Concrete fracture models: testing and practice

Zdeněk P. Bažant *

Departments of Civil Engineering and Materials Science, The Technological Institute, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208-3109, USA

Abstract

The existing fracture models for concrete and the testing methods for fracture energy and other fracture characteristics are reviewed and some new results on the relationship between fracture testing and size effect are presented. The limitations of the cohesive crack model are discussed. The discrepancy between the fracture energy values measured by Hillerborg’s work-of-fracture method and the size effect method is explained and mathematically described by the recently proposed broad-range size effect law. The implications of the recently identified large statistical scatter of the fracture energy values measured by the work of fracture, compared to those obtained by the size effect method or Jenq–Shah method, are discussed. Merits of various testing methods are analyzed. A testing procedure in which the maximum loads of notched beams of only two different sizes in the ratio 2:1 and two different notches depths are tested is proposed and a least-square procedure for calculating the fracture parameters is given. A simplified testing procedure with an empirical coefficient, in which only the maximum loads of identical notched and unnotched beams of one size are tested, is also proposed as an alternative. To improve the size effect description for small sizes, the small-size asymptotics of the cohesive crack model is determined and a formula matching this asymptotics, as well as the large-size linear elastic fracture mechanics asymptotics, is presented. Finally, various arguments for introducing fracture mechanics into concrete design practice are reviewed and put into the perspective of safety factors. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Fracture; Concrete; Rock; Size effect; Scaling; Quasibrittle materials; Testing methods; Design practice; Fracture characteristics; Fracture process zone

1. Early history, objectives and scope

Initiated in 1961 by Kaplan [66], the study of fracture mechanics has progressed by the turn of the century quite far. Kesler et al. [67] showed that the classical linear elastic fracture mechanics (LEFM) of sharp cracks was inadequate for normal concrete structures. This conclusion was supported in 1972 by the results of Walsh [104,105], who tested geometrically similar notched beams of different sizes and plotted the results in a double logarithmic diagram of nominal strength versus size. Without attempting a mathematical description, he made the point that this diagram deviates from a straight line of slope −1/2 predicted by LEFM.
A major advance in concrete fracture was made in 1976 by Hillerborg et al. [59]. Inspired by the softening and plastic models of fracture process zone (FPZ) initiated in the works of Barenblatt [2,3] and Dugdale [52] and developed earlier for materials other than concrete by Rice [89], Smith [96], Knauss [69,70], Wnuk [109], Palmer and Rice [79], and Kfouri and Rice [68]. Hillerborg et al. [59–62] improved and adapted to concrete the cohesive crack model. Their finite element analysis showed that the cohesive crack model (also called the fictitious crack model) predicts, for the flexural failure of unnotched plain concrete beams, a deterministic size effect, different from the Weibull statistical size effect. This conclusion was strengthened and the model further refined by Petersson [80].

An analytical study of the size effect due to localization of distributed cracking was begun in 1976 [6]. Later, a simple formula for the size effect, which describes the size effect for quasibrittle failures preceded by large stable crack growth and allows determination of material fracture parameters from maximum load tests, was derived [8,9] and the crack band model [7,8], providing an almost equivalent alternative to the cohesive crack model, was developed. This model was shown to be in good agreement with the basic fracture data and size effect data [37], and has been found convenient for programming. It is nowadays the main concrete fracture model used in industry and commercial codes; e.g., code DIANA [93], SBETA [49,50] and ATENA.

Beginning with the mid 1980s, the stream of fracture and size effect studies swelled to a torrent. Many researchers made significant contributions; to name but a few: Petersson [80], Hillerborg [61,62], Elices et al. [53,54], Planas et al. [81–87], Guinea et al. [55–58], Carpinteri [43], Mihashi et al. [74], and Hu and Wittmann [63,64]. The size effect has become a major theme at conferences on concrete fracture [1,75,18,100,108].

An intriguing idea was injected in 1994 into fracture mechanics of concrete by Carpinteri et al. [44–48]. They studied the effect of the invasive fractal character of crack surface and the lacunar fractal character of microcracking in the FPZ. Although significant doubts were raised with regard to proposals for a fractal explanation of the size effects in structures [16,17,26], the fractal approach might be one useful way to describe the effects of roughness and disorder on the fracture energy of concrete.

The purpose of this article is to survey the basic models for concrete fracture, highlight some recent results, particularly those concerning the size effect, discuss the fracture testing methods, especially the currently debated question of a testing standard, indicate an effective testing procedure, and point out some arguments relevant to practical applications.

Exhaustive coverage obviously cannot be claimed. The references in this survey will be selective. Admittedly, the results from the writer’s home institution (with which he is most familiar) will be overemphasized. A detailed textbook-style exposition of most of the subjects touched in this survey is found in the recent book by Baţant and Planas [40].

### 2. Relative FPZ size and the paramount role of scale

Concrete as well as rock and many other materials—including various fiber composites and particulate composites, coarse-grained or toughened ceramics, ice (especially sea ice), cemented sands, grouted soils, bone, paper, wood, wood-particle board, etc.—requires a different kind of fracture mechanics than metals. In both metal and concrete structures of normal sizes, fracture mechanics is nonlinear due to the development of a sizable nonlinear zone that develops at fracture front. But whereas, in ductile–brittle metals, most of this zone involves hardening plasticity or perfect yielding, and the FPZ, defined as the zone in which the material undergoes softening damage (tearing), is quite small, in concrete and rock fracture the plastic flow is next to nonexistent and the nonlinear zone is almost entirely filled by the FPZ (Fig. 1, from Ref. [12]). Such materials are now commonly called quasibrittle.
As a second difference from metals important for concrete and geotechnical structures, the length of the FPZ, which is equal or proportional to the so-called characteristic length (or material length), \( l \), may occupy a much larger portion of the cross-section of the structure. Often the FPZ may encompass the whole cross-section.

The scale and size govern almost everything in fracture. In normal concrete or coarse-grained rock such as bresca, typically the FPZ length \( l \approx 0.5 \) m; in dam concrete with extra large aggregate, \( l \approx 3 \) m, and about the same holds for horizontally propagating fractures in sea ice; in a grouted soil mass, \( l \approx 10 \) m is possible; and in a mountain with jointed rock (with the joints imagined as continuously smeared), \( l \approx 50 \) m may be typical. On the other hand, for drilling into an intact granite block between two adjacent joints, \( l \approx 1 \) cm; in a fine-grained silicon oxide ceramic, \( l \approx 0.1 \) mm, and in a silicon wafer, \( l \approx 10–100 \) nm. Depending on structure size \( D \), understood as the dimension of the cross-section, different theories are appropriate for analyzing failure. They may be approximately delineated as follows:

\[
\begin{align*}
\text{For } D/l & \geq 100: & \text{LEFM} \\
\text{For } 5 \leq D/l < 100: & \text{nonlinear quasibrittle fracture mechanics} \\
\text{For } D/l < 5: & \text{nonlocal damage, discrete element models, plasticity}
\end{align*}
\]

In the last case, to be more precise, the strength-based plastic limit analysis gives only crude engineering estimates of the small-size behavior, while accurate analysis, at least in theory, calls for nonlocal damage models [25,40], particle models [41], or lattice models.

Thus, for example, a massive rock slide in a 1000 m tall mountain, consisting of a rock with joints spaced about 5 m apart, may be analyzed by LEFM; so can the breaching of a concrete arch or gravity dam 100 m tall, the fracture of a 5 m thick concrete slab of a LNG tank, or the fracture of a fine-grained ceramic part 2 cm large. On the other hand, the punching failure of a concrete slab 50 cm thick, the fracturing of granite at the tip of a drill bit, the failure of a tunnel in a rock mass with joints 3 m apart, the breakout of a deep borehole or mining stope, the fracture of a ceramic part 1 mm large in a micromechanism, the fracture of a silicon wafer 100 nm thick—all necessitate quasibrittle fracture mechanics. The punching failure of a concrete slab 15 cm thick can conveniently be handled according to design code specifications based on plastic limit analysis.

3. Cohesive crack and crack band models

3.1. Material characterization

Conceptually the simplest model to characterize the behavior of a finite-size FPZ is the cohesive crack model (for concrete also known under the name ‘fictitious crack model’ given by Hillerborg). This model is simple enough to be understood even by someone who has no knowledge of fracture mechanics. This is doubtless one reason for its popularity.
The basic hypothesis of the cohesive crack model is that, for mode I fracture, the FPZ of a finite width can be described by a fictitious line crack that transmits normal stress $\sigma(x)$ and that this stress is a function (monotonically decreasing) of the separation $w$ (called also the opening displacement, or opening width);

$$\sigma = f(w)$$

(Fig. 2b, d and e). By definition, $f(0) = f_0'$ = direct local tensile strength of concrete (in ACI notation) or rock. The terminal point of the softening curve $f(w)$ is denoted as $w_f$; $f(w_f) = 0$. In materials science, the cohesive portion of FPZ, in which the cohesive stresses represent the forces transmitted across an almost formed crack due to wedging and pullout of fragments and aggregates from one crack face, is distinguished from the microcracking zone, in which no distinct crack can yet be discerned. For continuum mechanics, though, this distinction is irrelevant.

Function $f(w)$ first descends very steeply and then, roughly at $\sigma \approx 0.15f_0' - 0.33f_0'$, the descent becomes slow (see Fig. 2d and e, for slope change at $f_0'/3$); following Wittmann et al., the CEB-FIP code uses $0.15f_0'$ [40, Fig. 7.2.3b], and Rokugo et al. proposed $0.25f_0'$. The tail of the descending curve is very long. This poses severe problems for the measurement of fracture energy $G_f$ corresponding to the area under the entire curve $f(w)$.

In the initial work of Hillerborg and collaborators in 1976, the softening curve $f(w)$ was described as a decaying exponential with a horizontal asymptote below axis $w$. Later, following Petersson [80], a simple

![Cohesive Crack Diagram](image)

Fig. 2. Stress distributions and softening curves: (a,b) cohesive crack model for ductile–brittle metals; (c,d) cohesive crack model for quasibrittle materials (concrete); (e,f) crack band model for quasibrittle materials.
bilinear form has generally been adopted (Fig. 3a). The stress at slope change is generally considered to be between 0.15$f'_t$ (as recommended by CEB-FIP model code in 1990) and $f'_t/3$.

The area under the entire softening stress-separation curve $f(w)$ represents the total energy dissipated by fracture per unit area of the crack plane, $G_F$ (dimension J/m²), as the crack faces are completely separated at a given point. From this, it is often inferred, without any proof, that the area under $f(w)$ also represents the energy dissipated per unit area of crack plane as the FPZ moves forward by $da$ (the stress profile $\sigma(x)$ within the FPZ changing only infinitesimally). That this is so, however, requires a mathematical proof, which was given by Rice [89] on the basis of his $J$-integral. This is theoretically important since it is the latter energy that must be equal to the energy release rate of the structure.

In the light of the generally accepted bilinear approximation (Fig. 3a) of the softening curve, $f(w)$, the cohesive crack model is characterized by two fracture energies:

$$G_F = \int_0^\infty f(w) \, dw, \quad G_f = \frac{f'_t^2}{2\sigma'_0} = \frac{w_0^2\sigma'_0}{2} \quad \left(\sigma'_0 = \frac{df(0)}{dw}\right)$$  \hspace{1cm} (3)

$G_F$ corresponds to the area under the entire curve $f(w)$, while $G_f$ to the area under the initial tangent of slope $\sigma'_0$, i.e., under the initial steep segment extended down to the $w$ axis (within a realistic size range, it is solely $G_f$ which controls the maximum loads of structures and thus the size effect, as noticed by Planas et al. [85]). For the scaling and size effect, it is important to realize that the fracture energy and material strength imply, according to dimensional analysis, the existence of a fracture characteristic length as a material property [64a]. In view of the bilinear approximation of $f(w)$ (Fig. 3a), concrete possesses two fracture characteristic lengths:

$$\ell_1 = \frac{E'G_f}{f'_t^2} \quad \text{and} \quad \ell_{ch} = \frac{E'G_F}{f'_t^2}$$ \hspace{1cm} (4)

This expression for $\ell_{ch}$, in the sense of the length of FPZ, was introduced already in 1958 by Irwin. While some other parameters of the dimension of length can be formed, a systematic use of $\ell_{ch}$ for concrete was initiated in 1976 by Hillerborg et al. [59]. Typically

$$G_F \approx 2.5G_f, \quad \ell_{ch} \approx 2.5\ell_1$$ \hspace{1cm} (5)
which was concluded by Planas et al. [85] and Guinea et al. [56,57] (see also Ref. [40]), and has recently been confirmed by Bazănt and Becq-Giraudon [23] as optimal for a database involving 238 test series from different laboratories. If the stress at slope change is fixed, the values $G_F$ and $G_f$ suffice to determine the shape of the bilinear softening curve of the cohesive crack model. Assuming the slope change to occur at stress $\psi f'_t$, we have

$$w_f = \frac{2}{\psi f'_t} [G_F - (1 - \psi)G_f], \quad \sigma'_0 = \frac{df(0)}{dw} = \frac{f'_t^2}{2G_f}$$

(6)

Consider, e.g., $\psi = 0.25$ (as proposed by Wittmann et al. [107], and Rokugo et al. [94]; see Ref. [40, Fig. 7.2.3b]) and $G_F/G_f = 2.5$; then

$$w_f \approx \frac{14G_f}{f'_t} \quad \text{or} \quad w_f \approx \frac{5.6G_F}{f'_t}$$

(7)

and the intercept of the initial tangent with the axis $w$ is $w_0 = (1/7)w_f$. One could go on debating endlessly whether, in theory, three parameters, i.e.,

$$G_F, G_f, f'_t \quad \text{or} \quad \ell_{ch}, \ell_1, f'_t$$

(8)

suffice to characterize the cohesive crack model, but practically this is a moot point. The inevitable material variability, scatter, error of testing methods, and mathematical ill-conditioning of the identification problem (yet to be discussed) conspire to make it almost impossible to identify more than three parameters (and in normal testing more than two).

