the characteristic length of the material, which may be interpreted as the radius of a circle on which the uniform distribution $\alpha = 1$ has the same volume as the normal distribution function extending to infinity in the plane.

If the body is finite, the normal distribution function obviously extends beyond the boundary. This is handled by deleting the region outside the body

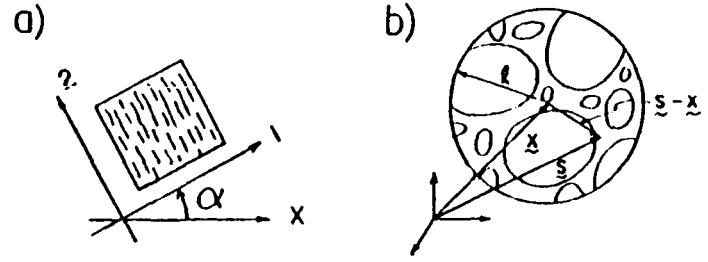


FIG. 2. *(a) Distributed Cracking of Single Orientation; and (b) Representative Volume of Material Compared to Aggregate Size*

from the integration domain V, both for the calculation of the average (Eq. 6) and for the calculation of the representative volume (Eq. 7). This fact causes V_r to depend on location $\mathbf{x}$.

In finite-element calculations, the integrals in Eqs. 6 and 7 are approximately evaluated as finite sums over all the integration points of all the elements in the structure. However, the integration points whose distances from the point $\mathbf{x}$ exceed approximately $2l$ may be omitted since for them the value of α is negligible.

In preliminary calculations, Eq. 6 did not involve the symbol $\langle\ \rangle$, i.e., the averaging was made directly over ω rather than over $\langle\epsilon_{11}\rangle$. However, this had to be abandoned since ω does not grow for $\epsilon_{11} > \epsilon_0$, which meant that the average of ω could never attain full damage ($\omega = 1$) unless the damage was the same in all of the elements. Therefore, the nonlocal damage $\bar{\omega}$ is considered to be a function of the nonlocal normal strain $\bar{\epsilon}_{11}$, i.e.,

$$(1 - \bar{\omega})^{-1} = \phi(\bar{\epsilon}_{11}) \quad \ldots\ldots (9)$$

The nonlocal constitutive equation with local strain takes the form:

$$\begin{Bmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \gamma_{12} \end{Bmatrix} = \begin{bmatrix} C_{1111}/(1-\bar{\omega}) & C_{1122} & 0 \\ C_{2211} & C_{2222} & 0 \\ 0 & 0 & C_{1211}/\beta \end{bmatrix} \begin{Bmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{12} \end{Bmatrix} \quad \ldots\ldots (10)$$

In matrix form, Eq. 10 may be written as $\sigma' = \mathbf{D}'\epsilon'$ in which $\mathbf{D}' = \mathbf{C}'^{-1}$ and $\mathbf{C}'$ = compliance matrix, referring to the local coordinate axes which are set to lie normal to the maximum principal strain direction when the strain softening begins. In finite-element programs, Eq. 10 must of course be transformed to the global coordinates, taking the form $\sigma = \mathbf{D}\epsilon$, in which $\mathbf{D} = \mathbf{T}^T\mathbf{D}'\mathbf{T}$ and $\mathbf{T}$ = rotation matrix.

In each loading step, the finite-element calculations utilize the direct stiffness iteration (secant method) and proceed according to the following algorithm:

1. First we assemble the global secant stiffness matrix $\mathbf{K}$ using, for every finite element, the nonlocal material compliance matrix $\mathbf{C}'$ (Eq. 10) which was obtained either in the preceding step (for the first iteration of this step) or in the previous iteration (*for the subsequent iterations*). *In the first loading step*, $\mathbf{C}'$ is the elastic compliance matrix of the material. Considering the load and displacement increments prescribed for this loading step, we solve the column matrix $\Delta\mathbf{u}$ of the nodal displacement increments and calculate $\mathbf{u}^F = \mathbf{u}^I + \Delta\mathbf{u}$ where F and I are the labels for the initial and final values in the loading step. Then we calculate ϵ^F for all of the integration points of all finite elements. For each integration point we check whether cracking (damage) has started. If not, we calculate for this point the local coordinate direction x_1' (angle α) as the direction of the maximum principal strain ϵ_{11}^F. Otherwise we keep the old direction x_1'. Then we transform ϵ^F into ϵ'^F for the local coordinate.

2. With the help of a spatial averaging subroutine which approximates the integrals in Eqs. 6 and 7 by finite sums over the integration points, we calculate the nonlocal principal strains $\bar{\epsilon}_{11}^F$ and evaluate $(1 - \bar{\omega})^{-1}$ for all of the integration points of all elements (Eq. 9). Then we update matrices $\mathbf{C}'$ and $\mathbf{D}'$ and calculate the final stresses, $\sigma'^F = \mathbf{D}'\epsilon'^F$. This yields for each integration point of each

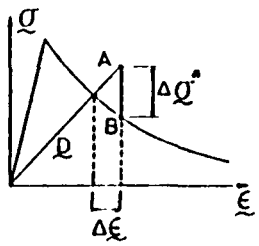


FIG. 3. Adjustment of Stress State (from point A to point B) to Stress–Strain Curve, as Implemented in Iteration of Loading Step

element the stress state corresponding to point B in Fig. 3 (in which point A shows the stress state based on the initial $\mathbf{C}'$ for this step). This stress state now satisfies the constitutive equations exactly but is not exactly in equilibrium with the loads because $\Delta\mathbf{u}$ was solved in step 1 on the basis of the old $\mathbf{C}'$ rather than the newly calculated $\mathbf{C}'$ (the initial equilibrium state is shown as point A in Fig. 3). Then we calculate the equivalent nodal forces $\mathbf{f}^E$ from σ'^F according to the principle of virtual work and evaluate the residual nodal forces $\mathbf{f}^R$ as $\mathbf{f}^R = \mathbf{f}^F - \mathbf{f}^E$ where $\mathbf{f}^F$ are the applied final nodal forces.

3. If the ratio $\Sigma|\mathbf{f}^R|/\Sigma|\mathbf{f}|$ is not less than a certain prescribed tolerance (for which the value 0.5% was used), then we return to step 1 and start the next iteration of this loading step. Otherwise we start at step 1 the first iteration of the next loading step.

Alternative II. Nonlocal Model with Rotating Smeared Cracks

The practice of keeping the crack direction constant and not allowing new cracks to form at oblique angles with this direction (as done in step 2 despite a possible rotation of the principal strain direction) is standard in the classical smeared cracking approach although not quite realistic. Recently, a better agreement with some measurements on cracked reinforced concrete panels (Gupta and Akbar 1984) has been obtained under the assumption that the cracks are always normal to the maximum principal strain, ϵ_1, and rotate as ϵ_1 rotates.

Although the hypothesis that the cracks rotate would be physically inadmissible if taken literally (Bažant 1983), it might nevertheless describe adequately a situation where the cracks of one direction close (and lock in shear) while cracks of another direction form. The main appeal of this hypothesis, however, is its simplicity.

A general smeared cracking model based on the assumption of rotating cracks has apparently not yet been formulated, and we will do so now, formulating it in a nonlocal form. We work now in Cartesian coordinates x_1, x_2 of fixed directions and denote as n_i the direction cosines of the current unit normal $\mathbf{n}$ to the crack planes. As already implied in Eq. 1, we consider the fracture strain (cracking strain) ϵ_{ij}^f to be a uniaxial strain. This means we assume the formation of cracks to affect neither the normal strains in the directions normal to $\mathbf{n}$ nor the shear strains on the crack plane. Therefore, the total stress–strain relation and the nonlocal cracking strain are:

$$\epsilon_{ij} = C_{ijkm}\sigma_{km} + \bar{\epsilon}^f_{ij}; \qquad \bar{\epsilon}^f_{ij} = n_i n_j \bar{\epsilon}^f \quad (11)$$

where $\bar{\epsilon}^f$ = nonlocal fracture strain magnitude and C_{ijkm} = initial elastic compliance tensor, which is isotropic and for two dimensions is the same as defined in Eq. 1 (with $\beta = 1$). Generally, for three dimensions, $C_{ijkm} = [(1 + \nu)(\delta_{ik}\delta_{jm} + \delta_{im}\delta_{jk}) - 2\nu\delta_{ij}\delta_{km}]/2E$.

For nonrotating cracks, Eq. 11 must coincide with Eq. 10. Therefore, if we set $n = (1,0,0)$, we must have:

$$\frac{C_{1111}}{1 - \tilde{\omega}}\sigma_{11} = C_{1111}\sigma_{11} + \bar{\epsilon}^f \quad (12)$$

From this we obtain $\bar{\epsilon}^f = [\tilde{\omega}/(1 - \tilde{\omega})]C_{1111}\sigma_{11}$ where $C_{1111} = 1/E'$; $E' = E$ for plane stress as well as for three dimensions. Noting that $\sigma_{11} = n_i n_i \sigma_{11}$, we obtain for a general crack direction the generalization:

$$\bar{\epsilon}^f = \frac{\tilde{\omega}}{(1 - \tilde{\omega})E'} n_k n_m \sigma_{km} \quad (13)$$

Thus the total stress–strain relation for Alternative II is:

$$\epsilon_{ij} = \left[C_{ijkm} + \frac{\tilde{\omega}}{(1 - \tilde{\omega})E'} n_i n_j n_k n_m\right]\sigma_{km} \quad (14)$$

as recently proposed by Bažant. The bracketed term represents the secant compliance tensor. Note that, for loading, $\tilde{\omega}/(1 - \tilde{\omega}) = \phi(\bar{\epsilon}_t)$, because of Eq. 9; $\bar{\epsilon}_t$ is the nonlocal maximum principal strain; $\epsilon_t = n_i n_j \epsilon_{ij}$. Here we imply again the assumption that the orientation of cracks is decided by the principal direction n_i of nonlocal strain $\bar{\epsilon}_{ij}$ rather than the local strain ϵ_{ij}. The vector n in general rotates and must be evaluated at each stage of computation for each integration point of each finite element.

The computational algorithm is analogous to that we described for Alternative I but simpler in that one need not keep track of the coordinate axes for cracking at each integration point and need not carry out the corresponding rotation transformations. The finite-element results for the example that follows happen to be almost exactly the same because the principal stress directions remain nearly constant in this fracture test.

The decision in favor of Alternative I or II will necessitate further testing, since at present there apparently exist no test data for strain softening (progressive fracture) of plain concrete at rotating principal stress directions. A fully realistic model no doubt calls for allowing partial or full cracking in all directions, as is done in the microplane model (Bažant and Prat 1988).

It should be pointed out that function ϕ might depend separately on the average, ϵ_L, of the lateral normal strains in the directions normal to n; $\epsilon_L = (\epsilon_{ll} + \epsilon_{mm})/2 = (\epsilon_{kk} - n_i n_j \epsilon_{ij})/2$. The reason is that compressive lateral stresses alone can produce axial splitting cracks, as known for example from the uniaxial compression test. However, they also produce tensile strain ϵ_t, and thus it might be that the dependence of ϕ on $\bar{\epsilon}_t$ approximately accounts for the compression splitting cracks as well. Further study and testing is needed.

Fracture Specimens and Size Effect Due to Nonlocal Aspect

Without a nonlocal formulation, the finite-element codes with strain softening cannot describe the results of fracture tests. In particular, such finite-

element codes yield no size effect. Yet the size effect is the principal, experimentally verified feature of fracture mechanics.

To demonstrate that the nonlocal smeared crack model correctly describes fracture and especially the size effect, we analyze the three-point-bend specimens shown in Fig. 4. Three different specimen sizes of size ratios 1:2:4, with geometrically similar shapes, were considered. Their finite element meshes are shown in Fig. 5 and their refinements in Fig. 6. The problem was solved both for linear strain softening and exponential strain softening, shown in Fig. 7.

This type of specimen was tested at Northwestern University (Bažant and Pfeiffer 1987), using both concrete with maximum aggregate size d_a = 0.5 in. (12.7 mm) and mortar with d_a = 0.19 in. (4.83 mm). The maximum loads P were measured for each specimen of each size, from which the values of the nominal stress at failure σ_N were calculated; $\sigma_N = P/bd$ where b = specimen thickness (1.5 in. or 38.1 mm). The test results are shown in

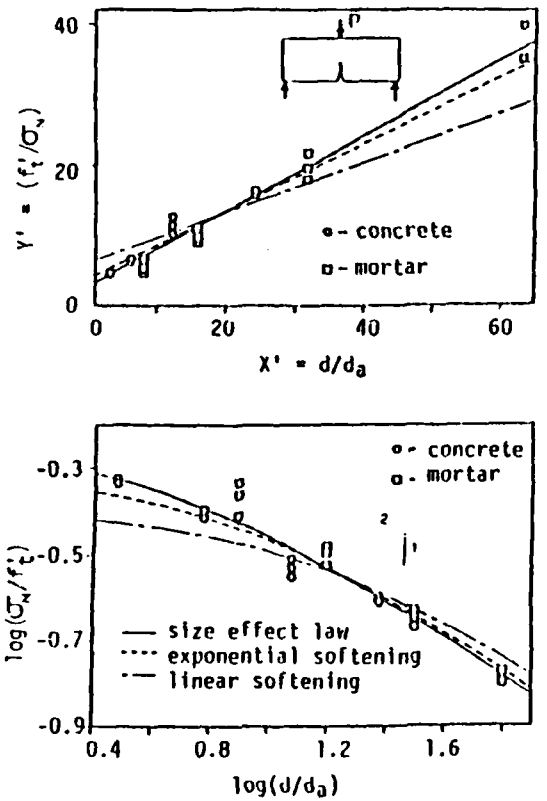


FIG. 4. Test Results of Bažant and Pfeiffer (1987) in Size Effect Plot (bottom) and Linear Regression Plot (top) Compared to Size Effect Law as Well as to Finite-Element Results for Linear and Exponential Strain Softening (Three-Point-Bent Specimen)

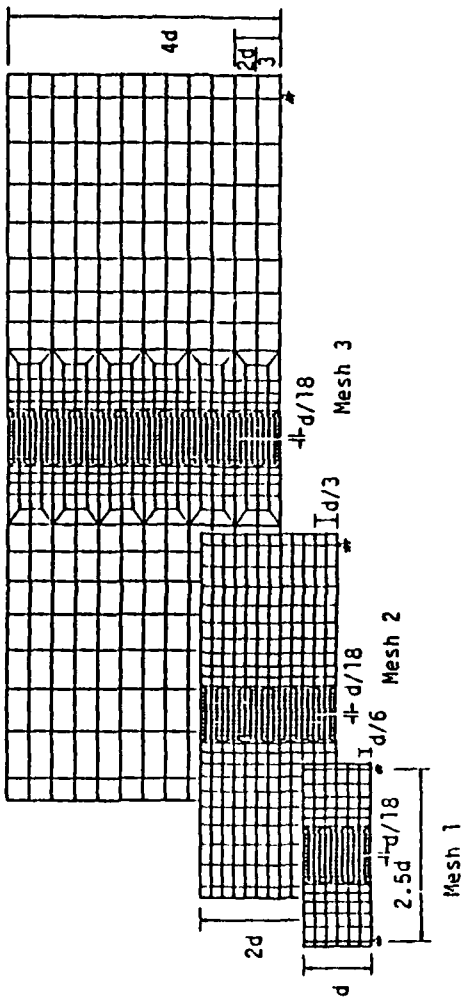


FIG. 5. Finite-element Meshes Used for Nonlocal Analysis of Specimens of Three Sizes

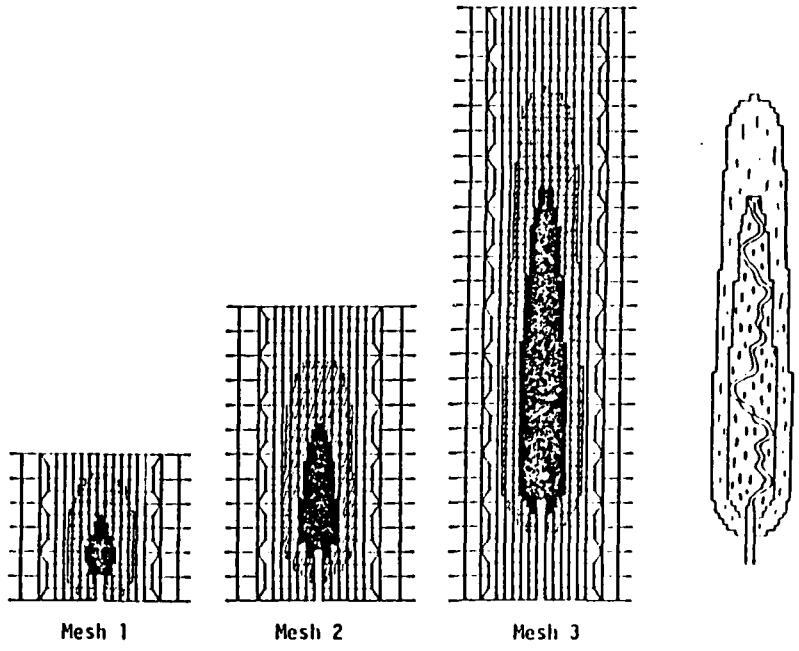


FIG. 6. Extents of Softening Zone (cross-hatched) and Fracture Zone (black) Obtained by Nonlocal Finite-Element Analysis

Fig. 4 in the size effect plots (bottom) and the corresponding linear regression plots (top), where f'_t = direct tensile strength of concrete. The solid lines in this figure represent the optimum fit of these test results with Bažant's size effect law for blunt fracture (Bažant 1984; Ba'ant and Pfeiffer 1987):

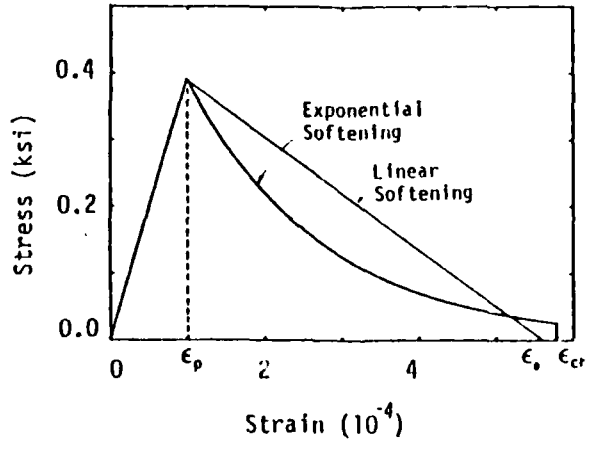


FIG. 7. Stress–Strain Diagram Used in Nonlocal Finite-Element Analysis

$$\sigma_N = Bf_t' \left[1 + \left(\frac{d}{\lambda_0 d_a}\right)^r\right]^{-1/2r} \quad \ldots\ldots (15)$$

in which the coefficient r may be approximately taken as 1, and B and λ_0 are two empirical constants which can be determined from the test data in Fig. 4. To this end, Eq. 15 may be algebraically rearranged to the linear regression equation $Y = AX + C$, in which $X = (d/d_a)^r$, $Y = (f_t'/\sigma_N)^{2r}$, $C = B^{2r}$, and $A = C\lambda_0^r$. The linear regression plots, shown for the measured data in Fig. 4 [taken from Bažant and Pfeiffer (1987)], yield the coefficients B and λ_0 from the slope A and the Y-axis intercept C. Furthermore, E = 3976 ksi (27420 MPa), ν = 0.18, and f_t' = 390 psi (2.69 MPa). The remaining material parameters were determined so as to obtain with the finite-element program and the meshes in Figs. 5 and 6 optimum fits of the test data in Fig. 4. These values were, for the linear strain softening, E_t = −726 ksi (−5007 MPa), and l = 2.3 in. (58.42 mm), and for the exponential strain softening, c = 690 psi (4.76 MPa), a = 5730 psi (39.52 MPa), l = 3.2 in. (81.28 mm). The corresponding optimum values of the fracture energy were G_f = 0.230 lb/in. (40.29 N/m) for the linear strain softening, and 0.249 lb/in. (43.62 N/m) for the exponential strain softening.

The fracture energy is equal to the area under the stress–strain diagram in Fig. 7 times the effective width of the fracture process zone, w_c. The value of w_c for the classical (local) crack band model is defined clearly; it coincides with the width of the crack band or the size of the finite element. But w_c is harder to define for the nonlocal formulation since the effective width of the fracture process zone varies (Fig. 6), depending on the influence of the boundaries and other factors. The actual width of the fracture process zone is roughly equal to 2.7 l (Bažant and Pijaudier-Cabot 1988), but since the density of cracking decreases from the middle of the crack band to its boundary, the effective width of the fracture process zone based on uniform distribution of the energy dissipation (or of cracking density) is about l. The aforementioned values of G_f correspond to the area under the stress–strain diagram times l.

The most important test to determine whether the finite-element model correctly describes fracture is the size effect. For structures that follow the strength criterion (or yield criterion), there is no size effect, i.e., the plot in Fig. 4 is represented by a horizontal line. The linear elastic fracture mechanics yields the strongest possible size effect which corresponds to the straight inclined line of slope −1/2 shown in Fig. 4 (bottom). For materials that exhibit distributed cracking, the size effect plot represents a smooth transition from the horizontal line for the strength criterion to the inclined straight line for the linear elastic fracture mechanics, as shown in Fig. 4 (bottom).

The load-displacement curves calculated by the nonlocal finite-element program for the three meshes shown in Figs. 5 and 6 are plotted in Fig. 8. The maximum load values, P_{max}, were fitted by the size effect law (Eq. 15), for which A = 0.384, C = 6.169 for the linear strain softening, and A = 0.483, C = 3.971 for the exponential strain softening. The agreement with the size effect law was rather accurate, with the error (deviation from Eq. 15) not exceeding 5%. The curves of the size effect law (Eq. 15), which were obtained by the optimum fits of the values of P_{max} calculated by finite elements for the three specimen sizes, are shown in Fig. 4 as the dashed

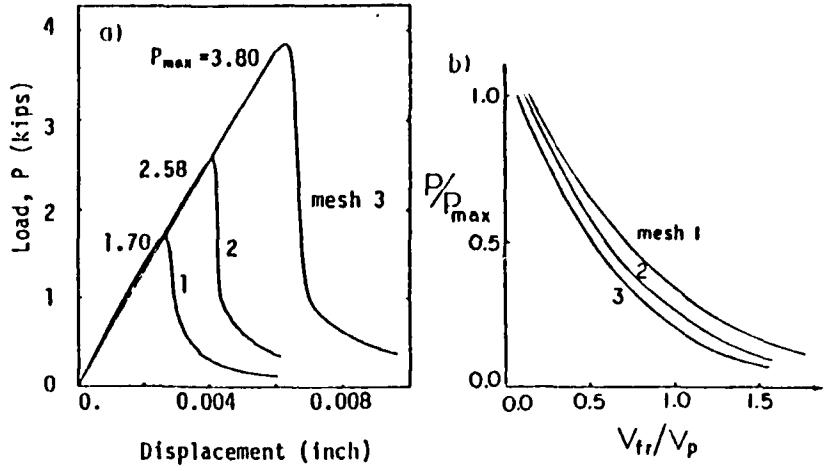


FIG. 8. (*a*) Diagrams of Load Versus Load-Point Displacement for Three-Point-Bend Specimens of Three Sizes Obtained with Meshes in Figs. 5 and 6; (*b*) Corresponding Diagrams of Load Versus Load-Point Fracture Displacement

lines for the exponential strain softening and the dash–dot lines for the linear strain softening.

An interesting feature is now noticed from Fig. 4: the shape of the strain-softening diagram significantly influences the shape of the size effect curve. This fact has already been noticed in other studies for different reasons. *Evidently, the exponential strain softening agrees with the test results much* better than the linear strain softening, which was originally used in the crack band model (Bažant and Oh).

Another interesting plot is that of load P versus the part, v_{fr}, of the load-point displacement v that is due to fracture and is defined as $v_{fr} = v - C_e P$, in which C_e is the elastic compliance of the specimens associated with load P. The plots of P versus displacement for the three specimen sizes and three finite-element meshes are shown in Fig. 8(*a*). We see that, for increasing specimen size, the initial postpeak drop of the load gets larger and larger (v_p = load-point displacement at maximum load). This indicates an increasing brittleness of the specimen as its size increases, in accordance with the size effect law.

If the presently considered specimens of three different sizes are analyzed with geometrically similar meshes using the classical (local) smeared cracking model based on the strength criterion, the resulting maximum load values show no size effect, i.e. the plot in Fig. 4 is horizontal. Thus, the classical smeared cracking is incapable of describing the size effect. This is obviously unsatisfactory, in contradiction to fracture tests as well as tests of various brittle failures of concrete structures. On the other hand, if these three specimen sizes are analyzed by a finite-element code for linear elastic fracture mechanics, the results match the straight line of slope −1/2 shown in Fig. 4 (bottom), for which the size effect is too strong.

Fig. 6 shows for the three specimen sizes and the corresponding meshes the extent of the strain-softening zone (cross-hatched area) and of the fully

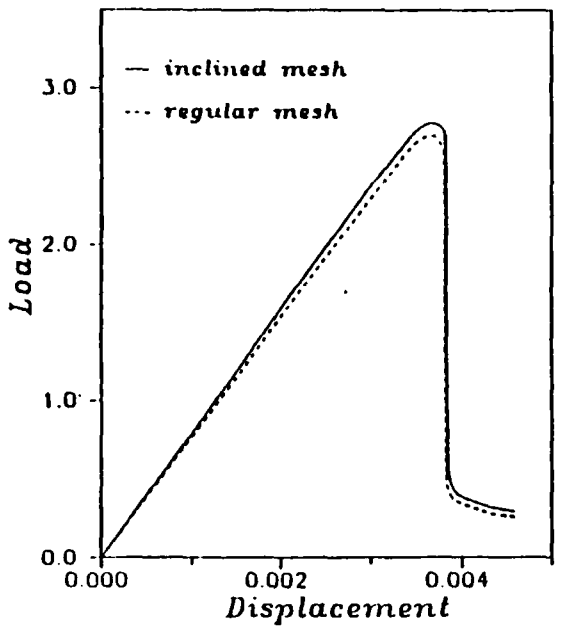


FIG. 9. Load Versus Load-Point Displacement Diagram Obtained for Three-Point-Bent Specimen for Orthogonal and Slanted Meshes Shown in Fig. 10.

fractured zone (black area). The former zone is defined as the locus of the material states for which the peak-stress point has been passed (i.e., $\epsilon_{11} > \epsilon_p$), and the latter zone is defined as the zone for which the stress σ_{11} has been reduced to less than $0.06\sigma_p$.

It should be kept in mind that normally the final fracture surface is not plane but highly tortuous. As an approximation, it may be assumed that the final crack surface randomly meanders through the black zone, oscillating on its way between the zone boundaries (Fig. 6, right). The resolution that can be obtained with the nonlocal model for element sizes that are much smaller than the aggregate size is of course meaningful only in a statistical sense since on such a scale the material behavior is highly random. The boundaries of the zones in Fig. 6, as well as the stress and strain values obtained for the fine meshes shown, must be regarded as merely statistical averages which correspond to smoothing of the random fields of microstresses and microstrains, and to averaging over a large number of tests of statistically identical specimens. The black zone corresponds to dense microcracking (normally invisible to the eye) and the strain-softening zone to sparse microcracking (Fig. 6, right).

The results shown in Fig. 6 for the three meshes have another interesting implication: the width of the fracture process zone as well as the softening zone is not constant but varies. In consequence, the rate of energy dissipation per unit fracture extension is variable, too. To capture such variation might be one important purpose of refining the mesh, possibly even below the aggregate size.

As a further implication, the present nonlocal model does not correspond to a unique stress-displacement diagram if the strain across the softening band is lumped into a crack opening displacement, equivalent to that used in models of Hillerborg's type. Rather, the equivalent discrete crack model would have to use somewhat different stress-displacement softening diagrams depending on the distance of the fracture front from the notch and from the specimen boundaries.

The computer running times for the nonlocal analyses are usually shorter than for the corresponding local analyses. Apparently, the additional computer time needed to calculate the spatial averages over the mesh is more than offset by improvement of the convergence of iterations, due to the stabilizing influence of the nonlocal formulation.

For other types of nonlocal continuum with local elastic strain, convergence at mesh refinement has been demonstrated before (Bažant et al. 1987; Bažant and Pijaudier-Cabot 1987, 1988; Pijaudier-Cabot and Bažant 1987). The fact that, despite mesh refinement, the energy dissipation cannot localize in these nonlocal models into a zone of zero volume has also been mathematically proven in general (Bažant and Pijaudier-Cabot 1987, 1988).

Elimination of Directional Bias of Mesh

With the crack band model in which the strain-softening band is one element wide, it has been difficult to simulate fractures that propagate at a small angle ϕ with regard to the mesh line, although some artificial measures (Pan

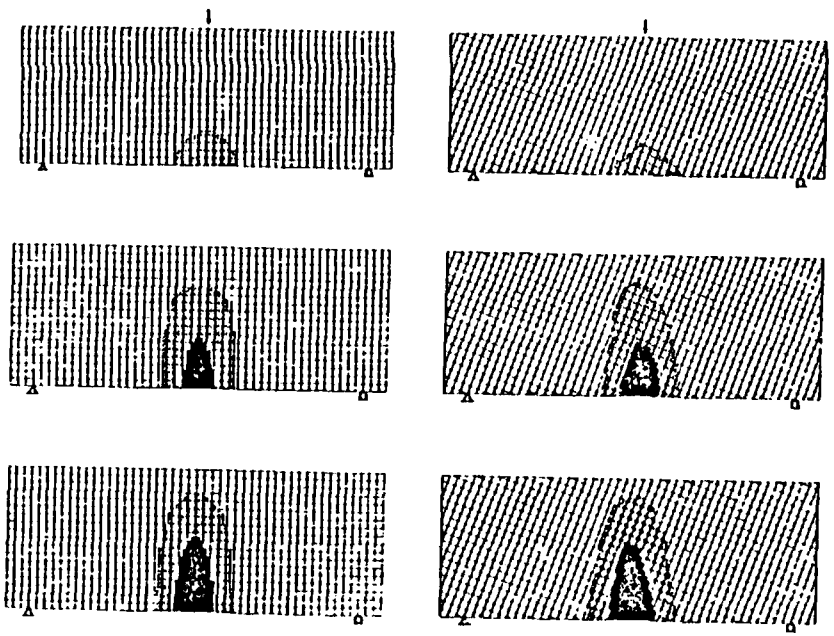


FIG. 10. Orthogonal Mesh and Slanted Mesh Used in Nonlocal Finite-Element Analyses, with Softening and Fracture Zones Obtained in Both Cases

et al. 1984) were found to offer partial help. To study this aspect, the presently considered specimen of the smallest size has been analyzed by finite elements using the slanted square mesh shown in Fig. 10 (right), whose inclination is 20°. The softening and fracture zones obtained are shown in Fig. 10 on the left for the orthogonal mesh and on the right for the slanted mesh. Success is now noticeable. The zone boundaries obtained for the orthogonal and slanted meshes are nearly the same, and the propagation for the slanted mesh still occurs in the correct vertical direction.

The problem of simulating propagation of crack bands in directions inclined with regard to the mesh line has recently been studied by Droz (1987) and for the analogous problem of shear bands by Leroy and Ortiz (1988). (The classical approach, in which the smeared cracking or the shear band tends to localize into a band of single-element width, fails since a localization limiter such as the nonlocal continuum concept cannot be used.) These investigators achieved propagation in oblique directions by superimposing on a continuous displacement field of the finite elements a certain additional discontinuous field which simulates the discontinuity of the zig-zag crack or shear band running in an inclined direction. However, the programming of this approach for general finite-element programs seems to be quite difficult. By contrast, the present approach is quite simple, although it does require mesh refinements such that there are at least three finite elements across the width of the cracking band. The present approach could also be adopted for the modeling of inclined shear bands.

Conclusions

1. The well-known inobjectivity, spurious mesh sensitivity, and incorrect convergence of the classical smeared cracking model can be eliminated by adopting a nonlocal formulation in which cracking (or strain softening) is characterized by nonlocal damage. The nonlocal damage may be considered to be a function of the spatial average of the positive part of the maximum principal strain. The nonsoftening (elastic) response is treated as local.

2. The finite-element results closely agree with the results of fracture tests, in particular with the observed influence of specimen size on the maximum load, which is described by Bažant's size effect law.

3. The nonlocal approach eliminates the directional bias of the mesh, provided the mesh is sufficiently refined. In particular, the results obtained with orthogonal and slanted meshes give approximately the same direction of fracture propagation and the same boundaries of the strain-softening zone and the fracture zone.

4. The nonlocal model is numerically efficient and easily programmed. It merely requires introducing into an existing code a subroutine for the spatial averaging of the maximum principal strains.

5. The nonlocal smeared model is formulated in two alternatives: the classical one in which the crack direction is fixed when the cracking starts, and a novel one in which the cracks are imagined to rotate along with the maximum principal strain direction. The latter alternative is simpler.

Acknowledgment

Partial financial support for the general nonlocal concept has been obtained under AFOSR contract No. F49620-87-C-0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu, and for the detailed formulation and finite-element programming under NSF Grant MSM-9877830 to Northwestern University. Supercomputer funds (on CRAY X-MP/48) were provided under NSF Grant 00000-MSM8211642.

Appendix. References

Bažant, Z. P. (1976). "Instability, ductility and size effect in strain-softening concrete." *J. Engrg. Mech.*, ASCE, 102(2), 331–344. Discussions, 103, 357–358, 775–777; 104, 501–502 (based on *Struct. Engrg. Report No. 74-8/640*, Northwestern University, August 1974).

Bažant, Z. P. (1983). "Comment on orthotropic models for concrete and geomaterials." *J. Engrg. Mech.*, ASCE, 109(3), 849–865.

Bažant, Z. P. (1984). "Size effect in blunt fracture: Concrete, Rock, Metal." *J. Engrg. Mech.*, ASCE, 110(4), 518–535.

Bažant, Z. P. (1986). "Mechanics of distributed cracking." *Appl. Mech. Rev.*, ASME, 39(5), 675–705.

Bažant, Z. P. (1987). "Snapback instability at crack ligament tearing and its implication for fracture micromechanics." *Cement and Concrete Res.*, 17(6), 951–967.

Bažant, Z. P., Belytschko, T. B., and Chang, T. P. (1984). "Continuum theory for strain-softening." *J. Engrg. Mech.*, ASCE, 110(12), 1666–1692.

Bažant, Z. P., and Cedolin, L. (1979). "Blunt crack band propagation in finite element analysis." *J. Engrg. Mech.*, ASCE, 105(2), 297–315.

Bažant, Z. P., and Cedolin, L. (1980). "Fracture mechanics of reinforced concrete." *J. Engrg. Mech.*, ASCE, 106(6), 1287–1306. Discussion and Closure (1982), 108, 464–471.

Bažant, Z. P., and Gambarova, P. (1981). "Rough cracks in reinforced concrete." *J. Struct. Engrg.*, ASCE, 106(4), 819–842. Discussions and Closure, 2579–2581, 1377–1388.

Bažant, Z. P., Lin, F. B., and Pijaudier-Cabot G. (1987). "Yield limit degradation: Nonlocal continuum model with local strain." Preprints, *Int. Conf. On Computational Plasticity*, held in Barcelona, D. R. J. Owen, E. Hinton, and E. Onate, eds. Univ. of Wales, Swansea, 1757–1780.

Bažant, Z. P., and Oh, B. H. (1983). "Crack band theory for fracture of concrete." *Materials and Structures*, RILEM, Paris, France, 16, 155–177.

Bažant, Z. P., and Oh, B. H. (1984). "Rock fracture via strain-softening finite elements." *J. Engrg. Mech.*, ASCE, 110(7), 1015–1035.

Bažant, Z. P., and Pfeiffer, P. A. (1987). "Determination of fracture energy from size effect and brittleness number." *ACI Mat. J.* 84 (Nov.–Dec.), 463–480.

Bažant, Z. P., and Pijaudier-Cabot, G. (1987). "Modeling of distributed damage by nonlocal continuum with local strain." Preprints. *4th Intern. Conf. on Numerical Methods in Fracture Mechanics*, A. R. Luxmore, D. R. J. Owen, and M. F. Kanninen, eds., held in San Antonio, Texas, Mar., 411–432.

Bažant, Z. P., and Pijaudier-Cabot, G. (1988). "Nonlocal continuum damage, localization instability and convergence." *J. Appl. Mech.*, ASME, 55(June), 287–293.

Bažant, Z. P., and Prat, P. C. (1988). "Microplane model for brittle-plastic material—I. Theory, II. Verification." *J. Engrg. Mech.*, ASCE, 114(10), 1672–1702.

Cedolin, L., and Bažant, Z. P. (1980). "Effects of finite element choice in blunt crack band analysis." *Computer Methods in Appl. Mech. and Engrg.* 24(3), 305–316.

Darwin, D. (1985). "Concrete crack propagation—Study of model parameters." *Proc. Finite Element Analysis of Reinforced Concrete Structures*. C. Meyer and H. Okamura, eds. ASCE, New York, 184–203.

De Borst, R. (1984). "Application of advanced solution techniques to concrete cracking and non-associated plasticity." *Numerical methods for non linear problems*, C. Taylor et al., eds. Vol. 2, Pineridge Press, Swansea, United Kingdom, 314–325.

De Borst, R, and Nauta, P. (1984). "Smeared crack analysis of reinforced concrete beams and slabs failing in shear." *Proc. Int. Conf. on Computer-aided Analysis and Design of Concrete Structures*, Split, Yugoslavia, 261–274.
De Borst, R., and Nauta, P. (1985). "Non-orthogonal cracks in a smeared finite element model." *Eng. Computation*, 2, 35–46.
Droz, P. (1987). "Numerical modeling of the non-linear behavior of massive structures of unreinforced concrete," doctoral thesis No. 682, Ecole Polytéchnique Fédérale de Lausanne, Switzerland.
Eringen, A. C., and Edelen, D. G. B. (1972). "On nonlocal elasticity." *Int. J. Eng. Sci.*, No. 10, 233–248.
"Finite element analysis of reinforced concrete." (1982). *State of the Art Report*, prepared by Task Committees, chaired by A. H. Nilson, ASCE, New York, Chapter 4, 204–233.
Gupta, A. K. and Akbar, H. (1984). "Cracking in reinforced concrete analysis." *J. Struct. Engrg.*, ASCE, 110(8), 1735.
Hand, F. R., Pecknold, D. A., and Schnobrich, W. C. (1973). "Nonlinear layered analysis of RC plates and shells, *J. Struct. Engrg.*, ASCE, 99(7), 1491–1505.
Hillerborg, A. (1985). "Numerical methods to simulate softening and fracture of concrete." *Fracture mechanics of concrete: Structural application and numerical calculation*, G. C. Sih and A. DiTomasso, eds. Martinus Nijhoff, Doordrecht–Boston, 141–170.
Hillerborg, A., Modéer, M., and Petersson, P. E. (1976). "Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements." *Cement and Concrete Res.* 6(6), 773–782.
Ingraffea, A. (1985). "Fracture propagation in rock." *Mechanics of geomaterials: Rocks, concretes, soils*, Z. P. Bažant, ed. John Wiley and Sons, Inc., New York, N.Y., 219–258.
Ingraffea, A. R., and Saouma, V. (1985). "Numerical modeling of discrete crack propagation in reinforced and plain concrete." *Fracture mechanics of concrete: Structural application and numerical calculation*, G. C. Sih and A. DiTomasso, eds. Martinus Nijhoff, Doordrecht–Boston, 171–225.
Leroy, Y., and Ortiz, M. (1988). "Finite element analysis of strain localization in frictional materials." Report, Div. of Engrg., Brown Univ., Providence, R.I.
Lin, C.-S., and Scordelis A. (1975). "Nonlinear analysis of RC shells of general form." *J. Struct. Engrg.*, ASCE, 101(3), 523–538.
Marchertas, A. H., Kulak, R. F., and Pan, Y. C. (1982). "Performance of blunt crack approach within a general purpose code." *Nonlinear numerical analysis of reinforced concrete*, L. E. Schwer, ed. ASME, New York; N.Y., 107–123.
Ngo, D., and Scordelis A. C. (1967). "Finite element analysis of reinforced concrete beams." *ACI J.* 64(3), 152–163.
Pan, Y. C., Marchertas, A. H., and Kennedy, J. M. (1983). "Finite element analysis of blunt crack propagation, a modified J-integral approach." *Trans. 7th Int. Conf. on Structural Mechanics in Reactor Technology*, Paper H5/3, North-Holland, Amsterdam, 235–292.
Pan, Y. C., Marchertas, A. H., Pfeiffer, P. A., and Kennedy, J. M. (1984). "Concrete cracking simulations for nuclear applications." *Theoret. Appl. Fracture Mech.*, 2(1), 27–38.
Pijaudier-Cabot, G., and Bažant, Z. P. (1986). "Nonlocal damage theory." *Report No. 86-8/428n*, Center for Concrete and Geomaterials, Northwestern Univ., Evanston, Ill.
Pijaudier-Cabot, G., and Bažant, Z. P. (1987). "Nonlocal damage theory." *J. Engrg. Mech.*, ASCE, 113(10). 1512–1533.
Rashid, Y. R. (1968). "Analysis of prestressed concrete pressure vessels." *Nuclear Engrg. and Design* 7(4), 334–355.
Suidan, M., and Schnobrich, W. C. (1973). "Finite element analysis of reinforced concrete." *J. Struct. Engrg.*, ASCE, 99(10), 2109–2122.
Yuzugullu, O., and Schnobrich, W. C. (1973). "A numerical procedure for the determination of the behavior of a shear wall frame system." *ACI. J.* 70(7), 474–479.

APPENDIX C3

NONLOCAL DAMAGE THEORY

By Gilles Pijaudier-Cabot,[1] S. M. ASCE and Zdeněk P. Bažant,[2] F. ASCE

ABSTRACT: In the usual local finite element analysis, strain softening causes spurious mesh sensitivity and incorrect convergence when the element is refined to vanishing size. In a previous continuum formulation, these incorrect features were overcome by the imbricate nonlocal continuum, which, however, introduced some unnecessary computational complications due to the fact that all response was treated as nonlocal. The key idea of the present nonlocal damage theory is to subject to nonlocal treatment only those variables that control strain softening, and to treat the elastic part of the strain as local. The continuum damage mechanics formulation, convenient for separating the nonlocal treatment of damage from the local treatment of elastic behavior, is adopted in the present work. The only required modification is to replace the usual local damage energy release rate with its spatial average over the representative volume of the material whose size is a characteristic of the material. Avoidance of spurious mesh sensitivity and proper convergence are demonstrated by numerical examples, including static strain softening in a bar, longitudinal wave propagation in strain-softening material, and static layered finite element analysis of a beam. In the last case, the size of the representative volume serving in one dimension as the averaging length for damage must not be less than the beam depth, due to the hypothesis of plane cross sections. It is also shown that averaging of the fracturing strain leads to an equivalent formulation, which could be extended to anisotropic damage due to highly oriented cracking.

INTRODUCTION

Progressive damage due to distributed cracking has to be treated in structural analysis as strain softening. This behavior is typical of concrete and is observed also in many other brittle heterogeneous materials such as rocks, stiff clays, two-phase ceramics, various composites, ice, wood and wood particle board, paper, filled elastomers, polymers and asphalt concretes, fiber-reinforced concrete, etc. Strain softening can be modeled by various types of constitutive laws, including endochronic theory, plastic-fracturing theory, plasticity with decreasing yield limit, bounding surface theory, and most recently continuum damage theory (4).

Application of these models in finite element programs and other methods of structural analysis, however, runs into severe difficulties. As shown rigorously already in 1974 (2), when the stress-strain diagram exhibits a negative slope, and more generally when the matrix of tangential moduli ceases to be positive definite, the strain-softening damage tends to localize in a zone of vanishing volume (a line or a surface). This is true not

[1]Grad. Res. Asst., Northwestern Univ., Evanston, IL 60201; on leave from Laboratoire de Mécanique et Technologie, Cachan, France.

[2]Prof. Dept. of Civ. Engrg., and Dir., Ctr. for Concrete and Geomaterials, Northwestern Univ. Tech 2410, Evanston, IL 60201.

Note. Discussion open until March 1, 1988. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on September 2, 1986. This paper is part of the *Journal of Engineering Mechanics*, Vol. 113, No. 10, October, 1987. ©ASCE, ISSN 0733-9399/87/0010-1512/$01.00. Paper No. 21882.

only of statics, but also of dynamics (5,14,15). Because the energy dissipation per unit volume is finite, the vanishing of the damage zone volume causes the indicated structure to fail at zero energy dissipation. This is physically unrealistic, although mathematically the solutions of at least some structural problems with strain softening do exist and represent the limit of the finite element solution for a vanishing element size (4,5,8,10–12).

The consequence of strain localization due to strain softening is that finite element solutions exhibit strong spurious mesh sensitivity, becoming unobjective with regard to the analyst's choice of the mesh (2). These features have first been documented by numerical examples for simple strain-softening constitutive laws (4,8,9,11,16). Recently, however, it has been shown (12,13) by numerical examples of beams that the same spurious mesh sensitivity due to strain localization occurs with the models of continuum damage mechanics (17,26–28,33–38,40,46).

The simplest, but admittedly a crude method to avoid strain localization and the associated mesh sensitivity is to impose a lower limit on the element size (2), as introduced in the crack band model (8,9,11,16). Alternatively, a strain-softening band of a width that is a material property may be embedded in the finite elements, as is done in the composite damage models of Pietrusczak and Mróz (41), and Willam, et al. (44,45) (see also Ref. 4).

Still another possibility is to consider a line of damage (i.e., a crack), characterized by a stress-displacement rather than stress-strain relation, as developed for concrete in the model of Hillerborg, et al. (21,22). The line-crack stress-displacement formulation, however, is not generally applicable. It cannot cope with problems in which the damage zone does not localize fully, because the spacing of line cracks is arbitrary and makes the response of this model unobjective in such situations (4).

A more general and fundamental way to avoid strain localization to a zero volume and to overcome spurious mesh sensitivity is to adopt a nonlocal continuum approach (3,6,7,14). The idea of a nonlocal continuum, introduced during the 1960s by Kröner (29), Kunin (31), Krumhansl (30), Eringen and Edelen (19), and others (4), is to consider the stress to be a function of the mean of the strain from a certain representative volume of the material centered at that point. Introducing certain essential modifications of the classical nonlocal theory, Bažant, and Bažant, et al. (3,7,13) developed a nonlocal model that can handle strain softening in finite element analysis in a consistent manner, while at the same time avoiding spurious mesh sensitivity and precluding strain localization to a vanishing volume. The nonlocal formulation, which may be considered as the limit case of an imbricated system of finite elements (7,13), has nevertheless two properties that are inconvenient for practical applications: (1) The behavior is nonlocal for all situations, including the elastic- or plastic-hardening response; and (2) an overlay with a local continuum must be introduced in order to avoid certain zero-energy periodic modes of instability. The objective of the present study is to develop a modified nonlocal formulation which avoids these two inconvenient properties.

To this end it is useful, although not requisite, to adopt a formulation in which damage is described by a separate variable, distinct from the variables that describe the elastic behavior. This convenient feature is

provided by continuum damage mechanics (17,25,27,28,33–37,40,46). The principal idea of this paper is to use the nonlocal concept only for the variables that control damage and not for the strains or stresses in the constitutive relation. Such an approach will represent a basic departure from the classical formulation of a nonlocal continuum.

The purpose of introducing the nonlocal concept for damage is to limit the localization of strain softening to a zone of a certain minimum size that is a material property, i.e., to serve as a localization limiter. A variety of mathematical formulations have recently been introduced to serve, in effect, as localization limiters. The simplest formulation of localization limiters is involved in the crack band model (2,8,9,11,16), as well as in the composite-damage finite element models (40,44,45). An alternative approach to localization limiters is to make the material strength or yield limit a function of the strain gradient. This idea, which was perhaps first proposed by L'Hermite and Grieu (32) for shrinkage cracking and was developed to describe the size effect in bending of concrete beams with assumed plane cross section (26), has recently been adopted for strain softening in general by Schreyer and Chen (43), Floegl and Mang (20), and Mang and Eberhardsteiner (39). Since the introduction of strain gradient is equivalent to comparing to the strength limit the stress value at a certain small distance from the point, this approach is also essentially nonlocal. Another variant of the nonlocal approach to localization limiters is to introduce into the constitutive equation higher-order spatial derivatives, particularly the Laplacian, which arise from a Taylor series expansion of the nonlocal averaging integral (3,1). In this study, which is based on a 1986 report (42), we will show that nonlocal damage is a very general and effective formulation of localization limiters.

Continuum Damage Theory

For the sake of simplicity, attention will be restricted to isotropic materials with isotropic damage. Time-dependent material response (creep), as well as the effect of temperature and humidity will be neglected. The material will be assumed incapable of plastic behavior, which implies that unloading and reloading follow straight lines passing through the origin. Thus, our formulation of damage theory will be a special case of the general theory for continuum damage and plasticity (17,33). For the sake of simplicity, the anisotropic aspect of damage will be neglected, i.e., the general tensor $\boldsymbol{\Omega}$ will be approximated by the spherical tensor $\boldsymbol{\Omega} = \Omega\,\mathbf{I}$, where Ω is the scalar and $\mathbf{I}$ the identity matrix. The present idea of nonlocal damage could nevertheless be implemented for anisotropic damage as well.

In the theory of continuum damage mechanics, first proposed by Kachanov (25) for creep and applied by Lemaitre, and Lemaitre and Chaboche (36,37), Krajcinovic (28) and others (17,40,46) to quasi-static behavior of strain-softening materials, the state of the material is characterized by strain tensor $\boldsymbol{\varepsilon}$ and damage Ω is evaluated at the given point of the continuum. The strains are assumed to be small. From the thermodynamic viewpoint, the state of the material may be characterized by its free energy density ψ, defined as

$$\rho\psi = \tfrac{1}{2}\boldsymbol{\sigma} : \boldsymbol{\varepsilon} \quad \text{(1)}$$

in which ρ = mass density; the colon denotes the tensor product contracted on two indices; and $\boldsymbol{\sigma}$ = stress tensor, which may be expressed as

$$\boldsymbol{\sigma} = \frac{\partial(\rho\psi)}{\partial\boldsymbol{\varepsilon}} \quad \text{(2)}$$

As a result of microcracking, the net resisting area of the material which transmits stress diminishes. Consequently the true stress $\boldsymbol{\sigma}'$ in the undamaged material between the cracks becomes larger than the macroscopic stress $\boldsymbol{\sigma}$, which is in damage theory described by the relation (25):

$$\boldsymbol{\sigma}' = \frac{\boldsymbol{\sigma}}{(1-\Omega)} \quad \text{(3)}$$

in which Ω = damage (we use Ω instead of the usual notation ω because Ω will be later defined as nonlocal while the notation ω is reserved for the usual local damage); Ω represents an internal (hidden) variable which can never decrease ($\dot{\Omega} \geq 0$), has the initial value $\Omega = 0$, and always $\Omega \leq 1$ with $\Omega = 1$ representing the idealized asymptotic failure of a homogeneously deformed material ($\boldsymbol{\varepsilon} \to \infty$, $\boldsymbol{\sigma} \to 0$). In practice, failure occurs earlier, at $\Omega < 1$, which is obtained through analysis of localization, such as crack-tip singularity.

Assuming, for the sake of simplicity, that the strain in the undamaged matrix of the material is equal to the macroscopic strain, and that the material exhibits no plasticity, Lemaitre and Chaboche (38) and others expressed the true stress as

$$\boldsymbol{\sigma}' = \mathbf{C} : \boldsymbol{\varepsilon} \quad \text{(4)}$$

Substituting Eq. 4 into Eq. 3, we have

$$\boldsymbol{\sigma} = (1-\Omega)\mathbf{C} : \boldsymbol{\varepsilon} \quad \text{(5)}$$

and substituting this into Eq. 1 we obtain

$$\rho\psi = \frac{1-\Omega}{2}\boldsymbol{\varepsilon} : \mathbf{C} : \boldsymbol{\varepsilon} \quad \text{(6)}$$

Since damage comprises the creation and propagation of cracks and voids, growth of damage dissipates energy. By differentiation of Eq. 6, the energy dissipation rate is:

$$\varphi = -\frac{\partial(\rho\psi)}{\partial t} = -\frac{\partial(\rho\psi)}{\partial\Omega}\frac{\partial\Omega}{\partial t} = Y\dot{\Omega} \quad \text{(7)}$$

in which we introduce the quantity:

$$Y = -\frac{\partial(\rho\psi)}{\partial\Omega} = \frac{1}{2}\boldsymbol{\varepsilon} : \mathbf{C} : \boldsymbol{\varepsilon} \quad \text{(8)}$$

called the damage energy release rate (38). According to the second law of thermodynamics (or Clausius-Duhem inequality), $\varphi \geq 0$. Since Y is a positive definite function of $\boldsymbol{\varepsilon}$ and $\dot{\Omega} \geq 0$, this condition is always verified.

From this definition, however, one must exclude reverse stiffening phenomena such as crack healing. This phenomena could be described by a second damage variable (31,43).

The damage properties of the material may be generally characterized, on the basis of experiments, by the damage evolution equation of the form $\dot{\Omega} = f_1(\Omega, \boldsymbol{\varepsilon}:\boldsymbol{\sigma})$. Although the present idea can be applied in general, we will consider, for the sake of simplicity, only the special case where this evolution equation is integrable for monotonic damage growth and reduces to the functional form $\Omega = f(Y)$. Loading and unloading are distinguished by means of the loading function:

$$F(Y) = Y - \kappa(Y) \quad \text{(9)}$$

in which $\kappa(Y)$ = hardening-softening parameter. Its initial value is given as Y_1. After Y first exceeds Y_1, $\kappa(Y)$ is assumed to be equal to the largest value of Y reached thus far at the given point. The damage evolution at a point of the continuum for loading and unloading may then be defined as follows:

$$\text{If} \quad F(Y) = 0 \quad \text{and} \quad \dot{F}(Y) = 0, \quad \text{then} \quad \Omega = f(Y) \quad \text{(10a)}$$

$$\text{If} \quad F(Y) < 0, \quad \text{or if} \quad F(Y) = 0 \quad \text{and} \quad \dot{F}(Y) < 0, \quad \text{then} \quad \dot{\Omega} = 0 \ldots \quad \text{(10b)}$$

It must be emphasized that Eqs. 5 and 8–10 serve only as a simple prototype of damage formulation. Obviously, more complicated damage evolution laws and damage loading functions are required to distinguish between damage in tension and compression, or in other triaxial states (33,46).

Nonlocal Continuum Damage

The foregoing standard formulation of damage theory is local. It has been demonstrated (2,5–10,12,13) that, in a local theory, the damage or strain softening can, and often does, localize into a zone of vanishing volume. Since φ is finite, the total energy dissipation rate in this volume tends to zero, i.e.

$$\lim_{V\to 0}\int_V \varphi\, dV = 0 \quad \text{(11)}$$

In particular, the dynamic failure due to wave propagation in a material represented by the tensile stress-strain curve shown in Fig. 1(*a*) occurs right at the start of softening (i.e., at the maximum stress point), without any stable progressive accumulation of damage before failure and without any dissipation of energy (5). This physically unrealistic property is also borne out by numerical finite element computations when the finite element subdivision of the structure is refined (5,14). As already mentioned, it is proposed to circumvent this problem by defining damage in a nonlocal manner. This can be done by replacing the local definition, $\Omega = f(Y)$, with the nonlocal definition:

$$\Omega = f(\bar{Y}) \quad \text{(12)}$$

in which $\bar{Y}$ represents the mean of Y over the representative volume V, of the material centered around the given point and Ω the nonlocal damage.

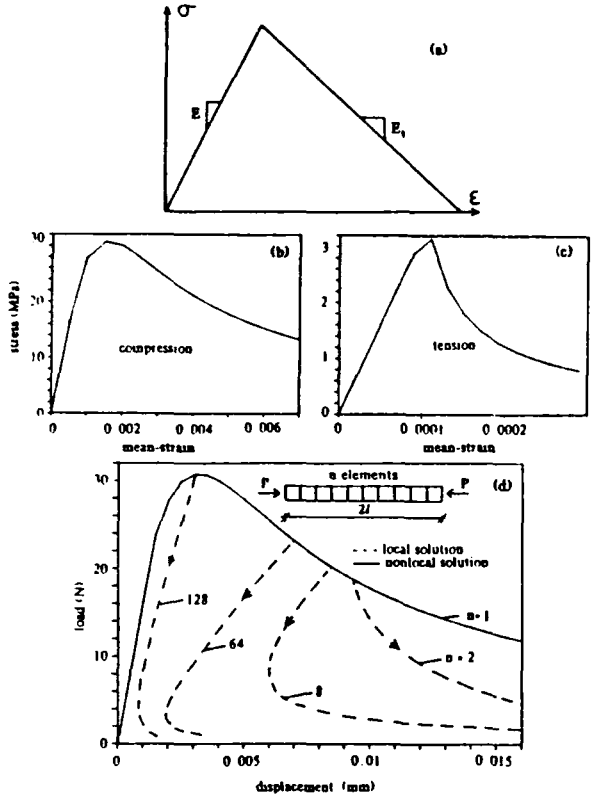


FIG. 1. (*a*) Simplified Triangular Stress-Strain Curve for Strain-Softening Materials; (*b*), (*c*) Stress versus Mean-Strain Curve for Nonlocal Model; (*d*) Local and Nonlocal Computation of Load-Displacement Curve in Compression

We will use

$$\bar{Y} = \frac{1}{V_r}\int_{V_r} Y\, dV \qquad (13)$$

More generally, a weighting function could be introduced in the integrand of this equation, the same as in Ref. 3. The loading function for damage must also be formulated in terms of nonlocal quantities, i.e., Eq. 9 is replaced by

$$F(\bar{Y}) = \bar{Y} - \kappa(\bar{Y}) \qquad (14)$$

and Y is replaced by $\bar{Y}$ in Eq. 10. The representative volume in one, two, or three dimensions may be taken as a line segment of length l, or a circle or sphere of diameter l where l = characteristic length of the material. That spatial averaging becomes physically realistic if we note that the macroscopic equivalence between our assumed continuum with damage and the real heterogeneous material such as concrete cannot be achieved within a domain smaller than a few aggregate sizes. The material parameter l cannot be identified from tests of specimen whose strain is (or is assumed to be)

homogeneous. Identification of l will require both tests in which the damage localizes (as in fracture tests) and tests in which damage does not localize. This problem is beyond the scope of this paper.

Due to spatial averaging, a special treatment is required for points located at the boundary or so close to it that a part V^* of the material representative volume V_r protrudes outside the boundary. For such points, which occupy a boundary layer of a certain characteristic thickness that is a material property, Eq. 13 must be modified as

$$\bar{Y} = \frac{1}{(V_r - V^*)}\int_{V_r - V^*} Y\, dV \qquad (15)$$

This means that the protruding part of the representative volume is chopped off and averaging is carried out only through the part of the representative volume that lies inside the body. The existence of a boundary layer is an inevitable consequence of the nonlocal continuum concept.

Special Uniaxial Formulation

In one dimension, the representative volume reduces to a line segment whose length is the characteristic length, l, provided that no weighting function is used. The mean damage energy release rate is defined as

$$\bar{Y}(x) = \frac{1}{(x_2 - x_1)}\int_{x_1}^{x_2} \frac{1}{2} E\varepsilon^2(x)\, dx \qquad (16)$$

where E = initial Young's elastic modulus. If the one-dimensional bar occupies the interval $0 \le x \le L$ such that $L > l$, the limits of the characteristic length are defined as

$$x_1 = x - l/2,\ x_2 = x + l/2 \quad \text{if} \quad x_1 \ge 0; \quad \text{else} \quad x_1 = 0,\ x_2 = x + l/2 \qquad (17a)$$

$$x_1 = x - l/2,\ x_2 = x + l/2 \quad \text{if} \quad x_2 \le L; \quad \text{else} \quad x_1 = x - l/2,\ x_2 = L \qquad (17b)$$

The imbricate finite element scheme developed before to treat the nonlocal behavior could be chosen to calculate averages directly from nodal displacements. This approach, however, would be less efficient numerically since averaging, required only for computation of damage, can here be explicitly implemented with the usual, nonimbricated mesh on the basis of local strains.

To obtain a realistic shape of the stress-strain diagram with strain-softening, we use the following definition of damage evolution:

$$\text{If} \quad F(\bar{Y}) = 0 \quad \text{and} \quad \dot{F}(\bar{Y}) = 0, \quad \text{then} \quad \Omega = 1 - \frac{1}{1 + b(Y - Y_1)^n} \qquad (18)$$

$$\text{If} \quad F(\bar{Y}) < 0, \quad \text{or} \quad F(\bar{Y}) = 0 \quad \text{and} \quad \dot{F}(\bar{Y}) < 0, \quad \text{then} \quad \dot{\Omega} = 0 \qquad (19)$$

in which b, n, and Y_1 represent the material damage parameters, Y_1 being the damage threshold. In numerical computations, different thresholds, Y_1, were used for tension and compression ($\sigma > 0$ or $\sigma < 0$). The parameters for tension were $b = 9.27 \times 10^{-3}$, $Y_1 = 180.5$ MPa, $n = 1$, and for compression $b = 2.05 \times 10^{-5}$, $Y_1 = 8540$ MPa, $n = 1$. Note that

generalization of Eq. 18 to multiaxial stress states would require certain precautions analyzed elsewhere (17,33).

It may be noted that, according to Eq. 18, the area under the complete uniaxial stress-strain diagram $\sigma = (1 - \Omega) E\varepsilon$ up to $\varepsilon \to \infty$ is infinite if $n = 1$, and finite if $n > 1$. For the modeling of complete failure due to distributed cracking, the stress-strain diagram should be integrable, however for the present purpose of demonstrating objectivity of the theory this condition is unimportant, and $n = 1$ has been used for the sake of simplicity. The formula in Eq. 18 for $n = 1$ has the advantage that its parameters b and E can be identified from given test data by linear regression if ε/σ is plotted versus $Y - Y_1$; the slope of the regression line of this plot is b/E, and its vertical axis intercept is $1/E$.

Obtained from Eq. 18, the stress versus mean strain (displacement divided by length) of a homogeneously deformed specimen of length l and a united cross-sectional area in tension and compression are shown Figs. 1(*b*) and (*c*). The shapes of these curves are obviously quite realistic for concrete.

To verify proper convergence of nonlocal damage formulation, various one-dimensional problems have been solved by finite elements. Only constant-strain finite elements have been used, for two reasons: (1) The numerical implementation of damage averaging is simpler than for higher-order elements; and (2) large higher-order elements cannot represent the discontinuities due to cracking as well as small constant-strain elements. The center of the averaging line segment having characteristic length l is made to coincide with the center of each element, whose coordinate is x^c. To implement the averaging, we need to first determine the numbers $n(i, j)$, $j = 1, \ldots, n_e(i)$ of all finite elements which lie entirely or partly within the characteristic length l, and their total number $n_e(i)$. We also compute the portion $h_{i,j}$ of the element length h_e, which lies within the characteristic length, i.e., within the interval $(x_i^c - l/2, x_i^c + l/2)$, x_i^c being the coordinate of the center of the element number i. For elements that lie entirely within the characteristic length, $h_{i,j} = h_e$. At each loading step or time step of the computation, the mean value Y_i of Y is calculated for each element number i from the following averaging formula:

$$Y_i = \sum_{j=1}^{n_e(i)} g_{i,j} Y_{n(i,j)} \qquad (20a)$$

$$a_i = \sum_{j=1}^{n_e(i)} h_{i,j} \qquad (20b)$$

$$g_{i,j} = \frac{h_{i,j}}{a_i} \qquad (20c)$$

which is applicable in general, whether or not the characteristic length reaches beyond the boundary. It is, of course, again advantageous to calculate the values of $h_{i,j}$ and a_i at the beginning of the computation and store them, provided that the computer's storage capacity is not overtaxed.

The averaging rule in Eqs. 20*a–c* has been implemented in a nonlinear finite element code based on continuum damage theory. First the tension

or compression of a bar consisting of N elements has been calculated for various element subdivisions of constant element size, with $N = 8, 64, 128$. The bar length is $L = 2l$ and the characteristics length is given as $l = 1$. For the nonlocal approach, calculated curves of stress versus mean strain in the bar are the same for various N [Figs. 1(*b*) and (*c*)]. For comparison, computations were also made for the classical local approach. In that case, it is found that at a certain point of the descending branch [the branching points in Fig. 1(*d*)], uniform deformation of the bar becomes unstable and further softening localizes into a single element whereas the other elements unload. The subsequent unstable portion of the response curve is numerically obtained by prescribing the strain into the softening element while the remaining part of the structure still behaves elastically. (According to Ref. 2, softening must localize into a single element if the formulation is local.) After the branching point (bifurcation point), only the response curve which descends sharply is possible. For sufficiently small elements, this curve exhibits snap-back [see Fig. 1(*d*); calculated for compression]. The fact that element subdivisions $N = 8, 64, 128$ give very different results confirms that local damage theory is unobjective with regard to element choice, i.e., it exhibits spurious mesh sensitivity and improper convergence, as established in Refs. 2 and 8. For the nonlocal solution [the smooth curve in Fig. 1(*d*)], no instability develops for this beam, which is relatively short compared to the characteristic length l. Note also that when the beam is loaded through a spring, which is typical of the situation in a testing machine, then instability (bifurcation) occurs earlier and the response curve after the critical point shifts to the left.

Longitudinal Wave in Strain-Softening Materials

An instructive example to study dynamic strain localization due to strain softening is the interference of two longitudinal waves propagating in opposite directions (4). Consider a bar of length $L = 4l$, and beginning with the instant $t = 0$ we impose as the boundary conditions constant outward velocities $\dot{u} = -c$ at left boundary point $x = 0$ and $\dot{u} = c$ at right boundary point $x = L$ ($c > 0$). Velocities are such that the initial inward waves are elastic, consisting of a step of strain of magnitude $\varepsilon = c/v$, where $v = (E/\rho)^{1/2}$ = elastic wave velocity and ρ = mass per unit length = 2,500 kg/m^3. At time $t = L/2v$, the two elastic waves meet at midlength ($x = L/2$), which causes strain to instantly double if the behavior remains elastic. However, we assume the boundary velocities to be sufficiently large so that superposition of the two wave fronts at midlength would produce strain softening at the midlength point. In particular, we consider that c is such that $\varepsilon = c/v = 0.75\varepsilon_p$ (ε_p strain at peak stress).

For local behavior, this problem has an exact solution (5), which is unique and may be used for a convergence check of the finite element solution. We solve the problem with a uniform finite element subdivision, using both local and nonlocal damage models. The explicit algorithm, as stated in detail in Refs. 7 and 12, is used, with time step $\Delta t = 0.2 \times 10^{-6}$ sec. For numerical calculations we assume $Y_1 = 180.5$ MPa, $\varepsilon_p = 10^{-4}$, $E = 32{,}000$ MPa, $c = 8.5$ mm/s.

Fig. 2 shows a comparison of nonlocal finite element solutions for different numbers N of finite elements with the exact local solution (5). A good

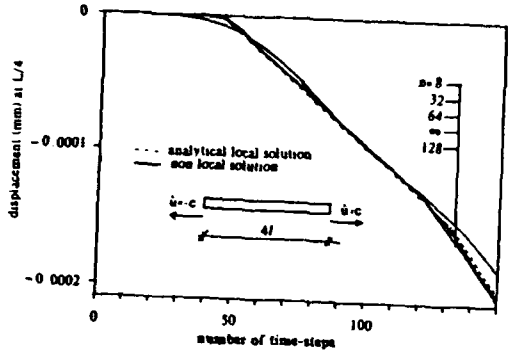


FIG. 2. Interference of Constant Strain Waves in Rod; Convergence of Displacements at $x = 1/4$

convergence may be noted. This is further apparent from Fig. 3, which shows the strain and stress histories at midlength for the nonlocal finite element solution, compared to the exact analytical solution. Despite large numerical noise, which is inevitable in this type of problem with a step wave front, the histories in Figs. 3(*a*) and (*b*) again indicate that convergence may be taking place for the overall response, the peak stress value excepted [Fig. 3(*b*)].

The profiles of strain ε and damage Ω at various times t are plotted in Figs. 4(*a*) and (*b*) for a mesh of $N = 64$ elements. The peak stress is reached [Fig. 3(*b*)] at time $t = 0.00185$ sec (93 time steps), and at $t = 0.0019$ sec (95 time steps) the damage has already started to develop and the strain profiles begin to exhibit a peak at midlength ($x = L/2$). It is noteworthy that the size (length) of the strain-softening zone is approximately equal to the characteristic length l. As it appears, l is the smallest length on which damage can exist [this observation lends further justification to the crack-band model (11)]. It may be also observed from the strain and damage profiles (Fig. 4) that the strain-softening zone gradually expands up to a size approximately $2l$. At the same time, damage accumulates within the entire strain-softening zone. Failure occurs for a high number of time steps in the damaged part of the bar.

The differences between local and nonlocal solutions are best revealed by Figs. 5(*a*)–(*d*). It shows the strain and damage profiles at time 3×10^{-5} sec for different numbers of elements, N. The strain profiles obtained for the local formulation [Fig. 5(*b*)] conspicuously display progressively sharper localization of the strain-softening damage zone, which apparently exhibits physically incorrect convergence to a Dirac delta function. By contrast, the nonlocal solution obviously converges [Figs. 5(*a*) and (*c*)] to a distribution with a finite-size strain-softening damage zone. This behavior is of the same type as previously achieved with the nonlocal imbricate continuum, in which the nonlocal concept was applied not only to damage but to total strains (7). The corresponding stress profiles are shown in Figs. 5(*e*) and (*f*), and again convergence to a strain-softening zone of finite size may be noted. The comparison in Fig. 5 provides the most compelling argument for the nonlocal approach to strain-softening damage.

In comparison to the previous nonlocal solution for the same problem based on the imbricate continuum model (7), it may be observed from the

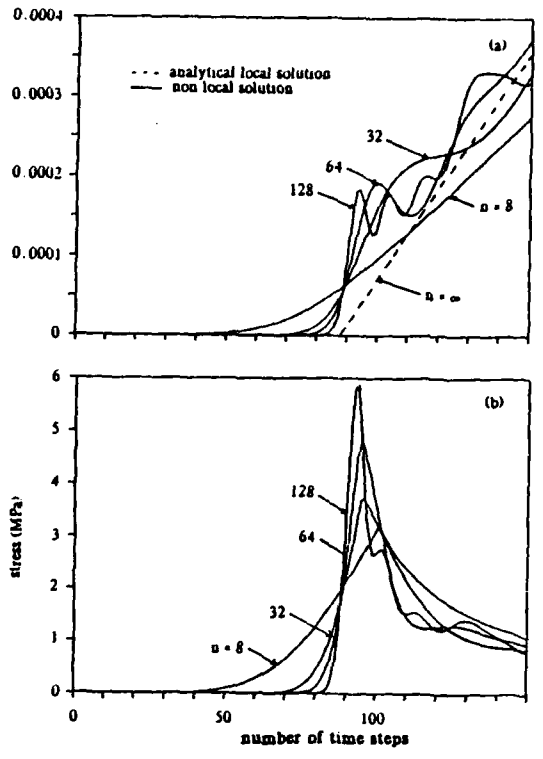


FIG. 3. Interference of Constant Strain Waves in Rod: (*a*), (*b*) Convergence of Stress and Strain Histories for Mesh Refinements

stress profiles at various times in Figs. 5(*e*) and (*f*) that the numerical solution behaves better near the boundary. In the previous solution, stress profiles near the boundary showed poor convergence, exhibited much numerical noise, and their boundary slopes were especially scattered. No doubt this was due to the difficulty in imposing physically reasonable boundary conditions when the total strain, including the elastic one, is treated in a nonlocal manner. In the present solution, characterized by local treatment of the elastic part of the strain, convergence and general behavior of the solution near the boundary is as good as it is in the interior of the bar.

The most important aspect of strain-softening damage is energy dissipation. According to Ref. 7 (Eq. 34), the total energy dissipated in the bar, W, may be calculated from the equation

$$W(t_{r+1}) = W(t_r) + \sum_{n=1}^{N} \frac{h_e}{2} \times [\sigma_{n,r}\varepsilon_{n,r} - \sigma_{n,r+1}\varepsilon_{n,r+1} + (\sigma_{n,r} + \sigma_{n,r+1})(\varepsilon_{n,r+1} - \varepsilon_{n,1})] \quad \text{..... (21)}$$

in which subscript n refers to element numbers and subscript r refers to discrete time $t_r (r = 1, 2, 3 \ldots)$. Fig. 6 shows the values of energy dissipated up to time $t = 3 \times 10^{-5}$ sec (150 time steps) for both local and nonlocal solutions. As before (7), we see that for the local solution, W

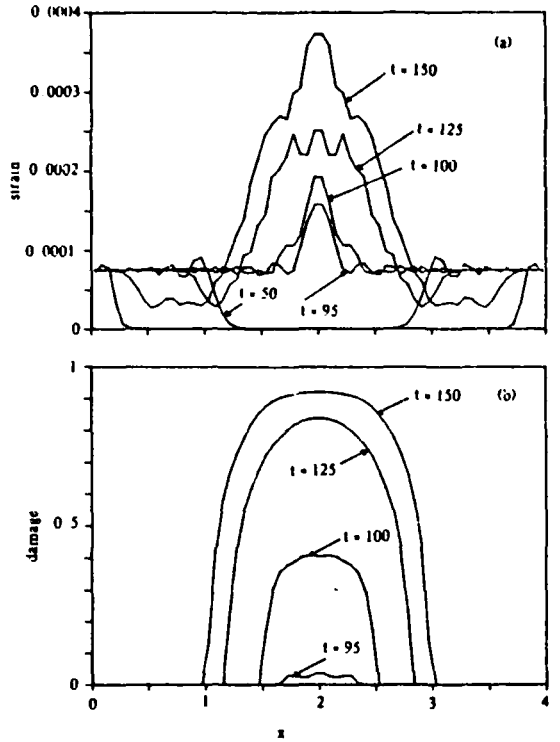


FIG. 4. Interference of Constant Strain Waves in Rod: (*a*), (*b*) Evolution of Strain and Damage Profiles

decreases as the number of elements N increases, and apparently converges to zero as $N \to \infty$. This property is physically unacceptable. By contrast, the present nonlocal solution, the same as the imbricate nonlocal solution in Refs. 7 and 10, yields about the same energy dissipation values for all element subdivisions and converges to a finite value as $N \to \infty$.

As another example (Fig. 7), we solve by finite elements a bar which is free at the left end ($x = 0$) and clamped at the right end ($x = L$). We assume that $L = 2l$. The boundary condition at $x = 0$ is $\sigma(t) = \sigma_0 H(t)$ where $H(t)$ is the Heaviside step function. At the right end, $u = 0$ at all times. Material properties are the same as in the previous example.

A wave with a step-function stress profile, with wave-front magnitude σ_0, propagates toward the clamped end, and is reflected at time $t = L/v$, at which the stress is doubled according to the elastic solution. By choosing $\sigma_0 = 1.8$ MPa, the strain at wave arrival to the clamped end exceeds ε_p, and strain softening is produced at the clamped end.

For the local solution, strain-softening damage again does not propagate from the clamped end, i.e., it remains concentrated in a single element, the boundary element. Fig. 7(*a*) shows the histories of displacement u at bar midlength ($x = L/2$) for various element subdivisions N. The nonlocal solutions apparently converge as N increases. Fig. 7(*b*) shows the stress

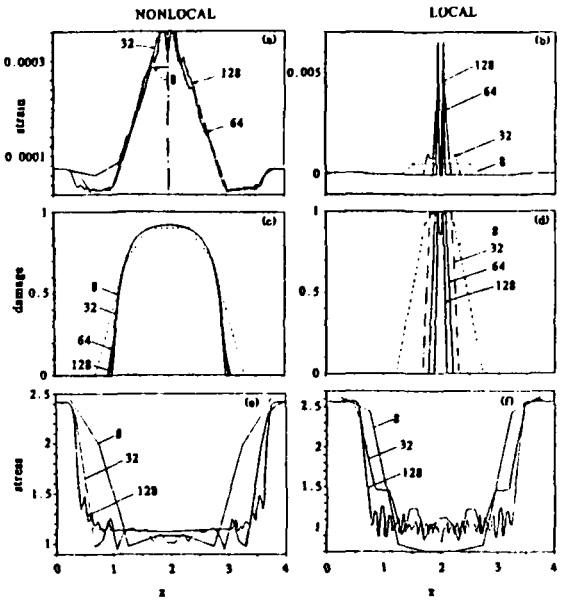


FIG. 5. Interference of Constant Strain Waves in Rod: (*a*) Convergence of Strain, (*c*) Damage and (*e*) Stress Profiles; (*b*), (*d*), (*f*) Comparison with Local Solution

history in the element at the clamped end for various N. Again, the nonlocal solution apparently converges, including peak stress value.

It may be noted that the present clamped bar responds overall in the same way as one-half of the bar in the previous example. However, comparison of Fig. 7(*b*) to Fig. 3(*b*) shows that for the same element size, peak stress values are different [4.7 MPa in Fig. 3(*b*) and 3.7 MPa in Fig. 7(*b*)], for element length $0.0625l$.

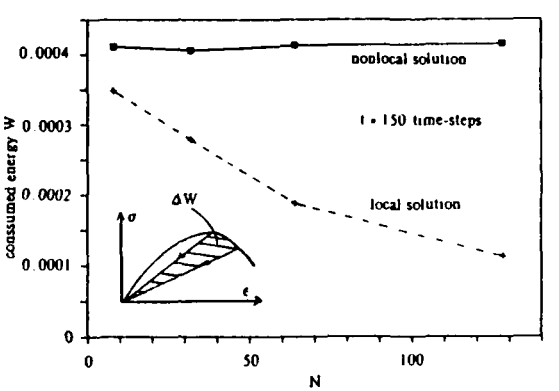


FIG. 6. Interference of Constant Strain Waves in Rod: Convergence of Energy Dissipated Due to Damage

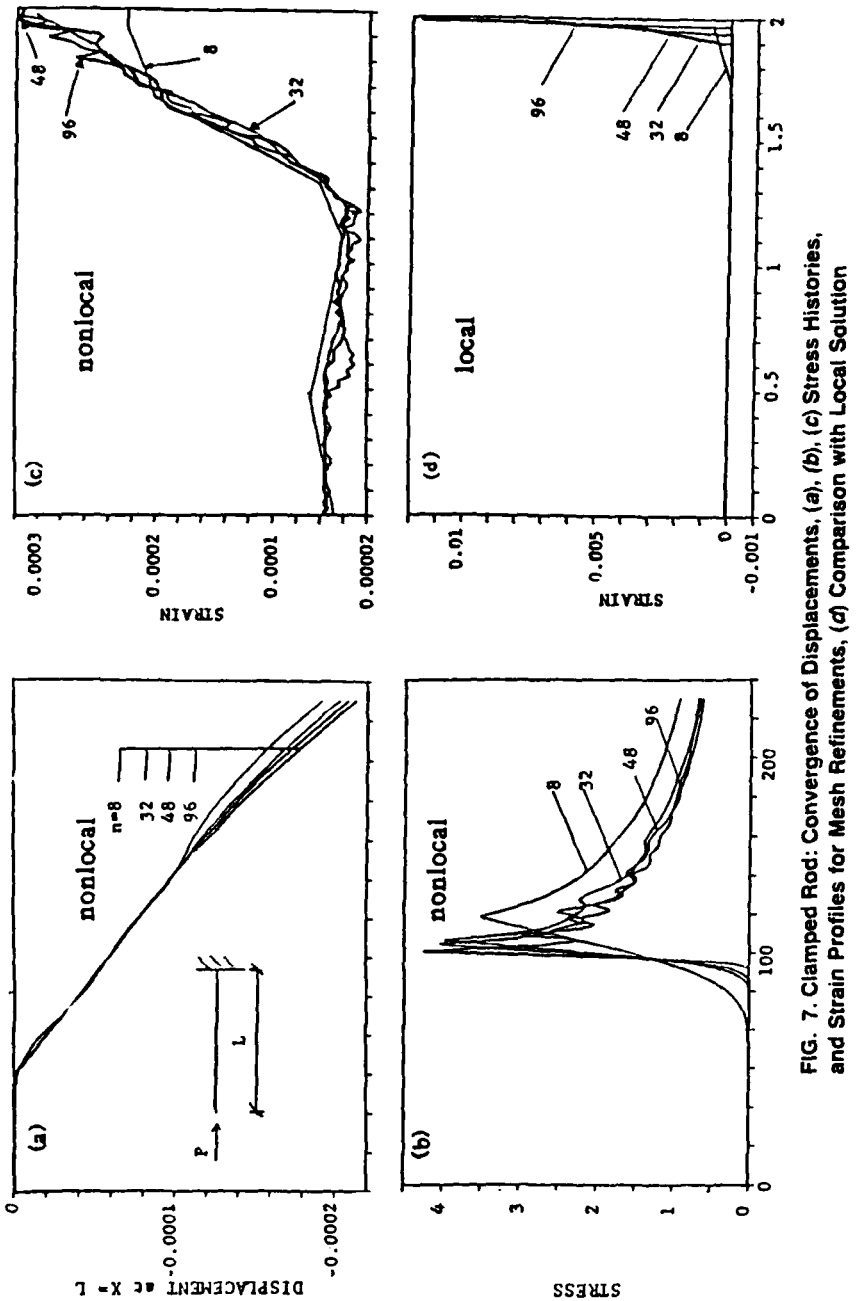


FIG. 7. Clamped Rod: Convergence of Displacements, (*a*), (*b*), (*c*) Stress Histories, and Strain Profiles for Mesh Refinements, (*d*) Comparison with Local Solution

Computations also showed [Fig. 7(*d*)] that convergence to finite unique values of u, ϵ, and σ is not achieved with the local formulation. Especially, the peak strain strongly depends on the element size.

Static Problem of Bending Theory

In a previous study (12,13) it was shown that solutions of beams and frames by the layered finite element technique also exhibit physically incorrect convergence and strong spurious sensitivity to element size when the material is assumed to exhibit strain softening and the formulation is local. Due to kinematic assumption of the preservation of planeness of cross sections and to Saint Venant's principle, the minimum length of finite elements is not only related to the characteristic length l of the material, but cannot be less than the beam depth h (12,13). This lower bound limitation on the element length may be directly used in a local-type finite element analysis similar to the multidimensional crack-band theory, in which a lower limit of the element size is imposed (2,8,11). With such a limitation, however, detailed stress and deflection distribution cannot be resolved, and numerical approximation is mathematically unfounded since the continuum limit at mesh refinement is left undefined.

This limitation may again be circumvented by the present concept of nonlocal damage. In the case of bending theory, however, the definition of spatial averaging (Eq. 6) must be slightly modified. Averaging length needs to be introduced as

$$l' = \max(l, h) \quad \text{..... (22)}$$

in which l is the previously introduced material characteristic length, and h = beam depth. For concrete, $l \approx 3d_a$ where d_a = maximum aggregate size, and since concrete beams are always deeper than three aggregate sizes, the beam depth limitation decides, i.e., $l' = h$. Thus, Eq. 16 for averaging of the damage energy release rate Y must be modified as

$$\bar{Y}(x, z) = \frac{1}{h}\int_{x-h/2}^{x+h/2} Y(\xi, z)\, d\xi, \quad Y(\xi, z) = \tfrac{1}{2} B\epsilon^2(\xi, z) \quad \text{..... (23)}$$

in which x = length coordinate of the beam. Because damage localization in the vertical direction is impossible, due to the constraint of planeness of the cross section, we ignore averaging over the depth coordinate z. However, to take into account the well-known effect of strain gradient on the apparent strength or yield limit in concrete beams analyzed according to the bending theory (26), we could also introduce averaging over the depth:

$$\bar{Y}(x, z) = \frac{1}{hl}\int_{x-h/2}^{x+h/2}\int_{z-l/2}^{z+l/2} Y(\xi, \zeta)\, d\zeta\, d\xi \quad \text{..... (24)}$$

In the present study, which is concerned only with numerical problems of convergence and not with material properties, only the one-dimensional averaging in Eq. 23 has been used. Needless to elaborate, when the averaging length h (or l) protrudes beyond the boundary of the beam, the same treatment as before must be used (i.e., the averaging domain outside the beam is ignored). In the layered finite element model, Eq. 23 is applied

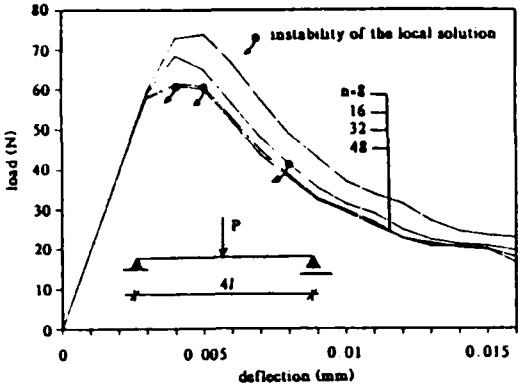


FIG. 8. Three-Point Bending; Convergence of Nonlocal Solution for Load-Deflection Curve

to each layer of each beam element, Y being assumed constant within each layer.

For demonstration, we solve one of the examples previously analyzed in Refs. 12 and 13. It is a simply supported beam of length $4h$, with a concentrated load at midspan (Fig. 8) and a square cross section. Material properties are the same as in our previous examples. Each beam element is divided into 10 layers. Step-by-step loading technique is used, with a load-point displacement controlled (prescribed) and the load calculated as a reaction. The direct iteration method based on the secant modulus is used (12,13). Numerical results for the load-displacement curve of the beam are plotted in Fig. 8 for various numbers N of elements along the beam, which all have the same length. We see that the solution converges well. By contrast, the local solution (12,13) showed a strong dependence on the number of elements N and produced instabilities of the snap-back type in the post-peak response, as discussed in Ref. 13. The occurrence of these instabilities depended on element size, not on beam properties. In the present solution, snap-back instability can occur only as a function of the l/h ratio, when its value is sufficiently high.

Alternative Viewpoint: Mean Fracturing Strain

Among various possibilities of spatial averaging that do not affect elastic strain, another is to average the fracturing strain. This strain, imagined to represent the cumulative overall contribution of microcrack openings, was used as the basis of the crack band model (11) and similar formulations (41,44,45). The use of fracturing strain offers a simpler alternative to anisotropic damage in which the oriented aspect of damage due to cracking can be taken into account. Assuming all cracks to be parallel and denoting their normal direction as z, we write:

$$\varepsilon_z = \varepsilon_z^{el} + \bar{\varepsilon}_z^{fr} \qquad (25a)$$

$$\bar{\varepsilon}_z^{fr} = \frac{\sigma_z}{E}(1 + \bar{\gamma}) \qquad (25b)$$

in which ε_z^{el} = elastic strain determined from stresses σ_x, σ_y, and σ_z; and $\bar{\varepsilon}_z^{fr}$ = mean fracturing strain in the z-direction, specified as a function of strain (possibly also stress and internal variable), and $\bar{\gamma}$ = mean specific fracturing strain. Strains ε_y and ε_x in the directions parallel to the crack are elastic. According to the basic idea advanced here, $\bar{\gamma}(x)$ is obtained by spatial averaging:

$$\bar{\gamma}(x) = \frac{1}{l}\int_{-l/2}^{l/2} \gamma(x + s)\, ds \qquad (26)$$

By choosing

$$\gamma(x) = \left(\frac{b}{2} E\varepsilon_x^2\right)^n \quad \text{for loading} \qquad (27a)$$

$$\gamma = 0 \quad \text{for unloading and reloading} \qquad (27b)$$

the formulation becomes in one dimension identical to the foregoing damage model (Eq. 16). Indeed the damage model used in our examples yields for monotonic loading (and constant E) the uniaxial stress-strain relation

$$\sigma(x) = \frac{E\varepsilon(x)}{1 + \left\{\dfrac{b}{2l} E \displaystyle\int_{-l/2}^{l/2} [\varepsilon(x + s)]^2\, ds\right\}^n} \qquad (28)$$

and Eqs. 25–27 reduce for uniaxial stress to the same equation. By contrast, the classical local approach is equivalent, for monotonic loading, to the stress-strain relation

$$\sigma(x) = \frac{E\varepsilon(x)}{1 + \left[\dfrac{b}{2l} E\varepsilon^2(x)\right]^n} \qquad (29)$$

Alternatively, if the expression of the mean fracturing strain γ is a linear function of ε, averaging γ is equivalent to averaging the total strain ε, and so Eq. 28 is equivalent to:

$$\sigma(x) = \frac{E\varepsilon(x)}{1 + \left\{\dfrac{b}{2l} E\left[\displaystyle\int_{-l/2}^{l/2} \varepsilon(x + s)\, ds\right]\right\}^{2n}} \qquad (30)$$

Fig. 9 shows for the dynamic example treated (Figs. 2–6), the result of this strain-averaging method. With the same overall material behavior, we find that the response diagrams are graphically undistinguishable from those obtained with the energy-averaging formulation.

Instead of Eq. 27, the expression for $\bar{\varepsilon}_z^{fr}$ may be taken as linear, which corresponds to the triangular stress-strain relation used in the original crack-band model [Fig. 1(*a*)]. By introducing the function $\gamma = \exp[-k\varepsilon_z^{fr}] - 1$ in which k, n = constants, we obtain for $\sigma(x)$ a function which generalizes the local stress-strain relation $\sigma = E\varepsilon \exp(-k\varepsilon^n)$ also used in the past for the crack-band model.

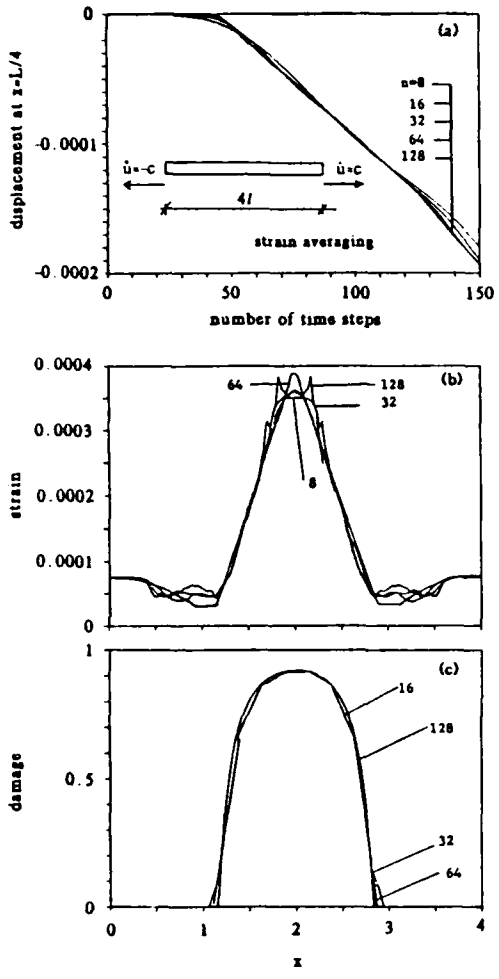


FIG. 9. Strain-Averaging Solution: (*a*) Convergence of Displacement History, (*b*), (*c*) Strain and Damage Profiles

Finally, we should comment on the failure mode. The true failure mode in tension no doubt consists of unstable (runaway) crack propagation. In the multiaxial analysis just illustrated, the transverse crack propagation cannot be modeled; however, by ensuring that $l(\int \sigma\, d\varepsilon)$ = fracture energy of the material, our approach is equivalent to the fracture mechanics approach at least in the sense of total energy balance, which is most important. In local continuum softening analysis, energy balance is incorrect as dissipated energy vanishes. Compression failures can be described by our uniaxial analysis only to the extent that $l(\int \sigma\, d\varepsilon)$ approximates the energy dissipated by axial splitting and shear cracks localized into a segment of specimen length. Such localization can occur only in long

compression specimens. A detailed analysis of compression failure must, of course, be three-dimensional.

Conclusions

1. Although existing imbricate nonlocal formulation for strain-softening materials does insure proper convergence and eliminates spurious mesh sensitivity, it has two inconvenient features: (1) All the behavior is formulated as nonlocal, including the elastic part of strain; and (2) an overlay with a local continuum must be introduced to suppress certain periodic zero-energy modes. These inconveniences are circumvented by the present nonlocal damage theory.