Thus the cohesive crack model is in practice hardly more than a three-parameter model. Although sophisticated experimental techniques have been used to estimate four parameters (e.g., by Guinea; see Ref. [40, Sec. 7.3, p. 171]), the factor $\psi$ and especially the ratio $G_F/G_f$ appear to be quite uncertain. Since a large enough set of test data could rarely be available in engineering applications, the values of $\psi$ and $G_F/G_f$ need to be guessed, and so it appears that the cohesive crack model is in practice essentially a two-parameter model (just like the size effect model and the Jenq–Shah model to be discussed later).

The association of $G_f$ with the initial tangent of the softening law cannot be taken in a strict mathematical sense. One can, for example, check that the bilinear and staircase softening curves in Fig. 3a and b lead in finite element computations to equivalent results. Thus the problem of identifying the softening curve from test data is ill-conditioned unless some constraints are imposed. It must be concluded that the initial straight line that is associated with $G_f$ is not exactly the initial tangent of the softening stress-separation curve but a straight line that gives the least-square fit of the adopted softening curve within approximately the first half or two-thirds of the descent from $f'_t$ to 0 (note that the situation is analogous to the problem of identifying the relaxation spectrum from creep tests; that problem, too, is ill-conditioned unless some constraints are imposed).

In the literature one can also find a function $f(w)$ that is assumed to first rise from some initiation stress $\sigma_0(<f'_t)$ before the peak $f'_t$ is reached and the descent begins (Fig. 3c) [e.g. 78,102]. In that case, however, the fracture model is not an objective material model, satisfying Hillerborg’s premise that a cohesive crack would open at any place at which the maximum principal stress attains the tensile strength, $f'_t$. If a point on the cohesive crack still corresponds to the rising portion, the stress at sufficiently close points on the flanks of the crack must exceed $\sigma_0$, too. This would imply a paradox—many cohesive cracks would have to form next to each other, infinitely close. Any displacement nonlinearity due to hardening before $f'_t$ is reached should properly be modeled by means of inelastic strains within a certain zone of finite width (this is automatically captured by the crack band model and nonlocal models).
3.2. Mathematical formulation and solution

The cohesive crack model leads to the following two integral equations on the crack line with coordinate \( x \):

\[
w(x) = g[\sigma(x)] = -b \int_{a_0}^{a} C(x, \xi)\sigma(\xi) \, d\xi + C_P(x)P
\]

\[
b \int_{a_0}^{a} \kappa(x)\sigma(x) \, dx + P\kappa_P = 0
\]

(9) (10)

where \( g \) is the inverse function to \( f \); \( x \in (a_0, a) \) is the FPZ; \( a \), the tip of the cohesive crack; \( a_0 \), the tip of the notch or pre-existing stress-free crack; \( b \), body thickness in the third dimension; \( C_P(x) \) or \( \kappa(x) \) and \( C(x, \xi) \) or \( \kappa_P \) are the crack-face displacement at point \( \xi \) and the stress intensity factor at \( a \) caused by a pair of unit normal forces applied to crack faces at \( x \) or by unit applied load \( P \); compliance function \( C(x, \xi) \) can be seen as a Green’s function. If the load-point displacement \( u \) is prescribed, one must introduce another integral equation relating \( u, \sigma(x) \) and \( P \) [40].

Eq. (9) is the condition of compatibility of elastic deformation of the body with the given softening stress–displacement law. The second equation means that the total stress intensity factor, \( K_{\text{tot}} \), caused jointly by applied load \( P \) and cohesive stresses \( \sigma(x) \), must vanish. This condition, which ensures that the stress at the cohesive crack tip is finite and that the crack faces close smoothly at the tip, has sometimes been omitted, but this is generally incorrect because it causes the mathematical problem to become indeterminate.

For numerical computations, the crack line is subdivided by nodes, and the compliance (or Green’s) functions are approximated by compliance matrices. To calculate the maximum load, the classical approach has been to consider advances of the crack tip from one node along the crack line to the next, and, for each crack tip location, solve iteratively the load, displacement and opening profile from the matrix approximation of Eqs. (9) and (10), which are nonlinear if the relation between \( \sigma \) and \( w \) is nonlinear. To determine in this manner the effect of structure size \( P \) on the maximum load \( P_{\text{max}} \), the whole calculation may be repeated for many structure sizes \( D \). There is, however, a much shorter way.

The size effect plot can be solved directly if one inverts the problem, seeking the size \( D \) for which a given relative crack length \( \alpha = a/D \) corresponds to \( P_{\text{max}} \). This leads to the equations [73,31]:

\[
D \int_{\alpha_0}^{\alpha} C^{\alpha}(\xi, \xi')v(\xi') \, d\xi' = -g'[\sigma(\xi)]v(\xi), \quad P_{\text{max}} = \frac{\int_{\alpha_0}^{\alpha} v(\xi) \, d\xi}{D \int_{\alpha_0}^{\alpha} C^{\alpha}(\xi)v(\xi) \, d\xi}
\]

(11)

where the first represents an eigenvalue problem for a homogeneous Fredholm integral equation, with \( D \) as the eigenvalue and \( v(\xi) \) as the eigenfunction; \( \xi = x/D, x = \) coordinate along the crack (Fig. 2); \( \alpha = a/D \), \( \alpha_0 = a_0/D \); and \( C^{\alpha}(\xi, \xi'), C^{\alpha}(\xi) \) = dimensionless compliance functions of structure for crack surface force and given load \( P \). Choosing a sequence of \( \alpha \)-values, for each one obtains from Eq. (11) the corresponding values of \( D \) and \( P_{\text{max}} \) [73]. Eq. (11) has also been generalized to obtain directly the load and displacement corresponding, on the load–deflection curve, to a point with any given tangential stiffness, including the

---

1 In an ongoing study at Northwestern University, it is shown that, if the smeared-tip method [40] is used, the cohesive crack model can be very accurately represented for large structure sizes by a continuous \( K \)-density profile along the crack line [42,21,22]. Asymptotically for very large sizes, the correspondence is one-to-one; from the softening curve \( \sigma = f(w) \) one can get the \( K \)-density profile, and from that profile one can recover exactly the same softening curve. It may be that the material fracture characterization in terms the \( K \)-profile instead of the softening function \( f(w) \) could be a viable alternative model for the entire practical size range. The \( K \)-profile is more convenient for analytical solutions than function \( f(w) \).
displacement at the snapback point (point of a vertical tangent), which characterizes the ductility of the structure.

Finite element computations with the cohesive crack model illustrate the transition from failure with a relatively large FPZ for the case of small structures, to the failure with a relatively small FPZ for the case of large structures; see Fig. 4 where the stress profile is normalized by \( f_0' \) and then scaled by factor \( \sigma_N/\sigma_0 \) according to the size effect law in Eq. (17). The process zone for each stress profile stretches from the notch tip, which is at \( a/D = 0.2 \), to the maximum stress point seen on each profile. Note how the maximum stress points move, in relative coordinates, closer to the tip of the notch if the structure size \( D \) is increased. For an infinite size the maximum point approaches the crack tip. This means that the FPZ becomes, in relative coordinates, a point, and so the LEFM must apply. These results confirm that, for large sizes, the LEFM behavior is approached.

Another noteworthy feature of the numerical results with the cohesive crack model in Fig. 4 is that the stress profiles for all the sizes within this very broad size range have a vertical (or almost vertical) stress drop at the notch tip (point \( a/D = 0.2 \)). This documents that the tail of the softening curve \( f(w) \) (i.e., the portion for \( \sigma < \psi_0' \)) is not yet entered at maximum load by the material ahead of the ligament; it has no discernible effect on the maximum load and the corresponding stress profile. This further explains that for a normal size range the maximum load predictions depend only on the initial tangent of the softening curve \( f(w) \) and thus only on \( G_f \), being independent of \( G_F \) (this was shown by Planas et al. [85], and in detail by Guinea et al. [56,57]; see also Ref. [40, p. 172, Fig. 7.2.5, p. 273]).

It follows that, for computing the maximum loads of structures with the cohesive crack model, one needs to know only \( G_f \), or the initial slope of the softening curve \( f(w) \). The tail of the softening curve \( f(w) \) is, for this purpose, irrelevant, and the tail of \( G_F \) is needed only for computing the far postpeak deflections. Vice versa, it follows that size effect measurements can provide only \( G_f \), but not \( G_F \) (although \( G_F \) can be estimated by accepting the relation \( G_F \approx 2.5G_f \)).

3.3. Crack band, nonlocal and lattice models

The crack band model [1,37,38] has been the model most widely used in practice for analyzing the distributed cracking and fracture of concrete and geomaterials, and has been incorporated in a number of
commercial finite element codes (e.g., DIANA, SBETA, ATENA [93,49,50]). Aside from programming convenience, its advantage over the cohesive crack models is that it can take into account triaxial stresses in the FPZ, particularly the normal and shear stresses acting in the directions parallel to the crack plane.

For analyzing the common notched fracture specimens, the crack band model gives essentially the same results (with differences of the order of 1%) as the cohesive crack model and can represent all the fracture test data (peak loads, load–deflection curves) as well as the cohesive crack model [37]. The reason is that the width of a notch or crack has very little effect on the computational results as long as the energy dissipation per unit area of the plane of the band is correct. This is, for example, evidenced by the fact that replacing a notch in a typical fracture specimen with a band of square finite elements of zero stiffness and size about 1/15 of the cross-section dimension gives the stress intensity factor with an error under 1% [24a].

In the crack band model, the ratio of the effective element size, \( h_{\text{ef}} \), to the characteristic length, \( \ell \), is used to adjust the average slope of postpeak softening \( \sigma(\epsilon) \) curve (Fig. 2e and f), so as to make the energy dissipation per unit advance of the crack band independent of the element size. The user specifies the stress–strain relation for the basic case \( h_{\text{ef}}/\ell = 1 \), for which the propagating localized damage band is single-element wide. As far as possible, the size of the finite elements is kept equal to \( \ell \). But if, for computational efficiency, the finite elements need to be larger, then the postpeak portion of the master constitutive diagram (Fig. 2), to which a tensorial constitutive law may be related, is scaled horizontally (i.e., in the direction of strain axis) by the factor \( h_{\text{eq}}/\ell \). This makes the strain-softening steeper and achieves that the energy dissipated by the cracking band per unit advance remain the same (note that the postpeak portion is the difference, for the same stress, of the strain at the postpeak diagram for strain softening from the strain at the unloading diagram emanating from the peak, not from a vertical line dropping down from the peak). In the simple case of a triangular stress–strain diagram with linear strain-softening in uniaxial tension (which is analogous to a cohesive crack with a linear softening), characterized by tangent modulus \( E_i < 0 \), the adjusted postpeak tangent modulus is given by

\[
\frac{1}{E_i} = \frac{2G_f}{h_f^{1/2}} \frac{1}{E}
\]

where \( E \) is the initial Young’s modulus and \( G_f = h_{\text{ef}}(E^{-1} - E_i^{-1})(f'_f/2)^2 \). To better approximate a line fracture, one can also introduce a very narrow cracking band, with \( h_{\text{eq}} < \ell \). The same scaling rule for the postpeak diagram can again be applied, making the strain-softening less steep.

The increase of finite element size, however, has a limit. For a certain large enough critical element size, \( h_{\text{crit}} \), the adjusted postpeak diagram develops a snapback, which would cause spurious instabilities and divergence in static analysis. To avoid the spurious snapback, one can use a vertical stress drop and scale down the strength limit \( f'_f \) to some equivalent value \( f'_{eq} \) such that \( h_{\text{ef}} \) times the area under the stress–strain diagram with the vertical drop remain equal to \( G_f \). It turns out that generally \( f'_{eq} \propto 1/\sqrt{h_{\text{ef}}} \), and for the special case of a triangular stress–strain diagram \( f'_{eq} = f'_f/\sqrt{2\ell_1/h_{\text{ef}}} \) [24a] where \( \ell_1 = E^0G_f/f'_f^2 \) is characteristic length (for extremely large sizes, though, \( \ell_1 \) must be replaced by \( \ell_{\text{ef}} \) because the tail affects the peak load).

The crack band model, same as the cohesive crack model, performs best if the path of the crack band (or the fracture to be approximated by it) is known in advance and if the mesh is laid out so that a mesh line would coincide with this path. If the path is not known a priori, either one can repeat the numerical solutions, updating the mesh layout after each (the accuracy of which is unknown and probably poor), or one can use an orientation-dependent bandwidth correction factor [10,12,49,50]. This factor adjusts the effective value of \( h_{\text{eq}} \) depending on the angle of the crack band with the mesh line, and makes the width different from the average width of the obliquely propagating zig-zag crack band.

In a study in progress at Northwestern University, the microplane constitutive model for concrete (model M4, Ref. [24]) has been adapted for the crack band approach with \( h_{\text{ef}} \neq \ell \). This is achieved by means
of an adjustment of the postpeak stress–strain relation on the microplane level. The adjustment, dependent on the element size, is based on the specified fracture energy of concrete.

A caveat needs to be mentioned with respect to those infrequent situations where distributed cracking does not localize (being stabilized, for example, by a heavy enough reinforcement net or by an adjacent layer of compressed material). In such situations (whose treatment by cohesive crack model requires certain refinements [40, Sec. 8.7.2]), the postpeak strain-softening behavior must not be rescaled. So, in using the crack band approach, the user must separately assess whether it is reasonable to expect localization. This can be done by running the solution twice—first with the strain-softening rescaled, and second not rescaled. Whichever gives a steeper postpeak softening of the structure is more realistic.

There are two other, computationally more intensive, approaches to the modeling of quasibrittle fracture: (1) the nonlocal approach (either of integral type or gradient type), which utilizes strain-softening stress–strain relations (like the crack band model) and converges at arbitrary mesh refinement, and (2) the random lattice model or random particle model. Like the crack band model, these models should, in principle, be able to take into account the triaxial stresses in the FPZ (especially the $T$-stresses), but on top of that, they should also be able to capture changes in the FPZ width due to proximity of the boundaries and the stress field around the FPZ. As far as the geometrical representation of the fracture process goes, the nonlocal and crack band models are more realistic at the beginning, and the cohesive crack model at the end. The nonlocal and lattice approaches, however, are beyond the scope of the present exposition.

3.4. Limitations of cohesive crack model

The cohesive crack model has been widely regarded as a fundamental model providing a yardstick for evaluating the soundness of all other models. This is not quite true, though, for four reasons:

(1) The cohesive crack model is a strictly uniaxial model. This seems just fine for standard notched fracture specimens but is questionable for general applications in structures, in which triaxial stresses arise. For instance, large compressive stresses parallel to the crack, either in the direction of propagation (called sometimes the $T$-stresses) or in the third (orthogonal) direction, often exist in the FPZ (Fig. 5). Such compressive stresses have a strong effect on fracture, and when they exhaust the compression strength $f'_c$, they can even cause splitting fracture by themselves. In a more realistic cohesive crack model, the softening stress–displacement curve would have to depend on the parallel compressive stresses (and perhaps tensile as well)—in presence of a significant parallel compressive stress $\sigma_x$, it should start from a lower strength limit and be scaled down, and for $\sigma_x = f'_c$, the tensile strength limit of the cohesive crack model should be zero. In fact, the softening curve should also depend on the shear stresses $\tau_{xy}$ and $\tau_{xz}$ transmitted across the crack plane. It so happens that the standard fracture specimens used so far give no information on such situations. They all happen to involve negligible parallel compressive stresses.

(2) In the cohesive crack model, a very tortuous crack with the adjacent zone of microcracking and frictional slips (Fig. 5) is replaced by an ideal straight line crack, which introduces some error. Most of the

![Fig. 5. Triaxial stresses in the FPZ and its nonlocal behavior.](image-url)
energy dissipation that goes into the fracture energy $G_f$, assumed to represent the area under the stress–displacement curve of the cohesive crack model, does not in fact occur on the crack line. Rather, it occurs over a rather wide zone at the fracture front, most of it away from the crack line.

(3) The reason for a finite width $h$ of the energy dissipation zone (or the FPZ) probably is not so much the energy dissipated by microcracking on the side of the final crack path but mainly the energy dissipated by frictional slips, which represents at least 50% and perhaps 80% of $G_f$. This is, for example, confirmed by the fact that the displacement does not return to zero upon unloading to zero stress (Fig. 2d); in this regard, note the large difference between $G_f$ and the pure fracture energy, discussed in Ref. [15].

(4) If it is recognized that the true FPZ has a finite width $h$, then it becomes questionable to assume that stress $\sigma(x)$ at point $x$ depends only on the displacement $w(x)$ at the same point $x$ (Fig. 5, top right) and not at all on the displacements at the adjacent points $\xi$ (this is the same weakness as in replacing an elastic half-space by the Winkler foundation). Consequently, a nonlocal generalization of the cohesive crack model may be needed. In that generalization, a concentrated force applied at a point of the boundary of the FPZ, considered as a two-dimensional body isolated from the rest of the structure, would have to cause a displacement not only at that point but also at adjacent points, as sketched in Fig. 5 (left and bottom right).

The last point means that a more realistic stress–displacement relation for the cohesive crack model would have the form:

$$\sigma(x) = \int_{\Omega} Z(x, \xi)f(w(\xi)) \, d\xi$$

(13)

where $\Omega$ is an interaction domain (a line segment) centered at $x$; and $Z(x, \xi)$ is a bell-shaped interaction function (which, in theory, is given by the solution of the stochastic boundary value problem of stress and strain fields within the isolated FPZ).

Numerical stochastic finite element solutions, employing some form of a nonlocal approach, or non-linear fracturing random lattice models, could probably be used to simulate the behavior of the FPZ, in order to check whether there is sufficient justification for a local form of the cohesive law and a unique stress-separation curve.

4. Adaptations of LEFM

The cohesive and crack band models are nonlinear. LEFM per se obviously cannot be directly applied to quasibrittle fracture, but various adaptations of LEFM involving, in addition to fracture energy (or to fracture toughness $K_c$), at least one material parameter, provide useful approximations.

4.1. Equivalent LEFM

The simplest adaptation of LEFM, co-opted from metal fracture, is the so-called ‘equivalent LEFM’, in which the tip of an equivalent LEFM crack—a sharp traction-free crack of length $a$—is not located at the beginning of the FPZ but some distance $\Delta a$ ahead;

$$a_{eq} = a_0 + \Delta a$$

(14)

where $a_0$ is the length of notch or the traction-free portion of crack. Based on the fact that (in contrast to ductile metals) the cohesive stress distribution throughout the FPZ is close to linear, $\Delta a$ is approximately one-half of the FPZ length. The length of the FPZ for an infinitely large specimen, denoted as $2c_f$, is a shape-independent parameter (a material constant), and so $\Delta a = c_f$ for $D \rightarrow \infty$, in general. For any $D$, a constant FPZ length may be assumed as an approximation, and so $\Delta a \approx c_f$ in general (for an analysis of the variation of $\Delta a$, see Refs. [40, ch. 5; 71; 95]).
The maximum load may be characterized in terms of the nominal strength of structure, \( \sigma_N = P_{\text{max}}/bD \). Applying LEFM to the equivalent crack, one has

\[
\sigma_N = \frac{K_c}{\sqrt{Dg(a_0 + \Delta a/D)}}
\]  

(15)

where \( a_0 = b_0/D \); \( g(x) \), where \( x = a/D \) is the dimensionless energy release rate function, is defined as

\[
g(x) = k^2(x), \quad k(x) = bK_f \sqrt{D/P} = \text{dimensionless stress intensity factor} \quad (K_f = \text{stress intensity factor caused by} \ P, \ \text{given for many specimen geometries in handbooks}); \ K_c = \text{fracture toughness} = \text{critical value of} \ K_f \text{ required for propagation}, \ K_c = \sqrt{E'G_f} \ (\text{Irwin relation}), \ E' = E \text{ for plane stress and} \ E' = E/(1 - v^2) \text{ for plane strain}, \ v = \text{Poisson ratio}.

Nallathambi and Karihaloo [77,72] evaluated \( \Delta a \) from compliance measurements on notched three-point bend beams. They proposed an ‘effective crack model’, in which \( \Delta a \) is given by an empirical formula obtained from their tests of notched three-point bend beams. That formula, however, contains the geometrical parameters of notched beams and thus does not have general applicability. We will now briefly outline two generally applicable models.

4.2. Size effect model

Fracture analysis can be made both simple and general by tying it to a model for size effect (understood as the dependence of \( \sigma_N \) on the size \( D \) of geometrically similar structures). Although various fine theoretical points can be treated more accurately [40], the large-size asymptotic laws for the size effect may be easily obtained from Eq. (15) by substituting for \( \Delta a \) the FPZ half-length \( c_f \). Since function \( g(x) \) is smooth, one may expand this function into Taylor series;

\[
g(x_0 + c_f/D) = g(x_0) + g’(x_0) \frac{c_f}{D} + \frac{1}{2!} g''(x_0) \left( \frac{c_f}{D} \right)^2 + \cdots
\]  

(16)

which is an asymptotic power series in terms of \( D \). With regard to the truncation of this series, it is helpful to distinguish the following two basic cases: (1) if \( x_0 \) is large, the Taylor series may be truncated after its second (linear term), (2) if \( x_0 \) vanishes, which occurs at the start of propagation from a smooth surface, then \( g(x_0) = 0 \) and so the series may be truncated only after its quadratic term. Substitution into Eq. (15) then yields two asymptotic size effect laws [40] (the second one after some further approximations):

(a) Large crack: \[ \sigma_N = \frac{\sigma_0}{\sqrt{1 + \beta}}, \quad \beta = \frac{D}{D_0} \]  

(17)

\[
\sigma_0 = \frac{K_c}{c_f g(x_0)}, \quad D_0 = c_f g'(x_0)/g(x_0)
\]  

(18)

(b) Initiating crack: \[ \sigma_N = \sigma_\infty \left(1 + \frac{rD_b}{D} \right)^{1/r} \]  

(19)

\[
\sigma_0 = \frac{K_c}{c_f g'(0)}, \quad D_b = c_f \frac{-g''(0)}{4 g'(0)}
\]  

(20)

where exponent \( r \) has been introduced because it offers a useful degree of freedom while having no effect on the first two terms of the expansion. For geometrically similar structures, \( \sigma_\infty \) and \( D_b \) are constants, and if the cracks or notches are similar, so are \( \sigma_0 \) and \( D_0 \). Parameter \( \beta \), called the brittleness number, provides a geometry independent measure of brittleness of structure, understood as the proximity of behavior to the LEFM. The second law (19) gives the size effect for modulus of rupture (flexural strength).

It so happens that these size effect laws also satisfy the asymptotic conditions for \( D \rightarrow 0 \) which stem from failure criteria in terms of stress and strain (this would not be the case if further terms in the expansion were
including these laws provide what is known as the asymptotic matching [4,5]. Hence, they are suitable approximations for the entire size range.

Knowing $G_f$ and $c_f$ (which implies $K_c = \sqrt{E'G_f}$), the size effect law implies an $R$-curve, and thus a complete fracture model from which the entire load–deflection of structure can be calculated. Such a calculation was presented for an arbitrary size effect law by Bazant et al. [30] and is graphically illustrated in Fig. 6 (left); the energy release curves for $P = P_{\text{max}}$ must contact the $R$-curve, and so the $R$-curve represents the envelope of the $G$-curves for all sizes. Based on this fact, the following equations defining parametrically the $R$-curve that corresponds to the size effect law in Eq. (17) have been derived [29,40]:

$$R(\Delta a) = G_f \frac{g'(z) \Delta a}{c_f} \left( \frac{g'(z_0)}{g'(z)} - \left( z - z_0 \right) \right) $$

(21)

(where $z = a/D = (a_0 + \Delta a)/D$). To plot the $R$-curve, choose a series of values of $z$; for each, evaluate first $\Delta a/c_f$, and then $R(\Delta a)$. The $R$-curve obtained depends, of course, on the structure geometry.

4.3. Jenq–Shah model

Noting a formal analogy with previous models of metals despite differences in physical mechanism, Jenq and Shah [65] adapted to concrete the Wells–Cottrell model [106,50a]. The Jenq–Shah model has been called the ‘two-parameter model’, although the size effect model (without the ‘broad-range’ extension to be discussed later) is also a two-parameter model (so is the cohesive crack model—except if sophisticated experimental techniques that could yield estimates of further parameters are applied). The Jenq–Shah model utilizes the LEFM equation for the opening $w_T$ at $a_0$ of a crack of length $a_0 + \Delta a$, which has the general form $w_T = f(z_0, \Delta a) \sigma_N D/E'$. The basic hypothesis is that failure occurs when the crack-tip opening displacement $w_T$ reaches a certain critical value, $\delta_{\text{CTOD}}$. If this value and $K_c$ are known, Eq. (15) along with another LEFM equation for the opening displacement $w$ at $a_0$ of an LEFM crack of length $a = a_0 + \Delta a$ are, in this model, solved for $\Delta a$ and $\sigma_N$. To determine the material properties $K_c$ and $\delta_{\text{CTOD}}$ experimentally, one needs to measure the peak load of a notched fracture specimen, and also the
compliance for unloading that begins from a state close after the peak load. This compliance decides the value of $a$.