2. The key idea is that nonlocal treatment should be applied only to those variables that cause strain softening, and *not* to the elastic behavior. Thus, nonlocal theory should reduce to a local theory when the response is purely elastic. (The treatment of the plastic-hardening part of the response could no doubt also be local if this is convenient to do).

3. Among various constitutive models which can describe strain softening, the continuum damage mechanics appears ideal for the present purpose since all strain softening is controlled by a single variable, the damage. The essential attribute of the presently proposed nonlocal damage theory is that the damage energy release rate is averaged over the representative volume of the material whose size is a characteristic of the material. However, an alternative formulation in which spatial averaging is applied to the fracturing part (i.e., damage part) of the strain works equally well and may be just as efficient.

4. The characteristic length l (size of the representative volume) is a material property which must be determined by experiments and should be corroborated by micromechanics methods. No doubt, length l is related to the size of material inhomogeneities. However, when nonlocal damage theory is applied to beams, the assumption of preservation of the planeness of cross section requires that the averaging length not be smaller than the beam depth.

5. Numerical finite element computations indicate that nonlocal damage theory avoids spurious mesh sensitivity and that the calculated distributions of strain, stress, damage, and displacement exhibit proper convergence as the finite element subdivision is refined. Most importantly, the energy dissipated due to strain-softening damage converges to a finite value, while for the usual local finite element codes with strain softening this energy converges to zero as the mesh is refined, which is physically meaningless and unrealistic.

6. Since averaging is required by the present theory only in those domains of the structure that exhibit strong nonlinear behavior, the computation is more efficient than with previous imbricate nonlocal formulations. Averaging of damage can be easily introduced in any nonlinear finite element code with a strain-softening model.

Acknowledgments

Partial financial support under U.S. Air Force Office of Scientific Research contract number F49620-87-C-0030DEF with Northwestern Uni-

versity, monitored by Spencer T. Wu, is gratefully acknowledged. Nadine Pijaudier-Cabot is thanked for her outstanding secretarial assistance.

Appendix I. References

1. Aifantis, E. C., "On the Microstructural Origin of Certain Inelastic Models," *Journal of Engineering Materials Technology*, Vol. 106, 1984, pp. 326–330.
2. Bažant, Z. P., "Instability Ductility and Size Effect in Strain-Softening Concrete," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 102, No. EM2, 1976, pp. 331–344; discussions Vol. 103, pp. 357–358, 775–777; Vol. 104, pp. 501–502 (based on Struct. Eng. Report No. 74-8/640, Northwestern University, August, 1974).
3. Bažant, Z. P., "Imbricate Continuum and Its Variational Derivation," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 110, No. 12, 1984, pp. 1693–1712.
4. Bažant, Z. P., "Mechanics of Distributed Cracking," "Applied Mechanics Reviews", ASME, Vol. 39, No. 5, May, 1985, pp. 675–705.
5. Bažant, Z. P., and Belytschko, T. B., "Wave Propagation in a Strain-Softening Bar: Exact Solution," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 111, No. 3, 1985, pp. 381–389.
6. Bažant, Z. P., and Belytschko, T. B., "Localization Limiters for Strain-Softening," Report No. 86-8/428L, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL, August, 1986.
7. Bažant, Z. P., Belytschko, T. B., and Chang, T. P., "Continuum Theory for Strain-Softening," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 110, No. 12, 1984, pp. 1666–1692.
8. Bažant, Z. P., and Cedolin, L., "Fracture Mechanics of Reinforced Concrete," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 106, No. EM6, Dec., 1980, pp. 1287–1306; Discussion and Closure, Vol. 108, 1982, pp. 464–471; see also "Blunt Crack Band Propagation in Finite Element Analysis," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 105, No. 2, 1979, pp. 297–315.
9. Bažant, Z. P., and Cedolin, L., "Finite Element Modeling of Crack Band Propagation," *Journal of Structural Engineering*, ASCE, Vol. 108, No. ST2, Feb., 1982.
10. Bažant, Z. P., and Chang, T. P., "Nonlocal Finite Element Analysis of Strain-Softening Solids," *Journal of the Engineering Mechanics Division*, ASCE, Vol. 113, No. 1, 1987, pp. 84–105.
11. Bažant, Z. P., and Oh, B. H., "Crack Band Theory for Fracture of Concrete," *Materials and Structures*, Vol. 16, No. 93, May-June, 1983, pp. 155–177.
12. Bažant, Z. P., Pan, J., and Pijaudier-Cabot, G., "Softening in Reinforced Concrete Beams and Frames," Report No. 86-7/428s, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL, July, 1986.
13. Bažant, Z. P., Pijaudier-Cabot, G., and Pan, J., "Ductility, Size Effect and Redistribution in Softening Frames," Report No. 86-7/428d, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL, July, 1986.
14. Belytschko, T. B., Bažant, Z. P., Hyun, Y. W., and Chang, T. P., Strain-Softening Materials and Finite Element Solutions," *Computers and Structures*, Vol. 23, No. 2, 1986, pp. 163–180.
15. Belytschko, T. B., Wang, X., Bažant, Z. P., and Hyun, H., "Transient Solutions for One-Dimensional Problems with Strain-Softening," Report, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL, June, 1986.
16. Cedolin, L., and Bažant, Z. P., "Effect of Finite Element Choice in Blunt Crack Band Analysis," *Computer Methods in Applied Mechanics and Engineering*, Vol. 24, No. 3, Dec., 1980, pp. 305–316.
17. Cordebois, J. P., "Critères d'instabilité plastique et endommagement ductile en grandes déformations," Thèse de Doctorat d'Etat ès Sciences Physiques, Université Paris VI, France, 1983.
18. De Borst, R., and Nauta, P., "Non-orthogonal Cracks in a Smeared Finite Element Model," *Engineering Computations*, No. 2, 1985, pp. 35–46.
19. Eringen, A. C., and Edelen, D. G. B., "On Nonlocal Elasticity," *International Journal of Engineering Science*, Vol. 10, 1972, pp. 233–248.
20. Floegl, H., and Mang, H. A., "On Tension-Stiffening in Cracked Reinforced Concrete Slabs and Shells Considering Geometric and Physical Nonlinearity," *Ingenieur-Archiv.*, Vol. 51, No. 314, 1981, 215–242.
21. Hillerborg, A., Modeer, M., and Petersson, P. E., "Analysis of Crack Formation and Crack Growth in Concrete by Means of Fracture Mechanics and Finite Elements," *Cement and Concrete Research*, Vol. 6, 1976, pp. 773–782.
22. Hillerborg, A., "Numerical Methods to Simulate Softening and Fracture of Concrete," *Fracture Mechanics Applied to Concrete Structures*, edited by G. C. Sih, Martinus Nijhoff, The Hague, 1984.
23. Ingraffea, A. R., "Fracture Propagation in Rock," *Mechanics of Geomaterials: Rocks, Concretes, Soils*, edited by Z. P. Bazant, Wiley, Chichester and New York, 1983, pp. 219–258.
24. Ingraffea, A. R., *Fracture Mechanics Applied to Concrete Structures*, edited by G. C. Sih and A. Carpinteri, Martinus Nijhoff, The Hague, 1984.
25. Kachanov, L. M., "Time of Rupture Process Under Creep Conditions" (in Russian), *Izvestia Akademii Nauk*, USSR, No. 8, 1958, pp. 26–31.
26. Karsan, D., and Jirsa, J. O., "Behavior of Concrete Under Varying Strain Gradients," *Journal of Structural Mechanics Division*, ASCE, Vol. 96, No. ST8, 1970, pp. 1675–1696.
27. Krajcinovic, D., and Fonseka, G. V., "The Continuous Damage Theory of Brittle Materials," *Journal of Applied Mechanics*, ASME, Vol. 48, Dec., 1981, pp. 809–815.
28. Krajcinovic, D., "Constitutive Equations for Damaging Materials," *Journal of Applied Mechanics*, ASME, Vol. 50, Jun., 1983, pp. 355–360.
29. Kröner, E., "Elasticity Theory of Materials with Longrange Cohesive Forces," *International Journal of Solids and Structures*, Vol. 3, 1968, pp. 731–742.
30. Krumhansl, J. A., "Some Considerations of the Relations Between Solid State Physics and Generalized Continuum Mechanics," *Mechanics of Generalized Continua*, edited by E. Kröner, Springer-Verlag, Heidelberg, Germany, 1968, pp. 298–331.
31. Kunin, I. A., "The Theory of Elastic Media with Microstructure and the Theory of Dislocations," *Mechanics of Generalized Continua*, edited by E. Kröner, Springer-Verlag, Heidelberg, Germany, 1968, pp. 321–328.
32. L'Hermite, R., and Grieu, J. J., "Etude expérimentales récentes sur le retrait des ciments et des bétons," Annales I.T.B.T.P. (Paris), Vol. 5, No. 52–53, 1952, pp. 494–514.
33. Ladevèze, P., "On an Anisotropic Damage Theory" (in French), Internal Report No. 34, Laboratoire de Mécanique et Technologie, France, 1983.
34. Leckie, F. A., "Constitutive Equations of Continuum Creep Damage Mechanics," *Philosophical Transactions of Royal Society, London*, Ser. A, Vol. 288, 1978, pp. 27–47.
35. Leckie, F. A., and Onat, E. T., "Tensorial Nature of Damage Measuring Internal Variables," *Physical Nonlinearities in Structural Analysis*, edited by J. Hult and J. Lemaitre, Proceedings of the May, 1980, IUTAM Symposium Held at Senlis, France, Springer-Verlag, Berlin, 1981, pp. 140–155.
36. Lemaitre, J., "How to Use Damage Mechanics" (Division lecture presented at SMiRT7, Chicago, 1983), Internal Report No. 40, Laboratoire de Mécanique et Technologie, Ecole Normale Supérieure de l'Enseignement Technique, Paris, France, 1983.
37. Lemaitre, J., and Chaboche, J-L., "Aspect phénoménologique de la rupture par endommagement," *Journal de Mécanique Appliquée* (Paris), Vol. 2, 1978, pp. 317–365.

38. Lemaitre, J., and Chaboche, J. L., "*Mécanique des materiaux solides*," Dunod-Bordas, Paris, 1985.
39. Mang, H., and Eberhardsteiner, J., "Collapse Analysis of Thin R. C. Shells on the Basis of a New Fracture Criterion," *U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures*, Tokyo, May, 1985, Preprints, pp. 217–238.
40. Mazars, J., "Application de la mécanique de l'endommagement au comportement non-linéaire et à la rupture du béton de structure," Thèse de Doctorat d'Etat ès Sciences Physiques, Université Paris VI, France, 1984.
41. Pietruszczak, ST., and Mróz, Z., "Finite Element Analysis of Deformation of Strain-Softening Materials," Vol. 17, 1981, pp. 327–334.
42. Pijaudier-Cabot, G., and Bažant, Z. P., "Nonlocal Damage Theory," Report No. 86-8/428n, Center for Concrete and Geomaterials, Northwestern University, Evanston, Il., August, 1986.
43. Schreyer, H. L., and Chen, Z., "The Effect of Localization on the Softening Behavior of Structural Members," Proceedings of Symposium on Constitutive Equations: Micro, Macro, and Computational Aspects, ASME Winter Annual Meeting, New Orleans, Dec., 1984, edited by K. Willam, ASME, New York, 1984, pp. 193–203.
44. Willam, K. J., Bicanic, N., and Sture, S., "Constitutive and Computational Aspects of Strain-Softening and Localization in Solids," Proceedings of Symposium on Constitutive Equations: Micro, Macro and Computational Aspects, ASME, Winter Annual Meeting, New Orleans, Dec., 1984, edited by K. J. Willam, ASME, New York, 1984.
45. Willam, K. J., Hurlbut, B., and Sture, S., "Experimental, Constitutive and Computational Aspects of Concrete Failure," Preprints, *U.S.-Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures*, Tokyo, May, 1985, pp. 149–172.
46. Zaborski, A., "An Isotropic Damage Model for Concrete," (in French), Internal Report No. 55, Laboratoire de Mécanique et Technologie, E.N.S.E.T., Cachan, France, 1985.

NONLOCAL MICROPLANE MODEL FOR FRACTURE, DAMAGE AND SIZE EFFECT IN STRUCTURES

By Zdeněk P. Bažant,[1] F. ASCE, and Jožko Ožbolt[2]

ABSTRACT

A generalized microplane model, which was previously developed to describe tensile cracking as well as nonlinear triaxial response of brittle-plastic materials in compression and shear, is implemented in a finite element code. To limit localization instabilities due to strain softening and the consequent spurious mesh-sensitivity, the recently proposed concept of nonlocal continuum with local strain (nonlocal damage) is adopted and is combined with the microplane model. An effective numerical algorithm permitting large loading steps is developed, applying the idea of exponential algorithms previously used for creep. Problems due to nonsymmetry of the tangential stiffness matrix are avoided by using the initial elastic stiffness matrix in the incremental force-displacement relations. Numerical results demonstrate that the microplane model, which has previously allowed an excellent description of the existing test data on nonlinear triaxial behavior of concrete as well as unidirectional and multidirectional crack or crack shear, is endowed (in its nonlocal generalization) with the capability of modeling also tensile fracture. The model yields the correct transitional size effect observed in concrete and agrees with the size effect law proposed by Bažant. The formulation is applicable to brittle-plastic materials in general.

INTRODUCTION

At present there exist basically two kinds of material constitutive models: (1) the macroscopic tensorial models, in which the constitutive relation is formulated in terms of the macroscopic continuum stress and strain tensors and their invariants; and (2) the microscopic nontensorial models, in which the micromechanics

[1]Prof. of Civil Eng., Center for Advanced Cement-Based Materials, Northwestern University, Evanston, Illinois 60208.

[2]Visiting Scholar, Northwestern University; Assoc. Prof. on leave from Techn. Univ. of Zagreb.

of deformation is described by some suitably simplified model and the constitutive material properties are characterized by a relation between the stress and strain components (or forces and displacements) on the microlevel. The former is the classical approach, which has led to important advances in the modeling of concrete (e.g. Chen and Chen, 1975; Bažant and Bhat, 1976; Willam and Warnke, 1974; Gerstle, 1980, 1981; Dougill, 1976; Bažant and Kim, 1979; Ortiz, 1985; Cedolin, dei Poli and Iori, 1983; Lin et al., 1987). However, the more realistic macroscopic constitutive models for concrete are rather complicated and are very difficult to identify from test data since they contain many material parameters, typically over fifteen of them, while at the same time the agreement with test data is still less than satisfactory. Extensive recent efforts have not succeeded in overcoming these drawbacks, and it seems that the macroscopic approach has already yielded what it can and that further great efforts can only provide diminishing returns.

More promising seems to be the second kind of material models, which use micromechanics ideas and exist now basically in two versions: the random particle system modeling of the aggregate structure, which however poses forbidding demands on computer time, and the microplane model. The latter still requires more computer time than even the most sophisticated macroscopic models, but with the present computer capabilities is manageable even in larger finite element programs. Compared to the macroscopic constitutive models, the microplane model trades computer time for conceptual simplicity. The modeller does not need to bother with tensorial invariance restrictions and needs to deal only with a few stress and strain components on planes of various orientations, called the microplanes. A simple, nontensorial relation between these components defines the constitutive properties. The macroscopic stress and strain tensors are related to the microplane stresses by a virtual work relation and by a static or kinematic constraint.

The basic idea of microplane models, which is due to G. I. Taylor (1938), was initially applied only to metal plasticity and went under the name "slip theory." Extensions to nonsoftening behavior of rocks and soils were made by Zienkiewicz

and Pande (1977) and Pande (1982). Some major conceptual modifications were needed to make Taylor's idea applicable to strain-softening brittle behavior (Bažant and Oh, 1983, 1985; Bažant and Gambarova, 1984) and most recently to strain-softening brittle-plastic behavior, typical of concrete (Bažant and Prat, 1988). This latest formulation provided so far the most complete material model for concrete under tensile as well as compressive triaxial stress states. Since the deformation on various planes in the microstructure could no longer be called "plastic slip," the more general term "microplane" was coined (Bažant, 1984).

The stress and deformation interactions in the microstructure are in essence of two kinds: (1) interactions over a distance, and (2) interactions between orientations. The microplane models take into account only the latter. The former can be most simply taken into account by the concept of nonlocal continuum (Krŏner, 1968; Eringen and Edelen, 1972; cf. also Bažant, 1986), in which the strain at a point depends not only on the strain at the same point but on the entire strain field in a certain neighborhood of the point. The nonlocal concept has already been proven to be a useful cure to the problems with localization instabilities which are endemic for local continuum models. Applications of the nonlocal concept, however, have so far been confined to the macroscopic tensorial constitutive model. The objective of the present paper is to marry this concept with the more powerful microplane approach, thus gaining a model that can handle in a rather realistic manner both fracture (or damage localization) and nonlinear triaxial deformations.

First we need to describe precisely the form of the microplane model to be used. Then we develop its nonlocal generalization. In the process, we also formulate a new, very effective algorithm (of exponential type) for the analysis of loading steps.

REVIEW OF MICROPLANE MODEL

(a) Basic Hypotheses and Strain Components.

Hypothesis I. - The strains on any microplane are the resolved components of the macroscopic strain tensor ε_{ij}. This hypothesis, which represents a kinematic constraint, yields the relations (Fig. 1a):

$$(\vec{\varepsilon}_R)_i = n_i \varepsilon_{ij}; \quad \varepsilon_N = n_i n_j \varepsilon_{ij}; \quad \varepsilon_D = n_i n_j \varepsilon_{ij} - \varepsilon_V \tag{1}$$

$$(\varepsilon_T)_i = n_k \varepsilon_{ik} - n_i (n_j n_k \varepsilon_{jk}) = \tfrac{1}{2}(n_k \delta_{ij} + n_i \delta_{ik} - 2 n_i n_j n_k)\varepsilon_{jk}; \quad \varepsilon_T = \sqrt{\varepsilon_{Ti} \varepsilon_{Ti}}. \tag{2}$$

Here latin lower case subscripts refer to cartesian coordinates x_i ($i = 1,2,3$); repetition of subscripts implies summation; $\vec{\varepsilon}_R$ is the strain vector on a microplane whose unit normal is n_i, $\vec{\varepsilon}_T$ and ε_T are the tangential vector component of $\vec{\varepsilon}_R$ and its magnitude, ε_{Ti} are the components of $\vec{\varepsilon}_T$, and ε_N is the normal strain component on the microplane (Fig. 1a). The normal strain vector is separated into volumetric strain $\varepsilon_V = \varepsilon_{kk}/3$ and deviatoric strain ε_D, i.e. $\varepsilon_N = \varepsilon_V + \varepsilon_D$. This split, proposed by Bažant, makes it possible to obtain any value of elastic Poisson ratio between -1 and 0.5 and, more importantly, has been found by experience to be essential for allowing good representation of post-peak softening for both tension and compression.

Hypothesis II. - Each microplane resists not only normal strain ε_N, but also shear strains, ε_T, whose vector $\vec{\varepsilon}_T$ is assumed to keep the same direction as the vector of shear stresses $\vec{\sigma}_T$. In the initial formulation by Bažant and Oh (1983, 1985), the resistance to $\vec{\varepsilon}_T$ was neglected, which sufficed for modeling multidirectional tensile cracking. However, if compression response should be modeled simultaneously, resistance to $\vec{\varepsilon}_T$ must be considered (or else the ratio of compression and tensile strengths is incorrect and snapback instability develops soon after the peak stress in compression).

Hypothesis III. - The response on each microplane depends on the main lateral strain ε_L, which is equivalent to the dependence on the volumetric strain $\varepsilon_V = \varepsilon_{kk}/3$. This feature was found to be necessary for modeling triaxial test data for very high confining pressures (but not other data).

Hypothesis IV. - The stress-strain curves for each microplane are path-independent as long as there is no unloading on this microplane. During each unloading and reloading, which is defined separately on each microplane, the curves of the stress and strain differences from the state at the start of unloading are also path-independent. Thus, all the macroscopic path-dependence is produced by various combinations of loading and unloading on all the microplanes. It may be noted that some

microplanes may get unloaded even for macroscopically monotonic or virgin loading, thus making the response path-dependent. The number of possible macroscopic path directions is enormous (for 21 microplanes there are 2^{21} possible tangential stiffness matrices in each loading step, due to all possible combinations of loading and unloading).

Hypothesis V. - The volumetric, deviatoric and shear responses on each microplane are mutually independent. This of course greatly simplifies data fitting and was shown to suffice to fit each data.

(b) Microplane Stress-Strain Relations and Damage Rule.- In this work we adopt the stress-strain relations of Bažant and Prat (1988), with only some minor simplifications. According to the foregoing hypotheses, the behavior on the microplane level can be described by nontensorial path-independent total stress-strain relations of the form: $\sigma_V = F_V(\varepsilon_V)$, $\sigma_D = F_D(\varepsilon_D)$, $\sigma_T = F_T(\varepsilon_T)$. For two reasons, namely representation of unloading and application of the nonlocal damage concept, it is convenient to cast the total stress-strain relation in the form of continuum damage mechanics:

$$\sigma_V = C_V\varepsilon_V, \quad \sigma_D = C_D\varepsilon_D, \quad \sigma_T = C_T\varepsilon_T \tag{3}$$

in which, except for volumetric compression,

$$C_V = C_V^0(1-\omega_V), \quad C_D = C_D^0(1-\omega_D), \quad C_T = C_T(1-\omega_T). \tag{4}$$

The last relations are co-opted from continuum damage mechanics, whose application on the microplane level is simpler than on the macro-continuum level because all the variables on the microplane are scalar. C_V, C_D, C_T represent the secant moduli which are variable; $C_V = F_V(\varepsilon_V)/\varepsilon_V$, $C_D = F_D(\varepsilon_D)/\varepsilon_D$, $C_T = F_T(\varepsilon_T)/\varepsilon_T$; C_V^0, C_D^0, C_T^0 are the initial values of C_V, C_D, C_T corresponding to elastic behavior; and ω_V, ω_D, ω_T are the volumetric, deviatoric and shear damage on the microplane level. Fitting of selective test data by Bažant and Prat led to the following approximation for virgin loading:

$$\text{for } \varepsilon_V \geq 0: \quad \omega_V = 1 - \left(\exp - \left|\frac{\varepsilon_V}{e_1}\right|^m\right) \tag{5}$$

$$\text{for } \varepsilon_D \geq 0: \quad \omega_D = 1 - \exp\left(-\left|\frac{\varepsilon_D}{e_1}\right|^m\right) \tag{6a}$$

$$\text{for } \varepsilon_D < 0: \quad \omega_D = 1 - \exp\left(-\left|\frac{\varepsilon_D}{e_2}\right|^n\right) \tag{6b}$$

$$\omega_T = 1 - \exp\left(-\left|\frac{\varepsilon_T}{e_3}\right|^k\right) \tag{7}$$

where e_1, e_2, e_3, m, n, k are empirical material constants. Note that ε_T cannot be negative since it represents a magnitude (description of shear by means of shear strain magnitude ε_T is made possible by hypothesis II). In the volumetric behavior there is no damage ($\omega_V = 0$) and the response for virgin loading is described by

$$\text{for } \varepsilon_V < 0: \quad C_V = C_V^0[(1 + |\varepsilon_V/a|)^{-p} + |\varepsilon_V/b|^q] \tag{8}$$

where a, b, p, q are empirical constants.

For the description of unloading and reloading, we introduce a simpler rule than that in Bažant and Prat (1988), assuming that the tangential moduli for unloading or reloading are equal to the initial elastic moduli. This simplification, of course, would not suffice for a good description of large unloading cycles but seems adequate to represent the moderate degree of unloading that occurs on some microplanes during strain softening.

Virgin loading for ε_V occurs when $\varepsilon_V \Delta\varepsilon_V \geq 0$ and $(\varepsilon_V - \varepsilon_V^{max})(\varepsilon_V - \varepsilon_V^{min}) \geq 0$ where ε_V^{max}, ε_V^{min} are the maximum and minimum values of ε_V that have occurred so far; otherwise unloading or reloading takes place. The loading criteria for ε_D and ε_T are analogous. The increment of ε_T must for this purpose be defined as $\Delta\varepsilon_T = |\underline{\varepsilon}_T + \Delta\underline{\varepsilon}_T| - |\underline{\varepsilon}_T|$, not as $|\Delta\varepsilon_T|$ (which is always positive).

(c) Incremental Macroscopic Stress-Strain Relations. - For incremental solutions with incremental loading, the total stress-strain relations need to be differentiated to yield incremental relations: $d\sigma_V = C_V d\varepsilon_V + \varepsilon_V dC_V$, $d\sigma_D = C_D d\varepsilon_D + \varepsilon_D dC_D$, $d\sigma_T = C_T\, d\varepsilon_T + \varepsilon_T dC_T$. For iterative solutions, it is convenient to introduce incremental moduli $\hat{C}_V$, $\hat{C}_D$, $\hat{C}_T$ which may be equal or larger than C_V, C_D, C_T. The foregoing incremental relation may be rewritten in the form:

$$\Delta\sigma_V = C_V\Delta\varepsilon_V - \Delta\sigma_V''; \quad \Delta\sigma_D = \hat{C}_D\Delta\varepsilon_D - \Delta\sigma_D''; \quad \Delta\sigma_T = \hat{C}_T\Delta\varepsilon_T - \Delta\sigma_T'' \tag{9}$$

in which

$$\Delta\sigma_V'' = -\varepsilon_V \Delta C_V + (\hat{C}_V - C_V)\Delta\varepsilon_V; \quad \Delta\sigma_D'' = -\varepsilon_D \Delta C_D + (\hat{C}_D - C_D)\Delta\varepsilon_D; \quad \Delta\sigma_T'' = -\varepsilon_T \Delta C_T + (\hat{C}_T - C_T)\Delta\varepsilon_T. \tag{10}$$

(4)

$\Delta\sigma_V''$, $\Delta\sigma_D''$, $\Delta\sigma_T''$ are formally treated as inelastic stress increments in a step-by-step iterative solution.

As shown in the previous work, stress equilibrium between the micro- and macro-levels may be approximately enforced by the virtual work equation:

$$\frac{4\pi}{3}\Delta\sigma_{ij}\delta\varepsilon_{ij} = 2\int_S (\Delta\sigma_N \delta\varepsilon_N + \Delta\sigma_{Tr}\delta\varepsilon_{Tr})F(\vec{n})dS \tag{11}$$

where $\vec{n}$ are the unit vectors normal to the microplanes; $\delta\varepsilon_{ij}$, $\delta\varepsilon_N$ and $\delta\varepsilon_{Tr}$ represent the small variations on the macro- and micro-levels. The left-hand side of Eq. 11 represents the macroscopic work in a unit sphere of material, while the right-hand side represents the microscopic work over the surface of the same sphere. The factor 2 is due to the fact that the integral needs to extend only over a hemisphere surface, S. $F(\vec{n})$ is a weighting function of the normal directions $\vec{n}$ which can introduce anisotropy of the material in its initial state. We assume $F(\vec{n}) = 1$, which implies isotropy. Substituting Eqs. 1, 7 and 9, one can get from Eq. 11 the incremental macroscopic stress-strain relation:

$$\Delta\sigma_{ij} = C_{ijkm}\Delta\varepsilon_{km} - \Delta\sigma_{ij}'' \tag{12}$$

(5)

in which

$$C_{ijkm} = \frac{3\pi}{2}\int_S [(\hat{C}_D - \hat{C}_T)n_i n_j n_k n_m + \tfrac{1}{3}(\hat{C}_V - \hat{C}_D)n_i n_j \delta_{km} + \tfrac{1}{4}\hat{C}_T(n_i n_k \delta_{jm} + n_i n_m \delta_{jk} + n_i n_k \delta_{im} + n_j n_m \delta_{ik})]F(\vec{n})dS \tag{13}$$

(6)

$$\Delta\sigma_{ij}'' = \frac{3\pi}{2}\int_S [n_i n_j(\Delta\sigma_V'' + \Delta\sigma_D'') + \tfrac{1}{2}(n_i\delta_{rj} + n_j\delta_{ri} - 2n_i n_j n_r)\Delta\sigma_{Tr}'']F(\vec{n})dS. \tag{14}$$

(7)

C_{ijkm} is the macroscopic incremental material stiffness tensor, and $\Delta\sigma_{ij}''$ are the associated macroscopic inelastic stress increments. To make the calculations efficient, the tensorial coefficients in Eqs. 13 and 14, such as $n_i n_j n_k n_m$, $n_j n_k \delta_{im}$, are generated only once at the beginning of the analysis and are stored in the computer memory.

Tensor C_{ijkm} can have different values depending on the choice of $\hat{C}_V$, $\hat{C}_D$, $\hat{C}_T$. There are two basic choices:

1. One choice is to set $\hat{C}_V = C_V$, $\hat{C}_D = C_D$, $\hat{C}_T = C_T$ for the integration points that are loading, and $\hat{C}_V = C_V^0$, $\hat{C}_D = C_D^0$, $\hat{C}_T = C_T^0$ for those that are unloading or reloading. The resulting tensor C_{ijkm} changes from step to step and is anisotropic. It is also nonsymmetric, except when $\hat{C}_V = \hat{C}_D$. In the case there is no unloading, the resulting C_{ijkm} represent the secant stiffness tensor.

2. Another choice is to set $\hat{C}_V = C_V^0$, $\hat{C}_D = C_D^0$, $\hat{C}_T = C_T^0$ for both loading and unloading-reloading. In this case C_{ijkm} is always symmetric and equal to the initial stiffness tensor C_{ijkm}, which is here considered to be isotropic; $C_{ijkm} = (K - 2G/3)\delta_{ij}\delta_{km} + G(\delta_{ik}\delta_{jm} + \delta_{im}\delta_{jk})$ where K, G = initial elastic bulk and shear moduli; $K = E/3(1 - 2\nu)$ where E = Young's modulus, ν = Poisson ratio.

The computational efficiency is usually higher for the second choice because of the following advantages: Tensor C_{ijkm} is always the same, and so the structural stiffness matrix need not be recalculated at each iteration of each loading step. All the inelastic effects are then represented by $\Delta\sigma''_{ij}$, a tensor of fewer components than C_{ijkm}. The iterative procedure in this case coincides with the well-known initial stiffness method (e.g., Owen and Hinton, 1980). When we discuss nonlocal generalization we will see another advantage: tensor C_{ijkm} is always local and symmetric.

The first choice usually increases computational requirements since C_{ijkm} need to be recalculated in each loading step. However, the iterative process generally converges faster. In the interest of stability of computations it is advisable to keep the values of C_{ijkm} for all the iterations of the same step the same as in the first iteration, in which C_{ijkm} are determined on the basis of the stresses and strains solved in the previous load step.

There are also instances where it is necessary to calculate tangential stiffness tensor C^t_{ijkm}, i.e., the stiffness tensor providing the relation $d\sigma_{ij} = C^t_{ijkm} d\varepsilon_{km}$. For example, in plane stress analysis (in plane x_1, x_2) the stiffness coefficients C^t_{33km} are needed to ensure that the transverse stresses σ_{33} vanish.

These coefficients can be calculated as the σ_{33}-values when each of ε_{ij} is 1, while all the others are 0 (and $\Delta\sigma_{ij} = 0$). This procedure is of course equivalent to calculating C_{ijkm} from Eq. 12 with $\hat{C}_V = C_V^t$, $\hat{C}_D = C_D^t$, $\hat{C}_T = C_T^t$, where for virgin loading $C_V^t = dF_V/d\varepsilon_V$, $C_D^t = dF_D/d\varepsilon_D$, $C_T^t = dF_T/d\varepsilon_T$ (tangential moduli for the microplane), and $\hat{C}_V = C_V^0$, $\hat{C}_D = C_D^0$, $\hat{C}_T = C_T^0$ for unloading and reloading.

Another case where C_{ijkm}^t needs to be calculated is when the equilibrium response path of the structure is suspected to exhibit bifurcations. Contrary to widespread opinion, stability of iterations does not guarantee that bifurcations are absent or that stable post-bifurcation path is followed. Criteria involving C_{ijkm}^t and the lowest eigenvalue of the associated tangential stiffness of the structure need to be checked (Bažant, 1988a,b).

(d) Material Parameters and Poisson's Ratio. - Matching of Eq. 11 to Hooke's law yields the Young's modulus and Poisson ratio (Bažant and Prat, 1988):

$$E = (1-2\nu)C_V^0, \quad \nu = \frac{5 - 2\eta - 3\xi}{10 + 2\eta + 3\xi}, \quad \eta = \frac{C_D^0}{C_V^0}, \quad \xi = \frac{C_T^0}{C_V^0}. \tag{15}$$

Inverting these relations, one gets:

$$\frac{C_T^0}{C_V^0} = \frac{10(1-2\nu)}{9(1+\nu)}\,\eta. \tag{16}$$

If $\eta \to 0$ and $\xi \to 0$, then $\nu \to 0.5$; and if $\eta \to \infty$ or $\xi \to \infty$, or both, then $\nu \to -1$. We see that ν satisfies the well-known thermodynamic restriction, namely $-1 < \nu < 0.5$, and that the entire range of ν can be obtained with the present microplane model (this was not the case for the previous microplane models in which the idea of splitting ε_N into ε_V and ε_D did not exist). In the fitting of test data, it is convenient to take the given (measured) values of E (Young's modulus) and ν (Poisson's ratio) as the basic parameters. The value of the ratio $\eta = C_D^0/C_V^0$ may then be chosen as any real positive number. Then Eq. 15 yields $C_V^0 = E/(1-2\nu)$, and Eq. 16 provides

$$C_T^0 = \frac{1}{3}\left(\frac{5(1-2\nu)}{1+\nu} - 2\eta\right)C_V^0. \tag{17}$$

Bažant and Prat (1988) found that a good choice for fitting of all the test data on concretes is $\eta = 1$. The initial Poisson's ratio (ν) and Young's modulus (E) for

concrete is known or determined in advance, and so parameters a, b, p, q, e_1, e_2, e_3, m, n, k have to be found by data fitting. Parameters a, b, p and q can be found independently of the others by fitting hydrostatic pressure test data for compression ($\varepsilon_V < 0$). The exponents m, n and k can be taken as constants. Therefore, only three free parameters (e_1, e_2 and e_3) have to be found by fitting triaxial test data other than the hydrostatic pressure curve. This is not a difficult task.

(e) Integration over a Hemisphere. - The microplane model trades conceptual simplicity for high computer time requirements. Thus, numerical efficiency is very important. The integrals in Eq. 13 and 14 need to be evaluated numerically. This must be done a great number of times - in every load step, in every iteration of the step, and in every integration point of every finite element. Therefore, the evaluation of these integrals over a hemisphere must be as efficient as possible. This problem has been studied in detail. Efficient Gaussian-type formulas for integration over a unit hemisphere have been identified (Stroud, 1971) and some new superior ones found (Bažant and Oh, 1985, 1986). These formulas generally approximate the integrals in the form:

$$\frac{4\pi}{3}\int_S \approx 6\sum_{\alpha=1}^{n} w_\alpha F_\alpha \tag{18}$$

in which subscript α refers to a certain discrete set of microplanes characterized by the spatial discretisation of their normals associated with points on a unit hemisphere, and w_α are the weights (numerical integration coefficients) for these directions. In the present study, Bažant and Oh's (1985, 1986) 21-point integration formula is used (Fig. 1b). Any known formula with a lesser number of points does not give sufficient accuracy in the strain-softening regime. When there are symmetries (plane strain, axisymmetry) the number of integration points can of course be reduced; e.g., for plane strain or stress, the 21-point formula reduces to a 4-point formula.

EXPONENTIAL ALGORITHM FOR LOAD STEPS

In every iteration of r-th load or time step, we can use the known macrostrains $\varepsilon_{ij(r)}$ and their known increments $\Delta\varepsilon_{ij(r)}$ to calculate the strains and strain incre-

ments on each microplane based on the kinematic constraint (Eq. 1). Then we can use the known values of $\varepsilon_N = \varepsilon_V + \varepsilon_D$, ε_{Ti}, $\Delta\varepsilon_N = \Delta\varepsilon_V + \Delta\varepsilon_D$, and $\Delta\varepsilon_{Ti}$, to calculate the stresses on each microplane by solving the differential equations 9. Each of these equations could be solved by using a central difference approximation. However, such an approximation often appears unstable when the stress-strain relation has a negative slope (strain softening), and, even if the computations remain stable, a large error is usually accumulated. The result is that the stress is not reduced exactly to zero at very large strains.

These drawbacks can be eliminated by the so-called exponential algorithm, initially developed for aging creep of concrete (Bažant 1971, Bažant and Wu 1973) as a refinement of the viscoelastic algorithms of Zienkiewicz, Watson and King (197) and Taylor, Pister and Goudreau (197). Later, Bažant and Chern (1985) extended this algorithm to creep with strain softening. We will now adapt their exponential algorithm to the microplane model.

The basic idea of the exponential algorithm is that the integration formula is the exact solution of the differential equation obtained for the current loading step under the assumption that the material properties, the loads and the prescribed inelastic strain rates are constant in time. To develop the formula of exponential algorithm for our case, we proceed similarly to Bažant and Chern (1985), using analogy with Maxwell's spring-dashpot model. The response in each step is then similar to stress relaxation, and the advantage is that stress relaxation always ultimately yields a zero stress value if the load step is very long. Even though our problem is time-independent, we associate with the load steps a fictitious time t. Then, for each microplane and for both (normal and tangential) directions, it is possible to rewrite the stress-strain relation so that it formally looks like the equation for Maxwell's model. Since the relations for ε_V, ε_D, and ε_T are the same, we now drop subscripts V, D and T. From $\sigma = C\varepsilon$ we have $\dot{\sigma} = C\dot{\varepsilon} + \dot{C}\varepsilon$, i.e.

$$\dot{\sigma} + \frac{\sigma}{E} = C_a\dot{\varepsilon} \tag{19}$$

where the superior dots denote time derivatives, $E = -C_a/\dot{C}$, $\dot{C} = (dC/d\varepsilon)\dot{\varepsilon}$, $\dot{\varepsilon} = \Delta\varepsilon/\Delta t$ and $C_a = (C_r + C_{r+1})/2$ for time step number r. In every time step $\Delta t_r = t_{r+1} - t_r$. E can be approximated as a constant, and then the exact solution of Eq. 18 is of the form

$$\sigma(t) = A\, e^{-(t - t_r)/E} + C_a E \dot{\varepsilon} \tag{20}$$

where A is an integration constant. From the initial condition $\sigma = \sigma_r$ at $t = t_r$, it follows that

$$\sigma(t) = \sigma_r\, e^{-(t - t_r)/E} + (1 - e^{-(t - t_r)/E}) C_a E\, \Delta\varepsilon. \tag{21}$$

For the end of the time step, $t = t_{r+1} = t_r + \Delta t$, we get

$$\sigma_{r+1} = \sigma_r + \Delta\sigma = \sigma_r\, e^{-\Delta z} + (1 - e^{-\Delta z}) C_a \frac{\Delta\varepsilon}{\Delta z} \tag{22}$$

in which $\Delta z = \Delta t/E = -\Delta C/C_a$. Eq. 22 can be rewritten in the form of a pseudo-elastic stress-strain relation on the level of each microplane stress component

$$\Delta\sigma = D\, \Delta\varepsilon - \Delta\sigma'' \tag{23}$$

where $D = C_a(1 - e^{-\Delta z})/\Delta z$, $\Delta\sigma'' = D\, \Delta\varepsilon''$ and $\Delta\varepsilon'' = (1 - e^{-\Delta z})\sigma_r/D$; $D\, \Delta\varepsilon = \Delta\sigma_{el}$, $\Delta\sigma_{el}$ and $\Delta\sigma''$ are the elastic and inelastic stress increments.

This algorithm, as well as the similar previous ones, is called exponential because its formula typically involves an exponential function. Attaching subscripts V, D and T, one may use the foregoing formulas to determine the stress increments on each microplane and in both (normal and tangential) directions. The stress increment on the macro-level is then determined using numerical integration over the unit hemisphere (with 21 integration points), as already described. Numerical experience with the present algorithm indicated excellent stability and good accuracy, as checked by time step refinement.

The present integration scheme has also the advantage that any creep law for concrete can be easily incorporated into the microplane model because numerical integration for the creep law of concrete is similar and entails the same kind of numerical stability problems as in the case of strain softening.

NONLOCAL GENERALIZATION OF MICROPLANE MODEL

Macroscopic behavior of concrete requires a constitutive model which exhibits strain softening, i.e. a decrease of stress at increasing strain, or in general a loss of positive definiteness of the tangential moduli (stiffness) matrix of the material. In concrete, strain softening is due to degradation of material stiffness caused by microcracking. In classical, local continuum analysis by finite element method, strain localization engenders problems of instability, inobjectivity and spurious mesh sensitivity (Bažant, 1976, 1986; Bažant and Belytschko, 1987; Bažant, Belytschko and Chang, 1984). These problems can be avoided by the use of nonlocal continuum.

Nonlocal continuum is a concept in which the stress at a point depends not only on the strain at the same point (local dependence) but also on the strains within a certain neighborhood of the point; in particular, on a certain weighted spatial average of the strains from the neighborhood. This concept, introduced in elasticity by Kröner (1968), Eringen and Edelen (1972), Krumhansl (1968), Levin (1971) and others, has been applied to strain softening continuum by Bažant, Belytschko and Chang (1984). However, the early form, called the imbricate continuum, proved too cumbersome.

A more effective form, in which only the variables that are associated with strain softening are nonlocal and all the other variables, especially the strain as a kinematic variable and the elastic strain, are local, was proposed by Bažant and Pijaudier-Cabot (1987) and Pijaudier-Cabot and Bažant (1987) and further developed in Bažant and Pijaudier-Cabot (1988) and Bažant and Lin (1988). In this formulation, called the nonlocal damage theory, the nonlocal (spatially averaged) quantity is the energy dissipated due to the strain softening. An important advantage of this formulation is that, in contrast to the classical, fully nonlocal continuum theory, the equations of equilibrium and the boundary conditions are of the same form as in the local continuum theory, which is a useful feature for finite element programming. Another advantage over the fully nonlocal continuum is that

no zero-energy periodic modes of instability exist.

In the case of the present microplane model. the variables that control strain softening are the damage variables ω_V, ω_D and ω_T. These variables are calculated from macrostrain ε_{ij} by means of the microplane strains components ε_V, ε_D and ε_T as indicated in Eqs. 4-7. The nonlocal generalization is achieved merely by replacing ω_V, ω_D and ω_T in Eqs. 4-7 with nonlocal damage variables $\bar{\omega}_V$, $\bar{\omega}_D$ and $\bar{\omega}_T$,which are calculated from nonlocal macrostrains $\bar{\varepsilon}_{ij}$ by means of the resolved microplane components of nonlocal strain, $\bar{\varepsilon}_V$, $\bar{\varepsilon}_D$ and $\bar{\varepsilon}_T$. For the load increments in which unloading or reloading takes place, the local formulation is used, i.e., the strains ε_V, ε_D, ε_T in the relations $\sigma_V = C_V\varepsilon_V$, $\sigma_D = C_D\varepsilon_D$, and $\sigma_T = C_T\varepsilon_T$ are local, and so is ε_{ij} in Eqs. 12-14.

The nonlocal macrostrain $\bar{\varepsilon}_{ij}$, which is needed at each integration point of each finite element in each loading step, is calculated from the spatial averaging integral

$$\bar{\varepsilon}_{ij}(\underset{\sim}{x}) = \frac{1}{V_r(\underset{\sim}{x})}\int_V \alpha(\underset{\sim}{s}-\underset{\sim}{x})\varepsilon_{ij}(\underset{\sim}{s})d\underset{\sim}{x} = \int_V \alpha'(\underset{\sim}{s},\underset{\sim}{x})\ \varepsilon_{ij}(\underset{\sim}{s})d\underset{\sim}{x}. \tag{24}$$

The superimposed bar denotes the averaging operator; $\underset{\sim}{x}$, $\underset{\sim}{s}$ = coordinate vectors, $\alpha(\underset{\sim}{x})$ = given empirical weighting function, V = volume of the entire structure, and

$$V_r(\underset{\sim}{x}) = \int_V \alpha(\underset{\sim}{s}-\underset{\sim}{x})dV(\underset{\sim}{s}), \quad \alpha'(\underset{\sim}{s},\underset{\sim}{x}) = \frac{\alpha(\underset{\sim}{s}-\underset{\sim}{x})}{V_r(\underset{\sim}{x})}. \tag{25}$$

In numerical computation, the spatial averaging integral is approximated by a finite sum over all the integration points (whose number is denoted as N) of all the finite elements of the structure. For convenience of programming, the integration points of all elements are included in the sum even though for some points the value of $\alpha'(\underset{\sim}{s}-\underset{\sim}{x})$ may be zero. The values of $\alpha'(\underset{\sim}{s}-\underset{\sim}{x})$ are generated in advance and stored as an N × N square matrix. $V_r(\underset{\sim}{x})$ has approximately the same meaning as the representative volume in the statistical theory of heterogeneous materials (Krǒner, 1968). When the averaging volume protrudes outside the boundary of the body, the points outside the body must be chopped off and the weights scaled up from α to α' so that their sum over the points inside the body would be equal to the original (unchopped) averaging volume. Due to this adjustment, α' depends separately on $\underset{\sim}{x}$

and $\underset{\sim}{s}$ whereas α depends only on the distance $|\underset{\sim}{s} - \underset{\sim}{x}|$. The scaling from α to α' causes the matrix of weights α' to become nonsymmetric.

Although a weighting function that is uniform over a certain finite representative volume could be used, it is computationally more efficient (and probably more realistic) to use a smooth weighting function of a bell shape. The existing experimental data on the consequences of nonlocal behavior, such as the size effect, are at present insufficient to decide which shape of function $\alpha(\underset{\sim}{x})$ is more correct, partly because of limited scope of the data, partly because of their random scatter. The initial studies of Bažant and Pijaudier (1987) and Bažant and Lin (1988) used for weighting function α the Gaussian distribution function. This function extends with nonzero values to infinity, although its values are negligible beyond a certain distance from point $\underset{\sim}{x}$ under consideration. Recently Bažant proposed for the spatial averaging the following bell-shaped function (Fig. 1c) which vanishes for distances greater than R:

$$\alpha(\underset{\sim}{x}) = [1 - (\rho/\rho_1)^2]^2 \quad \text{for } \rho < \rho_1; \quad \alpha = 0 \quad \text{for } \rho \geq \rho_1 \tag{26}$$

where $\rho = r/k\ell$, $\rho_1 = R/k\ell$, $\underset{\sim}{x} = x$ or (x,y) or (x,y,z) for one, two or three dimensions (1D,2D,3D); x,y,z = cartesian coordinates; and

$$\begin{aligned} &\text{for 1D:} \quad r^2 = x^2, \quad k = 15/16 = 0.9375 \\ &\text{for 2D:} \quad r^2 = (x^2 + y^2), \quad k = \sqrt[3]{3/4} = 0.9086 \\ &\text{for 3D:} \quad r^2 = (x^2 + y^2 + z^2), \quad k = (105/192)^{1/3} = 0.8178. \end{aligned} \tag{27}$$

The above values of k were determined from the condition that volume under function $\alpha(\underset{\sim}{x})$ should be equal to the volume under function $\bar{\alpha}(\underset{\sim}{x})$ that is uniform, $\bar{\alpha}(\underset{\sim}{x}) = 1$, for $r < \ell$, with $\bar{\alpha}(\underset{\sim}{x}) = 0$ for $r \geq \ell$, where ℓ represents the characteristic length of nonlocal continuum (which gives the size of the representative volume).

A point to note is that the assembled structural stiffness matrices of finite element systems for the nonlocal continuum with local strains are in general nonsymmetric (Bažant and Pijaudier-Cabot, 1988). The nonsymmetry has two reasons: (1) the fact that, for points close to the boundary, the averaging volume of size R protrudes outside the boundary of the body, and (2) the fact that the averaging

volume for a point that is loading (or unloading) may comprise other points that are unloading (or loading). The consequence is that: (1) if point m is loading and point n is unloading, the strain at point n affects the response at point m (because the response at point n is local and at point m is nonlocal); (2) if point m is outside and point n inside a boundary layer of thickness R, the influence of strain at n on the response at m is stronger than the influence of strain at m on the response at n (because the size of the averaging volume for point n is smaller than that for point m, due to truncation of the averaging volume that protrudes outside the boundary). For an infinite body in which no point is unloading, the stiffness matrices are symmetric. Compared to the Gaussian distribution function, which is nonzero everywhere in the body, the weighting function in Eq. 24 mitigates the incidence of nonsymmetry because the weighting function is nonzero only in a limited volume.

It may be noted that the original, fully nonlocal (imbricate) model was giving symmetric stiffness matrices. In fact, it was shown (Bažant, 1984) that the assumption of symmetry in the derivation of the field equation from the virtual work equation inevitably leads to the imbricate model. However, this model is too cumbersome to program and suffers by zero-energy periodic modes of instability which must be artificially suppressed. The nonsymmetry of the tangential stiffness matrix in the present nonlocal formulation was initially deemed suspect, but after considerable numerical experience (e.g. Bažant and Lin, 1988a,b) it does not appear to cause any problems. The reason probably consists in the fact that the nonsymmetry arises only from damage, a phenomenon for which physics actually imposes no symmetry requirements.

In the present calculations, nonsymmetry of the structural stiffness matrix was altogether avoided by always using for the iterations the initial elastic stiffness matrix. Thus, all the symmetry-breaking nonlocal response was incorporated into the inelastic nodal forces.

ITERATIVE NUMERICAL ALGORITHM IN EACH LOAD STEP

1. Before the first load step, evaluate and store the elastic structural stiffness matrix $\underset{\sim}{K}$. Also, calculate the integration formula coefficients used in the microplane model and the values of the nonlocal averaging function α', and store them all in the computer memory.

2. For each microplane at each integration point of each element, use the mean (midstep) strains and stresses from the previous iteration or, in the case of the first iteration, the final strains and stresses from the previous load step, to decide whether there is virgin loading, unloading or reloading and to evaluate the material parameters for each microplane. If at least one microplane at a certain point undergoes virgin loading, calculate for that point the nonlocal macrostrains $\bar{\varepsilon}_{ij}$. For each microplane, use $\bar{\varepsilon}_{ij}$ to calculate $\bar{\varepsilon}_V$, $\bar{\varepsilon}_D$, $\bar{\varepsilon}_T$ if there is virgin loading or ε_{ij} to calculate ε_V, ε_D, ε_T if there is unloading or reloading. Then, assuming that that $\hat{C}_V = C_V^0$, $\hat{C}_D = C_D^0$, $\hat{C}_T = C_T^0$, calculate $\Delta\sigma''_{ij}$ for each integration point of each element, and then calculate the column matrix $\Delta\underset{\sim}{f}''$ of inelastic nodal force increments. Adding to $\Delta\underset{\sim}{f}''$ the matrix of nodal force increments $\Delta\underset{\sim}{f}^L$ due to prescribed load increments, obtain the column matrix of the total nodal force increments, $\Delta\underset{\sim}{f}$, and implement the condition of prescribed displacement increments.

3. Except in the first iteration, use the values of the column matrix $\Delta\underset{\sim}{u}$ of nodal displacement increments from the previous iteration to calculate the residual forces $\Delta\underset{\sim}{f}^r = \underset{\sim}{K}\Delta\underset{\sim}{u} - \Delta\underset{\sim}{f}$ where $\underset{\sim}{K}$ = initial elastic stiffness matrix. If $\|\Delta\underset{\sim}{f}^r\|$ is less than the prescribed tolerance, return to 2 and start the first iteration of the next loading step.

4. Solve the linear equation system $\underset{\sim}{K}\Delta\underset{\sim}{u} = \Delta\underset{\sim}{f}$ where $\underset{\sim}{K}$ = initial elastic stiffness matrix, which is the same in every load step and every iteration. From displacement increments $\Delta\underset{\sim}{u}$ in every load step and every iteration, calculate the local strain increments $\Delta\varepsilon_{ij}$ for all the integration points of all elements. Then, after evaluating $\Delta\varepsilon_V$, $\Delta\varepsilon_D$, $\Delta\varepsilon_T$ and deciding whether there is virgin loading, unloading or reloading, calculate the increments $\Delta\sigma_V$, $\Delta\sigma_D$, $\Delta\sigma_T$ for each microplane at each inte-

gration point of each element and add them to the final stresses from the previous load step to get the final microplane stress σ_V, σ_D, σ_T for this load step. Return to 2 and start the next iteration of the same loading step.

Numerical finite element analysis using the present nonlocal formulation was found to exhibit good convergence, excellent stability and no spurious mesh sensitivity at mesh refinements, even in the strain-softening regime. The rate of convergence of the present nonlocal analysis in the softening range was better than for a corresponding local formulation obtained by deleting the spatial averaging integrals. Actually, despite the computer time needed to calculate the spatial averages, the overall running time was faster than for the corresponding local finite element code provided the number of elements was large. This agrees with the experience of Bažant and Lin (1988) from nonlocal solutions with over 3000 degrees of freedom.

NUMERICAL STUDIES AND SIZE EFFECT

Two examples of finite element analysis using nonlocal microplane models will now be presented. In both, calculations have been made using four-node isoparametric finite elements, with one integration point in example 1, and four integration points in example 2.

Example 1. - Concrete specimen with depth-length ratio d/h = 600/1400 mm, and thickness b = 100 mm is loaded in tension (Fig. 2) in the longitudinal direction. Only one quarter of the specimen is modeled, and plane strain state is assumed. Strain localization is initiated using weaker elements, as shown by the shaded finite elements in Fig. 2, by decreasing initial modulus of elasticity by 5%. The microplane model parameters for concrete are taken as follows: $a = 0.005$, $b = 0.035$, $p = 1.0$, $q = 1.85$, $e_1 = 0.00006$, $e_2 = 0.0004$, $e_3 = 0.0004$, $m = 1.2$, $n = 1.1$, $k = 1.1$, $E_0 = 30{,}000$ N/mm^2, and $\nu = 0.18$. Using these parameters, the uniaxial tensile strength of concrete is calculated as $f_t' = 2.05$ N/mm^2. In nonlocal analysis, the characteristic length is assumed to be $\ell = 150$ mm.

To demonstrate mesh sensitivity in the softening regime, two different finite

element meshes are generated (Fig. 2), the first one with 21 and the second one with 84 finite elements. The results of the analysis using local continuum (Fig. 3a) indicate strong mesh sensitivity. Fig. 3b shows the results of the analysis using nonlocal continuum, and it is evident that there is no mesh sensitivity.

To demonstrate stability of the exponential algorithm, the results for different load steps are plotted in Fig. 3c (for a mesh of 21 finite elements, nonlocal). For case II the load increment is twice as large as for case I, but the results are still in good agreement.

Example 2. - To demonstrate that the microplane model using nonlocal approach correctly describes the fracture of concrete and especially the size effect, three-point-bend specimens are considered. Three different specimen sizes of size ratios 1:2:4, with geometrically similar shapes, were used. The finite element meshes and the geometry of the specimens are shown in Fig. 6. The depth of the notch was always 1/6 of the depth of the beam. This type of specimen was tested by Bažant and Pfeiffer (1987) using concrete with maximum aggregate size $d_a = 12.7$ mm; the depth of the smallest specimen was $d = 7.62$ cm. The characteristic length is taken as $\ell = 3d_a$. The microplane model parameters used in the finite element analysis are partly adjusted by fitting test data for the smallest test specimen, and are as follows: $a = 0.005$, $b = 0.035$, $p = 1.0$, $q = 1.85$, $e_1 = 0.000075$, $e_2 = 0.0012$, $e_3 = 0.0012$, $n = 0.5$, $m = 1.2$, $k = 1.2$, $E_0 = 35{,}000$ N/mm^2 and $\nu = 0.18$. From these parameters, the tensile strength of concrete is calculated as $f_t' = 3.27$ N/mm^2 (Fig. 4).

The maximum loads P, which were previously measured by Bažant and Pfeiffer (1987), are now calculated for each specimen of each size. The values of the nominal stress at failure (σ_N) are calculated as $\sigma_N = P/bd$ where b = specimen thickness (38.1 mm) and d = depth of the specimen. The previous experimental results and the present results of numerical analysis are shown in Figs. 5a,b. The solid and dashed curves represent the optimum fit of the test results with the present model and the fits by the size effect law proposed by Bažant (1984):

$$\sigma_N = B f_t' [1 + (d/d_0)]^{-\frac{1}{2}}. \tag{28}$$

The optimum values for parameters B and d_0 are obtained by linear regression. For the previous test results, these parameters are: B = 0.448 and d_0 = 12.04 mm (Fig. 5a). Figs. 5a and 5b demonstrate a good agreement between the test results and the calculated results.

In addition, Fig. 6 shows the contours of the fracture process zone at the peak load for all the specimens, calculated from the finite element results. The radius of the fracture process zone is found to be roughly equal to the characteristic length ℓ, and to increase slightly as the specimen size increases. The figure also shows the contours of the hardening nonlinear zone around the fracture process zone and of the zone in which the stresses as a function of strains are elastic with an error less than 5% of f_t'.

CONCLUSIONS

1. As confirmed numerically, the nonlocal generalization of microplane model (1) is free of spurious mesh sensitivity and localization instabilities, (2) can be implemented in a finite element code, (3) allows an efficient solution procedure, and (4) can handle both fracture (or damage localization) and nonlinear triaxial behavior with or without strain softening.

2. The nonsymmetry of the tangential stiffness matrix, arising from both the microplane formulation and the nonlocality, seems to pose no computational problems if the load-step algorithm is based on the initial elastic stiffness matrix.

3. The idea of exponential algorithm, in which the load-step stress-strain relation is derived from a certain exact solution of the differential equations of the microplane model, appears to increase computational efficiency. Although the computer work per step becomes larger, the exponential algorithm makes it possible to take much larger loading steps. It also guarantees that the stress at the tail of strain-softening behavior gets reduced exactly to zero, regardless of previous numerical errors.

4. The microplane model which has already allowed an excellent description of

the existing test data on nonlinear triaxial behavior, as well as unidirectional and multidirectional smeared cracking, is now shown capable of modeling at the same time the tensile fracture.

5. The model is free of spurious mesh sensitivity and exhibits correct transitional size effect which was previously experimentally verified for concrete specimens or structures and was approximately described by the size effect law proposed by Bažant.

6. According to the present model, the shape and width of the fracture process zone in concrete varies, albeit mildly, with the specimen size.

7. The method of nonlocal generalization, the exponential algorithm, and the capability to describe fracture with transitional size effect is not limited to concrete but should be applicable to brittle-plastic materials in general (e.g., ceramic composites, rocks and stiff soils).

Acknowledgments. - Financial support under AFOSR contract No. F49620-87-C-0050DEF with Northwestern University is gratefully acknowledged. Further funds for the numerical algorithm studies were also obtained under NSF grant MSM-8700830. Thanks are also due to T. Hasegawa of Shimizu Institute of Technology, Tokyo, for some valuable discussions. Partial funding for the material modeling has also been obtained from the NSF Science and Technology Center on Advanced Cement-Based Materials at Northwestern University.

APPENDIX.-REFERENCES

1. Bazant, Z.P. (1971), "Numerically stable algorithm with increasing time steps for integral-type aging creep", Proc., First Int. Conf. on Struct. Mech. in Reactor Technology, West Berlin, T.A. Jaeger, ed., Vol. 4, Part H 119-126.

2. Bazant, Z.P., and Wu, S.T. (1974), "Rate-type creep law of aging concrete based on Maxwell Chain", Materials and Structures (RILEM, Paris) 7 (37), 45-60.

3. Bazant, Z.P., and Phat, P.D. (1976), "Endochronic theory of inelasticity and failure of concrete", ASCE J. of Engng. Mech. Division 102 (4), 701-722.

4. Bazant, Z.P. (1976), "Instability, ductility and size effect in strain-softening concrete", ASCE J. Engng. Mech. 102 (2), 331-334; Discussions Vol. 103, pp. 357-358, 775-777, Vol. 104, pp. 501-502.

5. Bazant, Z.P., and Kim, S.S. (1979), "Plastic-fracturing theory for concrete", ASCE J. of Engng. Mech. Div. 105 (3), 407-428.

6. Bazant, Z.P., and Oh, B.H. (1983), "Crack band theory for fracture of concrete", Materiaux et Constructions (Materials and Structures), RILEM, Paris, 16 (93), pp. 155-177

7. Bazant, Z.P. (1984), "Size effect in blunt fracture: concrete, rock, metal", ASCE J. of Engng. Mech. 110 (4), 518-535.

8. Bazant, Z.P. (1984), "Microplane model for strain-controlled inelastic behavior", Mechanics of Engineering Materials, C.S, Desai and R.H. Gallagher, eds., John Wiley & Sons, New York 45-59.

9. Bazant, Z.P. (1984) "Imbricate continuum and its variational derivation", ASCE J. of Engng. Mech. 110 (12), 1693-1712.

10. Bazant, Z.P., and Gambarova, P.G. (1984), "Crack shear in concrete: crack band microplane model'" ASCE J. of Struct. Engng., 110 (9) 2015-2035.

11. Bazant, Z.P., Belytschko, T.B., and Chang, T.P. (1984), "Continuum model for strain-softening", ASCE J. of Engng. Mech. 110 (12), 1666-1692.

12. Bazant, Z.P. (1985), "Mechanics of distributed cracking", ASME Applied Mechanics Reviews 39 (5), 675-705.

13. Bazant, Z.P., and Oh, B.H. (1985), "Microplane model for progressive fracture of concrete and rock", ASCE J. of Engng. Mech. 111 (4), 559-582.

14. Bazant, Z.P., and Chern, J.C. (1985), "Strain-softening with creep and exponential algorithm", ASCE J. Engng. Mech. 111 (3), 381-390.

15. Bazant, Z.P., and Oh, B.H. (1986), "Efficient numerical integration on the surface of a sphere", Zeitschrift fur Angewandte Mathematik und Mechanik (ZAMM) 66 (1), 37-49.

16. Bazant, Z.P., and Pfeiffer, (1987), "Determination of fracture energy from size effect and brittleness number", ACI Materials Journal 84 (Nov.) 463-480.

17. Bazant, Z.P., and Pijaudier-Cabot, G. (1987), "Modeling of distributed damage by nonlocal continuum with local strain", Preprints, Fourth Int. Conf. on Numerical Methods in Fracture Mechanics, held in San Antonio, Texas, March 1987, publ. by Pineridge Press, Swansea, U.K., A.R. Luxmore et al., 411-432.

18. Bazant, Z.P., and Lin, F.-B. (1988), "Nonlocal yield-limit degradation", Int. J. for Numerical Methods in Engineering 26, 1805-1823.

19. Bazant, Z.P., and Pijaudier-Cabot, G. (1988), "Nonlocal continuum damage, localization instability and convergence", ASME J. of Applied Mechanics 55, 287-293.

20. Bazant, Z.P., and Prat, P.C. (1988), "Microplane model for brittle plastic material: I. Theory, II. Verification", ASCE J. of Engng. Mech. 114 (1672-1702).

21. Bazant, Z.P. (1988a), "Stable states and paths of structures with plasticity or damage", ASCE J. of Engng. Mech. 114 (12), 2013-2034.

22. Bazant, Z.P. (1988b), "Stable states and stable paths of propagation of damage zones and interactive fractures in cracking and damage'", Proc., France-US Workshop held at ENS, Cachan, France, September 1988, J. Mazars and Z.P. Bazant, eds., Elsevier, London, 183-206.

23. Bazant, Z.P., and Belytschko, T. (1988), "Strain-softening continuum damage: Localization and size effect", Proc., Second Int. Conf. on Constitutive Laws of Engineering Materials: Theory and Applications, held in Tucson, AR., Jan. 1987, C.S. Desai et al., eds. Elsiviere, New York, 11-33.

24. Cedolin, L., dei Poli, S., and Iori, I. (1983), "Experimental determination of the fracture process zone in concrete", Cement and Concrete Research 13, 557-567.

25. Chen, A.C.T., and Chen, W.F. (1975), "Constitutive relations for concrete", ASCE J. of Engng. Mech. 101 (4), 465-481

26. Dougill, J.W. (1976), "Unstable progressively fracturing solids", J. of Appl. Math. and Physics (ZAMP) 27, 423-436.

27. Eringen, A.C., and Edelen, D.G.D. (1972), "On nonlocal elasticity", Int. J. of Engng. Sci. 10, 233-248.4.

28. Gerstle, K.H., et al. (1980), "Behavior of concrete under multi-axial stress states", ASCE J. of Engng. Mech. Division 106 (6), 1383-1403.

29. Gerstle, K.H. (1981), "Simple formulation of biaxial concrete behavior", ACI Journal 78 (1), 62-68; also "Simple formulation of triaxial concrete behavior", 78 (Sept.).

30. Kröner, E. (1968), "Interrelations between various branches of continuum mechanics", in Mechanics of Generalized Continua, E. Klörner, ed., Springer, W. Berlin, 330-340.

31. Krumhansl, J.A. (1968), "Some considerations of the relations between solid state physics and generalized continuum mechanics", in Mechanics of Generalized Continua, E. Kröner, ed., Springer, West Berlin, 298-331.

32. Levin, V.M. (1971), "The relation between mathematical expectation of stress and strain tensors in elastic micro-heterogenous media", Prikl. Mat. Mekh. 35, 694-701 (in Russian).

33. Lin, F.P., et al. (1987), "Concrete model with normality and sequential identification", Computers and Structures 26 (6), 1011-1026.

34. Ortiz, M.A. (1985), "A constitutive theory for the inelastic behavior of concrete", Mechanics and Materials 4, 67-93.

35. Pande, G.H., and Xiong, W. (1982), "An improved multi-laminate model of jointed rock masses", Proc., First Int. Symposium on Numerical Models in Geomechanics 218-226.

36. Pijaudier-Cabot, G., and Bazant, Z.P. (1987), "Nonlocal damage theory", ASCE J. of Engng. Mech. 113, 1512-1533.

37. Stroud, A.H. (1971). "Approximate calculation of multiple integrals", Prentice-Hall, Englewood Cliffs, NJ, 296-302.

38. Taylor, G.I. (1938), "Plastic strain in metals", Journal of Inst. of Metals 62 , 307-324.

39. Taylor, R.L., Pister, K.S., and Goudreau, G.L. (1970), "Thermo-mechanical analysis of visco-elastic solids", Int. J. of Numerical Methods in Engineering 2, 45-60.

40. Willam, K.J., and Warnke, E.P. (1974), "Constitutive model for the triaxial behavior of concrete", in "Seminar on concrete structures subjected to triaxial stresses", International Association of Bridge and Structural Engineering Conference, Bergamo, Italy, May

41. Zienkiewicz, O.C., Watson, M., and King, I.P. (1968), "A numerical method of visco-elastic stress analysis", Int. J. of Mech. Sci. 10, 807-827.

42. Zienkiewicz, O.C. and Pande, G.N. (1977), "Time-dependent multi-laminate model of rocks -- a numerical study of deformation and failure of rock masses", International Journal of Num. and Analytical Meth. in Geomechanics 1, 219-247.

Figure Captions

Fig. 1. (a) Strain Components on a Microplane, (b) Numerical Integration Points, and (c) Weighting Function for Nonlocal Averaging.

Fig. 2. Rectangular Panel in Example and Finite Element Meshes Used.

Fig. 3. Load-Displacement Curves of Panel Obtained with the Meshes Shown in Fig. 2.

Fig. 4. Three-Point Bend Fracture Specimen, Meshes Used for Its Three Different Sizes, and Uniaxial Stress-Strain Curve.

Fig. 5. Nominal Stresses σ_N at Maximum Load Obtained with the Meshes in Fig. 5 for Three Specimen Sizes d, Compared to Test Data and to Optimum Fit of Data by Size Effect Law.

Fig. 6. Fracture Process Zones in the Specimens Shown in Fig. 5.

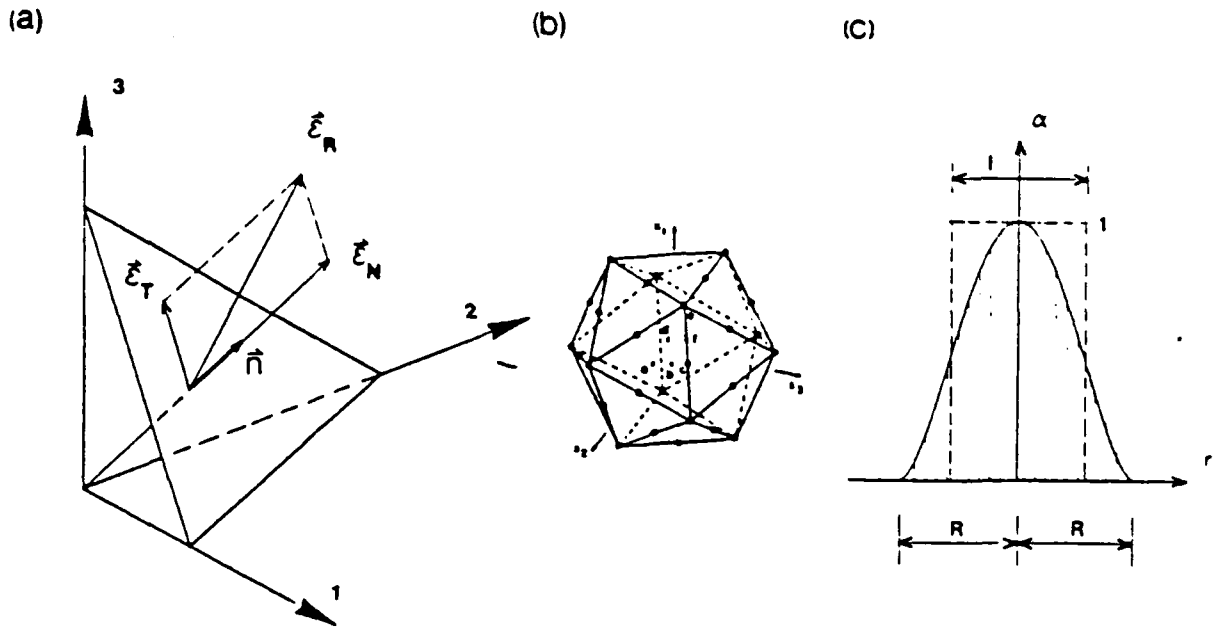

(a)
(b)
(c)
3
2
1
$\vec{\varepsilon}_R$
$\vec{\varepsilon}_N$
$\vec{\varepsilon}_T$
$\vec{n}$
α
1
1
R
R

fig.1

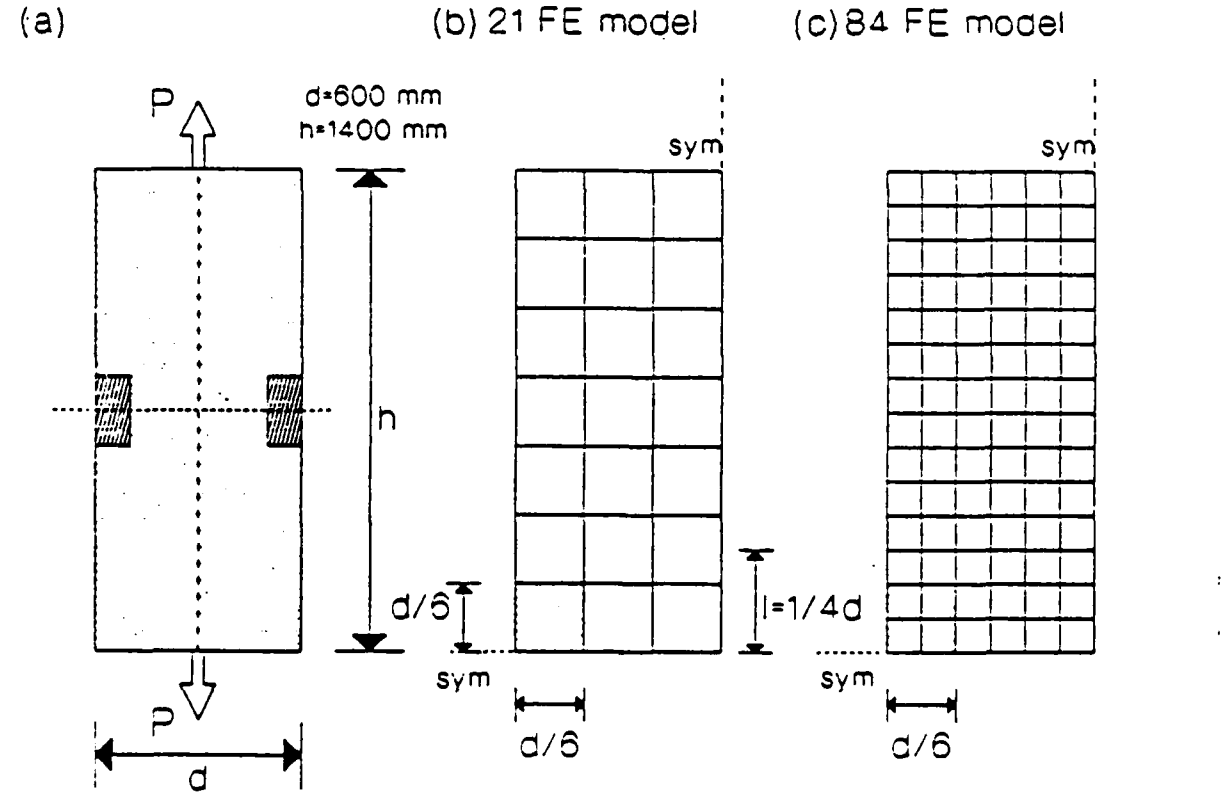

(a)
(b) 21 FE model
(c) 84 FE model
P
d=600 mm
h=1400 mm
h
d/6
P
d
sym
sym
d/6
l=1/4d
sym
d/6

fig.2

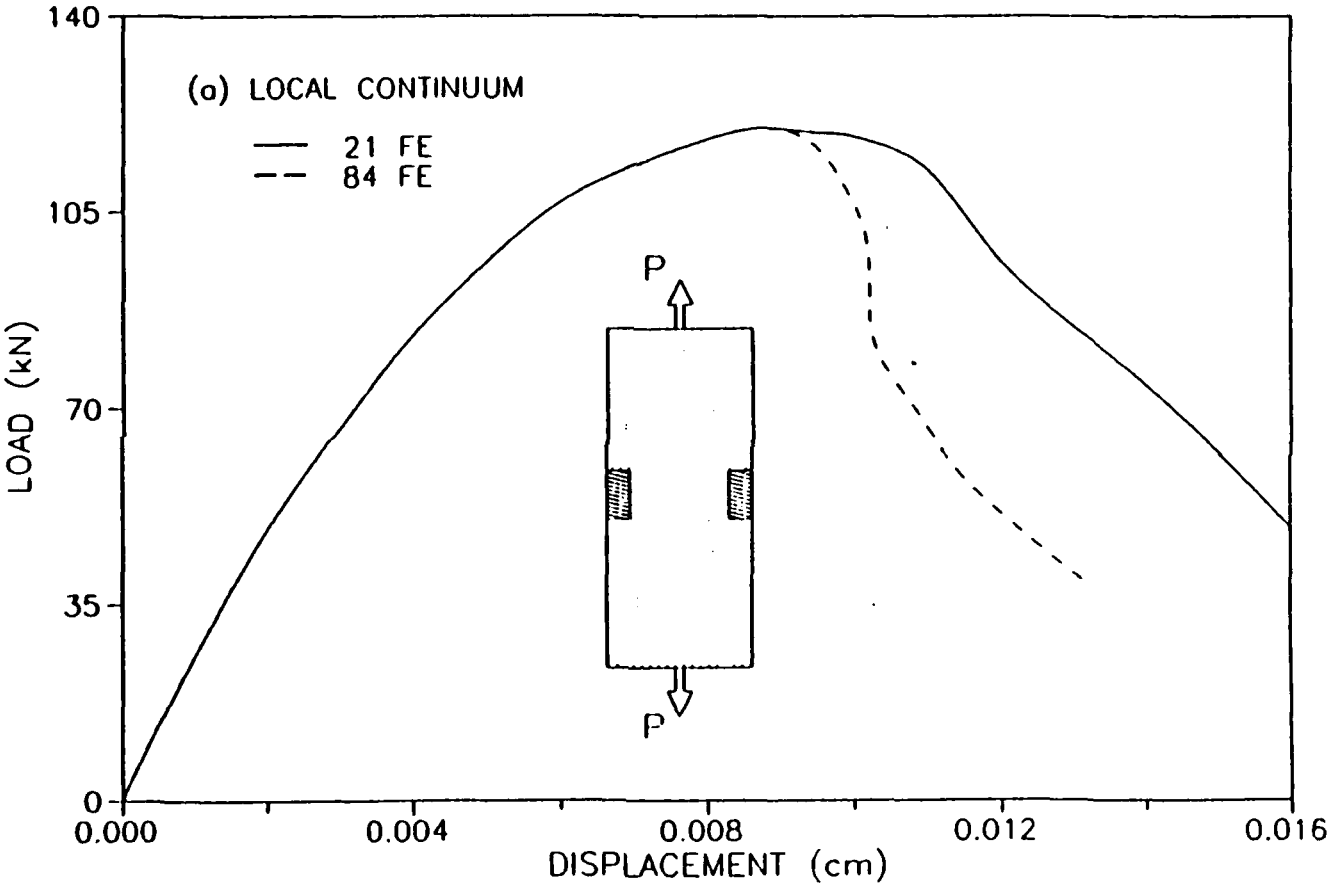

140
105
70
35
0
LOAD (kN)
(a) LOCAL CONTINUUM
21 FE
84 FE
P
P
0.000
0.004
0.008
0.012
0.016
DISPLACEMENT (cm)

fig. 3a

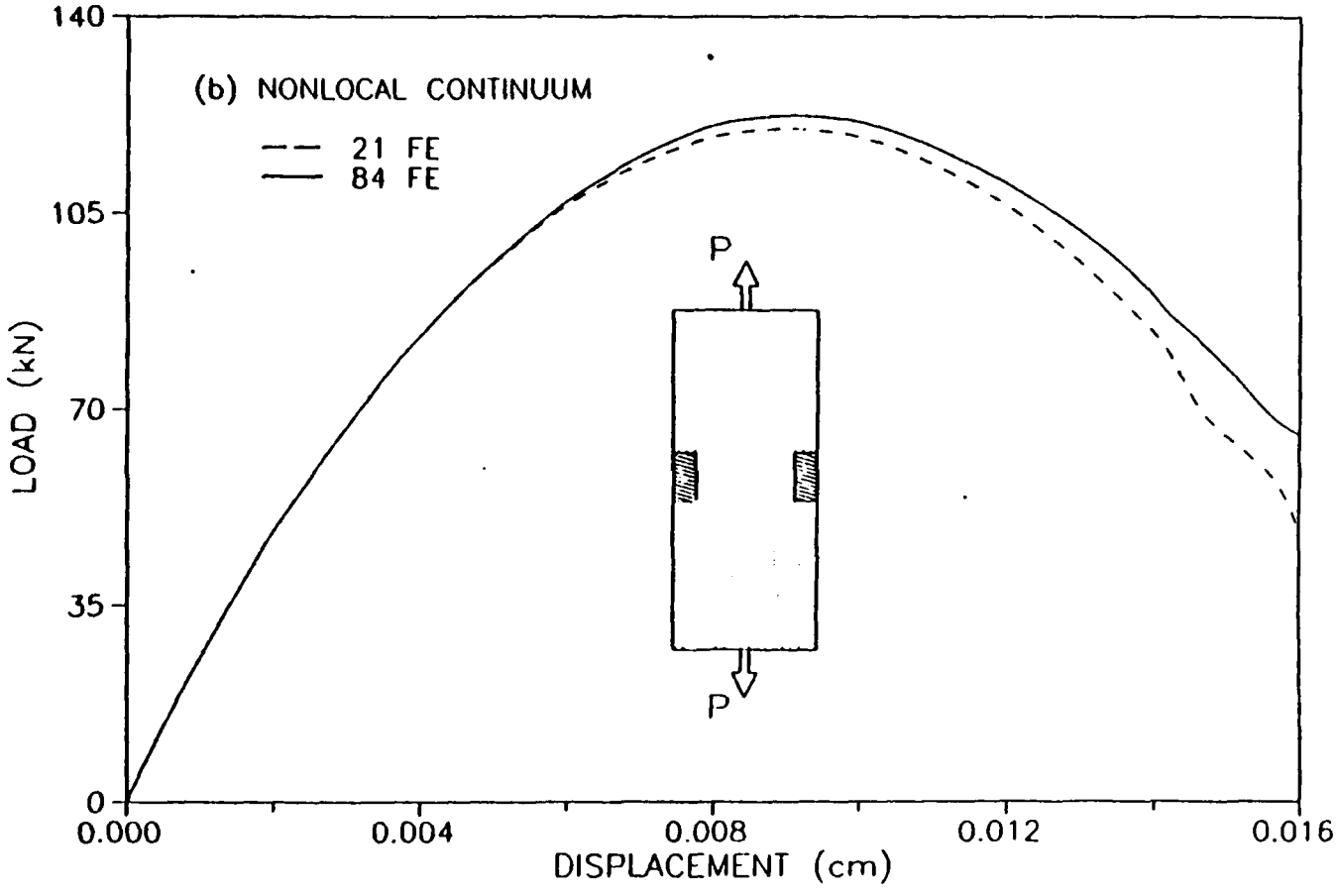

140
105
70
35
0
LOAD (kN)
(b) NONLOCAL CONTINUUM
21 FE
84 FE
P
P
0.000
0.004
0.008
0.012
0.016
DISPLACEMENT (cm)

fig. 3b

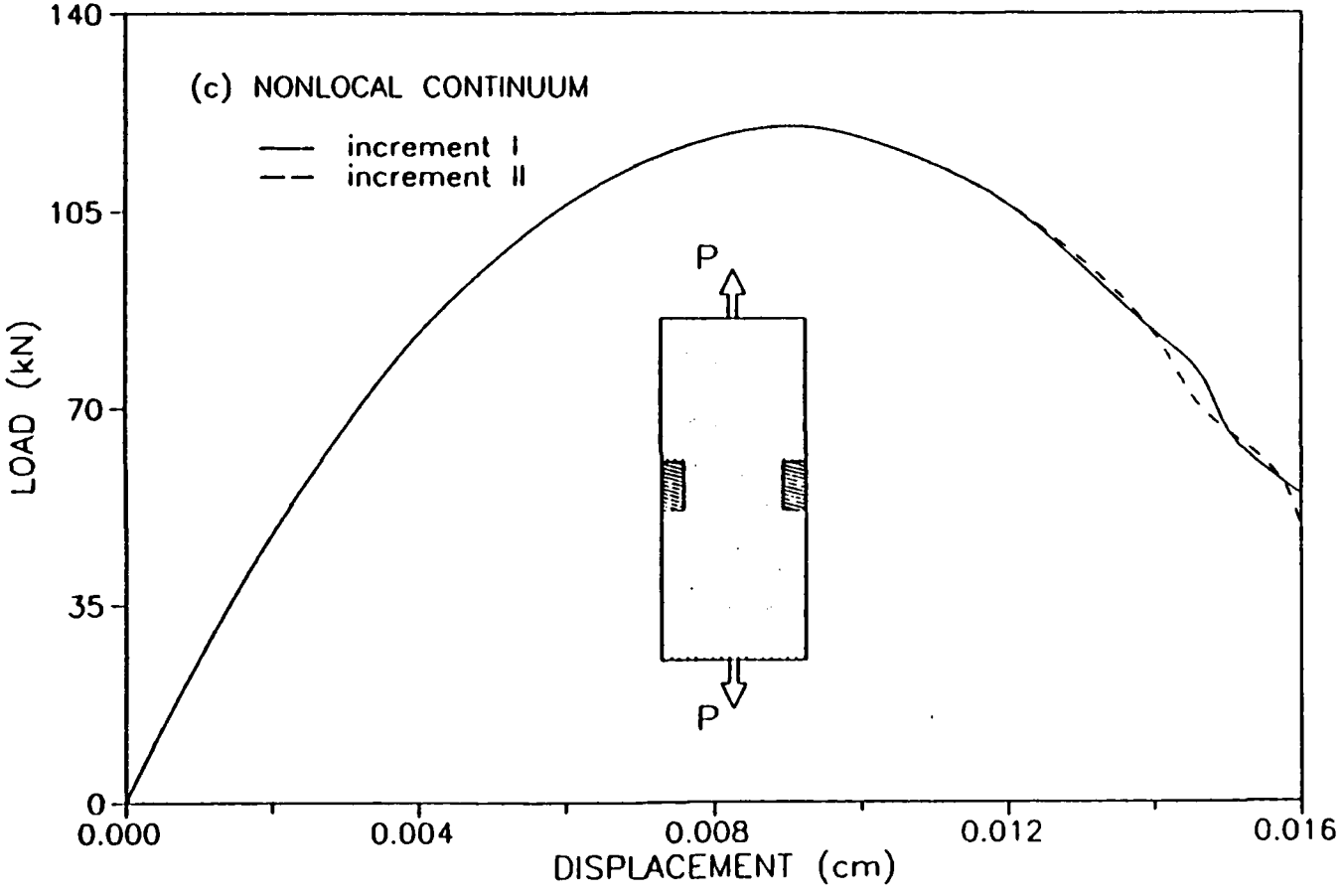

(c) NONLOCAL CONTINUUM
increment I
increment II
140
105
70
35
0
LOAD (kN)
P
P
0.000
0.004
0.008
0.012
0.016
DISPLACEMENT (cm)

fig. 3c

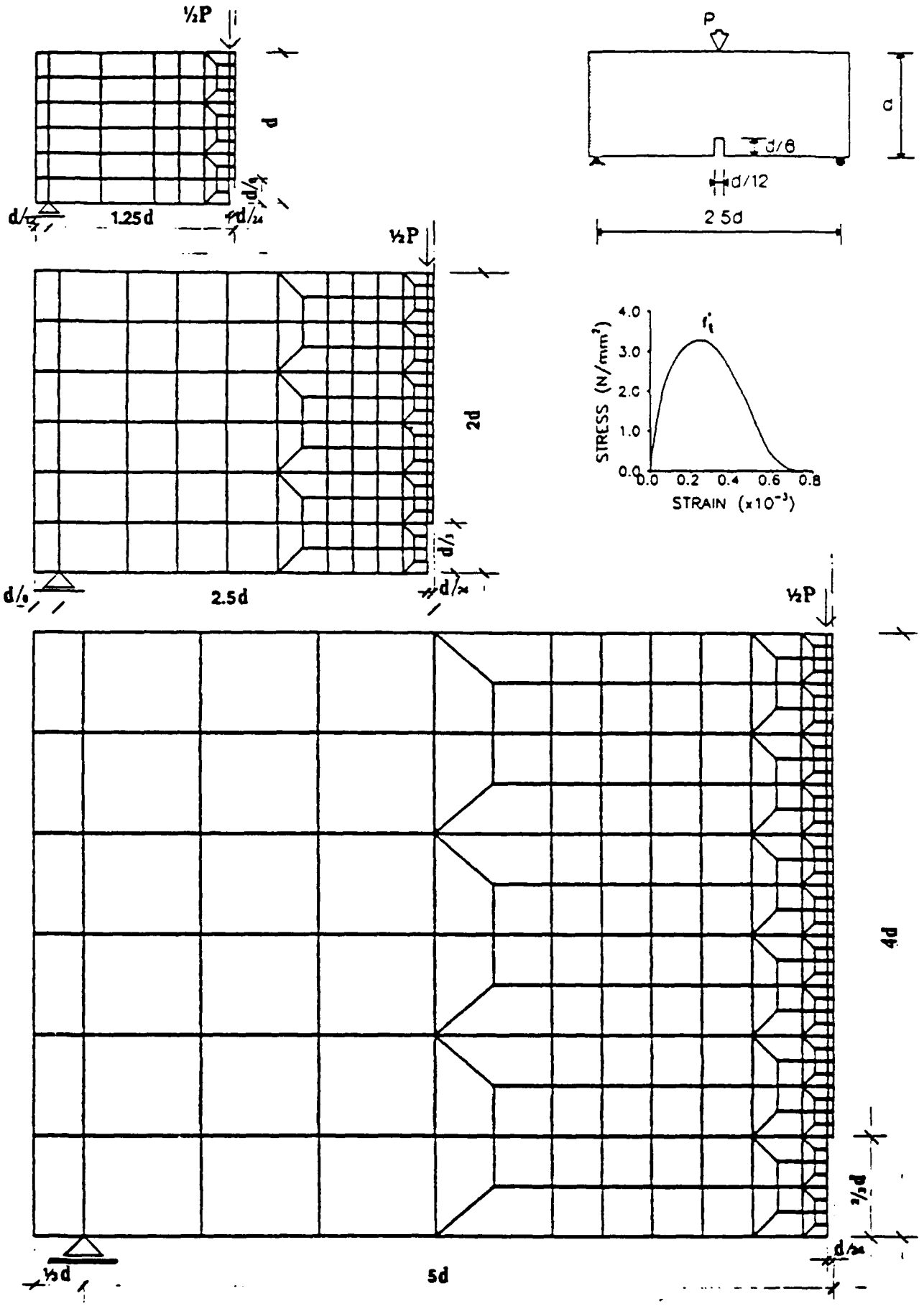

½P
d
d/8
d/24
1.25d
P
d
d/6
d/12
2.5d
½P
2d
d/8
d/24
2.5d
STRESS (N/mm²)
4.0
3.0
2.0
1.0
0.0
0.0
0.2
0.4
0.6
0.8
STRAIN (x10⁻³)
½P
4d
½d
d/24
½d
5d

fig.4

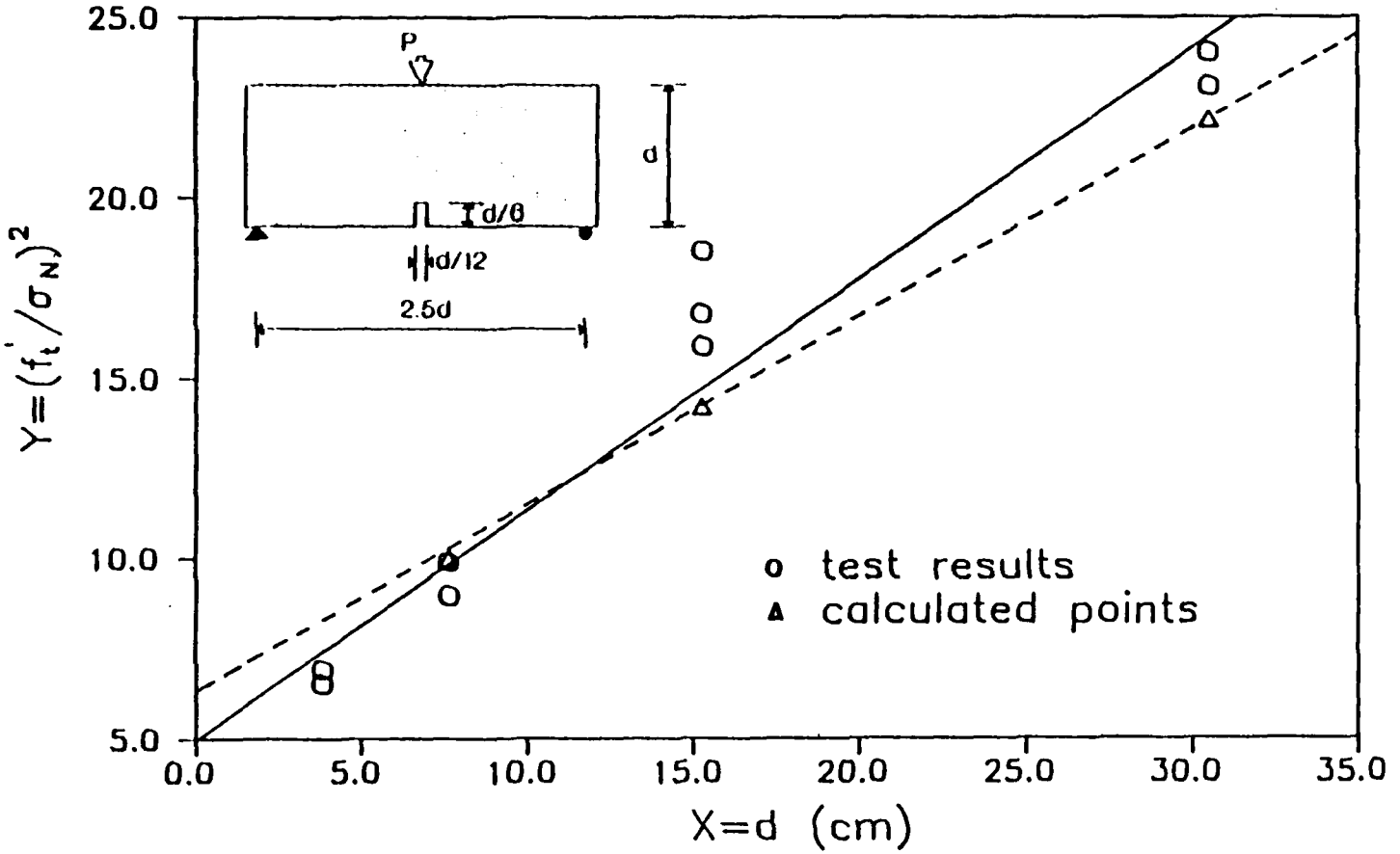

25.0
20.0
15.0
10.0
5.0
0.0
5.0
10.0
15.0
20.0
25.0
30.0
35.0
P
d
d/8
d/12
2.5d
$Y=(f_t'/\sigma_N)^2$
X=d (cm)
o test results
▲ calculated points

fig.5a

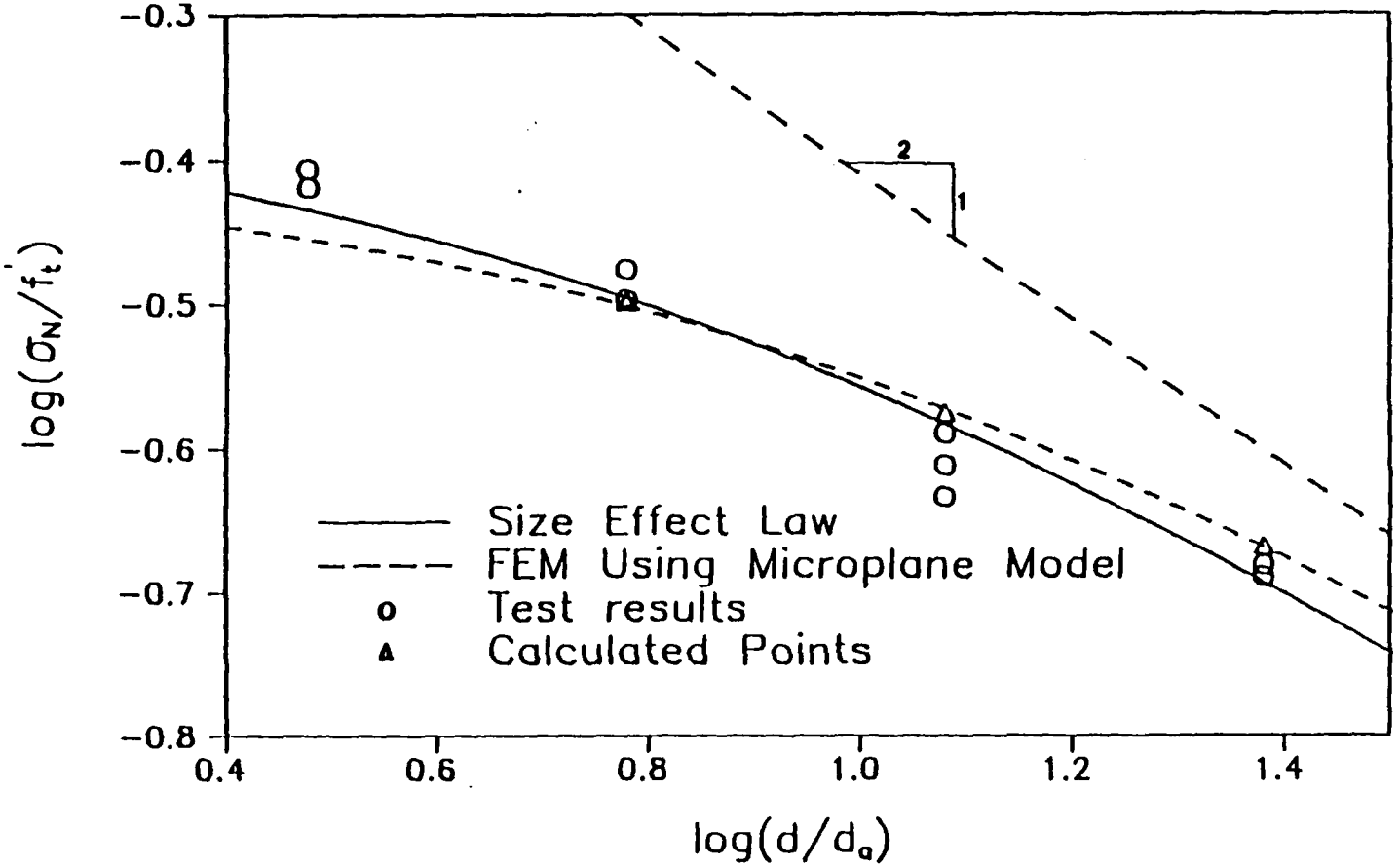

−0.3
−0.4
−0.5
−0.6
−0.7
−0.8
0.4
0.6
0.8
1.0
1.2
1.4
log(σN/f't)
log(d/da)
8
2
1
Size Effect Law
FEM Using Microplane Model
Test results
Calculated Points

fig. 5b

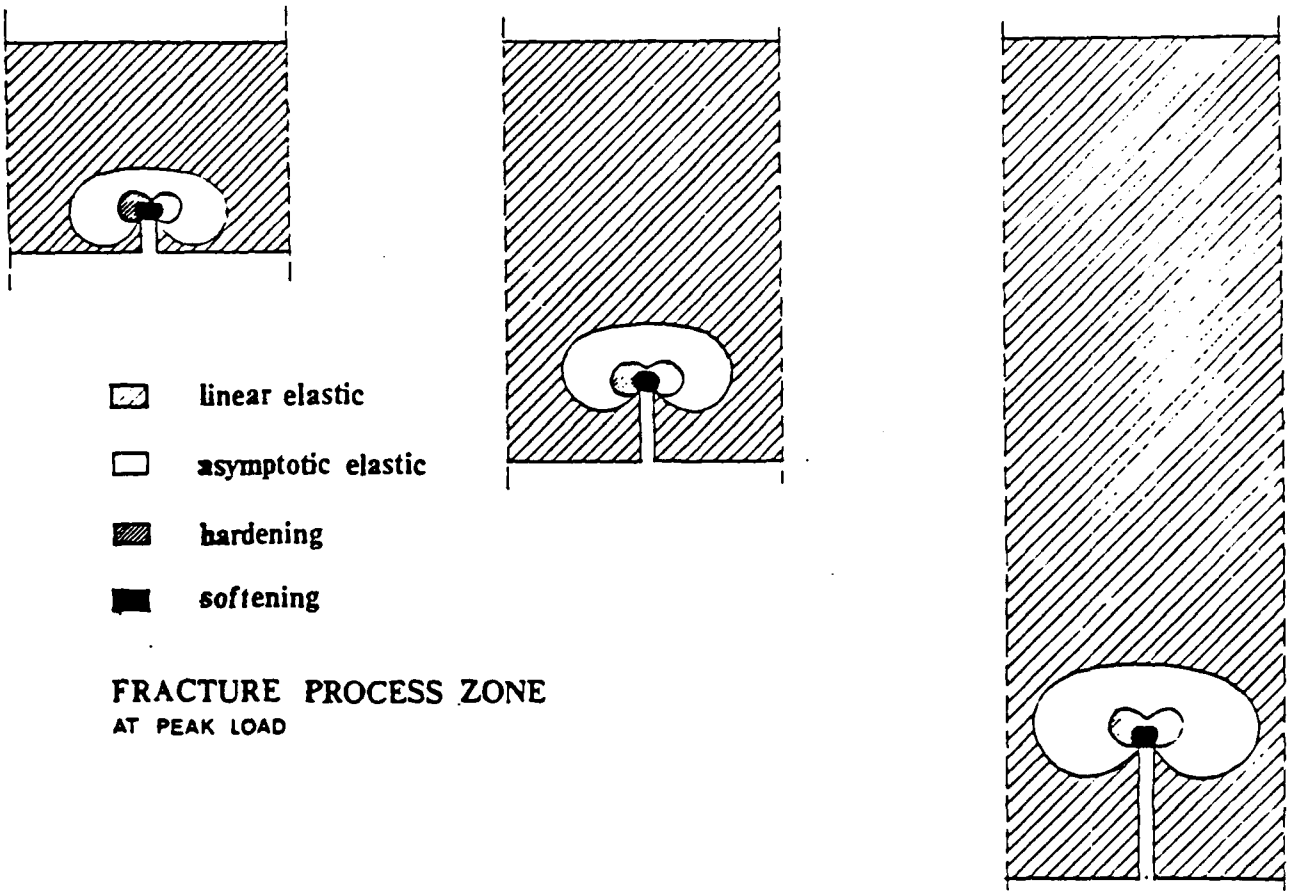

linear elastic
asymptotic elastic
hardening
softening
FRACTURE PROCESS ZONE
AT PEAK LOAD

fig.6

Nonlocal Continuum Damage, Localization Instability and Convergence

Zdeněk P. Bažant
Professor of Civil Engineering.
Mem. ASME

Gilles Pijaudier-Cabot[1]
Graduate Research Assistant.