The Jenq–Shah model and the size effect model give very similar numerical results [98]. They are asymptotically identical up to the first two terms of the expansion of size effect in terms of powers of $1/D$ (as demonstrated analytically in Ref. [14]), and are approximately equivalent throughout the entire size range. The values of $K_c$ (or $G_f$) in both models are of course the same. The second parameters of each model are related by [14,28]:

$$\delta_{CTOD} = \frac{\sqrt{32}}{\pi} \frac{K_c}{E} \sqrt{e_f}$$  \hspace{1cm} (22)

5. Broad-range size effect and tail of softening stress-separation curve

The extrapolation to infinite size has often been invoked as a means to define the fracture characteristics unambiguously. However, certain aspects suggesting an update of this view emerged recently. Consider the following points, which bear on the philosophy of testing methods.

1) First, note that the simple size effect law proposed by Bažant et al. [8,9,29] has never been compared to test data exceeding the range of 1:160 (tests of sea ice [51]), and that it cannot be expected to be accurate for a size range exceeding about 1:20 (as already noted on the basis of finite element results discussed in Ref. [11]). Thus, the extrapolation to infinity cannot be taken literally. What should be understood is an extrapolation of the nominal strength values to sizes significantly larger, but not infinitely larger, than the largest size tested (in practice perhaps about 20 to 50 times larger).

2) Consider now the stress profiles ahead of the notch tip that exist at maximum load $P_{\text{max}}$ of specimens of different sizes; see for example the finite element results in Fig. 4. They show that, through the entire size range of 1:100, there is a sudden, or almost sudden, drop of stress at the notch tip. In other words, the stress profile does not begin at the notch tip from zero but from some finite value (Fig. 7, top). This means that the tail of the softening stress-separation curve of the cohesive crack model has no effect on the maximum load, within the size range of 1:100. So there can be no way to identify the tail from size effect measurements over a practical size range. The size effect on the nominal strength of structure depends only on the initial tangent of the softening stress-separation law, i.e., only on $G_f$, and is independent of the tail, i.e., of $G_F$.

![Normal Size](image1)
![Extreme Size](image2)

Fig. 7. Stress profiles at maximum load ahead of notch tip at maximum load for test specimens of normal size and extreme (practically unattainable) size.
Only for extreme sizes (for concrete probably unattainable), the stress at the notch tip at $P_{\text{max}}$ approaches zero (Fig. 7, bottom).

(3) In relative coordinates $x_i/D$, the FPZ in an infinitely large specimen represents a point. So, regardless of the geometry of the structure (with the loading arrangement), the elastic displacement and stress fields surrounding the FPZ must be the asymptotic near-tip fields of LEFM, which depend only on $K_t$ and are independent of the geometry of the structure that has been expanded to infinite size (Fig. 8). The FPZ can respond only to the field to which it is exposed at its boundary. So, for $D \to \infty$, the state of the FPZ must also become independent of structure geometry and size, and the FPZ must dissipate the same amount of energy.

(4) The contact of the fracture equilibrium curve for $P = P_{\text{max}}$ with the $R$-curve must occur, in an infinitely large specimen, on its final horizontal plateau (Fig. 6, left). The fact that $G_f$ is less, in fact much less, than $G_F$ means that, for a size range encountered in normal testing and structural design (not more than 1:100), the fracture equilibrium curve touches the $R$-curve in its rising part, still quite far to the left of the horizontal plateau corresponding to $G_F$.

(5) Since the location of the asymptote of slope $-1/2$ of Bažant’s size effect formula (17) uniquely determines the fracture energy $G_f$ (see Eq. (18)), one must conclude from the preceding points that, for a many-fold, say 1000-fold, size increase, the plot of $\log \sigma_N$ versus $\log D$ must lie significantly to the right of that asymptote (Fig. 9).

In view of these points, one needs to devise a generalized size effect formula which, in the logarithmic scale, would give asymptotes located more and more to the right (Fig. 9) as the size range is extended by each 1 to 2 orders of magnitude. Two such formulae, each with different advantages, were proposed (with any reasoning and discussion omitted) in Ref. [20] (Eqs. (26) and (25); see also Ref. [22]); they may be written as

\[
\sigma_N^2 = \frac{\sigma_0^2}{(1 + D/D_0)G_f} \sum_{k=0}^{n} \frac{\Gamma_k D}{H_k + D} \quad (D_0 < H_1 < H_2 < \cdots < H_n)
\]  

(23)

\[
\sigma_N^2 = \frac{\sigma_0^2}{(1 + D/D_0)G_f} \sum_{k=0}^{n} \Gamma_k (1 - e^{-D/H_k})
\]  

(24)

where $\Gamma_k$ are nonnegative constants of the dimension of fracture energy, and $H_k$ are positive constants of the dimension of length. Since these expressions must reduce for $n = 0$ to the classical size effect law, Eq. (17),

\[2\text{ Eq. (26), as printed in Ref. [20], unfortunately contained a misprint; } D_i^p \text{ in the numerator should be replaced by } D^p.\]
\[ \Gamma_0 = G_f, \quad H_0 = 0 \]

As we will justify later (by discretization of a continuous spectrum of fracture energies), \( H_k \) may always be taken in the form:

\[ H_k = \lambda^k D_0 \]

where \( \lambda \) is a constant (>1) whose value may be expected to be between 10 and 50. For a discussion of some interesting mathematical aspects relevant to the determination of \( H_k \), see Appendix A.

The equation of the large-size asymptotic size effect for the laws in Eq. (23) or Eq. (24) is

\[ \sigma_N^2 = \frac{\sigma_0^2 D_0}{G_f D} \sum_{k=0}^{n} \Gamma_k \]

and for the classical law (17) is

\[ \sigma_N^2 = \frac{\sigma_0^2 D_0}{D} \]

provided that a fixed geometry is considered, i.e., function \( g(x) \) is the same for all \( D \). Equating these expressions to the equation for the small-size asymptotic size effect \( \sigma_0 = \) constant, which may be expressed in terms of \( g(x) \) as \( \sigma_0^2 = E'G_f/c_f g(x_0) \), we find that the inclined small-size and large-size asymptotes intersect the horizontal asymptote at points (Fig. 9):

\[ D_{0n} = \Upsilon G_{Vn}, \quad D_0 = \Upsilon G_f \]

with the constants defined as

\[ G_{Vn} = G_f + \Gamma_1 + \Gamma_2 + \cdots + \Gamma_n = \sum_{k=0}^{n} \Gamma_n, \quad \Upsilon = \frac{g'(x_0)E'}{g^2(x_0)\sigma_0^2}, \]

So, in the standard size effect plot of \( \log \sigma_N \) versus \( \log D \), the intersections of the horizontal asymptote with the asymptotes of slope \(-1/2\) are located at \( \log D_{0n} = \log \Upsilon + \log G_{Vn} \) and at \( \log D_0 = \log \Upsilon + \log G_f \),
respectively (Fig. 9). The additional terms \( k = 1, 2, \ldots, n \) in the size effect law have the effect of pushing the inclined asymptote to the right of the classical size effect law \((k = 0)\) by the distance

\[
\Delta \log G = \log D_{0n} - \log D_0 = \log \frac{G_{F_n}}{G_f} = \log \left(1 + \frac{\Gamma_1 + \Gamma_2 + \cdots + \Gamma_n}{G_f}\right)
\]  

\( (31) \)

The fact that this shift of the asymptote to the right has not been seen in experiments implies that \( H_1 \) must be much larger than \( D_0 \), at least an order of magnitude larger. Setting \( G_F = G_{F_n} \), and considering the experimental observation that \( G_F/G_f \approx 2.5 \), one concludes that the shift between the first and last inclined asymptote is

\[
\Delta \log G = \log \frac{G_F}{G_f} \approx \log 2.5 = 0.398
\]  

\( (32) \)

(Fig. 9) which is not much larger than the width of the scatter band of most size effect data. Within the range of normal testing (typically \( D \lesssim 5D_0 \)), the difference between the original and broad-range size effect curves is negligible compared to inevitable random scatter (Fig. 9). Aside from the fact that \( H_1 \gg D_0 \), this is doubtless the second reason why the size effect testing per se has not indicated any need for adding further terms to the size effect law. The evidence comes from the work-of-fracture tests, particularly the discrepancy between \( G_F \) and \( G_f \), and from numerical simulation [11].

Since experimental evidence gives no more information than the value of \( \Delta \log G \), we may further simply assume that \( \Gamma_1 = \Gamma_2 = \cdots = \Gamma_n = \Gamma \). The logarithmic plot of \( \log \sigma_N \) versus \( \log D \) for \( \lambda = 10 \) and \( n = 2 \) is shown in Fig. 9, along with the case \( n = 0 \) corresponding to the classical size effect law. Note that the difference between the two, compared to the inevitable data scatter, is visually indistinguishable up to sizes \( D = 10D_0 \), which is why the broad-range size effect law cannot be deduced from size effect tests alone.

The broad-range size effect, however, can be obtained numerically with the cohesive crack model. Fig. 10 (left) compares the classical size effect law (17) (solid curve) to the (adapted) numerical results of Li and Bažant [73] (dashed curve) obtained by eigenvalue size effect analysis according to Eq. (11) for Peterson’s [80] bilinear softening law with \( w_f/w_0 \) and slope change at \( 0.3f^* \) chosen such that \( G_F/G_f = 1.595 \). Note (1) that in the normal size range, the difference between the solid and dashed curves, corresponding to the second linear segment in the softening curve shown, is insignificant compared to inevitable data scatter, and

![Fig. 10. Left: Broad-range size effect numerically computed for Peterson’s bilinear softening law, compared to classical size effect law (17) corresponding to linear softening along the initial tangent. Right: Spectrum of fracture energies.](Image)
that a size range of two orders of magnitude is far from sufficient for approaching the large-size asymptote (dashed). These numerical results also demonstrate that for the normal size range the tail of the softening curve is unimportant for size effect computations. Further numerical results with the cohesive crack model have been reported by Hillerborg (discussed in Refs. [10–12]) and Jirásek (private communication by Milan Jirásek, EPFL, Lausanne, 1999), and their fits are shown in Fig. 11; these results indicate the ratio $G_F/G_f \approx 5$ rather than 2.5.

From now on, we consider only limited size ranges for which the broad-range size effect law is superfluous.

6. Statistical scatter of testing methods and approximate prediction of fracture parameters

The test results obtained in one laboratory on one particular concrete have a much lower scatter and can be interpreted much more easily and unambiguously than the aggregate of the test results obtained at various laboratories on various concretes. However, the latter inevitably provides a far broader range, which is a great advantage for statistical studies. But the differences of the test results from their mean value reflect the systematic differences between various concretes and cannot be regarded as the statistical scatter of random errors.

To get a picture of the scatter of random errors in the huge data set including the test results from the literature obtained on many different concretes, one must first eliminate from the data their systematic (deterministic, mean) trends. In other words, one must first find the approximate formulae optimally describing the mean trends of the data. The statistical errors are then approximately the deviations from these formulae.

The deviations of the data from the prediction formula have two causes: (1) an inevitable random scatter of material properties, and (2) an error of the testing method. The former may be expected to be about the same for different testing methods, and so the differences in the error statistics of the prediction formula are caused primarily by differences in the testing method. Thus they may be used for an approximate comparison of the testing methods.

Such a comparison is of course contaminated by an imperfect form of the prediction formula taking into account the differences in concrete composition, etc. But for comparing the errors of different testing methods there is at present no good alternative. A comparably large data set on the fracture characteristics
obtained by different testing methods will hardly ever be accumulated for one particular concrete, and if it will, another question would arise: Can the same error statistics be expected for other concretes?

A recent study [23] was devoted to this problem. It showed that mean \( G_f \), \( G_F \) and \( c_f \) (and thus also \( K_c \) and \( \delta_{C(0)} \)) can be approximately predicted from the standard compression strength, maximum aggregate size, water–cement ratio and aggregate type (river or crushed), with somewhat smaller errors than thought possible. A very large database, consisting of 238 test series, was extracted from the literature and tabulated. Optimization of the fits of this data set led to new approximate prediction formulae, which read:

\[
G_f = a_0 \left( \frac{f'_t}{0.051} \right)^{0.46} \left( 1 + \frac{d_u}{11.27} \right)^{0.22} \left( \frac{w}{c} \right)^{-0.30} \quad \omega_{G_f} = 17.8\%
\]

\[
c_f = \exp \left[ \gamma_0 \left( \frac{f'_t}{0.022} \right)^{-0.019} \left( 1 + \frac{d_u}{15.05} \right)^{0.72} \left( \frac{w}{c} \right)^{0.2} \right] \quad \omega_{c_f} = 47.6\%
\]

\[
G_F = 2.5G_f, \quad \omega_{G_F} = 29.9\%
\]

where \( a_0 = \gamma_0 = 1 \) for rounded aggregates, while \( a_0 = 1.44 \) and \( \gamma_0 = 1.12 \) for crushed or angular aggregates; \( \omega_{G_f} \) and \( \omega_{G_F} \) are the coefficients of variation of the ratios \( G_f^{\text{test}}/G_f \) and \( G_F^{\text{test}}/G_F \), for which a normal distribution may be assumed, and \( \omega_{c_f} \) is the coefficient of variation of \( c_f^{\text{test}}/c_f \), for which a lognormal distribution should be assumed. The standard deviation of the errors of the new formula for fracture energy, compared to the 238 test series from the literature, is lower than that of the older formula in the CEB-FIP Model Code (1990), which was of course developed from a much smaller database, as available in the 1980s.