Center for Concrete and Geomaterials,
Northwestern University,
Tech-2410,
Evanston, IL 60208

A recent nonlocal damage formulation, in which the spatially averaged quantity was the energy dissipated due to strain-softening, is extended to a more general form in which the strain remains local while any variable that controls strain-softening is nonlocal. In contrast to the original imbricate nonlocal model for strain-softening, the stresses which figure in the constitutive relation satisfy the differential equations of equilibrium and boundary conditions of the usual classical form, and no zero-energy spurious modes of instability are encountered. However, the field operator for the present formulation is in general nonsymmetric, although not for the elastic part of response. It is shown that the energy dissipation and damage cannot localize into regions of vanishing volume. The static strain-localization instability, whose solution is reduced to an integral equation, is found to be controlled by the characteristic length of the material introduced in the averaging rule. The calculated static stability limits are close to those obtained in the previous nonlocal studies, as well as to those obtained by the crack band model in which the continuum is treated as local but the minimum size of the strain-softening region (localization region) is prescribed as a localization limiter. Furthermore, the rate of convergence of static finite-element solutions with nonlocal damage is studied and is found to be of a power type, almost quadratic. A smooth weighting function in the averaging operator is found to lead to a much better convergence than unsmooth functions.

Introduction

Prediction of damage and failure of brittle heterogeneous materials such as concrete or rock requires a mathematically correct and physically realistic description of the strain-softening behavior (Bažant, 1986; Mazars and Pijaudier-Cabot 1986). Although it has been argued that strain-softening does not exist on the continuum level (Read and Hegemier, 1984), the macroscopic result of distributed microcracking or void growth is a behavior whose continuum description must incorporate strain-softening. Numerous attempts to describe this type of behavior by local inelastic continuum theories such as plasticity or continuum damage mechanics have been unsatisfactory because the phenomenon of strain localization caused by strain-softening cannot be captured objectively (Bažant, 1986). The principal fault of the local continuum models is that the energy dissipated at failure is incorrectly predicted to be zero, and the finite-element solutions converge to this incorrect, physically meaningless, solution as the mesh is refined.

To remedy the situation, one must introduce some form of a localization limiter (Bažant and Belytschko, 1987). Its simplest form is obtained by imposing a lower bound on the finite-element size, as is done in the crack band model. As a better localization limiter which permits arbitrary mesh refinement, one may adopt the nonlocal concept. Introduced into continuum mechanics long ago by Kröner (1967); Eringen (1972); Krumhansl (1968) and others, this concept was recently successfully applied to strain-softening (Bažant, et al., 1984). However, the formulation, in which all the state variables were nonlocal, turned out to be quite complicated. It required additional boundary and interface conditions, led to a nonstandard form of the differential equations of equilibrium, and the finite-element implementation required imbrication of the elements.

In the preceding study (Pijaudier-Cabot and Bažant, 1986) to be further expanded here, a new idea which turns out to bring considerable simplification was introduced. It was shown that it suffices to consider as nonlocal only the strain-softening damage, while the elastic behavior (including unloading and reloading) should be treated as local. This formulation was shown to require no element imbrication and no overlay with local continuum, which had to be previously used by Bažant, et al., (1984) in order to suppress certain periodic

[1]On leave from Laboratoire de Mécanique et Technologie, Cachan, France.

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of The American Society of Mechanical Engineers.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March 9, 1987; final revision, November 13, 1987.

Paper No. 88-WA/APM-18.

zero-energy modes in the imbricate, fully nonlocal solutions. The purpose of the present study is to show that, more generally, the key attribute of the nonlocal formulation for strain-softening is that the strain as a kinematic variable should be defined as local, and to study various aspects yet unexplored, including: (1) Static strain-localization instability; (2) Symmetry of the field operator; (3) Influence of various types of spatial averaging; (4) Convergence at mesh refinement.

Before embarking on our analysis, it should be mentioned that other forms of localization limiters are possible and deserve to be studied. The oldest idea, proposed already by L'Hermite (1952) in a study of concrete cracking due to shrinkage, is to introduce a dependence of the strength or yield limit on the strain gradient. In a general form, which may involve introduction of the strain gradient into the yield function, this idea has recently been developed by Floegl and Mang (1981), Schreyer and Chen (1984) and Mang and Eberhardsteiner (1986). Introduction of higher-order gradients into the differential equations of equilibrium or into the definition of strength was studied by Aifantis (1984) and also by Bažant (1984) and Bažant and Belytschko (1987). As another approach, Sandler (1984) as well as Needleman (1987) showed that the introduction of viscosity, whether real or artificial, may act in certain problems as a localization limiter, although this can be true only for a limited time period of response.

Nonlocal Generalization of Continuum Damage Mechanics

The principal idea for the treatment of softening is that only those variables which cause softening may be considered as nonlocal while the model must reduce to a local one for the special case of elastic response, which also includes unloading and reloading. This condition can be satisfied by a nonlocal formulation in which the strain, when used as a kinematic variable, is local. To implement this condition it is convenient, albeit not necessary, to use continuum damage mechanics because in this theory the strain-softening is characterized by a distinct single variable ω, called damage. For the nonlocal generalization, we adopt the simple, scalar damage formulation, although the same concept could be implemented in a similar manner in the anisotropic damage models, derived, e.g., by Mazars and Pijaudier-Cabot (1986) and Ladevèze (1983), which are more realistic. Various other constitutive theories can also be generalized to such a nonlocal form, particularly those which use fracturing strain or a degrading yield limit. (Bažant, et al. 1987). Unimportant from the viewpoint of the type of localization studied here, plastic strains will be omitted from the formulation; their inclusion, however, would require no conceptual changes.

As usual in continuum damage mechanics, we may introduce the relation between the strains ϵ_{ij} and the stresses σ_{ij} in the form (Lemaitre and Chaboche, 1985)

$$\sigma_{ij} = (1-\Omega)C_{ijkm}\epsilon_{km} \tag{1}$$

in which C_{ijkm} are the elastic constants of the material and Ω is the damage. We assume Ω to be nonlocal, defined by spatial averaging as follows

$$\Omega(\mathbf{x}) = \bar{\omega}(\mathbf{x}) = \frac{1}{V_r(\mathbf{x})}\int_V \alpha(\mathbf{s}-\mathbf{x})\omega(\mathbf{s})dV(\mathbf{s}) \tag{2}$$

in which

$$V_r(\mathbf{x}) = \int_V \alpha(\mathbf{s}-\mathbf{x})\mathrm{d}V(\mathbf{s}). \tag{3}$$

Superimposed bar denotes the spatial averaging operator, $\mathbf{x}$ and $\mathbf{s}$ are the coordinate vectors, V = volume of the body, and α = given weighting function. Initially $\Omega = \omega = 0$, and always $0 \le \Omega \le 1$, $0 \le \omega \le 1$.

As the simplest form of the weighting function, one may consider $\alpha = 1$ within a certain representative volume V_r centered at points $\mathbf{x}$, and $\alpha = 0$ outside of it. As we will see, however, a uniform weighting function over a finite domain does not yield the best convergence. A much better convergence on mesh refinement is obtained when function $\alpha(\mathbf{x}-\mathbf{s})$ decays smoothly with the distance from point $\mathbf{x}$. A suitable form is the Gaussian (normal) distribution function, previously used for nonlocal elasticity by Eringen (1972);

$$\alpha(\mathbf{x}) = \exp[-(k|\mathbf{x}|/l)^2] \tag{4}$$

in which we have, for one, two and three dimensions

$$\begin{aligned} &\text{1D:} \quad |\mathbf{x}| = x && k = \pi^{1/2} \\ &\text{2D:} \quad |\mathbf{x}| = (x^2+y^2)^{1/2} && k = 2 \\ &\text{3D:} \quad |\mathbf{x}| = (x^2+y^2+z^2)^{1/2} && k = (6\sqrt{\pi})^{1/3}. \end{aligned} \tag{5}$$

x, y, z are the Cartesian coordinates.

The expressions in equation (5) have been determined from the condition that the integral of $\alpha(\mathbf{x})$ for an infinite body be equal for one dimension to the length l (line segment), for two dimensions to the area of a circle of diameter l, and for three dimensions to the volume of a sphere of diameter l.

The function in equation (4) decays so rapidly that for points $\mathbf{s}$ whose distance from point $\mathbf{x}$ exceeds $2l$ one may set $\alpha = 0$. Calculation of $V_r(\mathbf{x})$ according to equation (3) is needed for the treatment of boundaries. Function α in general extends beyond the boundaries of the body. The domain beyond the boundary is simply deleted from integration, but the weighting function is scaled so that the integral of all effective weights $\alpha'(\mathbf{x}, \mathbf{s}) = \alpha(\mathbf{s}-\mathbf{x})/V_r(\mathbf{x})$ over the body should be exactly 1 for any $\mathbf{x}$. In numerical programming, the integrals in equations (2) and (3) are approximated as finite sums. The values of $\alpha'(\mathbf{x}, \mathbf{s})$ are generated for all combinations of all integration points of all elements in advance of the finite-element analysis. The programming of the averaging integral, which was demonstrated by Bažant, et al. (1987) in a study of cave-in of a tunnel due to compressive strain-softening for a nonlocal finite-element system with up to 3248 unknowns, is easier when the integral in equation (2) extends over the entire body. When a finite averaging domain is used, it is important to closely match by finite elements the precise boundaries of this domain, but this is difficult to implement (Saouridis, et al., 1987).

Length l, called the characteristic length, represents a material property and is of the same order of magnitude as the maximum size of material inhomogeneities. Length l has been determined experimentally (Bažant and Pijaudier-Cabot, 1987*a*) by comparing the responses of specimens in which the damage (e.g., microcracking) remains distributed with the responses of fracture specimens, in which damage localizes. For concrete, such experiments indicated that l is roughly equal to 2.7-times the maximum aggregrate size.

Length l can be determined also by micromechanics analysis. In a parallel study (Bažant, 1987), a local homogeneous elastic continuum, containing an array of growing circular cracks with periodic spacing l and quasi-periodic crack sizes, was analyzed. It was shown that by applying the usual homogenization conditions, one obtains a nonlocal continuum with local strain and a nonlocal cracking strain, for which the weighting function is uniform and the characteristics length is equal to crack spacing. For the case when the crack spacing is not constant but randomly distributed, this indicates that one should superpose at each point a set of uniform weighting functions with various characteristics lengths l. This, however, appears to be approximately equivalent to use a nonuniform weighting function.

For a uniform weighting function, volume V_r has a similar

meaning as the representative volume in the statistical theory of heterogeneous materials. However, its size l_r is smaller than that of the representative volume and is too small to obtain statistical averages of the microstructure. The representative volume in the statistical theory must be several times larger.

The specific free energy $\rho\psi$ per unit volume (with ρ being the material mass density) and the energy dissipation rate ϕ may be expressed as

$$\rho\psi = \tfrac{1}{2}(1-\Omega)C_{ijkm}\epsilon_{ij}\epsilon_{km} \tag{6}$$

$$\phi = -\partial(\rho\psi/\partial t) = -\dot{\Omega}\partial(\rho\psi)/\partial\Omega = \dot{\Omega}Y$$

in which Y, called the damage energy release rate, is

$$Y = -\partial(\rho\psi)/\partial\Omega = \tfrac{1}{2}C_{ijkm}\epsilon_{ij}\epsilon_{km}. \tag{7}$$

The damage evolution is characterized in general by an equation of the type $\dot{\omega} = f(\epsilon_{ij},\omega)$. An integrable special form of the damage evolution law has been used in all computations;

$$\omega = g(Y) = 1 - [1 + b(Y - Y_1)]^{-n} \tag{8}$$

in which b, n = positive material constants, $n > 2$, and Y_1 = local damage threshold.

Damage is assumed to grow only for loading. For unloading or reloading, $\dot{\omega} = 0$, which means the response is elastic. The loading criterion and the nonlocal damage Ω are defined as follows

$$\left.\begin{aligned} &\text{If } F(\tilde{\omega}) = 0 \text{ and } \dot{F}(\tilde{\omega}) = 0, \text{ then } \dot{\Omega} = \dot{\tilde{\omega}} \\ &\text{If } F(\tilde{\omega}) < 0, \text{ or if } F(\tilde{\omega}) = 0 \text{ and } F(\dot{\tilde{\omega}}) < 0, \text{ then } \dot{\Omega} = 0. \end{aligned}\right\} \tag{9}$$

Function $F(\tilde{\omega})$ represents the loading function and is defined as $F(\tilde{\omega}) = \tilde{\omega} - \kappa(\tilde{\omega})$, where $\kappa(\tilde{\omega})$ is a softening parameter which is set to be equal to the maximum value of $\tilde{\omega}$ achieved up to the present. The initial value of $\kappa(\tilde{\omega})$ is zero. The damage expression in equation (8) was found to approximate closely the behavior of concrete, provided that different local damage thresholds Y_1 are introduced for tension and compression (Mazars and Pijaudier-Cabot, 1986).

The formulation of the loading function automatically satisfies the dissipation inequality. The density of the energy dissipation rate due to damage is $\phi = \dot{\Omega}Y$, and since $\dot{\Omega} \geq 0$, we have $\phi \geq 0$.

In view of the use of a loading function, it might be more appealing to introduce some potential function, similar to plasticity. However, a formulation with a potential function has not yet been developed in the literature on continuum damage mechanics of concrete or geomaterials.

The fact that equation (1) uses nonlocal rather than local damage, and that the unloading condition is stated in terms of the nonlocal damage, represents all that is different from the classical local damage theory.

The original formulation of the nonlocal damage theory in the previous study by Pijaudier-Cabot and Bažant (1986) used a somewhat different definition of nonlocal damage. The averaged quantity was the damage energy release rate Y rather than the local damage, and so the nonlocal damage was determined as $\Omega = g(\bar{Y})$. This original formulation can be regarded as a simplified approximation of the present formulation. The numerical results for these two different definitions of nonlocal damage appear to be quite close.

The strains as well as ω have at least a C_0-continuity (i.e., they could consist of Dirac delta functions). According to equations (9) and (2) and the fact that $\phi = Y\dot{\Omega}$, the dissipation rate density is given by a spatial averaging integral over ω. Consequently, ϕ must have at least a C_1-continuity. This simple argument proves that the energy dissipation cannot localize into a zone of zero volume, and numerical solutions confirm it.

In general, one could show that by introducing any variable which causes the dissipation rate to be nonlocal (e.g., nonlocal plastic-strain path in plasticity with softening yield limit Bažant, et al., 1987), the energy dissipation is prevented to localize into a vanishing volume.

Differential Equation of Equilibrium and Boundary Conditions

In the original formulation of the nonlocal continuum for strain-softening in which nonlocal strains $\bar{\epsilon}_{ij}$ are used (Bažant, et al., 1984), the virtual work expression for a body of volume V_b and surface S_b must involve $\delta\bar{\epsilon}_{ij}$ rather than $\delta\epsilon_{ij}$

$$\delta W = \int_{V_b} (\sigma_{ij}\delta\bar{\epsilon}_{ij} - f_i\delta u_i)\,dV - \int_{S_b} p_i\delta u_i\,dS \tag{10}$$

in which u_i = displacement components in Cartesian coordinates x_i ($i = 1,2,3$), and f_i, p_i = distributed volume and surface forces. The fact that $\delta\bar{\epsilon}_{ij}$ involves the spatial averaging operator defined by equation (2) complicates the derivation of the differential equation of equilibrium and boundary conditions and yields a nonstandard form of these equations (Bažant, 1984).

For the present nonlocal theory, ϵ_{ij} is local and so $\delta\bar{\epsilon}_{ij}$ in equation (10) must be replaced by $\delta\epsilon_{ij}$. But then the variational derivation of the differential equation of equilibrium and the boundary conditions is the same as usual (Bažant and Pijaudier-Cabot, 1987). This shows that, as long as the strains are defined as local, the differential equations of equilibrium, as well as the boundary conditions or the interface conditions, will have the standard form. This further means that the finite-element discretization can be of the same type as for the usual, local continuum. Therefore we may conclude that the key simplifying feature of a continuum model for damage is the use of a nonlocal continuum in which the strains are local. In other words, the continuum should not be fully, but only partially nonlocal. It is also clear that if the strain is local then the elastic behavior, including the behavior at unloading and reloading, will be local.

One-Dimensional Strain-Localization Instability

Consider for the sake of simplicity the one-dimensional problem of a bar loaded through two springs of spring constant C (Fig. 1). This problem was used by Bažant (1976) to demonstrate the localization instability due to strain-softening and was recently solved exactly with a fully nonlocal theory (Bažant and Zubelewicz, 1986). The length coordinate is x, the

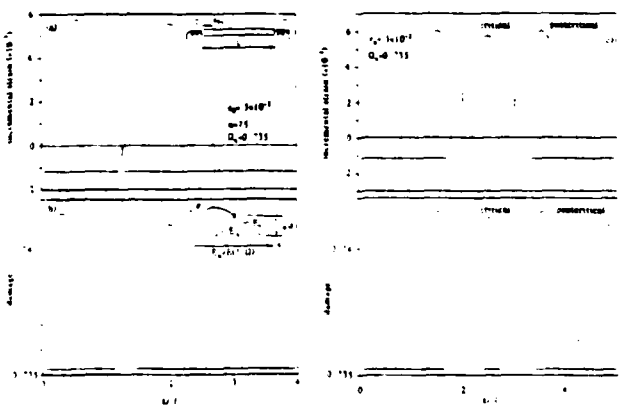


Fig. 1 Strain (a) and damage (b) localization profiles in the bar center and at the boundary, calculated for a one-dimensional bar; critical and postcritical strain and damage localization profiles (c, d)

bar length is L, and the bar ends are $x=0$ and $x=L$. The bar is initially in a state of uniform strain ϵ_0 and stress σ_0, with uniform damage $\Omega_0=\omega_0$. The initial state is in the strain-softening range and satisfies the relation $\sigma_0=(1-\Omega_0)E\epsilon_0$. If the constitutive relation for damage is given in the form of equation (8), we have $\omega_0=1-[1+b(\frac{1}{2}E\epsilon_0^2-Y_1)]^{-n}$.

We now consider a small deviation from the initial state caused by incremental variation of the load at the bar ends. Let $\delta\epsilon(x)=\eta(x)$ = incremental strain and $\delta\sigma=\tau$ = incremental stress. To maintain equilibrium, it is necessary that τ = constant along the bar. The compatibility condition for fixed supports requires that

$$\int_0^L \eta(x)dx+\frac{2\tau}{C}=0. \tag{11}$$

Taking the one-dimensional form of the constitutive law in equation (1) with the averaging according to equation (2) and the loading criterion according to equation (9), we find

For $\int_0^L \alpha(s-x)\eta(s)ds>0$: $\sigma_0+\tau$

$$=\left[1-\frac{1}{l'(x)}\int_0^L \alpha(s-x)\omega(s)ds\right]E(\epsilon_0+\eta(x)); \tag{12}$$

Otherwise: $\tau=(1-\omega_0)E\eta(x)$

in which $l'(x)=\int_0^L \alpha(s-x)dx$ = effective averaging length for point x. For infinitely small increments $\eta(x)$, the equations can be simplified by incremental linearization. To this end, we may introduce the linear approximation $\omega(x)=\omega_0+\omega_\epsilon\eta(x)$, with $\omega_\epsilon=\partial\omega/\partial\epsilon$ for $\omega=\omega_0$. Substituting into equation (12), neglecting the quadratic term $\eta(x)\eta(s)$, and subtracting the equation $\sigma_0=E(1-\Omega_0)\epsilon_0$, we can reduce equation (12) to the form

$$\tau=(1-\omega_0)E\eta(x)-\frac{k}{l'(x)}<\int_0^L \alpha(s-x)\eta(s)ds> \tag{13}$$

with $k=E\epsilon_0\omega_\epsilon$. The symbol $<\ >$, which introduces the loading criterion, is defined as $<x>=x$ if $x>0$ and otherwise $<x>=0$. For the special damage constitutive law in equation (8), we may evaluate $\omega_\epsilon=E\epsilon_0\omega_Y$, with $\omega_Y=bn[\frac{1}{2}E\epsilon_0^2-Y_1]^{n-1}[1-b(Y-Y_1)]^{-2n}$.

Equation (13) has the form of a linear integral equation of the second kind. However, the problem is not simply that of solving an integral equation, because equation (13) is an integral equation only for those x for which loading takes place. Outside the loading region, equation (12) reduces to the usual linear elastic differential equation for displacements and the behavior is then local, elastic, and unaffected by the strain-localization in another segment of the bar. If we find one solution such that there exist elastic segments of finite length at both ends of the bar, then we can obtain another solution simply by shifting the softening segment along with the softening solution profile as a rigid body. This shift is arbitrary provided the entire original length of the softening region, along with a small neighborhood of points located just outside the softening zone, remains within the bar. Due to this fact, we can calculate the length of the softening region and the solution profile through it by analyzing any bar of a shorter length, provided the bar length exceeds the length of the softening region. The actual boundary conditions and the compatibility condition may be disregarded in such analysis and satisfied afterward. Obviously, one can have infinitely many solutions. Arbitrary shifts of the softening region, however, do not affect the overall response of the bar. The length of the softening region and the solution profile through it is nevertheless unique. The localization profiles terminating at the boundary points are different from the interior one and require a special analysis. For real materials, the actual location of the strain-softening segment that does not reach the boundary is decided by inevitable random fluctuations of material strength along the bar.

Now consider the alternative formulation previously proposed by Pijaudier-Cabot and Bažant (1986), in which the averaged variable is Y rather than ω. For small strain increments, equation (2) can be approximated as $\Omega(x)=\omega_0+\Omega_Y[Y(x)-Y_0]$, with $\Omega_Y=[\partial\Omega/\partial Y]_{Y_0}=\omega_Y$ in which Y_0, Ω_0 and ϵ_0 are the values of Y, Ω and ϵ in the initial state. Linearization for small increments is then accomplished by writing

$$Y(x)=\frac{1}{l'(x)}\int_0^L \alpha(s-x)\frac{E}{2}(\epsilon_0^2+2\eta(s))\epsilon_0 ds. \tag{14}$$

Substituting this into the basic relation $\tau=(1-\Omega)E\eta(x)$, one obtains for $\eta(x)$ an equation identical to equation (13), in which $k=2Y_1\Omega_Y$. An equation of the same form would be obtained for various other possible types of averaging. For problems with nonuniform initial strain, however, different types of averaging do not yield the same field equations.

Equation (13) can be easily solved numerically. If we subdivide the bar length into N equal elements of length $\Delta x=L/N$, use for integration the trapezoidal rule, express τ from equation (11) and substitute it into equation (13), we may approximate the resulting equation as follows

$$\sum_{j=1}^{N} K_{ij}\eta_j=0 \quad \text{with}$$

$$K_{ij}=(1-\omega_0)E\delta_{ij}-\frac{k\Delta x}{l_i'}I_i\alpha(x_j-x_i)+\frac{C}{2}. \tag{15}$$

Subscripts i, j refer to centroids of elements number i or j, and $I_i=1$ if $\int_0^L \alpha(x_i-s)\eta(s)ds=\Sigma_i\alpha(x_j-x_i)\eta_j\Delta x\geq 0$; otherwise $I_i=0$. Equation (15) represents a system of homogeneous linear algebraic equations for η_j. The critical state occurs at such ϵ_0 for which $\det(K_{ij})=0$. The loading segment, characterized by $I_i=1$, is of course unknown in advance, and so it must be determined iteratively. To search for the critical state with the smallest ϵ_0, the values of ϵ_0 were incremented in small loading steps and, for each ϵ_0, the following algorithm was used:

1 In the first iteration, we assume that only one element i undergoes loading (the central element). In the subsequent iterations, we increase the number of elements that undergo loading by one (either on the left or on the right), unless the number of loading elements already equals N, in which case we start a new loading step with a larger ϵ_0.

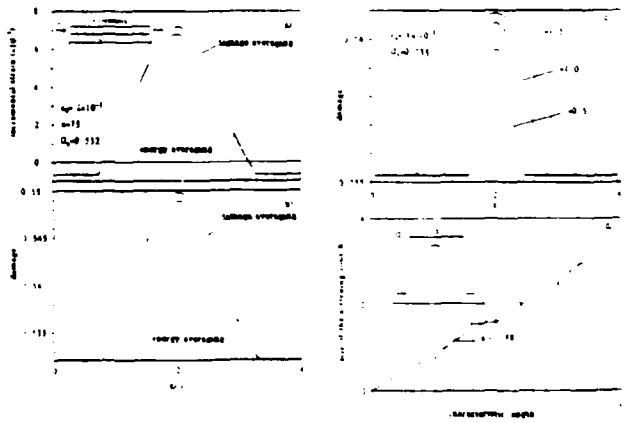


Fig. 2(a, b) Comparison of localization profiles of strain and damage for the damage and energy release ra... ...raging; (c, d) influence of the characteristic length l o... ...e damage localization profiles, and the size of the softening zone

2 For each assumed set of elements with loading, we calculate the lowest eigenvalue λ_1 of K_{ij} and the corresponding eigenvector $\eta_i^{(1)}$. If the positive values of $\eta_i^{(1)}$ in this eigenvector do not agree with the assumed set of loading elements, we return to 1 and repeat the calculation for the same ϵ_0 and a larger number of softening elements. If they agree but $\lambda_1 > 0$, we also return to 1 (after storing the λ_1-value) and commence the next load step with increased ϵ_0; otherwise ($\lambda_1 \leq 0$) the critical state $\epsilon_0 = \epsilon_0^{cr}$ has just been passed. In that case, we interpolate with regard to the λ_1-value in the preceding loading step, in order to estimate more closely the value of ϵ_0 for which $\lambda_1 = 0$. Then we print this ϵ_1-value, calculate the corresponding eigenvector $\eta_i^{(1)}$ and stop.

Lack of Symmetry. As we can see from equation (15), the tangential stiffness matrix K_{ij} is nonsymmetric ($K_{ij} \neq K_{ji}$). The nonsymmetry is inevitable for the present formulation and is due to the loading condition, brought in through parameter I_i, as well as to the variability of I_i'. If I_i as well as I_i' have the same values for all elements, K_{ij} would be symmetric. Similarly, the part of matrix K_{ij} for which I_i and I_i' are constant is symmetric. Matrix K_{ij} also becomes symmetric for $\epsilon_0 \to 0$ (elastic behavior). For small ϵ_0 it is weakly nonsymmmetric. The fact that K_{ij} is nonsymmetric implies that the integral operator in equation (13) is nonsymmetric, too, which means it is not self-adjoint. This further implies that a minimizing functional associated with equation (15) does not exist.

By contrast, the operators and stiffness matrices for the previously formulated imbricate nonlocal continuum (Bažant, et al., 1987) are symmetric. In fact, the requirement of symmetry was shown to inevitably lead to the imbricate nonlocal model (Bažant, 1984). Convenient though the symmetry is for numerical solutions, this advantage is nevertheless more than offset by other disadvantages of the imbricate model, as already mentioned.

Numerical Studies. Typical results of numerical solutions of the strain-localization are plotted in Figs. 1-3. It is assumed that $E = 32000$ MPa, $b = 20.5$, $Y_0 = 8540 \times 10^{-6}$; these values are based on certain measurements (Nazars and Pijaudier-Cabot, 1986), $l = 0.25\ L$. The solutions are calculated for the averaging of ω as well as for the averaging of Y, and also for the uniform weighting function. The loading increment is $\tau = -0.1$ MPa. The number of elements along the bar is $N = 75$.

The profile $\eta(x)$ depends only on the critical state ϵ_0, Ω_0 and the boundary conditions. The strain-localization profile is symmetric if the localization region does not touch the boundary points of the bar. One also can obtain solutions with several waves as shown in Fig. 1(*c*, *d*); these solutions, however, could be reached only if the bar were prevented to fail at smaller strains. Figure 2(*a*, *b*) shows that the present damage averaging formulation and the original energy averaging formulation gives very close results. Note from the solutions that the damage is continuous but the slope of the damage profile may be discontinuous at the boundary of the localization region. Also note that the values of the strain increment can be negative within the loading (i.e., softening) zone. This is not objectionable since the unloading is defined on the basis of the increments of Ω rather than ϵ. This property, which is inevitable if the strains should be continuous and is the physical reason for the lack of symmetry and self-adjointness of equation (13), is revealed by equation (15).

The characteristic length controls the width of the softening zone. It also has a great effect on the smallest strain value ϵ_0 at which instability is possible. Figure 2(*c*) shows the damage profiles for various values of l. The relationship between the length of the softening zone and the characteristic length l is shown in Fig. 2(*d*), and we see that the length h of the softening zone is approximately $h = 1.88\ l$ (at constant strain ϵ_0). This result, which is close to that obtained before for wave propagation problems (Pijaudier-Cabot and Bažant, 1986) has been exploited for measurement of the characteristic length (Bažant and Pijaudier-Cabot, 1987*a*). It also lends simple physical interpretation to the charactristic length, l.

Figure 3(*a*) shows the dependence of the critical value of the tangent modulus E_t on L/l for various values of C; E_t is understood as the effective tangential modulus for further loading (softening) and is evaluated on the basis of local damage (for nonlocal damage the value of E_t cannot be defined uniquely). Figure 3(*b*) shows the numerical results for dependence of the ductility ratio ϵ_0/ϵ_p on the ratio of the bar length L to the characteristic length l, for various values of the spring constant C (A = bar cross section = 1 in the calculation, ϵ_p = strain at peak stress).

Comparison with the previously given local solution (Bažant, 1976) is shown in Fig. 3(*c*). In a local formulation, however, the ratio h/L needs to be fixed in advance, as is done in the crack band theory. In a nonlocal formulation, h is a solution obtained on the basis of nonlocal damage theory; $h = \beta(\epsilon_0)l$ where β is a coefficient, mildly depending on the critical strain; for $\epsilon_0 = 3 \times 10^{-3}$ we find $\beta = 1.88$, as shown in Fig. 2(*d*). We also see that the local and nonlocal solutions are very close to each other. However, such close agreement need not extend to general three-dimensional situations with irregular damage zones.

Figure 3(*c*) compares the nonlocal solutions obtained for the present damage averaging and the previously used energy release rate averaging. The comparisons of solutions for various choices of the weighting function $\alpha(x)$ are presented in Fig. 3(*d*). Both types of averaging and different weight functions yield rather close results.

Energy Analysis of Localization Instability

Although a minimizing functional does not exist for the present nonlocal theory, stability can be checked by calculating the incremental work ΔW that needs to be done on the bar to produce small strain variations $\delta\epsilon(x) = \eta(x)$. Thus the stability condition is $\Delta W > 0$ for all $\eta(x)$ where $\Delta W = \delta W + \delta^2 W$. Since $\delta W = 0$ if the initial state is an equilibrium state, the stability condition is $\delta^2 W > 0$. We have $\delta^2 W = \frac{1}{2}\int \eta(x)\delta\tau dx - \frac{1}{2}C\delta u_2^2$. Upon substituting for $\delta\tau$ according to equation (13), we get the stability condition in the form

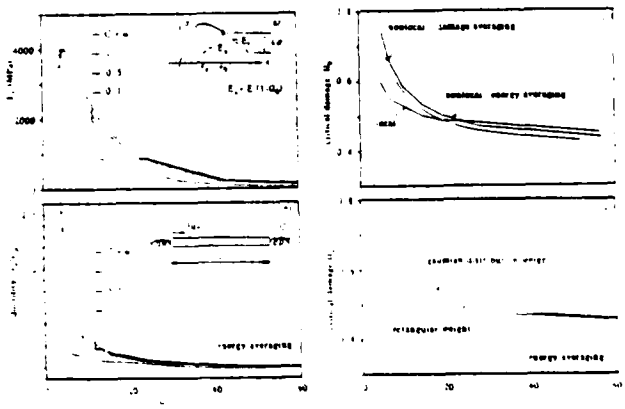


Fig. 3 Stability limits in terms of critical tangential modulus (*a*: ductility (*b*); critical damage obtained from the damage and energy averaging formulations as function of the length of the bar (*c*); influence of the weighting function (*d*)

$$\delta^2 W = \tfrac{1}{2}\int_0^L \eta(x)[(1-\omega_0)E\eta(x)$$

$$-\frac{E\epsilon_0\omega_0}{l'(x)} < \int_0^L \alpha(s-x)\eta(s)ds >]dx + \frac{C}{2}\delta u_L^2 > 0 \qquad (16)$$

subject to the compatibility condition $\int \delta\eta(x)dx + 2\delta\tau/C = 0$. As expected, $\delta^2 W$ vanishes for the critical state, as indicated by equation (13). From this, it follows that for $\epsilon_0 < \epsilon_0^{cr}$, $\delta^2 W > 0$, and for $\epsilon_0 > \epsilon_0^{cr}$, $\delta^2 W < 0$. Numerical evaluations of equation (16) confirm it. Thus the bar is stable for $\epsilon_0 < \epsilon_0^{cr}$, and unstable for $\epsilon_0 > \epsilon_0^{cr}$.

Note again that the quadratic functional in equation (16) is not symmetric. This is due to the presence of operator $< >$ as well as the variability of $l'(x)$.

Convergence at Mesh Refinement

The chief problem with the local solutions for strain-softening has been their convergence to physically incorrect solutions with zero-energy dissipation and the consequent spurious mesh sensitivity. For dynamic response, a rapid convergence has already been numerically demonstrated by Pijaudier-Cabot and Bažant (1986) but it still remains to examine convergence and determine how it is affected by the choice of the weighting function and other factors. Rapid and correct convergence is particularly important for strain-localization problems since strain-localization may often necessitate a very fine discretization in the damage zones. It is therefore important to capture strain-localization with the smallest possible number of elements per representative length of the material.

We now consider the formulations with the averaging of both the local damage ω and the damage energy release rate Y.

In numerical calculations, we compare the results for a uniform weight ($\alpha = 1$ over the length l), used in the initial work, with various nonuniform weights, such as the Gaussian distribution and the triangular function defined as $\alpha(s) = (2/l)(1-(2/l)|s|)$ for $|s| < l/2$ and $\alpha(s) = 0$ for $|s| \geq l/2$. In contrast to the original, fully nonlocal model (Bažant, et al., 1984), the weighting function, used for the strain-softening only, need not satisfy specific stability conditions stated in terms of the Fourier transform by Bažant and Chang (1984).

For each weighting function, the strain and damage localization profiles for various times were calculated, for the same bar as before, using different numbers, N, of constant-length, two-node finite elements. The number of elements per characteristic length l ranged from 3 to 25. Figure 4(*a–c*) shows the convergence of the strain profiles for a bar initially uniformly strained up to $\epsilon_0 = 3\times 10^{-3}$. We see that, for the uniform weight function, the convergence is relatively slow but improves drastically for the continuous weight functions, especially for the smooth Gaussian distribution function. For this function, only a few elements (3 to 6) over the characteristic length are required in order to determine with a practically sufficient accuracy the failure due to unstable strain-softening. This is also corroborated by Fig. 4(*c*) which shows the dependence of the size of the softening zone on the number of finite elements. Evidently, a relatively crude mesh suffices for good failure prediction if the Gaussian distribution function is used.

Let us now explore numerically the rate of convergence of the energy W dissipated due to damage. The finite-element solutions usually exhibit power-type convergence. This means that for a large enough number, N, of elements the error decreases as a certain power of the finite-element size h. Therefore the values of W calculated for various meshes should (in the asymptotic sense for $h \to 0$) vary as

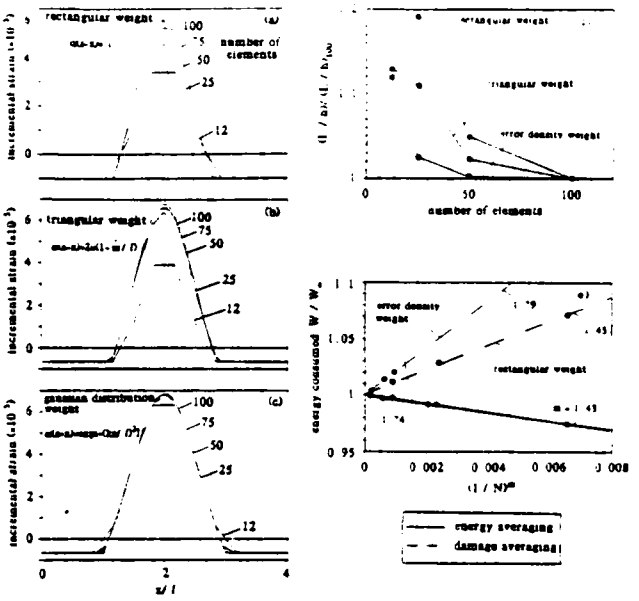


Fig. 4(a–c) Convergence of the strain localization profiles for various weighting functions; (*d*) convergence of the size of the softening zone; (*e*) convergence of the energy dissipated

$W = AN^{-m} + W_\infty$, in which A, W_∞ and m are certain constants. For problems of linear elasticity, the convergence is quadratic, i.e., $m = 2$. The value W_∞ represents the limiting solution for $N \to \infty$, which may be considered as the exact solution. For nonlinear problems, the convergence is generally slower than quadratic; i.e., $m < 2$.

The values of W_∞, A and M have been identified by fitting with the formula for W the results on at least three different meshes using Levenberg-Marquardt nonlinear optimization subroutine.

Instead of the static localization problem studied here so far, the dynamic localization problem studied previously by Pijaudier-Cabot and Bažant (1986) has been analyzed numerically. An initially stress-free bar is subjected for times $t > 0$ to constant outward velocities at both ends, which are such that the inward waves are elastic but reach strain-softening when they meet at the middle. Prior to softening, the inward waves are step functions but after softening a complex localization profile develops. For the local case, this problem has been solved exactly (cf. Bažant, 1986), but for the nonlocal case, a numerical solution is necessary. The results are plotted in Fig. 4(*e*) which shows that the convergence rate, as indicated by exponent m, ranges from $m = 1.43$ for the uniform weight function to 1.79 for the Gaussian weight function ($W_\infty = 0.0003925$). Moreover, the magnitude of error is by far the smallest for the Gaussian weight function. We therefore may conclude that this function should be preferred in practical applications.

Comparisons have also been made between the two types of averaging used in this paper. The convergence was found to be distinctly better for the averaging of the local damage ω than of the local damage energy release rate Y.

Conclusions

1 The key feature of a simple nonlocal theory of strain-softening is that the strain should be local. This implies that the elastic behavior, including unloading after strain-softening as well as reloading, must be local. Nonlocal must be the variables that determine damage, the phenomenon responsible for strain-softening.

2 Owing to the continuity properties introduced by the nonlocal concept, the energy dissipation due to damage can-

not localize into a region of vanishing volume. The present solutions of static strain-localization confirm this result, and so do the previous wave propagation solutions.

3 The field operator in the equilibrium for the present theory is nonsymmetric, except for the elastic part of response, and discretization yields a nonsymmetric tangential stiffness matrix. The lack of symmetry is caused by applying the spatial averaging operator only to damage. A symmetric operator can be achieved with the previously proposed imbricate nonlocal formulation, but that formulation has other more serious disadvantages.

4 Analysis of the strain-localization instability in a bar that is initially under uniform strain leads to an eigenvalue problem for an integral equation. This equation can be accurately solved numerically. The solution yields the size h of the localization zone. This size is found to be approximately proportional to the characteristic length l of the nonlocal continuum ($h = 1.88\ l$).

5 The magnitude of the critical strain for static localization instability in a strain-softening bar decreases as the bar length increases or as the support stiffness decreases. This agrees with previous solutions in which the size of the strain-softening region was considered to be a material property. In the present theory, this size is calculated and depends mainly on the characteristic length of the nonlocal continuum.

6 Finite-element analysis for the presently proposed nonlocal damage formulation appears to exhibit power-type convergence which is almost quadratic in energy, even if the entire bar is initially in the strain-softening state. A smooth weighting function, such as the Gaussian distribution function (normal distribution), yields a subtantially better convergence than unsmooth weighting functions.

Acknowledgment

Financial support under U. S. Air Force Office of Scientific Research Grant No. F49620-87-C-0030DEF to Northwestern University, monitored by Dr. Spencer T. Wu, is gratefully acknowledged.

References

Aifantis, E. C., 1986, "On the Microstructural Origin of Certain Inelastic Models," *J. Eng. Mat. Tech.*, No. 106, pp. 326–330.

Bažant, Z. P., 1976, "Instability, Ductility and Size Effect in Strain-Softening Concrete," *Journal of the Engineering Mechanics Division, ASCE*, Vol. 102, No. EM2, pp. 331–344; discussions Vol. 103, pp. 357–358, 775–777, Vol. 104, pp. 501–502 (based on Struct. Eng. Report No. 74-8/640, Northwestern University, Aug. 1974).

Bažant, Z. P., 1984, "Imbricate Continuum and its Variational Derivation," *Journal of the Engineering Mechanics Division, ASCE*, Vol. 110, No. 12, pp. 1693–1712.

Bažant, Z. P., 1986, "Mechanics of Distributed Cracking," ASME *Applied Mechanics Reviews*, Vol. 39, No. 5, May, pp. 675–705.

Bažant, Z. P., 1987, "Why Continuum Damage is Nonlocal: Justification by Quasiperiodic Crack Array," Concrete and Geomaterials Report No. 87-8/498W, Northwestern University, Evanston, IL (August); also *Mechanics Research Communications*, Vol. 14, 1987, pp. 407–419.

Bažant, Z. P., and Belytschko, T. B., 1987, "Strain-Softening Continuum Damage: Localization and Size Effect," Preprints, International Conference on Constitutive Laws for Engineering Materials, Vol. 1, eds. Desai, C. S, et al., held in Tucson, AZ., Jan., Elsevier, NY, pp. 11–33.

Bažant, Z. P., Belytschko, T. B., and Chang, T. P., 1984, "Continuum Theory for Strain-Softening," *Journal of the Engineering Mechanics Division, ASCE*, Vol. 110, No. 12, pp. 1666–1692.

Bažant, Z. P., and Chang, T. P., 1984, "Instability of Nonlocal Continuum and Strain Averaging," *Journal of Engineering Mechanics, ASCE*, Vol. 110, No. 10, pp. 1441–1450.

Bažant, Z. P., Lin, F. B., and Pijaudier-Cabot, G., 1987, "Yield Limit Degradation: Nonlocal Continuum with Local Strains," Preprints, *International Conference on Computational Plasticity*, held in Barcelona, Eds., Oñate, E., Owen, R., and Hinton, E., University of Wales, Swansea, pp. 1757–1779.

Bažant, Z. P., and Pijaudier-Cabot, G., 1987, "Modeling of Distributed Damage by Nonlocal Continuum with Local Strain," *Proceedings 4th International Conference on Numerical Methods in Fracture Mechanics*, held in San Antonio, TX, Mar., Eds., Luxmore, A. R., et al., Pineridge Press, Swansea, U.K., pp. 411–432.

Bažant, Z. P., and Pijaudier-Cabot, G., 1987a, "Measurement of Characteristic Length of Nonlocal Continuum," Report 87-12/498 m, Center for Concrete and Geomaterials, Northwestern University, November. Also *Journal of Engineering Mechanics*, ASCE, in press.

Bažant, Z. P., and Zubelewicz, A., 1986, "Strain-Softening Bar and Beam: Exact Nonlocal Solution," *Report*, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL; also *International Journal of Solids and Structures*, in press.

Eringen, A. C., 1972, "Linear Theory of Nonlocal Elasticity and Dispersion of Plane Waves," *International Journal of Engineering Science*, No. 10, pp. 425–435.

Floegl, H., and Mang, H. A., 1981, "On Tension-Stiffening in Cracked Reinforced Concrete Slabs and Shells Considering Geometric and Physical Nonlinearity," *Ingenieurm Archiv*, Vol. 51, No. 314, pp. 215–242.

Kröner, E., 1967, "Elasticity Theory of Materials with Long-range Cohesive Forces," *International Journal of Solids and Structures*, No. 3, pp. 731–742.

Krumhansl, J. A., 1968, "Some Considerations of the Relations Between Solid State Physics and Generalized Continuum Mechanics," *Mechanics of Generalized Continua*, Eds., Kröner, E., Springer-Verlag, W. Berlin, W. Germany, pp. 298–331.

L'Hermite, R., and Grieu, J. J., 1952. "Etude expérimentales récentes sur le retrait des ciments et des bétons," *Annales I.T.B.T.P.* (Paris), Vol. 5, No. 52–53, pp. 494–51.

Ladeveze, P., "On an Anisotropic Damage Theory," (in French), *Internal Report*, No. 34, Laboratoire de Mécanique et Technologie, Cachan, France.

Lemaitre, J., Chaboche, J. L., "Mécanique des matériaux solides," Dunod-Bordas, Paris.

Mang, H., and Eberhardsteiner, J., 1986, "Collapse Analysis of Thin R. C. Shells on the Basis of a New Fracture Criterion," *Finite Element Analysis of Reinforced Concrete (Proceedings of the U.S.-Japan Seminar Held in Tokyo*, May 1985), Eds., Meyer, C., and Okamura, H., ASCE, New York, pp. 579–601.

Mazars, J., and Pijaudier-Cabot, G., 1986, "Continuum Damage Theory: Application to Concrete," *Internal Report*, Laboratorie de Mécanique et Technologie, Cachan, France, also *Journal of Engineering Mechanics*, ASCE, in press.

Needleman, A., 1987, "Material Rate Dependence and Mesh Sensitivity in Localization Problems," *Report*, Division of Engineering, Brown University, Providence, RI, January.

Pijaudier-Cabot, G., and Bažant, Z. P., 1986, "Nonlocal Damage Theory," *Journal of Engineering Mechanics, ASCE*, Vol. 113, No. 10, Oct., pp. 1512–1533, see also, Report No. 86-8/428n, Northwestern University, Evanston, IL.

Read, H. E., and Hegemier, G. P., 1984, "Strain-softening of Rock, Soil and Concrete—A Review Article," *Mech. Mat.*, Vol. 3, No. 4, pp. 271–294.

Sandler, I. S., 1986, "Strain-Softening for Static and Dynamic Problems," *Proceedings, Symposium on Constitutive Equations: Micro, Macro and Computational Aspects*, ASME Winter Annual Meeting, New Orleans, Dec., Eds., William, K. J., ASME, New York, pp. 217–231.

Saouridis, C., Mazars, J., and Billardon, R., 1987, "Evolution nonlocale de l'endommagement et limiteurs de localisation en éléments finis," *Proceedings 8th Congress of the Association Universitaire de Mécanique*, held in Nantes, Sept. A.U.M., Ed., pp. 147–148.

Schreyer, H. L., and Chen, Z., 1984, "The Effect of Localization on the Softening Behavior of Structural Members," *Proceedings, Symposium on Constitutive Equations: Micro, Macro, and Computational Aspects*, ASME Winter Annual Meeting, New Orleans, Ed., William, K., ASME, New York, December, pp. 193–203; also ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, December 1986, pp. 791–797.

APPENDIX C6

Measurement of Characteristic Length of Nonlocal Continuum

By Zdeněk P. Bažant,[1] Fellow, ASCE, and Gilles Pijaudier-Cabot,[2] Associate Member, ASCE

Abstract: The characteristic length of a heterogeneous brittle material such as concrete represents a material property that governs the minimum possible width of a zone of strain-softening damage in nonlocal continuum formulations or the minimum possible spacing of cracks in discrete fracture models. This length is determined experimentally. The basic idea is to compare the response of two types of specimens, one in which the tensile softening damage remains distributed and one in which it localizes. The latter type of specimen is an edge-notched tensile fracture specimen, and the former type of specimen is of the same shape but without notches. Localization of softening damage is prevented by gluing to the specimen surface a layer of parallel thin-steel rods and using a cross section of a minimum possible thickness that can be cast with a given aggregate. The characteristic length l is the ratio of the fracture energy (i.e., the energy dissipated per unit area, dimension N/m) to the energy dissipated per unit volume (dimension N/m^3). Evaluation of these energies from the present tests of concrete yields $l = 2.7$ times the maximum aggregate size.

Introduction

From a number of recent studies (e.g., Bažant 1986; Bažant and Pijaudier-Cabot 1987a, 1987b) it transpires that mathematical modeling of softening damage, such as cracking or void formation, necessitates a certain length parameter that is a property of the material and is called the characteristic length l. Although the characteristic length has been used in various nonlocal models for strain-softening, its value has so far been estimated only indirectly, probably with considerable uncertainty. No measurements of this essential parameter have yet been made. To present such measurements is the objective of this study [which is based on a report and a conference paper by Bažant and Pijaudier-Cabot (1987b, 1988)].

Depending on the type of material model, the characteristic length can have two interpretations:

1. If the softening damage is modeled by means of a strain-softening continuum, it is necessary to use some type of a localization limiter which prevents the damage zone from localizing to a zero volume. Thus, there must be a certain minimum thickness of the localization region that is related to the characteristic length of the continuum. This is in general described by a nonlocal continuum, in which certain strain variables are determined by averaging over the neighborhood of a point, the size of the neighborhood being determined by the characteristic length. Without the characteristic length, the strain-softening models for damage can correctly describe only the situations where the damage remains

[1]Prof. of Civ. Engrg., Northwestern Univ., Evanston, IL 60208.
[2]Res. Assoc., Dept. of Civ. Engrg., Rensselaer Polytech. Inst., Troy, NY 12181; formerly, Grad. Res. Asst., Northwestern Univ., Evanston, IL.

Note. Discussion open until September 1, 1989. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on March 10, 1988. This paper is part of the *Journal of Engineering Mechanics*, Vol. 115, No. 4, April, 1989. ©ASCE, ISSN 0733-9399/89/0004-0755/$1.00 + $.15 per page. Paper No. 23365.

distributed, not the situations where the damage tends to localize.

2. The discrete crack model with a softening cohesive zone at the crack front, introduced for concrete by Hillerborg et al. (1976), can correctly describe localized cracking or damage, but may yield ambiguous results for situations where damage or cracking remains distributed (Bažant 1986). Such situations include, in statics, certain problems of cracking due to cooling or shrinkage, or problems where a compression zone lies closely ahead of the cracking front (as in bending), and problems of cracking in reinforced materials. In dynamics, localization of damage into a single crack may be prevented for a certain period of time by inertia effects. The same may occur in statics when the strain-rate sensitivity of the material is introduced. To describe such situations, the discrete crack models must contain a material length parameter that indicates the minimum possible spacing of the cracks. This parameter is equivalent to the aforementioned characteristic length l, and the discrete and continuum models give roughly the same results if the displacement used in the stress-displacement softening relation for the cohesive zone of a crack is made equal to $l\epsilon$, where ϵ = the strain used in the softening stress-strain relation of a nonlocal continuum.

To determine the characteristic length, the basic idea is to measure the response of two types of specimens that are as similar as possible but such that in one type of specimen the damage, such as cracking, remains nearly homogeneously distributed, while in the other type, it localizes to the minimum volume that is possible for the given material. We will describe the development of such specimens and testing results, and then proceed to analyze the test results to obtain the characteristic length of concrete.

Test Specimens and Measurement Method

The fact that a tensile specimen that is forced by an elastic restraint to deform homogeneously exhibits a gradual post-peak strain-softening and has a higher strength than an unrestrained tensile specimen was discovered by L'Hermite (1960). With his coworkers, he tested specimens made by casting concrete into a steel pipe with an internal thread. Bonded to concrete by the thread, the steel pipe was loaded in tension and transmitted the tensile force to the concrete. Since the pipe was elastic, the force it carried was easily determined by measuring the deformation of the pipe, and the remainder of the tensile force could be ascribed to the concrete. The steel pipe no doubt prevented the tensile cracking from localizing into a single major crack. Companion tests in which concrete was bonded to the pipe only near the grips revealed significant differences in the strength limit and the post-peak behavior.

L'Hermite's tests, however, had a drawback due to the fact that the steel envelope has a higher Poisson ratio than concrete. As a result, the concrete in the pipe must have been subjected in these tests to significant lateral compressive stresses producing a confinement effect. Therefore, these tests cannot be regarded as uniaxial, and the presence of triaxial stresses complicates interpretation of the measurements.

As a modification of L'Hermite's idea, one might think of achieving a homogeneous strain state in tension by testing tensile specimens reinforced

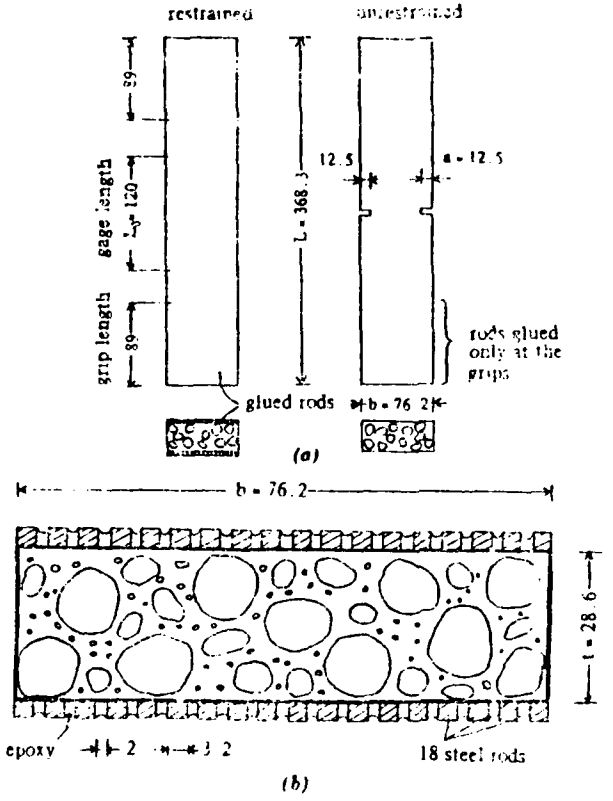


FIG. 1. (a) Notched and Unnotched Specimen Tested; (b) Arrangement of Steel Bars at Surface of Concrete (All Dimensions in mm)

by longitudinal steel bars instead of the pipe. However, due to the fact that the interface area between steel bars and concrete is relatively small and the bond is imperfect, a uniform strain state in reinforced tensile specimens is usually not achieved. A further source of nonuniformity may again be the difference of Poisson ratios between concrete and steel.

Alternatively, distributed cracking can be more easily achieved in specimens subjected to bending, due to the restraining effect of the compressed concrete layer in bent specimens. The nonuniformity of strain across the beam, however, complicates interpretation of the results. Moreover, the presence of a transverse strain gradient causes the tensile strength in bent specimens to be higher than in a uniform tensile strain field (this effect is actually explicable by the nonlocal character of the material).

To achieve perfect bond, one might cast concrete first and then glue to its surface metallic sheets. However, the mismatch in Poisson ratios would still produce lateral stresses.

In view of the foregoing considerations, the restrained and unrestrained specimens shown in Fig. 1 were selected. As is seen in Fig. 1(*b*), the longer

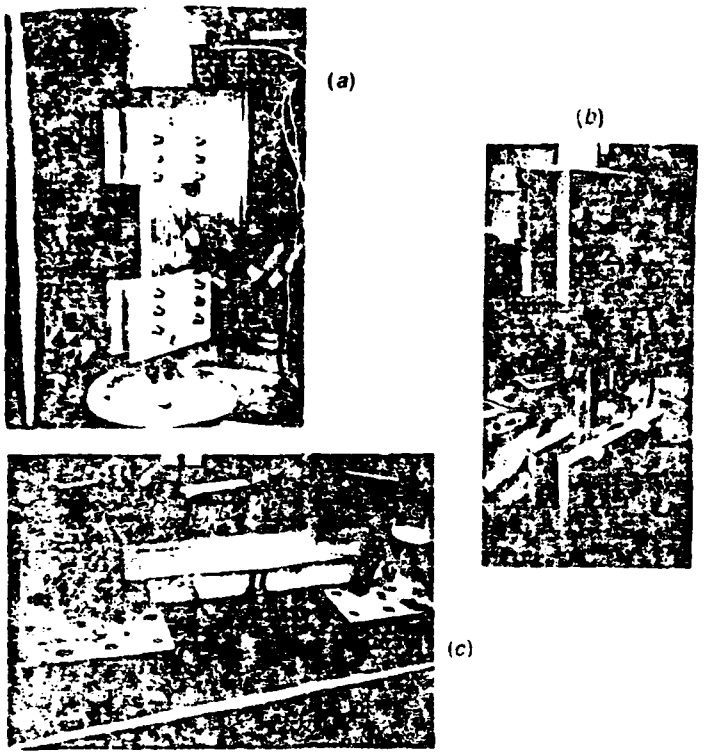


FIG. 2. Photographs of Specimen Tested

sides of the rectangular cross section are restrained by gluing to them with epoxy a system of regularly spaced, thin steel rods, which have relatively large gaps between them. These gaps, filled by epoxy, are quite deformable, because the elastic modulus of epoxy is much lower than that of concrete. Consequently, the set of thin steel rods cannot develop any significant transverse stresses and thus cannot interfere appreciably with the Poisson effect in concrete. Furthermore, by choosing the cross sections of the rods to be much smaller than the maximum aggregate size, the thin rods cannot affect the nonlocal properties of the material in the transverse direction. The rectangular cross section is elongated, so as to minimize the influence of the wall effect and the local stresses near the short sides of the cross section.

The thickness of the cross section was chosen to be only three times the maximum aggregate size. The reason for this was to ensure that the restraint due to the steel rods glued to the surface affects the entire thickness of the specimen. For much thicker specimens, the restraint of the interior would be incomplete and the strains could localize in the middle of the thickness. Selection of a circular cross section of a diameter only three aggregate sizes would yield a specimen whose behavior would no doubt be statistically more scattered. By using an elongated rectangular cross section, the number of aggregate pieces in the cross section is considerably increased, causing the

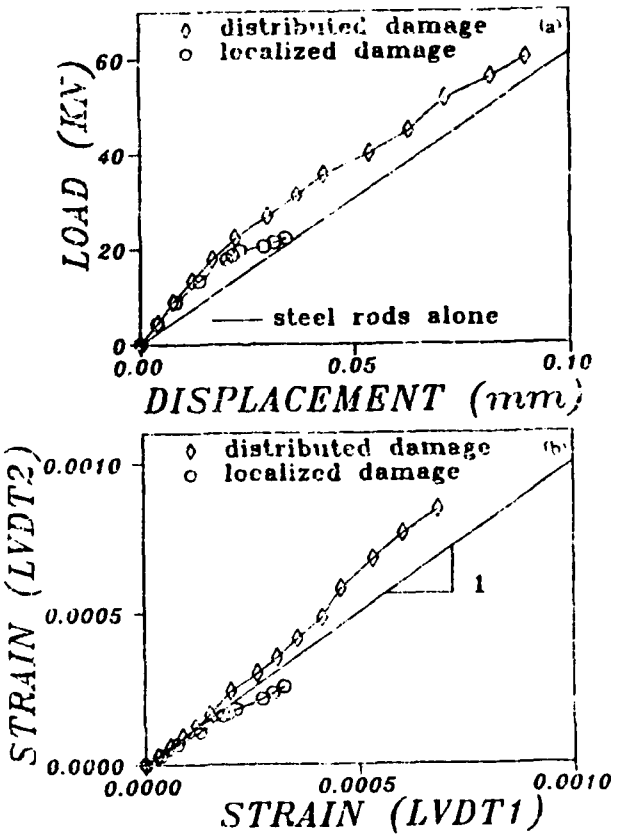


FIG. 3. (*a*) Load-Displacement Curves for Distributed and Localized Cracking Specimens; (*b*) Strains Measured on Each Side of Specimen

measured response, which roughly corresponds to the average property of the cross section, to be less scattered.

The dimensions of the specimen are all shown in Fig. 1 (in mm). The weight ratio of water-cement-sand-gravel in the mix was 1:2:2:0.6, and the maximum size of the aggregate was d_a = 9.53 mm. ASTM Type I cement was used. The specimens were removed from their plywood forms at 24 hours after casting and were then cured for 28 days in a room of relative humidity 95% and temperature 80° F.

At the ends of the specimen, metallic grips were glued by epoxy to the surface of the steel rods. In the companion unrestrained specimens, the surface steel rods were glued to the grips only within the area under the grips. To ensure that the tensile crack forms away from the grips and runs essentially normal to the axis, notches (of thickness 2.5 mm) were cut by a saw into the unrestrained specimens (Fig. 1).

The combined total cross section of the steel rods was selected so as to ensure that the tangential stiffness of both the restrained and the unrestrained specimen would always remain positive. Consequently, the stability of the

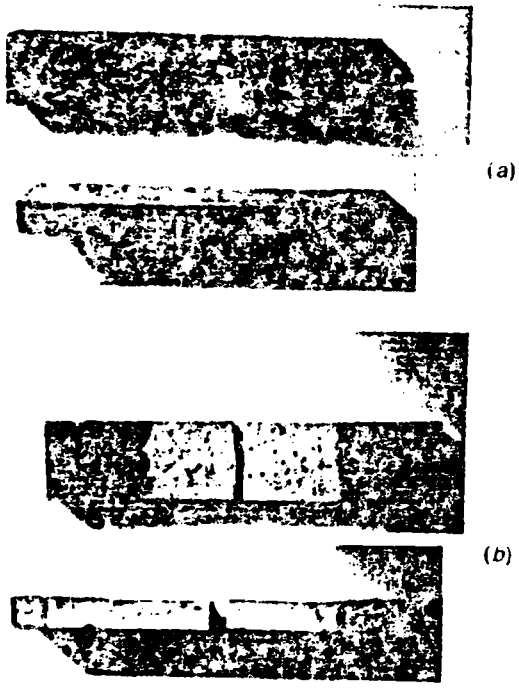


FIG. 4 Specimens after Failure: (*a*) Distributed Cracking; (*b*) Localized Cracking

specimen and strain localization could not depend on the stiffness of the testing machine.

The specimens were tested in tension in a closed-loop 100-kN MTS testing machine (Fig. 2). The loading was stroke controlled and was made at a constant displacement rate of 0.0038 mm/s, which produces the mean strain rate 2×10^{-5}/s. The relative displacements on a base length of 120 mm (Fig. 1) were measured by two symmetrically placed LVDT gages mounted on one face of the specimen and attached to the steel fibers (Fig. 2). Three specimens of each type were tested, but only two tests on unrestrained specimens could be exploited because of technical difficulties in setting up one of these experiments. Due to the small number of specimens, only the mean response is presented, and even the means might have considerable uncertainty. No conclusions on statistical variability can be drawn.

The plot of the average load versus displacement (mean of the measurements from the two LVDTs for each test) for the restrained and unrestrained specimens, which exhibit distributed and localized cracking, are shown in Fig. 3(*a*). The results confirm that the incremental stiffness has always been positive. It is seen that for the unrestrained specimens (unbonded rods), the load-displacement curve quickly approaches that for the steel rods alone. On the other hand, for the restrained specimens, the load-displacement curve remains for a long time significantly higher than that for the steel rods alone.

No macrocracks were observed on the short exposed sides of the restrained specimens (see Fig. 4), but a series of tiny microcracks could be detected.

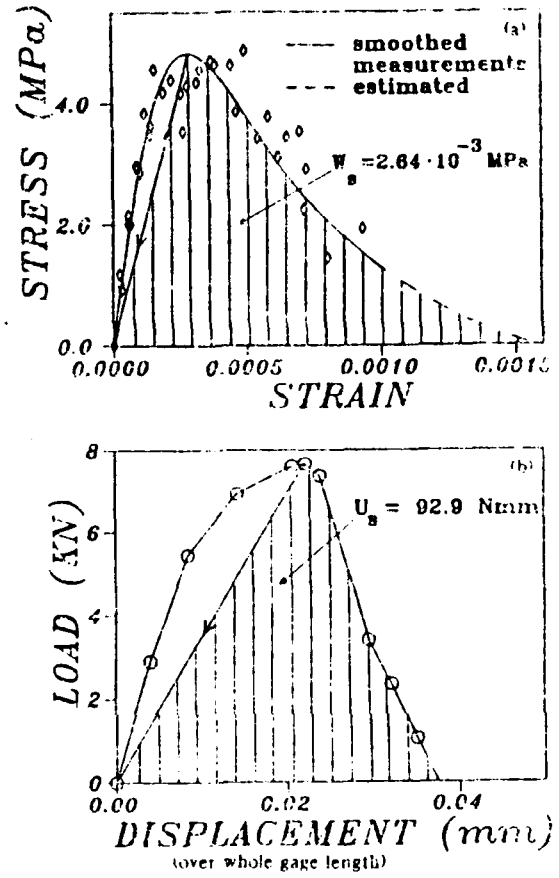


FIG. 5. (*a*) Stress-Strain Curve from Restrained Specimens; (*b*) Load-Displacement Curve from Unrestrained Specimens

It was also noticed that microcracking was somewhat more extensive farther away from the surface of the specimen. This may be explained by the restraint of the steel bars.

The results for the restrained specimens from Fig. 3 were converted to stress-strain curves for the restrained specimens with distributed damage [see Fig. 5(*a*)]. The final portion of the softening curve (shown dashed) had to be estimated by analogy with other test data. It may be noted that the relative scatter of the results in Fig. 5(*a*) is increased by the fact that the force in the steel is subtracted from the measured force values. This may explain why the measured response curve is not very smooth.

Fig. 3(*b*) shows the plot of the strains measured by the LVDT gages mounted on the left and right sides of the specimen. It may be noted that in the strain softening range, these two strains become markedly different, so that the specimen deformation becomes asymmetric, i.e., the specimen undergoes

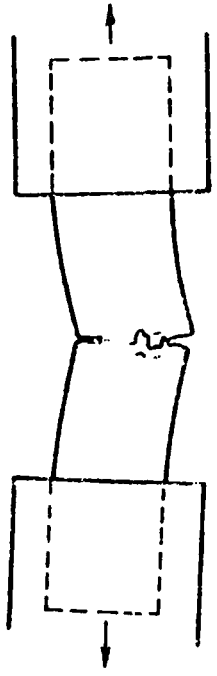


FIG. 6. Failure Mode of Unrestrained Specimens

bending, as shown in Fig. 6. This is not merely a consequence of random imperfections but is a manifestation of an instability that must occur in tensile fracture specimens of this type, as shown by Rots et al. (1987) and Hordijk et al. (1987). This implies that, in the unrestrained notched specimens, the cracking must propagate from one side of the cross section to the other, rather than happening simultaneously. The propagation of the cracking zone from one side of the specimen to the other helps to explain why the measurements indicate the maximum force carried by concrete in the unrestrained specimens to be considerably smaller than that in the restrained ones. However, even if the cracking zones propagated symmetrically, the fact that the failure cannot be simultaneous and that the specimen is notch-sensitive would require the maximum load to be less than that for the restrained specimen.

The test results for the unrestrained specimens with localized damage are plotted in Fig. 5(*b*) in terms of the average force P, carried by concrete versus the relative displacement u over the gage length. Interpretation in terms of the stress-strain curve is not attempted because the specimen state was inhomogeneous.

Calculation of Characteristic Length from Test Results

Taking the continuum viewpoint, we may consider the distributions of the macroscopic longitudinal normal strain along the gage length of the specimen to be as shown in Fig. 7, i.e., uniform for the restrained specimen, and localized, with a piecewise constant distribution, for the unrestrained specimen. Further, we assume that this localization begins to develop right at the peak stress point, at which the strain is $\epsilon = \epsilon_p$, as shown in the stress-strain diagram in Fig. 7. Although the nonlocalized states are stable up to a certain point on the descending stress-strain curve (Bazant 1976, 1986), the localization must begin, in a stable manner, precisely at the peak stress. This can be proven thermodynamically, on the basis of maximization of internal entropy (Bažant 1987). The fact that localization begins precisely at the peak stress point has also been documented by some recent measurements of deformation distributions in tensile specimens (Raiss 1986).

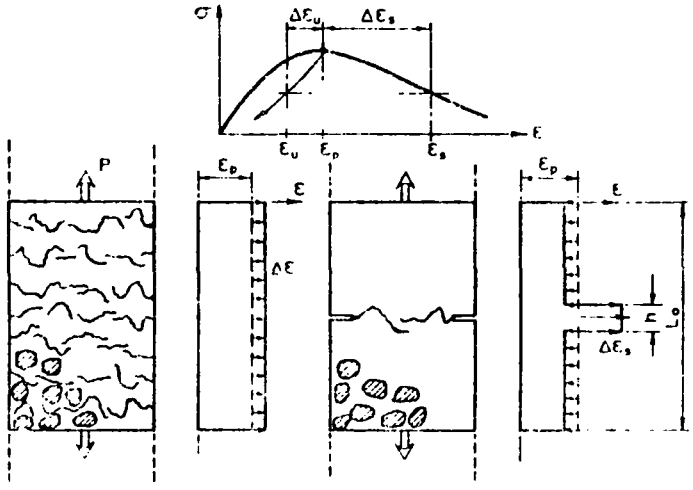


FIG. 7. Assumed Strain Distribution with and without Localization

The energy U_1 that is dissipated due to fracturing in the unbonded specimen with localized strain is represented by the area under the curves of axial force P versus axial relative displacement u for loading and for unloading from the peak-stress point, as shown by the cross-hatched area in Fig. 5(*b*). This energy should not include the blank area between the loading stress-strain curve and the unloading stress-strain curve, as proven in Bažant (1982). The energy U_2 dissipated by fracturing in the bonded specimen without strain localization must be determined in the same manner. This energy, representing the energy density, is shown by the cross-hatched area in Fig. 5(*a*).

We may now define the effective width h (Fig. 7) of the localized strain profile to be such that the stress-strain diagram for the localization zone (fracture process zone) would be the same as that for the bonded specimen with a homogeneous strain distribution. Thus, the balance of energy requires that $G_f = hW_u$, where $G_f = U_1/A_n$ = fracture energy of the material (dimension N/m), and A_n = the area of the net (ligament) cross section of the notched specimen. From this, the effective width of the localization zone is obtained as

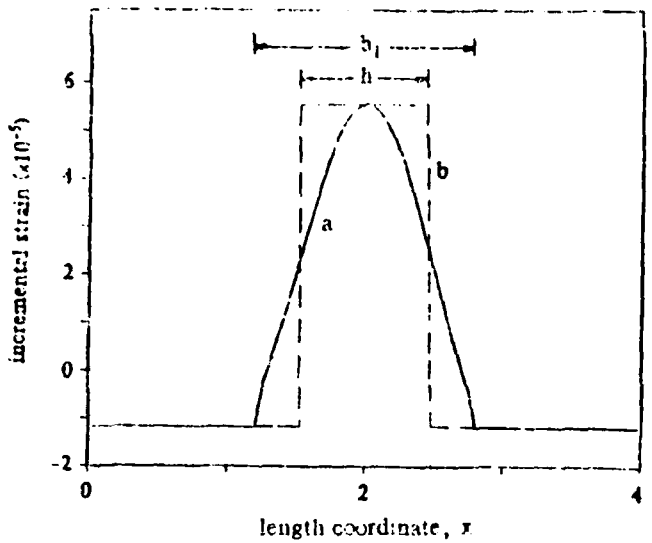


FIG. 8. Analytical Strain Profile at Localization Instability (after Bažant and Pijaudier-Cabot 1987)

$$h = \frac{G_f}{\bar{w}_t} \qquad (1)$$

The characteristic length of the nonlocal continuum can now be determined; however, the precise formulation of the nonlocal continuum must be specified. We consider the nonlocal continuum formulation from Bažant and Pijaudier-Cabot (1987b), in which nonlocal averaging is applied only to damage, and all other variables are treated as local. For this nonlocal damage theory, the profile of the continuum strain within the localization zone of a tensile specimen has been calculated, as shown by the graph in Fig. 8 taken from Bažant and Pijaudier-Cabot (1987b). This nonlocal formulation is equivalent in terms of energy if the area under the strain profile (curve a in Fig. 8) is the same as the area under the rectangular strain profile (curve b), which was implied in evaluating the test results according to Eq. 1. From the shape of this curve, one finds that the areas are equal if

$$h_1 = \alpha l \qquad (2)$$

in which $\alpha = 1.23$; and h_1 = width of the zone of localized damage. Furthermore, for this nonlocal continuum model (Bažant and Pijaudier-Cabot 1987c), it has been shown that

$$h_1 = \beta l \qquad (3)$$

in which $\beta = 1.09$. From Eqs. 1–3 it follows that

$$l = \frac{\alpha G_f}{\beta \bar{w}_t} \qquad (4)$$

Coefficients α and β are particular to the chosen type of nonlocal continuum formulation; for the one considered here, they yield $\alpha/\beta = 1.02$. Thus, the characteristic length is essentially the same as the width of the strain-softening zone, obtained under the assumption of a uniform strain within the zone. This is approximately true also for other variants of the nonlocal continuum formulation, and so we have, approximately

$$l = \frac{G_f}{W_t} \qquad (5)$$

G_f and W_t were evaluated from the results plotted Fig. 5. It was assumed that unloading followed a straight line passing through the origin, in accordance with the nonlocal model selected.

Based on the present test results, Eq. 4 yields

$$l = 2.7 d_a \qquad (6)$$

in which d_a = the maximum aggregate size. This is one basic result of the present study. However, verification by further tests with different aggregate sizes is desirable.

The value of the characteristic length obtained in Eq. 6 for the present experiments is consistent with previous estimates obtained when the crack-band model was proposed by Bažant and Oh (1983). At that time, the characteristic length was inferred indirectly, by using only the fracture test results and optimizing their fits for specimens of various geometries and sizes, made from different concretes. In that study, the optimum fit was obtained approximately for $l = 3d_a$.

The fracture energy, which is obtained from the unbonded test with strain localization according to the relation $G_f = U_t/A_0$, is, of course, an average value, obtained under the simplifying assumption that the fracture energy does not vary as the crack band propagates along the ligament cross section. From the value of the cross-hatched area in Fig. 5(*b*), one gets, for the present experiments, $G_f = 0.0635$ N/mm. This value corresponds to the work-of-fracture definition of fracture energy, as introduced for concrete by Hillerborg (1976) and used in the present RILEM recommendation (1985). From the size-effect tests, a distinctly lower value, namely $G_f = 0.039$ N/mm, has been previously obtained by Bažant and Pfeiffer (1987). The difference is no doubt due to the fact that the fracture energy obtained by the size-effect method has a somewhat different meaning, representing the fracture energy for the limiting case of an infinitely large specimen, in which the fracture process zone can develop without interference with the specimen boundaries, and is the same for any specimen geometry.

It should be also noted that Irwin's (1953) estimate of the length of the nonlinear zone, namely

$$l_c = \frac{EG_f}{\sigma_p^2} \qquad (7)$$

in which σ_p = peak stress (strength); and E = Young's elastic modulus, is sometimes also called the characteristic length (Hillerborg et al. 1976). From the presently measured values, we get

$$l_c = 8.63'' \qquad (8)$$

This value is much larger than the characteristic length l for the nonlocal continuum. The meaning of these two characteristic lengths is, of course, entirely different: l_r characterizes the length l_p of the fracture process zone ahead of the crack tip in the direction of crack propagation. (Hillerborg and Petersson have shown that $l_p = l_r/2$.) On the other hand, l characterizes the effective width of the fracture process zone, governed by material heterogeneity. In the continuum description, l represents the minimum possible width of the softening zone, and in a discrete model for cracks in concrete, l represents the minimum possible spacing of discrete cracks.

It may also be noted that the minimum possible spacing of discrete cracks has been lacking from the discrete crack models used so far. This did not matter for the analysis of localized fracture, with a single major crack; however, for general applications, a length characterizing the minimum possible spacing of discrete cracks needs to be added to the discrete crack models of Hillerborg type (Bažant 1986).

Since the term "characteristic length" has long been established in the atistical and nonlocal theories of heterogeneous media, l_r for concrete should etter be called the length (or size) of the fracture process zone or of the nonlinear zone. For metals, l_r is called the size of the yielding zone, as introduced by Irwin (1958).

Conclusions

1. The characteristic length of a nonlocal continuum that models strain-softening damage can be experimentally determined by measuring the energy dissipation of specimens in which the softening damage remains distributed and those in which it becomes localized.

2. The characteristic length of a nonlocal continuum is the ratio of the fracture energy (i.e., the energy dissipated per unit area, dimension N/m) to the energy dissipated per unit volume of the material (dimension N/m^3) (Eq. 5).

3. Distributed softening damage can be experimentally observed on concrete specimens of a rectangular cross section, whose width is approximately the minimum possible width allowed by the maximum aggregate size and whose long sides are restrained by a layer of many thin longitudinal steel rods glued to the specimen surface.

4. The characteristic length of concrete, which may be calculated from Eqs. 4 or 5, is obtained to be about 2.7 times the maximum aggregate size.

Acknowledgments

The experiments were supported by the U.S. Air Force Office of Scientific Research under contract F49620-87-C0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu, and the theoretical work was supported by the U.S. National Science Foundation under Grant MSM 8700830, monitored by Dr. Albert S. Kobayashi.

Appendix. References

Bažant, Z. P. (1976). "Instability, ductility, and size effect in strain softening concrete." *J. Engrg. Mech. Div.*, ASCE, 102(2), 331–344 (closure 104, 501–502).
Bažant, Z. P. (1983). "Crack band model for fracture of geomaterials." *Proc., 4th Int. Conf. on Numerical Methods in Geomechanics*, Edmonton, Canada, Z. Eisenstein, ed., Vol. 3, 1137–1152.
Bažant, Z. P. (1986). "Mechanics of distributed cracking." *Appl. Mech. Rev.*, ASME, 39(5), 675–705.
Bažant, Z. P. (1987). "Stable states and paths of structures with plasticity or damage." *Structural Engineering Report No. 87-10/606s*, Dept. of Civil Engineering, Northwestern Univ., Evanston, Ill.; also (1988), *J. Engrg. Mech.*, ASCE, 114(12), 2013–2034.
Bažant, Z. P., and Oh, B. H. (1983). "Crack band theory for fracture of concrete." *Mater. Struct.*, RILEM, Paris, France, 16, 155–177.
Bažant, Z. P., and Pfeiffer, P. (1987). "Fracture energy of concrete: Its definition and determination from size effect tests." *Concrete durability, Proc. K. and B. Mather Int. Conf.*, J. Scanlon, ed., Vol. 3, 89–109.
Bažant, Z. P., and Pijaudier-Cabot, G. (1987a). "Modeling of distributed cracking by nonlocal continuum with local strain." *Proc., 4th. Int. Conf. on Numerical Methods in Fracture Mechanics*, San Antonio, Texas, A. R. Luxmoore et al., eds., Pineridge Press, Swansea, U.K., 411–432.
Bažant, Z. P., and Pijaudier-Cabot, G. (1987b). "Measurement of characteristic length of nonlocal continuum." *Report No. 87-12/498 m*, Center for Concrete and Geomaterials, Northwestern University, Evanston, Ill.
Bažant, Z. P., and Pijaudier-Cabot, G. (1987c). "Nonlocal damage: Continuum model and localization instability." *Report No. 87-2/428n-1*, Center for Concrete and Geomaterials, Northwestern University, Evanston, Ill; also *J. Appl. Mech.*, ASME, 55 (1988), 287–293.
Bažant, Z. P., and Pijaudier-Cabot, G. (1988). "Nonlocal continuum damage and measurement of characteristic length." *Mech. of Composite Materials—1988*, G. J. Dvorak and N. Laws, eds., AMD Vol. 92, ASME, 79–85 (presented at Joint ASME/SES Appl. Mech. and Engr. Sci. Conf., Berkeley, Calif., June, 1988).
Hillerborg, A., Modéer, M., and Petersson, P. E. (1976). "Analysis of crack formation and crack growth in concrete by means of fracture mechanics and finite elements." *Cement Concr. Res.*, 6(6), 773–782.
Hordijk, D. A., Reinhardt, H. W., and Cornelissen, H. A. W. (1987). "Fracture mechanics parameters of concrete from uniaxial tensile tests as influenced by specimen length." *Preprints of the SEM/RILEM Int. Conf. on Fracture of Concrete and Rock*, Houston, Tex., S. P. Shah and S. Swartz, eds., SES (Society of Experimental Mechanics), 138–149.
Irwin, G. R. (1958). "Fracture." *Encyclopaedia of physics*, Vol. VI, Springer, Berlin, Germany.
L'Hermite, R. (1960). "Volume changes of concrete." *Proc., 4th Int. Symp. on the Chemistry of Cement*, Washington, D.C., 659–702.
Raiss, M. E. (1986). "Observation of the development of fracture process zone in concrete under tension," thesis presented to Imperial College, London, U.K., in partial fulfillment of the requirements for the degree of Doctor of Philosophy.
RILEM Committee 50-FMC (1985). "Determination of the fracture energy of mortar and concrete by means of three-point-bend tests on notched beams," RILEM Draft Recommendation, *Mater. Struct.* (RILEM, Paris), 18(106), 285–290.
Rots, J. G., Bordijk, D. A., and de Borst, R. (1987). "Numerical simulation of concrete fracture in direct tension." *Proc., 4th Int. Conf. on Numerical Methods in Fracture Mechanics*, San Antonio, Tex., A. R. Luxmoore et al., eds., Pineridge Press, Swansea, U.K., 457–471.

Softening Instability: Part I —Localization Into a Planar Band

Zdeněk P. Bažant
Professor,
Department of Civil Engineering,
Northwestern University,
Evanston, IL 60208
Mem. ASME

Distributed damage such as cracking in heterogeneous brittle materials may be approximately described by a strain-softening continuum. To make analytical solutions feasible, the continuum is assumed to be local but localization of softening strain into a region of vanishing volume is precluded by requiring that the softening region, assumed to be in a state of homogeneous strain, must have a certain minimum thickness which is a material property. Exact conditions of stability of an initially uniform strain field against strain localization are obtained for the case of an infinite layer in which the strain localizes into an infinite planar band. First, the problem is solved for small strain. Then a linearized incremental solution is obtained taking into account geometrical nonlinearity of strain. The stability condition is shown to depend on the ratio of the layer thickness to the softening band thickness. It is found that if this ratio is not too large compared to 1, the state of homogeneous strain may be stable well into the softening range. Part II of this study applies Eshelby's theorem to determine the conditions of localization into ellipsoidal regions in infinite space, and also solves localization into circular or spherical regions in finite bodies.

Introduction

Distributed damage such as cracking or void nucleation and growth may be macroscopically described by a constitutive law for a continuum that exhibits strain-softening (Bažant, 1986). In the softening range, the matrix of incremental moduli of the material is not positive-definite. After Hadamard (1903) pointed out that such a condition implies an imaginary wave speed, strain-softening has been considered an unacceptable property for a continuum and some scholars have argued that strain-softening simply does not exist (Read and Hegemier, 1984; Sandler, 1984).

The various mathematical arguments for the nonexistence of strain-softening, however, overlooked one crucial experimental fact pointed out in 1974 by Bažant (1986): A material in the strain-softening state also has available to it another matrix of incremental moduli which applies for unloading and is positive-definite. This fact makes an essential difference. It causes that the material in a strain softening-state can propagate unloading waves, and that solutions to various dynamic and static problems with strain-softening exist, even for the classical, local continuum. Some solutions are unique, representing limits of finite-element discretizations, which converge quite rapidly (Bažant and Belytschko, 1985, 1987; Belytschko et al., 1986). In some cases, strain-softening apparently leads to chaos (Belytschko et al., 1986). Generally, the dynamic strain-softening problems (for a local continuum) belong to the class of "ill-posed" initia. boundary-value problems, which are well known and accepted as realistic in other branches of physics (e.g., Joseph et al., 1985; Joseph, 1986, Yoo, 1985). Therefore, there is nothing fundamentally wrong with the concept of strain-softening from the mathematical viewpoint.