It must be admitted that the aforementioned coefficients of variation of prediction errors, including that for \( G_f \), are rather high. Therefore, a statistical approach to design is appropriate when these formulae are used. These coefficients of variation are nevertheless not higher than those in the widely used prediction formulae for concrete creep and shrinkage. But note that underestimation of fracture load is usually much more dangerous than underestimation of creep. Therefore it cannot be overemphasized that the present simple prediction formulae are intended only for preliminary design, and only for structures of not too high fracture sensitivity. The final analysis of important and sensitive structures should, of course, always be made on the basis of notched specimen tests performed on the local type of concrete used in the structure.

Fig. 12 shows the diagrams of measured versus predicted values of \( G_f \) and \( G_F \). The data points represent 77 tests relevant to \( G_f \) and 161 relevant to \( G_F \), extracted from the literature. Particularly noteworthy is the comparison of the coefficients of variation \( \omega \) of the differences of test data from the predicted fracture energy values, indicated in the figures;

\[
\omega_{G_F} = 1.67 \omega_{G_f}
\]

This large difference in error statistics confirms that the tail of the softening curve of the cohesive crack model is much more scattered than the peak load.

This is one reason why the choice of a fracture test based on the peak loads should receive priority.

7. Size effect method of testing fracture parameters

7.1. Alternative 1: Geometrically similar notched specimens of different sizes

Inverting the relations in Eq. (17), one can calculate the fracture energy and the half-length of FPZ from the size effect law parameters [29,40]:

\[\text{equation}\]
\[ G_f = \frac{\sigma_0^2}{E} D_0 g(z_0), \quad c_f = D_0 \frac{g(z_0)}{g'(z_0)} \] (37)

The parameters of Jenq–Shah method then follow from Eq. (22) and \( K_c = \sqrt{E/\sigma_f} \).

Eq. (37) provides the basis of the original size effect method [40,92]. In this method, one needs to measure only the maximum loads of notched specimens that have a sufficient size range and are geometrically similar, i.e., have the same values of \( g(z_0) \) and \( g'(z_0) \). The testing is very simple and does not require a servo-controlled testing machine. The measured \( \sigma_N \) values are then fit by the size effect law (17). The fitting can be reduced to linear regression in the plot of \( \sigma_N^2 \) versus \( D \), although nonlinear optimization in the plot of \( \log \sigma_N \) versus \( \log D \) is statistically preferable [40] because of more realistic weighting. The use of linear regression makes it easy to obtain the statistics of errors, particularly the coefficients of variation of \( G_f \) and \( c_f \) (these statistics are automatically provided by standard library subroutines and even hand calculators; the formulae are given in Ref. [40] or Ref. [92]).

Naturally, the size effect model parameters, \( G_f \) and \( c_f \), can alternatively be obtained by using Jenq and Shah’s testing procedure. Their procedure does not require specimens of different sizes nor different notches but is less simple, requiring more sophisticated gauges and a stiff servo-controlled testing machine. The statistics of errors of \( K_c \) and \( \delta_{CTOD} \) are not directly obtainable by linear regression. Their estimates, requiring nonlinear statistical analysis, are therefore harder to calculate and less reliable. This is a drawback in view of the large uncertainty in fracture parameters.

Should the broad-range size effect law be used to identify \( G_f \) from \( \sigma_N \) data? It could be done but the only result would be increased complexity. Fitting maximum load data with the broad-range size effect law would of course yield directly \( G_f \), rather than \( G_f \). But what is needed for maximum load calculations is \( G_f \), which would have to be estimated as \( G_f \approx 0.4G_f \).

7.2. Alternative 2: Size effect method for geometrically dissimilar notched specimens

Consider now that specimens of different sizes and different geometries are tested for maximum loads. Then formula (17) to be fit to the test data must be rewritten as
\[ \sigma_i = \sqrt{\frac{E G_f}{g'_i c_f + g_i D_i}} \]

where integer subscript \( i \) refers to specimen number and the following abbreviated notations are introduced:

\[ \sigma_i = \sigma_{Ni}, \ g_i = g_i(z_{0i}) \text{ and } g'_i = g'_i(z_{0i}) = \frac{d g_i}{dz} \text{ at } z = z_{0i}, \ D_i \text{ is the size of specimen } i; \] and \( g_i(z) = k_i^2(z) = \) dimensionless energy release function of specimen \( i \), \( k_i(z) \) being the dimensionless stress intensity factor.

Least-square fitting of Eq. (38) to the measured \( \sigma_i \) values provides the values of \( G_f \) and \( c_f \). This is a nonlinear problem which can be easily handled, for instance, by applying the standard library subroutine for the Levenberg–Marquardt optimization algorithm. However, conversion to a linear regression problem is desirable in the interest of statistics as well as clarity. This can for example be achieved by rearranging formula (38) as follows:

\[ F_i(G_f, c_f) = g_i D_i + g'_i c_f - E G_f \sigma_i^{-2} = 0 \]

where \( F_i(G_f, c_f) \) is a function of \( G_f \) and \( c_f \). The right-hand side is zero only for theoretically perfect data. In practice it is nonzero and its square should be minimized. Thus the condition for the optimum values of \( G_f \) and \( c_f \) may be written as

\[ \Phi = \sum_{i=1}^{N} F_i^2(G_f, c_f) = \min \]

where \( \Phi \) is a function of \( G_f \) and \( c_f \). The minimizing conditions are \( \partial \Phi/\partial G_f = 0 \) and \( \partial \Phi/\partial c_f = 0 \). They provide the following linear regression equations:

\[ A_{11} G_f + A_{12} c_f = C_1 \quad A_{21} G_f + A_{22} c_f = C_2 \]

in which

\[ A_{11} = E^2 \sum_i \sigma_i^{-4}, \quad A_{12} = A_{21} = -E' \sum_i g'_i \sigma_i^{-2}, \quad A_{22} = \sum_i g_i^2 \]

\[ C_1 = E' \sum_i g_i D_i \sigma_i^{-2}, \quad C_2 = -E' \sum_i g_i g'_i D_i \]

Solution of the linear equation system provides \( G_f \) and \( c_f \).

Multiplying the expression for \( F_i \) in Eq. (39) by \( \sigma_i \) or \( \sigma_i^2 \), or dividing it by \( g_i \) or \( g'_i \), one can obtain different linear regression equations. For perfect data, they would of course give the same results but, because of scatter in the real data, the results are slightly different. The reason is that these different regressions imply different weights for different \( i \). Compared to the case with the expression for \( F_i \) multiplied by \( \sigma_i^2 \) (or divided by \( g_i \)), the regression defined by Eqs. (42) and (43) gives higher weights for the domain of larger specimen sizes and not too short notches. This is desirable because the experimental scatter in that domain is generally lower.

7.2.1. Decreasing required size range by using dissimilar notched specimens

The regression would fail if the range of data would not be commensurate to the scatter width. The wider the scatter, the broader the required range, and the broader the range, the better the accuracy. If all the specimens are geometrically similar, then the typical scatter in concrete fracture tests requires the size range to be at least 1:4.

If the specimens are not geometrically similar, then the range of data is not measured merely by sizes \( D_i \). Rather, it must be measured by the range of the brittleness numbers \( \beta_i = D_i / D_{0i} = \gamma_i (D_i / c_f) \) with \( \gamma_i = g_i / g'_i \) [13,40]. Since \( c_f \) is the same for all the specimens, the range can be measured as the range of the effective specimen sizes.
\[ D_i = \gamma_i D_t, \quad \text{with } \gamma_i = \frac{g_i}{g_t} \]

Thus it transpires that, by using different relative notch depths \( z_0 = a_0/D_t \), one can broaden the range [29]. To determine how much, one needs to consider the diagrams of brittleness number parameter \( \gamma(x) = g(x_0)/g'(x_0) \) as a function of the relative notch depth \( x_0 \). For typical fracture specimens, such diagrams are plotted in Fig. 13 [33], in which 3PB denotes notched three-point bend beams of span-to-depth ratio \( L/D \); the ‘pure-bend’ is the case of a uniform bending moment; and EC the eccentrically compressed notched specimen. The broadest range can be obtained with the last, but the fact that the FPZ is forced to be small might be a disadvantage, and so we leave that case out of consideration.

The most popular at present seem to be the notched three-point bend specimens. A zero notch length gives the modulus of rupture test, whose relation to fracture tests leads to a more intricate problem [32,33]. Very small notches lie in a transition range in which the size effect is complicated (see the universal size effect law, e.g. Refs. [14a] and [40]).

The shortest notch for which standard fracture analysis is justified probably is \( z_0 = 0.1 \). According to Fig. 13, the ratio of the values of \( H(x) \) at the maximum, which occurs at \( x \approx 0.35 \), and at \( x = 0.1 \) is about 2, for the three-point bend specimens. So the size range can be approximately doubled by varying the relative notch depth. This point was experimentally demonstrated by Tang et al. [97].

Therefore, it is proposed here to use two different beam depths, with \( D_2/D_1 = 2 \), for each of which two different relative notch depths, \( z_0 = 0.35 \) and \( z_0 = 0.1 \), are used. This gives four combinations, of which the small specimen with the shorter notch can perhaps be omitted. This leaves at least three types of specimens to test and achieves the range 1:4 in terms of the brittleness number.

Merely by varying the notch length in specimens of one size, accurate enough results cannot be achieved. This fact can be deduced from the diagrams in Fig. 13. It can also be graphically explained by Fig. 6, as follows.

The \( G \)-curves are the curves of energy release rate at \( P = P_{\text{max}} \) as a function of total crack length \( a = a_0 + c \) where \( c = \Delta a = \) crack extension from the notch. These curves envelop the \( R \)-curve, and the complete envelope for specimens of given geometry completely characterizes their fracture behavior. The \( G \)-curves at \( P = P_{\text{max}} \) for specimens of sizes \( D = D_1, D_3 \) and similar notches are shown on the left of Fig. 6, and those for various notch depths \( a_0 \) at constant size \( D = D_2 \) on the right. The test specimens are placed so that the notch tip line up with the origin of \( R \)-curve.
Fracture testing defines the several $G$-curves for $P_{\text{max}}$, and the family of $G$-curves for various $D$-values or various $a_0$ values is an envelope of the $R$-curve. Now, each $G$-curve begins from 0 at the point lined up in the figure with the crack mouth. It is now graphically clear that, by varying the notch length at constant $D$, the zero point of the $G$-curves can be moved only little (not more than the specimen length), thus providing only a rather limited range $\Delta R_a$ of the contact points on the $R$-curve. In other words, the family of $G$-curves for various $a_0$ values in a specimen of one size envelops only a small portion of the $R$-curve. By contrast, by varying the specimen size $D$, the zero point of the $G$-curves can be moved to the left or right arbitrarily. Thus the $G$-curve can be made to contact the $R$-curve at an arbitrarily small slope if the specimen is large enough, and arbitrarily small slope if the specimen is small enough. Although the practical maximum size is limited, a much broader range $\Delta R_D$ of the contact points on the $R$-curve can be achieved on the left of Fig. 6 than on the right.

The relation of the proposed testing with $D_2/D_1 = 2$ and $a_0 = 0.1, 0.35$ to the $R$-curve is explained by Fig. 14 where portions of three-point bend specimens are placed below the $R$-curve so that the notch tip would line up with the origin. The zero points of the $G$-curves again line up with the crack mouths. For span-to-depth ratio 2.5, the $G$-curves are almost straight (see the calculations plotted in Fig. 4 in Ref. [33]). The range between the upper two contact points is obtained by doubling the size, and the range between the lower two contact points is obtained by testing with a shorter notch for the smaller size. The entire range $\Delta R$ achieved is about the same as in testing geometrically similar beam (with similar notches) of size range 1:4, which is known to be sufficient.

![Diagram](image)

Fig. 14. Curves of energy release $G$ and range $\Delta R$ of $R$-curve envelope achieved with proposed testing with two specimen sizes and two relative notch depths.
7.3. Alternative 3: Combining one-size notched and unnotched maximum load tests?