From the physical viewpoint, however, strain-softening in a classical local continuum is an unacceptable concept because structures are predicted to fail with zero energy dissipation. The reason is that the zone of softening damage often localizes into a line or surface while the energy dissipation per unit volume is finite. This difficulty, however, may be overcome by introducing some form of localization limiters (Bažant and Belytschko, 1987)—mathematical formulations that prevent the strain-softening damage to localize into a region of zero volume. Among numerous possibilities, the limitation to localization may be best introduced by treating the softening damage as nonlocal while the elastic behavior, including unloading, is treated as local (Pijaudier-Cabot and Bažant, 1986; Bažant, Lin, and Pijaudier-Cabot, 1987; Bažant and Pijaudier-Cabot, 1987). Physical justification of this nonlocal formulation has been given on the basis of homogenization of a quasi-periodic microcrack array (Bažant, 1987). Other possibilities are, e.g., the use of strain gradient or higher-order derivatives (or gradients) of the yield function in the constitutive equation (e.g., Floegl and Mang, 1981; Mang and Eberhardsteiner, 1985; Schreyer and Chen, 1986; Triantafyllidis and Aifantis, 1986).

In this paper we will seek analytical solutions to

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March 13, 1987; final revision, February 1, 1988. Paper No. 88-WA/APM-34.

multidimensional localization. They appear to be possible with the well-known approach which uses the simplest and earliest type of localization limiter proposed in 1974 by Bažant (1976) and later implemented in the finite-element crack band model (Bažant and Cedolin, 1979, 1980; Bažant, 1982; Bažant and Oh, 1983, 1984). Despite the presence of strain-softening, the constitutive equation is in this approach treated as local, but the softening damage is not allowed to localize into a region whose size is less than a certain characteristic length l of the material. This length is determined empirically by fitting test data and roughly corresponds to the size of the representative volume used in statistical theory of heterogeneous materials (Kröner, 1967). For concrete, l is of the order of the maximum aggregate size. Although this approach is, admittedly, crude and does not permit resolving the distribution of damage density throughout the softening region, it has given some surprisingly good results for the overall structural response. This was confirmed by comparisons with extensive fracture test data (Bažant and Oh, 1983; Bažant 1982, 1984) and later also by comparisons with accurate nonlocal finite-element solutions (Bažant and Pijaudier-Cabot, 1987; Bažant, Pijaudier-Cabot, and Pan, 1986; Bažant and Zubelewicz, 1986; see the Appendix).

The method of analysis of strain-localization instability due to softening and its thermodynamic basis were formulated in 1974 (Bažant, 1976) and the stability conditions for one-dimensional localization in a bar and flexural localization in a beam were derived. The equivalence to failure analysis on the basis of fracture energy was also demonstrated (Bažant, 1982). Several other studies reanalyzed the one-dimensional localization from other viewpoints yielding equivalent results (e.g., Ottosen, 1986).

Rudnicki and Rice (1975) and Rice (1976) made important pioneering studies of localization into a planar band in an infinite space. They focused their studies primarily on localization caused by the geometrically nonlinear effects of finite strain before the peak of the stress-strain diagram (i.e., in the plastic hardening range), but they also obtained some critical states for negative values of the plastic hardening modulus. Rudnicki (1977) further analyzed, on the basis of Eshelby's theorem, an infinite space that contains a weakened zone of ellipsoidal shape, determined for localization into such a zone the critical neutral equilibrium states of homogeneous stress and strain, and demonstrated various cases of localization instabilities in the hardening as well as softening regime. Although the studies of Rudnicki and Rice represented an important advance, they did not actually address the stability conditions but were confined to neutral equilibrium conditions for the critical state. They did not consider a general incremental stress-strain relation but were limited to von Mises plasticity (Rudnicki and Rice 1975) or Drucker-Prager plasticity (Rudnicki, 1977), in some cases enhanced with a vertex hardening term. They also did not consider bodies of finite dimensions, for which the size of the localization region usually has a major influence on the critical state and, in the case of planar localization bands, did not consider unloading to occur outside the localization band, which is important for finite bodies. The present study attempts to take all these conditions and effects into account.

The purpose of the present study is to obtain exact analytical solutions for some multidimensional localization problems with softening, using the method introduced in 1974 by Bažant (1976, 1979), which is based on a local continuum with an imposed lower bound on the size h of the softening region. In this paper, we will analyze localization of strain into a planar band, both without and with geometrical nonlinearity of strain and, in a subsequent companion paper, we will analyze in a similar manner the localization of strain into ellipsoidal regions, including the special case of spherical and circular regions. The solution for ellipsoidal regions will be based on the celebrated Eshelby's theorem for eigenstrain in an ellipsoidal inclusion in an infinite elastic solid.

The present study (including Part II) will deal only with stability of equilibrium states, and not with bifurcations of the equilibrium path. As is known from Shanley's column theory, such bifurcations can occur at increasing load and do not necessarily coincide with the limit of stable equilibrium.

Before embarking on our analysis, comments on several related aspects are in order. It has been widely believed that softening is properly treated by fracture mechanics, in particular, by line-crack fracture models with a cohesive zone characterized by a softening stress-displacement relation. One problem with this approach may be that the stress-displacement relation might not be unique and might depend on the fracture specimen geometry. This is suggested by nonlocal finite-element results as well as the difficulties in fitting test data for various specimen geometries on the basis of the same material properties. Another problem is that this approach is limited to single fractures or noninteracting multiple fractures, and becomes unobjective (with regard to the choice of mesh) when multiple interacting cracks are present (Bažant, 1986). The formulation could be made objective by imposing a certain minimum admissible spacing of cracks as a material property, but this runs into difficulty when the interacting cracks are not parallel. (Line cracks with a fixed minimum spacing h are, of course, macroscopically equivalent to distributed cracking if the cumulative cracking strain over distance h is made to be equal to the crack opening displacement.)

Softening Band Within a Finite Layer or Infinite Solid

Let us analyze the stability of softening that is localized in an infinite layer which is called the softening band (or localization band) and forms inside an infinite layer of thickness L ($L \geq h$); see Fig. 1. The minimum possible thickness h of the band is assumed to be a material property, proportional to the characteristic length l. The layer is initially in equilibrium under a uniform (homogeneous) state of strain ϵ_{ij}^0 and stress σ_{ij}^0, assumed to be in the strain-softening range (Fig. 2). Latin lower case subscripts refer to Cartesian coordinates x_i ($i = 1, 2, 3$) of material points in the initial state. The initial

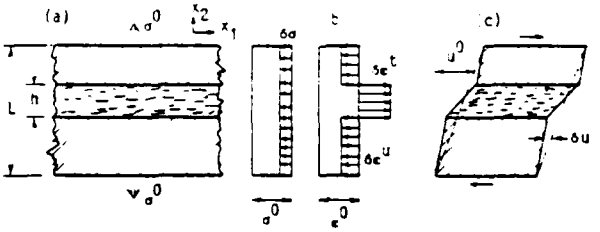


Fig. 1 Planar localization band in a layer

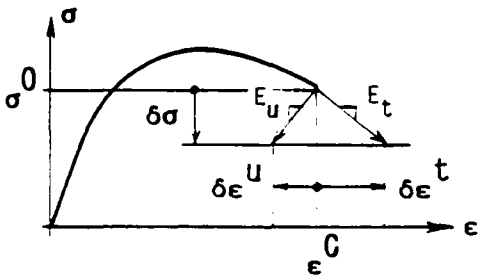


Fig. 2 Stress-strain diagram with softening and unloading

equilibrium is disturbed by small incremental displacements δu_i whose gradients $\delta u_{i,j}$ are uniform both inside and outside the band (subscripts preceded by a comma denote derivatives). The values of $\delta u_{i,j}$ inside and outside the band are denoted as $\delta u^l_{i,j}$ and $\delta u^u_{i,j}$ and are assumed to represent further loading and unloading, respectively, and so the increments $\delta u_{i,j}$ represent strain localization. We choose axis x_2 to be normal to the layer (Fig. 1). As the boundary conditions, we assume that the surface points of the layer are fixed during the incremental deformation, i.e., $\delta u_i = 0$ at the surfaces $x_2 = 0$ and $x_2 = L$ of the layer. Assuming homogeneous strains inside and outside of the band, compatibility of displacements at the band surface requires that

$$\delta u^u_{i,1} = \delta u^l_{i,1} = 0, \quad \delta u^u_{i,3} = \delta u^l_{i,3} = 0 \tag{1}$$

$$h\delta u^l_{i,2} + (L-h)\delta u^u_{i,2} = 0 \quad (i = 1,2,3). \tag{2}$$

We assume that the incremental material properties are characterized by incremental moduli tensors D^l_{ijpq} and D^u_{ijpq} for loading and unloading (Fig. 2). These two tensors may be either prescribed by the given constitutive law directly as functions of ϵ^0_{ij} and possibly other state variables, or they may be implied indirectly. The latter case occurs, e.g., for continuum damage mechanics. The crucial fact is that D^u_{ijpq} differs from D^l_{ijpq} and is always positive-definite, even if there is strain-softening. Noting that $\epsilon_{ij} = (u_{i,j} + u_{j,i})/2$ and $D_{ijkm} = D_{jimk}$, we may write the incremental stress-strain relations as follows (repetition of subscripts implies Einstein's summation rule)

$$\delta\sigma_{ij} = D^l_{ijkm}\delta\epsilon^l_{km} = D^l_{ijkm}\delta u^u_{k,m} \quad \text{for loading} \tag{3}$$

$$\delta\sigma^u_{ij} = D^u_{ijkm}\delta\epsilon^u_{km} = D^u_{ijkm}\delta u^u_{k,m} \quad \text{for unloading.} \tag{4}$$

Stability of the initial equilibrium state may be decided on the basis of the work ΔW that must be done on the layer per unit area in the (x_1, x_3)-plane in order to produce the increments δu_i. This work, which represents the Helmholtz free energy under isothermal conditions and the total energy under adiabatic conditions, may be expanded as $\Delta W = \delta W + \delta^2 W + \ldots$ where δW = first variation (first-order work done by σ^0_{ij} on $\delta u_{i,j}$) and $\delta^2 W$ = second variation (the second-order work). If the initial state is an equilibrium state, δW must vanish. We, therefore, need to calculate only $\delta^2 W$. Using equations (3), (4), and (1), as well as the relation $\delta u^l_{i,2} = -\delta u^u_{i,2}(L-h)/h$ which follows from equation (2), we get

$$\delta^2 W = \frac{h}{2}\delta\sigma^l_{ij}\delta u^l_{i,2} - \frac{L-h}{2}\delta\sigma^u_{ij}\delta u^u_{i,2}$$

$$= \frac{h}{2}(\delta\sigma^l_{ij} - \delta\sigma^u_{ij})\delta u^l_{i,2} = \frac{h}{2}(D^l_{2j12}\delta u^l_{j,2} - D^u_{2j12}\delta u^u_{j,2})\delta u^l_{i,2} \tag{5}$$

$$\delta^2 W = \frac{h}{2}(D^l_{2j12} - \frac{h}{L-h}D^u_{2j12})\delta u^l_{j,2}\delta u^l_{i,2}. \tag{6}$$

We may denote

$$Z_{ij} = D^l_{2j12} + \frac{h}{L-h}D^u_{2j12} \tag{7}$$

or

$$\mathbf{Z} = [Z_{ij}] = \begin{bmatrix} D^l_{2112}, & D^l_{2122}, & D^l_{2132} \\ D^l_{2212}, & D^l_{2222}, & D^l_{2232} \\ D^l_{2312}, & D^l_{2322}, & D^l_{2332} \end{bmatrix} + \frac{h}{L-h}\begin{bmatrix} D^u_{2112}, & D^u_{2122}, & D^u_{2132} \\ D^u_{2212}, & D^u_{2222}, & D^u_{2232} \\ D^u_{2312}, & D^u_{2322}, & D^u_{2332} \end{bmatrix} \tag{8}$$

Matrix $\mathbf{Z}$ is symmetric ($Z_{ij} = Z_{ji}$) if and only if $D^l_{2j12} = D^l_{i212}$ and $D^u_{2j12} = D^u_{i212}$. This is assured if D^l_{ijkm} as well as D^u_{ijkm} has a symmetric matrix. Our derivation, however, is valid in general for nonsymmetric Z_{ij} or D^l_{ijkm}. The necessary stability condition may be stated according to equation (6) as follows

$$\delta^2 W = \frac{h}{2}\delta u^l_{i,2}Z_{ij}\delta u^l_{j,2} > 0 \text{ for any } \delta u^l_{i,2}. \tag{9}$$

The expression in equation (9) is a quadratic form. If $\delta^2 W$ is positive for all possible variations $\delta u^l_{i,2}$, no change of the initial state can occur if no work is done on the body, and so the initial state of uniform strain is stable. On the other hand, if $\delta^2 W$ is negative for some variation $\delta u^l_{i,2}$, the change of the initial state releases energy, which is ultimately dissipated as heat, thus increasing the entropy S of the system by $\Delta S = -\delta^2 W/T$ where T = absolute temperature. Hence, due to the second law of thermodynamics, such a change of initial state must happen spontaneously. Therefore, we must conclude that the intial state of homogeneous strain is unstable if $\delta^2 W$ (or Z_{ij}) ceases to be positive-definite. Positive-definiteness of the 3×3 matrix Z_{ij} is a necessary condition of stability. (We cannot claim equation (9) to be sufficient for stability since we have not analyzed all possible localization modes. However, changing $>$ to $<$ yields a sufficient condition for instability.)

Positive-definiteness of the 6×6 matrix $\mathbf{Z}^* = \mathbf{D}_l + \mathbf{D}_u h/(L-h)$ (where $\mathbf{D}_l$ and $\mathbf{D}_u$ are the 6×6 matrices of incremental moduli for loading and unloading) implies stability. However, the body can be stable even if $\mathbf{Z}^*$ is not positive-definite.

If the softening band is infinitely thin ($h/L \to 0$), or the layer is infinitely thick ($L/h \to \infty$), we have $Z_{ij} = D^l_{2j12}$, and so matrix $\mathbf{Z}$ loses positive-definiteness when the (3×3) matrix of D^l_{2j12} ceases to be positive-definite. This condition, whose special case for von Mises plasticity was obtained by Rudnicki and Rice (1975) and Rice (1976), indicates that instability may occur right at the peak of the stress-strain diagram. However, for a continuum approximation of a heterogeneous material, for which h must be finite, the loss of stability can occur only after the strain undergoes a finite increment beyond the peak of the stress-strain diagram. For $L \to h$, the softening band is always stable.

Strain-localization instability in a uniaxially stressed bar represents a one-dimensional problem. It also may be obtained from the present three-dimensional solution as the special case for which the softening material is incrementally orthotropic, with $D^l_{2222} = E_l$ (< 0) and $D^u_{2222} = E_u$ (> 0) as the only nonzero incremental moduli. Equations (7) and (9) then yield the stability condition presented in 1974 (Bažant, 1976)

$$-\frac{E_l}{E_u} < \frac{h}{L-h} \quad \text{or} \quad \frac{E_l}{h} - \frac{E_u}{L-h} > 0. \tag{10}$$

This simple condition clearly illustrates that for finite L, h the localization instability can occur only at a finite slope E_l, i.e., some finite distance beyond the peak of the stress-strain diagram.

If the end of the bar at $x = L$ is not fixed but has an elastic support with spring constant C_s, one may obtain the solution by imagining the bar length to be augmented to length L', the additional length $L' - L$ having the same stiffness as the spring; i.e., $L' - L = C_s E_u$. Therefore, the stability condition is $-E_l E_u < h(L' - h)$, which yields the condition given before (Bažant, 1976).

The stability condition for a layer whose surface points are supported by an elastic foundation (Fig. 3) may be treated similarly, i.e., by adding to the layer of thickness L another layer of thickness $L' - L$ such that its stiffness is equivalent to the given foundation modulus.

The case when the outer surfaces of the layer are kept at

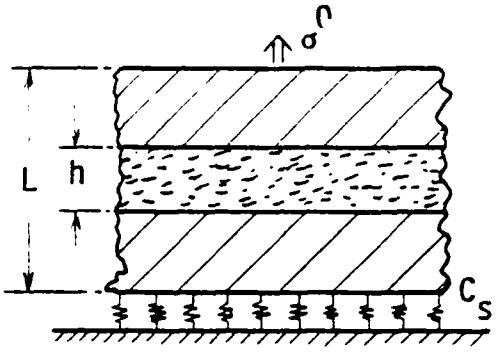


Fig. 3 Layer with localization band and an elastic foundation

constant load $p_i^0 = \sigma_{ij}^0 n_j$ during localization is equivalent to adding a layer of inifinite thickness $L' - L$. Therefore, the necessary stability condition is that the 3×3 matrix of D^t_{2j2} be positive-definite, i.e., no softening can occur.

Another simple type of localization may be caused by softening in pure shear, Fig. 2(c). In this case, $\delta u_{1,2} \neq 0$, $\delta u_{2,2} = \delta u_{3,2} = 0$, and $D^t_{2112} = G_t$, $D^u_{2112} = G_u$ are the shear moduli. According to equations (7) and (9), the necessary stability condition is

$$-\frac{G_t}{G_u} < \frac{h}{L-h}. \tag{11}$$

Consider now the limiting case of an infinite space. While in a layer an infinitely long softening band must be parallel to the layer surface, for an infinite space, the softening band can have any orientation. Since $L \to \infty$, we have $Z_{ij} = D^t_{2j2}$, and if the band is normal to axis x_2, stability requires that $\delta^2 W = \delta u^t_{i,2} D^t_{2j2} \delta u^t_{j,2} h/2$ = positive-definite. To generalize this condition to a band of arbitrary orientation, we may carry out an arbitrary rotation transformation of coordinates from x_i to x'_i. The transformation relations are $x_i = c_{ij} x'_j$ where c_{ij} are the direction cosines of the old coordinate base vectors in the new coordinates. According to the rules of transformation of tensors, we now have $2\delta^2 W = h(c_{ik} c_{2m} \delta u'_{k,m})\ (c_{2p} c_{iq} c_{jr} c_{2s} D^t_{pqrs})\ (c_{jv} c_{2n} \delta u'_{v,n})$ where the primes refer to the new coordinates x'_i. Noting that $c_{ik} c_{iv} = \delta_{kv}$, $c_{jr} c_{jv} = \delta_{rv}$, we obtain

$$2\delta^2 W = h \delta a_{pq} D^t_{pqsr} \delta a_{sr} \quad \text{with} \quad \delta a_{pq} = c_{2p} \delta u_{2,m} c_{2m}. \tag{12}$$

For arbitrary rotations, δa_{pq} can have any values. Thus the 6×6 matrix of moduli D^t_{pqsr}, and also D^t_{ijkm}, must be positive-definite in order to insure that the strain cannot localize into an infinite planar band of any orientation. The same requirement was stated by Hadamard (1903), who derived it from the condition that the wave speed would not become imaginary. Hadamard's analysis, however, implied that $\mathbf{D}^u = \mathbf{D}^t$.

Finally, consider another case of boundary conditions: The case when the plane surfaces of an infinite layer slide freely over rigid bodies during the localization. The boundary conditions at $x_2 = 0$ and $x_2 = L$ now are $\delta\sigma_{21} = \delta\sigma_{23} = 0$. Similar to equation (6), we now have

$$\delta^2 W = \frac{h}{2}\left(D^t_{2222} + \frac{h}{L-h} D^u_{2222}\right)\delta u^t_{2,2} \delta u^t_{2,2} + \tfrac{1}{2}(\delta N_{11} \delta u_{1,1} - \delta N_{11} \delta u_{1,1} - 2\delta N_{13} \delta u_{1,3}) \tag{13}$$

in which δN_{11}, δN_{33}, δN_{13} are homogeneously distributed in-plane incremental normal and shear force resultants over the whole thickness of the layer. Overall equilibrium now requires that $\delta N_{11} = \delta N_{33} = \delta N_{13} = 0$. Hence, the necessary condition of stability against localization is

$$D^t_{2222} + \frac{h}{L-h} D^u_{2222} > 0. \tag{14}$$

Extension to Geometrically Nonlinear Effects in Finite Strain

Strain localization in an infinite planar band can be also easily solved even when the geometrically nonlinear effects are taken into account. The geometrical nonlinearity we consider is due to finite strains or finite rotations, not to changes in the geometrical configuration of the structure. The deviations δu_i and $\delta\epsilon_{ij}$ from the initial state, characterized by homogeneous stresses σ^0_{ij}, are again considered to be infinitely small, and the energy expression ΔW that governs stability is second-order small. However, the contribution to $\delta^2 W$, which arises from the geometrically nonlinear finite-strain expression, is also second-order small, and so it must be included. This may be done if the material stress increment $\delta\sigma_{ij}$ in the work expression is replaced by the mixed (first) Piola-Kirchhoff stress increment $\delta\tau_{ij}$, which is referred to the initial state and is nonsymmetric (Malvern, 1969). As is well known, $\delta\tau_{ij} = \delta T_{ij} - \sigma^0_{ik}\, \delta u_{j,k} + \sigma^0_{ij}\, \delta u_{k,k}$ where δT_{ij} = material increment of the true (Cauchy) stress. Since neither $\delta\sigma_{ij}$ nor $\delta\tau_{ij}$ is invariant at coordinate rotations, one must use in the incremental stress-strain relation the objective stress increment $\delta\sigma_{ij}$ (representing the objective stress rate times the increment of time); $\delta\sigma_{ij}$ is symmetric.

The relationship between $\delta\sigma_{ij}$ and $\delta\tau_{ij}$ may be written (Bažant, 1971) in the general form

$$\delta\tau_{ij} = \delta\sigma_{ij} + R_{ij\,km\,rs}\sigma^0_{rs}\delta u_{k,m} = H_{ijkm}\delta u_{k,m} \tag{15}$$

$$H_{ijkm} = D_{ijkm} + R_{ij\,km\,rs}\sigma^0_{rs} \tag{16}$$

where we substituted $\delta\sigma_{ij} = D_{ijkm}\delta u_{k,m}$. Coefficients $R_{ij\,pq\,rs}$ are certain constants which take into account the geometrical nonlinearity of finite strain. The values of $R_{ij\,km\,pq}$ are different for various possible choices of the objective stress rate and the associated type of the finite-strain tensor. The expressions which are admissible according to the requirements of tensorial invariance and objectivity are (Bažant 1971)

$$R_{ij\,pq\,rs}\sigma^0_{rs} = \delta_{ip}\sigma^0_{qj} - \alpha(\delta_{ip}\sigma^0_{qj} + \delta_{iq}\sigma^0_{pj} + \delta_{jp}\sigma^0_{qi} + \delta_{jq}\sigma^0_{pi}) \tag{17}$$

where α can be an arbitrary constant. The case $\alpha = 0$ corresponds to Truesdell's objective stress rate and Green's (Lagrangian) finite-strain tensor; the case $\alpha = 1/4$ to the finite-strain theory of Biot; the case $\alpha = 1/2$ to the Jaumann's objective stress rate and the finite strain theory of Southwell, Biezeno-Hencky, and Neuber (and to the logarithmic strain); and the case $\alpha = 3/4$ to Cotter-Rivlin's convected stress rate (see, equations (14*a*), (15), (17*a*), and (22) in Bažant, 1971). In general, α can be any real number. For each different α-value, however, different values of incremental moduli D_{ijpq} must be used so as to obtain physically equivalent results; generally [cf. Bažant, 1971, equation (19)]

$$D_{ijpq} = [D_{ijpq}]_{\alpha=0} - \alpha(\delta_{ip}\sigma^0_{qj} + \delta_{iq}\sigma^0_{pj} + \delta_{jp}\sigma^0_{qi} + \delta_{jq}\sigma^0_{pi}). \tag{18}$$

From equation (5), in which $\delta\sigma_{ij}$ must now be replaced by $\delta\tau_{ij}$, the necessary stability condition is

$$\begin{aligned} 2\delta^2 W &= h\delta\tau^t_{2i}\delta u^t_{i,2} + (L-h)\delta\tau^u_{2i}\delta u^u_{i,2} \\ &= h(\delta\tau^t_{2i} - \delta\tau^u_{2i})u^t_{i,2} = t(H^t_{2i/2}\delta u^t_{i,2} - H^u_{2i/2}\delta u^u_{i,2})\delta u^t_{i,2} \\ &= h\delta u^t_{i,2} Z_{ij}\delta u^t_{j,2} > 0 \quad \text{for any } \delta u^t_{i,2} \end{aligned} \tag{19}$$

in which

$$\begin{aligned} Z_{ij} &= H^t_{2i/2} - \frac{h}{L-h} H^u_{2i/2} \\ &= D^t_{2i/2} - \frac{h}{L-h} D^u_{2i/2} - \frac{L}{L-h} R_{2i/2rs}\sigma^u_{rs}. \end{aligned} \tag{20}$$

In particular, if we use the Lagrangian (Green's) finite strain

which is associated with Truesdell's objective stress rate ($\alpha = 0$), we have

$$Z_{ij} = \left[D^t_{ijkl} - \frac{h}{L-h} D^u_{ijkl}\right]_{\alpha=0} - \frac{L}{L-h}\sigma^0_{ik}\delta_{lj} \quad (21)$$

If we use Jaumann's objective stress rate ($\alpha = 1/2$), we have

$$Z_{ij} = \left[D^t_{ijkl} - \frac{h}{L-h} D^u_{ijkl}\right]_{\alpha=\frac{1}{2}} - \frac{L}{2(L-h)}(\sigma^0_{ik}\delta_{lj} - \sigma^0_{jk}\delta_{li} - \sigma^0_{kl}\delta_{ij} - \sigma^0_{ij}). \quad (22)$$

It must be emphasized that the conditions of positive-definiteness of matrix $[Z_{ij}]$ for various possible values of α are all physically equivalent [even though for each different α-value the incremental moduli D^t_{ijkl} and D^u_{ijkl} are different, as dictated by equation (18)]. This fact often has been forgotten. The use of various types of objective stress rates often has been discussed in the literature on the basis of some imagined numerical convenience, and the fact that the incremental moduli cannot be the same for different objective stress rates has been overlooked (despite the analysis in Bažant, 1971). These practices, widespread in the finite-element literature, are of course incorrect.

As for the identification of the incremental moduli from test data, it can be done only in reference to a certain chosen objective stress rate. Different values of the moduli must result for different choices although the results are equivalent physically (see, Bažant, 1971).

The condition $\det [Z_{ij}] = 0$ obviously represents the critical state. The special case of this condition for which the constitutive law consists of von Mises plasticity enhanced by vertex hardening, $L/h \to \infty$ (infinite space), $\alpha = 1/2$ (Jaumann's rate), and the initial stress σ^0_{ij} is pure shear σ^0_{12}, was derived by Rudnicki and Rice (1975). Their analysis was concerned only with the critical state of neutral equilibrium, rather than with stability. The values of the unloading moduli D^u_{ijkm} and the fact that they are different from D^t_{ijkm} and positive-definite were irrelevant for their analysis.

Rudnicki and Rice (1975) showed that, due to geometrical nonlinearity, the critical state of strain localization can develop in plastic materials while the matrix of D^t_{ijkm} is still positive-definite, i.e., before the final yield plateau of the stress-strain diagram is reached. This may explain the formation of shear bands in plastic (nonsoftening) materials. The destabilizing effect is then due exclusively to geometrical nonlinearity. According to a geometrically linear analysis (small strain theory), strain localization could develop in plastic (nonsoftening) materials only upon reaching the yield plateau but not earlier.

From equation (20), it appears that the geometrical nonlinearity can have a significant effect on strain localization only if the incremental moduli for loading are of the same order of magnitude as the initial stresses σ^0_{ij}; precisely, if $\max |D^t_{ijkm}|$ and $\max |\sigma^0_{ij}|$ are of the same order of magnitude. (The unloading moduli D^u_{ijkm} of structural materials are always several orders of magnitude larger.) Thus the importance of geometrical nonlinearity depends on D^t_{ijkm}. For strain-softening type of localization, the geometrical nonlinearity can be important only if the instability develops at a very small negative slope of the stress-strain diagram, which occurs very close to the peak stress point. This can occur only if the layer thickness L is much larger than l. If $L - h \ll l$, instability occurs when the downward slope of the stress-strain diagram (Fig. 2) is of the same order of magnitude as the initial elastic modulus E, and this is inevitably orders of magnitude larger than the initial stresses σ^0_{ij}.

Therefore, we must conclude that the role of geometrical nonlinearity in localization due to strain-softening cannot be significant except if the localization occurs immediately after the peak point of the stress-strain diagram (Fig. 2). Then the values of the strains at localization instability obtained by geometrically nonlinear and linear analyses cannot differ significantly since they must both be close to the peak point of the stress-strain diagram. For the postcritical large deformations, however, geometrical nonlinearity is no doubt always important.

Conclusions

The condition of stability of a layer against localization of strain into a band reduces to the condition of positive-definiteness of a certain matrix which represents a weighted average of the matrices of loading and unloading moduli of the material. The weight of the loading (softening) moduli increases with the ratio of the band thickness to the layer thickness. If this ratio tends to zero (infinite space), the instability is determined solely by the loading moduli and occurs as soon as this matrix ceases to be positive-definite. When the band thickness is finite, instability does not occur until the initial strain exceeds the strain at peak stress by a finite amount. Geometrical nonlinearity of strain has a significant effect only for a very small band-to-layer thickness ratio, and unstable localization then occurs near the peak stress state.

Extension to localization that is not unidirectional and numerical examples are relegated to a subsequent companion paper.

Acknowledgment

Partial financial support under AFOSR Contract F49620-87-C-0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu, is gratefully acknowledged.

References

Bažant, Z. P., 1971, "A Correlation Study of Formulations of Incremental Deformation and Stability of Continuous Bodies," ASME JOURNAL OF APPLIED MECHANICS, Vol. 38, No. 4, pp. 919–928.

Bažant, Z. P., 1976, "Instability, Ductility and Size Effect in Strain-Softening Concrete," *Journal of Engineering Mechanics*, ASCE, Vol. 102, pp. 331–344; Discussions, Vol. 103, pp. 357–358, 775–777, Vol. 104, pp. 501–502; based on Structural Engineering Report No. 74-8/640, Northwestern University, Aug. 1974, and private communication to K. Gerstle, S. Sture, and A. Ingraffea, in Boulder, CO, June 1974.

Bažant, Z. P., 1982, "Crack Band Model for Fracture of Geomaterials," *Proceedings of the 4th International Conference on Numerical Methods in Geomechanics*, Edmonton, AB, Canada, June, Z. Eisenstein, ed., Vol. 3, (invited lectures), pp. 1137–1152.

Bažant, Z. P., 1984, "Mechanics of Fracture and Progressive Cracking in Concrete Structures," in *Fracture Mechanics Applied to Concrete Structures*, Sih, G. C., and DiTomasso, A., eds., Martinus Nijhoff, The Hague, pp. 1–94.

Bažant, Z. P., 1986, "Mechanics of Distributed Cracking," *Applied Mechanics Reviews*, Vol. 39, pp. 675–705.

Bažant, Z. P., 1987, "Why Continuum Damage is Nonlocal: Justification by Quasi-Periodic Microcrack Array," *Mechanics Research Communications*, Vol. 14, (5/6), pp. 407–420.

Bažant, Z. P., Belytschko, T. B., and Chang, T.-P., 1984, "Continuum Theory for Strain Softening," *Journal of Engineering Mechanics*, Vol. 110, pp. 1666–1692.

Bažant, Z. P., and Belytschko, T. B., 1985, "Wave Propagation in a Strain-Softening Bar: Exact Solution," *Journal of Engineering Mechanics*, ASCE, Vol. 111 (3).

Bažant, Z. P., and Belytschko, T. B., 1987, "Strain-Softening Continuum Damage: Localization and Size Effect," Preprints, International Conference on Constitutive Laws for Engineering Materials, Vol. 1, held in Tucson, AZ, Jan. 1987, ed. by C. S. Desai et al., University of Arizona, Elsevier, New York, pp. 11–33.

Bažant, Z. P., and Cedolin, L., 1979, "Blunt Crack Band Propagation in Finite Element Analysis," *Journal of Engineering Mechanics*, ASCE, Vol. 105, (2), pp. 297–315.

Bažant, Z. P., and Cedolin, L., 1980, "Fracture Mechanics of Reinforced Concrete," *Journal of Engineering Mechanics*, ASCE, Vol. 106, (6), pp. 1287–1306; Discussion, Vol. 108, 1982, pp. 464–471.

Bažant, Z. P., Lin, F.-B., and Pijaudier-Cabot, G., 1987, "Yield Limit Degradation: Nonlocal Continuum Model With Local Strain," Preprints, International Conference on Computational Plasticity, held in Barcelona, Apr. 1987.

R. Owen, and E. Hinton, (U. of Wales, Swansea) and E. Oñate, (Techn. U. Barcelona), eds.

Bažant, Z. P., and Oh, B. H., 1983, "Crack Band Theory for Fracture of Concrete," *Materiaux et Constructions* (Rilem, Paris), Vol. 16, (93), pp. 155–177.

Bažant, Z. P., and Oh, B. H., 1984, "Rock Fracture via Softening Finite Elements," *Journal of Engineering Mechanics*, ASCE, Vol. 110, (7), pp. 1015–1035.

Bažant, Z. P., and Panula, L., 1978, "Statistical Stability Effects in Concrete Failure," *Journal of Engineering Mechanics*, ASCE, Vol. 104, (5), pp. 1195–1212.

Bažant, Z. P., Pijaudier-Cabot, G., and Pan, J., 1986, "Strain Softening Beams and Frames," Report, Center for Concrete and Geomaterials, Northwestern University.

Bažant, Z. P. and Pijaudier-Cabot, G., 1987, "Nonlocal Damage: Continuum Model and Localization Instability," Report No. 87-2/498n, Center for Concrete and Geomaterials, Northwestern University.

Bažant, Z. P., and Zubelewicz, A., 1986, "Strain-Softening Bar and Beam: Exact Nonlocal Solution," *International Journal of Solids and Structures*, in press.

Belytschko, T., Bažant, Z. P., Hyun, Y.-W., and Chang, T.-P., 1986, "Strain-Softening Materials and Finite Element Solutions," *Computers and Structures*, Vol. 23, (2), pp. 163–180.

Floegl, H., and Mang, H. A., 1981, "On Tension Stiffening in Cracked Reinforced Concrete Slabs and Shells Considering Geometric and Physical Nonlinearity," *Ingenieur-Archiv*, Vol. 51, pp. 215–242.

Hadamard, J., 1903, *Lecons sur la propagation des ondes*, Hermann, Paris. Chapter VI.

Joseph, D. D., Renardy, M., and Saut, J.-C., 1985, "Hyperbolicity and Change of Type in the Flow of Viscoelastic Fluids," *Archive for Rational Mechanics and Analysis*, Vol. 87, (3), pp. 213–251.

Joseph, D. D., 1986, "Change of Type and Loss of Evolution in the Flow of Viscoelastic Fluids," *Journal of Non-Newtonian Fluid Mechanics*, Vol. 20, pp. 117–141.

Kröner, E., 1967, "Elasticity Theory of Materials with Long-range Cohesive Forces," *International Journal of Solids and Structures*, Vol. 3, pp. 731–742.

Malvern, L. E., 1969, *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall, Englewood Cliffs, NJ.

Mang, H., and Eberhardsteiner, J., 1985, "Collapse Analysis of Thin R.C. Shells on the Basis of a New Fracture Criterion," U.S.-Japan Seminar on the Finite Element Analysis of Reinforced Concrete Structures, Preprints, pp. 217–238 (and *Proceedings, ASCE*, New York 1986, C. Meyer, and H. Okamura, eds.

Ottosen, N. S., 1986, "Thermodynamic Consequences of Strain Softening in Tension," *Journal of Engineering Mechanics*, ASCE, Vol. 112 (11), pp. 1152–1164.

Pijaudier-Cabot, G., and Bažant, Z., 1986, "Nonlocal Damage Theory," Report No. 85-8/428n, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL; also ASCE *Journal of Engineering Mechanics*, Vol. 113 (1987) in press.

Read, H. E., and Hegemier, G. P., 1984, "Strain-Softening of Rock, Soil and Concrete," A Review Article, *Mechanics of Materials*, Vol. 3, (4), pp. 271–294.

Rice, J. R., 1976, "The Localization of Plastic Deformation," Preprints, 14th International Congress of International Union of Theoretical and Applied Mechanics held in Delft, Netherlands, W. Koiter, ed., North Holland, Amsterdam.

Rudnicki, J. W., and Rice, J. R., 1975, "Conditions for the Localization of Deformation in Pressure Sensitive Dilatant Materials," *Journal of the Mechanics and Physics of Solids*, Vol. 23, pp. 371–394.

Rudnicki, J. W., 1977, "The Inception of Faulting in a Rock Mass With a Weakened Zone," *Journal of Geophysical Research*, Vol. 82, No. 5, pp. 844–854.

Sandler, I. S., 1984, "Strain-Softening for Static and Dynamic Problems," *Proceedings, Symposium on Constitutive Equations: Micro, Macro and Computational Aspects*, ASME Winter Annual Meeting held in New Orleans, Dec., K. William, ed., ASME, New York, pp. 217–231.

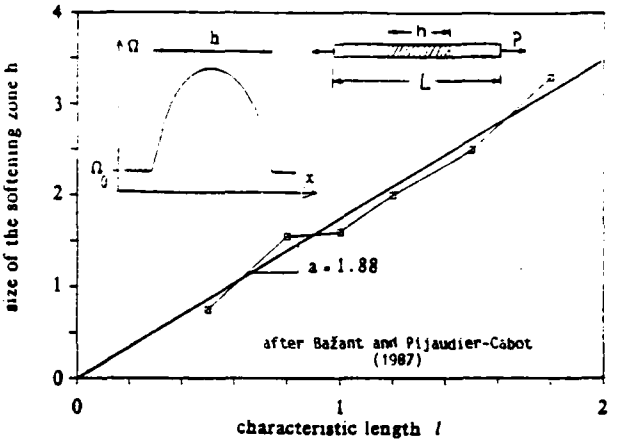


Fig. 4 Comparison of approximate local solution based on prescribed softening segment length *h* (dashed lines) with the exact nonlocal solution (solid lines), obtained for a bar of length *L* by Bazant and Pijaudier-Cabot (1987)

Schreyer, H. L., and Chen, Z., 1986, "One-Dimensional Softening With Localization," ASME JOURNAL OF APPLIED MECHANICS, Vol. 53, No. 4, pp. 791–797.

Triantafyllidis, N., and Aifantis, E. C., 1986, "A Gradient Approach to Localization of Deformation, I: Hyperelastic Materials," *Journal of Elasticity*, Vol. 16, pp. 225–237.

Yoo, J. Y., Ahrens, M., and Joseph, D. D., 1985, "Hyperbolicity and Change of Type in Sink Flow," *Journal of Fluid Mechanics*, Vol. 171, pp. 203–214.

APPENDIX

In the crack band model, the minimum possible size h of the softening zone in a local continuum is considered to be a material constant. The degree to which this approach can capture nonlocal behavior is illustrated by the comparison in Fig. 4, taken from Bažant and Pijaudier-Cabot (1987). The data points show the length h of the strain-softening segment in a bar of length $L > h$, made of a nonlocal strain-softening material with characteristic length l. Nonlocal is only the fracturing (or damage) strain while the elastic strain, including unloading, is local. The results represent accurate solutions of an integral equation for the static uniaxial localization instability in a bar that is initially strained uniformly, with a strain in the strain-softening range. In contrast to the present analysis, the length h of the localization segment is unknown and is solved as a function of the assumed characteristic length l. The results show that approximately $h = a$ where a = constant ($a = 1.88$). Since l is a material constant, h may also be approximately considered to be a constant, as assumed in the crack band model.

Softening Instability: Part II—Localization Into Ellipsoidal Regions

Zdeněk P. Bažant
Professor,
Department of Civil Engineering,
Northwestern University,
Evanston, IL 60208
Mem. ASME

Extending the preceding study of exact solutions for finite-size strain-softening regions in layers and infinite space, exact solution of localization instability is obtained for the localization of strain into an ellipsoidal region in an infinite solid. The solution exploits Eshelby's theorem for eigenstrains in elliptical inclusions in an infinite elastic solid. The special cases of localization of strain into a spherical region in three dimensions and into a circular region in two dimensions are further solved for finite solids – spheres in 3D and circles in 2D. The solutions show that even if the body is infinite the localization into finite regions of such shapes cannot take place at the start of strain-softening (a state corresponding to the peak of the stress-strain diagram) but at a finite strain-softening slope. If the size of the body relative to the size of the softening region is decreased and the boundary is restrained, homogeneous strain-softening remains stable into a larger strain. The results also can be used as checks for finite element programs for strain-softening. The present solutions determine only stability of equilibration states but not bifurcations of the equilibrium path.

Introduction

The strain-localization solutions in the preceding paper (Bažant, 1987) deal with unidirectional localization of strain into an infinite planar band. If the body is finite, localization into such a band does not represent an exact solution because the boundary conditions cannot be satisfied. In this paper, we will seek exact solutions for multidirectional localization due to strain-softening in finite regions. In particular, we will study localization into ellipsoidal regions, including the special cases of a spherical region in three dimensions and a circular region in two dimensions. All definitions and notations from the preceding paper (Bažant, 1987, Part I) are retained.

Softening Ellipsoidal Region in Infinite Solid

This type of strain-localization instability can be solved by application of Eshelby's (1957) theorem for ellipsoidal inclusions with uniform eigenstrain. Consider an ellipsoidal hole, Fig. 1(*a*), in a homogeneous infinite medium that is elastic and is characterized by elastic moduli matrix $\mathbf{D}_u$. We imagine fitting and glueing into this hole an ellipsoidal plug of the same material, Fig. 1(*a*), which must first be deformed by uniform strain ϵ^0 (the eigenstrain) in order to fit into the hole perfectly (note that a uniform strain changes an ellipsoid into another ellipsoid). Then the strain in the plug is unfrozen, which causes the plug to deform with the surrounding medium to attain a new equilibrium state. The famous discovery of Eshelby (1957) was that if the plug is ellipsoidal and the elastic medium is homogeneous and infinite, the strain increment ϵ^e in the plug which occurs during this deformation is *uniform* and is expressed as

$$\epsilon^e_{ij} = S_{ijkm}\epsilon^0_{km}. \tag{1}$$

S_{ijkm} are components of a fourth-rank tensor which depend only on the ratios a_1/a_3 and a_2/a_3 of the principal axes of the ellipsoid and, and for the special case of isotropic materials, on Poisson ratio ν_u; see, e.g., Mura (1982) and Christensen (1979). Due to symmetry of ϵ^0_{ij} and ϵ^e_{ij}, $S_{ijkm} = S_{jikm} = S_{ijmk}$, but in general $S_{ijkm} \neq S_{kmij}$. Coefficients S_{ijkm} are, in general, expressed by elliptic integrals; see also Mura (1982). Extension of Eshelby's theorem to generally anisotropic materials was later accomplished at Northwestern University by Kinoshita and Mura (1971) and Lin and Mura (1973).

It will be convenient to rewrite equation (1) in a matrix form

$$\boldsymbol{\epsilon}^e = \mathbf{Q}_u\boldsymbol{\epsilon}^0 \tag{2}$$

or

Contributed by the Applied Mechanics Division for presentation at the Winter Annual Meeting, Chicago, IL, November 28 to December 2, 1988, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.
Discussion on this paper should be addressed to the Editorial Department, ASME, United Engineering Center, 345 East 47th Street, New York, N.Y. 10017, and will be accepted until two months after final publication of the paper itself in the JOURNAL OF APPLIED MECHANICS. Manuscript received by ASME Applied Mechanics Division, March 13, 1987; revision, February 1, 1988.
Paper No. 88-WA/APM-35.

$$\begin{Bmatrix} \epsilon^e_{11} \\ \epsilon^e_{22} \\ \epsilon^e_{33} \\ 2\epsilon^e_{12} \\ 2\epsilon^e_{23} \\ 2\epsilon^e_{31} \end{Bmatrix} = \left[\begin{array}{ccc|c|cc} S_{1111} & S_{1122} & S_{1133} & S_{1112} & S_{1123} & S_{1131} \\ S_{2211} & S_{2222} & S_{2233} & S_{2212} & S_{2223} & S_{2231} \\ S_{3311} & S_{3322} & S_{3333} & S_{3312} & S_{3323} & S_{3331} \\ \hline 2S_{1211} & 2S_{1222} & 2S_{1233} & 2S_{1212} & 2S_{1223} & 2S_{1231} \\ \hline 2S_{2311} & 2S_{2322} & 2S_{2333} & 2S_{2312} & 2S_{2323} & 2S_{2331} \\ 2S_{3111} & 2S_{3122} & 2S_{3133} & 2S_{3112} & 2S_{3123} & 2S_{3131} \end{array}\right] \begin{Bmatrix} \epsilon^0_{11} \\ \epsilon^0_{22} \\ \epsilon^0_{33} \\ 2\epsilon^0_{12} \\ 2\epsilon^0_{23} \\ 2\epsilon^0_{31} \end{Bmatrix} \tag{3}$$

in which $\epsilon = (\epsilon_{11}, \epsilon_{22}, \epsilon_{33}, 2\epsilon_{12}, 2\epsilon_{23}, 2\epsilon_{31})^T$ and superscript T denotes the transpose of a matrix.

For isotropic materials, the only nonzero elements of matrix $\mathbf{Q}$ are those between the dashed lines marked in equation (3), which is the same as for the stiffness matrix. The factors 2 in matrix $\mathbf{Q}_u$ in equation (3) are due to the fact that the column matrix of strains is (6×1) rather than (9×1) and, therefore, must involve shear angles $2\epsilon_{12}$, $2\epsilon_{23}$ and $2\epsilon_{31}$ rather than tensorial shear strain components ϵ_{12}, ϵ_{23} and ϵ_{31}, or else $\sigma^T\delta\epsilon$ where $\sigma^T = (\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{23}, \sigma_{31})$ would not be a correct work expression. (The work expression $\sigma_{ij}\delta\epsilon_{ij}$, as well as the sum implied in equation (1) for each fixed i,j, has 9 terms in the sum, not 6.) For example, writing out the terms of equation (1) we have

$$\epsilon^e_{11} = \ldots + (S_{1112}\epsilon^0_{12} + S_{1121}\epsilon^0_{21}) + \ldots = \ldots + S_{1112}(2\epsilon^0_{12}) + \ldots \tag{4}$$

while the factors 2 arise as follows

$$\begin{aligned} 2\epsilon^e_{12} &= 2[\ \ldots + S_{1233}\epsilon^0_{33} + (S_{1112}\epsilon^0_{12} + S_{1121}\epsilon^0_{21}) \\ &\quad + (S_{1223}\epsilon^0_{23} + S_{1232}\epsilon^0_{32}) + \ldots\ ] \\ &= \ldots (2S_{1233})\epsilon^0_{33} + (2S_{1212})2\epsilon^0_{12} \\ &\quad + (2S_{1223})2\epsilon^0_{23} + \ldots \end{aligned} \tag{5}$$

The stress in the ellipsoidal plug, σ^e (which is uniform), may be expressed according to Hooke's law as

$$\sigma^e = \mathbf{D}_u(\epsilon^e - \epsilon^0). \tag{6}$$

After substituting $\epsilon^0 = \mathbf{Q}_u^{-1}\epsilon^e$, according to equation (2), we get $\sigma^e = \mathbf{D}_u(\epsilon^e - \mathbf{Q}_u^{-1}\epsilon^e)$ or

$$\sigma^e = \mathbf{D}_u(\mathbf{1} - \mathbf{Q}_u^{-1})\epsilon^e \tag{7}$$

where $\mathbf{1}$ is a unit 6×6 matrix. The surface tractions that the ellipsoidal plug exerts upon the surrounding infinite medium, Fig. 1(*a*), are $p^e_i = \sigma^e_{ij}\, n_j$, in which n_j denotes the components of a unit normal $\mathbf{n}$ of the ellipsoidal surface (pointed from the ellipsoid outward).

Consider now infinitesimal variations $\delta\mathbf{u}$, $\delta\boldsymbol{\epsilon}$, $\delta\boldsymbol{\sigma}$ from the initial equilibrium state of uniform strain ϵ^0 in an infinite homogeneous anisotropic solid (without any hole). The matrices of incremental elastic moduli corresponding to ϵ^0 are $\mathbf{D}_l$ for further loading and $\mathbf{D}_u$ for unloading, $\mathbf{D}_u$ being positive-definite. We imagine that the initial equilibrium state is disturbed by applying surface tractions δp_i over the surface of the ellipsoid with axes a_1, a_2, a_3, Fig. 1(*b*). We expect δp_i to produce loading inside the ellipsoid and unloading outside. We try to calculate the displacements δu_i produced by tractions δp_i at all loading points on the ellipsoid surface.

Let $\delta\epsilon^e_{ij}$, δu^e_i be the strain and displacement variations produced (by tractions δp_i) in the ellipsoid, and denote the net tractions acting on the softening ellipsoid as δp^e_i, and those on the rest of the infinite body, i.e., on the exterior of the ellipsoid, as δp^r_i. As for the distributions of δp^e_i and δp^r_i over the ellipsoid surface, we assume them to be such that $\delta p^e_i = \delta\sigma^e_{ij} n_j$ and $\delta p^r_i = \delta\sigma^r_{ij} n_j$ where $\delta\sigma^e_{ij}$ and $\delta\sigma^r_{ij}$ are arbitrary constants; $\delta\sigma^e_{ij}$ is the stress within the softening ellipsoidal region, which is uniform (and represents an equilibrium field), and $\delta\sigma^r_{ij}$ is a fictitious uniform stress in this region which would equilibrate δp^e_i. Stresses $\delta\sigma^e_{ij}$ in reality exist only on the outside of the ellipsoid surface.

Equilibrium requires that $\delta p_i = \delta p^e_i - \delta p^r_i$. The first-order work δW done by σ^0_{ij} must vanish if the initial state is an equilibrium state. The second-order work done by δp_i may be calculated as

$$\begin{aligned} \delta^2 W &= \int_S \frac{1}{2}\delta p_i \delta u^e_i\, dS \\ &= \frac{1}{2}\int_S \delta p^e_i \delta u^e_i\, dS - \frac{1}{2}\int_S \delta p^r_i \delta u^e_i\, dS \\ &= \frac{1}{2}\delta\sigma^e_{ij}\int_S n_j \delta u^e_i\, dS - \frac{1}{2}\delta\sigma^r_{ij}\int_S n_j \delta u^e_i\, dS \end{aligned} \tag{8}$$

where S = surface of the softening ellipsoidal region. Note that $\delta\sigma^e_{ij}$ are not the actual stresses in the solid but merely serve the purpose of characterizing the surface tractions δp^e_i. Applying Gauss' integral theorem and exploiting the symmetry of tensors $\delta\sigma^e_{ij}$ and $\delta\sigma^r_{ij}$, we further obtain

$$\begin{aligned} \delta^2 W &= \frac{1}{2}(\delta\sigma^e_{ij} - \delta\sigma^r_{ij})\int_V \delta u^e_{i,j}\, dV \\ &= \int_V \frac{1}{2}(\delta\sigma^e_{ij} - \delta\sigma^r_{ij})\frac{1}{2}(\delta u^e_{i,j} + \delta u^e_{j,i})\, dV \\ &= \int_V \frac{1}{2}(\delta\sigma^e - \delta\sigma^r)^T \delta\epsilon^e\, dV \end{aligned} \tag{9}$$

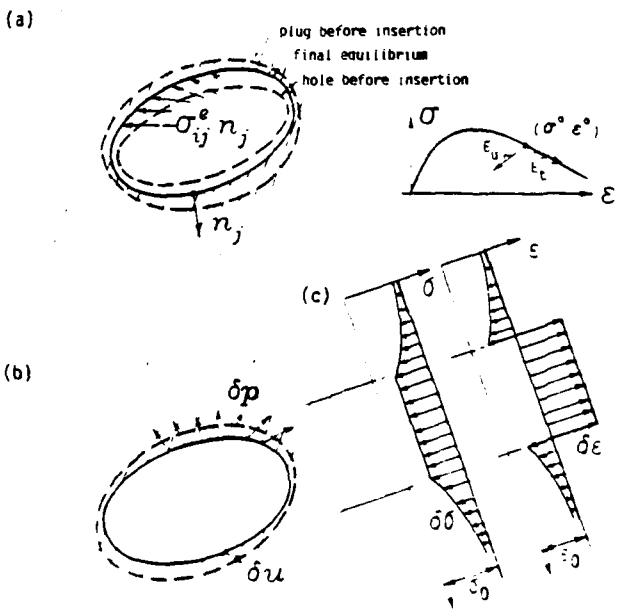


Fig. 1(a) Ellipsoidal plug (inclusion) inserted into infinite elastic solid, and (b) localization of strain into an elliptic region

where V = volume of the softening ellipsoidal region, and subscripts preceded by a comma denote partial derivatives. We changed here to matrix notation and also recognized that $\frac{1}{2}(\delta u^e_{i,j} - \delta u^e_{j,i}) = \delta\epsilon^e_{ij}$. Now we may substitute

$$\delta\sigma^t = \mathbf{D}_t \delta\epsilon^e \tag{10}$$

or $\delta\sigma^{tT} = \delta\epsilon^{eT}\mathbf{D}_t$ (assuming $\mathbf{D}_t^T = \mathbf{D}_t$). According to equation (7), we also have, as a key step

$$\delta\sigma^e = \mathbf{D}_u(1 - \mathbf{Q}_u^{-1})\delta\epsilon^e. \tag{11}$$

In contrast to our previous consideration of the elastic ellipsoidal plug made of the same elastic material, equations (2)–(7), the sole meaning of $\delta\sigma^e$ now is to characterize the tractions δp_i^t acting on the ellipsoidal surface of the infinite medium lying outside the ellipsoid. Noting that the integrand in equation (9) is constant, we thus obtain

$$\delta^2 W = \frac{1}{2}\delta\epsilon^{eT}\mathbf{Z}\delta\epsilon^e V = \frac{1}{2}Z_{ijkm}\delta\epsilon^e_{ij}\delta\epsilon^e_{km}V \tag{12}$$

in which $\mathbf{Z}$ denotes the following 6×6 matrix

$$\mathbf{Z} = \mathbf{D}_t + \mathbf{D}_u(\mathbf{Q}_u^{-1} - \mathbf{1}). \tag{13}$$

Equation (12) defines a quadratic form. If the initial uniform strain ϵ^0 is such that the associated $\mathbf{D}_t$ and $\mathbf{D}_u$ give $\delta^2 W > 0$ for all possible $\delta\epsilon^e_{ij}$, then no localization in an ellipsoidal region can begin from the initial state of uniform strain ϵ^0 spontaneously, i.e., without applying loads δp. If, however, $\delta^2 W$ is negative for some $\delta\epsilon^e_{ij}$, the localization leads to a release of energy which is first manifested as a kinetic energy and is ultimately dissipated as heat. Such a localization obviously increases entropy of the system, and so it will occur, as required by the second law of thermodynamics. Therefore, the necessary condition of stability of a uniform strain field in an infinite solid is that matrix $\mathbf{Z}$ given by equation (13) must be *positive-definite*.

The expressions for Eshelby's coefficients S_{ijkm} from which matrix $\mathbf{Q}_u$ is formed (see, e.g., Mura, 1982) depend on the ratios a_1/a_3, a_2/a_3 of the axes of the ellipsoidal localization region. They also depend on the ratios of the unloading moduli D^u_{ijkm}. If, e.g., the unloading behavior is assumed to be isotropic, they depend on the unloading Poisson ratio ν_u. Matrix $\mathbf{D}_u$ is determined by ν_u and unloading Young's modulus E_u. If, just for the sake of illustration, the loading behavior is assumed to be also isotropic, matrix $\mathbf{D}_t$ is determined by ν_t and E_t (Poisson's ratio and Young's modulus for loading). E_t, ν_t, E_u, ν_u, in turn, depend on the strain ϵ^0_{ij} at the start of localization. Since a division of $\mathbf{Z}$ by E_u does not affect positive-definiteness, only the ratio $\mathrm{E}_t/\mathrm{E}_u$ matters. Thus, $\mathbf{Z}$ is a function of the form

$$\mathbf{Z} = \mathrm{E}_u\hat{\mathbf{Z}}\left(\frac{a_1}{a_3}, \frac{a_2}{a_3}, \nu_u, \nu_t, \frac{\mathrm{E}_t}{\mathrm{E}_u}\right)$$

$$= \mathrm{E}_u\hat{\mathbf{Z}}\left(\frac{a_1}{a_3}, \frac{a_2}{a_3}, \epsilon^0_{ij}\right) \tag{14}$$

where $\hat{\mathbf{Z}}$ and $\hat{\mathbf{Z}}$ are nondimensional matrix functions.

Note that matrix $\mathbf{Z}$, which decides the localization instability, is independent of the size of the ellipsoidal localization region. This is the same conclusion as already made for a planar localization band in an infinite solid. No doubt, the size of the localization ellipsoid would matter for finite-size solids, same as it does for localization bands in layers.

The previously obtained solution for a planar localization band in an infinite solid must be a special case of the present solution for an ellipsoid. Localization in line cracks also must be obtained as a special case for $a_3 \to \infty$; however, the present solution is not realistic for this case since energy dissipation due to strain-softening is finite per unit volume and, therefore, vanishes for a crack (the volume of which is zero). For this case, it would be necessary to include the fracture energy (surface energy) in the energy criterion of stability, same as in fracture mechanics. In this study, though, we take the view that, due to material heterogeneity, it makes no sense to apply a continuum analysis to localization regions whose width is less than a certain length h proportional to the maximum size of material inhomogeneities.

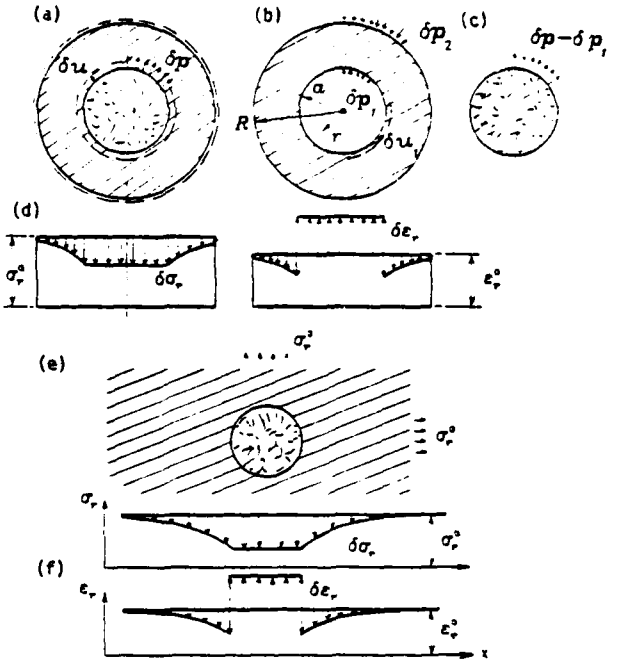


Fig. 2 Localization of strain into spherical and circular regions

Spherical Softening Region in a Sphere or Infinite Solid

Localization in a spherical region is a special case of the preceding solution for ellipsoidal regions. However, the solution may now be easily obtained even when the solid is finite. Consider a spherical hole of radius a inside a sphere of radius R, (Fig. 2(*a*)). We assume polar symmetry of the deformation field and restrict our attention to materials that are isotropic for unloading. As shown by Lamé (1852) (see, e.g., Timoshenko and Goodier, 1970, p. 395), the elastic solution for the radial displacements and the radial normal stresses at a point of radial coordinate r is

$$u = Ar + Dr^{-2}, \quad \sigma_r = \mathrm{E}_u(\bar{A} - 2\bar{D}r^{-3}) \tag{15}$$

where $\bar{A} = A/(1-2\nu_u)$, $\bar{D} = D/(1+\nu_u)$; E_u, ν_u = Young's modulus and Poisson's ratio of the sphere, and A, D = arbitrary constants to be found from the boundary conditions.

We now consider a solid sphere of radius R which is initially under uniform hydrostatic stress σ^0 and strain ϵ^0 ($\sigma^0 = \sigma^0_{kk}/3$, $\epsilon^0 = \epsilon^0_{kk}/3$), and seek the conditions for which the initial strain may localize in an unstable manner into a spherical region of radius a. Such localization may be produced by applying on the solid sphere at $r = a$ radial outward tractions δp (i.e. pressure) uniformly distributed over the spherical surface of radius a, Fig. 2(*a*). To determine the work of δp, we need to calculate the radial outward displacement δu_1 at $r = a$. We will distinguish several types of boundary conditions on the outer surface $r = R$.

(a) **Outer Surface Kept Under Constant Load.** As the boundary condition during localization, we assume that the initial radial pressure p_2^0 applied at outer surface $r = R$ is held constant, i.e., $\delta p_2 = 0$. For $\delta\sigma_r = -\delta p_1$ at $r = a$ and $\delta\sigma_r = -\delta p_2 = 0$ at $r = R$, equation (11) may be solved to yield $\bar{A} = a^3\delta p_1/\mathrm{E}(R^3 - a^3)$, $\bar{D} = \bar{A}R^3/2$, and from equation (15)

$$\delta u_1 = \frac{\delta p_1}{C_u},\quad \frac{1}{C_u} = \frac{a}{E_u(R^3 - a^3)}\left[(1-2\nu_u)a^3 + \frac{1+\nu_u}{2}R^3\right]. \tag{16}$$

The inner spherical softening region of radius a, Fig. 2(*c*), is assumed to remain in a state of uniform hydrostatic stress and strain, and its strain-softening properties to be isotropic, characterized by E_t and ν_t. Thus the strains for $r<a$ are $\delta\epsilon = \delta u_1/a$, the stresses are $\delta\sigma = 3K_t\,\delta u_1/a = C_t\delta u_1$ with $C_t = 3K_t/a$; K_t = bulk modulus for further loading (softening); $3K_t = E_t/(1-2\nu_t)$, with $E_t<0$, $K_t<0$, for softening. The surface tractions acting on the softening region, Fig. 2(*c*), are equal to $C_t\delta u_1$ where $C_t = 3K_t/a$. Hence, by equilibrium, the total distributed traction at surface $r=a$ must be $\delta_p = C_u\delta u_1 + C_t\delta u_1$ and the work done by δp is $\Delta W = 4\pi a^2(\frac{1}{2}\delta p\delta u_1) = 2\pi a^2(C_u + C_t)\delta u_1^2$. Thus the necessary condition of stability of the initial uniform strain ϵ^0 in the solid sphere is $C_u + C_t > 0$, which yields the necessary condition for stability

$$-\frac{E_t}{E_u} < \frac{2(1-2\nu_t)(R^3 - a^3)}{2(1-2\nu_u)a^3 + (1+\nu_u)R^3}. \tag{17}$$

Changing $<$ to $>$, we obtain the sufficient condition for strain-localization instability.

(*b*) **Outer Surface Kept Fixed.** In this case we assume that during localization $\delta u = 0$ at $r=R$, and $\delta\sigma_r = -\delta p_1$ at $r=a$. From equation (15), we may then solve $D = -AR^3$ and

$$A = -\frac{\delta p_1}{E_u}\left(\frac{1}{1-2\nu_u} + \frac{2R^3}{(1+\nu_u)a^3}\right)^{-1} \tag{18}$$

which yields

$$\delta p_1 = C_u\delta u_1,\quad C_u = \frac{E_u a^2}{R^3 - a^3}\left(\frac{1}{1-2\nu_u} + \frac{2R^3}{(1+\nu_u)a^3}\right). \tag{19}$$

The work done on the solid sphere by tractions $\delta p = C_u\delta u_1 + C_t\delta u_1$ applied at surface $r=a$ is $\Delta W = 2\pi a^2\,(C_u + C_t)\,\delta u_1^2$ where $C_t = E_t/(1-2\nu_t)a$ as before. Thus the necessary stability condition is $C_u + C_t > 0$, which can now be reduced to the condition

$$-\frac{E_t}{E_u} < \frac{1-2\nu_t}{R^3 - a^3}\left(\frac{a^3}{1-2\nu_u} + \frac{2R^3}{1+\nu_u}\right). \tag{20}$$

Assuming that $|E_t|$ increases continuously after the peak of the stress-strain diagram as ϵ_0 is increased, instability develops at the value $a = a_{cr}$ which minimizes $|E_t|$ under the restriction $h/2 \le a \le R$ where h is the given minimum admissible size of the strain-softening region, representing a material property. For the case of prescribed pressure at the boundary $r=R$, we find from equation (17) that $a_{cr} = R$, which corresponds to $E_t = 0$. So the sphere becomes unstable right at the start of strain-softening, i.e., no strain-softening can be observed when the boundary is not fixed. For the case of a fixed (restrained) boundary at $r=R$, equation (20), one can verify that Min $|E_t|$ is finite and occurs at a_{cr} = Min $a = h/2$ (provided that $\nu_u \ge 0$).

For $R/a \to \infty$, equation (20) yields the stability condition for the case of infinite solid fixed at infinity, Fig. 2(*e*, *f*)

$$-\frac{E_t}{E_u} < \frac{2(1-2\nu_t)}{1+\nu_u} \tag{21}$$

It is interesting that for $R/a \to \infty$ equation (17) yields the same condition, but this limit case is of questionable significance since we found that $a_{cr} = R$ when the boundary is not fixed. Note that the stability condition in equation (21) does not depend on the radius a of the softening region; yet, unlike the softening in a layer in infinite solid, solved before, instability does not begin at the peak of stress-strain diagram (where $E_t = 0$) but begins only at a certain finite negative slope of the stress-strain diagram. This slope can in fact be rather steep ($-E_t = 2E_u$ for $\nu_t = \nu_u = 0$).

Circular Softening Region in a Planar Disk or Infinite Plate

Working in two dimensions, consider a circular hole of radius a inside a homogeneous isotropic circular disk of radius R Fig. 2(*a*,*b*). We assume a plane-stress state and, then, according to Lamé's solution, the radial displacement u and the radial normal stresses are (see, e.g., Flügge, 1962, p. 37-13; or Timoshenko-Goodier, 1970, p. 70)

$$u = \frac{1-\nu_u}{E_u}\,\frac{a^2p_1 - R^2p_2}{R^2 - a^2}r + \frac{1+\nu_u}{E_u}\,\frac{R^2a^2(p_1 - p_2)}{(R^2 - a^2)r} \tag{22}$$

$$\sigma_r = \frac{R^2a^2(p_2 - p_1)}{(R^2 - a^2)r^2} + \frac{a^2p_1 - R^2p_2}{R^2 - a^2},$$

$$\epsilon_z = \frac{2\nu_u(p_2R^2 - p_1a^2)}{E_u(R^2 - a^2)} \tag{23}$$

where p_1 and p_2 are the pressures applied along the hole perimeter and along the outer perimeter of the disk, respectively, and ϵ_z is the transverse strain in the plate, which is independent of r. Depending on the boundary conditions, we distinguish three cases:

(*a*) **Outer Boundary Kept Under Constant Load.** As the boundary condition during the strain localization instability, we now assume that the initial radial pressure p_2 applied at the outer boundary $r=R$ of the disk is held constant, i.e., $\delta p_2 = 0$. Equation (22) then yields for $\delta u_1 = \delta u$ at $r = a$, Fig. 2(*b*), the relation

$$\delta u_1 = \frac{\delta p_1}{C_u},\quad \frac{1}{C_u} = \frac{a}{E_u}\left(\nu_u + \frac{R^2 + a^2}{R^2 - a^2}\right). \tag{24}$$

The inner circular softening region of radius a, Fig. 2(*c*), is assumed to remain in a uniform state of stress and strain. Thus the strains for $r<a$ are $\delta\epsilon_r = \delta u_1/a$. Assuming the plate to be thin compared to radius a, we may assume the strain-softening region to be also in a plane-stress state, and then $\delta\epsilon_r = \delta\sigma_r(1-\nu_t)/E_t$. Hence,

$$\delta\sigma_r = C_t\delta u_1,\quad C_t = \frac{1}{a}\,\frac{E_t}{1-\nu_t}. \tag{25}$$

Now, we consider uniformly distributed outward tractions δp to be applied along the circle $r=a$ on the solid disk (without the hole). By equilibrium, $\delta p = C_u\delta u_1 - C_t\delta u_1$ and the work done by δp is $\Delta W = 2\pi a(\frac{1}{2}\,\delta p\delta u_1) = \pi a(C_u - C_t)\delta u_1^2$. Thus the necessary condition of stability of the initial state of uniform strain is $C_u - C_t > 0$. According to equations (24) and (25), the stability condition for plane stress becomes

$$-\frac{E_t}{E_u} < \frac{1-\nu_t}{\nu_u + \dfrac{R^2 + a^2}{R^2 - a^2}}\quad \text{(thin plate)}. \tag{26}$$

As another limiting case, we may consider a long cylinder of radius R (and length $>> R$), in which case the softening region is forced to have along the cylinder axis the same strain ϵ_z as the unloading region. However, the softening region is not in a plane-strain state either. Assuming that the planes normal to the cylinder axis remain plane, consider now that, unlike before, the axial stresses σ_z are nonzero. We must impose the equilibrium condition that the resultants of σ_z' in the softening region and of σ_z'' in the unloading region cancel each other, i.e., $\pi(R^2 - a^2)\sigma_z' = -\pi a^2\sigma_z''$. We leave it to a possible user to work out the solution in detail and we now restrict our attention to the case $a << R$, for which, according to equation (20) ($\delta p_2 = 0$), we have $\delta\epsilon_z = 2\nu_u\delta p_1 a^2/(a^2 - R^2)E_u \approx 0$ (for $r \ge a$), while also $\delta\sigma_z = 0$. Therefore, we may assume for the incremental deformation in the strain-softening region a state of plane strain. The solution may then be obtained simply by

replacing E_t with $E_t' = E_t/(1-\nu_t^2)$, and ν_t with $\nu_t' = \nu_t/(1-\nu_t)$ (the unloading region remains in plane stress in this case). Equation (26) thus transforms to

$$-\frac{E_t}{E_u} < \frac{(1+\nu_t)(1-2\nu_t)}{\nu_u + \dfrac{R^2+a^2}{R^2-a^2}} \quad \text{(long cylinder, } a<<R\text{).} \tag{27}$$

When a/R is not very small, the solution may be expected to lie between equations (26) and (27).

(*b*) **Outer Boundary Kept Fixed.** In this case, we have during localization $\delta u = 0$ at $r = R$ and $\delta\sigma_r = -\delta p_1$ at $r = a$. Taking the variations of equations (22) at $r = R$ and $r = a$, we get

$$\frac{1-\nu_u}{E_u}\frac{a^2\delta p_1 - R^2\delta p_2}{R^2-a^2}R + \frac{1+\nu_u}{E_u}\frac{R^2a^2(\delta p_1-\delta p_2)}{(R^2-a^2)R} = 0 \tag{28}$$

$$\delta u_1 = \frac{1-\nu_u}{E_u}\frac{a^2\delta p_1 - R^2\delta p_2}{R^2-a^2}a + \frac{1+\nu_u}{E_u}\frac{R^2a^2(\delta p_1-\delta p_2)}{(R^2-a^2)a}. \tag{29}$$

Eliminating δp_2 from these two equations, we get the relation $\delta p_1 = C_u \delta u_1$ with

$$\frac{1}{C_u} = \frac{1}{E_u(R^2-a^2)}\left\{a[(1-\nu_u)a^2 + (1+\nu_u)R^2] - \frac{4R^2a^3}{(1-\nu_u)R^2+(1+\nu_u)a^2}\right\}. \tag{30}$$

By the same reasoning as before, the necessary condition for the stability of the initial uniform strain ϵ^0 is $C_u + C_t > 0$ where C_t is again given by equation (25). This condition yields

$$-\frac{E_t}{E_u} < \frac{(1-\nu_t)(R^2-a^2)}{R^2+a^2+\nu_u(R^2-a^2) - \dfrac{4R^2a^2}{(1-\nu_u)R^2+(1+\nu_u)a^2}} \quad \text{(thin plate).} \tag{31}$$

For the case of a long cylinder of length $>> R$ and with $a << R$, we may obtain the solution again by replacing E_t, ν_t with E_t', ν_t'. This yields

$$-\frac{E_t}{E_u} < \frac{(1+\nu_t)(1-2\nu_t)(R^2-a^2)}{R^2+a^2+\nu_u(R^2-a^2) - \dfrac{4R^2a^2}{(1-\nu_u)R^2+(1+\nu_u)a^2}} \quad \text{(long cylinder, } a<<R\text{).} \tag{32}$$

Instability develops at the value $a = a_{cr}$ which minimizes $|E_t|$ under the restriction that $h/2 \le a \le R$. For the case of prescribed load at outer boundary, we find from equation (26) or (27) that $a_{cr} = R$, which corresponds to $E_t = 0$. Thus the disk becomes unstable right at the start of strain-softening, i.e., no strain-softening can be observed. For the case of a fixed (restrained) boundary at $r = R$, we find that, for $a \to R$, $\lim(-E_t/E_u) = \infty$ (to verify it one needs to substitute $a = R-\delta$ and consider $\delta \to 0$); consequently Min $|E_t|$ is finite, and it is found to occur at $a_{cr} = \text{Min } a = h/2$.

For $R/a \to \infty$, equation (31) or (32) yields the stability condition for the case of infinite plate fixed at infinity, Fig. 2(*e*, *f*)

$$-\frac{E_t}{E_u} < \frac{1-\nu_t}{1+\nu_u} \quad \text{for thin plate} \tag{33}$$

$$-\frac{E_t}{E_u} < \frac{(1+\nu_t)(1-2\nu_t)}{1+\nu_u} \quad \text{for massive solid.} \tag{34}$$

It is interesting that for $R/a \to \infty$ equations (26) and (27) yield the same conditions, but these limits are of questionable significance since we found that $a_{cr} = R$ when the boundary is not fixed. Note that the stability conditions in equations (33) and (34) are independent of the size of the localization region,

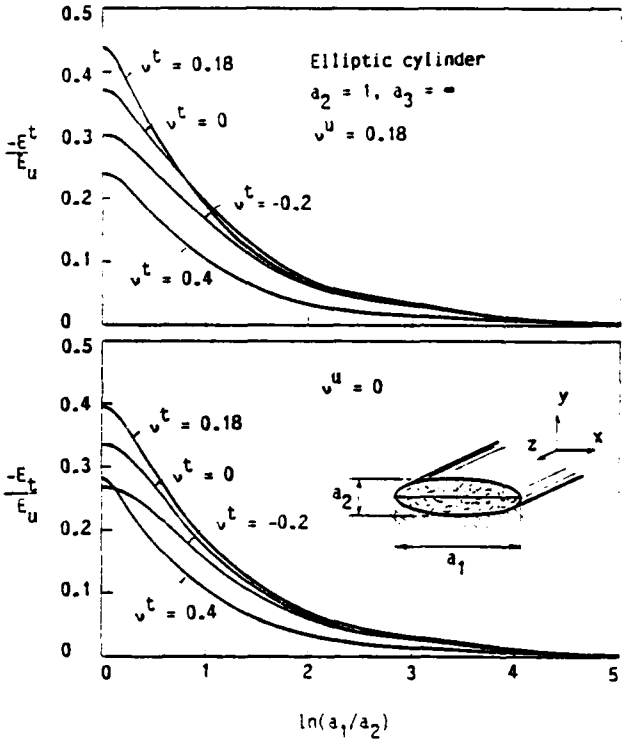


Fig. 3(*a*, *b*)

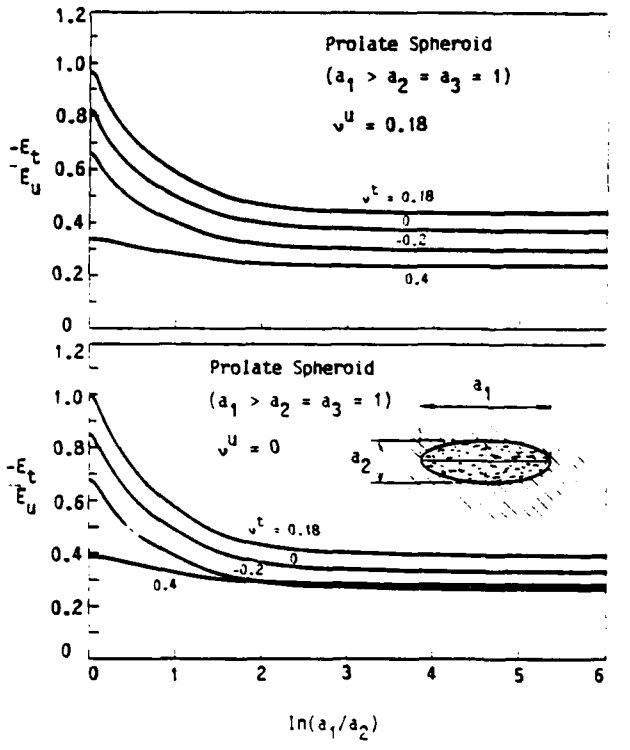


Fig. 3(*c*, *d*)

Fig. 3 Tangential modulus E_t at the limit of stability against localization into ellipsoidal region as a function of ratio a_1/a_2 of principal axes of ellipsoid

same as we found it for spherical localization regions and layers.

Softening Annulus or Shell

Solution on the basis of equations (22) and (23) or (25) is also possible for curved softening bands limited by two circles (an annulus) or by two spherical surfaces (a shell). In this case, the softening layer is not in a homogeneous state of strain and stress. The resulting formulas are more complicated. They represent a transition between the solution for a softening circle (or cylinder, sphere) and a softening band.

Numerical Examples and Discussion of Results

Figure 3 shows some numerical results for localization of strain into ellipsoidal domains in infinite space. The results were calculated for domains in the shape of infinitely long elliptic cylinder ($a_3 \to \infty$ and various a_1/a_2) as well as prolate spheroid ($a_1 > a_2 = a_3$, various ratios a_1/a_2). The material was assumed to be incrementally isotropic, with matrices $\mathbf{D}_t$ and $\mathbf{D}_u$ characterized by Young's moduli E_t and E_u, and Poisson ratios $\nu_u = 0.18$ with various values of ν_t. The assumption of incremental isotropy is here made for the sake of simplicity. In reality, the incremental moduli at strain-softening must be expected to be anisotropic, except when the initial state is a purely volumetric strain. A subsequent paper (Bažant and Lin, 1987) gives numerical results for incrementally anisotropic moduli corresponding to von Mises plasticity and nonassociated Drucker-Prager plasticity with a negative plastic modulus.

Matrix $\mathbf{Z}$, equation (13), was evaluated by computer on the basis of S_{ijkm} taken from Mura (1982, equations (11.22) and (11.29)). The smallest eigenvalue of matrix $\mathbf{Z}$ was calculated by a computer library subroutine. Iterative search by Newton method was made to find the value of E_t/E_u for which the smallest eigenvalue is zero and is about to become negative, which indicates loss of positive-definiteness.

The results are plotted in Figs. 3(*a–d*). For the infinite cylinder, Fig. 3(*a,b*), as a_1/a_2 increases, the localization instability occurs at smaller $|E_t/E_u|$. The case $a_1/a_2 \to \infty$ corresponds to an infinite planar band, and the results are identical to those obtained before for this case. In particular, $|E_t|$ tends to 0 as $a_1/a_2 \to \infty$; i.e., instability occurs right at the peak of the stress-strain diagram.

For the prolate spheroid, Fig. 3(*c, d*), the instability also occurs at decreasing $|E_t/E_u|$ as a_1/a_2 increases, but for $|E_t| \to \infty$, which corresponds to an infinite circular tube, a finite value of $|E_t|$, depending on Poisson's ratio, is still required for instability. This limiting case is equivalent to two-dimensional localization in a circular region, equation (34). On the other hand, the case $a_1/a_2 = 1$, Fig. 3(*c,d*), is equivalent to localization into a spherical region, equation (21).

Figures 4(*a–d*) shows the plots for various incremental Poisson's ratios of $|E_t/E_u|$ at incipient localization instability as a function of R/a for the following cases: (1) sphere, displacement fixed, equation (20); (2) localization in planar band in uniaxial extension (Equation (12) of Bažant, 1987, $R/a = L/h$); (3) localization in planar shear band (Bažant, 1987, equation (13), $R/a = L/h$).

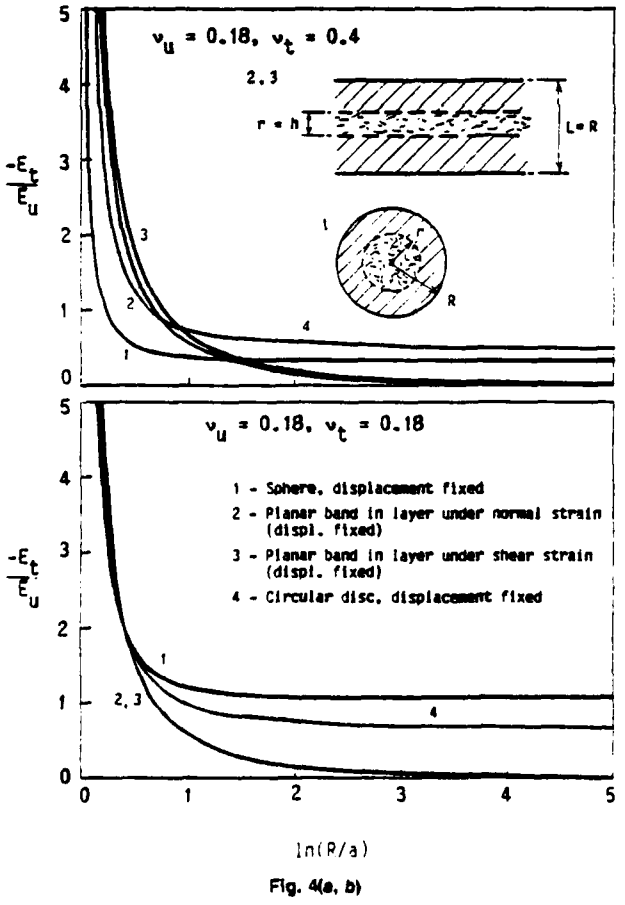


Fig. 4(*a, b*)

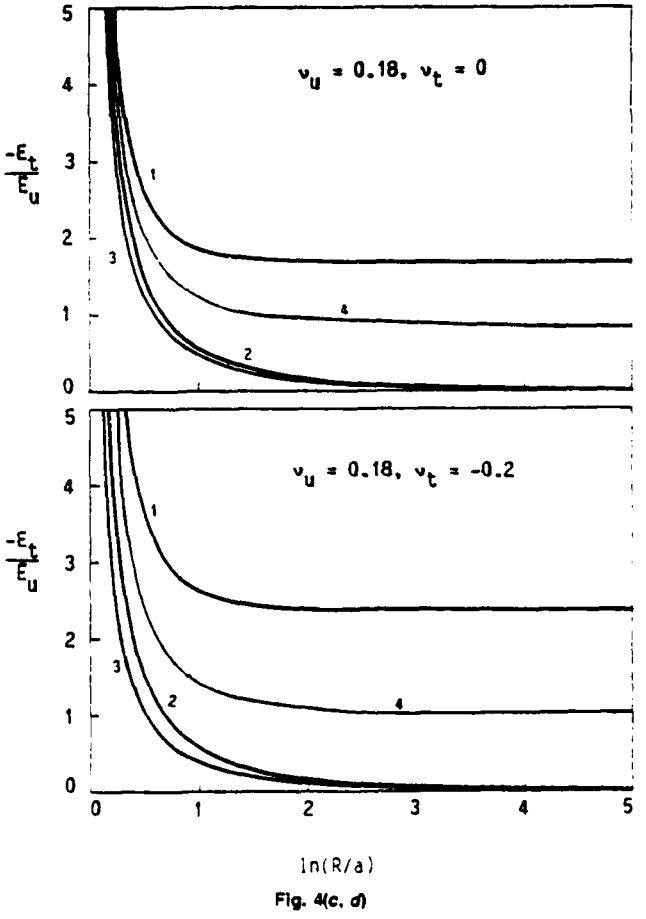


Fig. 4(*c, d*)

Fig. 4 Tangential modulus E_t at the limit of stability against localization into spherical and circular regions of radius r within a sphere or disk of radius R

We see from Fig. 4 that the value of E_t at instability depends strongly on the relative size $R(a)$ of the body as well as the Poisson's ratios. For infinite body size ($R/a \to \infty$), instability of planar bands occurs at $E_t = 0$, i.e., at the peak of stress-strain diagram. The same happens under load-controlled conditions for the cases of spherical or circular regions with $R/a = 1$ (i.e., the smallest possible body size for which the body remains at homogeneous strain). For $R/a > 1$, the spherical or circular regions generally require a finite slope $|E_t|$ to produce localization instability, provided the boundary is under prescribed displacement during the localization.

The results show that a localization instability in the form of a planar band always develops at a smaller $|E_t|$, and thus at a smaller initial strain, than the localization instability in the form of an ellipsoidal softening region. That does not mean, however, that the planar band would always occur in practice. A planar localization band cannot accommodate the boundary conditions of a finite solid restrained on its boundary, and a localization region similar to an ellipsoid may then be expected to form. It is remarkable how slowly the slope $|E_t|$ at instability decreases as a function of a_1/a_2. The value of the aspect ratio that is required to reduce $|E_t|$ at instability from about 0.4 to about 0.04 of E_u is $a_1/a_2 = e^3 = 20$. This means that if a very long planar softening band cannot be accommodated within a given solid, the deformation at softening instability is considerably increased.

The present solutions represent upper bounds on $|E_t|$ at actual localization. When the stability condition for some of the previously considered softening regions is violated, instability with such a region is possible and must occur since it leads to an increase of entropy. However, it is possible that localization into some form of region that we could not solve would occur earlier, at a smaller initial strain. For this reason, the present stability conditions are only necessary rather than sufficient. However, their opposites (i.e., $<$ changed to $>$) represent sufficient conditions for instability.

In the preceding analysis of ellipsoidal softening regions, we solved only the case of infinite solids and were unable to examine the effect of the boundary conditions at infinity. Now, from the fact that spherical and circular softening regions are special cases of ellipsoidal ones, we must conclude that our solution for ellipsoidal region is applicable only if the ellipsoidal region is of finite size, which is guaranteed only if the infinite body is fixed at infinity rather than having prescribed loads at infinity. Otherwise the limit cases $a_{cr} = R$ for spherical or circular softening regions discussed after (equations (20) and (32)) would not be satisfied.

Conclusion

The solutions to multidirectional localization problems with ellipsoidal, spherical and circular localization regions indicate that, in general, a loss of positive-definiteness of matrix $\mathbf{D}$ of tangential moduli for loading does not necessarily produce instability. Rather, the stability criterion requires positive-definiteness of a certain weighted average of the incremental moduli matrices $\mathbf{D}_t$ and $\mathbf{D}_u$ for loading and unloading. The weights depend on the relative size of the body. Not only for finite but also for infinite bodies restrained at the boundary, unstable strain localization into finite-size ellipsoidal regions cannot take place for a certain range of nonpositive-definite tangential moduli matrices. By contrast, unstable strain localization into an infinitely long planar band of finite thickness occurs in an infinite space as soon as $\mathbf{D}_t$ loses positive-definiteness.

The present results can be used to check the correctness of finite-element programs for strain-softening.

Acknowledgment

Partial financial support under AFOSR Contract F49620-87-C-0030DEF with Northwestern University, monitored by Dr. Spencer T. Wu, is gratefully acknowledged. Graduate research assistant F.-B. Lin deserves thanks for his valuable assistance in numerical calculations.

References

Bažant, Z. P., 1987, "Softening Instability: Part I: Localization Into a Planar Band," Report No. 87-2/498s, Center for Concrete and Geomaterials, Northwestern University, Evanston, IL.

Bažant, Z. P., and Lin, F.-B., 1987, "Localization of Softening in Ellipsoids and Bands: Parameter Study," Report No. 87-7/4981s, Center for Concrete and Geomaterials, Northwestern University, July, ASCE *Journal of Engineering Mechanics*, in press.

Christensen, R. M., 1979, *Mechanics of Composite Materials*, John Wiley, New York.

Eshelby, J. D., 1957, "The Determination of the Elastic Field of an Ellipsoidal Inclusion and Related Problems," *Proceedings of the Royal Society of London*, Vol. A241, pp. 376.

Flügge, W., Ed., 1962, *Handbook of Engineering Mechanics*, McGraw-Hill, New York.

Kinoshita, N., and Mura, T., 1971, "Elastic Fields of Inclusions in Anisotropic Media," *Phys. Stat. Sol.* (*a*), Vol. 5, pp. 759–768.

Lamé, G., 1852, *Leçons sur la theorie de l'élasticite*, Gauthier-Villars, Paris.

Lin, S. C., and Mura, T., 1973, "Elastic Fields of Inclusions in Anisotropic Media (II)," *Phys. Stat. Sol.* (*a*), Vol. 15, pp. 281–285.

Mura, T., 1982, *Micromechanics of Defects in Solids*, Martinus Nijhoff Publishers, The Hague.

Timoshenko, S. P., and Goodier, J. N., 1970, *Theory of Elasticity*, 3rd Ed., McGraw-Hill, New York.

Proof 1143 Pergamon Press SS MS. 626

Int. J. Solids Structures Vol. 00, No. 00, pp. 0000–0000, 1989
Printed in Great Britain.

0020-7683/89 $3.00 + .00
© 1989 Maxwell Pergamon Macmillan plc

STABILITY AGAINST LOCALIZATION OF SOFTENING INTO ELLIPSOIDS AND BANDS: PARAMETER STUDY

ZDENĚK P. BAŽANT
Department of Civil Engineering, Northwestern University, Evanston, IL 60208, U.S.A.

and

FENG-BAO LIN
Department of Civil and Environmental Engineering, Polytechnic University, Brooklyn, NY 11201, U.S.A.

(*Received* 10 *October* 1987; *in revised form* 29 *November* 1988)

Abstract—Strain-localization instabilities due to strain-softening which result from distributed damage such as cracking in heterogeneous brittle materials are analysed. Attention is restricted to the stability problem of equilibrium states. This problem is not equivalent to bifurcation of the equilibrium path, which may occur before stability of equilibrium is lost. The continuum is local but is enhanced by the localization limiter used in the crack band model, consisting of a lower bound on the minimum dimension of the strain-localization region, which is regarded as a material property. Presented are derivations of the critical state conditions for localization of initially uniform strain into ellipsoidal domains within an infinite continuum and into a planar band within a layer of finite thickness. These derivations are simpler than the previous Bažant's derivations of the general stability conditions for these localizations. A numerical parameter study of the critical states is made for a broad range of material properties as well as various initial stress states and relative sizes of the strain-softening region. The material is described by Drucker–Prager plasticity with strain-softening that is caused by yield limit degradation. The flatter the ellipsoidal domain, or the larger the size of the body (layer thickness), the smaller is found to be the strain-softening slope magnitude at which the critical state is reached. A softening Drucker–Prager material is found to be stable against planar-band localizations in infinite continuum for a certain range of softening material parameters.

INTRODUCTION

Distributed damage such as cracking or void growth can be macroscopically described as strain-softening, a behavior in which the stress declines at increasing strain, or more precisely the matrix of incremental elastic moduli ceases to be positive definite. Strain-softening causes the strain as well as the energy dissipation to localize. This localization represents a stability problem.

For one-dimensional localization that describes the tensile or compressive failure of a uniaxially stressed bar as well as the development of a planar localization band (e.g. crack band) in infinite space, the stability conditions in terms of the strain-softening properties of the material were derived in a 1974 report and a follow-up paper by Bažant (1976). From this stability analysis it transpired that, in the usual local continuum, the strain as well as the energy dissipation localize into a region of zero volume. This implies the structure to fail with a zero energy dissipation and the failure to occur right after the first onset of localization. This would indicate every strain-softening state to be unstable and therefore unobservable. In reality, structures of course fail with a finite energy dissipation and strain-softening states in which the energy dissipation is not localized to a zero volume do exist, as evidenced e.g. by measurements of the locations of sound emission sources.

The simplest remedy that eliminates this physically unrealistic situation is to impose a lower bound on the minimum cross-section dimension of the strain-softening region. This *ad hoc* measure was introduced by Bažant (1976) in the finite element crack band model, and was subsequently developed first for sudden softening (Bažant and Cedolin, 1979, 1980) and later for gradual softening (Bažant, 1982; Bažant and Oh, 1983). The last version has been shown to be in good agreement with all the basic concrete fracture data, and to exhibit the correct size effect, which is transitional between the size effect of plastic limit analysis and the size effect of linear elastic fracture mechanics. The crack band model has

further been refined and implemented in some large finite element codes (de Borst, 1984; de Borst and Nauta, 1984, 1985; Darwin, 1985; Bažant, 1986). It will therefore be assumed in this study that strain-softening cannot localize into a region whose minimum cross-section dimension is less than a certain characteristic length h, which is considered to be a material property.

The strain localization instability was explained and analyzed in 1974 by Bažant (1976) for one-dimensional strain-softening in bars of infinite or finite length with rigid or elastic supports, as well as for flexural softening in beams; see also Bažant's (1986) review. Rudnicki and Rice (1975) formulated the condition of localization into a planar band in an infinite space (see also Rice, 1976). This condition turned out to be identical to the conditions of uniqueness and of shear band formation, previously obtained by Hill (1962) (see also Mandel, 1966 and Mróz, 1966). The study of Rudnicki and Rice (1975) was focused primarily on localization instabilities caused by geometrically nonlinear effects of strain, which they showed to be possible already before the peak of the stress–strain diagram (i.e. in the plastic-hardening range). However, some critical states were also obtained for negative values of the plastic-hardening modulus, i.e. in the softening range. These studies, which were restricted to nonassociated Drucker–Prager plasticity, in some cases enhanced with a vertex hardening term, did not consider bodies of finite dimensions for which the size of the localization region usually has paramount influence on the critical state, and did not generally treat unloading outside the localization band, which is essential for the loss of stability in finite bodies. Subsequently, Rudnicki (1977) studied localization into ellipsoidal regions in a uniformly stressed infinite space and showed examples of such instabilities in the hardening as well as softening regime. This study was also limited to nonassociated Drucker–Prager plasticity (without or with a vertex hardening term), and dealt only with the critical state of neutral equilibrium while stability was not analyzed.

Recently Bažant (1988a, b) formulated in a closed form the conditions of critical state as well as stability for the localization of strain-softening into planar bands or ellipsoidal regions. This study included formulation of the stability conditions for strain-softening localization into a planar band that forms within a layer of finite thickness. These stability conditions allowed for completely general material properties characterized by arbitrary tensors of incremental moduli for loading and unloading (with general anisotropy), which makes it possible to determine the important effect of the type of constitutive law on these instabilities. The numerical examples, however, dealt only with the special case of isotropic incremental elastic moduli tensors, which is simple to treat but of course not too realistic.

Our objective will be to apply these previously derived stability conditions assuming more realistic, incrementally anisotropic material properties based on Drucker–Prager plasticity (Fung, 1965; Owen and Hinton, 1980; Chen, 1982). At the same time, we will present a direct derivation of the conditions of critical state (stability limit) which is considerably simpler than the previous derivation of the stability conditions. Compared to the work of Rudnicki and Rice (1975) and Rice (1976), the present study as well as the preceding ones by Bažant (1988a, b) extends the localization condition to arbitrary constitutive laws and to bodies of finite dimensions for which unloading is important and the body size matters.

It needs to be emphasized that, as in Bažant's (1988a, b) preceding study, the present study deals only with the critical state of the loss of stability of equilibrium, and not with bifurcation of the equilibrium path as the load is increased. The problems of stability loss and bifurcation are not equivalent. For uniaxially stressed bars or planar localization bands, the critical state for stability differs from the state of bifurcation (Bažant, 1988c). More detailed comments on this subject will be made later in this paper.

To avoid localization of energy dissipation into a region of zero volume, one generally needs to introduce the so-called localization limiters (Bažant and Belytschko, 1987). The imposition of a lower bound on the size of this region, adopted in this study, is the simplest but crudest localization limiter. Limitation of localization in general calls for adopting the nonlocal continuum approach (Bažant *et al.*, 1984; Bažant, 1987). The latest form of this approach, which is easily implemented in large finite element codes, is the nonlocal continuum with local strain (Bažant and Pijaudier-Cabot, 1987, 1988; Pijaudier-Cabot and Bažant, 1987; Bažant and Lin, 1988). The nonlocal approach, however, does not seem

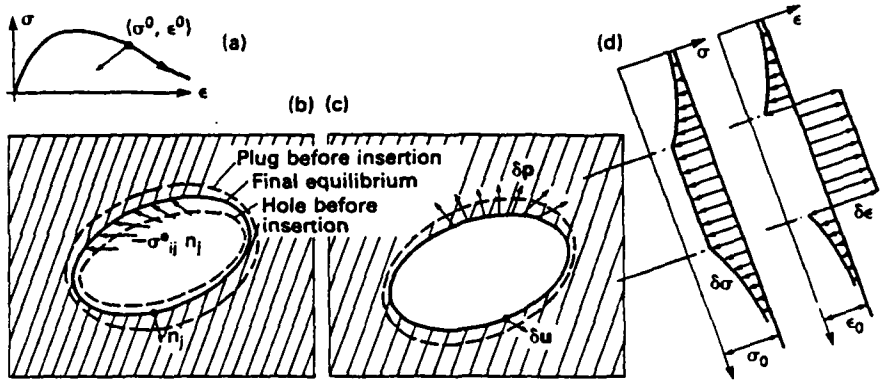


Fig. 1. Ellipsoidal plug (inclusion) inserted into infinite elastic solid, and localization of strain into an elliptic region.

amenable to closed-form expressions for the stability conditions and therefore will not be pursued here.

Compared to the general nonlocal formulation, the use of a local continuum with the simple imposition of a lower limit on the size of the localization zone, made in the crack band model and adopted here, might seem too simplistic. The numerical results obtained with this model have nevertheless been found to be rather close to the nonlocal solutions (Bažant and Pijaudier-Cabot, 1988; Pijaudier-Cabot and Bažant, 1987). This is a supporting argument for the present approach.

CRITICAL STATE CONDITIONS FOR LOCALIZATION INTO ELLIPSOIDS

First we consider an ellipsoidal hole (Fig. 1b) in an infinite homogeneous elastic continuum characterized by the elastic moduli matrix $\mathbf{D}_u$. We imagine to fit and glue into this hole an ellipsoidal plug made of the same material. To fit, the plug must first be deformed by a uniform strain, ε^0, called the eigenstrain (note that an ellipsoid can be transformed to any other ellipsoid by uniform strain). The eigenstrain is then imagined unfrozen, which causes the plug with ellipsoidal hole to undergo strain increment ε^e in order to establish equilibrium with the surrounding infinite continuum. According to the celebrated Eshelby's theorem (Eshelby, 1957; Christensen, 1979; Mura, 1982), the strain ε^e in the plug is uniform and is expressed as

$$\varepsilon^e_{ij} = S_{ijkm}\varepsilon^0_{km} \tag{1}$$

in which S_{ijkm} are the components of a fourth-rank tensor which depend only on the aspect ratios a_1/a_2, a_3/a_2 of the principal axes of the ellipsoid, as well as on the elastic moduli. Eshelby's coefficients S_{ijkm} are generally calculated as elliptic integrals (Mura, 1982). Always, $S_{ijkm} = S_{jikm} = S_{ijmk}$, but in general $S_{ijkm} \neq S_{kmij}$. For generally anisotropic material properties, the expressions for coefficients S_{kmij} were derived by Kinoshita and Mura (1971) and Lina and Mura (1973).

For convenience we rewrite eqn (1) in a matrix form:

$$\begin{Bmatrix} \varepsilon^e_{11} \\ \varepsilon^e_{22} \\ \varepsilon^e_{33} \\ 2\varepsilon^e_{12} \\ 2\varepsilon^e_{23} \\ 2\varepsilon^e_{31} \end{Bmatrix} = \left[\begin{array}{ccc:ccc} S_{1111} & S_{1122} & S_{1133} & S_{1112} & S_{1123} & S_{1131} \\ S_{2211} & S_{2222} & S_{2233} & S_{2212} & S_{2223} & S_{2231} \\ S_{3311} & S_{3322} & S_{3333} & S_{3312} & S_{3323} & S_{3331} \\ \hdashline 2S_{1211} & 2S_{1222} & 2S_{1233} & 2S_{1212} & 2S_{1223} & 2S_{1231} \\ 2S_{2311} & 2S_{2322} & 2S_{2333} & 2S_{2312} & 2S_{2323} & 2S_{2331} \\ 2S_{3111} & 2S_{3122} & 2S_{3133} & 2S_{3112} & 2S_{3123} & 2S_{3131} \end{array}\right] \begin{Bmatrix} \varepsilon^0_{11} \\ \varepsilon^0_{22} \\ \varepsilon^0_{33} \\ 2\varepsilon^0_{12} \\ 2\varepsilon^0_{23} \\ 2\varepsilon^0_{31} \end{Bmatrix} \tag{2}$$

or

$$\varepsilon^e = \mathbf{Q}_u \varepsilon^0,$$

in which ε^0 and ε^e are the column matrices of the eigenstrains and the equilibrium strains in the ellipsoidal region. These column matrices (with the factors 2) are defined in such a manner that $\sigma^T \varepsilon$ is the correct expression for work density; $\sigma = (\sigma_{11}, \sigma_{22}, \sigma_{33}, \sigma_{12}, \sigma_{23}, \sigma_{31})^T$ = column matrix of stresses (superscript T denotes the transpose of a matrix). For isotropic materials, the only nonzero elements of matrix $\mathbf{Q}$ of Eshelby's coefficients are those between the dashed lines marked in eqn (2).

According to Hooke's law, the stress in the ellipsoidal plug, σ^e, which is uniform, may be expressed as $\sigma^e = \mathbf{D}_u(\varepsilon^e - \varepsilon^0)$. Upon substituting $\varepsilon^0 = \mathbf{Q}_u^{-1}\varepsilon^e$ [according to eqn (2)], we obtain

$$\sigma^e = \mathbf{D}_u(\mathbf{1} - \mathbf{Q}_u^{-1})\varepsilon^e \tag{3}$$

in which $\mathbf{1}$ = unit 6×6 matrix.

Consider now that an infinite continuum (without any hole) is in an initial equilibrium state of uniform initial strain ε and uniform initial stress σ balanced by loads applied at infinity. We seek the condition under which the initial state loses stability in a mode in which the strain localizes into an ellipsoidal region (Fig. 1) without changing the prescribed stresses (or the prescribed displacements) at infinity. If these variations can happen while maintaining equilibrium, we have a state of neutral equilibrium which represents the limit of stability, i.e. the critical state.

Due to localization, the strain and stress in the infinite continuum outside the ellipsoidal region becomes nonuniform, while according to Eshelby's theorem the stress and strain inside the ellipsoidal region will remain uniform. The strains can become discontinuous across the ellipsoidal surface, as shown in Fig. 1d. On the other hand, the normal and shear stress components acting on this surface must remain continuous in order to maintain equilibrium (while the normal stresses parallel to the ellipsoidal surface need not be continuous). If the ellipsoidal region undergoes strain-softening, the outside undergoes unloading and therefore behaves elastically. According to eqn (3), the stress variations immediately outside the ellipsoidal region are $\delta\sigma^e = \mathbf{D}_u(\mathbf{1} - \mathbf{Q}_u^{-1})\,\delta\varepsilon^e$ in which $\mathbf{D}_u$ must now be interpreted as the matrix of elastic moduli for unloading, corresponding to the initial strain ε. This matrix is positive definite, and it is also isotropic if the material is isotropic. The uniform stress variations inside the ellipsoidal region are $\delta\sigma' = \mathbf{D}_t\,\delta\varepsilon^e$, in which $\mathbf{D}_t$ is the matrix of incremental elastic moduli for further loading, corresponding to the initial strain ε. This matrix is not positive definite if the initial state ε is in the strain-softening range, and generally it must be assumed to be anisotropic. Equilibrium of the ellipsoidal region with its exterior is maintained if $\delta\sigma' = \delta\sigma^e$. From this we obtain the following conditions of neutral equilibrium (i.e. critical state):

$$[\mathbf{D}_t - \mathbf{D}_u(\mathbf{1} - \mathbf{Q}_u^{-1})]\,\delta\varepsilon^e = 0. \tag{4}$$

It may be noted that for each point of ellipsoidal surface, at which the unit normal is $\mathbf{n}$, equilibrium requires only that $\delta\sigma'\mathbf{n} = \delta\sigma^e\mathbf{n}$. However, due to the fact that this condition must hold for various $\mathbf{n}$, it is necessary that $\delta\sigma' = \delta\sigma^e$.

Equation (4) represents a system of six homogeneous linear algebraic equations for the six components of $\delta\varepsilon^e$. A nonzero solution, which represents the strain localization instability mode, is possible if and only if the determinant of this equation system vanishes, i.e.

$$\text{Det } \mathbf{Z} = 0, \quad \text{with } \mathbf{Z} = \mathbf{D}_t - \mathbf{D}_u(\mathbf{1} - \mathbf{Q}_u^{-1}). \tag{5}$$

This is the same result as obtained previously in Bažant (1988b), in which it was further shown that the initial state of uniform strain ε in unstable if matrix $\mathbf{Z}$ is not positive definite, and is stable (with regard to the presently assumed localization mode) if this matrix is positive definite. This result was obtained (Bažant, 1988b) by analyzing the sign of the

second-order work needed to produce an equilibrium localization increment. This is the work that is done by the interfacial tractions applied on the ellipsoidal surface during an equilibrium localization increment (this work can be shown to be equal to $-T(\Delta S)_{in}$ where $(\Delta S)_{in}$ is the increment of internally produced entropy of the body; see Bažant 1988c).

The special case of the critical state condition in eqn (5) for associated or nonassociated Drucker–Prager plasticity was previously derived by Rudnicki (1977).

CRITICAL STATE CONDITIONS FOR LOCALIZATION INTO PLANAR BAND

Consider now strain localization into an infinite planar band of thickness h that forms inside an infinite layer of thickness L ($L \geqslant h$). We choose axis x_2 to be normal to the layer (Fig. 2). The layer is assumed to be initially in equilibrium in a state of uniform (homogeneous) strain ε_{ij} and stress σ_{ij} (the latin lower-case subscripts refer to Cartesian coordinates x_i, $i = 1, 2, 3$). We imagine the initial equilibrium state to be disturbed by infinitesimal displacement variations δu_i whose gradients $\delta u_{i,j}$ have the values $\delta u^l_{i,j}$ and $\delta u^u_{i,j}$ inside and outside the band, representing further loading or unloading, respectively. These variations are uniform within the band and also outside the band, with discontinuous jump across the surface of the band. As the boundary conditions, we consider that the surface points of the layer are fixed during the incremental deformation, i.e. $\delta u_i = 0$ at the surfaces $x_2 = 0$ and $x_2 = L$ of the layer. The stress variations inside the band and outside the band are:

$$\delta\sigma^l_{2i} = D^l_{2ijm}\,\delta\varepsilon^l_{jm} = D^l_{2ijm}\,\delta u^l_{j,m} = D^l_{2ij2}\,\delta u^l_{j,2} \tag{6}$$

$$\delta\sigma^u_{2i} = D^u_{2ijm}\,\delta\varepsilon^u_{jm} = D^u_{2ijm}\,\delta u^u_{j,m} = D^u_{2ij2}\,\delta u^u_{j,2} \tag{7}$$

in which again D^l_{kijm} and D^u_{kijm} represent the incremental elastic moduli for further loading and for unloading, corresponding to the initial strain ε_{ij}. In writing eqns (6)–(7) we assume that $\delta u^u_{i,1} = \delta u^l_{i,1} = 0$, $\delta u^u_{i,3} = \delta u^l_{i,3} = 0$. Without these conditions, there would be a slip continuity between the inside and outside of the band.

Compatability of deformations over the thickness of the layer requires that $h\delta u^l_{j,2} + (L-h)\,\delta u^u_{j,2} = 0$ ($j = 1, 2, 3$). Expressing $\delta u^u_{j,2}$ from this equation and taking into account the condition of equilibrium at the band surfaces, $\delta\sigma^u_{2i} = \delta\sigma^l_{2i}$ ($i = 1, 2, 3$), we obtained from eqns (6)–(7) the following condition of neutral equilibrium (critical state):

$$\left(D^l_{2ij2} + \frac{h}{L-h}D^u_{2ij2}\right)\delta u^l_{j,2} = 0. \tag{8}$$

This represents a system of three homogeneous linear algebraic equations. A nonzero solution exists if and only if

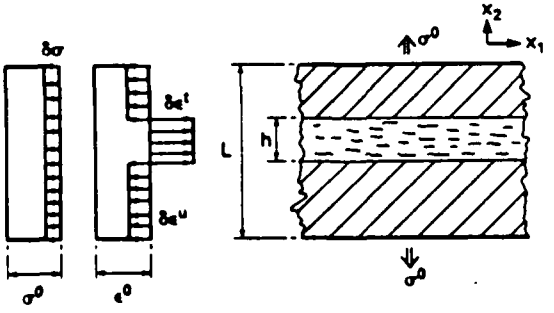


Fig. 2. Planar localization band in a layer.

$$\text{Det } Z_{ij} = 0, \quad \text{with } Z_{ij} = D'_{2ij2} + \frac{h}{L-h} D''_{2ij2}. \tag{9}$$

This critical state condition is again the same as obtained previously by Bažant (1988a), in which it was also proven (by analysis of the second-order work) that the states for which the matrix Z_{ij} is not positive definite are unstable, and the states for which it is positive definite are stable with regard to localization into a planar band. A special case of eqn (9) is the condition for uniaxial or shear localization instability derived in 1974 by Bažant (1976).

The special case of eqn (9) for nonassociated Drucker–Prager plasticity without or with vertex hardening and for infinite space (for which $L \to \infty$ and unloading $\mathbf{D}''$ plays no role) was obtained by Rudnicki and Rice (1975) and by Rudnicki (1977).

BASIC PROPERTIES OF THE CRITICAL STATES

The critical state conditions in eqns (5) and (9) were discussed in detail by Bažant (1988a, b). Briefly, according to eqn (5) which refers to an infinite continuum, the critical state condition for localization into an ellipsoidal region is independent of the region's size, and depends only on the aspect ratios a_1/a_2 and a_3/a_2 of the ellipsoid. By contrast, according to eqn (9), which applies to a finite body, the critical state condition does depend on the size of the localization region, in this case the thickness h of the band. The smaller the thickness of the band, the more stringent is the stability condition, i.e. the smaller is the magnitude of the strain-softening slope (tangent modulus) at which instability occurs. If $h \to 0$ were allowed, instability would always occur no later than at the peak stress point, and so strain-softening could never be observed. The fact that it can be observed implies that the thickness of the localization band, h, cannot be considered to be smaller than a certain finite length. This length must be considered to be a material property. It is roughly equal to the characteristic length l of the material.

It is now proper to comment on the difference from Hill's (1962) condition of bifurcation of the equilibrium path, $D'_{2ij2}\delta u'_{j,2} = 0$. This equation coincides with Rudnicki and Rice's (1975) condition and is obtained from eqn (8) with $L/h \to \infty$. As a bifurcation condition, however, this equation has been used [in contrast to eqn (8)] for a finite body of any size. In that case there is no unloading in the body, same as for a layer of infinite thickness. So the moduli D'_{ijkm} may be assumed to apply everywhere, which means the analysis can be made on the basis of a linear elastic body; see Hill's concept of a linear comparison solid. It must be noted, however, that the eigenmode $\delta u^*_{i,j}$ obtained from the equation $D'_{2ij2}\,\delta u'_{j,2} = 0$ violates the boundary condition of fixed displacements at the layer surface. Therefore, neutral equilibrium does not exist in the layer at the bifurcation state, and the calculated eigenmode $\delta u^*_{i,j}$ cannot actually occur as an instability. This means that the bifurcation must occur at increasing rather than constant boundary displacements. If the boundary displacement increase is controlled, such a mode of deformation does not represent stability loss, whose investigation is the objective of the present study. Rather, the bifurcation state as well as the states immediately after the bifurcation are stable.

The foregoing discussion reflects the fact that one must distinguish between the concepts of stable state and stable path, as shown in Bažant's (1988c) study. It was proven (for the special case of uniaxial stress) that after the bifurcation point, which occurs according to Hill (1962) or Rudnicki and Rice (1975) at the peak stress state, the strain must begin to localize if the material is deformed in an equilibrium manner and if the tangential moduli vary continuously as the layer is loaded in a displacement-controlled fashion. However, it was also shown that if a nonlocalized uniform post-peak state is somehow reached, this state is stable if eqn (8) is satisfied. Such a state can be reached in an equilibrium loading process if the tangential moduli decrease discontinuously during loading (or drop suddenly due to heating or other effects). Even for continuously varying moduli, such a state can be reached dynamically, or if certain temporary restraints are applied, then also statically.

Whether the aforementioned distinction between the stability limit and the bifurcation state exists also for finite ellipsoidal regions is not immediately clear. Further study is planned.

The critical state conditions in eqns (5) and (9) were examined in the previous works (Bažant, 1988a, b) for the effect of material parameters. Consideration was limited to incrementally isotropic tensors D_{ijkm}. For the ellipsoidal localization regions, it was found that the critical slope E_t^{cr} of the stress–strain diagram (<0) at which instability occurs decreases in magnitude as the ellipsoid becomes flatter, more elongated in the x_1 and x_3 directions, and becomes zero as the ellipsoid approaches an infinite band, i.e. for $a_1 \to \infty$, $a_3 \to \infty$. Rather large ratios a_1/a_2 and a_3/a_2, however, are needed to make $|E_t^{cr}|$ small. For example, for $|E_t^{cr}|$ to become less than about 5% of the unloading slope E_u, the aspect ratios a_1/a_2 and a_3/a_2 need to exceed approximately 200. This is a surprisingly large value.

Such an elongated ellipsoidal region (whose thickness for concrete may not be less than several times the aggregate size) often cannot be accommodated within a finite body representing a typical concrete structure. This ellipsoidal region would have to be contained within a larger region that is sufficiently remote from the surface of the body so that the boundary conditions of infinite space around the ellipsoidal region would be approximately applicable. Therefore, real structures should often be stable even for strain-softening slopes whose magnitude is much larger than that predicted by the formula for the planar band [eqn (9) and also Bažant, 1976] using a characteristic length of the same order of magnitude as the maximum size of inhomogeneity in the material.

Equations (5) and (9) can be used to study the influence of material parameters on the localization instabilities for more realistic constitutive laws, for which the tensor D^t_{ijkm} is anisotropic. We will do so now for some plastic constitutive models.

PARAMETER STUDY FOR NONASSOCIATED DRUCKER–PRAGER PLASTICITY

The plastic stress–strain relations for isotropic materials have the form:

$$ds_{ij} = 2G(de_{ij} - de^p_{ij}), \quad d\sigma^v = 3K(d\varepsilon^v - d\varepsilon^p) \tag{10}$$

in which G, K = shear and bulk elastic moduli, $\varepsilon^v = \varepsilon_{kk}/3$ = volumetric strain, $e_{ij} = \varepsilon_{ij} - \delta_{ij}\varepsilon^v$ = deviatoric strains (δ_{ij} = Kronecker delta), $\sigma^v = \sigma_{kk}/3$ = volumetric stress, $s_{ij} = \sigma_{ij} - \delta_{ij}\sigma^v$ = deviatoric stresses, and ε^p, e^p_{ij} = volumetric and deviatoric plastic strains. The plastic strain increments are given by the flow rule:

$$d\varepsilon^p_{ij} = d\lambda \frac{\partial g}{\partial \sigma_{ij}} \tag{11}$$

in which $d\lambda$ is the plastic strain parameter and g is the plastic potential function. According to the plasticity theory proposed by Drucker and Prager (see Fung, 1965; Chen, 1962), the loading function and the plastic potential function are introduced as follows:

$$f(\sigma, \kappa) = \bar{\tau} + \psi(\sigma^v) - \kappa = 0 \tag{12}$$

$$g(\sigma, \kappa) = \bar{\tau} + \phi(\sigma^v) - \kappa = 0 \tag{13}$$

in which $\bar{\tau} = J_2^{1/2} = (\frac{1}{2}s_{ij}s_{ij})^{1/2}$ = stress intensity, J_2 = stress deviator, κ = plastic hardening parameter, and ϕ, ψ = empirical material functions of the volumetric stress. As usual, we introduce the notations:

$$H = \frac{d\kappa}{d\lambda}, \quad \beta = \frac{\partial \phi}{\partial \sigma^v}, \quad \beta' = \frac{\partial \psi}{\partial \sigma^v}, \tag{14}$$

in which H = plastic modulus, which is positive for hardening and negative for softening,

β' = internal friction coefficient, and β = dilatancy ratio. The formulation satisfies the normality rule if $g \equiv f$, which means that $\phi = \psi$ (and also $\beta = \beta'$). In this case, the material is said to be associated, and otherwise nonassociated. H, β and β' completely characterize the incremental properties of Drucker–Prager plastic material. Traditionally H has been considered positive (which describes plastic hardening, i.e. increase of the yield limit). However, allowing it to be negative provides a convenient model for strain-softening, describing degradation of the yield limit (a model of this type has been used in nonlocal analysis of tunnel cave-in; see Bažant and Lin, 1988).

On the basis of eqns (10)–(14), it can be shown by a well-known procedure that the tangential (incremental) moduli are

$$D^t_{ijkm} = D^e_{ijkm} - \frac{\left(\frac{G}{\bar{\tau}} s_{ij} + K\beta\delta_{ij}\right)\left(\frac{G}{\bar{\tau}} s_{km} + K\beta'\delta_{km}\right)}{G + K\beta\beta' + H}. \tag{15}$$

For nonassociated plasticity ($\beta' \neq \beta$) the tangential moduli for loading are nonsymmetric, while for associated plasticity ($\beta' = \beta$) they are symmetric (with respect to interchanging ij with km). The tangential moduli for unloading are assumed to be the same as for initial elastic loading, and so

$$D^u_{ijkm} = (K - \tfrac{2}{3}G)\,\delta_{ij}\,\delta_{km} + 2G\delta_{ik}\,\delta_{jm}. \tag{16}$$

Although the material is assumed to be isotropic, the tensor of incremental moduli for further loading [eqn (15)] exhibits stress-induced anisotropy.

After implementing the Drucker–Prager model in the computer program, it has first been checked whether the results for the case $a_1/a_2 \to \infty$, $a_3/a_2 \to \infty$ agree with the solution of Rudnicki and Rice (1975). They do.

The critical state condition in eqn (5) or (9) depends on the following four nondimensional material parameters

$$\nu,\ H/G,\ \beta,\ \beta' \tag{17}$$

in which ν = Poisson's ratio. In addition it depends on $s_{ij}/\bar{\tau}$, i.e. the ratios of the initial deviatoric stresses, and also on nondimensional geometric parameters such as L/h for localization in a band, and a_1/a_2, a_3/a_2 for localization in an ellipsoid. It does not depend on the elastic modulus E, however, because the division of eqn (5) or (9) by E has no effect on its eigenvalues.

The effect of the aforementioned parameters has been studied numerically. The eigenvalues of matrix $\mathbf{Z}$ in eqn (5) or matrix Z_{ij} in eqn (9) have been calculated on a computer for various combinations of the parameters in eqn (17). Using Newton's iterative method, material parameter combinations for which the smallest eigenvalue vanishes have been found, and their plots are shown in Figs 3–13. The initial stress states considered in these diagrams were: uniaxial compression, biaxial compression, pure shear ($s_{12} \neq 0$, other $s_{ij} = 0$), and uniaxial tension. Poisson's ratio has been considered as $\nu = 0.18$ for all the calculations.

Figures 3–5 show the effect of the ratio of the layer thickness L to the band thickness h on the value of the plastic modulus H at which the stability limit is reached. The states below each curve are stable, the states above it are unstable, and the states on the curve are critical. In Figs 3–5 on top we see the curves for various values of β' when $\beta = \beta'$ (i.e. for associated plasticity, for which the normality rule is satisfied). In Figs 3–5 at the bottom we see the curves for various values of the internal friction coefficient β' when the dilatancy ratio is fixed as $\beta = 1$. As might have been expected, the thicker the given layer, the lower is the magnitude of the plastic modulus H at which the layer becomes unstable.

To check the limit case $L/h \to 1$, we rewrite eqn (8) as $[D^u_{2j2} + D^t_{2j2}(L-h)/h]\,\delta u^t_{j,2} = 0$. Since the equation $D^u_{2j2}\,\delta u^t_{j,2} = 0$ has no nonzero solution (D^u_{2j2} is positive-definite), we

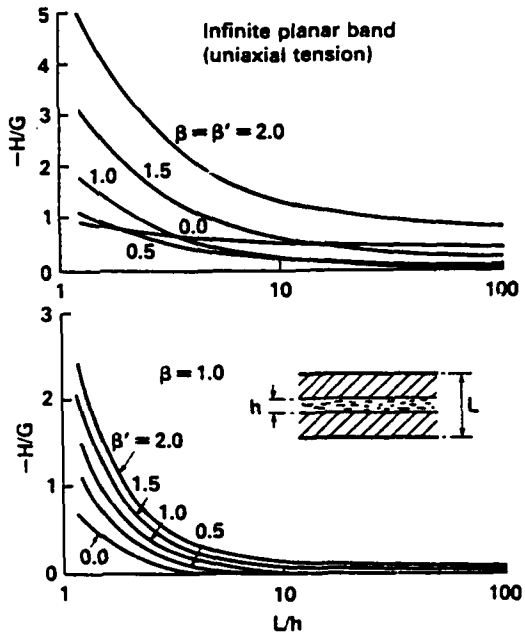


Fig. 3. Effect of size L/h on critical plastic hardening modulus H for infinite planar band in uniaxial tension.

conclude that at least some component of D'_{2j2} must tend to ∞ to have the posssibility of localization (nonzero $\delta u'_{j,2}$). According to eqn (15) this occurs if and only if $H \to -\infty$. Therefore, the curves in Figs 3–5 tend to ∞ as $L/h \to 1$.

Another check on the present general numerical solution is provided by the special case of a band in an infinite space solved by Rudnicki and Rice (1975), which corresponds to $L/h \to \infty$. Indeed, their eqn (16) yields numerically the same results as the present solution for $L/h \to \infty$ and the Drucker–Prager material. It should be noted that Rudnicki and Rice considered the band to have arbitrary orientation and used among the critical state

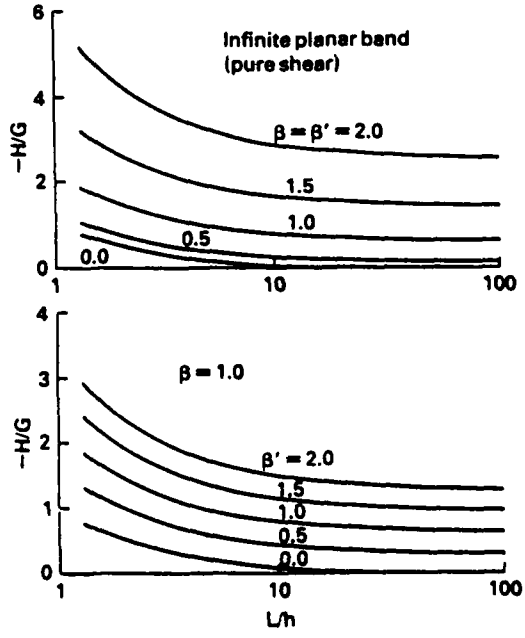


Fig. 4. Effect of size on critical H for infinite planar band in pure shear.

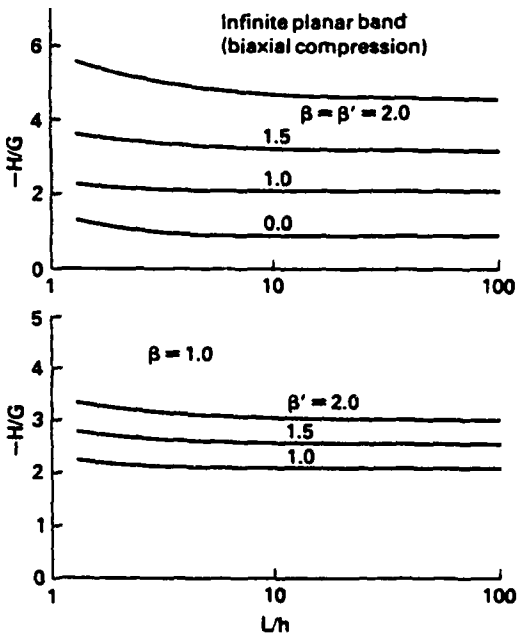


Fig. 5. Effect of size on critical H for infinite planar band in biaxial compression.

conditions for all possible orientations of axis x_2 the one which is most severe. In the present study, on the other hand, the orientation of the band (axis x_2) is fixed because the band of thickness h must be parallel to the surfaces of the layer of given thickness L (except when $L \to \infty$).

The asymptotic values of the curves are interesting. For an infinitely thick layer (i.e. infinite space), the strain localization instability does not occur at zero plastic modulus ($H \to 0$), as might have been expected, but at a finite negative value of the plastic modulus. This means that strain-softening according to the Drucker–Prager plasticity model is stable

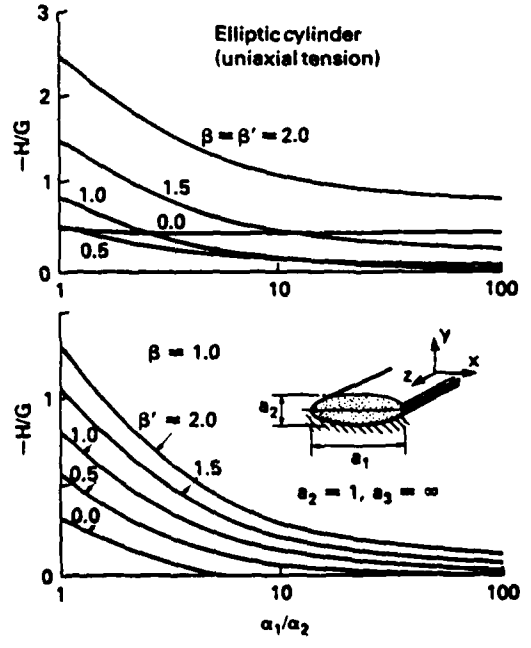


Fig. 6. Effect of ratio a_1/a_2 of ellipsoidal axes on critical H for elliptic cylinders in uniaxial tension.

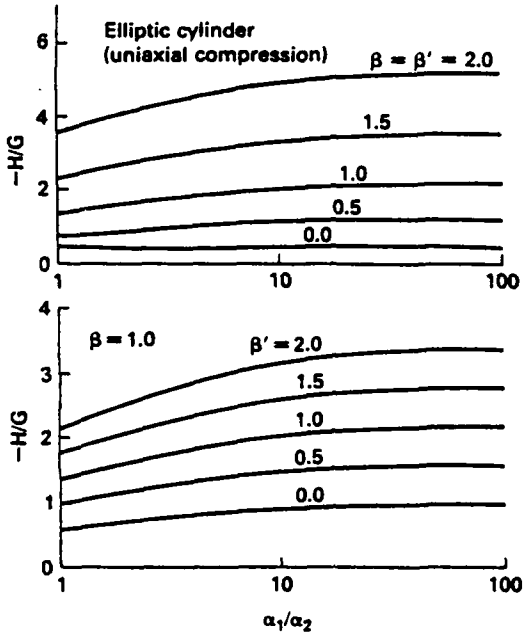


Fig. 7. Effect of axis ratio on critical H for elliptic cylinders in uniaxial compression.

against localization into a planar band even in an infinite space, provided that the plastic modulus magnitude does not exceed a certain limit.

By contrast, in the previous work (Bažant, 1988a), which considered only incrementally isotropic material properties, the value of the softening modulus E_t (which is analogous to H) always approaches zero for an infinitely thick layer. This implies that the strain-softening cannot be stable in an infinitely thick layer if the incremental modulus matrix is isotropic. Apparently, strain-softening of plasticity type (i.e. degradation of the yield limit) is not as destabilizing as some other forms of incremental material properties.

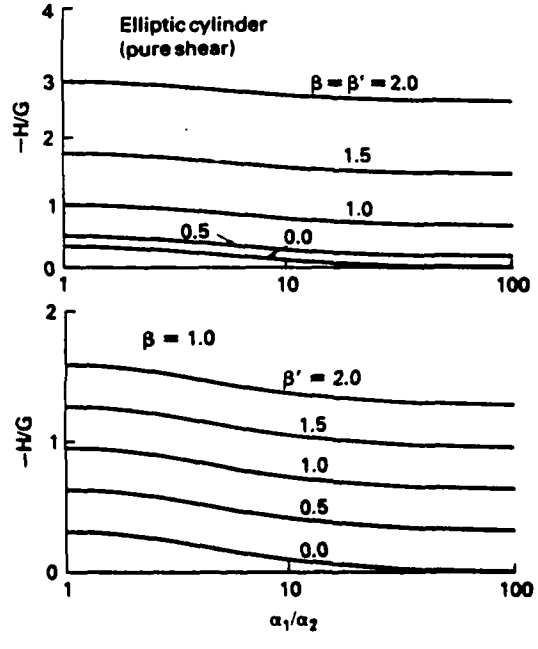


Fig. 8. Effect of axis ratio on critical H for elliptic cylinders in pure shear.

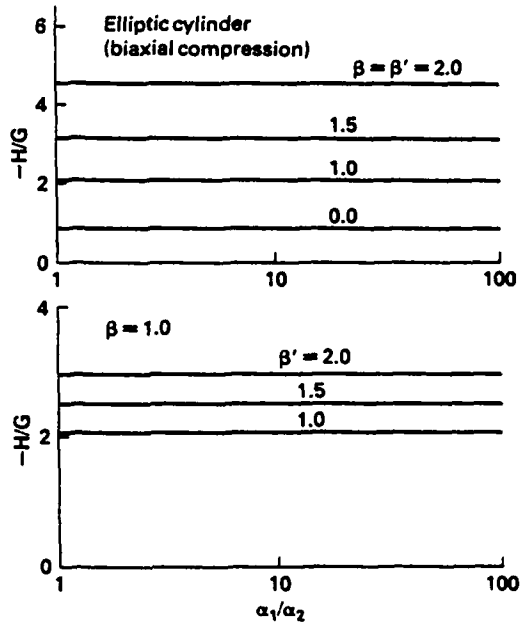


Fig. 9. Effect of axis ratio on critical H for elliptic cylinders in biaxial compression.

Figures 6–9 show the dependence of the critical plastic modulus on the aspect ratio a_1/a_2 of a strain-softening ellipsoid degenerated into an infinite elliptic cylinder, i.e. $a_3/a_2 \to \infty$. This limiting case is equivalent to localization in an infinite planar band within an infinite space. So the asymptotic values from Figs 6–9 and Figs 3–5 must coincide. They indeed do, which serves as a check on the correctness of the calculations.

Again it might have been expected that for an infinite aspect ratio a_1/a_2, the critical value of H should vanish. Surprisingly, it does not. The preceding comments apply here, too.

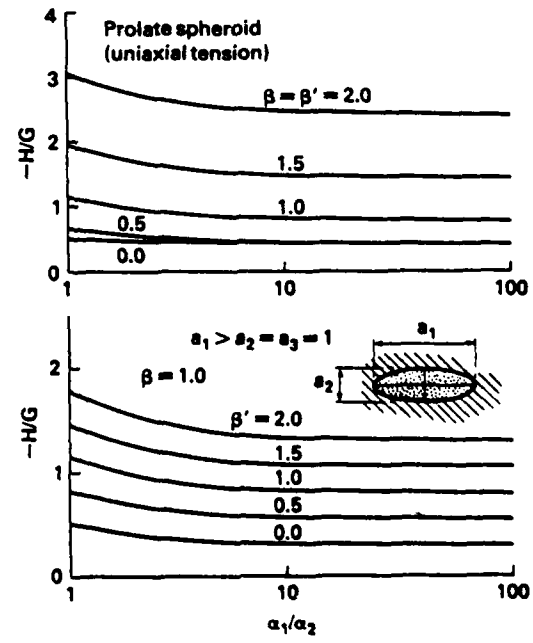


Fig. 10. Effect of axis ratio on critical H for prolate spheroids in uniaxial tension.

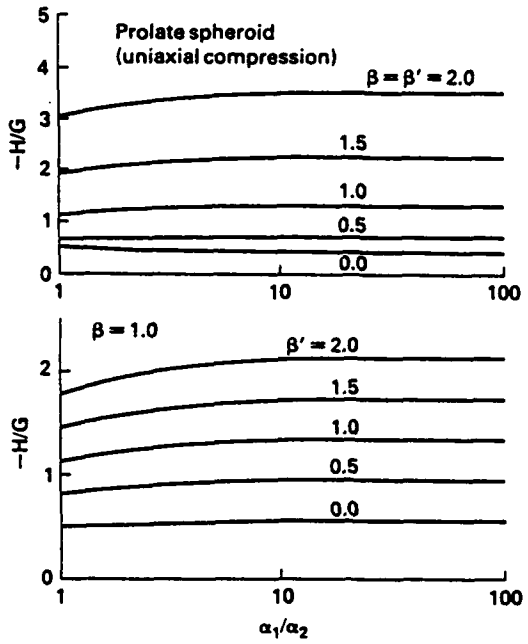


Fig. 11. Effect of axis ratio on critical H for prolate spheroids in uniaxial compression.

Furthermore, it should be noted that generally very high aspect ratios of the ellipsoid are needed to reduce the critical magnitudes of H near that for localization into a band in a layer. This of course means, similar to the conclusion of the previous study (Bažant, 1988a, b), that the localization stability limits for a planar band are not well applicable to finite bodies representing typical structures.

Figures 10–13 show the dependence of the critical magnitudes of H on the aspect ratio for prolate spheroids ($a_2 = a_3$). As the aspect ratio tends to infinity, the prolate spheroid approaches an infinite cylinder (a fiber). This is a different limiting case than before, and vanishing of the magnitude of H is expected. It is also noteworthy that the strain-softening

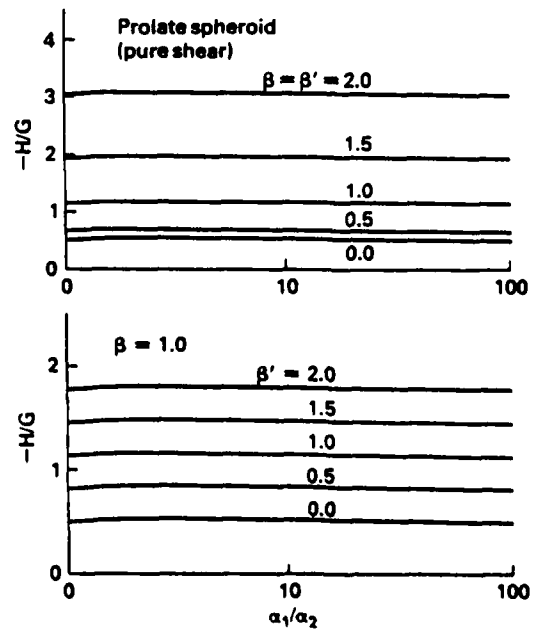


Fig. 12. Effect of axis ratio on critical H for prolate spheroids in pure shear.

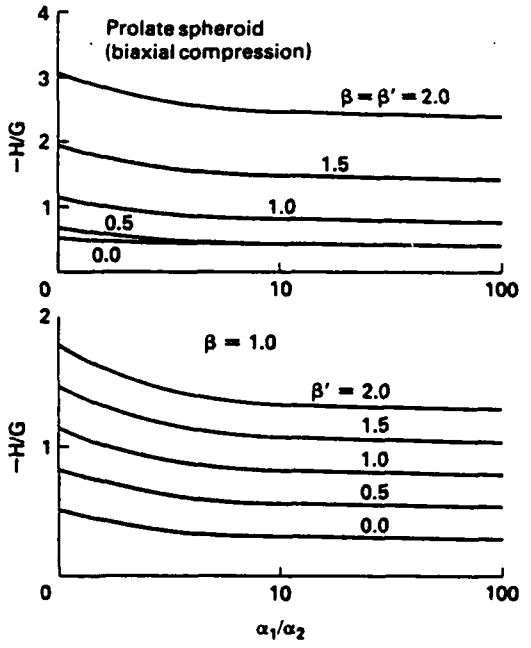


Fig. 13. Effect of axis ratio on critical H for prolate spheroids in biaxial compression.

in prolate spheroids is much more stable than in elliptic cylinders, i.e. the critical magnitudes of H are much larger in Figs 10–13 than in Figs 6–9.

The eigenvectors of matrix $\mathbf{Z}$ or Z_{ij} for the critical states for which the lowest eigenvalue is zero have been calculated too. For several typical cases, these eigenvectors are shown graphically in Fig. 14 for the volumetric (Rendulić) sections of the loading surface, and at bottom right also for the deviatoric cross-section (the values $\nu = 0.18$, $\beta' = 1.0$ and uniaxial tension have been assumed in these calculations). In plotting these eigenvectors, it is understood that the axes of ε^p and the deviatoric plastic strain intensity $\bar{\gamma} = (\frac{1}{2}e^p_{ij}e^p_{ij})^{1/2}$ are superimposed on the axes of the volumetric stress and the deviatoric stress intensity. It is interesting that in the volumetric cross-section these eigenvectors are always inclined away from the normal to the loading function and toward the deviatoric axis, for all the dilatancy ratios β. These eigenvectors, of course, need not be normal to the loading function even if normality ($\beta' = \beta$) is assumed (it is the plastic strain incremental vector which must be normal if $\beta' = \beta$).

CONCLUSIONS

Based on the general stability conditions for localization of strain-softening into planar bands and ellipsoidal domains, the critical values of the plastic modulus at stability limit are calculated for nonassociated and associated Drucker–Prager plasticity. Also presented in a direct derivation of the critical state conditions, which is simpler than the previous derivation of the stability conditions. It is found that strain-localization instabilities of the type considered do not occur for a certain range of material parameters even if strain-softening takes place. This is true also for strain-localization into a planar band in an infinite continuum, even though previous uniaxial analysis (Bažant, 1976) as well as the multiaxial analysis (Bažant, 1988a) indicated such situations to be always unstable for certain other constitutive models.

Acknowledgement—Partial financial support from the U.S. Air Force Office of Scientific Research under contract F49620-87-C-0030DEF with Northwestern University, monitored by Dr Spencer T. Wu, is gratefully acknowledged.

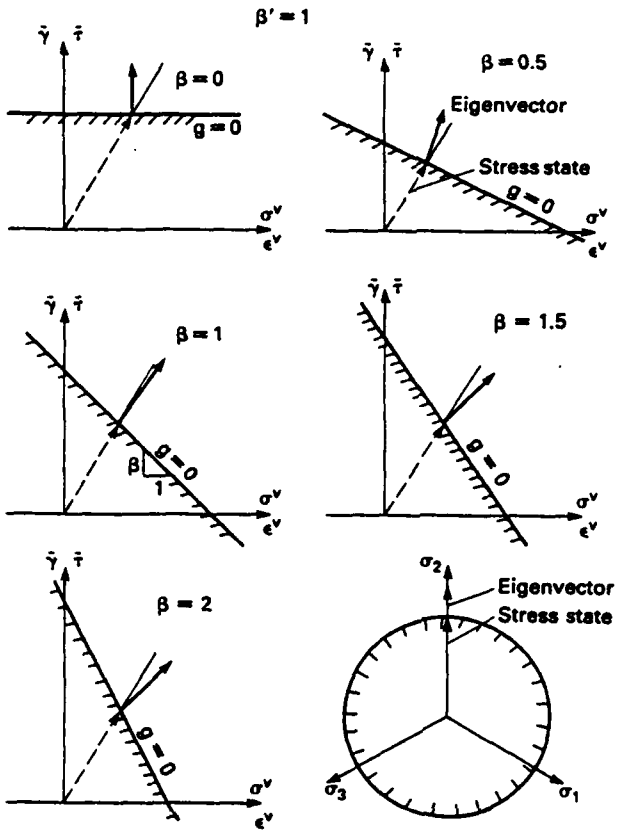


Fig. 14. Eigenvectors at the critical states in the volumetric and deviatoric sections.

REFERENCES

Bažant, Z. P. (1976). Instability, ductility and size effect in strain-softening concrete. *J. Engng Mech. Div., ASCE* **102** (EM2). 331–344; Discussions **103**, 357–358, 775–777; **104**, 501–502 (based on Struct. Engng Report No. 74-8/640, Northwestern University, August 1974).

Bažant, Z. P. (1980). Work inequalities for plastic fracturing material. *Int. J. Solids Struct.* **16**, 873–901.

Bažant, Z. P. (1982). Crack band model for fracture of geomaterials. *Proc. 4th Int. Conf. on Numerical Methods in Geomechanics*, Edmonton, Canada (Edited by Z. Eisenstein), **3**, 1137–1152.

Bažant, Z. P. (1986). Mechanics of distributed cracking. *Appl. Mech. Revs, ASME* **39** (5), 675–670.

Bažant, Z. P. (1987). Why continuum damage is nonlocal: justification by quasiperiodic microcrack array. *Mech. Res. Communs* **14** (5/6), 407–419.

Bažant, Z. P. (1988a). Softening instability: I. Localization into a planar band. *J. Appl. Mech., ASME* **55**, 517–522.

Bažant, Z. P. (1988b). Softening instability: II. Localization into ellipsoidal regions. *J. Appl. Mech., ASME* **55**, 523–529.

Bažant, Z. P. (1988c). Stable paths and states of structures with plasticity or damage. *J. Engng Mech., ASCE* **114** (Dec.

Bažant, Z. P. and Belytschko, T. B. (1987). Strain-softening continuum damage: localization and size effect. *Preprints, Int. Conf. on Constitutive Laws for Engineering Materials*, Tucson, AZ, Jan. 1987 (Edited by C. S. Desai *et al.*), Vol. 1, pp. 11–33. Elsevier, New York.

Bažant, Z. P., Belytschko, T. B. and Chang, T. P. (1984). Continuum theory for strain-softening. *J. Engng Mech. Div., ASCE* **110** (12), 1666–1692.

Bažant, Z. P. and Cedolin, L. (1979). Blunt crack band propagation in finite element analysis. *J. Engng Mech., ASCE* **105**(2), 297–315.

Bažant, Z. P. and Cedolin, L. (1980). Fracture mechanics of reinforced concrete. *J. Engng Mech. Div., ASCE* **106** (EM6), 1287–1306; Discussion and Closure (1982) **108**, 464–471.

Bažant, Z. P. and Lin, F. B. (1988). Nonlocal yield limit degradation. *Int. J. Numer. Meth. in Engng* **26**, 1805–1823; see also *Preprints, Int. Conf. On Computational Plasticity*, Barcelona, April 1987 (Edited by E. Oñate, R. Owen and E. Hinton), 1757–1780. University of Wales, Swansea.

Bažant, Z. P. and Oh, B. H. (1983). Crack band theory for fracture of concrete. *Materials and Structures* (RILEM, Paris) **16**, 155–177.

Bažant, Z. P. and Pijaudier-Cabot, G. (1987). Modeling of distributed damage by nonlocal continuum with local strain. *Preprints, 4th Int. Conf. on Numerical Methods in Fracture Mechanics*, San Antonio, Texas (Edited by A. P. Luxmore, R. J. Owen and M. F. Kanninen), 411–432.

Bažant, Z. P. and Pijaudier-Cabot, G. (1988). Nonlocal continuum damage, localization instability and conver-
:e. *J. Appl. Mech., ASME* **55**, 287–293.
Chen, W. F. (1982). *Plasticity in Reinforced Concrete*. McGraw-Hill, New York.
Christensen, R. M. (1979). *Mechanics of Composite Materials*. John Wiley, New York.
Darwin, D. (1985). Crack propagation in concrete—study of model parameters. In *Proc. U.S.–Japan Seminar on Finite Element Analysis of Reinforced Concrete Structures*, Tokyo (Edited by C. Meyer and H. Okamura), pp. 184–203. ASCE, New York.
de Borst, R. (1984). Application of advanced solution techniques to concrete cracking and non-associated plasticity. In *Numerical Methods for Non-linear Problems* (Edited by C. Taylor *et al.*), Vol. 2, 314–332. Pineridge Press, Swansea, U.K.
de Borst, R. and Nauta, P. (1984). Smeared crack analysis of reinforced concrete beams and slabs failing in shear. In *Proc. Int. Conf. on Computer-aided Analysis and Design of Concrete Structures*, Split, Yugoslavia (Edited by F. Damjanić *et al.*), pp. 161–174. Pineridge Press, Swansea, U.K.
de Borst, R. and Nauta, P. (1985). Non-orthogonal cracks in a smeared finite element model. *Engng Comput.* **2**, 35–46.
Eshelby, J. D. (1957). The determination of the elastic field of an ellipsoidal inclusion and related problems. *Proc. R. Soc. Lond.* **A241**, 376.
Fung, Y. C. (1965). *Foundations of Solid Mechanics*. Prentice-Hall, Englewood Cliffs, N.J.
Hill, R. (1962). Uniqueness criteria and extremum principles in self-adjoint problems of continuum mechanics. *J. Mech. Phys. Solids* **10**, 185–194.
Kinoshita, N. and Mura, T. (1971). Elastic fields of inclusions in anisotropic media. *Phys. Stat. Solidi* (a) **5**, 759–768.
Lin, S. C. and Mura, T. (1973). Elastic fields of inclusions in anisotropic media (II). *Phys. Stat. Solidi* (a) **15**, 281–289.
Mandel, J. (1966). Conditions de stabilité et postulat de Drucker. In *Rheology and Soil Mechanics* (Edited by J. Kravtchenko and P. M. Sirieyes), pp. 58–68. Springer, Berlin.
Mróz, Z. (1966). On forms of constitutive laws for elastic–plastic solids. *Arch. Mech. Stoss.* (Warszaw) **18**, 3–35.
Mura, T. (1982). *Micromechanics of Defects in Solids*. Martinus Nijhoff, The Hague.
Owen, D. R. J. and Hinton, E. (1980). *Finite Elements in Plasticity: Theory and Practice*. Pineridge Press, Swansea, U.K.
Pijaudier-Cabot, G. and Bažant, Z. P. (1987). Nonlocal damage: dynamic stability and convergence. Report No. 87-2/428n-II, Center of Concrete and Geomaterials, Northwestern University, Evanston, IL.; also *Computers and Structures*, in press.
Rice, J. R. (1976). The localization of plastic deformation. *Proc. 14th Int. Congress of Theoretical and Applied Mechanics*, Delft, The Netherlands, Vol. 1, 207–220.
Rudnicki, J. W. (1977). The inception of faulting in a rock mass with a weakened zone. *J. Geophys. Res.* **82**(5), 844–854.
Rudnicki, J. W. and Rice, J. R. (1975). Conditions for the localization of deformation in pressure-sensitive dilatant materials. *J. Mech. Phys. Solids* **23**, 371–394.

APPENDIX E1

STABLE STATES AND PATHS OF STRUCTURES WITH PLASTICITY OR DAMAGE

By Zdeněk P. Bažant,[1] Fellow, ASCE

ABSTRACT: Stability conditions for irreversible structural systems are formulated on the basis of the second law of thermodynamics. It is found that distinction must be made between stable equilibrium states and stable equilibrium paths. An equilibrium state is stable if any admissible deviation from it leads to a decrease of the internal entropy of the structure. Among all the equilibrium paths emanating from a bifurcation point, the stable path is that which maximizes the increment of internal entropy. These criteria are expressed in terms of the second-order incremental work, and distinction between load control and displacement control is made. Shanley's perfect elastoplastic column is analyzed as an example. It is found that the undeflected states of the column are stable up to the reduced modulus load. However, the undeflected stable states above the tangent modulus load are not reachable by a continuous loading process. The stable path is such that the deflection becomes nonzero as soon as the tangent modulus load is exceeded. Uniaxial strain-softening must localize right after the peak stress state even though the limit of stable homogeneous strain states occurs only later at a finite descending slope. The results indicate a need to include checks for path stability in inelastic finite element programs, especially those for damage with strain-softening.

INTRODUCTION

Although the conditions of uniqueness and bifurcation of the equilibrium path of an elastoplastic structure have been studied extensively (Bruhns and Mielniczuk 1977; Bruhns 1984; Hill 1959, 1961, 1962; Hutchinson 1974; Mróz 1966; Shanley 1947; Tvergaard 1985), no basic principles and methods seem to exist for stability of equilibrium states and paths, except for a few special solutions (e.g., Bažant 1976, 1977, 1985, 1986). The principle of minimum potential energy, unfortunately, does not apply to inelastic structures. As we will see, static stability analysis can nevertheless be accomplished on the basis of the second law of thermodynamics. We will show it first for the classical problem of buckling of an elastoplastic column. Based on a report by Bažant (1987), we will then formulate general criteria of stable state and stable path, and finally we will consider some broader implications, especially for softening structures.

SHANLEY'S ELASTOPLASTIC COLUMN

For the sake of simplicity, let us study the idealized perfect column considered in the epoch-making paper by Shanley (1947); see Fig. 1(*b–d*). The column is hinged and geometrically perfect. It consists of two rigid bars of lengths $l/2$, which are connected by a very short elastoplastic link (point hinge) of length $h << l$ and width h, having an ideal I-beam cross section of area A. The lateral deflection and the axial displacement at the load point (positive if shortening) are denoted as q_1 and q_2, respectively.

[1]Prof. of Civ. Engrg., Northwestern Univ., Evanston, IL 60208.

Note. Discussion open until May 1, 1989. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on November 18, 1987. This paper is part of the ***Journal of Engineering Mechanics***, Vol. 114, No. 12, December, 1988. ©ASCE, ISSN 0733-9399/88/0012-2013/$1.00 + $.15 per page. Paper No. 22993.

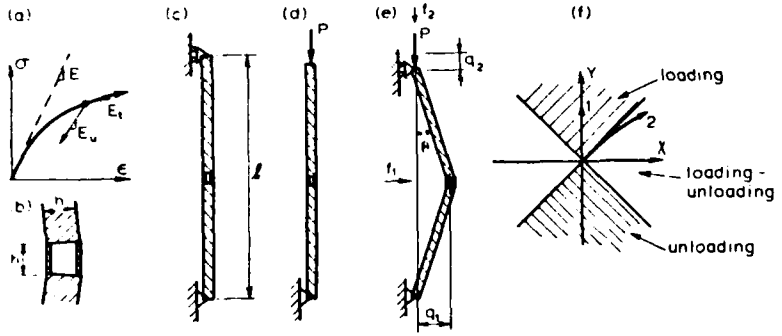


FIG. 1. Shanley's Elastoplastic Column

Initially the column is perfectly straight and is loaded by an axial centric load P (positive if compressive). The initial equilibrium under load $P = P_0$ at zero lateral deflection is disturbed by increasing the axial load to $P = P_0 + \delta f_2$ and applying a small lateral load δf_1. This causes in the flanges of the elasto-plastic link the strains $\delta\epsilon_1 = -\delta\theta - \delta q_2/h$ and $\delta\epsilon_2 = \delta\theta - \delta q_2/h$ (on the concave and convex sides, respectively); $\delta\theta = 2\delta q_1/l$ = rotation of the rigid bars, assumed to be small.

The incremental moduli for loading and unloading are denoted as E_t and E_u [Fig. 1(*a*)]; E_t is called the tangential modulus. Always $E_t < E_u$ (except when the material is elastic, in which case $E_t = E_u$). E_t is a given function of the initial uniform stress $\sigma = \sigma_1 = \sigma_2 = -P/A$. It is convenient to express the moduli at the concave and convex faces as $E_1 = \eta E_t$ and $E_2 = \eta\xi E_t$ and define the nondimensional displacements $X = \delta q_1/l$ and $Y = \delta q_2/2h$. The force variables associated by work are $f_X = l\delta f_1$ and $f_Y = 2h\delta f_2$. Based on the foregoing expressions for strains, one may then obtain for buckling to the right ($X > 0$) the following loading-unloading criteria [Fig. 1(*c*)]:

$$\text{for } Y \geq X \text{ (loading only):} \quad \xi = 1, \qquad \eta = 1 \qquad (1a)$$

$$\text{for } -X \leq Y \leq X \text{ (loading-unloading):} \quad \xi = \xi_u, \qquad \eta = 1 \qquad (1b)$$

$$\text{for } Y < -X \text{ (unloading only):} \quad \xi = 1, \qquad \eta = \xi_u \qquad (1c)$$

where $\xi_u = E_u/E_t$. Based on the incremental stresses $\delta\sigma_1 = E_1\delta\epsilon_1$ and $\delta\sigma_2 = E_2\delta\epsilon_2$ at the concave and convex faces, the moment and axial conditions of equilibrium at the midspan lead, for buckling to the right, to the equations:

$$\begin{Bmatrix} l\delta f_1 \\ 2h\delta f_2 \end{Bmatrix} = 2\eta P_t l \begin{bmatrix} 1 + \xi - 2P_0/\eta P_t, & 1 - \xi \\ 1 - \xi & 1 + \xi \end{bmatrix} \begin{Bmatrix} X \\ Y \end{Bmatrix} \qquad (2)$$

in which $P_t = E_t Ah/l$ = Shanley's tangent modulus load for modulus E_t (Bažant 1985, 1986; Bruhns and Mielniczuk 1977; Bruhns 1984; Chajes 1974). If $\delta f_2 = 0$ (P = constant), then the only nonzero solution of Eq. 2 is $P_0 = 2\xi_u P_t/(\xi_u + 1) = P_r$. P_r represents the reduced modulus load of Engesser and von Kármán (see e.g., Chajes 1974; Chen and Atsuta 1976; Chen and Lui 1987; Johnston 1976; Timoshenko and Gere 1961), at which there is neutral equilibrium [Fig. 2(*a*)]. For a straight segment of the $\sigma - \epsilon$ diagram,

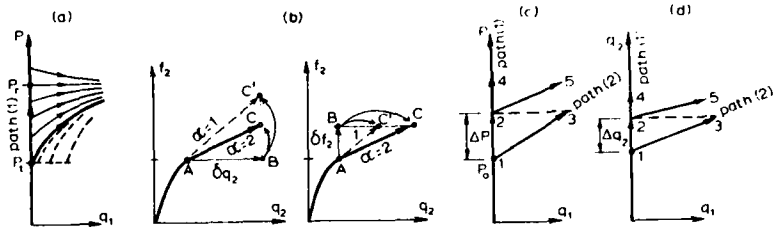


FIG. 2. Bifurcations of Equilibrium Paths

E_t, P_t, and P_r are constant. When E_t depends on σ (curved $\sigma - \epsilon$ diagram), P_t and P_r depend on $\sigma_0 = -P/A$ through $E_t(\sigma_0)$.

Because the inequalities in Eq. 1 offer several alternatives, Eq. 2 for $\delta f_1 = 0$ permits more than one equilibrium path. One path, called the primary path or path 1 (Fig. 2), is characterized by zero deflection and $\xi = \eta = 1$;

$$X^{(1)} = 0, \quad \delta f_2^{(1)} = 2P_t \frac{l}{h} Y^{(1)} \qquad (3)$$

For $P_0 \to P_t$, however, $X^{(1)}$ is arbitrary since the matrix of Eq. 2 becomes singular (but $X^{(1)} < Y^{(1)}$ because $\xi = 1$). Another path, called the secondary path or path 2 (Fig. 2), is characterized by positive deflection (at $f_1 = 0$), $\eta = 1$ and $\xi > 1$. For this path Eqs. 2 and 3 yield

$$X^{(2)} = \frac{\xi_u - 1}{\xi_u + 1 - \dfrac{2P_0}{P_t}} Y^{(2)}, \quad \delta f_2^{(2)} = \frac{P_t l}{h} \frac{4\xi_u P_t - 2(\xi_u + 1)P_0}{(\xi_u + 1)P_t - 2P_0} Y^{(2)} \qquad (4)$$

The superscripts (1) and (2) refer to paths 1 and 2. Path 2 according to Eq. 4 is possible only if $\xi > 1$, and $\xi > 1$ is possible only if $P_0 \geq P_t$. Since the solutions in Eqs. 3 and 4 exist for each point $P_0 \geq P_t$ [solid curves in Fig. 2(*a*)], the main path for $P \geq P_t$ represents a continuous sequence of bifurcation points. The first bifurcation occurs at $P_0 = P_t$, as shown by Shanley [Fig. 2(*a*)]. Note that for $P_0 = P_t$ we have $Y^{(2)}/X^{(2)} = 1$, i.e. the secondary path starts along the boundary of the loading-only sector in the (X, Y) plane.

To analyze stability, we calculate the second-order incremental work $\delta^2 W$ of small equilibrium forces δf_1 and δf_2 on arbitrary small incremental displacements δq_1 and δq_2. We have $\delta^2 W = 1/2(\delta f_1 \delta q_1 + \delta f_2 \delta q_2)$, and substitution of Eq. 2 provides:

$$\delta^2 W(X, Y) = \frac{P_t l \eta}{\xi + 1} \left\{ [(\xi + 1)Y - (\xi - 1)|X|]^2 + 4\xi \left(1 - \frac{(\xi + 1)P_0}{2\xi P_t \eta} \right) X^2 \right\} \qquad (5)$$

The absolute value $|X|$ is introduced here in order to make Eq. 5 valid for buckling to both right and left.

Eq. 5 can alternatively also be derived as $\delta^2 W = \delta^2 W_0 + \delta^2 W_\sigma$ in which $\delta^2 W_0 = (1/2)(\delta\sigma_1\delta\epsilon_1 + \delta\sigma_2\delta\epsilon_2)\, hA/2$ = second-order work of stresses, and $\delta^2 W_\sigma = -P_0 \Delta l$, with $\Delta l = l(1 - \cos\theta) \simeq l(\delta\theta)^2/2 = 2lX^2$ = second-order axial displacement at the load point. Note that $\delta^2 W_\sigma = (1/2)(G_{11}X^2 + 2G_{12}XY$

$= G_{22}Y^2$) where G_{ij} = geometric stiffness matrix; and $G_{11} = -4P_0l$, $G_{12} = G_{22} = 0$.

Eq. 5 applies in general if both the axial load and the column length can change during the incremental deformation. Under a displacement-controlled mode of loading, we have $\delta q_2 = 0$ during buckling, and then $\delta^2W = 1/2\ \delta f_1\delta q_1$, i.e. (for $Y \geq 0$):

$$\delta^2W = [\delta^2W]_{Y=\text{const}} = \delta^2W(X) = \eta P_t l\left(\xi + 1 - \frac{2P_0}{\eta P_t}\right)X^2 \quad \ldots\ldots (6)$$

On the other hand, under a load-controlled mode of loading ($\delta f_2 = 0$):

$$\delta^2W = [\delta^2W]_{P=\text{const}} = \delta^2W(X) = \eta P_t l\left(\frac{4\xi}{\xi + 1} - \frac{2P_0}{\eta P_t}\right)X^2 \quad \ldots\ldots (7)$$

To identify the stable path, we will also need the work δ^2W done along the equilibrium paths (for $f_1 = 0$). For path 2, $\delta^2W^{(2)} = 1/2\delta f_2^{(2)}\delta q_2^{(2)}$. Substituting Eq. 4 we get, in terms of Y,

$$\delta^2W^{(2)} = \frac{2\xi_u P_t - (\xi_u + 1)P_0}{(\xi_u + 1)P_t - 2P_0}\,2P_t l Y^2 \quad \ldots\ldots (8)$$

and in terms of δf_2

$$\delta^2W^{(2)} = \frac{(\xi_u + 1)P_t - 2P_0}{2\xi_u P_t - \xi_u - 1}\,\frac{h^2}{2P_t l}\,\delta f_2^2 \quad \ldots\ldots (9)$$

For path 1, $\delta^2W^{(1)} = 1/2\delta f_2^{(1)}\delta q_2^{(1)} = 2P_t l Y^2$ in terms of Y, or $\delta^2W^{(1)} = h^2\delta f_2^2/2P_t l$ in terms of δf_2, according to Eq. 3. The work difference is found to be, when Y is the same,

$$\delta^2W^{(1)} - \delta^2W^{(2)} = \frac{(P_0 - P_t)(\xi_u - 1)P_t l}{\left(1 + \dfrac{(\xi_u - 1)^2}{4\xi_u}\right)P_t - P_0}\,Y^2 \quad \ldots\ldots (10)$$

and when δf_2 is the same,

$$\delta^2W^{(1)} - \delta^2W^{(2)} = -\frac{(P_0 - P_t)(\xi_u - 1)h^2}{2(P_t - P_0)(\xi_u + 1)P_t l}\,\delta f_2^2 \quad \ldots\ldots (11)$$

General Criterion for Stability of State

The state of the structure is characterized by displacements q_k ($k = 1, 2, \ldots, n$). The total energy and Helmholtz free energy of the structure are defined as $U = W + Q$, $F = U - TS$ where $dW = \mathbf{P}\cdot d\mathbf{q}$ = work of loads $P_1, \ldots, P_n$ (applied forces), $\mathbf{q} = (q_1, \ldots, q_n)$ = vector of load-point displacements, $\mathbf{P} = (P_1, \ldots, P_n)$, T = absolute temperature (assumed to be uniform), dQ = heat that flows into the structure from the outside, and S = entropy of the structure, which is defined as $dS = (dQ/T) + dS_{in}$ where dS_{in} = internal entropy of the structure. From these relations, one can verify that

$$dF = \mathbf{P}\cdot d\mathbf{q} - SdT - TdS_{in} \quad \ldots\ldots (12)$$

$$dU = \mathbf{P}\cdot d\mathbf{q} + TdS - TdS_{in} \quad \ldots\ldots (13)$$

Considering isothermal ($dT = 0$) and isentropic ($dS = 0$) infinitesimal equilibrium deformations (which must be path-independent in the small), we have, according to Eqs. 12 and 13,

$$dF = \mathbf{f}_T\cdot d\mathbf{q}, \quad dU = \mathbf{f}_S\cdot d\mathbf{q} \quad \ldots\ldots (14)$$

where $\mathbf{f}_T$, $\mathbf{f}_S$ are the vectors of associated forces (reactions) that depend on $\mathbf{q}$ according to isothermal or isentropic material properties; $\mathbf{f}_T = (f_1, \ldots, f_n)_T$, $\mathbf{f}_S = (f_1, \ldots, f_n)_S$. These forces must be distinguished from the applied loads which follow their own law (e.g. gravity) that is independent of the response of the structure.

Because of the principle of conservation of energy,

$$\delta F = \int_V \boldsymbol{\sigma}_T{:}\delta\boldsymbol{\epsilon}\,dV, \quad \delta U = \int_V \boldsymbol{\sigma}_S{:}\delta\boldsymbol{\epsilon}\,dV \quad \ldots\ldots (15)$$

where V = volume of the body, $\delta\boldsymbol{\epsilon}$ = strain tensor increment, $\boldsymbol{\sigma}_T$, $\boldsymbol{\sigma}_S$ = current stress tensors calculated from the history of $\boldsymbol{\epsilon}$ on the basis of isothermal or isentropic material properties. In the case of geometric nonlinearity or buckling phenomena, only certain types of stress and strain tensors, known from finite-strain continuum mechanics, can be used for $\boldsymbol{\sigma}$ and $\boldsymbol{\epsilon}$ in order to ensure equivalence of the work of stresses (Eq. 15) to the work of reaction forces (Eq. 14) (see, e.g., Hill 1958, 1962; Maier 1971).

If the problem is geometrically linear, the equivalence of Eqs. 14 and 15 also follows from the principle of virtual work (Maier 1971; Maier and Drucker 1973) because $\boldsymbol{\sigma}$ is in equilibrium with $\mathbf{f}$ and $\boldsymbol{\epsilon}$ is compatible with $\mathbf{q}$. Note that the term TdS_{in} does not belong in Eqs. 14 and 15 because $\mathbf{f}_T$ and $\mathbf{f}_S$ as well as $\boldsymbol{\sigma}_T$ and $\boldsymbol{\sigma}_S$ must be evaluated to include both reversible and irreversible responses (see Eqs. 19 and 20 where $\mathbf{K}_{jk}$ and $\mathbf{C}_T$ include plastic stiffness). Because $dS_{in} = 0$ in Eqs. 14 and 15, we have $dS = dQ/T$, and so the isentropic ($dS = 0$) and adiabatic ($dT = 0$) material properties are equivalent for the determination of $\mathbf{f}_S$ and $\boldsymbol{\sigma}_S$.

It is convenient to introduce the Helmholtz free energy $\mathcal{F}$ and the total energy $\mathcal{U}$ of the structure-load system; $\mathcal{F} = F - W$ and $\mathcal{U} = U - W$. For an elastic structure, $\mathcal{F}$ and $\mathcal{U}$ reduce to the potential energy. But for an inelastic structure they do not represent potentials in the mathematical sense since they are irreversible, path-dependent variables. According to Eqs. 12–15 we obtain

$$d\mathcal{F} = -SdT - TdS_{in} = \int_V \boldsymbol{\sigma}_T{:}d\boldsymbol{\epsilon}\,dV - \mathbf{P}\cdot d\mathbf{q} = \mathbf{f}_T\cdot d\mathbf{q} - \mathbf{P}\cdot d\mathbf{q} \quad \ldots\ldots (16)$$

$$d\mathcal{U} = TdS - TdS_{in} = \int_V \boldsymbol{\sigma}_S{:}d\boldsymbol{\epsilon}\,dV - \mathbf{P}\cdot d\mathbf{q} = \mathbf{f}_S\cdot d\mathbf{q} - \mathbf{P}\cdot d\mathbf{q} \quad \ldots\ldots (17)$$

Consider now a change from the initial equilibrium state $\mathbf{q}^0$ to a neighboring state $\mathbf{q}^0 + \delta\mathbf{q}$. Since inelastic response must be path-independent in the small, we may introduce Taylor series expansion of function $\mathbf{f}(\mathbf{q})$. Assuming the applied loads $\mathbf{P}$ to be constant (dead loads) we thus obtain

$$\Delta\mathcal{F} = \int_{q_j^0}^{q_j^0 + \delta q_j} \sum_j \left[f_j^0 + \sum_k f_{j,k}^0(q_k - q_k^0)\right] \quad [\textit{Continued}]$$

$$+ \frac{1}{2}\sum_k \sum_m f^0_{j,km}(q_k - q^0_k)(q_m - q^0_m) + \ldots \Big] dq_j - \sum_j P_j \delta q_j \quad \ldots\ldots\ldots\ldots (18)$$

where f^0_j, $f^0_{j,k}$, $f^0_{j,km} + \ldots$ are the initial values of the equilibrium forces and their partial derivatives with respect to q_k and q_m. Because the initial state is an equilibrium state, we have $\mathbf{f}^0 = \mathbf{P}$, which means that the first-order work term $\delta W = \Sigma_j P_j \delta q_j$ cancels out.

Let us now assume that $f^0_{j,k}$ are nonzero. Consider that the change of the structure state is isothermal ($dT = 0$). Then, after integrating Eq. 18 up to second-order terms in δq_k and setting $\Sigma_k f^0_{j,k}\delta q_k = \delta f_j$ = equilibrium force changes, we obtain

$$\Delta \bar{\mathcal{F}} = -T\Delta S_{in} = \frac{1}{2}\,\delta\mathbf{f}\cdot\delta\mathbf{q} = \sum_j \frac{1}{2}\,\delta f_j \delta q_j = \sum_j \sum_k \frac{1}{2} K_{jk}(\boldsymbol{\nu})_T \delta q_j \delta q_k = \delta^2 W \quad (19)$$

where $K_{ij}(\boldsymbol{\nu})_T$ is the tangential stiffness matrix associated with $\mathbf{q}$, evaluated on the basis of isothermal material properties ($dT = 0$). This matrix depends in general on the direction vector $\boldsymbol{\nu}$ of the displacements $\delta\mathbf{q}$, as illustrated by Eq. 2 in which the values of ξ and η depend on the ratio Y/X.

From the principle of conservation of energy it further follows that

$$\Delta \bar{\mathcal{F}} = -T\Delta S_{in} = \int_V \frac{1}{2}\,\delta\boldsymbol{\sigma}_T : \delta\boldsymbol{\epsilon}\, dV + \delta^2 W_\sigma = \int_V \frac{1}{2}\,\delta\boldsymbol{\epsilon} : \mathbf{C}_T : \delta\boldsymbol{\epsilon}\, dV + \delta^2 W_\sigma \quad \ldots. (20)$$

Here $\mathbf{C}_T$ = tensor of isothermal tangential moduli of the material, which depends on whether the material loads or unloads; $\delta^2 W_\sigma = \Sigma\Sigma(1/2)\delta q_i G_{km}\delta q_m$; and G_{km} = geometric stiffness matrix, which depends (linearly) on the initial stress σ^0_{ij} (or P_0) but is independent of $\mathbf{C}_T$. In absence of geometrically nonlinear effects ($G_{ij} \approx 0$), Eq. 20 also follows from the principle of virtual work, since $\delta\mathbf{f}$ is in equilibrium with $\delta\boldsymbol{\sigma}$ and $\delta\boldsymbol{\epsilon}$ is compatible with $\delta\mathbf{q}$.

Similarly, for isentropic deformations ($dS = 0$):

$$\Delta \mathcal{U} = -T\Delta S_{in} = \frac{1}{2}\,\delta\mathbf{f}\cdot\delta\mathbf{q} = \sum_j \frac{1}{2}\,\delta f_j \delta q_j = \sum_j \sum_k \frac{1}{2} K_{jk}(\boldsymbol{\nu})_S \delta q_j \delta q_k = \delta^2 W \quad (21)$$

$$\Delta \mathcal{U} = -T\Delta S_{in} = \int_V \frac{1}{2}\,\delta\boldsymbol{\sigma}_S : \delta\boldsymbol{\epsilon}\, dV + \delta^2 W_\sigma = \int_V \frac{1}{2}\,\delta\boldsymbol{\epsilon} : \mathbf{C}_S : \delta\boldsymbol{\epsilon}\, dV + \delta^2 W_\sigma \quad \ldots. (22)$$

Here $K_{ij}(\boldsymbol{\nu})_S$ and $\mathbf{C}_S$ must be evaluated on the basis of adiabatic (equivalent to isentropic) material properties.

According to the second law of thermodynamics, the changes for which $\Delta S_{in} > 0$ will occur and those for which $\Delta S_{in} < 0$ cannot occur. Therefore, the structure is *stable* (i.e. remains in its initial state) if $\Delta S_{in} < 0$ for all possible δq_k or all possible δf_k. The structure is *unstable* if $\Delta S_{in} > 0$ for some δq_k or some δf_k. This criterion, whose essence for fluids was stated already by Gibbs, is the fundamental criterion of stability of the state of any system, including structures, whether reversible (elastic) or irreversible (inelastic). According to Eqs. 19–22, this criterion may also be stated as follows:

The structure is *stable* if the second-order work given by Eqs. 19–20 (for $dT = 0$) or Eqs. 20 and 21 (for $dS = 0$) is positive for all possible δq_k (or all possible δf_k), in other words, if it is positive definite. The structure is *unstable* if this second-order work is negative for some δq_k (or some δf_k).

The case when the second-order work is zero for some δq_k (or some δf_k) represents the critical state, defined as the limit of stable states. Whether the critical state is stable or unstable can be decided only on the basis of the higher-order terms of the Taylor series expansion in Eq. 18, which furnishes higher-order terms to be added to Eqs. 19–22.

If the loads $\mathbf{P}$ are variable, higher-order terms involving $\mathbf{P}$ need to be added to Eq. 18, and consequently also to Eqs. 19–22. The stability criterion $\Delta S_{in} < 0$ remains valid whether loads $\mathbf{P}$ are conservative or nonconservative, but it of course cannot detect dynamic instabilities such as flutter.

The special case of the foregoing criterion of stability of equilibrium states for plastic materials and for boundary conditions of fixed displacements was presented, without thermodynamic derivation, by Hill (1958).

Stable Equilibrium States of Elastoplastic Column

According to Eq. 19 or 21, the condition of stability is that the expressions in Eq. 5 (for load control) or Eq. 6 (for displacement control) must be positive for all possible X and Y (i.e., positive definite). From this we conclude that under load control Shanley's column is stable if $P_0 < P^L_{cr}$ and unstable if $P_0 > P^L_{cr}$ where

$$P^L_{cr} = P_r = \frac{2\xi_u}{\xi_u + 1} P_t \quad \ldots\ldots\ldots\ldots (23)$$

Under displacement control, the column is stable if $P_0 < P^D_{cr}$ and unstable if $P_0 > P^D_{cr}$ where

$$P^D_{cr} = \frac{\xi_u + 1}{2} P_t = \left(1 + \frac{(\xi_u - 1)^2}{4\xi_u}\right) P_r > P_r \quad \ldots\ldots\ldots\ldots (24)$$

For elastic columns, by contrast, $P^L_{cr} = P^D_{cr}$. The physical reason for P^D_{cr} to be higher than P^L_{cr} is that lateral deflection of an elastoplastic column at constant P is accompanied by axial displacement (Eq. 4). Note that P^D_{cr} is the P_0-value for which Eq. 2 with $\delta f_1 = Y = 0$ has a nonzero solution (at $\delta f_2 \neq 0$), while P^L_{cr} is the P_0-value for which Eq. 2 with $\delta f_1 = \delta f_2 = 0$ (P = constant) has a nonzero solution.

The main aspects of the present stability problem can be illustrated by the surfaces in Fig. 3. This figure shows (for $E_u = 3E_t$) three-dimensional views of the surfaces of $\delta^2 W$ or $T\Delta S_{in} = -\delta^2 W$ given by Eq. 5 as functions of X and Y. The equilibrium state is characterized by $\partial\delta^2 W/\partial X = 0$ and $\partial\delta^2 W/\partial Y = 0$, and accordingly, all the surfaces shown have zero slopes at the origin, for any direction.

Eq. 5, which can be written as $\delta^2 W = \Sigma\ 1/2\ K_{jk}\delta q_j \delta q_k$, appears to be a quadratic form but is not, because ξ and η depend on $\delta q_2/\delta q_1$. The surfaces in Fig. 3 consist of quadratic portions separated by the lines $X = \pm Y$ at which ξ or η changes discontinuously. These are lines of curvature discontinuity. Therefore, in contrast to the potential energy surfaces for elastic stability problems (Thompson 1982), the present surfaces are not smooth. They are nevertheless continuous. They also have continuous slopes, because the gradient represents the equilibrium forces ($f_k = \partial W/\partial q_k$). The curvatures, however, must be discontinuous because they represent the incre-

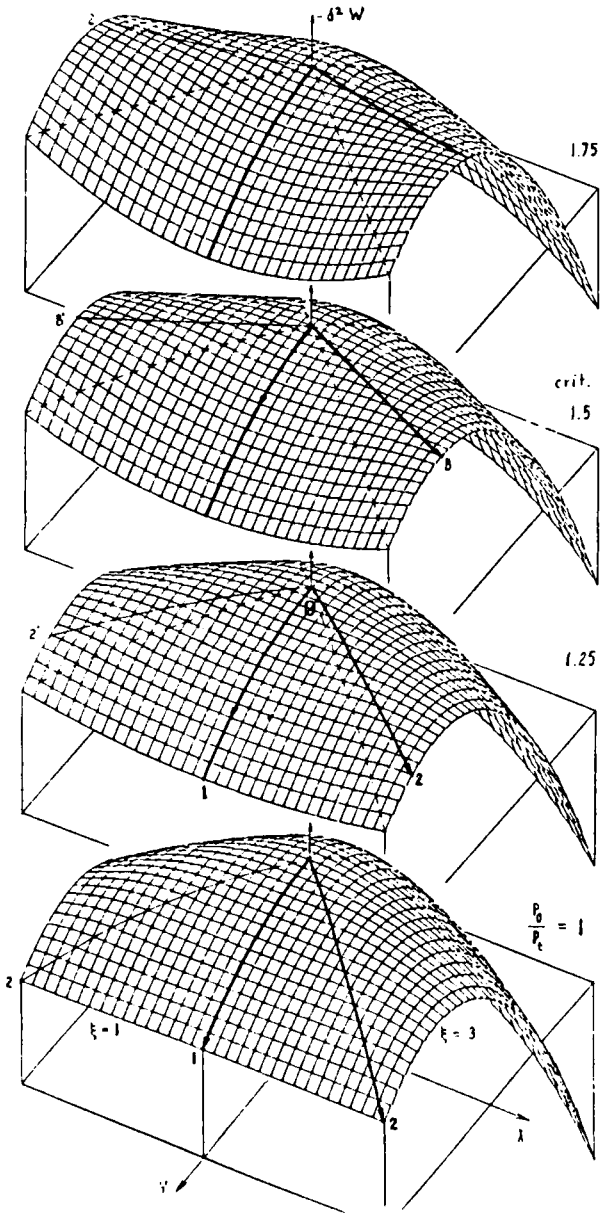


FIG. 3. Surfaces of Internal Entropy Increment $\Delta S_{in} = -\delta^2 W/T$ for Shanley's Column at Various Load Levels

mental stiffnesses of the structure, which change discontinuously (e.g. from E_t to E_u).

Note that the surfaces of $\delta^2 W$ describe path-independent behavior only for those paths for which the path direction at all points of the path belongs to the same sector of ν-directions in the $\mathbf{q}$-space. For other paths, there is path-dependence.

The stack of the surfaces for $P_0/P_t = 1$, 1.25, 1.5 (crit.), 1.75 shows how these surfaces evolve as the load is increased. The limit of stable states ($P_0 = 1.5P_t = P_{cr}$) is manifested on these surfaces by the existence of a horizontal path emanating from the origin 0 (lines B or B' in the figures). Instability is characterized by the existence of a path for which $\delta^2 W$ descends (or $-\delta^2 W$ rises) while moving away from the origin. Absence of any such path ensures stability. Note that for $P_0/P_t = 1.25$ a portion of the surface in Fig. 3 for $Y > |X|$ is a hyperbolic paraboloid, even though the state is stable.

General Criteria of Stable Equilibrium Path

It has often been alleged that the critical load of elastoplastic columns is between P_r and P_t, and P_t has sometimes been called the critical load. As we now see from Eqs. 23 and 24, P_t is not the critical load but merely the lowest bifurcation load. Doesn't this conflict with Shanley's proposition that columns should be designed for the tangent modulus load P_t? Why should perfect columns start to deflect at P_t rather than remaining straight up to $P = P_{cr}$?

To answer these questions, we note that an equilibrium path represents a series of infinitesimal deviations from equilibrium and its restorations. Therefore, we consider, for an arbitrary structure, a *small loading* step along the equilibrium path α ($\alpha = 1$ or 2) which starts at the bifurcation state A (at which the applied load is P_0, and $X = 0$) and ends at another state C on path α [Fig. 2(*b*)]. We decompose this step into two substeps, the first one (I) *away* from the initial equilibrium state A and ending at some intermediate nonequilibrium state B, and the second one (II) *toward* a new equilibrium state, ending on one of the equilibrium paths at state C; see Fig. 2(*b*). (This decomposition was introduced in Bažant (1985), which presented a special case of the analysis that follows.) The displacements or forces which are controlled are denoted as q_m, f_m. For our column, only one displacement or load is controlled, namely q_2 or f_2, i.e. $m = 2$. If q_m is controlled, we consider δq_m to be changed in the first substep [Fig. 2(*b*), left] while f_m are kept constant, which of course destroys equilibrium; δq_m are kept constant during the second substep in which δf_m are allowed to change so as to regain equilibrium. If f_m are controlled, we consider δf_m to be changed in the first substep [Fig. 2(*b*), right] while q_m are frozen (constant), which destroys equilibrium; δf_m are kept constant during the second substep in which δq_m are allowed to change so as to restore equilibrium.

First consider displacement control (i.e., δq_m prescribed) and isothermal conditions ($dT = 0$). Then δq_m are the same for all the paths α but the equilibrium force increments $\delta f_m^{(\alpha)}$ are different. According to Eq. 15, the increment of Helmholtz free energy of the structure over the entire step is, up to second-order terms, $\Delta F = \Sigma(f_m^0 + 1/2\ \delta f_m^{(\alpha)})\delta q_m = \Delta F_I + \Delta F_{II}^{(\alpha)}$. Here $\Delta F_I = \Sigma f_m^0 \delta q_m$ = increment of F over the first substep, which is the same for both paths $\alpha = 1$, 2, and $\Delta F_{II}^{(\alpha)} = \Sigma\ 1/2\ \delta f_m^{(\alpha)} \delta q_m = \delta^2 W^{(\alpha)}$ = increment

of F in the second substep in which q_m are constant, while the forces f_m change by $\delta f_m^{(\alpha)}$ to find new equilibrium on path α; $\delta^2 W^{(\alpha)}$ is the second-order work along path α. (For our column, there is only one term in the sum, $m = 2$, i.e. $\Delta F_{II}^{(\alpha)} = 1/2\ \delta f_2^{(\alpha)} \delta q_2$.) According to Eq. 12 with $dT = 0$, we have for the second substep (in which $\delta \mathbf{q} = 0$) $\Delta F_{II}^{(\alpha)} = -T\Delta S_{in}$. The second law of thermodynamics indicates that the structure will approach the equilibrium state for which ΔS_{in} is maximized (Guggenheim 1959), i.e. $\Delta F_{II}^{(\alpha)}$ is minimized. Hence, the path α which actually occurs is that for which

$$-T\Delta S_{in} = \delta^2 W^{(\alpha)} = \sum_m \frac{1}{2} \delta f_m^{(\alpha)} \delta q_m = \sum_m \sum_j \frac{1}{2} K_{mj}^{(\alpha)} \delta q_m \delta q_j = \min_{(\alpha)} \quad (25)$$

(if q_m is controlled). $K_{mj}^{(\alpha)}$ is the tangential stiffness matrix for path α, which must be based on isothermal incremental material properties.

Second consider displacement control (δq_m prescribed) and isentropic conditions ($dS = 0$). According to Eq. 15, the increment of the total energy of the structure over the entire step is, up to the second-order terms, $\Delta U = \Sigma(f_m^0 + 1/2\ \delta f_m^{(\alpha)}) \delta q_m = \Delta U_I + \Delta U_{II}^{(\alpha)}$. Here $\Delta U_I = \Sigma f_m^0 \delta q_m$, which is the same for both paths α, and $\Delta U_{II}^{(\alpha)} = \Sigma\ 1/2\ \delta f_m^{(\alpha)} \delta q_m = \delta^2 W^{(\alpha)}$ = increment of U in the second substep in which q_m are constant while f_m change by $\delta f_m^{(\alpha)}$ to find new equilibrium on path α. According to Eq. 13 with $dS = 0$, we have for the second substep (in which $\delta \mathbf{q} = 0$) $\Delta U_{II}^{(\alpha)} = -T\Delta S_{in}$. The second law of thermodynamics indicates that on approach to equilibrium $\Delta U_{II}^{(\alpha)}$ must be minimized (Guggenheim 1959). Hence, the path which occurs is again determined by Eq. 25, in which however $K_{mj}^{(\alpha)}$ must be based on isentropic (adiabatic) rather than isothermal material properties.

Third, consider load control and isothermal conditions. The proper thermodynamic function is now Gibbs' free energy, which is defined as $G = F - \mathbf{P} \cdot \mathbf{q}$. From Eq. 12 one obtains

$$dG = -\mathbf{q} \cdot d\mathbf{f} - SdT - TdS_{in} \quad (26)$$

Here $\mathbf{q} \cdot d\mathbf{f}$ represents the complementary work. In terms of equilibrium displacements q_{Ik} calculated from isothermal incremental material properties, we have (according to Eq. 26 for $dT = dS_{in} = 0$) $dG = -\mathbf{q}_I \cdot d\mathbf{f}$. Accordingly we have, for both substeps combined, $\Delta G = -\Sigma(q_m + 1/2\ \delta q_m^{(\alpha)}) \delta f_m = \Delta G_I + \Delta G_{II}$. Here $\Delta G_I = -\Sigma\ q_m \delta f_m$ = increment of G over the first substep, which is the same for both paths $\alpha = 1, 2$, and $\Delta G_{II}^{(\alpha)} = -\Sigma\ 1/2\ \delta f_m^{(\alpha)} \delta q_m = -\delta^2 \bar{W}^{(\alpha)}$ = increment of G over the second substep in which f_m are constant while q_m are allowed to change so as to restore equilibrium by reaching path α; $\delta^2 \bar{W}^{(\alpha)}$ is the second-order complementary work along path α. (For our column, there is only one term in the sum, $m = 2$, i.e. $\Delta F_{II}^{(\alpha)} = -1/2\ \delta f_2^{(\alpha)} \delta q_2$.) According to Eq. 26 with $dT = 0$, we have for the second substep (in which $\delta \mathbf{f} = 0$) $\Delta G_{II}^{(\alpha)} = -T\Delta S_{in}$. Based on the second law of thermodynamics, the approach to new equilibrium must maximize ΔS_{in}, i.e. minimize $\Delta G_{II}^{(\alpha)}$ (Guggenheim 1959). Hence, the path α which actually occurs is that for which

$$T\Delta S_{in} = \delta^2 \bar{W}^{(\alpha)} = \sum_m \frac{1}{2} \delta f_m \delta q_m^{(\alpha)} = \sum_m \sum_j \frac{1}{2} D_{mj}^{(\alpha)} \delta f_m \delta f_j = \max_{(\alpha)} \quad (27)$$

(if f_m is controlled). $D_{mj}^{(\alpha)}$ is the tangential compliance matrix for path α, which must be based on isothermal material properties. Note that, in contrast to Eq. 25, the path label (α) now appears with δq_m rather than δf_m.