The limit value \( \lim \sigma_N = \sigma_0 \) of the size effect law (17) for \( D \to 0 \), defines the horizontal asymptote, i.e., the case of no size effect. Thus, in theory, this value could be obtained by plastic limit analysis. However, there are two problems: (a) the value of material strength to be used in such analysis is not the standard material strength, and (2) it depends on the specimen shape [33,40, discussion below Eq. (9.3.19); also see Section 9]. This problem can be overcome by

- agreeing to fix one standardized shape of the notched specimen (e.g., the three-point bend specimen of span-to-length ratio 2.5 and notch depth 0.35\( D \)), and
- specifying, on the basis of experiments to be carried out before the testing standard is introduced, the ratio

\[
\rho = \frac{\sigma_0}{f_i^0(D)} \quad \text{with, e.g., } f_i^0(D) = \rho f_i(D) = \rho \frac{6}{bD^2} M_u(D)
\]

Here \( f_i(D) \) is the modulus of rupture of unnotched beams of the same size \( D \) and the same external shape as the notched beams, and \( M_u(D) \) is the maximum bending moment measured on these specimens. Alternatively, it might be preferable to use \( f_i^B(D) = \rho f_i^B(D) \) where \( f_i^B(D) = \) Brazilian split-cylinder strength (because, according to a private communication by J. Planas, 2001, the prepeak nonlinearity in the Brazilian tests is less sensitive to the type of concrete than it is in the modulus of rupture test).

This method would be the simplest but its general applicability hinges on the assumption that the \( \rho \) value (for a fixed specimen size and external shape) is independent or almost independent of the type of concrete. Plausible though this assumption may seem, it would have to be thoroughly checked by experiments before a testing standard could be introduced.

Should future testing reveal rather different \( \rho \) values for different types of concretes, one could define several classes of concrete and specify a different empirical \( \rho \) value for each. This could be done on a regional basis; for instance, if a Chicago firm wanted to use a fracture test for the quality control of pavements in the field, it could order tests to measure \( \rho \) for its particular concrete.

If the ratio \( \rho \) is known, the testing becomes extremely simple. The following equations [32], ensuing from Eq. (17), may be used:

\[
G_f = \frac{g(x_0)}{E} D_0 \left[ \frac{f_i^0(D)}{\sigma_N} \right]^2, \quad c_f = \frac{g(x_0)}{g(x_0)} D_0
\]

in which

\[
D_0 = \frac{1}{\left[ \frac{f_i^0(D)}{\sigma_N} \right]^2 - 1}
\]

Here \( \sigma_N \), or \( f_i(D) \), is the average of the nominal strength values, or the modulus of rupture values, for all the notched specimens tested. No regression is needed.

To improve accuracy, one could use a variant combining alternative 2 with alternative 3—tests with two notch lengths \( \alpha_{01} = 0.35D \) (\( i = 1 \)) and \( \alpha_{02} = 0.15D \) (\( i = 2 \)) in beams of the same depth \( D \), and the modulus of rupture test as the third test (\( i = 3 \)). In that case the evaluation can be carried out according to the linear least-square equations (41) and (42) with \( i = 1, 2, 3 \) (\( n = 3 \)) and \( \sigma_{N3} = \rho f_i(D) \), in which for \( i = 3 \) the effective value \( D_3 = 0 \) giving the strength (or plasticity) theory asymptote is formally substituted.
8. Comparison and choice of optimal testing method

Among the available experimental approaches to measuring the fracture properties of concrete [40, Sec. 7.2.2 and Sec. 7.3], the basic ones may be listed as follows:

(1) The work-of-fracture method [60–62,90], which infers the area $G_f$ under the stress-separation curve from the complete load–deflection curve of a notched specimen, as proposed for concrete by Hillerborg et al. [59] and introduced for ceramics by Nakayama [76] and Tattersall and Tappin [99].
(2) The Jenq–Shah method [65,92], which is an adaptation to concrete of Wells’ [106] and Cottrell’s [50a] approach to metals.
(3) The size effect method [13,39,92], which is implied by the size effect model [8,9].
(4) The direct tensile test method (Fig. 15).

The last is an attempt to measure directly the stress-separation curve of the cohesive crack model on a specimen whose cross-section undergoing fracture is nearly in a state of uniaxial tension. That idea (Fig. 15), however, has not met with success. If the direct tensile specimen is large enough to allow the FPZ to develop over its length and width without hindrance of the boundaries, the crack separation does not proceed simultaneously over the cross-section; the specimen inevitably flexes sideways and deforms non-uniformly, and the fracture does not occur simultaneously over the cross-section but propagates across. To avoid it, it has been attempted to use very small specimens bonded to very stiff platens in a very stiff frame (Fig. 15). Then (in the sense of statistical of scatter in the microstructure), a simultaneous separation can be achieved but the $\sigma(w)$ curve that is measured is not relevant for fracture of real structures because the FPZ development is hindered by the boundaries, affected by their geometry.

The effective $\sigma(w)$ curve characterizing the FPZ can, as a matter of principle, be unique only if the field of stress and strain around the FPZ is geometry independent. Consequently, this field must be the LEFM near-tip field, which exists only if the specimen is very large (infinitely large). But in the very small direct tension specimen, the field of stress and strain is different and is strongly sensitive to specimen geometry. Such a test yields information on the behavior of the small specimen but cannot provide a unique, geometry-independent, $\sigma(w)$ curve that could be regarded as a material property. Therefore, our attention will from now on be focused only on the other methods.

![Diagram](image)

Fig. 15. Left: The idea of direct measurement of the softening stress-separation curve. Middle: Specimen needed to achieve nearly simultaneous separation. Right: Field surrounding the FPZ in a large structure.
8.1. Choice of specimen type and T-stresses

The choice of specimen shape is a secondary issue, but the effect of large compressive T-stresses (i.e., crack-parallel normal stresses) is probably not. For metals, by contrast, the T-stresses are unimportant (except for the stability of a straight propagation path).

Evidently, the standard test should primarily yield information on the case of negligible T-stresses, which represents the reference case. All the usual notched fracture specimens generally have negligible T-stresses. For the Brazilian cylinder splitting tests, Rocco’s results [40, p. 182] suggest that the moderate T-stresses in these tests do not have a appreciable effect. Some limited test results [100] suggest that T-stresses less than about 0.7f’ have a negligible influence. It is nevertheless clear that when the T-stress becomes infinitely close to the compression strength f’ of the material, it alone can cause failure, and so the effective fracture energy for tensile mode I must vanish. Preliminary finite element results with model M4 indicate that large T-stresses can have a significant effect.

8.2. Work-of-fracture versus size effect and Jeng–Shah methods

Although the near-tip LEFM field does not exist around the FPZ in the test specimens of the size effect method, this field is implied by the extrapolation to infinity. So it is in the Jeng–Shah method, because the measurement of δctod implies asymptotic equivalence to the size effect model, and thus extrapolation to infinity.

As shown in the preceding analysis, generally three parameters, \( G_f, c_f \) and \( G_F \), are needed for predicting the structural response. Ideally, both \( G_f \) and \( G_F \) should be measured. Trading convenience for a larger error, one may nevertheless choose to measure directly only one of these two parameters. In that regard, note the following points:

1. For predicting the maximum load of notched test specimens or structures with pre-existing traction-free cracks (such as old cracks relaxed by creep, or fatigued cracks), \( G_F \) can be dispensed with. This corresponds to replacing the softening stress-separation curve by its initial tangent. However, \( G_f \) (or knowledge of the initial tangent) can never be dispensed with.
2. Because of the foregoing observation, and even more because \( G_f \) has a much smaller random scatter than \( G_F \) (as manifested in the coefficients of variation), a standard fracture test should primarily predict \( G_f \).
3. Measuring \( G_F \) alone, by the work-of-fracture method, and then assuming the shape of the softening curve (e.g., the relative location of the ‘knee’ point on the curve, or the relative slope of the initial tangent of the softening stress-separation curve), is tantamount to assuming the ratio \( G_F/G_f \approx 2.5 \). One implies the other. It has been a widespread misconception that testing the size effect on nominal strength gives no information on the complete softening curve. But, as a matter of fact, it probably gives as good information as the work of fracture test, and with less random scatter.
4. A standard procedure for the work-of-fracture method should better prescribe the use of two specimen sizes (with a minimum size ratio of 1:2), in order to better assess the coefficient of variation of errors.
5. Because of its high statistical scatter, a direct measurement of \( G_F \) by the work-of-fracture test might not yield more accurate results than the foregoing estimate based on measured maximum loads only.

For a standard experimental procedure to determine \( G_F \), there are two alternatives to choose from:

- measure the postpeak load–deflection diagram to obtain data for the work-of-fracture method (the peak load obtained in the process should always be used to check the initial tangent of the softening law by the size effect method);
• or do not proceed with loading beyond the peak and utilize the simple estimate $G_F \approx 2.5G_f$, from which the softening curve parameters $w_f$ and $w_0$ may further be calculated using Eq. (6).

The latter would obviously be more user friendly, especially for field testing in the industry [97].

8.3. Size effect versus Jenq–Shah method

For a standard experimental procedure to identify $G_F$ and $c_f$ (or, equivalently, $K_c$ and $\delta_{CTOD}$), two methods are available: (a) the size effect method (in three possible alternatives), and (b) the Jenq–Shah method. To compare them, the following points should be noted:

1. The Jenq–Shah method requires a stiff loading system, a servo-controlled testing machine, and precise deformation measurements, while the size effect method requires neither. Thus the experimental procedure for the latter is simpler, more robust, and foolproof.
2. Because of its simplicity, the size effect method may be expected to be less prone to testing errors. This is especially important for quality control tests in the field (such as those at the Texas Division of Highways; Ref. [97]). Such tests should be the principal objective of any standard.
3. If the empirical parameter $\rho$ for alternative 3 were calibrated by testing on many concretes, the size effect method would clearly be the optimal choice.
4. However, until $\rho$ is calibrated, the size effect method, unlike the Jenq–Shah method, requires producing specimens of two different sizes, having the size range at least 1:2 (alternative 2). But that does not seem to be a very serious obstacle.
5. There has been a misconception that the largest specimen for the size effect method must be much larger than that for the Jenq–Shah method or the work-of-fracture method. This is not really true. In fact, the largest specimen size appropriate for each is the same. As a justification, note that:

• The largest size in the size effect method is needed to approach the LEFM conditions. But this is needed equally for all the methods—in the Jenq–Shah method, to obtain $K_c$; in the size effect method to obtain $G_F$; and in the work-of-fracture method to have an FPZ that behaves without interference of the boundaries.
• The size effect method needs adding tests of smaller sizes (or shorter notches) in order to determine $c_f$, whose main purpose is to characterize the transition to nonbrittle behavior as the size is reduced.
• The Jenq–Shah method needs to measure $\delta_{CTOD}$ for exactly the same purpose; it can be said that the measurement of $\delta_{CTOD}$ replaces the need for testing smaller specimens (or shorter notches).

9. Quantitative two-sided asymptotic matching of cohesive crack model

The testing of fracture properties as well as analytical solutions of maximum loads of structures undergoing cohesive fracture could be further improved by anchoring an asymptotic matching formula quantitatively to a zero-size limit of the nominal strength corresponding to the cohesive crack model. The energetic size effect laws developed so far are anchored quantitatively only in the large-size asymptotic behavior. For the small-size limit, the asymptotic condition required has been the finiteness of the asymptotic small-size nominal strength corresponding to the strength theory of plasticity. But the value of that strength has not been imposed because it was felt that, for cross-sections smaller that about three times the size of material inhomogeneities, general solutions would be cumbersome, as they would have to be based on a lattice model of the microstructure (or, less realistically, on the continuum theory of plasticity), for which we do not have simple and general solutions.
The small-size limit, however, should be seen in a different light. The purpose of asymptotic matching is to approximate the behavior in the middle range between two asymptotic situations. An approximation for the middle range will be better if it satisfies both the small-size and large-size asymptotic properties of the theory applicable in that same range. This leads to a rethinking of the small-size asymptotics. It should be quantitatively based not on the lattice model (or the theory of plasticity) but on the same theory that governs the behavior in the middle size range—the cohesive crack model itself. In this light, the small-size asymptotics of the cohesive crack model will now be established (as outlined at a recent conference, Ref. [21]).

9.1. Small-size asymptotics of cohesive crack model

The static boundary value problem of linear elasticity is defined as follows:

\[ \sigma_{ij} = E_{ijkl}(u_{i,j} + u_{j,i}), \quad \sigma_{ij} + f_i = 0 \quad \text{(in } \tilde{\mathcal{V})} \]

\[ n_j \sigma_{ij} = p_i \quad \text{(on } \Gamma_s) \], \quad u_i = 0 \quad \text{(on } \Gamma_d) \]

Here, \( x_i \) are the Cartesian coordinates \( (i = 1, 2, 3) \); \( \sigma_{ij} \), stress tensor components; \( (1/2)(u_{i,j} + u_{j,i}) = \epsilon_{ij} \), strain tensor components; \( E_{ijkl} \), elastic moduli; \( f_i \), body forces; \( p_i \), surface tractions, prescribed on surface domain \( \Gamma_s \); \( n_i \), unit normal of the surface; and \( \Gamma_d \) is the surface domain where the displacements are fixed by supports.