Fourth, consider load control and isentropic conditions ($dS = 0$). The proper thermodynamic function is now the enthalpy, which is defined as $H = U - \mathbf{P} \cdot \mathbf{q}$. From Eq. 13 one obtains

$$dH = -\mathbf{q} \cdot d\mathbf{f} + TdS - TdS_{in} \quad (28)$$

In terms of equilibrium displacements $\mathbf{q}_S$ calculated from isentropic (adiabatic) incremental material properties we have (according to Eq. 28 for $dS = dS_{in} = 0$) $dH = -\mathbf{q}_S \cdot d\mathbf{f}$. Thus, for both substeps combined, $\Delta H = -\Sigma(q_n + 1/2\ \delta q_n^{(\alpha)}) \delta f_n = \Delta H_I + \Delta H_{II}$. Here $\Delta H_I = -\Sigma\ q_m \delta f_m$ = increment of enthalpy H over the first substep, which is the same for both paths $\alpha = 1, 2$, and $\Delta H_{II}^{(\alpha)} = -\Sigma\ 1/2\ \delta f_m^{(\alpha)} \delta q_m = -\delta^2 \bar{W}^{(\alpha)}$ = increment of H over the second substep in which f_m are constant while q_m are allowed to change so as to restore equilibrium. According to Eq. 28 with $dS = 0$, we have for the second substep (in which $\delta \mathbf{f} = 0$) $\Delta H_{II}^{(\alpha)} = -T\Delta S_{in}$. In view of the second law, the approach to new equilibrium must maximize ΔS_{in}, i.e. minimize $\Delta H_{II}^{(\alpha)}$. Hence, the path which occurs is again that indicated by Eq. 27, in which however $D_{mj}^{(\alpha)}$ must be based on isentropic rather than isothermal material properties.

Eq. 20 further shows that, along path α,

$$\delta^2 W^{(\alpha)} = \sum_m \frac{1}{2} \delta f_m^{(\alpha)} \delta q_m = \int_V \frac{1}{2} \delta \boldsymbol{\sigma}^{(\alpha)} : \delta \boldsymbol{\epsilon}^{(\alpha)} dV + \sum_m \sum_m \frac{1}{2} \delta \mathbf{q}_k G_{km} \delta \mathbf{q}_m \quad (29)$$

where $\delta \boldsymbol{\sigma}^{(\alpha)}$, $\delta \boldsymbol{\epsilon}^{(\alpha)}$ = stress and strain tensor increments along path α; and G_{km} = geometric stiffness matrix. For $\delta^2 \bar{W}^{(\alpha)}$, a similar expression generally does not exist because matrix G_{km} may be (and often is) singular. But in absence of nonlinear geometric effects ($G_{km} = 0$),

$$\delta^2 \bar{W}^{(\alpha)} = \sum_m \frac{1}{2} \delta f_m \delta q_m^{(\alpha)} - \int_V \frac{1}{2} \delta \boldsymbol{\sigma}^{(\alpha)} : \delta \boldsymbol{\epsilon}^{(\alpha)} dV \quad (30)$$

In finite element element programs, the sums of $\delta f_m \delta q_m$ are usually more efficient to calculate than the volume integrals.

It may also happen that some displacements, $q_i^{(\alpha)}$ ($i = 1, \ldots m$), and some loads, $f_j^{(\alpha)}$ ($j = m + 1, \ldots, n$), are controlled. For such mixed control, Legendre transformation may then be used to introduce a semi-complementary thermodynamic function, $Z = F - \Sigma_j f_j^{(\alpha)} q_j$. Then, for isothermal conditions,

$$dZ = -SdT + TdS_{in} + \Sigma_i f_i^{(\alpha)} dq_i - \Sigma_j q_j^{(\alpha)} df_j \quad (31)$$

A procedure similar to that which led to Eq. 25 (see Bažant 1988) then shows that the stable path is that for which

$$T\Delta S_{in} = \sum_i \frac{1}{2} \delta f_i \delta q_i^{(\alpha)} - \sum_j \frac{1}{2} \delta f_j^{(\alpha)} \delta q_j = \max_{(\alpha)} \quad (32)$$

among all the paths (α). The same result is obtained for isentropic conditions using the mixed thermodynamic function $\bar{Z} = U - \Sigma_j f_j^{(\alpha)} q_j$.

Analysis of Stable Equilibrium Path

Inspecting Eqs. 10 and 11 for Shanley's column, we now find that under displacement control (same Y) we always have $\Delta W^{(2)} < \Delta W^{(1)}$ if $P_0 > P_t$,

and under load control (same δf_2), we always have $\Delta W^{(2)} > \Delta W^{(1)}$ if $P_0 > P_t$. This means that, for $P_0 > P_t$, path 2 must occur and is, therefore, stable, while path 1 cannot occur and is, therefore, unstable. So the column must deflect for $P_0 > P_t$. Shanley's load P_t represents the maximum load of an undeflected column which can be achieved in a continuous loading process, provided E_t varies continuously.

What is then the meaning of the stable states of a perfect column for $P_0 > P_t$? They can be reached if temporary restraints are placed at the sides of the column to prevent it from buckling. The load may then be raised up to some value $P_0 > P_t$. If $P_0 < P_{cr}^D$ at axial displacement control, or if $P_0 < P_{cr}^L$ at axial load control, this column will not deflect when the lateral restraint is removed (provided that the column is perfect, of course). So the column is stable at such a load because the initial state does not change. Deflection occurs only if the load is increased further. If E_t decreases discontinuously (e.g. if the δ-ϵ diagram is bilinear or if the temperature suddenly increases), an undeflected equilibrium state $P_0 > P_t$ can be reached even without lateral restraints.

The equilibrium paths leading away from the origin are marked in Fig. 3 as 1, 2 and 2'. In the plots of ΔW, the structure follows the path that descends steeper with respect to Y (and less steeply for the plots of $-\Delta W$). The limit of stability of the main path is characterized by the fact that points 1 and 2 (of equal Y) are at equal altitude (Fig. 3). This happens on the surfaces for $P_0/P_t = 1$. Instability of the main path is characterized by the fact that point 2 in the plots of $-\delta^2 W$ lies at a higher altitude than point 1 (and a lower altitude for the plots of $\delta^2 W$). (Note also that the states on the main path $\overline{01}$ in Fig. 3 cannot be called metastable because for $P_0 > P_t$ it is not possible to move from point 1 to point 2.)

The static structural stability studies in the literature, even those conducted in the most general sense of catastrophy theory (e.g., Thompson 1982), have so far been confined to elastic structures which possess a potential. For inelastic structures, we have two crucial differences: (1) While the surface of the elastic potential is always smooth, the surface of ΔW is unsmooth, exhibiting lines of curvature discontinuity; and (2) while the elastic potential surface is path-independent, the surface of ΔW applies only to the radial outward paths (i.e., Y/X = constant) and some paths close to them.

As we have seen, the irreversible systems have two striking properties:

1. The first bifurcation point on the equilibrium path of an inelastic structure does not have to represent the limit of stability, i.e., the states on all the branches emanating from the bifurcation point can be stable (which cannot occur in elasticity).
2. Yet at the same time, the stable states on one branch beyond the first bifurcation point cannot be reached by a continuous loading process.

The basic cause for this behavior lies in the irreversibility of inelastic deformation. As illustrated for Shanley's column, after bifurcation at point 1 in Figs. 2(*c*) and (*d*), a subsequent prescribed increment of either axial load P or axial displacement q_2 can occur along two distinct equilibrium paths leading to points 2 and 3 (actually, if buckling to the left is also considered, there is also a third path $\overline{13}'$ but it need not be analyzed since it is symmetric to path $\overline{13}$). This is similar to elastic bifurcation. However, contrary to elastic bifurcation, the structure cannot move along path $\overline{23}$, not even in a non-

equilibrium (dynamic) manner, and cannot reach at point 3 the same values of q_1, q_2 and P. The cause is the irreversibility (path-dependence) of plastic strain, which prohibits reaching the same values of q_1, q_2, P as those reached along path $\overline{12}$.

An elastic structure, though, can move along path $\overline{23}$ in a nonequilibrium manner, and it does reach at point 3 the *same* q_1, q_2 and P values as does path $\overline{12}$. This is dictated by path-independence of elastic deformation. If the structure were elastic (reversible, path-independent) then admissibility of path $\overline{23}$ would cause the potential energy at point 2 to be non-positive definite. For an inelastic structure, on the other hand, the state (q_1, q_2, P) at point 3 [Figs. 2(*c*) and 2(*d*)] cannot manifest itself in the incremental work expression at point 2 since path 2 is kinematically inadmissible. It is for this reason that point 2 in Figs. 2(*c*) and 2(*d*) can be stable for inelastic structures but never for elastic structures.

For the same reason, point 2, even if it is infinitely close to point 1, can be a bifurcation state itself, permitting as the subsequent equilibrium paths both path $\overline{24}$ and path $\overline{25}$ in Figs. 2(*c*) and 2(*d*). Bifurcation states infinitely close to each other, occupying a continuous path (such as $\overline{124}$), are impossible for elastic structures (reversible systems).

The foregoing results force us to broaden the general concept of stability by distinguishing between: (1) A stable equilibrium *state;* and (2) a stable equilibrium *path*.

The stable path is that which (1) Consists entirely of stable states; and (2) maximizes ΔS_i compared to all other parts. So it is a narrower concept than a stable state.

For elastic (reversible) structures, both concepts are equivalent, and so this distinction does not exist. For irreversible systems, however, an equilibrium state can be stable while the equilibrium path on which it lies may be unstable. This stable state cannot, in reality, be reached. For such systems, examination of stable states is obviously insufficient.

Note also that stability of the state is decided on the basis of deviations away from equilibrium, while stability of the path is decided on the basis of approaches toward equilibrium.

The concept of a stable path does not quite fit the general definition of stability of solutions, as stated in the dynamic definition of stability in the sense of Poincaré and Liapunov (Leipholz 1970; Ziegler 1968). If an infinitely small disturbance (such as lateral load f_1) is introduced at the first bifurcation point [point 1 or 2 of Figs. 2(*c*) and 2(*d*)], it does not change path $\overline{124}$ to path $\overline{13}$ or $\overline{125}$; rather it excludes path $\overline{124}$ from the paths $\overline{124}$, $\overline{13}$, $\overline{125}$ which are possible in absence of any disturbance. Thus, instability of a path is not manifested by the creation of a second, distinct path, as a consequence of an infinitely small disturbance. It is manifested by the opposite, namely by the *exclusion* of one or two possible paths.

Criterion of Path Bifurcation

Consider now an arbitrary structure with an n-dimensional column matrix of displacements, $\mathbf{q} = (q_1, \ldots, q_n)$. Let L be the loading-only sector, U the adjacent loading-unloading sectors, and $\mathbf{K}^L$, $\mathbf{K}^U$ the corresponding $\mathbf{K}$. If there are two paths under load control, then $\mathbf{K}^L \delta\mathbf{q}^{(1)} = \delta\mathbf{f}$ and $\mathbf{K}^U \delta\mathbf{q}^{(2)} = \delta\mathbf{f}$ where $\delta\mathbf{f}$ is given. The direction $\nu^{(1)}$ of $\delta_{\mathbf{q}}^{(1)}$ always lies in sector L. Prior to the first

bifurcation, the direction $\nu^{(2)}$ of $\delta\mathbf{q}^{(2)}$ lies outside the corresponding sector U for all possible sectors U, i.e., no path 2 exists. After the first bifurcation, $\nu^{(2)}$ lies within the corresponding sector U at least for one U, and then path 2 exists.

Suppose now that the tangential material properties vary continuously along the loading path. Then the direction $\nu^{(2)}$ should also vary continuously. So, at first bifurcation, the direction $\nu^{(2)}$ must coincide with the boundary of sector L (as illustrated by Shanley's column). But then we must have not only $\mathbf{K}^t\delta\mathbf{q}^{(2)} = \delta\mathbf{f}$ but also $\mathbf{K}^L\delta\mathbf{q}^{(2)} = \delta\mathbf{f}$. Subtracting this from $\mathbf{K}^L\delta\mathbf{q}^{(1)} = \delta\mathbf{f}$ we get $\mathbf{K}^t(\delta\mathbf{q}^{(2)} - \delta\mathbf{q}^{(1)}) = \mathbf{0}$ where $\delta\mathbf{q}^{(2)} \neq \delta\mathbf{q}^{(1)}$. Consequently, the first bifurcation is indicated by singularity of matrix $\mathbf{K}^L$, i.e., by the fact that $\det \mathbf{K}^L = 0$, or that the smallest eigenvalue λ_1 of matrix $\mathbf{K}^L$ vanishes. This is the well-known condition of Hill (1961, 1962), and the solid corresponding to matrix $\mathbf{K}^t$ for which multiple solutions exist is called the linear comparison solid. If, however, the tangential material properties change along the loading path discontinuously, then the first bifurcation occurs when the value of λ_1 jumps from positive to negative, without $\mathbf{K}^L$ ever becoming singular.

The eigenvector $\mathbf{q}^*$ of the singular matrix $\mathbf{K}^L$ at the first bifurcation can lie either inside or outside sector L. If $\mathbf{q}^*$ lies inside L, then there exists path 2 such that $\mathbf{K}^L\delta\mathbf{q}^{(2)} = \mathbf{0}$ where $\delta\mathbf{q}^{(2)} \sim \mathbf{q}^*$. This means that there is neutral equilibrium, which represents the limit point instability (or snap-through). If $\mathbf{q}^*$ lies outside sector L (which is the case for Shanley's column), then $\delta\mathbf{q}^{(2)}$ cannot coincide with $\mathbf{q}^*$ but must lie at the boundary of sector L; then $\mathbf{K}^L\delta\mathbf{q}^{(2)} = \delta\mathbf{f}$ where $\delta\mathbf{f}$ is nonzero. This means that the secondary path at the first bifurcation occurs at increasing load, which represents the Shanley-type bifurcation.

If matrix $\mathbf{K}^t$ has a negative eigenvalue λ_1, we have $(\mathbf{K}^t - \lambda_1\mathbf{I})\mathbf{q}^* = 0$ where $\mathbf{I}$ = unit matrix. It follows that $\delta^2 W \sim \mathbf{q}^{*T}\mathbf{K}^t\mathbf{q} = \mathbf{q}^{*T}\lambda_1\mathbf{I}\mathbf{q} = \lambda_1 \mathbf{q}^{*T}\mathbf{q}^* < 0$. But this does not imply instability of state if the associated eigenvector $\mathbf{q}^*$ lies outside L (which has been ignored in some recent papers). However, the existence of negative λ_1 means that a bifurcation point must have been passed and that the state might not lie on a stable path.

If one displacement, say δq_n, is controlled and $\delta f_1 = \ldots = \delta f_{n-1} = 0$, one may take the foregoing case of load control for which $\mathbf{K}^L$ is singular at the first bifurcation point, and then scale δf_n and δq_n by a common factor so as to make δq_n for both paths mutually equal. Since such a scaling does not change the eigenvalues of $\mathbf{K}^L$, the condition $\det \mathbf{K} = 0$ also characterizes the first bifurcation point under displacement control (provided the tangential properties vary continuously).

Imperfections and Symmetry

The fact that real columns must start to deflect at P_r can be independently proven by analyzing the effect of imperfections. The simplest case of imperfect behavior of Shanley's column can be obtained by assuming that (1) The axial load P is raised up to the value $P = P_0 = P_r$ while $f_1 = 0$; (2) then a certain small lateral load f_1, representing the imperfection, is applied at constant P; and (3) then the axial load P is raised further while keeping f_1 constant. For this loading, and for the case that $\xi = \xi_u$ = constant during deflection (bilinear stress-strain diagram) and $\eta = 1$, Eqs. 2 and 3 yield (after elimination of Y) the solution:

$$P(q_1) = \frac{P_r(\xi_u - 1)h - \frac{1}{2}f_1 l(\xi_u + 1) + 4P_r\xi_u q_1}{(\xi_u - 1)h + 2(\xi_u + 1)q_1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (33)$$

The curves $P(q_1)$ are plotted for various values of the imperfection force in Fig. 2(*a*) as the dashed curves. It may now be easily checked that the limit of the curve $P(q_1)$ for $f_1 \to 0$ is Shanley's equilibrium deflection curve of perfect column 1 which begins at load P_r. Since the deflection curve of an imperfect column converges to the stable path (i.e. path 2) rather than to some unstable path, and since extremely small lateral disturbing loads are inevitable, the column must start to deflect at $P = P_r$, same as we concluded it from our analysis of stability. Similar conclusions have been reached before on the basis of numerical solutions for various other types of imperfections and other types of columns.

The bifurcations that we have illustrated correspond to a breakdown of symmetry. The perfect column has the symmetric choice of deflecting either left or right, but once it has deflected to the right it no longer has the choice of deflecting to the left, i.e., its symmetry has broken down. Structures without symmetry do not exhibit equilibrium path bifurcations. This is for example the case of our column if it is perturbed by a lateral load (Eq. 29). Symmetry of any system can be eliminated by introducing suitable imperfections. Does this render the preceding stability analysis useless? Hardly. Imperfect systems are in general harder to solve than the perfect (symmetric) systems, especially since in principle all the possible imperfections have to be considered. (In Eq. 33 we considered only one type of imperfection.)

Localization of Softening Damage

Path bifurcations frequently occur in structures which are destabilized by material softening (or damage, fracture). Consider a uniaxially loaded (tensioned or compressed) specimen of length $L = 2h$. The value of axial displacement u is controlled [Fig. 4(*a*)] and stress σ is the response. The material is assumed to follow the stress-strain curve $\overline{012345}$ for loading and curves such as $\overline{16}$, $\overline{27}$, $\overline{38}$, $\overline{49}$ for unloading [Fig. 4(*b*)]. For the post-peak strain values, there always exists a multiplicity of solutions. The strain ϵ can either increase uniformly or localize, increasing in one segment of length h (loading) and decreasing in another segment of length $L - h = h$ (unloading) in such a manner that stress σ in both segments remains the same [Fig. 4(*a*)].

We will assume that, due to material inhomogeneities, the strain cannot localize into a segment shorter than h where length h is a material constant. Fig. 4(*c*) shows various possible equilibrium paths of σ versus average strain $\bar{\epsilon} = u/2h$. Path $\overline{1234}$ is the main path ($\alpha = 1$) for which the strain is uniform, i.e., $\epsilon = \bar{\epsilon}$. After the peak point, there exist infinitely many possible bifurcation points, such as point 2 [Fig. 3(*c*)], at which the strain localization begins. The response curves for localizations with unloadings which begin at various points may be constructed by averaging the horizontal distances c and d or e and f as shown in Figs. 4(*b*) and (*c*).

The condition of stable equilibrium of the specimen, which was derived in Bažant (1976; see also Bažant 1985, 1986) and coincides with the general stability condition $\Delta W > 0$ derived here, indicates the limit of stable equilibrium states to be the snapback point. On the path $\overline{01234}$ corresponding

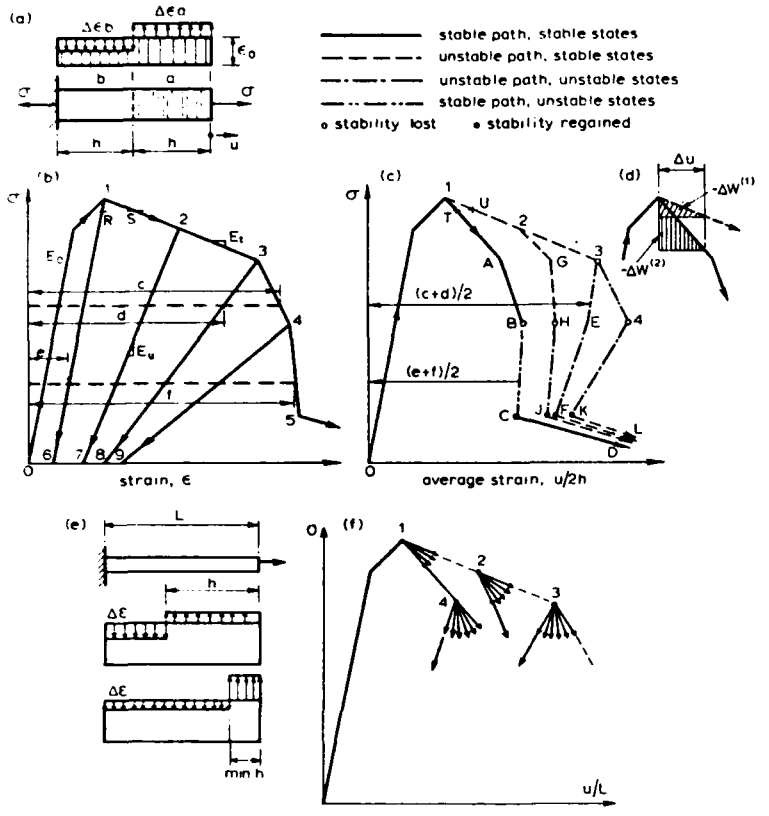


FIG. 4. Bifurcations of Equilibrium Path in Strain-Softening Uniaxial Test Specimen

to uniform strain, it is the point 3. The states of uniform strain on the segment $\overline{123}$ are stable, despite strain-softening (Bažant 1976, 1977). The limit of s ble states (point 3), however, is not the first bifurcation point on the main path. Rather it is the peak point 1. This is similar to the situation for Shanley's column [see the bifurcations in Fig. 2(*a*)]. The limit of stable states, i.e., the snapback point, is obviously analogous to the reduced modulus load. Does the peak stress point (point 1) correspond to Shanley's tangent modulus load? Does the stable path bifurcate at this point?

The answer is yes. This can be proven by considering the second-order work for prescribed Δu, which is represented by the areas $\Delta W^{(1)}$ and $\Delta W^{(2)}$ cross-hatched in Fig. 4(*d*). Their values are negative, and the smaller one corresponds to the stable path, as we have shown. Since $\Delta W^{(2)} < \Delta W^{(1)}$, we must conclude that the stable path bifurcates right at the peak stress point. The specimen, in a continuous loading, must therefore follow path $\overline{01ABCD}$ in Fig. 4(*c*), and not the path $\overline{013E}$. The stable states on the segment $\overline{123}$ are not reachable by a continuous loading process. (They would be reachable, though, if the specimen were temporarily forced to deform uniformly, e.g., by first gluing it to a stiff parallel steel plate, then moving

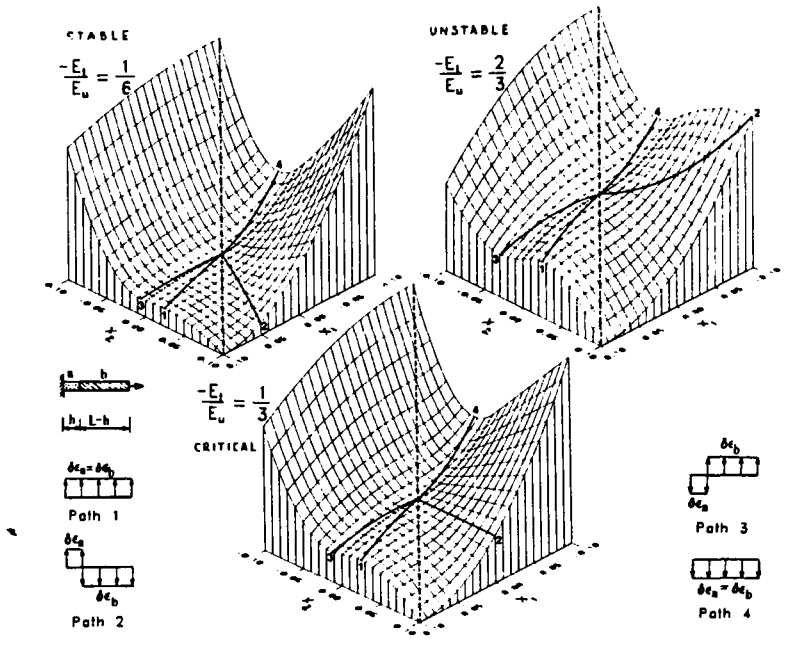


FIG. 5. Surfaces of $\delta^2 W$ for Strain-Softening Uniaxial Specimen ($X_1 = \delta u/L$, $X_2 = \delta v/L$)

the specimen ends and finally dissolving the glue while the ends are kept fixed; the specimen would then remain stable in a state such as point 2.)

To sum up, the thermodynamic criterion of stable path implies that localization of strain begins right after the peak point and the strain always remains localized to the shortest possible segment allowed by the nonlocal characteristics of the material.

The second-order work for the specimen in Fig. 4 is obtained as

$$\delta^2 W = \frac{1}{2} E_u \left[\frac{\xi}{h} \delta u + \frac{\eta}{L-h} (\delta u - \delta v)^2 \right] \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (34)$$

in which $\xi = 1$ for $\delta v \le 0$, $\xi = -k$ for $\delta v > 0$, $\eta = 1$ for $\delta u \le \delta v$, $\eta = -k$ for $\delta u > \delta v$; $k = -E_t/E_u$; u, v = axial displacements at the end and at the interface of segments h and $L - h$. The surfaces of $\delta^2 W$ are plotted in Fig. 5 for $L/h = 4$, and $k = 1/6$, $1/3$, and $2/3$, representing stable, critical and unstable states if v is controlled; 1, 2, 3, 4 label the paths, and the lines of curvature discontinuity are shown dashed. All the foregoing conclusions can be derived from these surfaces. Note that for stability of state at controlled displacement only positive definiteness of the cross section $u = 0$ matters, while for stability of path at controlled displacement the positive definiteness of the entire surface of $\delta^2 W$ matters.

Independently of thermodynamics, our result can also be proven by considering imperfections. Suppose that, for example, a specimen at peak stress state [point 1 in Fig. 4(*c*)] is subjected, at constant u, to a very small dis-

turbing axial load δf applied at the interface between the segments h in Fig. 4(a). This causes one segment to strain-soften to point S and the other segment to unload to point R [Fig. 4(b)]. As a result, the state in Fig. 4(c) moves from point 1 to point T (not to point U). Now a crucial property is that for the subsequent loading there is no bifurcation; the segments h can move in Fig. 4(b) from point S only toward point 2 and from point R only toward point 6, which means that the state in Fig. 4(c) moves from point T toward point A and cannot move toward point 2. Since an infinitesimal disturbance δf is inevitable, the bifurcated path $\overline{1TA}$ must be followed.

The length of the localization zone can in reality have any value between max $h = L$ and min $h = l$ where l is a material constant proportional to characteristic length. For an arbitrary h/L ratio [Fig. 4(e)], the horizontal coordinates of the response curve of σ versus u/L represent a weighted average of distances c and d of Fig. 4(a). As the ratio is varied, the response paths can occupy a continuous fan of directions as shown in Fig. 4(f). So there are infinitely many paths emanating from each bifurcation point, among which the extreme ones correspond to $h = l$ and $h = L$, and there is a continuous sequence of such bifurcation points. Still, however, the states on all these postbifurcation paths can be stable. This is the case, e.g., for points 1 and 2 in Fig. 4(f).

The limits of stable states have recently been determined for flexural softening in beams or frames (Bažant et al. 1987). Bifurcation of the stable path may again occur earlier than these limits.

Multiplicity of response paths that consist entirely of stable states no doubt exists in inelastic finite element systems with damage or softening (Bažant 1976, 1977, 1985, 1986) but often goes undetected. An example was published (Bažant 1985, 1986) where various alternatives to the growth of a strain-softening (cracking) zone permitted two different equilibrium paths of load versus displacement. Step-by-step loading with iterations in each step produced as output only one of these paths (unless the system was forced by a suitable small disturbance to follow the other path). Yet the path that the finite element program produced as output turned out to be impossible in reality since another undetected path was subsequently found to minimize the increments of the Helmholtz free energy (or maximize the values of internal entropy). In spite of this fact, the computations which led to the incorrect path appeared to be stable, the iterations converged well.

Obviously, without actually checking for path stability, the inelastic finite element analysis as currently practiced can be deceptive. It does not necessarily yield the correct response path and may miss the stable one. Further research on numerical implementation of the present criteria is needed.

Structure with Single Load

Determination of the stable path is simple when there is only one applied load, f, or one controlled displacement, q. Noting that the second-order work $(1/2)\delta f \delta q$ is represented by the cross-hatched triangles in Figs. 4 or 6, or that $\delta^2 W = (1/2)K\delta q^2$, $\delta^2 \bar{W} = 1/2\ \delta f^2/K$ where $K = df/dq$, we have the following theorem: If the initial state is stable, *the stable path is that for which the slope K, i.e., the tangential stiffness is minimum.*

Examples proving this simple theorem are shown in Fig. 6. The solid arrows (with the triangles of $\delta^2 W$ or $\delta^2 \bar{W}$) denote the stable paths, and the

dashed arrows the unstable paths. For the states 3, 4, 7 no path is stable because these states are unstable (as revealed by snapback).

Identification of Softening Stress-Strain Relations from Tests

The strain localization due to strain-softening complicates experimental evaluation of constitutive relations. Previous works have generally assumed the strain in small test specimens to remain uniform even after the peak stress. However, this can be assumed only for small specimens whose size is roughly equal to the characteristic length l of the material.

To simplify the material identification and avoid complicated finite element analysis of the test specimen, one may assume the post-peak specimen behavior to be approximately equivalent to a series coupling of a loading zone of volume fraction f which undergoes uniform strain-softening and an unloading zone of volume fraction $1 - f$, which undergoes unloading. The observed (measured) mean strain is $\bar{\epsilon} = f\epsilon + (1 - f)\epsilon_u$ where ϵ = true post-peak strain in the loading (softening) zone while ϵ_u = strain in the unloading

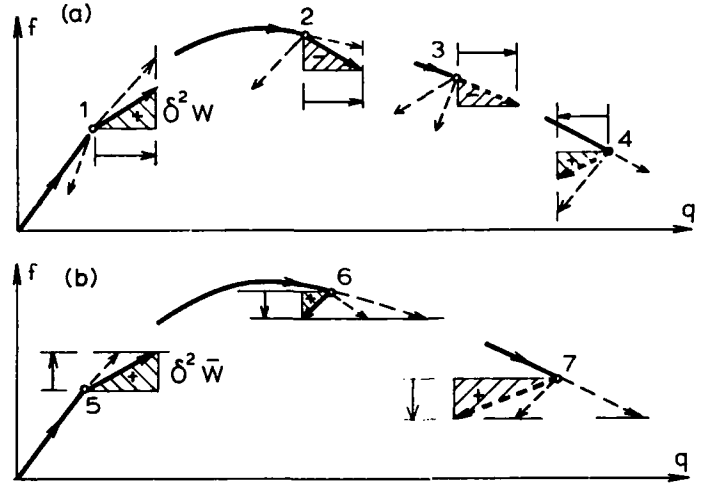


FIG. 6. Stable Post-Bifurcation Paths for Structures with Single Load P; (*a*) for Displacement Control, (*b*) for Load Control

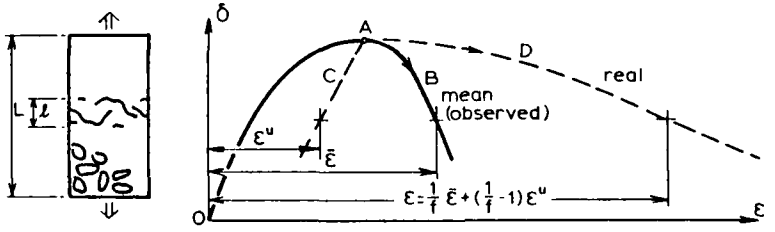


FIG. 7. Simple Approximate Estimation of Real Stress-Strain Relation from Observed Average Stress-Strain Relation

zone, which follows the unloading branch from the peak stress point (Fig. 7). From this,

$$\epsilon = \frac{1}{f}\bar{\epsilon} + \left(\frac{1}{f} - 1\right)\epsilon^u \quad \ldots\ldots\ldots\

Thompson, J. M. T. (1982). *Instabilities and catastrophes in science and engineering*. John Wiley and Sons, New York, N.Y.
Tvergaard, V. (1985). On bifurcation and stability under elastic-plastic deformation." In *Plasticity Today*. A. Sawczuk and G. Bianchi, eds., Elsevier Science Publishers, New York, N.Y.
Ziegler, H. (1968). *Principles of structural stability*. Blaisdell Publishing Co., Waltham, Mass.

Journal of Applied Mechanics,
submitted November 1989.

On the Crucial Role of Imperfections in Quasistatic Viscoplastic Solutions

by

T. Belytschko, B. Moran, M. Kulkarni

Department of Civil Engineering,

Northwestern University

Evanston, Illinois 60208, U.S.A.

Abstract

A closed form solution is developed for shear banding in strain-softening viscoplasticity in one dimension. The stability and structure of the shear band and how they relate to the imperfection is then studied. It is shown that the structure of the band depends on the structure of the imperfection: for continously differentiable imperfections, the shear band is continously differentiable, whereas when the imperfection is C^0 at the maximum, the shear band is C^0, and cusp-shaped. For step function imperfections, the shear band is shown to be a step function, but it is shown that this solution is unstable. Fourier analyses show that the width of the shear band depends directly on the width of the imperfection, suggesting that the imperfection scales the response of the viscoplastic material. New definitions of localization and stability in terms of a response function are also proposed.

material. New definitions of localization and stability in terms of a response function are also proposed.

1. Introduction

Considerable progress has been achieved in the past decade in the development of techniques for the study of material instabilities such as shear banding. Strain softening viscoplastic models have proven to be effective in capturing at least the gross characteristics of shear banding. Molinari and Clifton [1] have shown that viscoplastic material models exhibit what they call L_∞ localization; a definition which has proven useful in predicting conditions that lead to shear banding.

Pan [2] has shown that in the softening range, viscoplastic material models exhibit instability in the sense that Fourier perturbations grow for all wavelengths. It is of interest that this contrasts with the behavior of localization limiters, where perturbations with wavelengths above a threshold which depends on the localization parameter, do not grow unboundedly; Lasry and Belytschko [3]. Wu and Freund [4] using the method of characteristics, have shown that rate-independent softening models exhibit a phenomenon which they term "deformation-trapping" as a consequence of imaginary wave speeds. They then showed that they could obtain a solution numerically by adding rate-dependence (viscoplasticity) and coupled heat transfer. Bazant and Belytschko [5] obtained a closed form solution for rate-independent strain-softening. They showed that the strain softening was limited to a set of measure zero, which explains the difficulties of obtaining such solutions by numerical methods. They also explained that the rate-independent solution is physically unreasonable since the strain-softening, which is intended to model damage, is associated with no energy dissipation, because it occurs over a set of measure zero.

Needleman [6] showed that the ill-posedness of the rate-independent model could be eliminated by the addition of rate-dependence in the form of a viscoplastic model, these findings concur with those of Wu and Freund. Needleman also concluded that the rate-dependent model admits discontinuous solutions which are unique for a given imperfection. However, we will show that solutions associated with step imperfections which are discontinuous in the strain are unstable. Since such solutions are widely used, this is a significant point.

Wright and Batra [7] and Wright and Walter [8] have reported numerical solutions for viscoplastic dynamic problems coupled with heat transfer. Wright and Walter show that the shear strain distribution at the maximum is actually flat over a very small distance; however, extremely fine meshes with a logarithmic distribution about the center of the band and on the order of 10^3 mesh points were needed to achieve such solutions.

One factor which has been glossed over in most of these solutions is the role of the initial imperfections, or perturbations of the initial data. Most of the authors cited have used triangular imperfections or step imperfections. In this paper, we will show via a closed form solution that the imperfection plays a crucial role in the evolution of a quasistatic solution of shear banding.

We will adopt the classical mathematical viewpoint of dynamical systems theory in that an evolution solution is considered stable if a perturbation results in a solution which is in a finite "ball" or "hypersphere" about the unperturbed solution. By using this framework, we show that the homogeneous solution, for a strain-softening viscoplastic material is unstable. Solutions which are step functions with one sub-domain in a strain-softening regime, the other in a non-softening viscoplastic regime, are also shown to be unstable solutions. These solutions cannot be achieved physically and are meaningless.

The morphology of the shear band is shown to depend strongly on the structure of the imperfection. If the imperfection is smooth (C^1), then the shear band has the same character with a continous derivative at its maximum. If the imperfection is C^0 with a discontinuity in its derivative at the maximum, then the shear band also has a cusped structure with discontinuous derivatives at its maximum.

Furthermore we show that imperfections introduce a length scale into the evolving shear band. In particular, the spectrum of the shear band is related to the spectrum of the imperfection. It is possible that the viscoplastic model, when used correctly with imperfections that are representative of those which occur in nature, will reproduce the strain fields observed in experiments.

2. Problem Formulation

We consider a slab of length $2L$ subjected to a shear stress which rises instantaneously to a stress σ^* and then remains constant; see Fig.1. The governing equations are

the equilibrium equation

$$\sigma_{,x} = 0 \tag{1}$$

the strain displacement equation

$$\varepsilon = u_{,x} \tag{2}$$

and the viscoplastic constitutive equation

$$\dot{\sigma} = G\,(\dot{\varepsilon} - \dot{\varepsilon}^{vp}) \tag{3a}$$

$$\dot{\varepsilon}^{vp} = a_o \left(\frac{\sigma}{g(\varepsilon^{vp})} \right)^{\frac{1}{m}} \tag{3b}$$

In the above, σ is the shear stress, ε the corresponding shear strain, G, the elastic shear modulus, g is the viscoplastic response function, and a_o and m are material data, here $0 <$

$m \leq 1$ and $m \to 0$ represents the rate independent limit. Superposed dots denote material time derivatives and commas denote derivatives with respect to the variable which follows.

The boundary conditions are

$$\sigma(L,t) = \sigma^* H(t) \tag{4a}$$

$$u(-L)=0 \tag{4b}$$

where $H(t)$ is the Heaviside step function. Equations (1) and (4a) imply that

$$\dot{\sigma}= 0 \quad \text{for } x \in [\,-L, +L], \; t > 0 \tag{5a}$$

Hence

$$\sigma = \sigma^*, \; t > 0 \tag{5b}$$

$$\dot{\varepsilon} = \dot{\varepsilon}^{vp} \tag{5c}$$

The plastic response function for a softening material is taken to be

$$g(\varepsilon^{vp}) = \sigma_Y \left(\frac{\varepsilon_o + \varepsilon^{vp}}{\varepsilon_o} \right)^{-p} \tag{6a}$$

so that when $\varepsilon^{vp} = 0$, $g(\varepsilon^{vp}) = \sigma_Y$, here $p > 0$ is a material constant, σ_Y is a reference stress that corresponds to the yield stress and $\varepsilon_o = \frac{\sigma_Y}{G}$.

Let

$$e = \varepsilon_o + \varepsilon - \varepsilon^* \tag{6b}$$

where $\varepsilon^* = \frac{\sigma^*}{G}$ is the strain developed in the elastic response.

Using equation (6a), and since the elastic strain rate $\dot{e} = 0$, for $t > 0$, the viscoplastic response function may be written as

$$g(e) = \sigma_Y \left(\frac{e}{\varepsilon_o}\right)^{-p} \tag{6c}$$

This, with equation (3b) and equation (5b), gives

$$\sigma^* = g(e) \left(\frac{\dot{e}}{a_o}\right)^m \tag{7}$$

To represent initial imperfections, the yield stress σ_Y is perturbed as follows

$$\sigma_Y = \sigma_o \, (1 - \mu f(x)) \tag{8}$$

where μ is a small parameter representing the amplitude of the strength imperfection.

Combining equation (5) and equation (7) yields

$$\frac{\sigma}{\sigma_o \alpha(x)} = \left(\frac{e}{\varepsilon_o}\right)^{-p} \left(\frac{\dot{e}}{a_o}\right)^m \tag{9a}$$

where $\alpha(x) = (1 - \mu f(x))$. (9b)

Introduce the following dimensionless variables

$$\bar{t} = a_o t \tag{10a}$$

$$\gamma = \frac{e}{\varepsilon_o} \tag{10b}$$

$$\tau^* = \frac{\sigma^*}{\sigma_o \, \varepsilon_o{}^m} \tag{10c}$$

Equation (9a) may now be written in dimensionless form as

$$\tau^* = \alpha(x) \, \gamma^{-p} \left(\frac{\partial \gamma}{\partial \bar{t}}\right)^m \tag{11}$$

Integrating in time yields

$$\int_0^{\bar{t}} \left\{\frac{\tau^*}{\alpha(x)}\right\}^{\frac{1}{m}} d\zeta = \int_1^{\gamma} \eta^{-\frac{p}{m}} \left(\frac{\partial \eta}{\partial \bar{t}}\right) d\bar{t} \tag{12}$$

where η and ζ are dummy variables. After integration, the following expression is obtained for the strain field

$$\gamma(x,\bar{t}) = \left[\left(\frac{m-p}{m}\right) \left\{\frac{\tau^*}{\alpha(x)}\right\}^{\frac{1}{m}} \bar{t} + 1 \right]^{\frac{m}{m-p}} \tag{13}$$

Unless stated otherwise all results presented subsequently have been obtained using $\tau^* = 1$, $m=0.01$, $p = 0.02$, and all spatial profiles of field variables have been evaluated at $\bar{t} = 0.9$.

Stability of the strain field solution

For a perfect slab, i.e. $\alpha(x) = 1.0$, equation (13) yields the homogeneous solution

$$\gamma(x,\bar{t}) = \left[\left(\frac{m-p}{m}\right) \tau^{*\frac{1}{m}} \bar{t} + 1 \right]^{\frac{m}{m-p}} \tag{14}$$

However, this solution is unstable. Stability requires that if the initial data is perturbed by a small amount, then the solutions must differ by a small amount. To be specific, if solutions with initial data $\alpha_1(x)$ and $\alpha_2(x)$ are $u_1(x,\bar{t})$ and $u_2(x,\bar{t})$, then if

$$\| \alpha_1(x) - \alpha_2(x) \| < \mu \tag{15}$$

then $u_1(x)$ is a stable solution if

$$\| u_1(x,\bar{t}) - u_2(x,\bar{t}) \| < C\mu \quad \forall \bar{t} \tag{16}$$

where C is a constant. If

$$\alpha_1(x) = 1.0 \tag{17a}$$

$$\alpha_2(x) = 1 - \mu \cos\frac{\pi x}{L} \tag{17b}$$

then $\gamma_1(x,\bar{t})$ is a homogeneous solution given by equation (14), whereas $\gamma_2(x,\bar{t})$ is shown in Fig.2 ; the numerically integrated displacement fields $u_1(x,\bar{t})$ and $u_2(x,\bar{t})$ are shown in Fig.3. It can be seen that the two displacement fields do not satisfy (16), so the homogeneous solution $u_1(x,\bar{t})$ is unstable.

The strain-softening viscoplastic material has the idiosyncrasy that even for two perturbations which satisfy (15), the displacement fields may vary markedly. Thus, if we consider $\alpha_1(x)$ and $\alpha_2(x)$ as shown in Fig.4, the solutions, shown in Fig.5, do not satisfy the stability condition (16). It appears that once a perturbation is defined, its stability should be examined by considering small perturbations of this initial perturbation i.e. by considering $\alpha_3(x)$ where

$$\| \alpha_2(x) - \alpha_3(x) \| < \mu^2 \tag{18}$$

Another approach is to use a response function approach to stability, described subsequently.

A commonly used perturbation in the literature is the step function

$$\alpha(x) = 1.0 - \mu(H(x-b) - H(x+b)) \tag{19}$$

The solution corresponding to this perturbation is shown in Fig.6. This solution has the interesting property that even under the restriction (18), it is unstable. To illustrate, we consider the perturbation $\alpha_3(x)$ shown in Fig.7a. The solutions $u_2(x,\tilde{t})$ and $u_3(x,\tilde{t})$ do not satisfy (16), see Fig.7c, so $u_2(x,\tilde{t})$ is an unstable solution. Step-function imperfections are tacitly used in most finite element solutions. This analysis shows that the resulting solution is not stable or physically meaningful at least for this constitutive model.

Stability in terms of a Response Function.

A feature of the viscoplastic constitutive model under consideration here is that localization is triggered at the point at which $\alpha_1(x)$ is a minimum. As shown in Fig.5, the shear band occurs at the location of the minimum of $\alpha_1(x)$, which is quite random in a real body. Thus for an arbitrary, small imperfection, the location of the band may differ markedly. This randomness is in accord with experimental observations; unless notches or other tailored geometric imperfections are used to trigger localization, the location of the

shear band is quite unpredictable. However the overall response only depends on the occurrence of the band, not its location. Thus for engineering purposes, it appears worthwhile to consider an alternative definition of stability in terms of a response function, the appropriate choice of which will depend on the problem.

This definition of stability is stated as follows. Given two imperfections $\alpha_1(x)$ and $\alpha_2(x)$ satisfying (15), then if

$$\| r_1(\bar{t}) - r_2(\bar{t}) \| \leq C_0 \mu \tag{20}$$

where $r(\bar{t})$ is the response function, the solution is stable. For the particular problem considered here, the response function may be taken to be

$$r(t) = u(L,t),$$

so stability requires that

$$\| u_1(L,t) - u_2(L,t) \| = \left\{ \int_0^{\bar{t}} [u_1(L,t) - u_2(L,t)]^2 dt \right\}^{\frac{1}{2}} < C_0 \mu \tag{21}$$

Note that neither the homogeneous solution (14) nor the solution with the step function imperfection (19) meets this stability criterion. On the other hand, the solutions corresponding to the perturbations $\alpha_1(x)$ and $\alpha_2(x)$ as defined in Fig.4 meet this criterion, see Fig.5c.

The appropriate response function depends on the problem. In a problem with a prescribed displacement, the response function would be the resulting load. For engineering purposes, a solution which is stable in terms of a response function should be an adequate solution.

2.1 The Spatial Distribution of the strain field

In this section we characterize the effect of the initial imperfections on the spatial distribution of the strain field.

Let

$$\frac{m}{m-p} = n \tag{22a}$$

$$\left(\frac{m-p}{m}\right) \tau^{*\frac{1}{m}} = \xi \tag{22b}$$

Using the above constants, the strain field $\gamma(x,\bar{t})$ may be cast in the form

$$\gamma(x,\bar{t}) = \left[1 + \xi\, \bar{t}\, \alpha^{-\frac{1}{m}}\right]^n \tag{23}$$

It can be seen from the above, that if $\alpha(x)$ is a C^1 function, then $\gamma(x)$ is also a C^1 function. Thus for $\alpha(x) \in C^1$, the maximum of $\gamma(x,\bar{t})$ occurs where $\frac{\partial \gamma}{\partial x} = 0$

The first and second spatial derivatives of the strain field are given by

$$\frac{\partial \gamma}{\partial x} = -\left(\frac{\xi n\, \bar{t}}{m}\left(1 + \xi\, \bar{t}\, \alpha^{-\frac{1}{m}}\right)^{n-1} \alpha^{-\frac{1+m}{m}}\right)\frac{\partial \alpha}{\partial x} \tag{24}$$

$$\frac{\partial^2 \gamma}{\partial x^2} = \left(-\frac{\xi n\, \bar{t}}{m}\right)\beta\left(\frac{\partial^2 \alpha}{\partial x^2}\right) + \left(-\frac{\xi n\, \bar{t}}{m}\right)\left(\frac{\partial \beta}{\partial x}\right)\left(\frac{\partial \alpha}{\partial x}\right) \tag{25}$$

where

$$\beta = \left\{1 + \xi\, \bar{t}\, \alpha^{-\frac{1}{m}}\right\}^{n-1} \alpha^{-\frac{1+m}{m}}$$

From the above expressions it is clear that the derivative of the strain field vanishes only when the derivative of the imperfection function vanishes, i.e, $\frac{\partial \gamma}{\partial x} \Leftrightarrow \frac{\partial \alpha}{\partial x} = 0$. If $\alpha(x)$ is a C^0 or C^1, function, then $\gamma(x)$ is a C^0 or C^1 function respectively and $\gamma(x,\bar{t})$ is maximum

where $\alpha(x)$ is min..mum, see Fig.8 and Fig.9. If $\alpha(x)$ is a C^0 function, with a cusp-shaped maximum at x_0, then $\frac{d\alpha}{dx}(x_0-\Delta x) < 0$ and $\frac{d\alpha}{dx}(x_0+\Delta x) > 0$ for $\Delta x > 0$ i.e. the left-hand derivative is negative and the right-hand derivative is positive. Equation (24) then implies that the right spatial derivative of $\gamma(x,\bar{t})$ is positive and the left spatial derivative is negative since the quantities in the parentheses are positive. Thus the cusp-shaped spatial distribution for $\gamma(x,\bar{t})$ occurs only when $\alpha(x)$ is a C^0 function rather than C^1. Since many investigators have used triangular imperfections, the prevalence of cusp shaped strain-distributions reported in the literature is thus not unexpected.

In summary

a) The relative maximum of the strain field occurs where $\alpha(x)$ is a relative minimum.

b) Wherever the imperfection function has a constant magnitude, the strain field is constant, this further implies that if the imperfection function is a step function, then the strain field also contains a step function; see Fig.6.

3. Derivation of the Conditions For Localization

The concept of L_∞ localization proposed by Molinari and Clifton [1] provides a useful framework to examine the suitability of mathematical models to describe localization in one dimension. However the L_∞ definition of localization suffers from a drawback in that localization is defined on the basis of the behavior of the strain field at infinite time and not in terms of the current state variables. Since for a stress boundary condition the strain field tends to infinity in finite time, the L_∞ definition for this case is inadequate, though it does usefully describe the behavior of the strain field accurately for a velocity boundary condition. We propose a definition that accurately describes the behavior of the strain field for both velocity and stress boundary conditions.

Localization is defined to occur when the ratio of the maximum strain to the strains at all surrounding points is a convex upward function of time. This may alternatively be stated in terms of the peak strain as follows. The strain field is said to exhibit localization if the peak strain is a convex upward function of time. see Fig.10.

To determine necessary and sufficient conditions for localization, the strain field is expanded in a Taylor series about the point of maximum strain $\bar{x}$, and the behavior of $f(t) = \frac{\gamma(\bar{x},\bar{t})}{\gamma(\bar{x}+\Delta x,\bar{t})}$ is examined. Expanding in a Taylor Series about $\bar{x}$, where $\bar{x}$ is the point of maximum strain

$$\gamma(\bar{x}+\Delta x,\bar{t}) = \gamma(\bar{x},\bar{t}) + \frac{1}{2}\frac{\partial\gamma(\bar{x},\bar{t})}{\partial x}\Delta x + \frac{1}{2}\frac{\partial^2\gamma(\bar{x},\bar{t})}{\partial x^2}\Delta x^2 + \ldots \tag{26}$$

If the imperfection function is assumed to be C^1 then

$$\frac{\partial\alpha(\bar{x})}{\partial x} = 0$$

$$\frac{\partial\gamma(\bar{x})}{\partial x} = 0$$

$$\frac{\partial^2\alpha(\bar{x})}{\partial x^2} > 0$$

$$\frac{\partial^2\gamma(\bar{x})}{\partial x^2} = \left(-\frac{n\ \xi\ \bar{t}}{m}\right)\beta\frac{\partial^2\alpha(\bar{x})}{\partial x^2}$$

$$\gamma(\bar{x}+\Delta x,\bar{t}) = \gamma(\bar{x},\bar{t}) + \frac{1}{2}\left(-\frac{\xi\ n\ \bar{t}}{m}\right)\left\{1+\xi\ \bar{t}\ \alpha^{-\frac{1}{m}}\right\}^{n-1}\alpha^{-\frac{1+m}{m}}\frac{\partial^2\alpha(\bar{x})}{\partial x^2}\Delta x^2 \tag{27}$$

Dividing the above expansion for $\gamma(\bar{x}+\Delta x,\bar{t})$ by

$$\gamma(x,\bar{t}) = \left[1 + \xi\,\bar{t}\,\alpha^{-\frac{1}{m}}\right]^n$$

yields the following

$$\frac{\gamma(\bar{x}+\Delta x,\bar{t})}{\gamma(\bar{x},\bar{t})} = 1 - \frac{\phi(\Delta x)\,\bar{t}}{\left(1 + \xi\,\bar{t}\,\alpha^{-\frac{1}{m}}\right)} \tag{28}$$

where

$$\phi(\Delta x) = \frac{1}{2}\left(\frac{\xi\, n}{m}\,\alpha^{-\frac{1+m}{m}}\frac{\partial^2\alpha}{\partial x^2}\right)\Bigg|_{\bar{x}}\Delta x^2 \tag{29a}$$

Let

$$f(\Delta x,\bar{t}) = \frac{\gamma(\bar{x},\bar{t})}{\gamma(\bar{x}+\Delta x,\bar{t})} \tag{29b}$$

It may be noted that $\xi n > 0,\ \frac{\partial^2\alpha(\bar{x})}{\partial x^2} > 0,\ \alpha(\bar{x}) > 0,\ \Delta x^2 > 0,\ \Rightarrow\ \phi(\Delta x) > 0$

Equation (28) may then be written as

$$\frac{1}{f(\Delta x,\bar{t})} = 1 - \frac{\phi(\Delta x)\,\bar{t}}{\left(1 + \xi\,\bar{t}\,\alpha^{-\frac{1}{m}}\right)} \tag{29c}$$

$$\frac{\partial}{\partial\bar{t}}\left(\frac{1}{f}\right) = -\frac{\phi(\Delta x)}{\left(1 + \xi\,\bar{t}\,\alpha^{-\frac{1}{m}}\right)^2} \tag{29d}$$

$$\frac{\partial f}{\partial \bar{t}} = f^2 \frac{\phi(\Delta x)}{\left(1 + \xi\, \bar{t}\, \alpha^{-\frac{1}{m}}\right)^2} \tag{29e}$$

Irrespective of the sign of ξ, $\frac{\partial f}{\partial \bar{t}} > 0 \Rightarrow f(\Delta x, \bar{t})$ is a monotonically increasing function of time. For $\xi > 0$, $\frac{\partial f}{\partial \bar{t}}$ is not monotonically increasing with time for all Δx. For $\xi < 0$, $\frac{\partial f}{\partial \bar{t}}$ is monotonically increasing with time for all Δx, i.e. $f(\Delta x, \bar{t})$ is a convex upward function of time.

Thus a necessary and sufficient condition for localization as defined above, is, $\xi < 0$, i.e. $m < p$. It may be noted that $p < 0 \Rightarrow \xi > 0$, and therefore a hardening law does not exhibit localization. Fig.10 illustrates the behavior of $f(\Delta x, \bar{t})$ for the cases when $m < p$ and $m > p$. Fig.11 shows that the peak strain exhibits the same behavior as $f(\Delta x, \bar{t})$.

4. Scaling effect of the imperfection

For a step function perturbation, the width of the region of intense straining is set by the width of the imperfection, i.e. the width of the shear band is equal to the width of the imperfection. But as shown previously, this is an unstable solution which is not physically meaningful. Given the absence of a length scale in viscoplastic material models for quasistatic loading conditions, it has been of interest as to what determines the width of shear bands. To better understand the effect of of C^1 imperfections, the strain field given in (13) is expanded in a Fourier series and the effect of the width of the perturbation on the spectrum is studied. For the case of a parabolic imperfection it was found that the amplitude of the dominant mode was directly proportional to the width of the initial perturbation. From Fig.12 and Fig.13 it can be seen that the width of of the initial perturbation scales the amplitude of the dominant mode of the Fourier spectrum of the strain field.

Conclusions

A closed form solution for static shear banding in a viscoplastic one-dimensional problem has been presented. The solutions exhibit a strong dependence on the morphology of the initial imperfection. In particular:

1. In a homogeneous stress field, the shear band occurs at an extremum of the imperfection, so the position of the shear band is quite random.

2. The continuity of the strain field is identical to the continuity of the imperfection, i.e. for a C^1 imperfection, the shear strain field is C^1.

3. When the imperfection is C^{-1}, i.e. a step function, the strain field is is also C^{-1}. However, it has been shown that this solution is clearly unstable. This finding is of relevance in finite element solutions with constant strain elements, where the imperfections used are step functions, when, as is usually the case, the shear band is of the same size as the elements. Futhermore, when the imperfection is a step function, the morphology of the strain field differs significantly from that which arises from more realistic imperfections. Thus solutions based on step function imperfections cannot reveal the details of the strain field which are needed to understand material behavior at high strains.

Shear banding presents an interesting problem in stability. Even two imperfections which are close to zero can lead to strikingly different solutions. Therefore a definition of stability in terms of a response function has been proposed. This definition is particularly useful for engineering purposes since it overcomes difficulties which arise from random locations of shear bands.

A new definition of the localization process has also been proposed. This definition though similar to Molinari and Clifton's [1], L_∞ localization, circumvents some of its difficulties by avoiding infinite values of time or strain ratio.

Acknowledgement

The support of the Air Force Office of Scientific Research under Award F49620-87-C-0030 to Northwestern University is gratefully acknowledged.

References

[1] Molinari, A. and Clifton, R.J. (1987). "Analytical Characterization of Shear Localization in Thermoviscoplastic Materials.", *J. App. Mech.*, ASME, **54**, 806-812

[2] Pan, J. (1983). "Perturbation Analysis of Shear Strain Localization in Rate Sensitive Materials." *Int. J. Solids and Structures*, **19**, 153-164

[3] Lasry, D. and Belytschko, T. (1988). "Localization Limiters in Transient Problems." *Int. J. Solids and Structures*, **24**, 581-597

[4] Wu, F.H. and Freund, L.B. (1988). "Deformation Trapping due to Thermoplastic Instability in One-Dimensional Wave Propagation" *J . Mech. Phys.Solids*, **32**, 119-132

[5] Bazant, Z.P. and Belytschko, T. (1985) "Wave Propagation In a Strain Softening Bar : Exact Solution" *J . Engrg.Mech., ASCE,* **111**, 381-389

[6] Needleman, A. (1988) "Material Rate Dependence and Mesh Sensitivity in Localization Problems" *Comp. Meth. App.Mech. Engrg.*, **67**, 61-85

[7] Wright , T.W. and Batra, R.C. (1985) "The Initiation and Growth of Adiabatic Shear Bands" *Int. J. Plasticity.*, 1,205-212

[8] Wright , T.W. and Walter, J.W. (1987) "On Stress Collapse in Adiabatic Shear Bands" *J. Mech. Phys. Solids,* **35, 6,** 205-212

List of Figure Captions

Fig.1. A slab of viscoplastic material subjected to an instantaneous step stress.

Fig.2. Strain field response to the perturbation represented by equation (17b).

Fig.3. Displacement versus position.

Fig.4a. Initial imperfection to illustrate the random location of shear bands.

Fig.4b. Initial imperfection to illustrate the random location of shear bands.

Fig.5a. Strain field in response to imperfection shown in Fig.4a.

Fig.5b. Strain field in response to imperfection shown in Fig.4b.

Fig.5c. Displacement field in response to perturbations in Fig.4a and Fig.4b.

Fig.6. Strain versus position for a step function perturbation.

Fig.7a. Imperfection to illustrate the instability of step function perturbations.

Fig.7b. Strain field response to perturbation shown in Fig.7a.

Fig.7c. Displacement versus position for the imperfection shown in Fig.7a.

Fig.8 Strain field for a triangular imperfection.

Fig.9. Strain field for a parabolic imperfection.

Fig.10. Behavior of the ratio of peak strain to the strain at surrounding points as a function of time.

Fig.11. Peak strain as a function of time.

Fig.12. Effect of the width of the imperfection on the Fourier spectrum of the strain field.

Fig.13 Effect of width of imperfection on the dominant mode of the Fourier spectrum of the strain field.

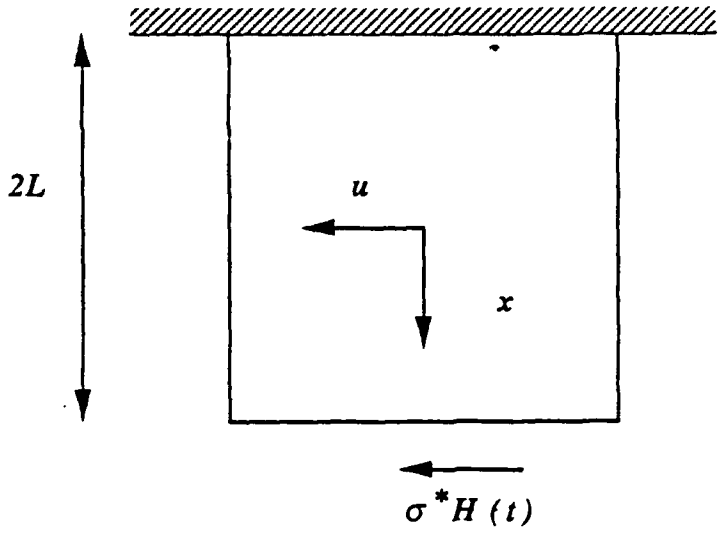


Fig.1. A slab of viscoplastic material subjected to an instantaneous step stress.

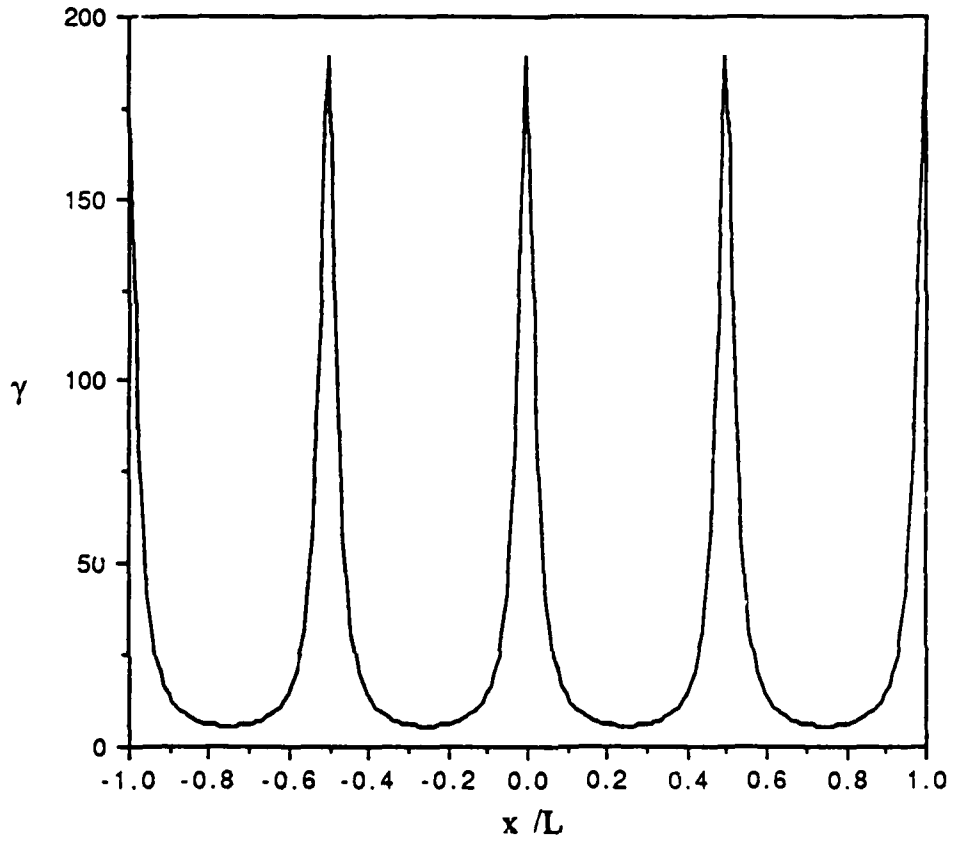


Fig. 2. Strain field response to the perturbation represented by equation (17b).

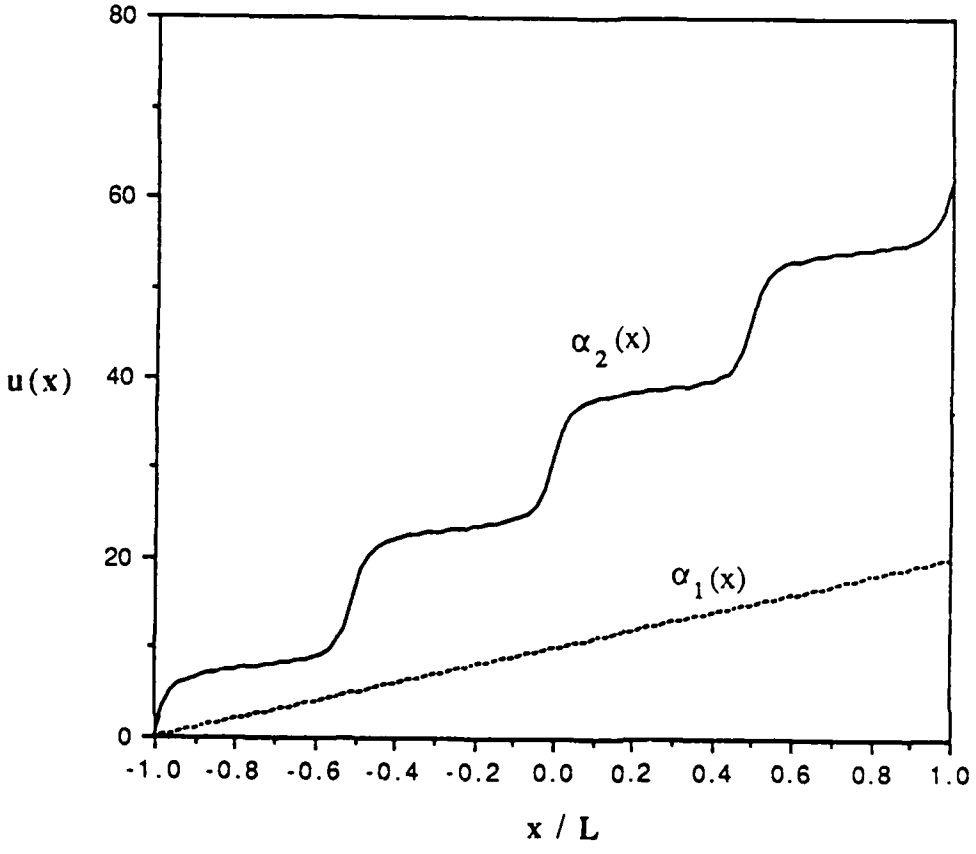


Fig. 3. Displacement versus position.

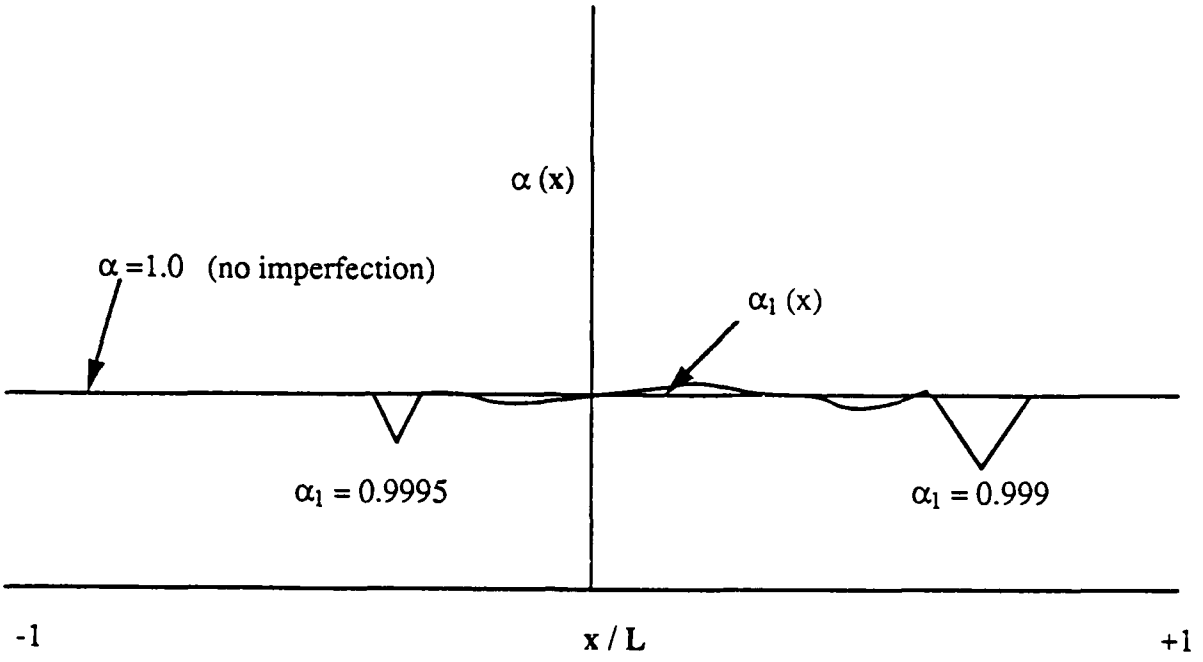


Fig. 4a. Initial imperfection to illustrate the random location of shear bands.

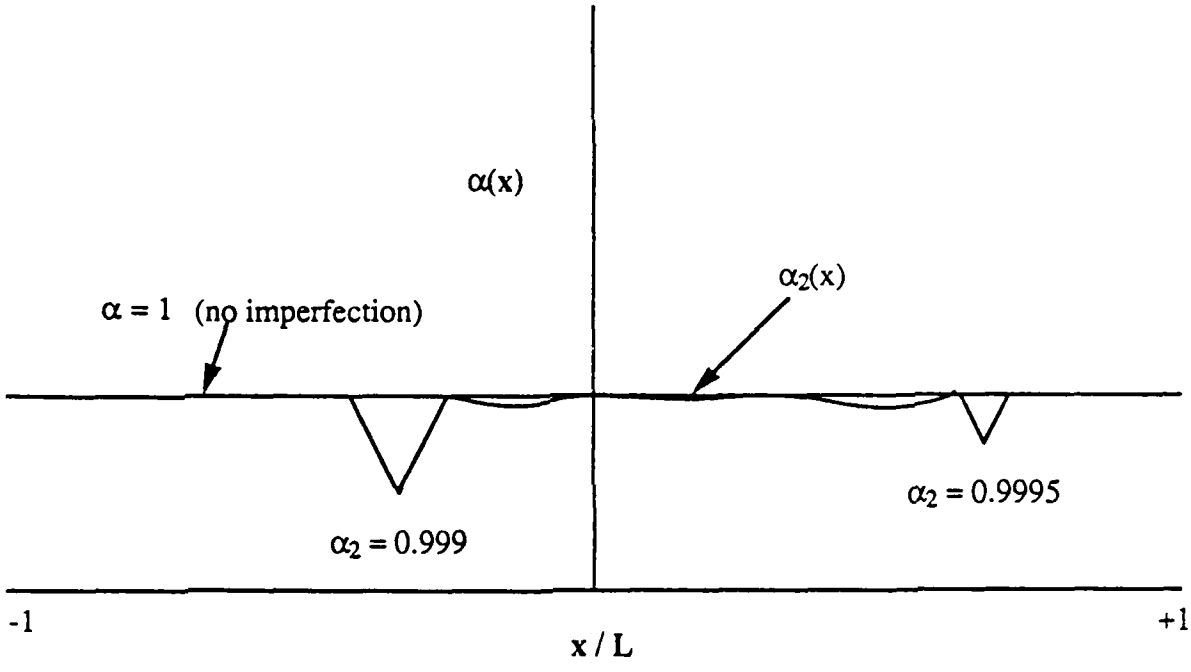


Fig. 4b. Initial imperfection to illustrate the random location of shear bands.

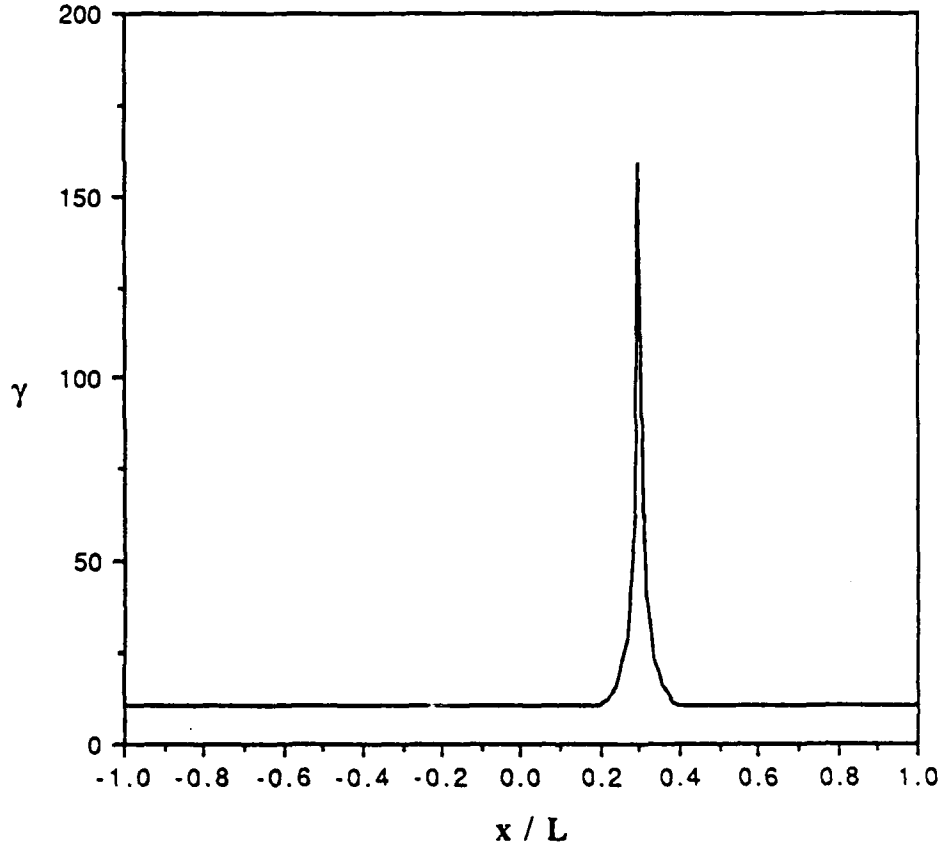


Fig. 5a. Strain field in response to imperfection shown in Fig. 4a.

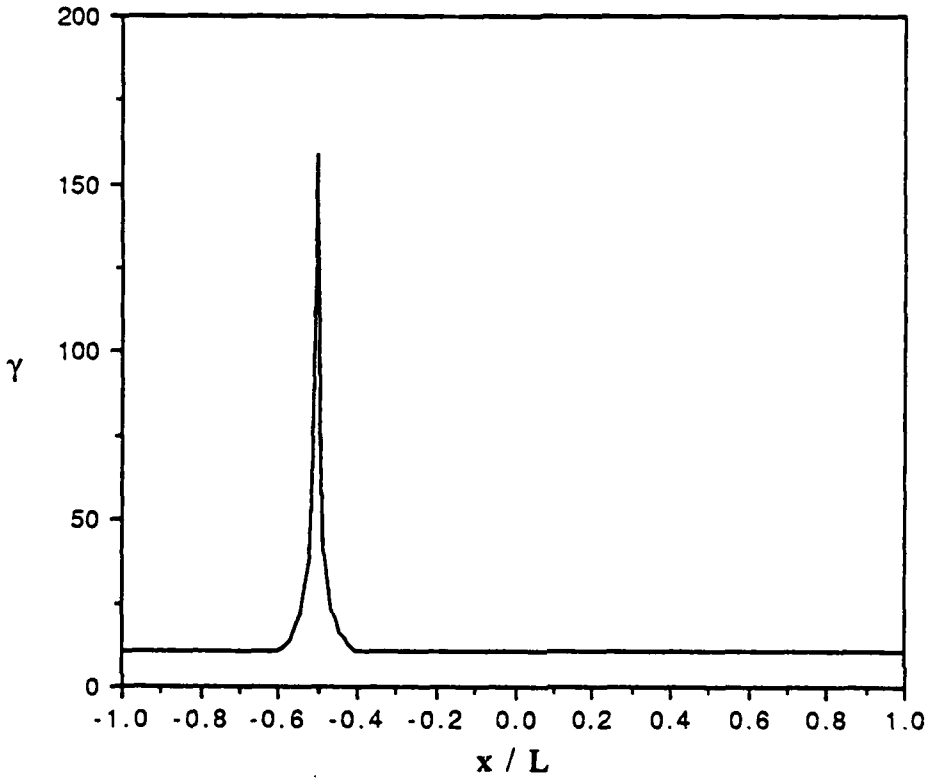


Fig. 5b. Strain field in response to the imperfection shown in Fig. 4b.

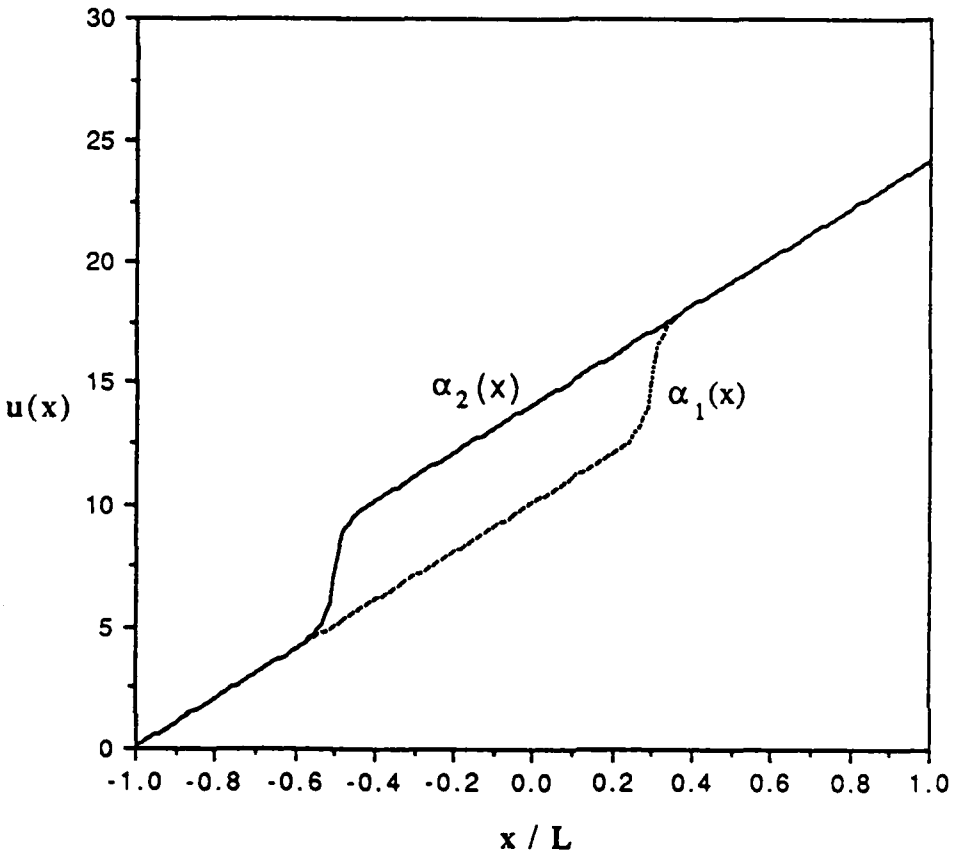


Fig. 5c. Displacement field in response to perurbations in Fig. 4a and Fig. 4b.

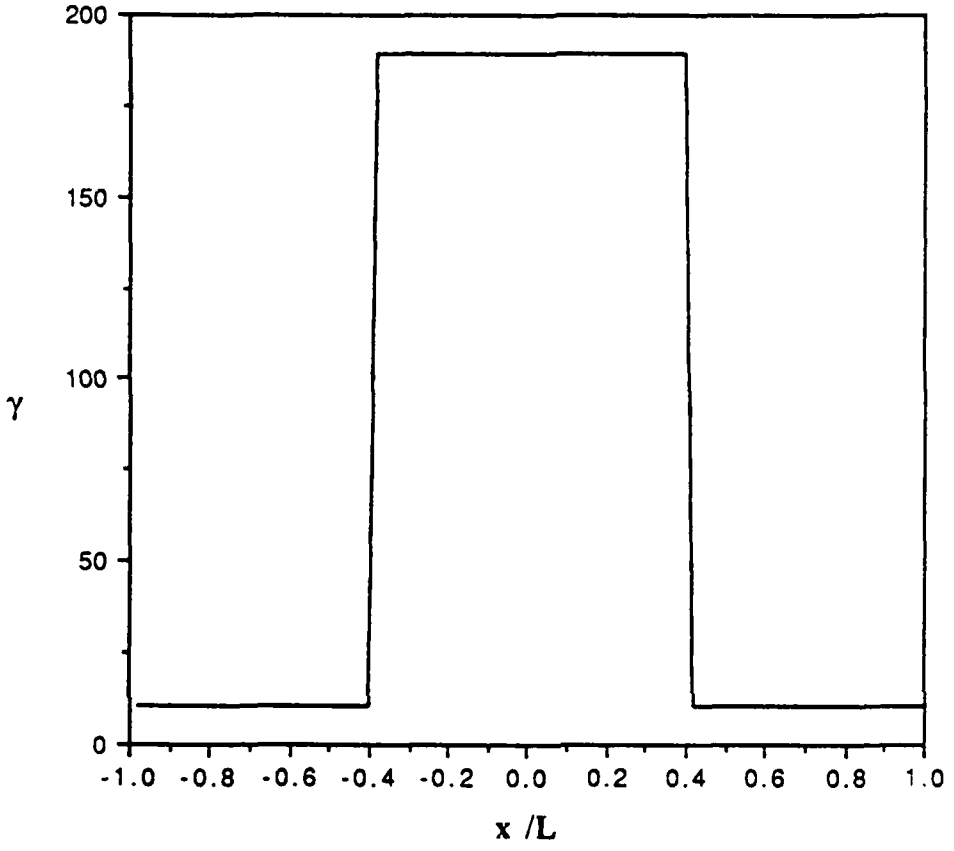


Fig. 6. Strain vs position for a step function perturbation.

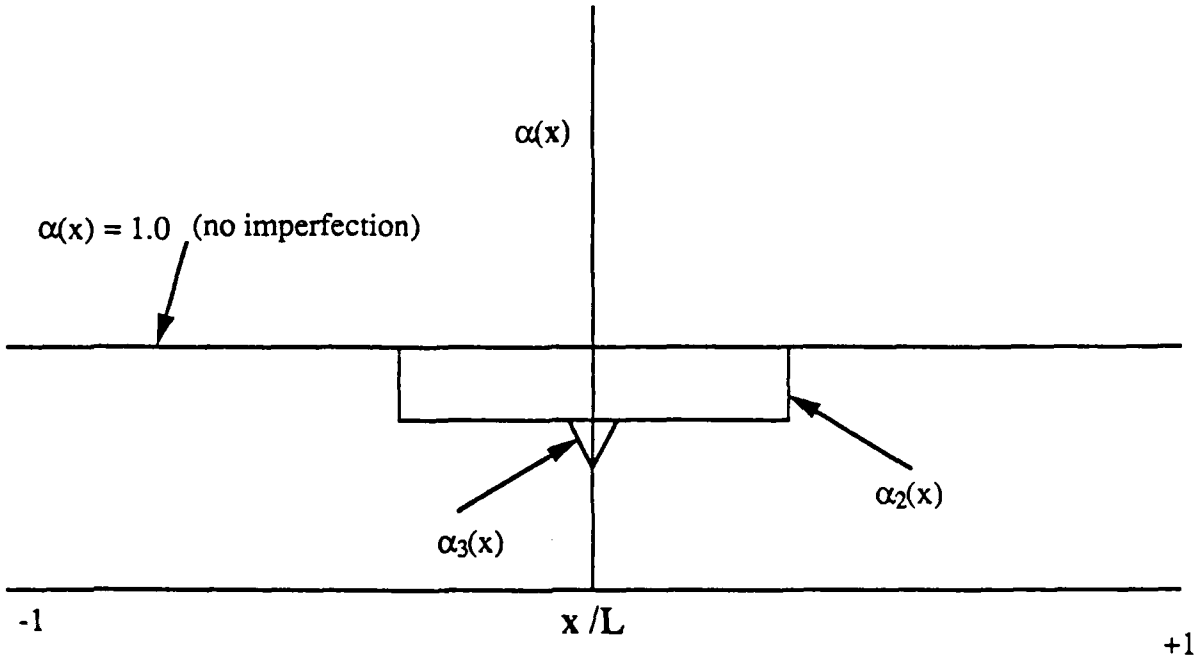


Fig. 7a. Imperfection to illustrate the instability of step function perturbations.

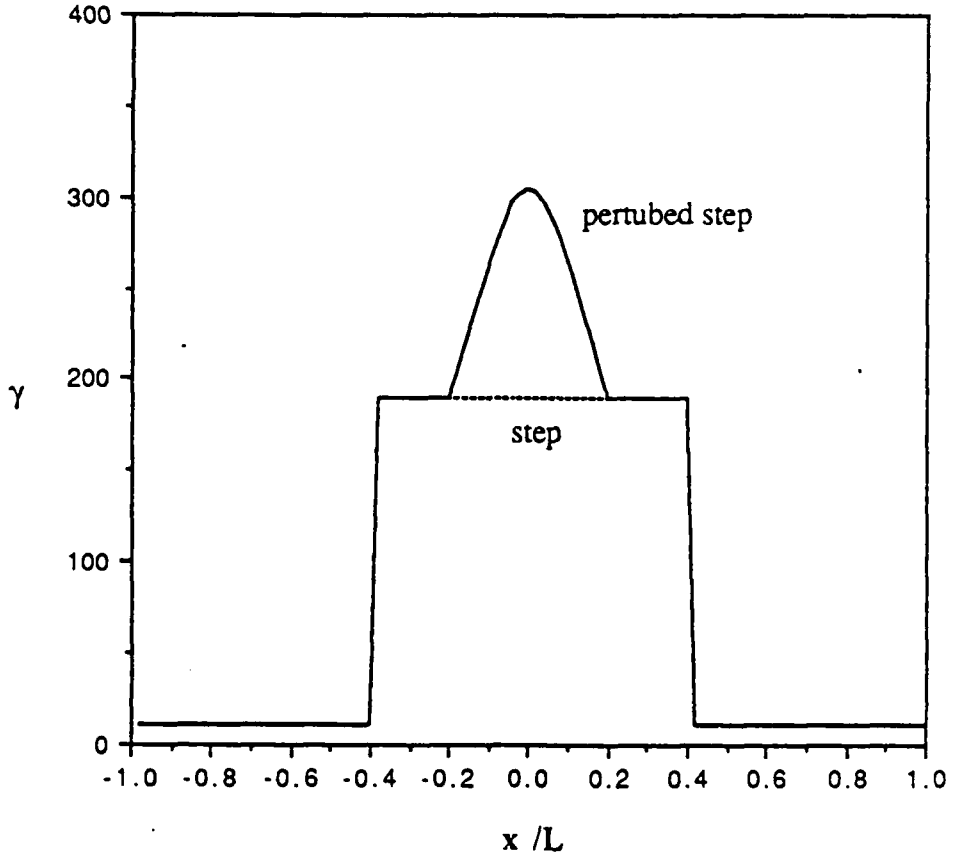


Fig. 7b. Strain field response to perturbation shown in Fig. 7a.

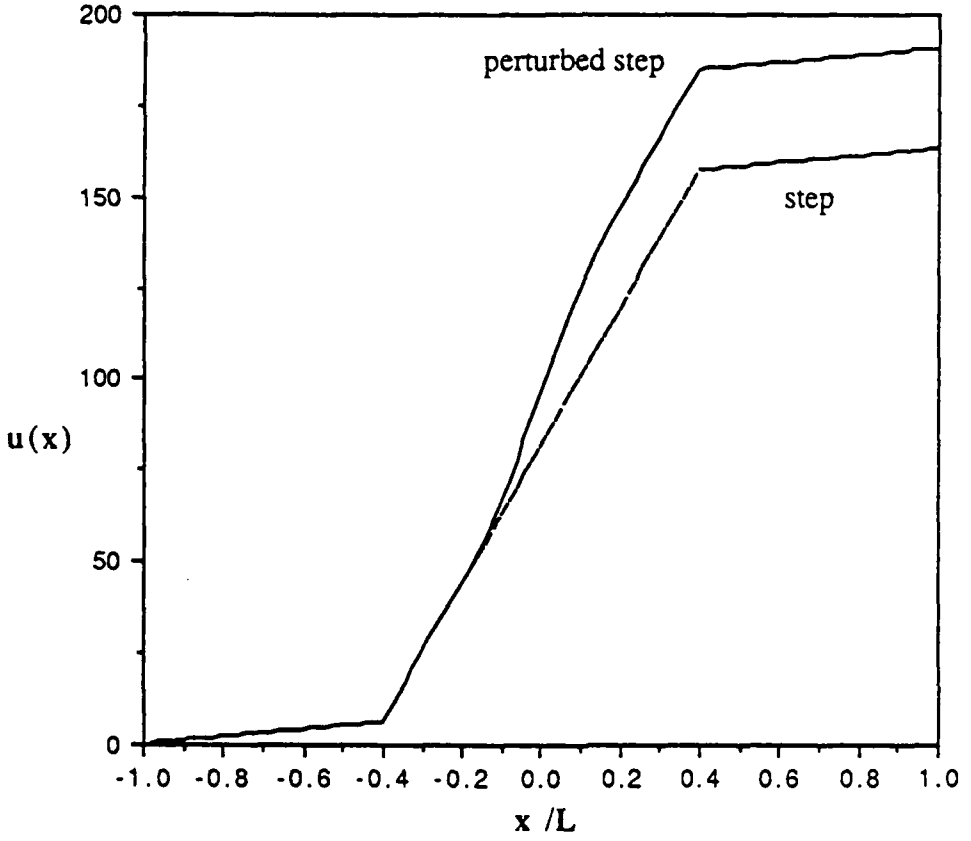


Fig. 7c. Displacement versus position for the imperfection shown in Fig. 7a.

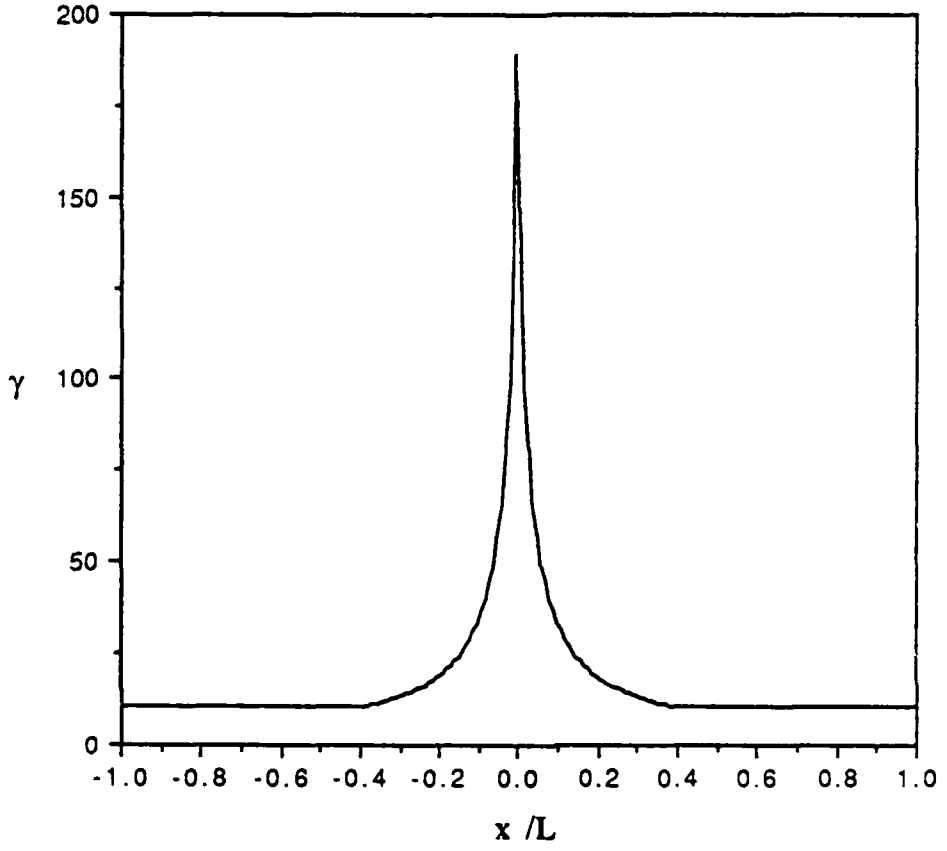


Fig. 8. Strain field for a triangular imperfection.

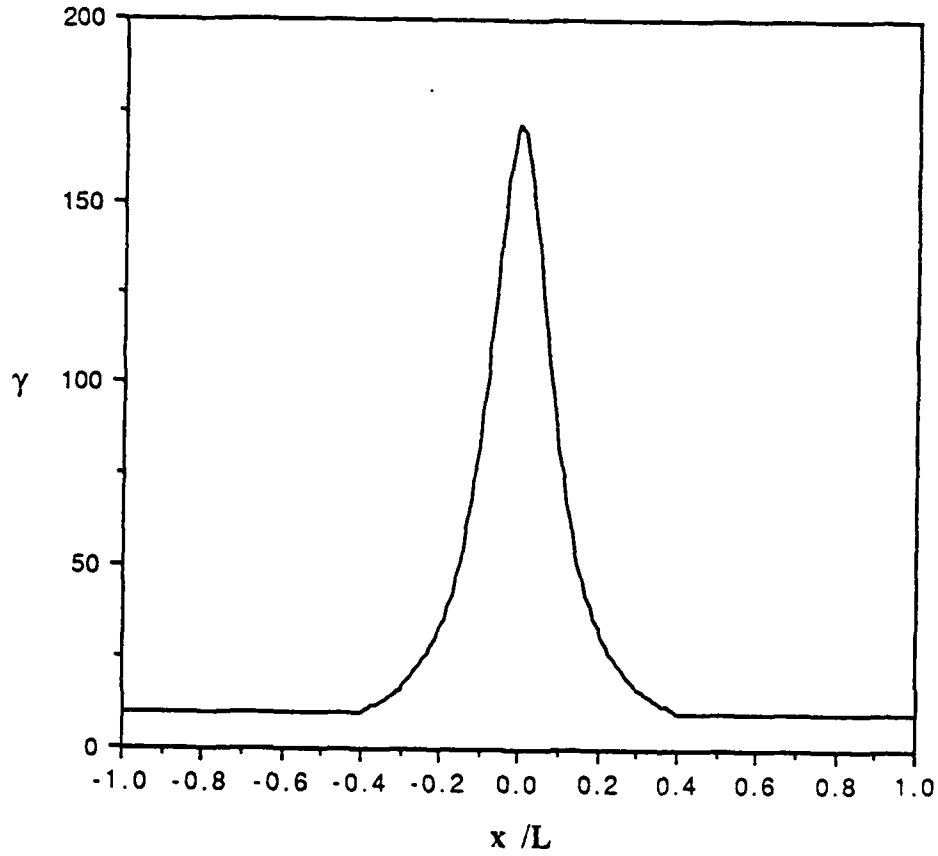


Fig. 9. Strain field for a parabolic imperfection.

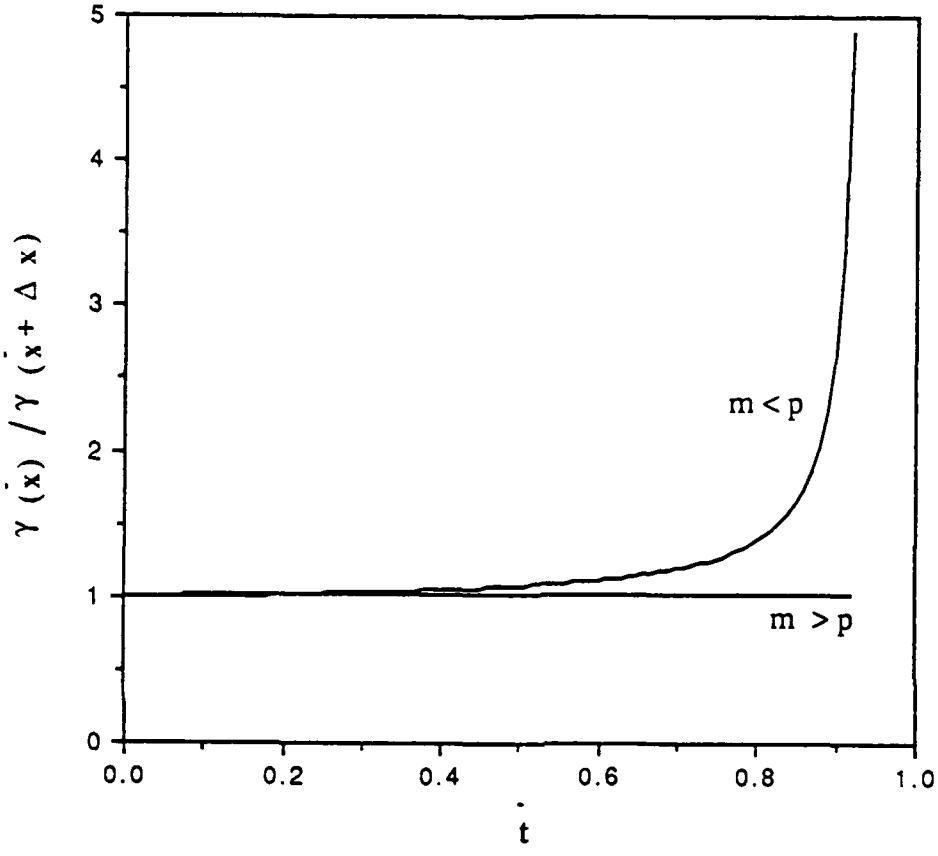


Fig. 10. Behavior of the ratio of peak strain to the strain at surrounding points as a function of time.

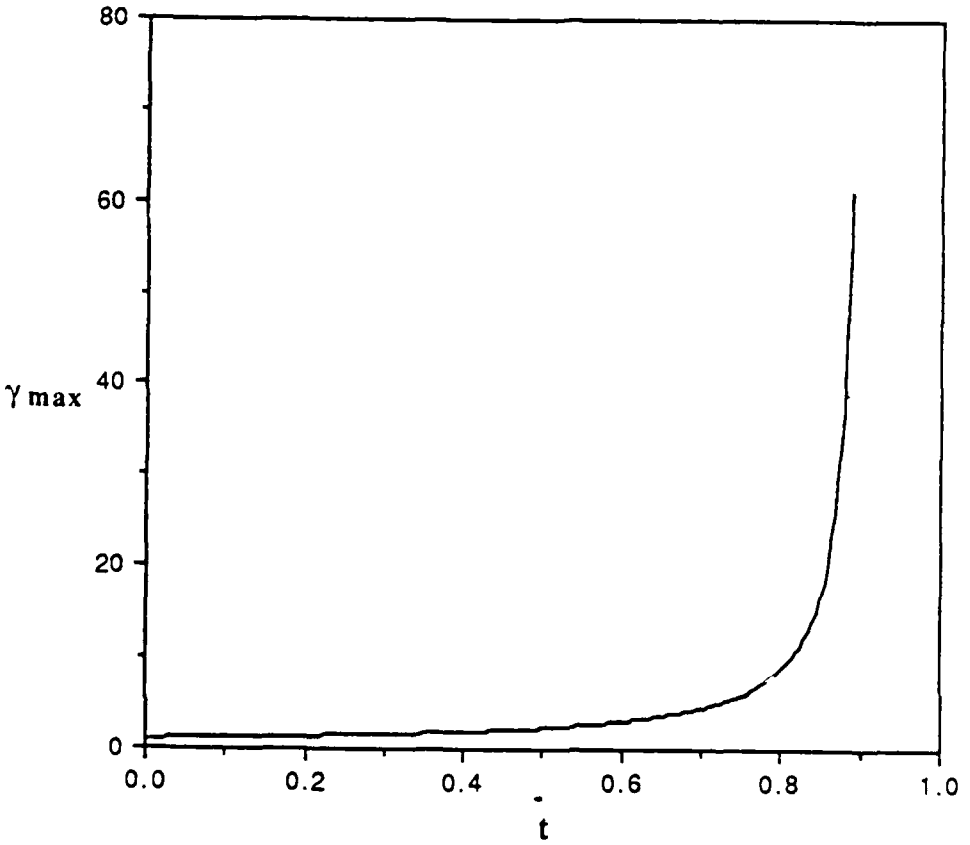


Fig. 11. Peak strain as a function of time.

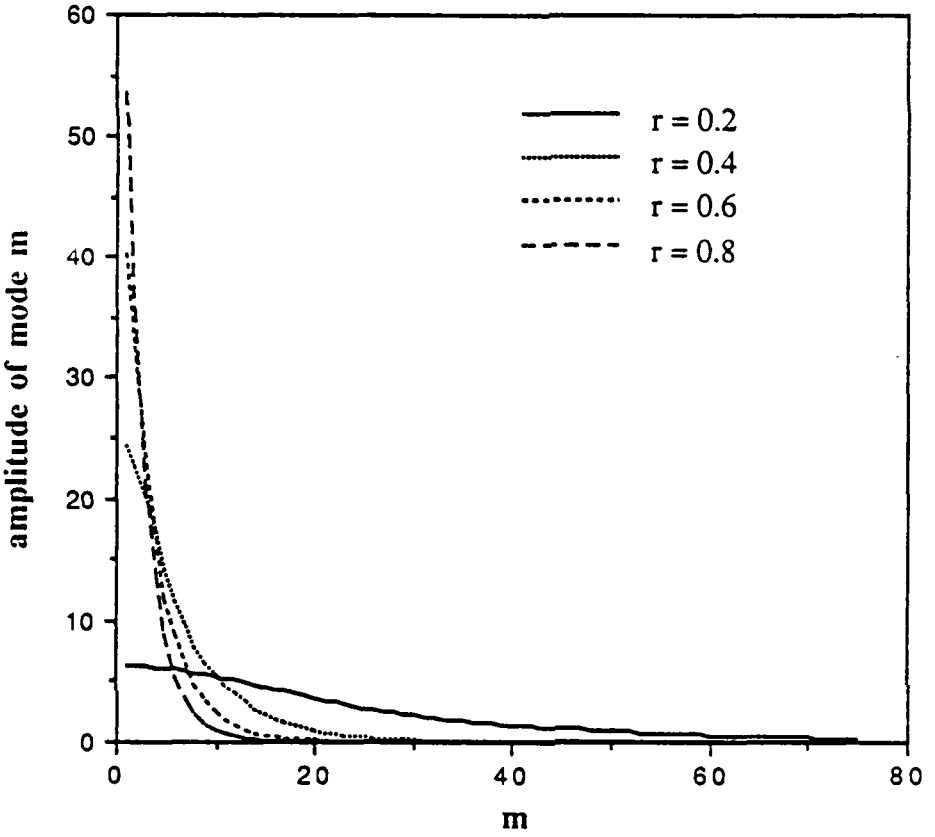


Fig. 12. Effect of the width of the imperfection on the Fourier spectrum of the strain field. Here r = w/2L where w is the width of the imperfection.

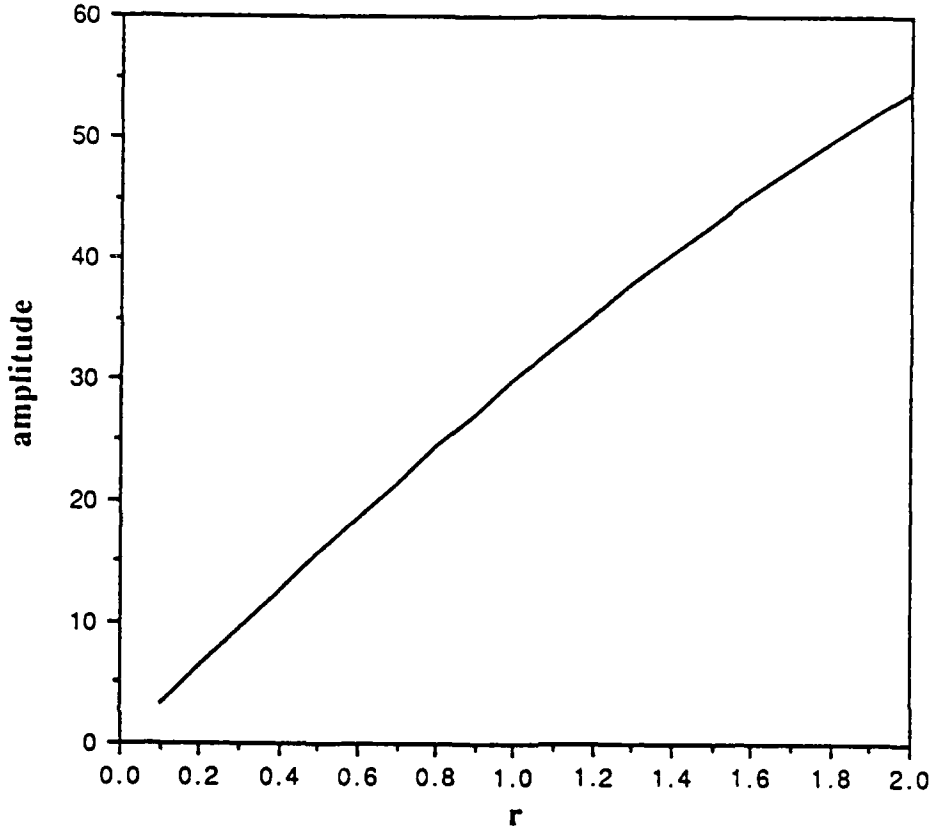


Fig. 13. Effect of width of imperfection on the amplitude of the Fourier spectrum. r = w/2 L where w is the width of the imperfection.

APPENDIX F2

Int. J. Solids Structures Vol. 24, No. 6, pp. 581–597, 1988
Printed in Great Britain.

0020–7683/88 $3.00 + .00
© 1988 Pergamon Press plc

LOCALIZATION LIMITERS IN TRANSIENT PROBLEMS

DAVID LASRY and TED BELYTSCHKO
Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, U.S.A.

(*Received* 29 *April* 1987; *in revised form* 10 *October* 1987)

Abstract—In materials with strain softening, closed form solutions have shown that the energy dissipation associated with the localization and damage processes vanishes. For finite element solutions, the dissipation depends on the size of the element in the localized zone. To remedy this situation, non-local and imbricate formulations have previously been proposed. In this paper, a simple remedy based on introducing additional higher order terms in the governing equations in the strain-softening portions of the domain is proposed. It is shown that the resulting formulation provides solutions which are independent of element size with a finite energy dissipation. Example solutions are given for one-dimensional rod problems and one-dimensional spherically symmetric problems.

1. INTRODUCTION

When deformed far enough into the inelastic range, certain materials exhibit narrow zones of intense straining. This localization of deformation may be linked in many cases, in particular for one-dimensional models, to a strain-softening effect in the constitutive behavior of the material, that is a negative slope in the stress–strain curve.

As will be seen subsequently, the appearance of localization in classical local continuum mechanics with rate independence is associated, from a mathematical point of view, with a change of type of the governing equations: loss of ellipticity in quasi-static problems; change from hyperbolic to elliptic type in the dynamic case. This change of type allows the prediction of the critical stress level which triggers localization; unfortunately, it leaves the size of the localization zone unspecified in static problems and gives infinite strains over a set of measure zero in dynamic problems[1, 2]. When incorporated into a computational model, strain-softening behavior therefore leads to severely mesh-dependent results, in which deformation localizes in one element irrespective of its size. Furthermore, the energy dissipated in the strain-softening domain tends to zero as the mesh is refined.

Several methods have been proposed to remedy this undesirable situation. In Ref. [3], a non-local formulation was used based on an averaging procedure, and for one-dimensional dynamic problems, a localization limited to a domain of finite size, and an energy dissipation that remains constant with mesh refinement were achieved. This averaging approach has also been used recently in Ref. [4], where the averaged quantities are the damage parameters in damage-type constitutive laws. Another procedure of a similar type consists of adding gradients of state variables: in Ref. [5], the gradient of the yield function was included in the constitutive equation. Triantafyllidis and Aifantis[6] used a second deformation gradient-dependent term in the expression of the strain–energy function to restore ellipticity, and obtained analytic solutions in the particular case of the Blatz–Ko material. Alternatively, Needleman illustrated in a recent paper[7] how, for a simple one-dimensional problem, the introduction of material rate dependence in the constitutive model via viscoplasticity can also result in momentum equations which remain elliptic and hyperbolic in the static and dynamic cases, respectively.

These methods, as well as others that have been proposed in recent years, ultimately try to eliminate mesh sensitivity in numerical calculations by ensuring that the size of the localization zone remains finite, and can therefore be referred to as localization limiters. In this paper, a localization limiter is proposed which is based on the introduction of an additional higher order term in the governing equations in the strain-softening portions of

the domain. Attention is focused on models for which localization is directly related to the onset of strain softening, namely dynamic one-dimensional problems.

The paper is organized as follows: Section 2 briefly reviews the mathematical aspects and difficulties underlying the localization phenomenon. Section 3 introduces the localization limiter based on the addition of high order terms to the strain expression. Examples are presented in Section 4, where the localization limiter is used in one-dimensional rod problems and spherically symmetric converging wave problems. This last type of problem is considered again in Section 5, which analyzes solutions obtained by using a material rate-dependent approach.

2. LOCALIZATION AND CHANGE OF TYPE IN THE GOVERNING EQUATIONS

From a mathematical point of view, the localization phenomenon is associated with a change of type of the governing partial differential equations. This change of type can be of various nature.

(1) In static problems, the onset of localization is connected with the loss of ellipticity in the equilibrium equations[8, 9]. In this approach, localization is viewed as an instability process: critical conditions are determined, at which the constitutive relations allow a bifurcation from a homogeneous state of deformation into a shear band with high strain values. This bifurcation is related to the loss of ellipticity of the incremental equilibrium equations. It should be pointed out that in multidimensional problems, the onset of localization is not necessarily linked with strain softening (a decline in the stress value with increasing strain), but indeed localization can occur with a positive slope in the stress–strain relation, in the presence of certain geometric or constitutive factors, such as non-associative flow rules[10, 11].

(2) In the wave propagation context, the emergence of localization is linked with a change of type in the governing wave equation, this time from hyperbolic to elliptic. Let us consider for example the one-dimensional wave equation

$$\sigma(\varepsilon)_{,x} = \rho u_{,tt}, \quad 0 \leqslant x \leqslant L \tag{1}$$

where ρ is the material mass density, u the displacement, σ the stress and ε the strain, related to the displacement through the compatibility condition

$$\varepsilon = u_{,x}. \tag{2}$$

Equations (1) and (2) can be written as a system of first-order differential equations

$$\begin{bmatrix} v \\ \varepsilon \end{bmatrix}_{,t} + \begin{bmatrix} 0 & -\dfrac{\sigma(\varepsilon)}{\rho} \\ -1 & 0 \end{bmatrix} \begin{bmatrix} v \\ \varepsilon \end{bmatrix}_{,x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3}$$

where the velocity is

$$v = u_{,t}. \tag{4}$$

If the function σ' is strictly positive, eqn (3) is a strictly hyperbolic system, and in the particular case of a linear elastic material (where $\sigma'(\varepsilon) = E_t$, the tangent Young's modulus), we find the usual expression for the wave speed c_0

$$c_0 = \sqrt{\left(\frac{E_t}{\rho}\right)}. \tag{5}$$

If, however, the tangent modulus $\sigma'(\varepsilon)$ decreases to zero for a certain strain ε_p, system (3) loses strict hyperbolicity and, further, if $\sigma'(\varepsilon) < 0$, it becomes elliptic. The corresponding initial value problem, now elliptic, is considered to be ill posed because the wave speed is imaginary. From a physical point of view, this change of type results in an accumulation of strain into a band or "deformation trapping", as described by Wu and Freund[12].

In both cases, static and dynamic, this change of type in the governing equations raises mathematical difficulties as well as numerical ones: when finite element solutions are computed, some of the results appear to be strongly dependent on the mesh refinement, since there is no finite length scale associated with the localization. Furthermore, the energy dissipation vanishes as the size of the element goes to zero, and in closed form solutions[1, 2], it is shown that the dissipation vanishes because it occurs over a set of measure zero. It is of interest to note that the imaginary character of the wave speed is not a critical shortcoming since in closed form solutions[1, 2] the strain softening is always limited to a set of measure zero.

(3) The difficulties due to a change of type in the governing equations are not confined to the area of solid mechanics. In fluid dynamics, classical transonic problems are governed by the transonic small disturbance equation[13]

$$[1 - M_\infty^2 - (\gamma + 1)M_\infty \phi_{,x}]\phi_{,xx} + \phi_{,yy} = 0 \tag{6a}$$

that can be rewritten as

$$A\phi_{,xx} + \phi_{,yy} = 0 \tag{6b}$$

where A is the non-linear coefficient appearing in eqn (6a), M_∞ the free stream Mach number, ϕ the velocity potential, and γ the ratio of specific heat coefficients; eqn (6b) is an equation of mixed type, it changes from hyperbolic to elliptic when A changes from negative (supersonic flow) to positive (subsonic flow). This change of type also causes numerical difficulties and is often treated by a regularization term consisting of higher order derivatives.

3. A LOCALIZATION LIMITER BASED ON HIGHER ORDER TERMS

As it has been already pointed out, the energy dissipation vanishes for dynamic solutions involving strain softening because the localization takes place over a set of measure zero. To remedy this situation, localization limiters have been introduced. A localization limiter should ensure that energy dissipation remains finite at arbitrary mesh refinement so that damage localization to a region of zero volume does not take place. One way this can be achieved is by inserting additional gradient terms into the momentum equation. More precisely, we will consider for the sake of clarity the one-dimensional equation of motion

$$\sigma_{,x} = \rho u_{,tt}. \tag{7}$$

We replace the usual strain–displacement relation

$$\varepsilon = u_{,x} \tag{8}$$

by the following expression involving a higher order term:

$$\hat{\varepsilon} = \varepsilon - x\varepsilon_{,xx} \tag{9a}$$

or

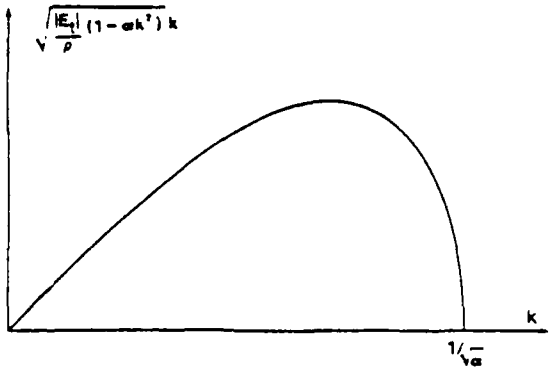


Fig. 1. Linearized analysis: plot of $\gamma(k) = \{(|E_t|/\rho)(1-\alpha k^2)\}^{1/2}k$ as a function of wavelength k.

$$\hat{\varepsilon} = u_{,x} + \alpha u_{,xxx} \tag{9b}$$

where α is a constant which will determine the extent of localization control.

We consider a linearized analysis of a disturbance δu, in which the material is in a strain-softening state over an interval $[x_1, x_2]$. Then

$$\delta\sigma(x) = -|E_t|\,\delta\varepsilon(x) \quad \text{for } x \text{ in } [x_1, x_2] \tag{10}$$

where the tangent modulus $E_t < 0$ is assumed constant. We will call the second term in the modified strain expression (9a) the *localization limiter*.

The equation of motion (7) becomes, for a small disturbance δu, using eqns (9b) and (10)

$$\delta u_{,tt} + \frac{|E_t|}{\rho}(\delta u_{,xx} + \alpha\delta u_{,xxxx}) = 0. \tag{11}$$

If we look for a possible wave solution of this last equation (using a von Neumann procedure[14])

$$\delta u(x, t) = A\, e^{ik(x-vt)} \tag{12}$$

where k is the wave number and v the phase velocity, we obtain

$$k^2v^2 = -\frac{|E_t|}{\rho}(1-\alpha k^2)k^2 \tag{13}$$

or

$$kv = i\left\{\frac{|E_t|}{\rho}(1-\alpha k^2)\right\}^{1/2}k. \tag{14}$$

We can plot (Fig. 1) $\gamma(k) = \{(|E_t|/\rho)(1-\alpha k^2)\}^{1/2}k$ as a function of the wave number k. For high frequencies, that is a wave number $k \geqslant 1/\sqrt{\alpha}$, $\gamma(k)$ is imaginary; therefore kv is real and the disturbance δu is bounded. For low frequencies ($k < 1/\sqrt{\alpha}$), kv is purely imaginary, and therefore δu grows unboundedly.

However, the region of the material where strain softening occurs is presumably small compared to a characteristic length of the problem (the total length of the rod); in fact in certain closed form solutions it is confined to a single point as $\alpha \to 0$. Therefore, large wavelengths will not be able to develop in the strain-softening region, and the above analysis shows that the higher order strain will bound any growth of short wavelength inputs. This

very simplified analysis illustrates the effectiveness of the additional higher order term in limiting the localization.

This effect can also be understood more qualitatively: near a peak in the local maximum in the strain function, the second derivative will be negative, therefore its presence will tend to reduce the strain or, equivalently, to increase the stress, since we are in the strain-softening branch of the stress–strain curve.

A simplified perturbation analysis sheds light on the relationship between the parameter α and the width δ of the localization zone. The modified equation of motion

$$u_{,tt} + \frac{|E_t|}{\rho}(u_{,xx} + \alpha u_{,xxxx}) = 0 \tag{15}$$

can be nondimensionalized by introducing $x^* = x/L$

$$\frac{\partial^2 u}{\partial t^2} + \frac{|E_t|}{\rho L^2}\left(\frac{\partial^2 u}{\partial x^{*2}} + \mu \frac{\partial^4 u}{\partial x^{*4}}\right) = 0 \tag{16}$$

where $\mu = \alpha/L^2$.

The localization limiter appears as a singular perturbation term, since the coefficient of the highest order derivative in the equation is much smaller than the other coefficients ($\mu \ll 1$). Hence, the effect of the additional term is negligible everywhere, except over a length δ where the second- and fourth-order terms should balance each other. We define a new scaling

$$\xi = \frac{x^*}{\mu^a}. \tag{17}$$

The balancing is given by

$$\frac{\partial^2 u}{\partial x^{*2}} \sim \mu \frac{\partial^4 u}{\partial x^{*4}} \tag{18a}$$

or

$$\mu^{-2a} u_{,\xi\xi} \sim \mu^{-4a+1} u_{,\xi\xi\xi\xi}. \tag{18b}$$

As μ tends to zero, the spatial derivative terms balance asymptotically when

$$a = \tfrac{1}{2}. \tag{19}$$

The size δ of the localization domain is therefore proportional to $\sqrt{\alpha}$, i.e.

$$\delta \sim p\alpha^{1/2}, \quad p = \text{constant}. \tag{20}$$

It should be pointed out that the localization limiter proposed in eqn (9b) is related to the Lax–Wendroff method[15] used in fluid mechanics in order to trigger shocks within a narrow confine of the finite difference grids. In fact, eqn (15) is a mathematical counterpart of the transient advection equation with artificial diffusion

$$\phi_{,t} - u\phi_{,x} - \alpha\phi_{,xx} = 0 \tag{21}$$

analyzed by Roache[16].

In another domain of application, higher order derivatives have also been used in the past for the backwards heat equation, known to be ill posed. In the context of phase

transformations in metal alloys (the "spinodal decomposition" problem). Cahn and Hilliard[17] defined a generalized diffusion equation

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(D(u)\frac{\partial u}{\partial x}\right) - 2K\frac{\partial^4 u}{\partial x^4}. \tag{22}$$

The additional fourth-order term in eqn (22), which accounts for the increase of energy due to the gradient of the concentration u, prevents the appearance of instability in the solution when the diffusion coefficient $D(u)$ becomes negative. Here again, a higher order term is introduced to deal with a change of type in the governing equations.

The localization limiter defined in eqn (9a) is related to the strain averaging technique used in the non-local theory through a Taylor expansion[18]. In the one-dimensional non-local theory, the non-local or average strain $\bar{\varepsilon}$ at a point x is defined as

$$\bar{\varepsilon}(x) = \frac{1}{l}\int_{-l/2}^{l/2} \varepsilon(x+s)\,\mathrm{d}s. \tag{23}$$

We can write an expansion around the center value $\varepsilon(x)$

$$\varepsilon(x+s) = \varepsilon(x) + s\varepsilon_{,x}(x) + \frac{s^2}{2}\varepsilon_{,xx}(x) + o(s^2) \tag{24a}$$

which, inserted in eqn (23), yields

$$\bar{\varepsilon}(x) = \varepsilon(x) + \frac{l^2}{24}\varepsilon_{,xx}(x) + o(l^2). \tag{24b}$$

The localization limiter $\hat{\varepsilon}$ appears as the lowest order differential approximation of the limiter based on a strain averaging procedure over a length l, with

$$l = \{24\alpha\}^{1/2} \tag{25}$$

in the one-dimensional case. It is of interest to compare expressions (20) and (25); evidently the domain of averaging l is related to the size of the domain of localization.

4. NUMERICAL IMPLEMENTATION AND EXAMPLES

The initial value problem, eqn (3) was solved by the finite element method. We used the usual weak (variational) form of eqn (1) (all natural boundary conditions are considered homogeneous)

$$\int_0^L [\delta u_{,x}\sigma + \delta u\rho u_{,tt}]\,\mathrm{d}x = 0. \tag{26}$$

This displacement field was then approximated by

$$u(x,t) = N_I(x)u_I(t) \tag{27}$$

where u_I are the nodal displacements, and repeated upper case indices are summed over the nodes. This yields the discrete equations

$$N_{I,x}\sigma \, dx + \int_0^L \rho N_I N_J u_{J,tt} \, dx = 0. \tag{28}$$

This can be written as

$$M_{IJ}u_{J,tt} + f_I^{int} = 0 \tag{29a}$$

where

$$f_I^{int} = \int_0^L N_{I,x}\sigma \, dx \tag{29b}$$

$$M_{IJ} = \int_0^L \rho N_I N_J \, dx. \tag{29c}$$

A tangent stiffness is not used in the numerical implementation since an explicit time integrator is used, but it is interesting to write it down since it indicates that the system is not self-adjoint. The tangent stiffness is defined by

$$f_{I,t}^{int} = K_{IJ}^{tan} u_{J,t} \tag{30}$$

so using eqns (29b), (27) and (9) we obtain

$$K_{IJ}^{tan} = \int_0^L N_{I,x} E_t (N_{J,x} + N_{J,xxx}) \, dx \tag{31a}$$

where E_t is the tangent modulus. This tangent stiffness matrix, that would appear in an implicit treatment of the transient problem, could be transformed by an integration by parts into a self-adjoint operator and an additional boundary term

$$K_{IJ}^{tan} = \int_0^L N_{I,x} E_t N_{J,x} \, dx - \int_0^L N_{I,xx} E_t N_{J,xx} \, dx + [N_{I,x} E_t N_{J,xx}]_0^L. \tag{31b}$$

The introduction of the second-order derivative in the strain expression, and therefore of a fourth-order displacement derivative in the equation of motion would require in theory the use of C^1 (Hermite) shape functions. As we wish only to illustrate the effectiveness of the method, we have used simple linear displacement (C^0) finite elements. The second derivative is computed by a centered finite difference approximation

$$(\varepsilon_{,xx})_j = (\varepsilon_{j+1} - 2\varepsilon_j + \varepsilon_{j-1})/\Delta x^2 \tag{32}$$

and we have used uniform (constant Δx) meshes. Calculations were conducted with an explicit time integrator, namely a central difference method with a lumped mass matrix. We used for most calculations a Courant number $CFL = 0.7$–0.9, although, in certain cases, a lower value 0.4 had to be used to obtain stable results.

An interesting aspect of the method is that the localization limiter is introduced only in the region where strain softening occurs: this is implemented by initializing at zero an array of values $\alpha(j)$, $j = 1, NELE$ where $NELE$ is the number of elements, and attributing the value $\alpha(j) = \alpha$ only when element j has reached a strain-softening state. Therefore, in the regions of the body where strain softening is not reached, the strain considered in the equation of motion is the usual local strain ε, without any higher order derivatives.

Two additional features were incorporated in the method. It should first be noted that

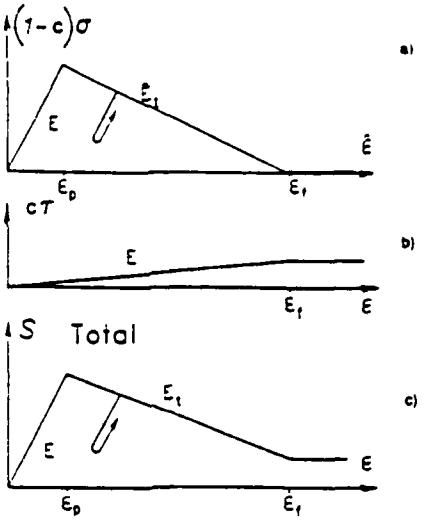


Fig. 2. Combination of stress–strain laws. For the one-dimensional rod problem, E is Young's modulus, for the spherically symmetric problem, E is replaced by K (bulk modulus), E_f by K_f, ε by ε_v (volumetric strain).

$$\varepsilon + \alpha \varepsilon_{xx} = 0 \tag{33}$$

admits as a solution

$$\varepsilon_1(x) = A \cos \frac{1}{\sqrt{\alpha}} x + B \frac{1}{\sqrt{\alpha}} x \tag{34}$$

where A and B can be arbitrarily large. This solution ε_1, called a zero-energy mode, can appear in the numerical solution; when superposed on the "real" solution, it leads to meaningless results. The same problem arises in the finite element implementation of non-local methods based on strain averaging[3], and has been resolved so far only by combining a local and a non-local law. Similarly, in our case, we will combine two stress–strain relations

$$\sigma = \hat{C}(\hat{\varepsilon})\hat{\varepsilon} \tag{35a}$$

$$\tau = C(\varepsilon)\varepsilon \tag{35b}$$

where σ is the stress resulting from the localization limiter $\hat{\varepsilon}$, τ the stress resulting from the usual strain ε. $\hat{C}(\hat{\varepsilon})$ and $C(\varepsilon)$ account respectively for a constitutive relation with strain softening and one without it, as summarized in Fig. 2.

The stress used in the equation of motion is taken as a combination

$$S = (1-c)\sigma + c\tau. \tag{36}$$

The term $c\tau$ acts as a stabilizing factor, preventing the development of zero energy modes. Parameter c is typically of the order of 0.1.

Second, in order to obtain solutions in the presence of the higher order derivatives, it is necessary to add damping. When wave propagation phenomena are analyzed by finite elements, oscillations from one element to the next appear ahead of and behind of the wave front, in both time and space (Gibbs' oscillations)[19]. These oscillations would lead to erratic second derivatives, and we therefore prevent them by introducing damping via

$$f_I^{\text{int}} = \int N_{I,x} S \, dx + \beta \int N_{I,x} \dot{S} \, dx \tag{37}$$

where parameter β controls the amount of damping introduced. The above represents a stiffness proportional damping, which acts on the high frequencies introduced by the Gibbs phenomenon, whereas mass proportional damping would affect low frequencies.

It should be pointed out that the *raison d'être* of these two additional features is purely related to numerical aspects, and that they are not intrinsic to the method itself. As far as damping is concerned for example, Sandler and Wright have shown[20] that in certain situations (wave propagation in a rod) damping alone can act as a localization limiter with a certain effectiveness. Reference [21] showed however the limitations of that approach, and its inappropriateness in other situations such as in the spherically symmetric problem described in Section 4.2.

The localization limiter introduced here can handle these two types of problems successfully, proving therefore that it is not the damping term alone that accounts for the localization limiting effect.

4.1. *Wave propagation in a rod*

This problem was considered in Ref. [3] (Fig. 3(a)). Equal and opposite velocities v_0 are applied to the two ends of a rod of length $2L$ made of a strain-softening material, so that tensile waves are generated at the ends. The magnitude of the strain is slightly less than the strain corresponding to the onset of strain softening. These tensile waves propagate elastically to the center; when they meet at the center, the stress would double if the behavior remained elastic, so that strain softening starts at this midpoint.

The analytical solution for this problem was proposed in Ref. [1]: localization occurs at the midpoint where the strain becomes infinite. The solution, symmetric about the midpoint $x = L$, is expressed for the left half as

$$u = -v_0\left\langle t - \frac{x}{c_0}\right\rangle - v_0\left\langle t - \frac{2L - x}{c_0}\right\rangle \tag{38a}$$

$$\varepsilon = \frac{v_0}{c_0}\left[H\left(t - \frac{x}{c_0}\right) - H\left(t - \frac{2L - x}{c_0}\right) + 4\langle c_0 t - L\rangle \delta(x - L)\right] \tag{38b}$$

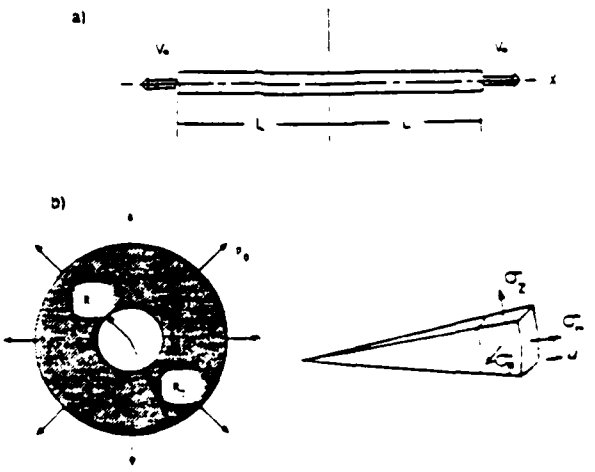


Fig. 3. Problem descriptions: (a) one-dimensional rod problem, $2L = 40$; (b) spherically symmetric problem, interior radius $R_1 = 10$, exterior radius $R_2 = 40$.

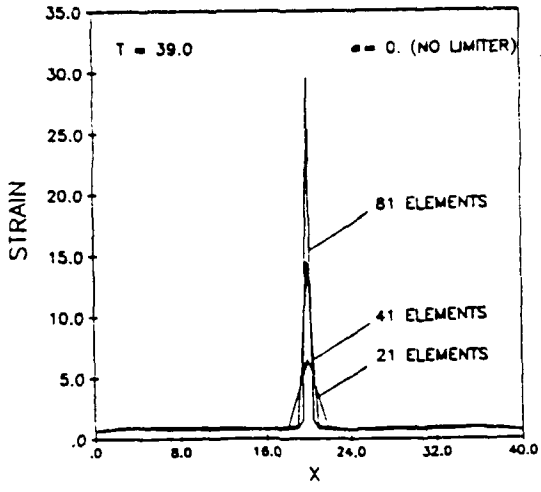


Fig. 4. One-dimensional rod problem, strain plots at time $T = 39.0$ for different meshes, no higher order term limiter ($\alpha = 0$).

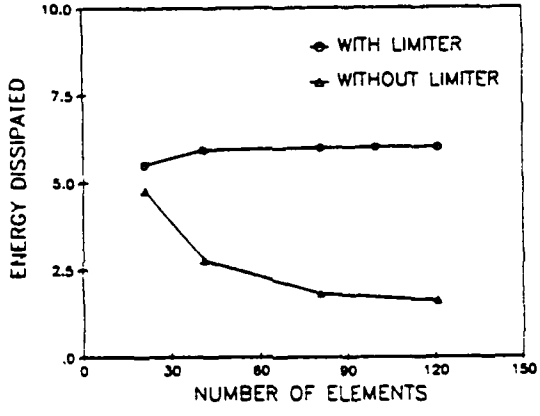


Fig. 5(a). One-dimensional rod problem, energy dissipated in the rod vs number of elements, with and without localization limiter.

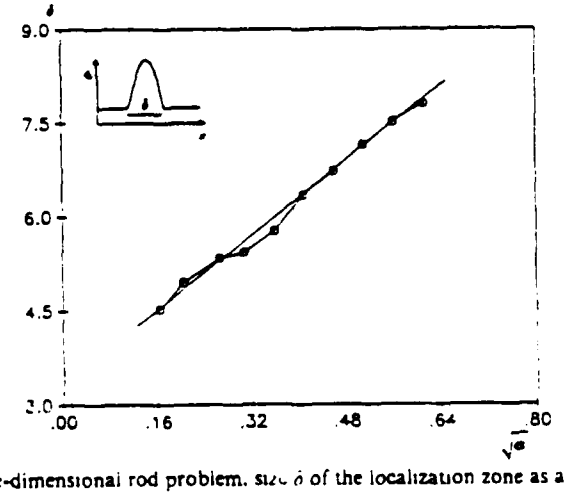


Fig. 5(b). One-dimensional rod problem, size δ of the localization zone as a function of $\sqrt{\alpha}$.

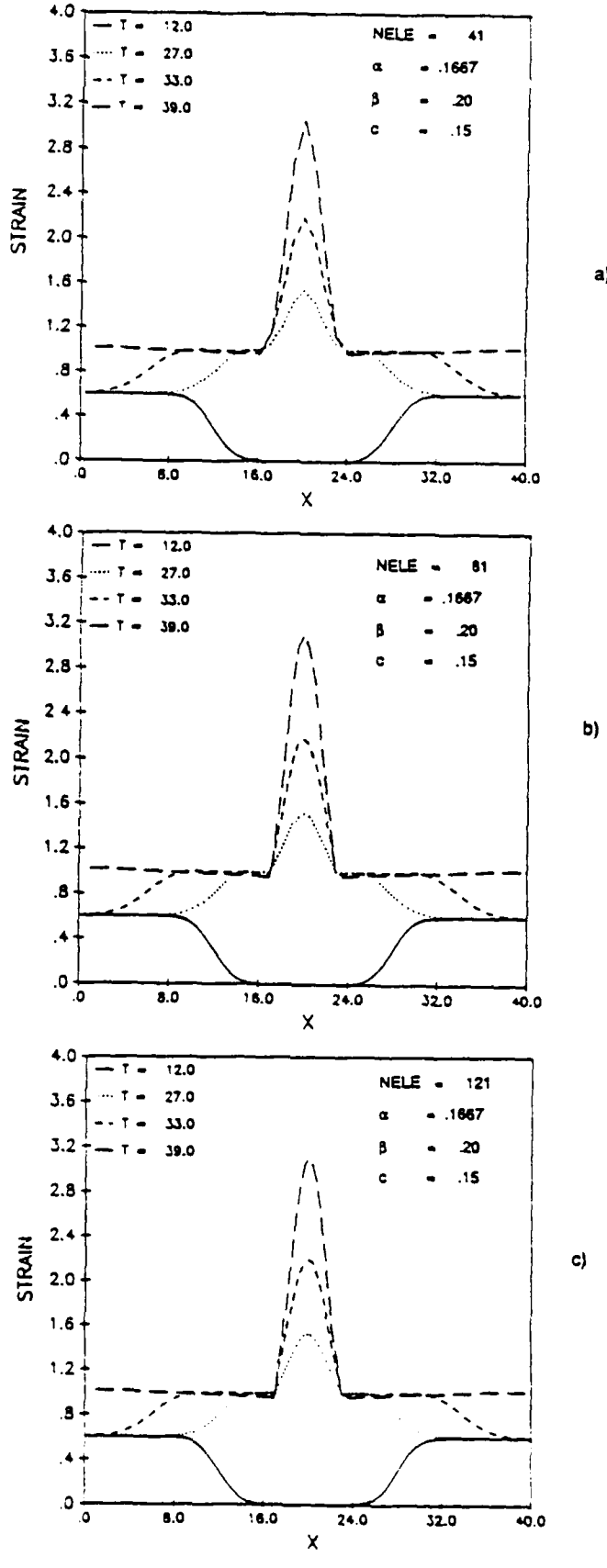


Fig. 6. One-dimensional rod problem, strain plots obtained using the localization limiter ($\alpha = 0.1667$): (a) 41 elements; (b) 81 elements; (c) 121 elements. For all plots, combination factor $c = 0.15$, damping coefficient $\beta = 0.20$.

where H is the Heaviside step function, $\langle A \rangle = A$ if $A > 0$, $\langle A \rangle = 0$ otherwise, δ is the Dirac-delta function, c_0 the elastic wave speed in the material. Numerical studies of this problem based on non-local approaches were conducted in Refs [3, 4]. Here we will use the localization limiter

$$\hat{\varepsilon} = \varepsilon + \alpha \varepsilon_{,xx}. \tag{9a}$$

The combination of stress–strain laws considered to stabilize the zero energy modes is illustrated in Fig. 2, and we used a damping factor $\beta = 0.20$, and a combination parameter $c = 0.15$, in accordance with the results of Ref. [3]. Other parameters used in the calculations were: density $\rho = 1$, end velocity $v_0 = 0.6$, and for the stress–strain relations in Fig. 3, $E = 1$, yield stress $\sigma_y = 1$, $E_s = -0.25$, $\varepsilon_f = 5$, nearly horizontal tail of slope $E_t = 0.001$ beyond ε_f.

It was first checked (Fig. 4) that, without introducing the localization limiter (that is, for $\alpha = 0$), the strain profiles are severely dependent on the mesh refinement, and the localization zone shrinks to one element, irrespective of its size. Furthermore, the total energy dissipated in the mesh tends to zero as the mesh is refined, as seen in Fig. 5(a). Convergence studies were then performed with the localization limiter, for a value $\alpha = 0.1667$, for different meshes with an increasing number of elements (Figs. 6(a)–(c)). They exhibit a localization limited to a finite zone, the length of that zone and the strain profiles being independent of the mesh refinement. Moreover, the total energy dissipated in the rod, computed as in Ref. [3] by

$$W(t_{r+1}) = W(t_r) + \sum_{j=1}^{NELE} \frac{h_j}{2}(1-c)[\sigma_{j,r}\varepsilon_{j,r} - \sigma_{j,r+1}\varepsilon_{j,r+1} + (\sigma_{j,r} + \sigma_{j,r+1})(\varepsilon_{j,r+1} - \varepsilon_{j,r})] \tag{39}$$

is independent of the mesh size, all other parameters remaining equal (in eqn (39), r refers to the time step, and h_j the size of element j). This is illustrated in Fig. 5(a).

Calculations were also conducted at fixed mesh size for different values of α (Fig. 5(b)), showing that the length of the localization zone is *linearly dependent on* $\sqrt{\alpha}$ (the line does not extrapolate to the origin (0, 0) since the presence of the damping coefficient β partially limits the localization[20], even for $\alpha = 0$). This is consistent with the results of Refs [3, 4], which found a linear dependence in l (averaging length), since the Taylor expansion (23) yielded a linear relation between α and l^2.

4.2. *Spherically symmetric problem*

This problem (Fig. 3(b)) was considered with strain-softening materials in Ref. [21]. A sphere made of a strain-softening material is loaded with a uniform traction on its exterior surface. To better appreciate the complexity of this problem, consider the load to be a ramp function in time. Before the onset of strain softening at an interior surface S, a portion of the stress will have passed through S. Due to the spherical geometry, the stresses in this wave are amplified as they pass to the center and trigger the formation of additional strain-softening surfaces. As conjectured in Ref. [21], it seems that an infinite number of *localization surfaces* will appear, although no analytical solution has been proposed so far.

The localization limiter defined previously in eqn (9a) was used to solve numerically this problem. We considered a sudden application of a uniform normal traction $\sigma_r = p_0 H(t)$ at the exterior surface, where H is the Heaviside step function. The applied surface pressure was chosen as $p_0 = 0.708$; for this boundary condition, the wave propagating from the outer surface remains elastic until the wavefront reaches 30% of the thickness $R_2 - R_1$. The other parameters used for the calculations were $c = 0.15$, $\beta = 0.10$, $R_1 = 10$, $R_2 = 100$, and the same material constants as in Section 4.1 were considered. The equations corresponding to this problem, adapted from Ref. [21], are summarized in the Appendix.

It was first noted (Fig. 7) that without the localization limiter (that is for $\alpha = 0$), as the number of elements is increased, several points of localization develop, and these points

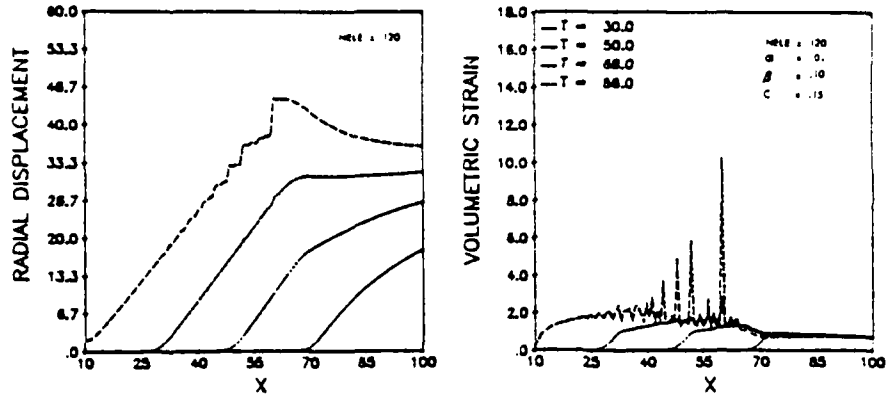


Fig. 7. Spherically symmetric problem, radial displacement and volumetric strain plots, without using the higher order term limiter ($\alpha = 0$).

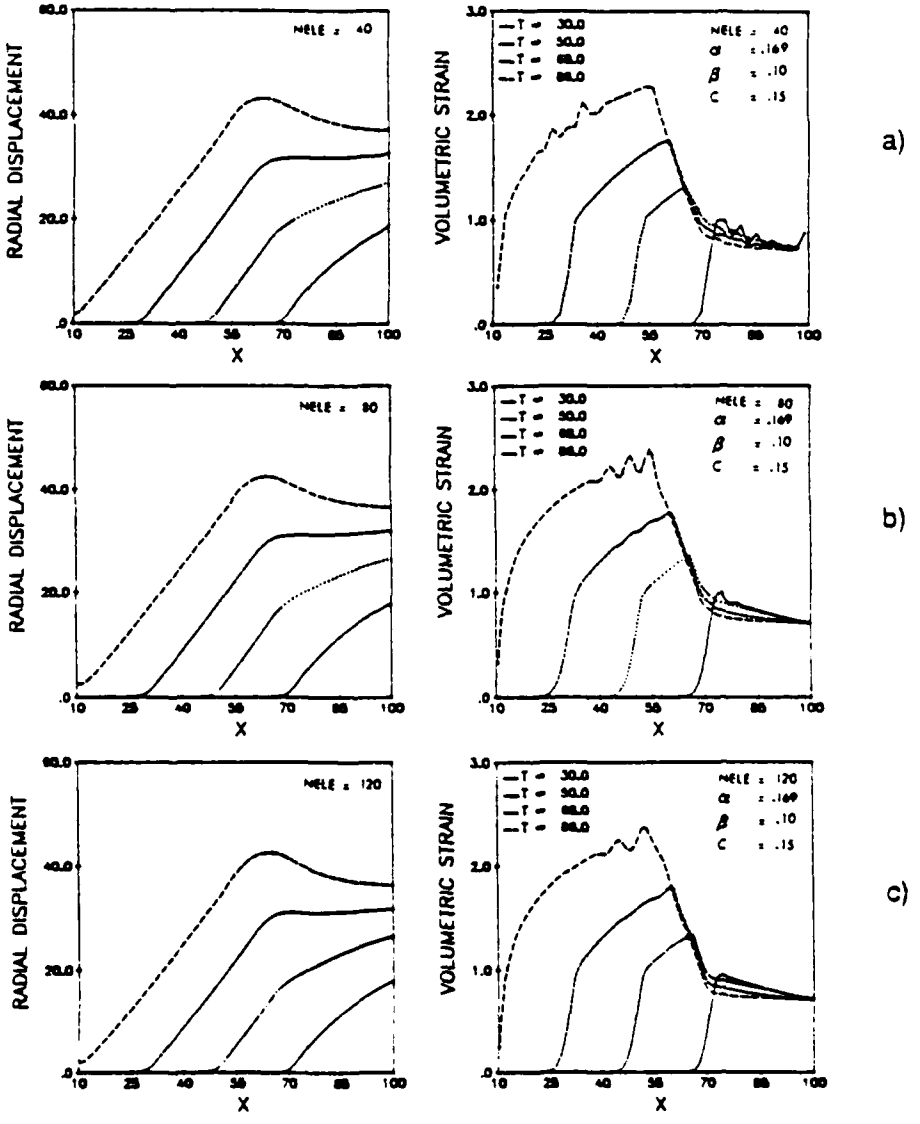


Fig. 8. Spherically symmetric problem, radial displacement and volumetric strain plots, with limiter ($\alpha = 0.169$): (a) 40 elements; (b) 80 elements; (c) 120 elements. For all plots, combination factor $c = 0.15$, damping coefficient $\beta = 0.15$.

change arbitrarily with mesh refinement, even in the presence of damping. These points of localization can be appreciated both in the volumetric strain plots, with the presence of spikes, and in the radial displacement plots, where sharp discontinuities indicate separation along a surface.

The next group of solutions (Figs 8(a)–(c)) examines the effect of the localization limiter. These solutions converge well with mesh refinement, and furthermore, they are very similar to those found with the imbricate element approach[21].

5. MATERIAL RATE DEPENDENCE SOLUTION TO THE SPHERICAL PROBLEM

It has been proposed recently by Needleman[7] that viscoplastic rate dependence eliminates the change of type associated with strain softening. His discussion, presented in the context of simple shearing, is equivalent to the tensile rod problem of Section 4.1. To summarize the formulation of Ref. [7], the incremental shear stress–shear strain relation is written as

$$\dot{\tau} = G(\dot{\gamma} - \dot{\gamma}^{\mathrm{p}}). \tag{40}$$

Material rate dependence is introduced by writing the plastic strain rate $\dot{\gamma}^{\mathrm{p}}$ in eqn (40) as

$$\dot{\gamma}^{\mathrm{p}} = \dot{a} f(\tau, g) \tag{41}$$

where $\dot{a}$ is a material parameter and g a hardness parameter that evolves with accumulating plastic strain; f is given by a power law

$$f(\tau, g) = (\tau/g)^{1/m} \tag{42}$$

where m is the strain rate hardening exponent. In the numerical applications, g is taken to be a function of the effective plastic strain

$$\bar{\gamma} = \int |\dot{\gamma}^{\mathrm{p}}| \, \mathrm{d}t \tag{43}$$

here, a simple linear strain function is used

$$\dot{g} = H_1 \dot{\bar{\gamma}} \tag{44}$$

with $g(0) = \tau_0$.

The non-dimensional quantities τ_0/G, H_1/τ_0, τ/τ_0, $\gamma/(\tau_0/G)$ defined in Ref. [7] were also used here. Furthermore, we considered numerical values of the different parameters leading, in the rate-independent limit case, to a stress–strain curve similar to the curve considered here (Fig. 2(c)). A step-by-step integration of constitutive relations (40)–(44) reveals the role of the parameter m, as shown in Fig. 9: the rate-independent limit is approached as $m \to 0$, whereas, for increasing values of m, that is, for an increasingly viscous solid, the stress–strain curve "overshoots" the strain-softening branch, delaying the loss of positive definiteness of the stress–strain modulus.

The spherical problem of Section 4.2 was then solved introducing this material rate dependence, for different values of m. Parameters γ, τ and G are now interpreted as the volumetric strain, hydrostatic stress and bulk modulus, respectively. The shear modulus is considered as in Section 4.2 to be negligibly small (10^{-6}).

Figure 10 illustrates that, as m becomes smaller, the solution gets closer to the ones obtained with our localization limiter Fig. 8(b). However, the calculations become increasingly difficult: as pointed out in Ref. [7], when the constitutive equations are updated with a simple forward Euler method, the stable time step decreases and eventually vanishes with m. For example, the strain profile for $m = 0.005$ was obtained by running the problem with

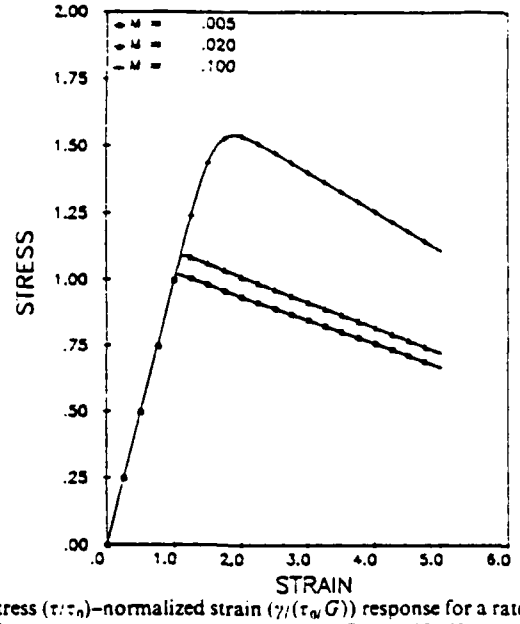


Fig. 9. Normalized stress (τ/τ_0)–normalized strain $(\gamma/(\tau_0/G))$ response for a rate-dependent material, for various values of the strain-hardening exponent m. $\tau_0/G = 0.008$, $H_1/\tau_0 = -10$, constant strain rate $\dot{\gamma}$ such that $\dot{\gamma}/a = 10^2$.

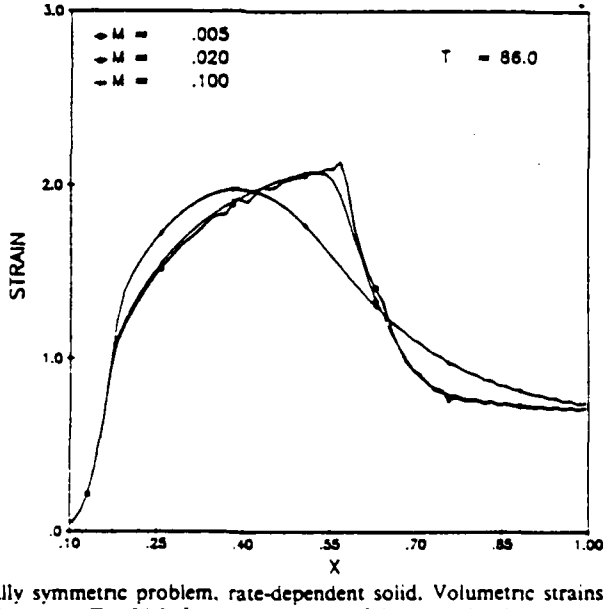


Fig. 10. Spherically symmetric problem, rate-dependent solid. Volumetric strains vs normalized radius x/R_2 at time $T = 86.0$, for various values of the strain-hardening exponent m.

a Courant number $CFL = 0.08$, but even this very small value leads to unstable calculations for $m = 0.001$.

6. CONCLUSIONS

The addition of a second derivative term to the strain expression leads to an efficient localization limiter for strain-softening materials, possessing the following desirable features: finite energy dissipation at arbitrary mesh refinement, development of a localized strain zone of finite size, independence of the calculations on mesh size. This theory may be considered nonlocal in that it emanates from a Taylor expansion of a non-local strain expression with a constant weight function.

It was shown by a Fourier analysis that this localization limiter yields a system of equations which attenuate short waves but cannot prevent the occurrence of instabilities in long waves. Since the initiation of strain softening appears to occur in very narrow zones, and the second derivatives are used only in the strain-softening (or localization) domains, this analysis implies that the growth of the strains in the localization domain will be bounded, rather than being unbounded as in the local, rate-independent medium. This was confirmed by numerical results.

A perturbation analysis reveals that the width of the zone in the strain-softening regime varies with the square root of α which determines the strength of the participation of the higher order derivative. This type of dependence is also indicated by a non-local interpretation of this limiter and was also verified numerically. This localization limiter was tested with success on one-dimensional rod problems, and on the more critical spherically symmetric problem, which emerges as a decisive test in evaluating the effectiveness of a localization limiter.

The localization limiting properties of the viscoplastic material model were also studied. It was shown that even in the spherically converging problem, this material model acts as a localization limiter. However, it is somewhat ineffective in explicit time integration problems since the stable time step becomes extremely small.

Several questions remain to be explored. A more thorough study of the additional boundary conditions introduced by the higher order terms is necessary. It would be desirable to relate the size of the localization zone (and therefore α) to the micromechanical behavior of the material, such as grain size, microcracking effects and other microstructural properties, and to give a physical interpretation of the additional terms.

Acknowledgement—The support of the U.S. Air Force Office of Scientific Research under Award F49620-87-C-0030 is gratefully acknowledged.

REFERENCES

1. Z. P. Bazant and T. B. Belytschko, Wave propagation in strain-softening bar: exact solution. *J. Engng Mech. ASCE* **111**, 381–389 (1985).
2. T. Belytschko, X.-J. Wang, Z. Bazant and H. Hyun, Transient solutions for one dimensional problems with strain-softening. *J. Appl. Mech.* **54**(3), 513–518 (1987).
3. Z. P. Bazant, T. B. Belytschko and T. P. Chang, Continuum theory for strain-softening. *J. Engng Mech. ASCE* **110**, 1666–1692 (1984).
4. G. Pijaudier-Cabot and Z. P. Bazant, Nonlocal damage theory. *J. Engng Mech. ASCE* **113**(10), 1512–1533 (1987).
5. H. L. Schreyer and Z. Chen, The effect of localization on the softening behavior of structural members. In *Constitutive Equations: Macro and Computational Aspects* (Edited by K. J. William), pp. 193–203. ASME, New York (1984).
6. N. Triantafyllidis and E. C. Aifantis, A gradient approach to localization of deformation. I. Hyperelastic materials. *J. Elasticity* **16**, 225–237 (1986).
7. A. Needleman, Material rate dependence and mesh sensitivity in localization problems. Report ONR (1987).
8. J. Hadamard, *Leçons sur la Propagation des Ondes*, Chap. VI. Hermann, Paris (1903); reprinted by Chelsea, New York (1949).
9. R. Hill, Acceleration waves in solids. *J. Mech. Phys. Solids* **10**, 1–16 (1962).
10. J. R. Rice, The localization of plastic deformation. In *Theoretical and Applied Mechanics. Proc. 14th Int. Congr. Theoret. Appl. Mech.* (Edited by W. T. Koiter), pp. 207–220. North-Holland, Amsterdam (1977).
11. J. W. Rudnicki and J. R. Rice, Conditions for the localization of deformation in pressure sensitive dilatant materials. *J. Mech. Phys. Solids* **23**, 371–394 (1975).
12. F. H. Wu and L. B. Freund, Deformation trapping due to thermoplastic instability in one-dimensional wave propagation. *J. Mech. Phys. Solids* **32**, 119–132 (1984).
13. P. R. Garabedian, *Partial Differential Equations*. Wiley, New York (1964).
14. K. W. Morton and R. D. Richtmeyer, *Difference Methods for Initial-value Problems*. 2nd Edn. Interscience, New York (1967).
15. P. Lax and B. Wendroff, Systems of conservations laws. *Commun. Pure Appl. Math.* **132**, 217–237 (1960).
16. P. Roache, On artificial viscosity. *J. Comp. Phys.* **10**(2), 169–184 (1972).
17. J. W. Cahn and J. E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy. *J. Chem. Phys.* **28**, 258 (1958).
18. Z. P. Bazant, Imbricate continuum and its variational derivation. *J. Engng Mech. ASCE* **110**, 1693–1712 (1984).
19. N. Holmes and T. B. Belytschko, Postprocessing of finite element transient response calculations by digital filters. *Comput. Struct.* **6**, 211–216 (1976).
20. I. Sandler and J. Wright, Summary of strain-softening. In *Theoretical Foundations for Large-scale Computations of Nonlinear Material Behavior*. DARPA-NSF Workshop (Edited by S. Nemat-Nasser), pp. 285–315 (1984).

21. T. B. Belytschko, Z. P. Bazant, Y. W. Hyun and T. P. Chang, Strain-softening materials and finite-element solutions. *Comput. Struct.* **23**(2), 163–180 (1986).

APPENDIX: GOVERNING EQUATIONS FOR THE SPHERICALLY SYMMETRIC PROBLEM

The strain–displacement relations are written as

$$\varepsilon_r = u_{,r}$$

$$\varepsilon_\theta = \varepsilon_z = \frac{u}{r}$$

$$\varepsilon_v = \varepsilon_r + 2\varepsilon_\theta$$

$$e_r = \tfrac{2}{3}(\varepsilon_r - \varepsilon_\theta)$$

$$e_\theta = \tfrac{1}{3}(\varepsilon_\theta - \varepsilon_r)$$

where u is the radial displacement (Fig. 3(b)), ε_r and ε_θ the radial and circumferential normal strains, ε_v the volumetric strain, and e_r and e_θ the deviatoric strains.

We define the modified quantities $\dot{\varepsilon}_r$, $\dot{\varepsilon}_\theta$, $\dot{\varepsilon}_v$, $\dot{e}_r$ and $\dot{e}_\theta$ in a way similar to eqn (9a). In particular, for the volumetric strain

$$\dot{\varepsilon}_v = \varepsilon_v - \alpha\varepsilon_{v,rr}.$$

The incremental stress–strain relations are given by

$$\dot{\sigma}_r = \hat{K}\dot{\varepsilon}_v + 2\hat{G}\dot{e}_r$$

$$\dot{\sigma}_\theta = \hat{K}\dot{\varepsilon}_v + 2\hat{G}\dot{e}_\theta$$

$$\dot{\tau}_r = K\dot{\varepsilon}_v + 2G\dot{e}_r$$

$$\dot{\tau}_\theta = K\dot{\varepsilon}_v + 2G\dot{e}_\theta.$$

The shear moduli $\hat{G}$ and G are assumed to be constant, while the tangent bulk moduli $\hat{K}$ and K and the stress–strain relations (σ, ε_v) and (τ, ε_v) are illustrated in Fig. 2. The total stress is given by

$$S_r = (1-c)s_r - c\tau_r$$

$$S_\theta = (1-c)\sigma_\theta + c\tau_\theta.$$

In the numerical applications, the shear moduli $\hat{G}$ and G are taken to be negligibly small (10^{-6}).

Computers & Structures Vol. 33, No. 3, pp. 707–715, 1989
Printed in Great Britain

0045-7949/89 $3.00 + 0.00
© 1989 Pergamon Press plc

APPENDIX F3

A STUDY OF LOCALIZATION LIMITERS FOR STRAIN-SOFTENING IN STATICS AND DYNAMICS

TED BELYTSCHKO and DAVID LASRY
Department of Mechanical Engineering, Northwestern University, The Technological Institute, Evanston, IL 60208, U.S.A.

(*Received* 16 *September* 1988)

Abstract—In localization problems involving strain-softening, localization limiters have been introduced to provide solutions where the deformation concentrates in regions of finite size. Limiters are here classified as integral, differential and rate limiters. The character of the governing partial differential equations is studied for strain-softening models with and without limiters. It is shown that, in simple cases, strain-softening causes the governing equations of statics and dynamics to lose respectively ellipticity and strict hyperbolicity. With the addition of a differential limiter, the dynamic equations change from hyperbolic to parabolic. Differential and integral limiters are shown to have similar effects on superimposed waves, with an unbounded growth occurring only for wavelengths greater than the characteristic localization bandwidth. Example solutions are given for one-dimensional models in statics.

1. INTRODUCTION

Solutions of problems involving strain-softening material laws are fraught with serious difficulties, both from mathematical and numerical points of view. In dynamics, these difficulties were illustrated in [1] using a simple one-dimensional wave propagation model: an elastic wave propagating in a bar travels with a velocity proportional to $\sqrt{(E_T)}$, where E_T is the tangent modulus. When E_T becomes negative, as in the case for a strain-softening material, the wave cannot propagate any more, giving rise to what Wu and Freund [2] called a deformation-trapping phenomenon: the deformation is trapped in a certain zone of the body and no information can be transmitted to the rest of the material. Bazant and Belytschko [1] have shown by a closed-form solution that strain-softening in transient problems is characterized by the appearance of infinite strains on a set of measure zero. This is reflected in numerical simulations by a strong dependency of the results upon the refinement of the mesh [3]. When the equations are discretized, by finite elements for example, the deformation will localize in the smallest discrete cell of material capable of representing that set of measure zero, namely one element in constant-strain elements in one dimension or a one-element-wide band in two dimensions. In other words, the problem of mesh dependency is not an intrinsically numerical one, but rather stems from the more fundamental loss of strict hyperbolicity of the equations of motion [4] upon attaining the strain-softening regime.

In statics, localization has been associated with the loss of ellipticity of the incremental equilibrium equations [5], and the existence of a bifurcation from a homogeneous state of deformation into nonhomogeneous ones and the appearance of multiple equilibrium paths. This approach provides the orientation of the localization band and the critical load for which localization may be triggered but does not provide any length parameter for the subsequent behavior; in this respect, it is somewhat different from the dynamic case, where a localization zone (albeit reduced to a single point or line) appears in the closed-form solutions [3]. This difference is also reflected in the numerical simulations. If a solid with no imperfections is submitted to a homogeneous state of deformation, the numerical solution for a static problem will follow that homogeneous deformation path even when it becomes unstable beyond the bifurcation point, provided the machine precision is sufficient to prevent round-off error from triggering an inhomogeneous mode.

To circumvent these difficulties, the concept of localization limiters has been proposed [3, 4]. The essential idea of these limiters is to change the character of the equations so that the region of localization does not degenerate to a set of measure zero. The limiters proposed in [3] and [4] were respectively of two distinct types: integral limiters based on nonlocal constitutive equations and differential limiters based on higher-order derivatives of the strain.

One purpose of this paper is to present analyses of the governing equations with and without limiters in one dimension and in the case of antiplane motion in two dimensions. It is shown that without limiters, the static equations lose ellipticity for strain-softening materials and nonassociated plastic laws, whereas the dynamic equations lose strict hyperbolicity. With the gradient-type localization limiter, the dynamic equations change from hyperbolic to parabolic, which introduces a length scale.

It is also shown in this paper that the two types of localization limiters, differential and integral, possess very similar characteristics. Both limiters (1) exhibit

a stable response to short-wavelength input and unstable response to long wavelengths, and (2) limit the localization to a width dependent strictly on the length parameter. The length of the localization zone is not sensitive to changes in the perturbation: for static solutions, numerical results show that the gradient limiter yields the same localization bandwidth for different scales of imperfection. It is noted that even with the limiter the discrete tangent stiffness does not maintain positive definiteness and the numerical difficulties associated with strain-softening in local materials also appear.

The paper is organized as follows: Sec. 2 deals with the relationship between localization and change of type in the governing equations. Section 3 classifies the different localization limiters, and in Sec. 4, a gradient approach to limiters for static problems and its numerical implementation is described. In Sec. 5, solutions are given for simple problems and the effect of the scale of the imperfection is studied.

2. CONDITION FOR LOCALIZATION AND CHANGE OF TYPE IN THE GOVERNING EQUATIONS

To understand better the difficulties associated with the localization phenomenon and the role of the gradient localization limiter, the relation between the onset of localization and a change of type of the governing equations is investigated.

2.1. *Change of type in statics and dynamics*

We will derive for a simple problem the condition for the onset of localization in statics and dynamics and relate it to the type of the system of PDEs governing the problem.

For the general three-dimensional case, the equations of motion are written

$$\sigma_{ij,i} + \mathbf{b}_j = \rho \mathbf{v}_{j,t}, \tag{2.1}$$

where σ is the Cauchy stress tensor, $\mathbf{v}$ the velocity vector, ρ the mass density and $\mathbf{b}$ the vector of body forces. Subscript indices preceded by a comma denote partial derivatives. The body forces intervene in the governing equation only as a forcing term, so we can omit them in the study of the character of the equations.

The constitutive law relating the stress and strain rates is written

$$\dot{\sigma}_{ij} = C_{ijkl}\dot{\epsilon}_{kl}, \tag{2.2}$$

where the tensor C_{ijkl} has minor symmetries $C_{ijkl} = C_{jikl} = C_{ijlk}$; ϵ is the strain tensor defined as

$$\epsilon_{kl} = \tfrac{1}{2}(u_{k,l} + u_{l,k}). \tag{2.3}$$

The equation will be analyzed for a simple antiplane shear problem, *but the main results remain* valid for the general three-dimensional case. The antiplane shear problem has been studied for a class of incompressible hyperelastic materials by Knowles [7] in statics and by Freund *et al.* [8] and Toulios [9] in dynamics. For this problem, the displacement and stress fields are as follows:

$$u_1 = u_2 = 0, \quad u_3 = u_3(x_1, x_2),$$

$$\sigma_{13} = \sigma_{13}(x_1, x_2), \quad \sigma_{23} = \sigma_{23}(x_1, x_2). \tag{2.4}$$

In statics, the equilibrium eqns (2.1) reduce to

$$\sigma_{13,1} + \sigma_{23,2} = 0. \tag{2.5}$$

The constitutive law reads

$$\dot{\sigma}_{13} = 2C_{1313}\dot{\epsilon}_{13} + 2C_{1323}\dot{\epsilon}_{23}$$

$$\dot{\sigma}_{23} = 2C_{2313}\dot{\epsilon}_{13} + 2C_{2323}\dot{\epsilon}_{23}. \tag{2.6}$$

We can make the simplifying assumption that stress is a single-valued function of strain. This holds for elastic–plastic laws as long as there is no unloading at any point. We can write

$$\sigma_{13,1} = \frac{\hat{c}\sigma_{13}}{\hat{c}\epsilon_{13}}\frac{\hat{c}\epsilon_{13}}{\hat{c}x_1} + \frac{\hat{c}\sigma_{13}}{\hat{c}\epsilon_{23}}\frac{\hat{c}\epsilon_{23}}{\hat{c}x_1}$$

$$= 2C_{1313}\epsilon_{13,1} + 2C_{1323}\epsilon_{23,1}, \tag{2.7}$$

and a similar relation for $\sigma_{23,2}$. We look for solutions that have discontinuities in $\epsilon_{\alpha\beta,\delta}$ along a line Γ defined by its local normal $\mathbf{n} = (n_1, n_2, 0)$. The tangent vector Γ at the current point is $\mathbf{s} = (s_1, s_2) = (-n_2, n_1, 0)$.

The governing equations along Γ (equilibrium, compatibility, directional derivatives) can be cast in a matrix form:

$$\begin{bmatrix} C_{1313} & C_{1323} & C_{2313} & C_{2323} \\ 0 & 1 & -1 & 0 \\ s_1 & s_2 & 0 & 0 \\ 0 & 0 & s_1 & s_2 \end{bmatrix} \begin{bmatrix} \epsilon_{13,1} \\ \epsilon_{13,2} \\ \epsilon_{23,1} \\ \epsilon_{23,2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \epsilon_{13,\delta} \\ \epsilon_{23,\delta} \end{bmatrix}. \tag{2.8}$$

This relation is of the form $\mathbf{A}\epsilon^D = \mathbf{c}$; for ϵ^D not to be unique, we require

$$\det \mathbf{A} = 0, \tag{2.9}$$

which yields in this case

$$-s_1^2 C_{2323} + s_1 s_2 (C_{1323} + C_{2313}) - s_2^2 C_{1313} = 0, \tag{2.10}$$

or in terms of the normal vector $\mathbf{n}$:

$$n_2^2 C_{2323} + n_1 n_2 (C_{1323} + C_{2313}) + n_1^2 C_{1313} = 0. \tag{2.11}$$

The above can be written

$$\det(n_i C_{ijkl} n_l) = \det(\mathbf{nCn}) = 0, \tag{2.12}$$

which is the classical localization condition [5, 10].

The loss of uniqueness corresponds to the loss of ellipticity of the governing equations, or in other words, to the appearance of real characteristics which are associated with equations of a hyperbolic type.

We focus now on t.. dynamic case. The equation of motion for the antiplane problem is

$$\sigma_{13,1} + \sigma_{23,2} = \rho v_{3,t}. \tag{2.13}$$

The cross-derivative relations:

$$\epsilon_{13,t} = v_{3,1}, \quad \epsilon_{23,t} = v_{3,2} \tag{2.14}$$

are combined with the equation of motion to yield a system of the first-order PDEs:

$$\begin{bmatrix} 0 & C_{1313} & C_{1323} \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_3 \\ \epsilon_{13} \\ \epsilon_{23} \end{bmatrix}_{,1} + \begin{bmatrix} 0 & C_{2313} & C_{2323} \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_3 \\ \epsilon_{13} \\ \epsilon_{23} \end{bmatrix}_{,2} + \begin{bmatrix} -\rho & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_3 \\ \epsilon_{13} \\ \epsilon_{23} \end{bmatrix}_{,t} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{2.15}$$

This system is of the form:

$$\mathbf{A}^1\mathbf{U}_{,1} + \mathbf{A}^2\mathbf{U}_{,2} + \mathbf{A}^t\mathbf{U}_{,t} = 0, \tag{2.16}$$

where $\mathbf{A}^t$ is nonsingular. The condition for $\Phi(x_1, x_2, t) = 0$ to be a characteristic surface of (2.15) is [11]:

$$\det(\mathbf{A}) = 0, \tag{2.17a}$$

where

$$\mathbf{A} = \mathbf{A}^1\Phi_{,1} + \mathbf{A}^2\Phi_{,2} + \mathbf{A}^t\Phi_{,t}, \tag{2.17b}$$

which yields here:

$$\Phi_{,t}[-\rho\Phi_{,t}^2 + (C_{1323} + C_{2313})\Phi_{,1}\Phi_{,2} + C_{1313}\Phi_{,1}^2 + C_{2323}\Phi_{,2}^2] = 0. \tag{2.18}$$

The extra factor $\Phi_{,t}$ in (2.18) corresponds to a characteristic surface with zero velocity and is a result of introducing an additional dependent variable by choosing strains and velocity as the dependent variables [12]. To understand better the meaning of (2.18), we define the constitutive matrix $\mathbf{D}$ such that

$$D_{kl} = C_{k3l3} \tag{2.19}$$

and select a new coordinate system $(\bar{1}, \bar{2})$ defined by the principal directions of $\mathbf{D}$, so that in the new coordinate system:

$$\mathbf{D} = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}. \tag{2.20}$$

Equation (2.18) can thus be written as

$$-\rho\Phi_{,t}^2 + D_1\Phi_{,\bar{1}}^2 + D_2\Phi_{,\bar{2}}^2 = 0. \tag{2.21}$$

Characteristic surfaces are cones of elliptic section, as illustrated in Fig. 1; the equation of the cone passing through a point $(\bar{x}_0, \bar{y}_0, t_0)$ is

$$(t - t_0)^2 = \frac{1}{c_1^2}(\bar{x} - \bar{x}_0)^2 + \frac{1}{c_2^2}(\bar{y} - \bar{y}_0)^2, \tag{2.22}$$

where

$$c_1 = \sqrt{\left(\frac{D_1}{\rho}\right)} \quad \text{and} \quad c_2 = \sqrt{\left(\frac{D_2}{\rho}\right)}. \tag{2.23}$$

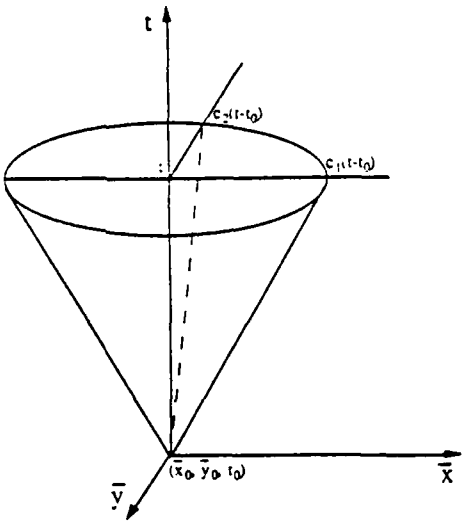


Fig. 1. Characteristic cone for the dynamic equations.

As $\mathbf{D}$ loses positive definiteness, say, for example, D_1 remains strictly positive and D_2 approaches zero, the cone collapses to a plane surface. Considered as a function of the variables $(\bar{y}, t)$, the system loses strict hyperbolicity, or, equivalently, real waves no longer propagate in every direction (in our case they stop propagating in the $\bar{y}$-direction). It is therefore seen that here the condition of strict hyperbolicity of the system of governing equations and the condition of strong ellipticity are equivalent.

It should be pointed out that when a viscoplastic constitutive law is used, the equations of motion do not lose hyperbolicity. This is readily seen by observing that for viscoplastic models the rate constitutive relation is written as

$$\dot{\sigma}_{ij} = C^t_{ijkl}\dot{\epsilon}_{kl} - R_{ij}(\sigma), \tag{2.24}$$

where C^e_{ijkl} is the elastic tensor and the inelastic part is embedded in the term R_{ij}. The type of the system of governing equations is determined by C^e_{ijkl}, so that it remains strictly hyperbolic, the inelastic effects appearing only as a forcing term.

2.2. *Relation between strain-softening and localization for elasto-plastic materials*

The rate constitutive relations for an elasto-plastic material are written in tensor form as:

$$\dot{\sigma} = C : \dot{\epsilon} = C^e : \dot{\epsilon} - \frac{\mathbf{C}^e : \mathbf{P}}{h + \mathbf{Q} : \mathbf{C}^e : \mathbf{P}} \mathbf{Q} : \mathbf{C}^e : \dot{\epsilon}, \quad (2.25)$$

where $\mathbf{P}$ and $\mathbf{Q}$ are symmetric first-order tensors giving respectively the direction of the plastic deformation and the outer normal to the yield surface, h is the rate of hardening, and $\mathbf{C}^e$ is the elasticity tensor:

$$C^e_{ijkl} = \lambda \delta_{ij} \delta_{kl} + G(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}), \quad (2.26)$$

and λ and G are Lamé's constants. The case $\mathbf{P} = \mathbf{Q}$ corresponds to plastic normality; $\mathbf{P} \neq \mathbf{Q}$ corresponds to a nonassociative flow rule.

We consider again the antiplane shear problem and focus on the relation between the localization condition and the strain-hardening modulus in the case of an elasto-plastic material model. For localization to occur on a plane of normal $\mathbf{n}$, condition (2.12) has to be met. We can write that condition in a set of *Cartesian axes* $\mathbf{n}$, $\mathbf{e}_3 \times \mathbf{n}$ *and* $\mathbf{e}_3$ ($\mathbf{e}_3$ *is the unit vector* in the '3' direction). With subscripts denoting components in that set of axes, and for the constitutive law, (2.25) and (2.26), the localization condition reduces to

$$C_{1313} = 2\left(G - \frac{4G^2 P_{13}}{h + 4P_{13} 2G Q_{13}} Q_{13}\right) = 0, \quad (2.27)$$

or equivalently

$$\frac{h}{G} = -2(2P_{13}Q_{13}). \quad (2.28)$$

This expression shows that if plastic normality holds (i.e. $\mathbf{P} = \mathbf{Q}$), then localization can only occur with negative h, that is in a strain-softening regime, whereas if normality does not apply, it is possible for localization to be triggered with a positive h.

This result, obtained for the particular case of the antiplane shear problem, is in fact general, as was shown by Rudnicki and Rice [13] and Rice [5]. In the three-dimensional case, (2.28) can be generated to (with α, β denoting components on Cartesian axes in the plane of localization)

$$\frac{h}{G} = -2P_{\alpha\beta}Q_{\alpha\beta} - \frac{2\lambda}{\lambda + 2G} P_{\alpha\alpha} Q_{\beta\beta}, \quad (2.29)$$

and the conclusions derived previously remain valid. An example of a material model where localization occurs for positive h can be found in [13].

3. LOCALIZATION LIMITERS

Localization limiters can be classified as follows:

(1) nonlocal or integral limiters, where the strain measure includes an integral of the deformation over a finite domain [3];

(2) differential limiters, where the strain or stress measures include derivatives of order higher than one [4, 14–17]; and

(3) rate limiters, where a time dependence is built into the equation [18].

The rationale underlying the nonlocal limiters is that a classical local theory does not take into account the influence of the length scale associated with a rapidly varying strain field on the stress distribution, an essential part of localization.

In the case of a one-dimensional rod with strain-softening, a nonlocal limiter is obtained by defining the stress field $\sigma(x)$ as a function of a nonlocal strain $\bar{\epsilon}(x)$ [3]:

$$\sigma(x) = \sigma[\bar{\epsilon}(x)] \quad (3.1)$$

with

$$\bar{\epsilon}(x) = \frac{1}{\ell} \int_{-\ell/2}^{\ell/2} \epsilon(x + s) w(s)\, ds, \quad (3.2)$$

where $[x - (\ell/2), x + (\ell/2)]$ is a domain around x, and $w(s)$ is a weighting function. For simplicity, we assume in what follows a uniform weighting function $w(s) = 1/\ell$.

The gradient-type limiter in a one-dimensional context is given by [4]

$$\bar{\epsilon}(x) = \epsilon(x) + \alpha \epsilon_{,xx}(x). \quad (3.3)$$

These two limiters are related through a Taylor expansion [4] and can actually differ by a function of order $o(\ell^4)$, provided that

$$\alpha = \frac{\ell^2}{24}. \quad (3.4)$$

In dynamic problems, the effect of the differential and rate limiters from a mathematical point of view is that the governing equations no longer become elliptic with the onset of strain-softening. This can be seen in a one-dimensional context for a path independent material, by combining the equation of motion and the compatibility condition into a system of first-order partial differential equations:

$$\begin{bmatrix} v \\ \epsilon \end{bmatrix}_{,t} + \begin{bmatrix} 0 & \dfrac{\sigma'(\epsilon)}{\rho} \\ -1 & 0 \end{bmatrix} \begin{bmatrix} v \\ \epsilon \end{bmatrix}_{,x} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (3.5)$$

where v, $\sigma(\epsilon)$, ϵ and ρ are respectively the particle velocity, stress, strain and density, and subscript comma denotes a partial derivative. This system is of the type

$$\mathbf{A}\mathbf{U}_{,t} + \mathbf{B}\mathbf{U}_{,x} = \mathbf{c}(\mathbf{U}), \tag{3.6}$$

where one of the matrices A or B, e.g. A, is nonsingular and c is a forcing vector. The mathematical nature of (3.6) is determined by the roots of the characteristic determinant $\det(\mathbf{B} - \lambda\mathbf{A})$ [or $\det(\mathbf{A} - \lambda\mathbf{B})$ if A is singular]. In the case of eqn (3.5),

$$\det(\mathbf{B} - \lambda\mathbf{A}) = \det(\mathbf{B} - \lambda\mathbf{I}) = \lambda^2 - \frac{\sigma'(\epsilon)}{\rho}, \tag{3.7}$$

so that the system becomes elliptic when $\sigma'(\epsilon) < 0$ (which corresponds to strain-softening) because the determinant (3.7) no longer possesses real roots.

When the differential limiter defined in (3.3) is included in the formulation, a modified system of PDEs is obtained:

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \epsilon \\ w_1 \\ w_2 \\ v \end{bmatrix}_{,t} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ \dfrac{\sigma'(\epsilon)}{\rho} & 0 & \dfrac{\alpha\sigma'(\epsilon)}{\rho} & 0 \end{bmatrix} \begin{bmatrix} \epsilon \\ w_1 \\ w_2 \\ v \end{bmatrix}_{,x} = \begin{bmatrix} w_1 \\ 0 \\ w_2 \\ 0 \end{bmatrix}, \tag{3.8}$$

where $w_1 = \epsilon_{,x}$ and $w_2 = \epsilon_{,xx}$. The characteristic determinant

$$\det(\mathbf{A} - \lambda\mathbf{B}) = -\frac{\alpha\sigma'(\epsilon)}{\rho}\lambda^4 \tag{3.9}$$

possesses four real roots, all equal, irrespective of the sign of $\sigma'(\epsilon)$, so that the system is parabolic. It should be pointed out that, when a rate-type limiter is used via a viscoplastic material model, the governing equations remain hyperbolic [18].

To understand better the behavior of the integral and differential limiters, a Fourier analysis by the method of frozen coefficients is useful. In this analysis a displacement disturbance δu is applied to the body, and the material is considered to be in a strain-softening state over an interval $[x_1, x_2]$:

$$\delta\sigma(x) = -|E_t|\delta\epsilon(x) \quad \text{for } x \text{ in } [x_1, x_2], \tag{3.10}$$

when the tangent modulus $E_t < 0$ is assumed constant. We then look for positive wave solutions of the form

$$\delta u(x, t) = A e^{ik(vt - x)} \tag{3.11}$$

for the equation of motion for δu:

$$\delta u_{,tt} - \frac{|E_t|}{\rho}\delta\epsilon^{\text{nonloc}}(x)_{,x} = 0. \tag{3.12}$$

The following dispersion relation was obtained in [4] with the gradient-type limiter $\bar{\epsilon}$ defined by eqn (3.3):

$$kv = i\left\{\frac{|E_t|}{\rho}(1 - \alpha k^2)\right\}^{1/2} \quad k = i\bar{\gamma}(k). \tag{3.13}$$

A similar analysis can be done for the integral limiter $\epsilon^{\text{nonloc}}(x) = \bar{\epsilon}(x)$ defined in eqn (3.2). Looking for wave solutions of the equation of motion

$$\delta u_{,tt} + \frac{|E_t|}{\rho}\frac{1}{\ell}\left[\int_{-\ell/2}^{\ell/2}\delta\epsilon(x + s)\,ds\right]_{,x} = 0, \tag{3.14}$$

the following relation is obtained:

$$kv = i\left\{\frac{|E_t|}{\rho}\frac{2k}{\ell}\sin\frac{k\ell}{2}\right\}^{1/2} k = i\bar{\gamma}(k). \tag{3.15}$$

The two functions $\bar{\gamma}(x)$ and $\bar{\gamma}(k)$ are plotted in Fig. 2, for values of α and ℓ related through eqn (3.4). In [4] the plot of $\bar{\gamma}(k)$ was interpreted to mean that the growth in short-wavelength inputs susceptible to develop in the narrow localization zones is bounded when the limiter is present.

It is interesting to notice that for small values of α, the expression of $\bar{\gamma}(k)$ can be expanded, and, using (3.4),

$$\bar{\gamma}(k) \approx \left\{\frac{|E_t|}{\rho}\frac{2k}{\ell}\left[\frac{k\ell}{2} - \frac{1}{3!}\left(\frac{k\ell}{2}\right)^3\right]\right\}^{1/2}$$

$$\approx \left\{\frac{|E_t|}{\rho}\frac{k}{\sqrt{(6\alpha)}}k\sqrt{6\alpha(1 - \alpha k^2)}\right\}^{1/2} = \tilde{\gamma}(k). \tag{3.16}$$

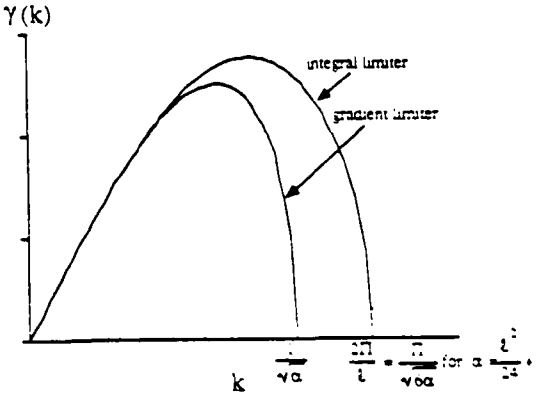


Fig. 2. Linearized analysis: $\bar{\gamma}(x)$ (gradient-type limiter) and $\tilde{\gamma}(k)$ (integral-type limiter) vs wavelength k.

A perturbation analysis [4] reveals that, when the differential limiter defined in (3.3) is used, the width of the zone in the strain-softening regime varies with the square root of the parameter α. Numerical simulations confirmed this type of dependence, and for the integral limiter they yield a zone size proportional to the averaging length ℓ [3], as can be expected from relation (3.4). Thus both localization limiters prevent the growth of waves of the scale of the localization bands which are generated by the presence of strain-softening.

As far as static problems are concerned, the only attempts to derive closed-form solutions using limiters known to the author are due to Schreyer and Chen [14], Coleman and Hodgdon [15] and Aifantis and co-workers [16, 17]. In the former approach, higher-order terms are included in the evolution equation of the flow stress, and in that sense it is quite similar to the work of Schreyer and Chen. In Coleman and Hodgdon [15], a second-order strain gradient is added directly into the constitutive equation without modifying the yield function. The common feature of these approaches is that they make the stress field dependent in some way on the spatial derivatives of the strain field. In the next section, we choose the method presented in [15] to develop a finite element formulation.