We will consider again geometrically similar structures of various sizes \( D \) and introduce the following dimensionless coordinates and variables, labeled by an overbar;

\[ \bar{x}_i = x_i / D, \quad \bar{u}_i = u_i / D, \quad \bar{\sigma}_{ij} = \sigma_{ij} / \sigma_0 \]

\[ \bar{p}_i = p_i / \sigma_N, \quad \bar{f}_i = f_i D / \sigma_N, \quad \bar{E}_{ijkl} = E_{ijkl} / \sigma_0 \]

\( \bar{p}_i \) is a size-independent distribution of the surface tractions on \( \Gamma_s \), and \( \bar{f}_i \) is a size-independent distribution of body forces in volume \( \tilde{\mathcal{V}} \). The surface normals \( n_i \) at homologous points are independent of size \( D \) (and thus need no overbar). Interchanging \( \sigma_0, \sigma_N \) and \( E, D \) and \( \ell \), etc., one could introduce other sets of dimensionless variables, but only the foregoing one leads to a result with simple physical interpretation.

Denoting \( \partial_i = \partial / \partial \bar{x}_i \) = partial derivatives with respect to the dimensionless coordinates, and noting that \( \partial / \partial \bar{x}_i = (1/D) \partial_i \), we can transform the foregoing equations to the following dimensionless form:

\[ \bar{\sigma}_{ij} = \bar{E}_{ijkl}(\partial_i \bar{u}_j + \partial_j \bar{u}_i), \quad \partial_i \bar{\sigma}_{ij} + \bar{f}_i / \sigma_0 = 0 \quad \text{(in } \tilde{\mathcal{V}}) \]

\[ n_j \bar{\sigma}_{ij} = \bar{p}_i \sigma_N / \sigma_0 \quad \text{(on } \bar{\Gamma}_s), \quad \bar{u}_i = 0 \quad \text{(on } \bar{\Gamma}_d) \]

where \( \tilde{\mathcal{V}} \) is the domain of structure volume in the dimensionless coordinates, and \( \bar{\Gamma}_s \) and \( \bar{\Gamma}_d \) are the surface domains in dimensionless coordinates corresponding to \( \Gamma_s \) and \( \Gamma_d \).

Let coordinates \( x_i \) be positioned so that the crack would lie in the plane \( (x_1, x_3) \) and that the tip of the cohesive crack (and not the notch tip) would lie at \( x_1 = 0 \). For a small enough \( D \), the crack-bridging stress \( \sigma > 0 \) along the whole crack length \( L \), and if \( D \) is small enough and if the compression strength is unlimited, the cohesive crack at maximum load will occupy the entire cross-section or, in the case of a notch, the entire ligament; then the dimensionless crack length \( \bar{L} = L / D \) = constant (if the compression strength is limited and the cross section is for instance subjected to bending, \( L / D \) will not necessarily be size independent but we may assume it to be such, as an approximation for small \( D \)).

In the case of cohesive fracture, Eqs. (52) and (53) must be supplemented by two conditions: (1) The dimensionless total stress intensity factor \( \bar{K}_t = K_t \sqrt{D} / \sigma_N \) produced jointly by the applied load and by the stresses \( \bar{\sigma} = D \bar{\sigma}_2 \) acting on the crack faces must vanish, i.e., \( \bar{K}_t = 0 \). (2) The cohesive (crack-bridging) stresses \( \sigma \) must satisfy the softening law of the cohesive crack, i.e., the curve relating \( \sigma \) to the opening displacement
$w = 2u_2$ on the crack plane. For the small-size asymptotics we need to consider only the initial softening, which can generally be of the form:

$$
\sigma = \sigma_0 [1 - (w/w_f)^p] \quad (\text{for } -L \leq x_1 < 0, x_2 = 0)
$$

where $p, w_f$ are positive constants, and $\sigma_0$ is the tensile strength (also denoted as $f_i$). In terms of the dimensionless variables corresponding to Eq. (50),

$$
\bar{\sigma} = 1 - (\bar{D} \bar{w})^p \quad \text{with} \quad \bar{\sigma} = \sigma/\sigma_0, \quad \bar{w} = w/D, \quad \bar{D} = D/w_f \quad (\text{for } -\bar{L}/D \leq \bar{x}_1 < 0, \bar{x}_2 = 0)
$$

Let us now assume that the dimensionless displacements, stresses and total stress intensity factor approach their limit for $\bar{D} \to 0$ as power functions of $\bar{D}$ with exponent $p$, and try to verify the correctness of this assumption. For small enough $\bar{D}$, we set:

$$
\begin{align*}
\sigma_N &= \sigma_{N0} + \sigma_{N1} \bar{D}^p, \\
\bar{\sigma}_{ij} &= \bar{\sigma}_{0ij} + \bar{\sigma}_{1ij} \bar{D}^p, \\
\bar{\sigma} &= \bar{\sigma}^0 + \bar{\sigma}^p \bar{D}^p
\end{align*}
$$

$$
\begin{align*}
\bar{u}_i &= \bar{u}_{0i} + \bar{u}_{1i} \bar{D}^p, \\
\bar{w} &= \bar{w}^0 + \bar{w}^p \bar{D}^p, \\
\bar{K}_r &= \bar{K}_{r0} + \bar{K}_{r1} \bar{D}^p
\end{align*}
$$

where $\bar{\sigma}_{0ij}, \sigma_{0N}, \bar{\sigma}_{0ij}, \sigma_{0N}, \ldots, \bar{K}_{r1}$ are size independent. These expressions may now be substituted into Eq. (55), (52), (53) and $\bar{K}_r = 0$. The resulting equations must be satisfied for various small sizes $\bar{D}$. For $\bar{D} \to 0$, the dominant terms in these equations are those of the lowest powers of $\bar{D}$, which are those with $\bar{D}^0$ and $\bar{D}^p$. By collecting the terms without $\bar{D}$ and those with $\bar{D}^p$, we obtain two independent sets of equations. It so happens that each of these two sets defines a physically meaningful boundary value problem of elasticity for a body with a crack (which proves our assumption made in Eqs. (56) and (57)).

**Elasticity problem I:** By isolating the terms that do not contain $\bar{D}$ (i.e., contain $\bar{D}^0$), we get:

$$
\begin{align*}
\bar{K}_i &= 0, \\
\bar{\sigma} &= 1 \quad (\text{for } -\bar{L} \leq \bar{x}_1 < 0, \bar{x}_2 = 0)
\end{align*}
$$

$$
\bar{\sigma}_{ij} = \bar{E}_{ijk} \left( \partial_{\bar{x}_k} \bar{u}_i^0 + \partial_{\bar{x}_i} \bar{u}_j^0 \right), \\
\partial_{\bar{x}_i} \sigma_{ij}^0 + f_{ij} \sigma_N^0/\sigma_0 = 0 \quad (\text{on } \bar{\Gamma}_s), \\
n_j \sigma_{ij}^0 = \bar{p}_j \sigma_{ij}^0/\sigma_0 \quad (\text{on } \bar{\Gamma}_d)
$$

**Elasticity problem II:** By isolating the terms that contain $\bar{D}^p$, we get:

$$
\begin{align*}
\bar{K}_i &= 0, \\
\bar{\sigma}' &= -(\bar{w}^0)^p \quad (\text{for } -\bar{L} \leq \bar{x}_1 < 0, \bar{x}_2 = 0)
\end{align*}
$$

$$
\begin{align*}
\bar{\sigma}'_{ij} &= \bar{E}_{ijk} \left( \partial_{\bar{x}_k} \bar{u}_i' + \partial_{\bar{x}_i} \bar{u}_j' \right), \\
\partial_{\bar{x}_i} \sigma_{ij}' + f_{ij} \sigma_N^0/\sigma_0 = 0 \quad (\text{on } \bar{\Gamma}_s), \\
n_j \sigma_{ij}' = \bar{p}_j \sigma_{ij}^0/\sigma_0 \quad (\text{on } \bar{\Gamma}_d)
\end{align*}
$$

Note that parameter $\bar{w}^0$ does not appear in this problem.

The role of stresses and displacements is played by $\bar{\sigma}_{ij}^0$ and $\bar{u}_i^0$ in problem I, and by $\bar{\sigma}_{ij}'$ and $\bar{u}_i'$ in problem II. In problem I, the crack faces are subjected to fixed uniform tractions equal to 1. In problem II, in which $\sigma'$ plays the role of the cohesive stress, the crack faces are subjected to tractions $-(\bar{w}^0)^p$ which vary along the crack faces but can be determined in advance from the $\bar{w}^0$-values obtained in solving problem I. The fact that isolation of the terms with the zeroth power and the $p$th power of $\bar{D}$ happens to yield two separate boundary value problems of elasticity is crucial. The rest of the argument is easy and may be stated as follows.

The magnitude of the loads (surface tractions and body forces) is proportional to $\sigma_{N0}$ in problem I, and to $\sigma_{N1}$ in problem II. These elasticity problems are known to have a unique solution. If $\sigma_{N0}$ were zero, i.e., if the applied load in problem I vanished, the crack face tractions equal to 1 would cause $K_{r0}$ to be nonzero, in violation of Eq. (58). Likewise, if $\sigma_{N1}$ were zero, i.e., if the applied load in problem II vanished, the non-uniform crack face tractions $-(\bar{w}^0)^p$ in problem II would cause $K_{r1}$ to be nonzero, in violation of Eq. (61).
If the loads for problems I and II were infinite, then \( K' \) or \( K' \) would be infinite as well, which would again violate Eq. (58) or Eq. (61). Therefore, the only possibility left is that both \( \sigma_N' \) and \( \sigma_N'' \) are finite.

9.1.1. Implications for some previous size effect formulae

As widely agreed, the softening cohesive law of concrete or rock begins its descent with a tangent of a finite slope; hence, \( p = 1 \). Therefore, according to Eq. (56), the size effect law must begin near \( D = 0 \) as a linear function of \( D \), which corresponds to an exponential in the logarithmic plot (because \( \ln \sigma_N - \ln \sigma_N^0 = \ln(1 + \sigma_N^0 D/\sigma_N^0) \approx (\sigma_N^0 / \sigma_N^0) e^{D/\sigma_N^0} \)).

The case \( p > 1 \) means that the softening law begins its descent from a horizontal initial tangent, which is reasonable to assume for ductile fracture of yielding materials. The case \( p < 1 \) means that the cohesive law begins its descent with a vertical tangent.

The condition that \( p = 1 \) for quasibrittle materials such as concrete happens to be satisfied by the classical size effect law for bodies with large and similar cracks proposed by Bažant in 1984 [9]. Indeed, \( \sigma_N \propto (1 + D/D_0)^{-1/2} \approx 1 - D/2D_0 \) for small \( D \) (\( D_0 = \) constant). But this condition is satisfied for none of the formulae:

\[
\sigma_N = \frac{Bf_t}{1 + \sqrt{D/D_0}}, \quad \frac{Bf_t'}{[1 + (D/D_0)]^{1/2}}, \quad \sigma_0 \sqrt{1 - e^{-D_0/D}}, \quad \sigma_0 \left( 1 - e^{-D_0/D} \right)
\]

(64)

(with \( D_0, r = \) positive constant, \( r < 1 \)), even though each of these three formulae (the first being a special case of the second) has the correct small-size and large-size asymptotes.

9.2. Two-sided cohesive size effect law and its potential for testing

Consider now the case of structures with a notch or a large stress-free (fatigued) crack formed before the maximum load. Various asymptotic matching formulae can be constructed to match the first two terms of both the small-size and large-size asymptotic expansions, which are as follows:

For \( D \to 0 \) : \( \sigma_N = \sigma_N^0 + \sigma_N^D + \cdots \)

For \( D \to \infty \) : \( \sigma_N^2 = \sigma_0 \frac{D_0}{D} \left( 1 - \frac{D_0}{D} + \cdots \right) \)

(65)

where \( D_0 = c_f g_0^t / g_0 \) is the transitional size for large-size asymptotic approximation, and \( \sigma_N^0 = E' g_0 D_0 \) is the small-size limit of the large-size two-term asymptotic approximation \( [g_0 = g(z_0), \ g_0' = dg(z_0) / dz] ; \sigma_N'' < 0 \). The formula must contain at least four free parameters to match the first two terms of each expansion. One such formula (Fig. 16) is:

![Fig. 16. Two-sided asymptotic matching of cohesive crack model for various values of \( \sigma_N^0 / \sigma_0 \).](image-url)
\[ \sigma_N = \left( p_0 + p_1 D + \frac{m_1 D^2}{m_0 - D} \right)^{-1/2} \]  

(66)

where

\[ m_0 = \frac{p_0 - s_0}{m_1}, \quad m_1 = p_1 - s_1, \quad p_0 = \frac{1}{\sigma_N^0}, \quad p_1 = \frac{-2\sigma_N^0}{\sigma_N^0}, \quad s_0 = \frac{1}{\sigma_0}, \quad s_1 = \frac{g_0 c_f}{E'G_f} \]  

(67)

To verify that Eq. (66) has the correct asymptotics, note that for small enough \( D \):

\[ \sigma_N \approx (p_0 + p_1 D)^{-1/2} = \left[ \sigma_N^0 \left( 1 - 2\sigma_N^0 D / \sigma_N^0 \right) \right]^{-1/2} \approx \sigma_N^0 (1 + \sigma_N^0 D / \sigma_N^0) = \sigma_N^0 + \sigma_N^0 D \]  

(68)

and for large enough \( D \):

\[ \sigma_N^{-2} = p_0 + p_1 D - (p_1 - s_1) D \left( 1 - \frac{p_0 - s_0}{p_1 - s_1} D \right)^{-1} \approx p_0 + p_1 D - (p_1 - s_1) D \left( 1 + \frac{p_0 - s_0}{p_1 - s_1} D \right) \]

\[ = p_0 + p_1 D - p_1 D + s_1 D - p_0 + s_0 = s_0 + s_1 D = \frac{g_0 c_f + g_0 D}{E'G_f} \]  

(69)

Eq. (66) is applicable only if the asymptotes \( \sigma_N^{-2} = p_0 + p_1 D \) and \( \sigma_N^{-2} = s_0 + s_1 D \) intersect at positive \( D \). This gives for the validity of Eq. (66) the condition \( (p_1 - s_1)(p_0 - s_0) < 0 \), which seems to cover most realistic situations.