4. GRADIENT APPROACH TO LOCALIZATION IN STATICS

We follow the approach of Coleman and Hodgdon [15], but do not limit our formulation to rigid plastic materials. The expression for the stress is given by

$$\sigma = \phi(\epsilon) - \alpha\nabla^2\epsilon, \tag{4.1}$$

where $\phi(\epsilon)$ is the usual elastoplastic constitutive law (stress–strain relationship) and $\alpha > 0$ is a coefficient having the dimensions of a force.

To use the finite element method, a weak (variational) form of the equilibrium equation is written

$$I = \int_0^L \delta\epsilon\sigma(x)\,dx - \delta\mathbf{d}^T\mathbf{f}^{ext} = 0. \tag{4.2}$$

The displacement field is approximated by

$$u(x, t) = \mathbf{N}(x)\mathbf{d}(t), \tag{4.3}$$

where $\mathbf{d}$ is the vector of nodal displacements. Equation (4.2) is a nonlinear equation resulting from the nonlinear constitutive relation. To develop a tangent stiffness matrix, we consider a linearization of (4.2), assuming that the external forces $\mathbf{f}^{ext}$ are independent of the current displacement $\mathbf{d}$:

$$\begin{aligned}\Delta I &= \int_0^L \delta\epsilon\ \Delta\sigma(x)\,dx \\ &= \int_0^L \delta\epsilon\ \Delta(\phi(\epsilon) - \alpha\epsilon_{,xx})\,dx \\ &= \int_0^L \delta\epsilon C_T\epsilon - \alpha\Delta(\epsilon_{,xx})\,dx,\end{aligned} \tag{4.4}$$

where C_T is the tangent (elastoplastic) modulus. The introduction of the second-order derivative in the strain expression would make impossible the use of standard finite elements with a linear interpolation for the displacement field. However, to examine the performance of the method, we use simple linear displacement finite elements and compute the second derivative of the strain in element e by a centered finite difference approximation. For a one-dimensional problem, this results in the following expression when uniform (constant Δx) meshes are used:

$$\begin{aligned}(\epsilon_{,xx})_e &= \frac{\epsilon_{e+1} - 2\epsilon_e + \epsilon_{e-1}}{\Delta x^2} \\ &= \frac{-d_{e-1,1} + 3d_{e,1} - 3d_{e,2} + d_{e+1,2}}{\Delta x^3},\end{aligned} \tag{4.5}$$

where $d_{e,i}$ for $i = 1, 2$ are the nodal displacements for element e.

For the formulation of the elementary tangent stiffness, this results in adding to the usual (2×2) element tangent stiffness an enlarged (2×4) matrix derived from (4.5)

$$\begin{aligned}&\int_0^{l_e} \delta\epsilon\ \Delta\sigma(x)\,dx \\ &= \int_0^{l_e} \delta\epsilon[C_T\epsilon - \alpha\Delta(\epsilon_{,xx})]\,dx \\ &= [\delta d_{e,1}\ \delta d_{e,2}] \int_0^{l_e} \left(\mathbf{B}^T C_T \mathbf{B} \begin{bmatrix}\Delta d_{e,1} \\ \Delta d_{e,2}\end{bmatrix} \right. \\ &\quad \left. - \alpha\mathbf{B}^T[-1 \quad 3 \quad -3 \quad 1] \begin{bmatrix}\Delta d_{e-1,1} \\ \Delta d_{e,1} \\ \Delta d_{e,2} \\ \Delta d_{e+1,2}\end{bmatrix} \right) dx \\ &= \delta\mathbf{d}^T\mathbf{K}_T^e[\Delta d_{e-1,1}\ \Delta d_{e,1}\ \Delta d_{e,2}\ \Delta d_{e+1,2}]^T,\end{aligned} \tag{4.6a}$$

where

$$\begin{aligned}\mathbf{K}_T^e = \int_0^{l_e} \mathbf{B}^T(&C_T[0 \quad \mathbf{B}_{1,2} \quad 0] \\ &- \alpha[-1 \quad 3 \quad -3 \quad 1])\,dx.\end{aligned} \tag{4.6b}$$

In the previous expression, $\mathbf{B}$ is the gradient matrix, its dimensions are 1×2 for two-node constant-strain elements in one dimension. The discrete equilibrium equations are then deduced from

$$I + \Delta I = 0. \tag{4.7}$$

which can be written, after assembly of the elementary tangent stiffness matrices $\mathbf{K}_T^e$ into a global tangent stiffness matrix $\mathbf{K}_T$,

$$\delta \mathbf{d}^T \mathbf{K}_T \Delta \mathbf{d} + \delta \mathbf{d}^T \int_0^{l_e} \mathbf{B}^T \sigma(x)\, dx - \delta \mathbf{d}^T \mathbf{f}^{ext} = 0$$

for every $\delta \mathbf{d}$, (4.8)

so that the iterative displacement vector $\Delta \mathbf{d}$ is the solution of a linear system of equations:

$$\mathbf{K}_T \Delta \mathbf{d} = \mathbf{f}^{ext} - \mathbf{f}^{int}, \tag{4.9}$$

where

$$\mathbf{f}^{int} = \int_0^{l_e} \mathbf{B}^T \sigma(x)\, dx. \tag{4.10}$$

As can be noticed from the expression of $\mathbf{K}_T^e$ in eqn (4.6b), the tangent stiffness matrix associated with the system is not self-adjoint and, moreover, it is not in general positive definite. When C_T loses positive definiteness, this result is expected since the governing system of equations even with the limiter does not possess a unique solution. This is readily verified by noting that, when the prescribed displacement d at the two ends of a perfect rod of length $2L$ is bigger than $L\epsilon_m$ (see Fig. 4), the homogeneous strain field d/L still satisfies the relevant equations (4.2), so at least two solutions exist. This is also true for an integral limiter.

5. NUMERICAL EXAMPLES

When conducting the numerical simulations, it was noticed that the introduction of the localization did not remove completely all the unpleasant features present in calculations involving strain-softening materials. More precisely, when the strain-softening regime is incipient, the Newton–Raphson procedure often results in iterations that oscillate between two or more states and fail to converge to one equilibrium state. From a numerical point of view, this is linked to the tangent stiffness $\mathbf{K}_T$ not remaining positive definite. In [19], a remedy for this difficulty was proposed; the problem is posed as the minimization of the length of the residual vector:

$$\text{Minimize: } \mathscr{F} = \mathbf{r}^T(\mathbf{d})\mathbf{r}(\mathbf{d}), \tag{5.1}$$

where

$$\mathbf{r}(\mathbf{d}) = \mathbf{f}^{ext} - \mathbf{f}^{int}, \tag{5.2}$$

and requires $\mathscr{F} = 0$ at the minimum.

This provides a more well-behaved problem for the line-search procedure, and the rate of convergence of the Newton method is improved substantially. The method was also adapted in [19], to combine it with arc-length procedures.

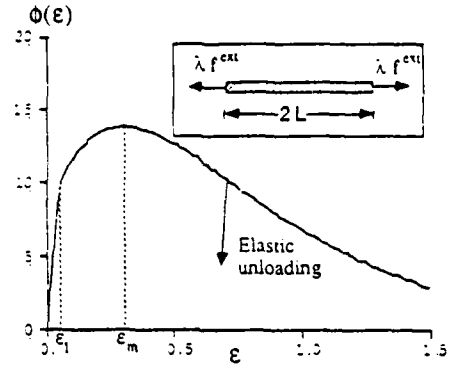


Fig. 3. Traction curve $\epsilon - \phi(\epsilon)$. Young's modulus $E = 200$, yield strain $\epsilon_1 = 0.05$, $\epsilon_m = 0.3$, parameter controlling the convexity $\delta_0 = 0.11$. In enclosed box, problem description: rod length $2L = 10$, 40 nodal points.

To test the effectiveness of the localization limiter and the solution strategy, we consider the problem of a one-dimensional rod, subjected to equal and opposite loading at its two ends, as illustrated in Fig. 3. One node in the mesh is held fixed, to prevent rigid body translations. To trigger the appearance of a nonhomogeneous strain distribution, a small imperfection is introduced. In the present example, this was accomplished by making the cross-section of the center element 1% smaller than the cross-section of all other elements.

Numerical studies of this problem were conducted based on the localization limiter defined previously in eqn (4.1), which in one-dimension reduces to

$$\sigma = \phi(\epsilon) - \alpha \epsilon_{,xx}. \tag{5.3}$$

The elasto-plastic strain–stress law considered is illustrated in Fig. 3. It consists of a linear elastic part, and an exponential branch including a strain hardening portion followed by a softening one. At any point the unloading is elastic with Young's modulus E. The physical parameters used for the calculation were: Young's modulus $E = 200$, yield strain $\epsilon_1 = 0.05$, $\epsilon_m = 0.3$, exponential branch:

$$\phi(\epsilon) = E\epsilon_1 \frac{(\epsilon + \delta_0)}{(\epsilon_1 + \delta_0)} \exp\left[g(\epsilon) - g(\epsilon_1)\right], \tag{5.4a}$$

where

$$g(\epsilon) = \left(1 - \frac{\epsilon + \delta_0}{\epsilon_m + \delta_0}\right). \tag{5.5b}$$

The parameter δ_0 controlling the convexity is taken to be equal to 0.11.

To solve this problem, the line-search technique combined with the arc-length method with a linearized constraint equation described in [19] was used. It was first checked that without introducing the localization limiter, that is, for $\alpha = 0$, the deformation localized in the element with imperfection, irrespective of its size, whereas all other elements unload

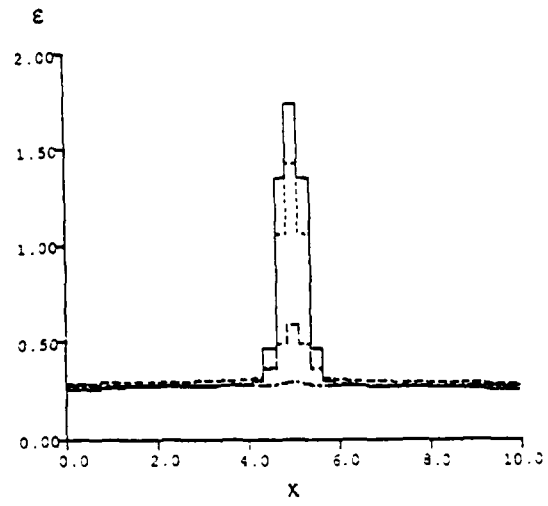


Fig. 4. Strain vs spatial coordinate profiles at four different load steps with higher-order term limiter ($\alpha = 0.5$).

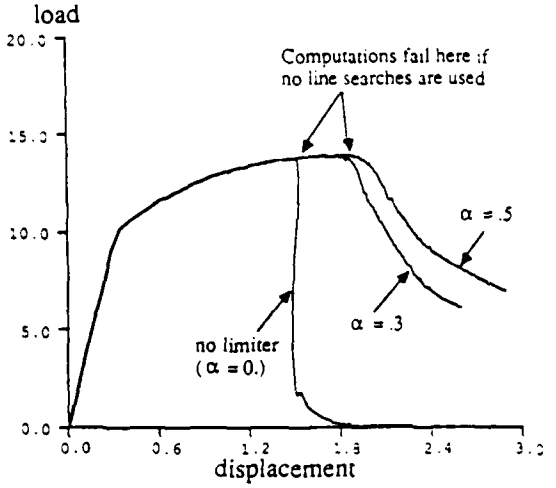


Fig. 5. Load–displacement curves for different values of the parameter α.

elastically. In a load–displacement curve, a sharp decrease is observed once strain-softening is attained, and even a snap-back behavior can be observed, which could not be captured with a pure displacement control strategy.

Calculations were then conducted with the localization limiter defined in eqn (5.3), for several values of the parameter α. The strain distribution along the rod for various load levels is given in Fig. 4. These strain-profiles are very close in shape to those obtained by Coleman and Hodgdon [15] in their study of the effects of the localization limiter (5.3) on strain-localization for a rigid plastic material with a parabolic law. Essentially, a finite localization zone emerges, practically constant in size, in which the strain increases but remains bounded, whereas in the rest of the rod, the material unloads elastically. In the finite element calculation, this localized zone spans a few elements of the mesh. The size of the zone is directly related to the value of α.

Load–displacement curves for various values of α are reported in Fig. 5. They exhibit a milder negative slope with increasing α. It should be pointed out that, without the use of the line procedure described above, the Newton–Raphson procedure fails to converge near the critical point.

As pointed out previously, a small imperfection is introduced in the center of the rod to create a nonhomogenous strain distribution. The sensitivity of the results to the imperfection is reported in Fig. 6, for several values of α. Imperfections of different lengths ℓ_{imp} are obtained by giving a smaller cross-section to several elements around the center of the rod. In Fig. 7, strain profiles for a given value of α and corresponding to the same end displacement are reported for different values of ℓ_{imp}. It can be observed from Figs 6 and 7 that the size of the localization zone is a function of α but is practically insensitive to changes in the imperfection size.

6. CONCLUSIONS

Localization limiters can be classified as nonlocal, differential and rate limiters. A Fourier analysis of the

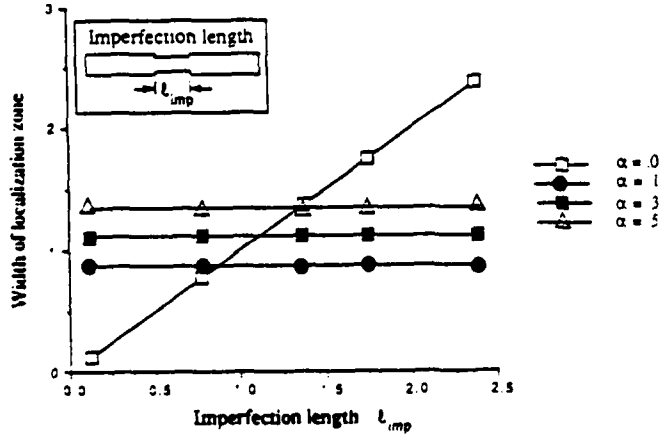


Fig. 6. Sensitivity of the band size to imperfection length ℓ_{imp}. For all points mesh of 80 elements, imperfection magnitude 2%.

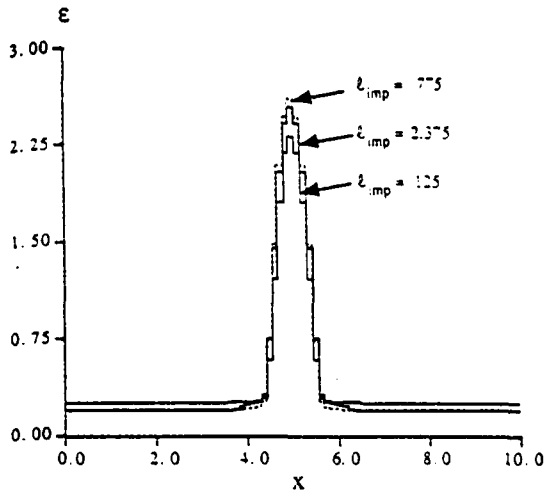


Fig. 7. Strain profiles for different values of the imperfection length ℓ_{imp} for the same end displacement. Mesh of 80 elements, localization limiting parameter $\alpha = 0.3$.

wave-propagation problem shows that the introduction of nonlocal or differential limiters lead to governing equations where short waves, which are likely to develop with the onset of strain-softening, have a bounded growth. In dynamic problems, strain-softening causes the governing equations to lose strict hyperbolicity; it was shown, for example, that they have become elliptic in at least one direction for the antiplane problem. With the gradient-type localization limiter, the dynamic equations change from hyperbolic to parabolic for the one-dimensional case. The character of the amplification spectrum of the integral and differential limiters is similar and they become identical in the limit as the magnitude of the parameter governing the limiter goes to zero.

The differential localization limiter proposed by Coleman and Hodgdon [15], based on the introduction of the second derivative of the strain in the stress expression, was implemented in the context of static problems. Numerical studies show that it allows for the development of a localized strain zone spanning several elements of the mesh. However, the addition of the limiter does not guarantee positive definiteness of the tangent matrix.

The differential limiter is not sensitive to the magnitude of the imperfection which triggers localization. For this reason, differential limiters are most suited for modeling phenomena where the size of the localization bands is relatively consistent. In fact, the localization bandwidth is governed by the α parameter, the coefficient of the higher-order term. This provides a phenomenological way of choosing α on the basis of experimental results.

Acknowledgements—The support of the Office of Naval Research Under Award N00014-85-K-0458 is gratefully acknowledged. The support of the Air Force Office of Scientific Research under contract no. F49620-87-C-0030 and the support of the Army Reseach Office under grant no. DAAL03-87-K-0035 to Northwestern University are also gratefully acknowledged.

REFERENCES

1. Z. P. Bazant and T. Belytschko, Wave propagation in strain-softening bar: exact solution. *J. Engng Mech. Div., ASCE* **111**, 381–389 (1985).
2. F. H. Wu and L. B. Freund, Deformation trapping due to thermoplastic instability in one-dimensional wave propagation. *J. Mech. Phys. Solids* **32**, 119–132 (1984).
3. Z. P. Bazant, T. Belytschko and T. P. Chang, Continuum theory for strain-softening. *J. Engng Mech. Div., ASCE* **110**, 1666–1692 (1984).
4. D. Lasry and T. Belytschko, Localization limiters in transient problems. *Int. J. Solids Struct.* **24**, 581–597 (1988).
5. J. R. Rice, The localization of plastic deformation. In *Theoretical and Applied Mechanics, Proc. 14th Int Congr. Theoret. Appl. Mech.* (Edited by W. T. Koiter), pp. 207–220. North-Holland, Amsterdam (1977).
6. T. Belytschko, J. Fish and B. E. Engelmann, A finite element with embedded localization zones. *Comput. Meth. appl. Mech. Engng* **70**, 59–89 (1988).
7. J. K. Knowles, The finite antiplane shear field near the tip of a crack for a class of incompressible elastic solids. *Int. J. Fract.* **13**, 611–639 (1977).
8. L. B. Freund, F. H. Wu and M. Toulios, Initiation and propagation of shear band in antiplane shear deformation. *Proc. Int. Symp. Plastic Instability*, Considère Memorial, Paris, Sept. 1985, pp. 125–134. Presses de l'Ecole Nationale des Ponts et Chaussées, Paris (1985).
9. M. Toulios, The wavefront induced in a homogeneously shearing solid by a localized material imperfection. *Q. J. appl. Math.* **43**, 225–235 (1985).
10. R. Hill, Acceleration waves in solids. *J. Mech. Phys. Solids* **10**, 1–6 (1962).
11. R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. II. Interscience, New York (1962).
12. R. J. Clifton, A difference method for plane problems in dynamic elasticity. *Q. J. appl. Math.* **25**, 97–116 (1967).
13. J. W. Rudnicki and J. R. Rice, Conditions for the localization of deformation in pressure sensitive dilatant materials. *J. Mech. Phys. Solids* **23**, 371–394 (1975).
14. H. L. Schreyer and Z. Chen, The effect of localization on the softening behavior of structural members. In *Constitutive Equations: Macro and Computational Aspects* (Edited by K. J. William), pp. 193–203. ASME, New York (1984).
15. B. D. Coleman and M. L. Hodgdon, On shear bands in ductile materials. *Archs Rational Mech. Anal.* **90**, 219–247 (1985).
16. N. Triantafyllidis and E. C. Aifantis, A gradient approach to localization of deformation. I. Hyperelastic materials. *J. Elast.* **16**, 225–237 (1986).
17. H. M. Zbib and E. C. Aifantis, On the post-localization of plastic deformation. I. The evolution of shear bands in plastic materials. Michigan Institute of Technology, MM Rep 14 (1987).
18. A. Needleman, Material rate dependence and mesh sensitivity in localization problems. *Comput. Meth. appl. Mech. Engng* **67**, 61–85 (1988).
19. T. Belytschko, D. Lasry and T. Lodygowski, A modified Newton procedure for post-bifurcation analysis, as in strain-softening. *Int. J. Numer. Meth. Engng* (submitted).

Computers and Structures,
submitted June 1989.

A General Finite Element Procedure for Problems with High Gradients

by

J. Fish and T. Belytschko
Northwestern University
Evanston, Illinois 60208, U.S.A.

Abstract

A unified approach for treating finite element solutions with embedded, enriched and superimposed fields is presented for large deformation problems. These methods are aimed at problems with high gradients. The unified approach allows us to combine various techniques developed previously by either merging them in separate domains, or in the same domain but in different directions. Results are obtained for a viscoplastic bar with shear bands.

1. Introduction

Much attention has been devoted in the past several decades to enhancing finite element solutions in regions of high gradients. In problems such as material failure due to localization and crack propagation, or shocks in fluid dynamics, the regions of high gradients are of several orders of magnitude smaller than the size of the domain of interest. Therefore, if a uniform finite element mesh is used, tremendous numbers of elements would be needed in order to accurately resolve the structure of the fields.

In order to improve the solution, a computational model must be enriched in the regions where high gradients are indicated by the solution. One procedure for enrichment are the h- and p- adaptive [1-4]. In the h-method, the finite element mesh is refined by keeping the elements of the same order and subdividing them, while in p-method, the same mesh is retained but the order of approximation is increased; it is possible to combine both into an h-p procedure. The p-method has an advantage over the h-method especially if hierarchical [5] or spectral [6] elements are used. An advantage of hierarchical elements is that the degrees of freedom and stiffness matrix of an element of order p are the subset of those for element of polynomial order p+1. In [6] the displacement field within each spectral element was represented as a higher order Lagrangian interpolant through Tchebysheff collocation points, and an exponential rate of convergence was obtained.

While high resolution of the fields can be obtained with both hierarchical and spectral elements, many degrees of freedom would be needed when the high gradients region is skewed in relation to the mesh. The techniques developed in [7-11] take advantage of this by enriching the approximation in the direction indicated by the solution. In [7-9] the shape of the strain field is assumed to be known a priori, and therefore it is possible to obtain a good estimate of the strain field with a modest increase in the number of unknowns. In [10,11] the field was considered completely unknown and interpolants able to resolve the field with high gradients in a single direction were used. A piecewise constant strain interpolation scheme has been developed in [10], while rapidly convergent spectral interpolants were used in [11].

In this paper a unified approach, which combines the ideas presented in [7-11] is developed in context of large deformation problems. The embedded element technique developed in [7-9] has been modified, so that the methods developed in [7-11] can be treated simultaneously. The unified approach allows us to combine both formulations by either merging them in separate domains, or in the same domain but in different directions.

The outline of this paper is as follows. In section 2, the superimposed fields are described with particular emphasis on the unifying the of techniques presented in [7-11].

Section 3 describes the compatibility condition which is needed to enforce homogeneous boundary conditions on the higher order deformation field. Section 4 gives the development of the discrete equations by either the displacement method or the Hu-Washizu variational principle and the equivalence between the two variational principles is discussed. Section 5 examine the convergence properties of the spectral methods and compares them to the finite-difference method. Linearization procedures for general nonlinear analysis are presented in Section 6. Finally some numerical examples of studies of shear banding are presented in Section 7.

2. Description of the method

In this section a unified approach in constructing finite element procedures for problems with high gradients is presented. This approach allows us to treat a variety of formulations presented in [7-11] together with some attractive new possibilities in a uniform manner. The common feature to all the techniques is that the deformation field is decomposed into two parts:

$$\mathbf{u} = \mathbf{u}^0 + \mathbf{u}^H \tag{2.1}$$

where $\mathbf{u}^0 \in \Omega$ is a lower order (macro) deformation field defined over the whole problem domain and is approximated by a C^o finite element interpolants

$$u_i^0 = N_{iA}^0 \, d_A \qquad A=1,2.....NHS \qquad \text{and} \quad i=1,...NSD \tag{2.2}$$

and $\mathbf{u}^H \in \Omega_H$ is a higher order (micro) deformation field which is defined on Ω_H - the regions where high gradients are indicated by the solution (see fig.1). Here d_A are nodal displacements; lower case subscripts indicate spatial components and NSD is the total number of space dimensions; upper case subscripts A,B,C indicate nodal degrees of freedom and NDG is the total number of nodal degrees of freedom. Standard tensorial notation is used with summation over the repeated indices. Note that in [7-9] the fields are defined within each element rather than on a superimposed band.

The condition that a higher order deformation field is subject to the homogeneous boundary conditions

$$\mathbf{u}^H = 0 \quad \text{on} \quad \Gamma_{H0} \cup \Gamma_u \tag{2.3}$$

is crucial for the simplicity of the method since it allows the treatment of complex boundary conditions $\mathbf{u}^*$ by the standard C^o finite element interpolants

$$\mathbf{u}^0 = \mathbf{u}^* \quad \text{on } \Gamma_u \tag{2.4}$$

where Γ is the boundary of the problem domain Ω, which consists of the prescribed displacement boundary Γ_u and the prescribed traction boundary Γ_t. Γ_{H0} is the boundary of Ω_H, which is not included in Γ.

In the following, attention will be focussed on the construction of interpolants for the higher order deformation field. It is of advantage in the approximation of $\mathbf{u}^H$ to exploit our knowledge of the higher order strain field ε_{ij}^H. For shear band localization problems, which are used here to demonstrate the essential features of the method, the structure of the "micro" strain field can be partially determined from a local bifurcation analysis[12-14], which yields

$$\varepsilon_{ij}^H = T_{ij}(\bar{x}_2)\,\gamma(\bar{x}_1) \tag{2.5}$$

where $T_{ij}(\bar{x}_2)$ defines the components of the "micro" strain field, which can vary in the direction $\bar{x}_2$ - the local coordinate axis aligned parallel to the localization band, and $\gamma(\bar{x}_1)$ is the magnitude of the "micro" strain field with $\bar{x}_1$ being a local coordinate axis in the direction normal to the band. It is tacitly assumed that the structure in the $\bar{x}_2$ direction is less rich so that it can be adequately dealt with by the lower order strain field obtained from C^0 finite element interpolants

$$\varepsilon_{ij}^0 = N_{(iA,j)}^0 d_A = B_{ijA}^0 d_A \tag{2.6}$$

The structure of $T_{ij}(\bar{x}_2)$ was extensively discussed in [10] and is given by

$$T_{ij} = \text{Sym}\,(m_i n_j) \tag{2.7}$$

where $\mathbf{n}$ and $\mathbf{m}$ represent the direction of the band and the localized strain field defined from a bifurcation analysis [12-14].

The numerical techniques developed for the purpose of resolving deformation fields with high gradients [7-11] can be classified into two major categories, depending on the discretization scheme which is used to approximate γ field. In the first category, advantage is taken of the fact that the shape of the "micro" strain field is known a priori, so γ can be expressed as a nonlinear function of several parameters a_I. For example, on the basis of what is known in one-dimensional solutions, the shape of the interpolated "micro"

strain field in rate-dependent materials can be assumed of a cusped structure (fig.2a) given by

$$\gamma = \frac{a_2 + a_1}{1 + \left(\frac{\bar{x}_1}{a_3}\right)^2} - a_1 \tag{2.8}$$

while for rate-dependent materials it can be assumed to be a C^{-1} plateau with the higher value in the localized region (fig.2b)

$$\gamma = \begin{cases} a_2 & \text{for } |\bar{x}_1| > b/2 \\ a_1 & \text{for } |\bar{x}_1| < b/2 \end{cases} \tag{2.9}$$

Note that for rate-independent materials, the width of the shear band b (or the hinge line in plates) cannot be determined from the solution, since the governing equations change type from ellyptic to hyperbolic. This approach has been used in [7-9] in the context of rate-independent materials and will be referred here as an *embedded* formulation. The embedded technique is very useful if the structure of the strain field in the high gradients region is known a priori, since it allows a good estimate of the magnitudes of the strains to be obtained with a modest increase in the number of unknowns.

In many problems involving complex micro-mechanical behavior, the deformation field cannot estimated a priori, and therefore the shape of the embedded strain field cannot be preselected. In this case, the "micro" strain or displacement field must be considered completely unknown and an interpolant able to accurately resolve fields with high gradients must be used. In [11] spectral interpolants were used to approximate the "micro" strain field

$$\gamma = \Phi_I(\bar{x}_1)\, a_I \qquad I=1,2..NHS \tag{2.10}$$

where Φ_I are the higher order (spectral) strain interpolants and a_I are the spectral degrees of freedom; upper case subscripts I,J,K indicate spectral degrees of freedom and NHS is the total number of spectral degrees of freedom. This choice of approximation originates from the desire to combine the generality of the finite element method with regards to complex shapes and boundary conditions with the exponential convergence of the spectral method.

For certain class of materials, such as slightly rate-dependent materials, the use of piecewise constant strain interpolants might be of advantage, since fewer integration points in the direction normal to the band are needed. This possibility has been considered in [10]. For the purpose of classification we will refer to [10,11] as an *enriched* strain/displacement finite element technique.

In order to simultaneously treat both embedded and enriched formulations the higher order strain interpolants will be defined as follows

$$\Phi_I = \frac{\partial \gamma}{\partial a_I} \qquad (2.11)$$

Consequently the variation and the rate of γ are given by

$$\delta\gamma = \Phi_I \, \delta a_I \qquad (2.12a)$$

and

$$\dot{\gamma} = \Phi_I \, \dot{a}_I \qquad (2.12b)$$

With this definition of $\delta\gamma$ and $\dot{\gamma}$ the major difference between the enriched and embedded formulations is that embedded formulation Φ_I is a nonlinear function of unknown parameters a_I (or the a_I's are fixed a priori), while with enriched formulation Φ_I is constant with respect to time.

Once the structure of "micro" strain field has been defined (2.5-2.10), one can obtain the expression for $\mathbf{u}^H$ by integrating (2.5) in $\bar{x}_1$ direction and transforming the resulting expression from the $\bar{\mathbf{x}}$ into the $\mathbf{x}$ system. The resulting higher order displacement field is given by

$$u_i^H = m_i \int_{-h/2}^{\bar{x}_1} \gamma(s)\, ds \qquad (2.13)$$

Homogeneous boundary conditions on Γ_{H0} (2.3) are satisfied if γ is subjected to the following constraint equation

$$\int_{-h/2}^{h/2} \gamma(\bar{x}_1)\, d\bar{x}_1 = 0 \qquad (2.14)$$

which in the following will be regarded as a compatibility condition.

Note that the minimal continuity requirement for u_i^H is C^0, while γ or ε_{ij}^H need to be C^{-1}. All the techniques presented in [7-11] can be further classified according to the continuity of the "micro" strain field. Various possibilities are summarized in fig.3. For example the C^{-1} embedded formulation, which was used in [7-9], corresponds to a C^{-1} γ field with its shape preselected a priori; the C^{-1} enriched formulation with piecewise constant strain interpolants was used in [10], while a C^0 enriched formulation with spectral basis functions was employed in [11]; a C^0 embedded formulation is another promising possibility, since it allows us to obtain a good estimate of the high-gradient strain field by adding only a single unknown for the case of rate-independent solid or two unknowns if a rate-dependent material is considered.

3. Compatibility condition

As mentioned previously, a compatibility condition (2.14) is required to impose the homogeneous boundary conditions on Γ_{H0} (2.3). In this section we will review some of the techniques presented in [7-11] to satisfy this condition.

With the embedded formulation it is usually possible to express one of the unknowns in the "micro" strain field as a function of others, so that the compatibility condition (2.14) is automatically satisfied. For rate-independent materials[7-9] this yields

$$a_2 = a_1 \left(\frac{h}{b} - 1 \right) \tag{3.1}$$

A similar expression can be obtained from equation (2.8) combined with (2.14) for rate-dependent materials

$$a_2 = a_1 \left(\frac{h/2a_3}{\tan^{-1}(h/2a_3)} - 1 \right) \tag{3.2}$$

Several techniques can be used to impose compatibility with an enriched formulation, depending on how the γ field is discretized. If a spectral approximation of γ is used, compatibility can be satisfied by making each of the basis functions satisfy the compatibility condition

$$\int_{-h/2}^{h/2} \Phi_I(\bar{x}_1)\, d\bar{x}_1 = 0 \tag{3.3}$$

Equation (3.3) can be systematically satisfied by constructing spectral basis functions Φ_I as a sum of one of the polynomial series T_I (or Fourier) with some constant C_I

$$\Phi_I = T_I + C_I \tag{3.4}$$

where C_I is chosen so that equation (3.3) is satisfied

$$C_I = -\frac{1}{h} \int_{-h/2}^{h/2} T_I(\bar{x}_1)\, d\bar{x}_1 \tag{3.5}$$

For example, if T_I represents a Fourier series (only cosIθ terms are used because of symmetry) than

$$C_I = 0 \tag{3.6}$$

whereas if a Tchebisheff polynomial is used

$$C_I = \frac{0.5\,[(-1)^I + 1]}{I^2 - 1} \qquad I=2,3.... \text{ NHS} \quad \text{and} \quad C_1 = 0 \tag{3.7}$$

In case of piecewise constant approximations of γ [10], it is convenient to define an additional function $a(\bar{x}_1)$, which is related to $\gamma(\bar{x}_1)$ by

$$\gamma(\bar{x}_1) = \frac{\partial a}{\partial \bar{x}_1} \tag{3.8}$$

This can be discretized by

$$\gamma_I = \frac{a_{I+1} - a_I}{\Delta \bar{x}_{1I}} \tag{3.9}$$

Using the above definition of γ we can integrate equation (2.13) to obtain the following expression for the higher order displacement field

$$u_i^H = m_i\, a(\bar{x}_1) \tag{3.10}$$

and the homogeneous boundary conditions on Γ_{H0} are satisfied if

$$a_1 = a_{NHS} = 0 \tag{3.11}$$

We will refer to equations (3.1), (3.2), (3.3) and (3.10) as a discrete compatibility condition.

4. Variational framework

In order to obtain the discrete equilibrium equations, a weak (variational) form of the governing equations is used. In this section, it will be shown that under certain conditions both the displacement method, which was used in [11], and a Hu-Washizu variational principle employed in [7-10] lead to identical discrete equilibrium equations.

We first consider the principle of virtual work, which serves as a vehicle for the finite element displacement method. This principle states that for a body Ω enclosed by $\Gamma = \Gamma_u + \Gamma_t$ if

$$u \in \mathcal{U},\ \mathcal{U} = \left\{ u \,|\, u \in C^o,\ u = u^* \text{ on } \Gamma_u \right\} \tag{4.1a}$$

and

$$\delta u \in \mathcal{U}_o,\ \mathcal{U}_o = \left\{ \delta u \,|\, \delta u \in C^o,\ u = 0 \text{ on } \Gamma_u \right\} \tag{4.1b}$$

then if

$$\int_\Omega \delta u_{(i,j)}\, \sigma_{ij}(u)\, d\Omega = \int_{\Gamma_t} \delta u_i\, t_i\, d\Gamma + \int_\Omega \delta u_i\, b_i\, d\Omega \qquad \text{for } \forall\ \delta u \in \mathcal{U}_o \tag{4.2}$$

the equilibrium and the natural (traction) boundary conditions are satisfied.

Substituting interpolants (2.2), (2.5) and (2.6) into (4.2) we obtain

$$\delta d_A \int_\Omega B^0_{ijA} \sigma_{ij}\, d\Omega + \delta a_I \int_{\Omega_H} T_{ij}\, \Phi_I\, \sigma_{ij}\, d\Omega =$$

$$\delta d_A \left\{ \int_{\Gamma_t} N^0_{iA} t_i\, d\Gamma + \int_\Omega N^0_{iA} b_i\, d\Omega \right\} + \delta a_I \left\{ \int_{\Gamma_t} N^H_{iA}\, t_i\, d\Gamma + \int_{\Omega_H} N^H_{iA}\, b_i\, d\Omega \right\} \tag{4.3}$$

We consider the following conditions, regarding external forces applied on the domain where the high gradients are indicated by the solution

$$b_i = 0 \quad \text{on} \quad \Omega_H \tag{4.4a}$$

$$t_i = 0 \quad \text{on} \quad \Gamma_{H0} \tag{4.4b}$$

which exist in many practical applications, such as localization problems where the Ω_H is placed along the localization band or in fracture mechanics applications where Ω_H encompasses the crack tip.

Equating the coefficients of δd_A and δa_I , we find that the discrete equilibrium equations are as follows

$$f_A^{ext} = f_A^{int} \tag{4.5a}$$

$$q_I = 0 \tag{4.5b}$$

where

$$f_A^{int} = \int_\Omega B^0_{ijA}\sigma_{ij}\, d\,\Omega \tag{4.6a}$$

$$f_A^{ext} = \int_{\Gamma_t} N^0_{iA}t_i\, d\Gamma + \int_\Omega N^0_{iA}b_i\, d\Omega \tag{4.6b}$$

$$q_I = \int_{\Omega_H} N^H_{(iI,j)}\sigma_{ij}\, d\Omega = \int_{\Omega_H} \Phi_I\,(\bar{x}_1)\,\sigma_{mn}\, d\Omega \tag{4.6c}$$

and

$$\sigma_{mn} = \mathbf{m}\cdot\sigma\cdot\mathbf{n} \tag{4.7}$$

Equation (4.5a) is an equilibrium condition at the nodes. Equation (4.5b) pertains to equilibrium of the higher order fields, and is a weak statement of the fact that the component σ_{mn} of the stress field is constant, or in other words, its jump across the shear band is equal to zero (since Φ_I does not contain the constant field).

Next we consider a Hu-Washizu variational principle, which states that if

$$u \in \mathcal{U} \tag{4.8a}$$

$$\sigma \in \mathcal{S},\ \mathcal{S} = \left\{\sigma \,|\, \sigma \in C^{-1}\right\} \tag{4.8b}$$

$$\varepsilon \in \mathcal{E},\ \mathcal{E} = \left\{\varepsilon \,|\, \varepsilon \in C^{-1}\right\} \tag{4.8c}$$

and

$$\delta u \in \mathcal{U}_o,\ \delta\sigma \in \mathcal{S},\ \delta\varepsilon \in \mathcal{E} \tag{4.9}$$

then if

$$\int_{\Omega} \delta\varepsilon_{ij}\, \Sigma_{ij}(\boldsymbol{\varepsilon})\, d\Omega - \int_{\Gamma_t} \delta u_i\, t_i\, d\Gamma - \int_{\Omega} \delta u_i\, b_i\, d\Omega$$

$$+ \delta \int_{\Omega} \sigma_{ij}\, (\varepsilon_{ij} - u_{(i,j)})\, d\Omega = 0 \qquad \text{for } \forall\ \delta u \in \mathcal{U}_0,\ \delta\sigma \in \mathcal{S} \text{ and } \delta\varepsilon \in \mathcal{E} \tag{4.10}$$

the equilibrium, strain-displacement and constitutive equations together with natural boundary conditions are satisfied. Here Σ is the stress tensor obtained from the constitutive function, which may be both time and path dependent and is obtained from the interpolated strain field.

Substituting interpolants (2.2), (2.5) and (2.6) into (4.10) we obtain

$$\delta d_A \int_{\Omega} B^0_{ijA}\sigma_{ij}\, d\Omega + \delta a_I \int_{\Omega_H} \Phi_I\, \sigma_{mn}\, d\Omega -$$

$$\delta d_A \Big\{ \int_{\Gamma_t} N^0_{iA} t_i\, d\Gamma + \int_{\Omega} N^0_{iA} b_i\, d\Omega \Big\} + \delta \int_{\Omega} \sigma_{mn}\, \gamma\, d\Omega = 0 \tag{4.11}$$

Then if σ_{mn} is discretized by

$$\sigma_{mn} = S_{mnI}\, s_I \tag{4.12}$$

For certain low order elements, the stress interpolants can be taken as a combination of constants and selected linear terms. In that case, as shown in [7], the last term in (4.11) vanishes due to the compatibility condition (2.14), and the resulting discrete form is identical to the one obtained from the displacement method under the assumption (4.4).

Note that the embedded formulation presented in references [7-9] differs from the approach adopted here because of the following:

1. In [7-9] no variation on the parameters in the strain field has been taken. Instead an internal equilibrium condition has been stated using a collocation method

$$n_i\, \sigma_{ij} = \text{constant} \qquad \text{for } j = 1\text{-NSD} \tag{4.13}$$

Furthermore, traction continuity is satisfied for all components , so that the components of $\mathbf{m}$, which are initially determined from a local bifurcation analysis, can change as the problem evolves. With this approach the number of Gauss points must be greater or equal to the number of a_I. For example, in the case of an embedded formulation used in the context of rate-independent materials, a single unknown in the strain field (after compatibility has been imposed) can be determined using a one-point integration

scheme.This scheme combined with stabilization [15] is also useful to avoid locking due to incompressibility.

Alternatively , one could maintain traction continuity in all its components by using Galerkin weak form

$$\int_{\Omega_H} \Phi_I \, n_i \, \sigma_{ij} \, d\Omega = 0 \tag{4.14}$$

2. In order to satisfy the criteria that any rigid body motion causes zero strain, the γ field in references [7-9] was defined by

$$\gamma = \alpha \, T_{kl} \, u_{(k,l)} \tag{4.15}$$

It will be shown in Section 6 that with the unified enriched/embedded formulation proposed here, the rigid body criteria is satisfied automatically and no additional precautions need to be taken.

3. Due to the definition of γ field in (4.15) and due to the combined Galerkin/Collocation variational scheme which was used in [7-9], the resulting tangent stiffness obtained by the linearizing the discrete equilibrium equations (4.5) is not symmetric. This is in contrast to the formulation presented in this paper (see Section 6), where the tangent stiffness is symmetric if an associative flow rule is used.

5. Notes on convergence

The accuracy of the embedded method depends solely on our success in anticipating the deformation field. If the shape of the deformation/strain field is known a priori, or in singular problems, if the order of singularity is known, the use of the embedded formulation might be of advantage over the other techniques.

The accuracy of enriched formulation, on the other hand, depends on the type of interpolants. In this section we will examine the rate of convergence of several spectral approximations. The accuracy of the spectral techniques is evaluated on the basis of comparison to a second order accurate finite-difference method, which was used in [10] to discretize the γ field.

We consider a $\gamma(\bar{x}_1)$ field where neither $\gamma(\bar{x}_1)$ nor its derivatives vanish at the boundaries of the spectral domain. For simplicity we assume $\gamma(\bar{x}_1)$ is a symmetric function,

so that only half of the spectral domain is considered. We first examine a Fourier expansion, where

$$T_I = \cos K\theta \tag{5.1a}$$

and

$$C_I = 0 \tag{5.1b}$$

So the γ field is interpolated by

$$\gamma(\theta) = \sum_{K=1}^{NHS} a_K \cos K\theta \tag{5.2}$$

where the $\bar{x}_1 \in [0;b/2]$ is mapped into $\theta \in [0;\pi]$. Note that all the terms in the cosine expansion, except for the one with K=0, which is a constant, satisfy the discrete compatibility condition (3.3). This term is omitted from (5.2) since the constant field can be obtained from the finite element approximation.

The spectral coefficients a_K can be found by Fast Fourier Transform (FFT):

$$a_K = \frac{2}{\pi} \int_0^{\pi} \gamma(\theta) \cos K\theta \, d\theta \qquad K = 1,2..NHS \tag{5.3}$$

This choice of basis functions does not provide an exponential convergence as shown in [16]. In order to illustrate this we integrate equation (5.3) by parts twice:

$$a_K = \frac{1}{K^2} \frac{2}{\pi} \gamma_{,\theta} \cos K\theta \Big|_0^{\pi} - \frac{1}{K^2} \frac{2}{\pi} \int_0^{\pi} \gamma_{,\theta\theta} \cos K\theta \, d\theta \qquad K = 1.2..NHS \tag{5.4}$$

The boundary term in (5.4) does not vanish because the spatial derivative of $\gamma(\theta)$ is not zero at the boundaries, as mentioned before. Hence, for all problems where the derivatives of the strain field do not vanish at the boundaries, the Fourier approximation, similar to the finite-difference method, provides only second order accuracy .

Next we will consider an interpolation of $\gamma(\bar{x}_1)$ by Tchebysheff polynomials

$$T_I(\xi) = \cos I\theta = \cos I(\cos^{-1}\xi) \tag{5.5}$$

which are modified by a constant C_I given in (3.7), so that γ field is approximated by

$$\gamma(\xi) = \sum_{I=1}^{NHS} a_I \, (\cos I\theta + C_I) \tag{5.6a}$$

where

$$\xi = \cos\theta \tag{5.6b}$$

and the discrete compatibility condition (3.3) is satisfied. Here the $\bar{x}_1 \in [0;b/2]$ is mapped into $\xi \in [-1;1]$. Following [11], we will refer to the expansion (5.6) as a Modified Tchebisheff Polynomial (MTP).

As in the Fourier series, the coefficients a_I can be obtained by an FFT by premultiplying (5.6a) by $\cos K\theta$ and then integrating

$$a_K = \frac{2}{\pi} \int_0^{\pi} \gamma(\xi) \cos K\theta \, d\theta = \int_{-1}^{1} \frac{\gamma(\xi) \, T_K(\xi)}{\sqrt{1-\xi^2}} \, d\xi \qquad K = 1,2..NHS \tag{5.7}$$

Note that C_I does not appear in equation (5.7) because of its orthogonality to $\cos K\theta$.

In order to estimate the accuracy of the MTP series we follow the same procedure as for the Fourier series. Integration by parts of (5.7) with respect to θ leads to

$$a_K = \frac{1}{K} \, \frac{2}{\pi} \int_0^{\pi} \bar{\gamma}(\xi) \sin K\theta \, d\theta \qquad K = 1,2..NHS \tag{5.8a}$$

where

$$\bar{\gamma}(\xi) = \gamma_{,\xi}(\xi) \sin\theta \tag{5.8b}$$

Additional integration by parts of eq. (5.8a) yields

$$a_K = -\frac{1}{K^2} \, \frac{2}{\pi} \, \bar{\gamma}(\xi) \cos K\theta \Big|_0^{\pi} - \frac{1}{K^2} \, \frac{2}{\pi} \int_0^{\pi} \bar{\gamma}(\xi) \sin\theta \cos K\theta \, d\theta \tag{5.9}$$

In contrast to the Fourier expansion, the boundary term in equation (5.9) vanishes because $\bar{\gamma}(-1) = \bar{\gamma}(1) = 0$. We can continue to integrate equation (5.9) by parts as long as the derivatives of γ are differentiable. Since the accuracy of the method is dominated by the

value of the first omitted coefficient , the accuracy of this method is of order r, where r is the number of non-zero spatial derivatives of $\gamma(\xi)$.

Although the spectral approximation with a MTP series provides a significantly better accuracy than the piecewise constant strain interpolation [10], the computational efficiency of the spectral approximation is not always superior. As mentioned previously, for certain classes of materials, such as a slightly rate-dependent solid with a rate sensitivity constant less than 0.01, the use of piecewise constant strain interpolants might be of advantage, since it involves fewer integration points. Furthermore, if singularities or discontinuous derivatives exist in the solution, say on Γ_{SING}, the exponential convergence of spectral techniques is maintained only if $\Gamma_{SING} \in \Gamma_{H0}$. In some cases, this restriction on the choice of Γ_{H0} may contradict our basic desire to have simply shaped spectral domains so that the spectral approximation could be effectively applied in multidimensions. On the other hand, if $\Gamma_{SING} \notin \Gamma_{H0}$, only second order accuracy is obtained, and therefore a simpler piecewise constant strain enrichment might be considered.

The MTP as well as the standard Tchebysheff Polynomial provide high resolution at the boundaries of the spectral domain. This is a very desirable property in the case of high gradients near the boundaries. However, if the γ field over half of the spectral domain is characterized by the high gradients only at one end, it is desirable to map ξ onto $\xi' \in [-1,1]$ so that higher resolution will be obtained only at one end. This technique with spectral interpolants is equivalent to grid refinement in a finite-difference method. Following [11], we use the mapping suggested in [17]

$$\xi = \frac{4}{\pi} \tan^{-1} \left\{ R \tan\left[\frac{\pi}{4} (\xi'-1) \right] \right\} + 1 = Q(\xi') \tag{5.10}$$

where R is a free parameter. In most practical applications, R ranges from one to ten, depending on the intensity of the strain gradients. Note that for R=1, $\xi = \xi'$. The mapping given in (5.10) can be seen as an additional modification of the spectral basis functions. For example, the fourth MTP term $\Phi_4(\xi') = \Phi_4\{Q^{-1}(\xi)\}$ is plotted in fig.4 as a function of ξ for various values of R. It can be seen that as R increases the function has higher resolution at the end where the high gradients exist.

Note that if (5.10) is employed, the expression for C_I is a function of R and equation (3.7) is no longer valid for $R \neq 1$. In order to evaluate C_I for the case of $R \neq 1$ (3.5) must be integrated numerically with respect to ξ'.

6. Incremental discrete equations for large deformation problems

The equilibrium equations (4.5) together with the specific definition of a constitutive model comprise a system of nonlinear equations for the discrete variables d_A and a_I. For rate-independent materials this system of equations has been solved by Newton method in [7-9], whereas for rate-dependent materials a forward Euler integration scheme has been employed in [10,11], where the precise definition of a tangent stiffness matrix is of great importance, since no equilibrium iterations are made.

The objective of this section is to derive the expression of the tangent stiffness matrix for the unified enriched/embedded formulation in context of general nonlinear analysis. Linearization of the nonlinear equations is achieved by taking the material time derivative of (4.5), which in case of rate-independent materials plays the role of a load parameter.

We start with the linearization of f_A^{int} given in (4.5a):

$$\frac{d}{dt} f_A^{int} = \frac{d}{dt} \int_{\Omega_0} \frac{\partial N_{iA}^o}{\partial X_k} \frac{\partial X_k}{\partial x_j} \sigma_{ij} J \, d\Omega$$

$$= \int_{\Omega_0} \frac{\partial N_{iA}^o}{\partial X_k} (\dot{F}_{kl}^{-1} \sigma_{il} J + F_{kj}^{-1} \dot{\sigma}_{ij} J + F_{kj}^{-1} \sigma_{ij} \dot{J}) \, d\Omega \qquad (6.1)$$

where x and X are the spatial and material coordinates, Ω_0 and Ω designate initial and current domains, σ is a Cauchy stress, F is a deformation tensor and J is a Jacobian transformation between the initial and the current domains.

A similar expression is obtained from the linearization of q_I in (4.5b), except that N_{iA}^o is replaced by N_{iA}^H.

To proceed, the constitutive equation is introduced, where the material objectivity is maintained by using Truesdell rate of Cauchy stress (although other objective rates could be used)

$$\dot{\sigma}_{ij} = \dot{\sigma}_{ij}^T - v_{l,l} \sigma_{ij} + v_{i,l} \sigma_{lj} + v_{j,l} \sigma_{il} \qquad (6.2)$$

A constitutive equation which simultaneously treats rate-dependent and rate-independent materials is employed

$$\dot{\sigma}_{ij}^T = C_{ijkl} v_{kl} - \alpha S_{ij} \qquad (6.3)$$

where for a rate-independent plastic solid, C_{ijkt} represents a tangent modulus and $\alpha = 0$, while for viscoplastic solids C_{ijkt} is a tensor of elastic moduli, $\alpha = 1$ and S_{ij} embodies the nonlinear response and is given by

$$S_{ij} = C_{ijkl}\,(\dot{\varepsilon}^{eff}\, R_{kl}) \tag{6.4}$$

where R_{kl} is a direction of the plastic flow, which is normal to the yield surface for an associated flow rule, and $\dot{\varepsilon}^{eff}$ is an effective viscoplastic strain given by a power law

$$\dot{\varepsilon}^{eff} = a\left(\frac{\sigma^{eff}}{g(\varepsilon^{eff})}\right)^{1/m} \tag{6.5}$$

where σ^{eff} is an effective stress, $g(\varepsilon^{eff})$ a uniaxial stress-viscoplastic strain function, m a rate sensitivity exponent and a is a material constant.

Combining eqs. (6.1-6.3) together with the following kinematical relations

$$\dot{J} = J\, v_{l,l} \tag{6.6}$$

$$\dot{F}^{-1}_{kl} = -\, F^{-1}_{kj}\, v_{j,l} \tag{6.7}$$

we obtain the linearized equilibrium equations

$$\begin{pmatrix} K_{AB} & K_{AJ} \\ K_{IB} & K_{IJ} \end{pmatrix} \begin{pmatrix} \dot{d}_B \\ \dot{a}_J \end{pmatrix} = \begin{pmatrix} \dot{f}^{ext}_A + \alpha V_A \\ \alpha V_I \end{pmatrix} \tag{6.8}$$

where

$$K_{AB} = \int_{\Omega} B^0_{ijA}\, D_{ijkl}\, B^0_{klB}\, d\,\Omega \tag{6.9}$$

$$K_{AJ} = \int_{\Omega} B^0_{ijA}\, D_{ijkl}\, T_{kl}\, \Phi_J\, d\Omega \tag{6.10}$$

$$K_{IJ} = \int_{\Omega} T_{ij}\, D_{ijkl}\, T_{kl}\, \Phi_I\, \Phi_J\, d\Omega \tag{6.11}$$

$$V_A = \int_{\Omega} B^0_{ijA}\, S_{ij}\, d\,\Omega \tag{6.12}$$

$$V_I = \int_{\Omega} T_{ij}\, S_{ij}\, \Phi_I\, d\Omega \tag{6.13}$$

$$D_{ijkl} = C_{ijkl} + \delta_{ik}\, \sigma_{jl} \tag{6.14}$$

and B^0_{ijA} has been redefined by

$$B^0_{ijA} = \frac{\partial N^o_{iA}}{\partial x_j} \tag{6.15}$$

Remarks:

1. In equations (6.10), (6.11) and (6.13) T_{ij} is a function of the current values of **m** and **n**. In deriving the linearized equilibrium equations (6.8-6.13), **m** has been treated as a constant. If **m** is a variable, equation (4.14) needs to be linearized instead of (4.5b) and the resulting equation will be nonsymmetric regardless the constitutive equation.

The current value of the normal to the band can be found as in [10] by

$$\mathbf{n}^{k+1} = \mathbf{n}^k + A_{ijl}\, v^{k+1/2}_{j,l}\, \Delta t \tag{6.16}$$

where

$$A_{ikl} = -\frac{n^*_p\, F^{*-1}_{pk}\, \delta_{il}}{(n^*_p\, F^{*-1}_{pk}\, n^*_s\, F^{*-1}_{sk})^{1/2}} + \frac{n^*_p\, F^{*-1}_{pi}\, n^*_s\, F^{*-1}_{sk}\, n^*_t\, F^{*-1}_{tl}}{(n^*_p\, F^{*-1}_{pk}\, n^*_s\, F^{*-1}_{sk})^{3/2}} \tag{6.17}$$

The (*) symbol indicates the configuration where the localization initially occurred, which serves as the reference frame. Details of the linearization procedure for the normal to the surface were presented in [10].

2. For rate-independent materials $\dot{a}_J$ is given by

$$\dot{a}_J = -\,K^{-1}_{JI}\,(K_{IB}\,\dot{d}_B) \tag{6.18}$$

and since $K_{IB}\,\dot{d}_B = 0$ because of $B^0_{ijB}\,\dot{d}_B = 0$ in rigid body motion, the criteria of zero strains in rigid body motion is satisfied by the unified enriched/embedded formulation.

7. Numerical examples

We consider a viscoplastic solid with an isotropic hardening law and von Mises yield criterion . A trilinear stress-viscoplastic strain law with softening in the last segment is used for $g(\varepsilon^{eff})$ where $g(\varepsilon_1^{eff} = 0.006) = 300.$, $g(\varepsilon_2^{eff} = 0.008) = 320.$, $g(\varepsilon_3^{eff} = 0.60) = 20$; see eq. (6.5). Other material parameters are: Young's modulus E = 20000., Poisson's ratio $\nu = 0.30$, m=0.05.

We consider a rectangular bar of length L=18 and width W=10 in a tension test. Because of symmetry, only one quarter of the bar is modelled and the spectral patch is placed over bifurcated elements as shown in fig.5a. An inhomogenity is introduced by a triangular imperfection in the thickness of the material which is of width W/80 and results in a maximum thickness change of 3%. This imperfection is rather large but it was difficult to trigger the bifurcation with a smaller imperfection of this width. We study the behavior of the proposed method as the number of spectral degrees of freedom is increased.

A typical deformed mesh is given in fig.5b. The force-displacement curve is given in fig.6 for various orders of spectral interpolations, designated by NSPEC. The evolution of peak effective plastic strain is presented in fig.7 and the typical effective plastic strain profiles are given in fig.9. A,B and C are points on the load-deflection curve which represent specific states in the specimen's evolution. As the maximum load is reached (point A in figures 6,7) the maximum effective plastic strain is only 5% of its final value (at point C). The overall response and the strain history for the various spectral approximations are almost identical until the high gradients in the strain field start to appear (point B, figs.6,7). This point indicates the onset of localization, which is characterized by the steep drop in the load-displacement and effective plastic strain-displacement curves. The peak strain at point B is still only 20% of its value at point C. From B to C, which is the interval of strain localization, the displacements increase only by 0.1% , while the maximum strains in the shear band are multiplied by more than five. Note that as the number of spectral interpolants is decreased, i.e. as resolution is decreased, the onset of localization is delayed and the peak strains decrease.

A profile of the effective plastic strain across the width of the spectral overlay (this direction is normal to the direction of the shear band) is shown in fig.8. It can be seen that the strain is extremely localized ; the width of the shear band is on the order of 50 microns.

which corresponds to widths reported in experiments. The width of the shear band depends on the imperfection, which in this case was 1000 microns wide.

Finally, a comparison between the spectral overlay method [11] and the piecewise constant enriched strain element [10] is presented. The same material model as in [10] is utilized with $g(\varepsilon_1^{eff} = 0.015) = 300.$, $g(\varepsilon_2^{eff} = 0.020) = 320.$, $g(\varepsilon_3^{eff} = 1.50) = 20.$ and other material constants and imperfection size as before. In fig.9 we compare force-displacement curves and effective plastic strain profiles obtained using spectral overlay technique with 15 spectral degrees of freedom and the enriched element with 12 piecewise constant interpolants in each element. As can be seen from fig.9 the results agree quite well, but in this case the enriched element technique involved far more unknowns. With the spectral overlay method only 15 additional unknowns were used, while with the enriched strain element the number of additional unknowns was 12xNENR, where NENR is the number of enriched elements in the mesh. On the other hand with the enriched element technique the additional degrees of freedom are eliminated(condensed) on the element level, and also fewer integration points are needed in the $\bar{x}_1$-direction.

8. Conclusions

A unified approach in treating finite element solutions with high gradients has been presented in context of large deformation problems. The embedded element technique developed in [7-9], has been modified, so that the enriched and embedded formulations can be treated simultaneously. The unified approach allows both formulations to be combined in the same problem. This merging can be performed either on separate domains, or on the same domain in different directions. For example, in fracture mechanics applications, the singularity near the crack tip can be effectively treated by an embedded formulation in the radial direction, especially if the order of singularity is known, while in the circumferential direction the enriched formulation can be employed. Furthermore, if in some other portions of the structure of high gradients are indicated by the solution due to localization or stress/strain concentration, an additional enriched or embedded approximation can be superimposed. These topics combined with adaptive methods are the subject of our further research.

Acknowledgement

The support of the Air Force Office of Scientific Research under Award F49620-87-C-0030 to Northwestern University is gratefully acknowledged.

References

1. G.F. Carey. A mesh-refinement scheme for finite element computations. *Comp. Meth. Appl. Mech. Engng.* 7, 93-105 (1976).
2. I. Babuska. The self-adaptive approach in the finite element method. *The Mathematics of Finite Elements and Applications. MAFELAP 1975*, Academic press, New-York (1976).
3. C.A. Fellipa. Optimization of finite elements grids by direct energy search. *Appl. Mat., Modelling 1*, 93-96 (1976).
4. M.R. Albert and K.O. Neill . The use of transfinite mappings with finite elements on a moving mesh for two-dimensional phase change. In I.Babuska , J.Chandra and J.E.Fhaherty (eds.) *Adaptive Computational Methods for Partial Differential Equations*. SIAM (1983)
5. M.P. Rossow and I.R. Katz. Hierarchichal finite elements and precomputed arrays. *Int. J. Num. Meth. Engng.* 12(6), 977-1000(1978).
6. A.T. Patera. Spectral element method for fluid dynamics: Laminar flow in a channel expansion. *Journal of Computational Physics* 54, 468-488 (1984).
7. T.Belytschko, J.Fish, B.E.Engelmann. A finite element with embedded localization zones. *Comp. Meth. Appl. Mech. Engng.* 70, 59-89 (1988).
8. J.Fish and T.Belytschko. Elements with embedded localization zones for large deformation problems. *Computers and Structures*. 30, No 1/2, 247-256 (1988).
9. T.Belytschko and J.Fish. Embedded hinge lines for plate elements. Comp. Meth. Appl. Mech. Engng. In press.
10. J.Fish and T.Belytschko. A finite element with a unidirectionally enriched strain field for localization analysis. To appear in *Comp. Meth. Appl. Mech. Engng.* (1989).
11. T.Belytschko, J.Fish and A.Bayliss. Spectral overlay on finite elements for problems with high gradients. To appear in *Comp. Meth. Appl. Mech. Engng.* (1989)
12. J. Hadamard. *Leçons sur la propagation des ondes et les equations de l'hydrodynamique*. Librairie Scientifique A Hermann, Paris (1903).
13. R.Hill. Acceleration waves in solids. *Journal of the Mechanics and Physics of Solids*, 10, 1-16 (1962).

14. M. Ortiz, Y. Leroy and A. Needleman. A finite element method for localization failure analysis. *Comp. Meth. Appl. Mech. Engng.* 61, 189-214 (1988).
15. D.P. Flanagan and T.Belytschko. A uniform strain hexahedron and quadrilateral with orthogonal hourglass control. *Int. J. Num. Meth. Engng.* 17, 679-706 (1981).
16. D. Gottlieb and S.A. Orszag. Numerical analysis of spectral methods: theory and applications. *NSF-CBMS Monograph , 26, Soc. Indus. Appl. Math., Philadelphia (1979)* .
17. A. Bayliss, D.Gottlieb, B.J. Matkowsky and M.Minkoff. An adaptive Pseudo-Spectral method for reaction diffusion problems. *Journal of Computational Physics* 81, 421-443 (1989).

List of figure captions

fig.1 Superposition of the finite element and the spectral approximations.

fig.2 Interpolated strain profiles across the shear band in rate-dependent and rate-independent materials

fig.3 Classification of various finite element techniques for localization analysis

fig.4 The forth MTP term $\Phi_4(\xi')$ vs. ξ for various values of R.

fig.5 Undeformed and deformed meshes for rectangular bar in tension.

fig.6 Force vs. displacement for various number of spectral unknowns.

fig.7 Effective plastic strain vs. displacement for various number of spectral unknowns.

fig.8 Strain profiles along normal to the shear band , $\bar{x}_1$, for various values of stretch of bar d_{BC}.

fig.9 Force vs. displacement (a) and effective plastic strain vs. displacement (b) for the Spectral Overlay[11] and Enriched element [10] methods.

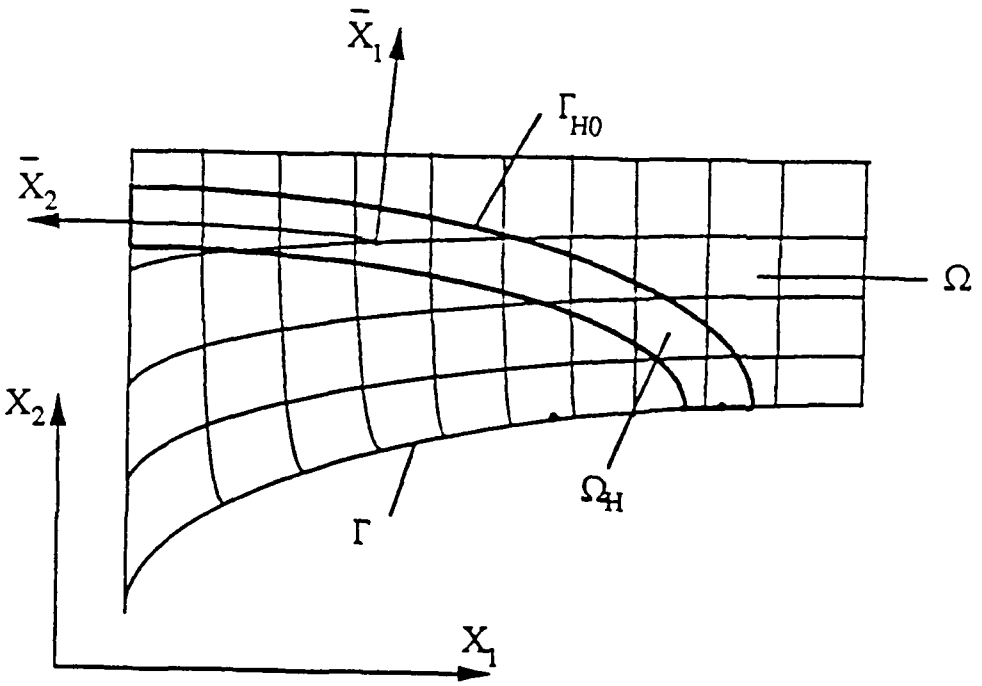
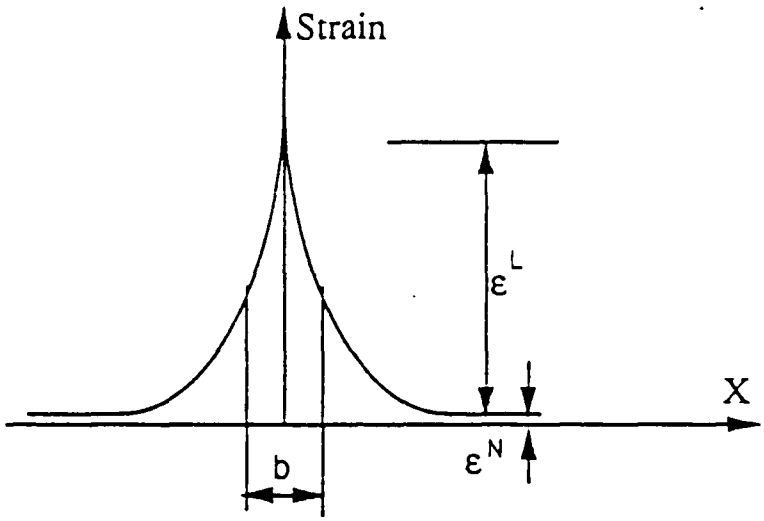
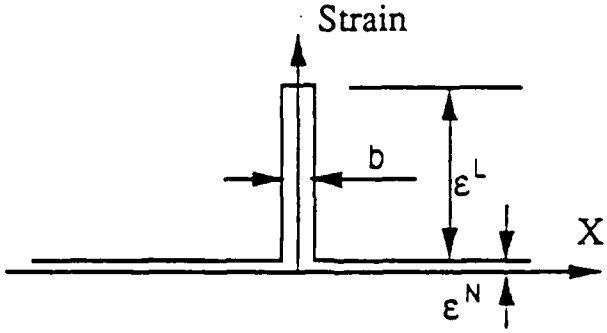


Figure 1



(a) Rate-dependent material



(b) Rate-independent material

Figure 2

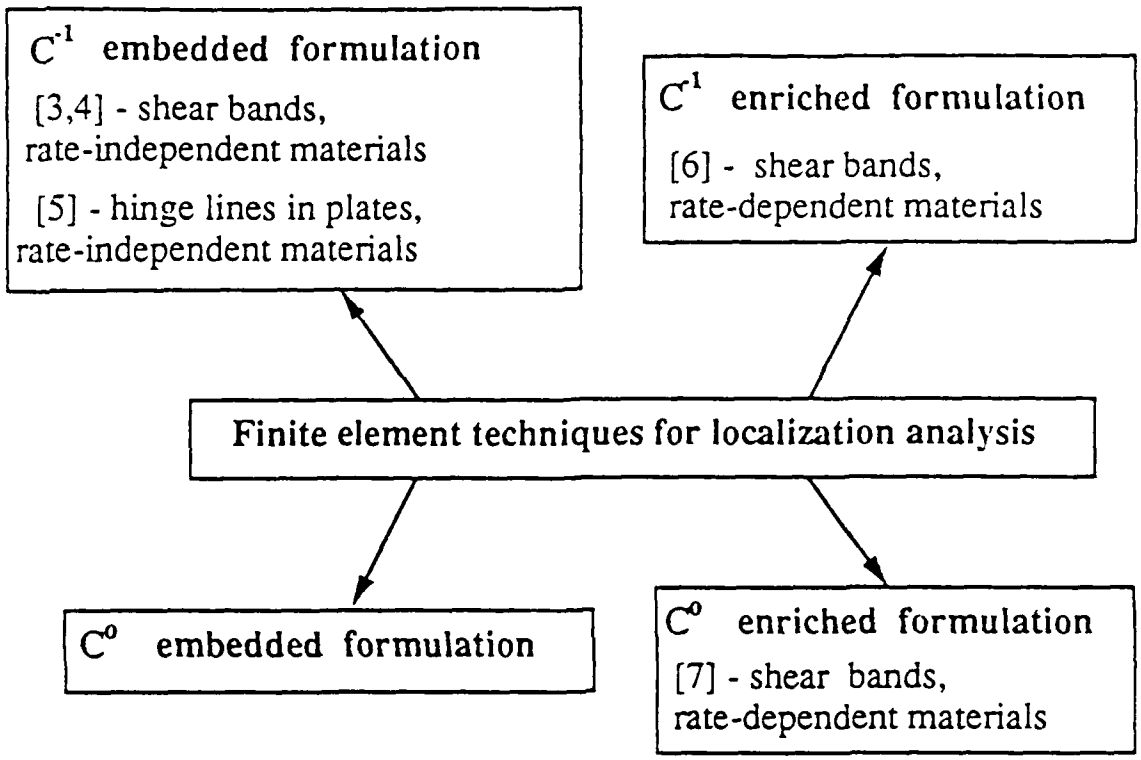

C^{-1} embedded formulation
[3,4] - shear bands, rate-independent materials
[5] - hinge lines in plates, rate-independent materials
C^{-1} enriched formulation
[6] - shear bands, rate-dependent materials
Finite element techniques for localization analysis
C^{0} embedded formulation
C^{0} enriched formulation
[7] - shear bands, rate-dependent materials

Figure 3

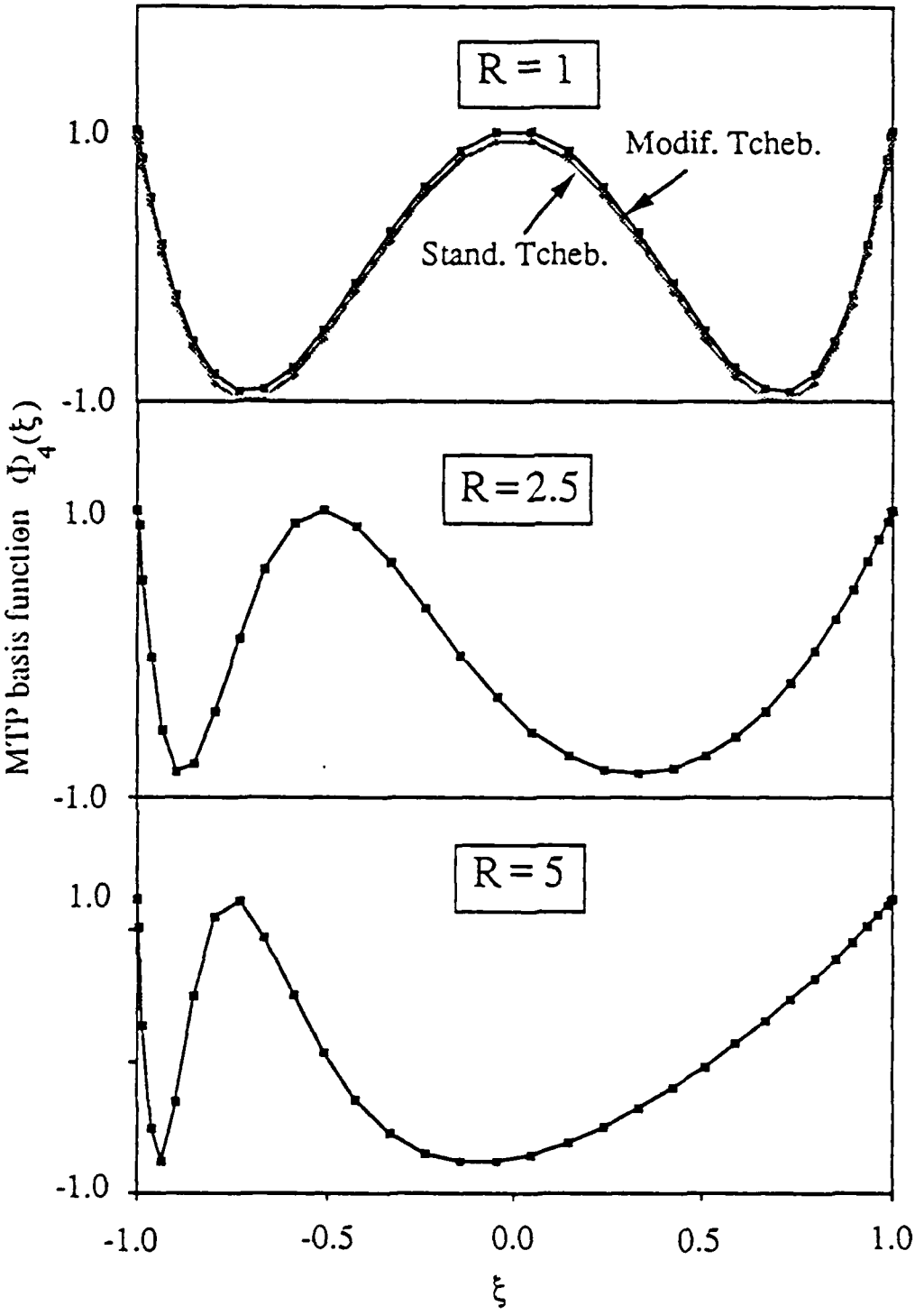

R = 1
Modif. Tcheb.
Stand. Tcheb.
R=2.5
R=5
1.0
-1.0
MTP basis function $\Phi_4(\xi)$
-1.0
-0.5
0.0
0.5
1.0
ξ

Figure 4

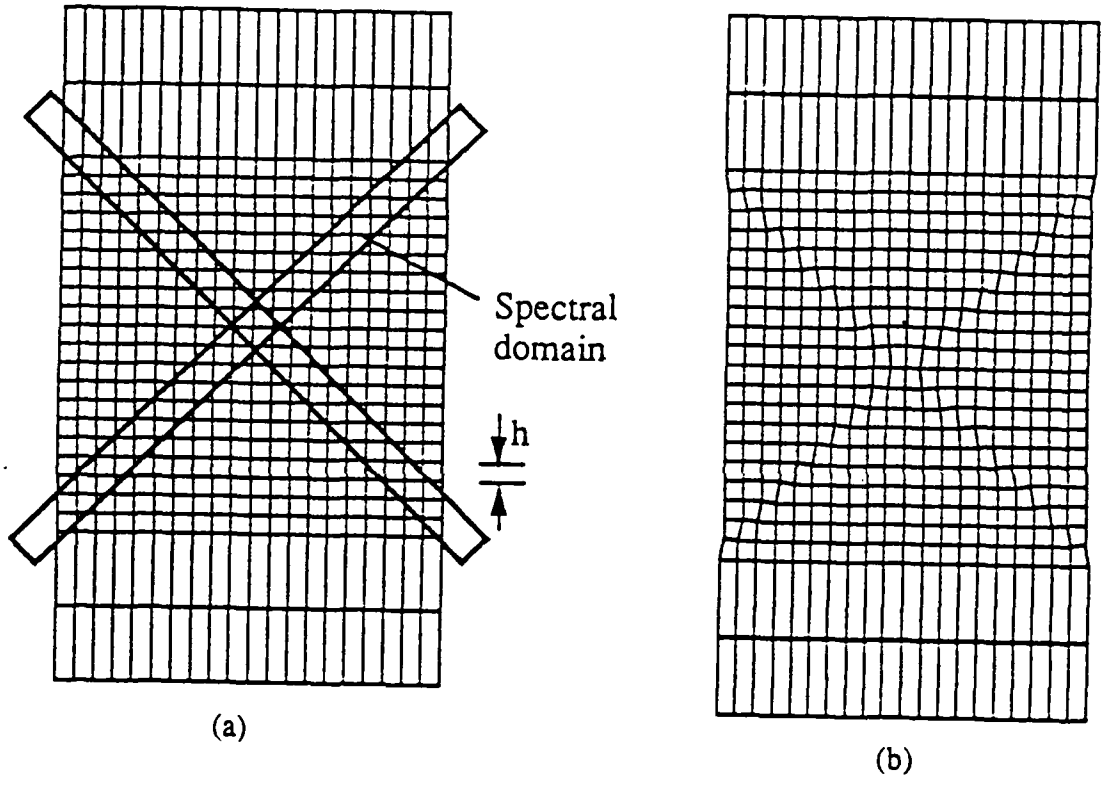


Figure 5

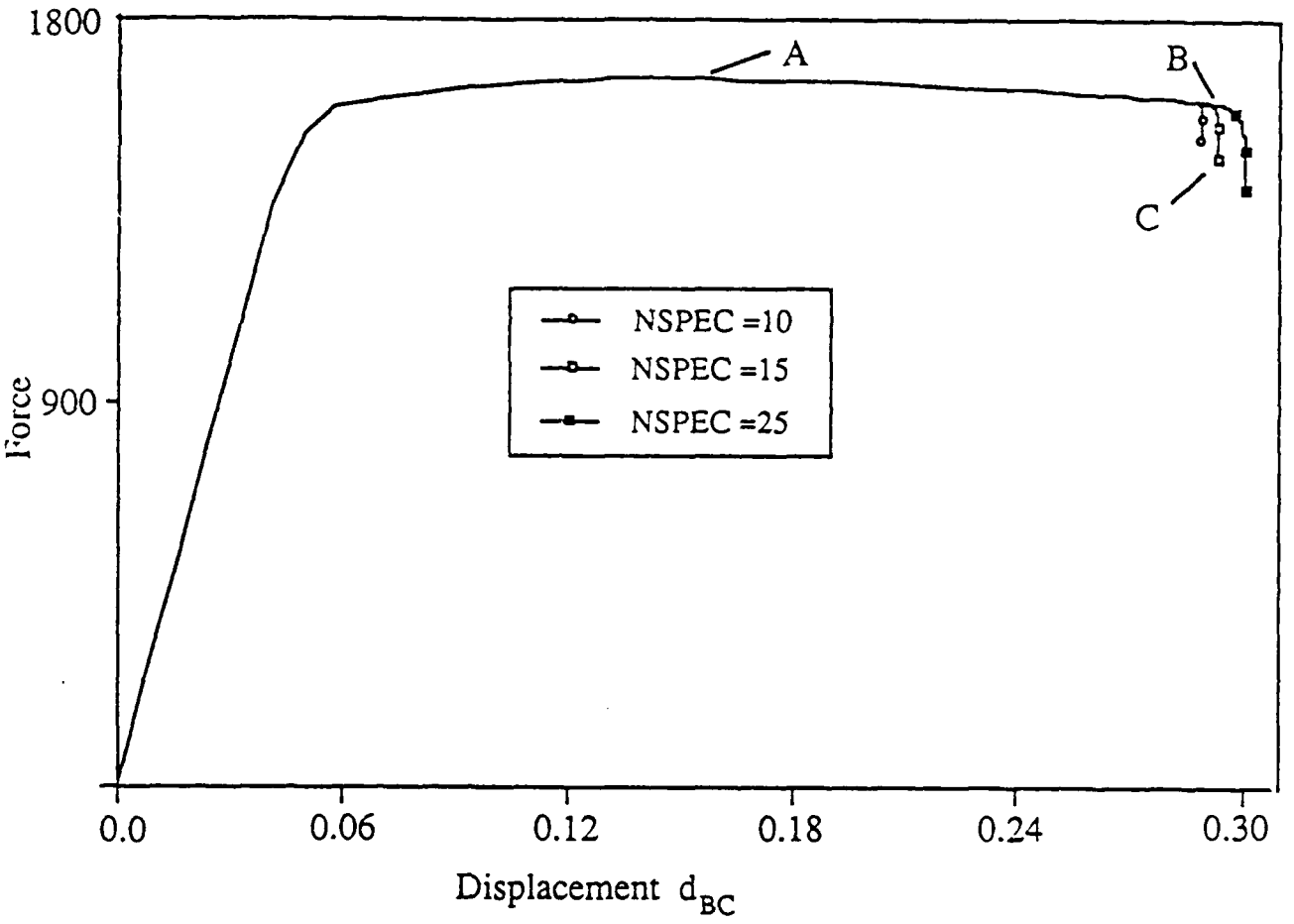

1800
900
Force
A
B
C
NSPEC =10
NSPEC =15
NSPEC =25
0.0
0.06
0.12
0.18
0.24
0.30
Displacement d_{BC}

Figure 6

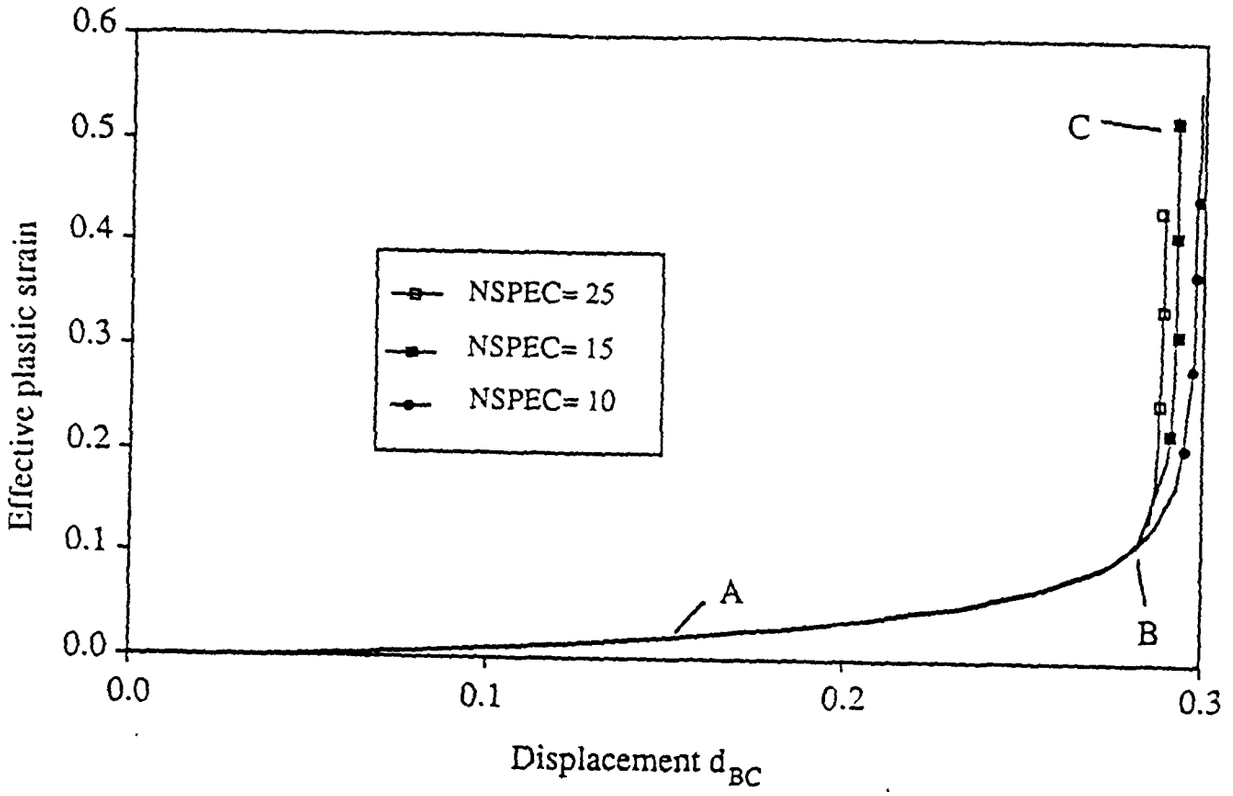

0.6
0.5
0.4
0.3
0.2
0.1
0.0
Effective plastic strain
NSPEC= 25
NSPEC= 15
NSPEC= 10
A
B
C
0.0
0.1
0.2
0.3
Displacement d_{BC}

Figure 7

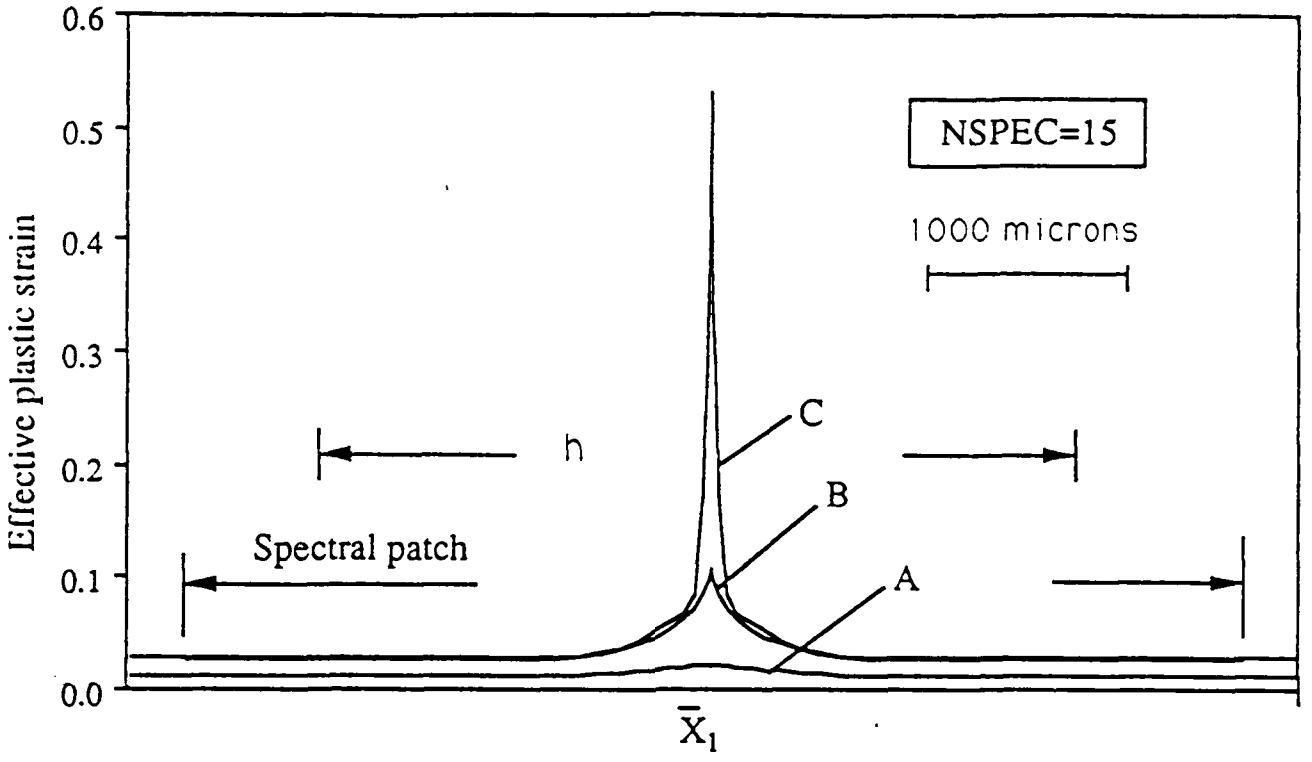
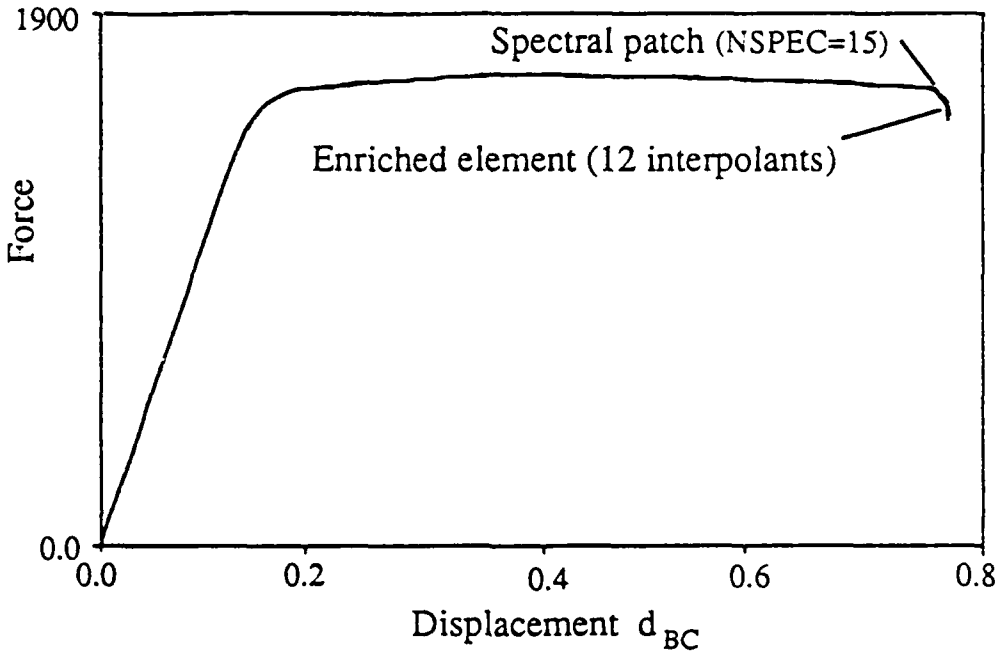

0.6
0.5
0.4
0.3
0.2
0.1
0.0
Effective plastic strain
NSPEC=15
1000 microns
h
C
B
A
Spectral patch
$\bar{X}_1$

Figure 8


1900
0.0
Force
Spectral patch (NSPEC=15)
Enriched element (12 interpolants)
0.0
0.2
0.4
0.6
0.8
Displacement d BC

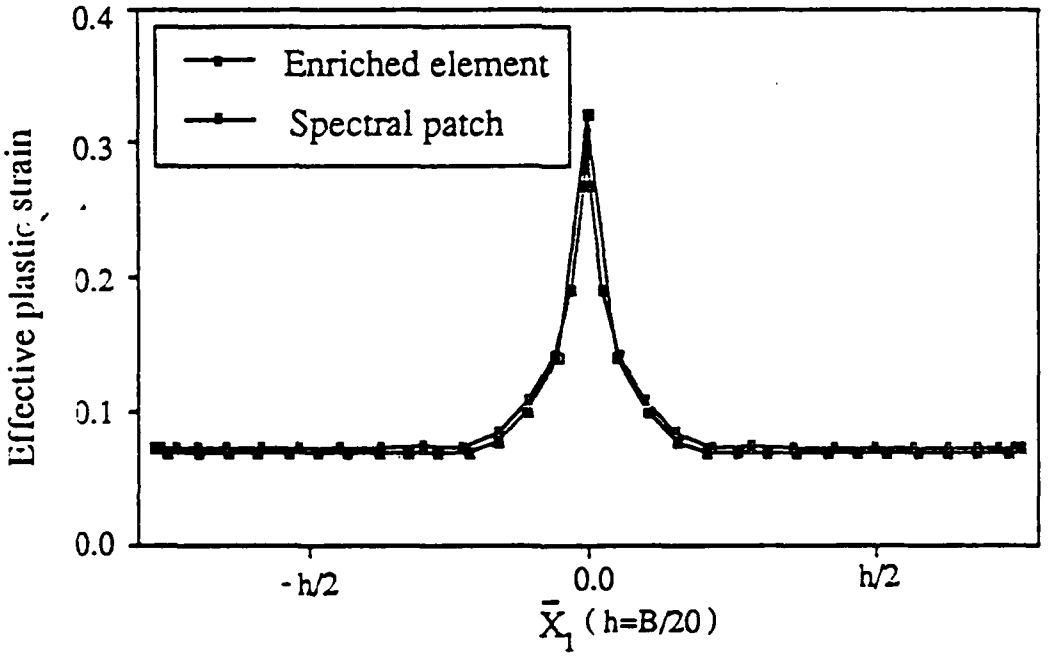

0.4
0.3
0.2
0.1
0.0
Effective plastic strain
Enriched element
Spectral patch
-h/2
0.0
h/2
X̄1 (h=B/20)

Figure 9