Since the transition from small-size to large-size behavior occurs in formula (66) within about to decades of the logarithmic time scale, the formula would need to be extended by further terms in order to behave like a broad-range size effect law, describing the behavior over many decades.

An important feature of formula (66) is that it captures not only the LEFM-type shape dependence (through \( g(z_0) \) and \( g'(z_0) \)), which dominates for large sizes, but also the plasticity-like shape dependence (through \( \sigma_N^0 \) and \( \sigma_N^0' \)), which dominates for small sizes. There is a gradual transition from near-fracture to near-plasticity shape dependence as the size diminishes.

Although experimental evaluations have not yet been undertaken, it is clear that the fitting of test results with Eq. (66) should make it possible to use smaller notched specimens. Since both the LEFM-type and plasticity-like shape dependencies are embodied in the formula, the use of rather small dissimilar specimens and, especially, unnotched specimens, becomes possible. These attractive possibilities deserve deeper study.

There exist some parameter combinations for which Eq. (66) gives a nonmonotonic plot or even imaginary values. At this moment it is not clear whether such combinations can occur for real material behavior.

10. Failures at crack initiation and role of strength randomness

The mathematical models discussed so far were focused on the notched specimens. However, as already mentioned, combining notched and unnotched specimens has the potential of the simplest testing method. In the case of unnotched specimens failing at crack initiation, as in modulus of rupture (flexural strength) tests, a different size effect law must be used. For \( D \to \infty \), this law must asymptotically approach Weibull-type size effect (as suggested by Petersson [80] and demonstrated by nonlocal Weibull theory, Ref. [34]); here = Weibull modulus (\( m \neq 12 \), as previously thought, but about 24 for concrete; Ref. [35]), and \( n = \) number of spatial dimensions in geometric similarity (1 or 2; \( n = 3 \) should not be used even if the geometric similarity is three dimensional). The formula that achieves asymptotic matching (of the first two terms of expansion) both to the deterministic size effect of cohesive crack model for \( D \to 0 \), and to the Weibull size effect for \( D \to \infty \), is [21]
\[ \sigma_N = \sigma_\infty \left[ \left( \frac{D_b}{\eta D_b + D} \right)^{rn/m} + \frac{rD_b}{\eta D_b + D} \right]^{1/r} \]  

(70)

(where \( \eta \) is a positive empirical parameter, and \( rn/m < 1 \) is required). The optimal fit of the existing test data is obtained for \( r \approx 1.15 \). The randomness of strength is not important for flexure of concrete beams less than about 30\( d_a \) in depth (\( d_a \) = maximum aggregate size). But it has a major effect for cross-section thickness on the order of 10 m, which is typical of arch dams [34,35].

The limit case of Eq. (70) for \( m \to \infty \) is a deterministic formula. The special case for \( r = 1 \) and \( m \to \infty \) (with \( \eta \geq 0 \)) is a formula empirically justified by Uchida et al. [103] and studied by Planas et al. [88]. For the special case with \( n \to \infty \) and \( \eta = 0 \) (with any \( r > 0 \)), which was derived from fracture mechanics in Ref. [18], the optimum fit of the existing data is obtained for \( r \approx 1.45 \). The special case for \( r = 1, \eta = 0, m \to \infty \) was derived in Ref. [33] from fracture mechanics, and in Refs. [32,33] from stress redistribution in the boundary layer of cracking. The special case for \( r = 2, \eta = 0, m \to \infty \) coincides with Carpinteri’s ‘multifractal scaling law’ or ‘MFSL’ [40, Section 12.7; 44–47]; however, the fit of the existing data with the deterministic formula for either \( r = 2 \) or \( r = 1 \) is distinctly worse than it is for the optimum, \( r \approx 1.45 \). A nonzero value of \( \eta \) is indicated by asymptotic matching of the cohesive crack model; but \( \eta \) can hardly be determined from tests since the values less than about 2 make an appreciable difference only for hypothetical specimen sizes less than about 3\( d_a \).

11. Reflections on design practice

11.1. Why have applications lagged?

In view of the great progress in concrete fracture theory, it may be appropriate to ponder the question why the applications have so far been so scant. It had been a long time since the birth of fracture mechanics in the work of Griffith [54a] until concrete fracture research got started. During 1940–1960, the spectacular fractures of the welded hulls of the Liberty ships, explosions of Comet airliners and failures of some large-span bridges and sports arena roofs caused by fatigue fracture of steel members, as well as concerns about the safety of nuclear reactors, precipitated fast progress in fracture mechanics of metals. After overcoming the difficulties caused by the large plastic yielding zone at fracture front, fracture mechanics of metals was rapidly adopted in the engineering practice.

Not so for concrete and geomaterials. Aside from the impossibility of simply transplanting the progress in steel to concrete, due to salient differences in the nature of the material, there have been two external reasons for the deplorably slow progress in practical applications.

One reason is that, in contrast to metals, the case for fracture mechanics applications in concrete and geotechnical engineering lacked some spectacular, politically compelling, impetus. The tragic disasters that did occur could not be easily interpreted. One reason is that, in contrast to aircraft engineering, the safety factors are so high that structural engineering disasters typically have a multitude of causes, and even if the size effect is one of them the others seem sufficient to explain the failure. Another reason is that even the largest concrete structures, in contrast to steel structures, are generally not large enough compared to the FPZ in order for the failures to be governed purely by fracture energy, with no role for material strength.

Recently, though, it transpired that there have been many catastrophes in which the fracture mechanics size effect must have been a large contributing factor, reducing the nominal strength to between 1/3 and 2/3 of the plastic limit analysis value used in design (the other previously identified causes of these disasters, though, are not disputed). Some examples of such catastrophes are given in Appendix B.

Now the second reason. Given that the fracture size effect in concrete structures is very strong, it might seem perplexing why, fortunately, structural catastrophes have not been much more numerous. This can be
explained by looking at the safety factors used in design. The dead load factor in the current codes is excessive, and its excessive value produces a hidden size effect for the design of large structures. According to the ACI Standard 318, the current value of the dead load factor is 1.4, but random errors in the own weight of large structures cannot justify a safety factor larger than about 1.05. Since large-span structures are totally dominated by their own weight, and for small-span structures the own weight is irrelevant, this discrepancy implies for large structures an overdesign in the ratio of up to about 1.3 (see Appendix C and, in detail, Ref. [27]). Doubtless this is what has, luckily, reduced the number of catastrophes but also offset the urgency of fracture mechanics approach.

To use an excessive dead load factor as a hidden substitute for the size effect is nevertheless inadequate, for six reasons:

1. For some brittle failures, 28% as the maximum capacity reduction due to size effect is far too small, for others, way excessive.
2. Relating dead load $D$ to structure size $D$, one finds that the overdesign ratio as a function of $D$ strongly deviates from the correct size effect law (17).
3. The hidden size effect implied by the current codes is the same for brittle failures (such as the diagonal shear failure), which do exhibit a size effect, and ductile failures (such as the bending failure due to tensile steel yielding), which do not.
4. Even for brittle failures alone, the own weight is very poorly correlated to the brittleness number which controls the fracture size effect. For example, a very tall column or pier might not be protected by the excessive dead load factor because the own weight might cause no significant bending moment; yet brittle failure due to compression crushing in flexure, which exhibits a size effect, may be caused mainly by horizontal loads such as wind or earthquake, for which the load factor is not excessive.
5. For prestressed concrete structures, which are generally lighter than unprestressed ones, the size effect implied by the code is generally weaker. Yet these structures are more brittle than unprestressed ones and thus exhibit stronger size effects.
6. In the case of bending fracture of unreinforced cross-sections thicker than about 1 m (e.g., vertical fracture of an arch dam due to flexure in a horizontal plane), there is a significant Weibull statistical size effect but the overdesign ratio gives a very different size effect.

The problem is not limited to simplified code formulae. Even if one uses a finite element code with realistic cracking and fracture model, which automatically captures the size effect, the result is a systematic overdesign and overreliability for large structures if the currently prescribed load factors are applied to the calculated load capacity. A much lower dead load factor, justified by statistics, should be used in conjunction with such realistic computational approaches. Otherwise they make no sense.

To develop a rational procedure for a design based on computer analysis of structures, the researchers in finite element fracture analysis of concrete structures and in structural reliability must work in synergy. The question of a possible reduction in the dead load factor cannot be separated from the question of size effect, and vice versa. The fracture experts and reliability experts must collaborate.

11.2. What are the arguments for adopting fracture mechanics in design?

Since the design practice in concrete and geotechnical engineering is still based on plastic limit analysis, and fracture mechanics is not yet generally accepted, it is appropriate to close by restating five potent arguments for its adoption, as advanced in Ref. [1].

1. Although fracture initiates as soon as the stress reaches the strength limit, it can grow only if the energy required to break up the material is supplied. The required energy can be supplied either at continued
external loading, in which case the fracture grows in a stable manner and no failure occurs, or at constant external load by spontaneous release of the energy stored in the structure, in which case the structure is unstable and fails. Therefore, material failure criteria in terms of stress or strain are insufficient and the energy release must be taken into account.

2. The results of numerical structural analysis should be objective with respect to the mesh choice. When the material exhibits postpeak softening and the material failure criteria are expressed in terms of stress and strain, the numerical results exhibit spurious mesh sensitivity and convergence to a physically incorrect solution corresponding to zero energy dissipation during failure. The energy criterion for crack propagation avoids this problem (as do the nonlocal strain-softening models).

3. When the load–deflection diagram exhibits no yield plateau, then (1) either the structure is losing stability due to buckling, or (2) the failure is not simultaneous along the whole failure surface, as required by plastic limit analysis, but is progressive, propagating. If the former is not the case, the latter must be happening. The propagating failure cannot be treated by means of the classical material failure criteria stated in terms of stresses and strains; it can be treated by the energy criterion of fracture mechanics (although it can also be treated by nonlocal strain-softening models).

4. The energy absorption capability of a structure, which is important for impact, blast and earthquake analysis, and the ductility of a structure, which is determined by the displacement at stability loss under displacement-controlled loading, are characterized by the equilibrium postpeak load–deflection diagram of structure. This diagram cannot be objectively calculated without fracture mechanics (or nonlocal models).

5. From the structural design viewpoint, the most compelling reason for adopting fracture mechanics is the size effect. Stable growth of large fractures prior to maximum load, typical for many problems of quasi-brittle materials (particularly reinforced concrete and rock), and the large size of FPZ, cause a strong (nonstatistical) size effect whose prediction requires fracture mechanics.

12. Concluding remarks

Fracture mechanics of concrete and rock—a theory which was in its infancy three decades ago—is now a mature theory. At present we have probably entered a golden period of research in which great rewards in engineering are within grasp.

The existing applications in structural engineering, however, are still far below the potential of the theory. The case for applying the available theory in practice needs to be made more convincingly.

Although many details relevant to structural design practice still remain to be researched and some fundamental questions settled, the theory now appears ready for applications.

Acknowledgements

Preparation of this paper was funded under grant 0740-350-A432 from Infrastructure Technology Institute of Northwestern University and the background research was funded under grant CMS-9732791 from the U.S. National Science Foundation to Northwestern University.

Appendix A. Spectrum of fracture energies of cohesive crack model

The series of exponentials used in Eq. (24) is known as the Dirichlet series (or Prony–Dirichlet series) and is used in viscoelasticity as a discrete approximation of the retardation (or relaxation) spectrum of material,
corresponding to the Kelvin chain rheological model (see the review in Refs. [91,101]). The constants \( H_k \) are analogous to the retardation times \( \tau_k \) of the Kelvin chain model. The plot of the elastic moduli \( \hat{E}_k \) of the Kelvin units versus \( \tau_k \) (\( k = 1, 2, 3, \ldots \)) is called the retardation spectrum.

It is well known that the retardation times cannot be determined from experimental data because their identification by least-squares leads to an ill-conditioned equation system (a tiny change in data leads to very different results). To avoid this problem, a continuous relaxation spectrum is introduced; it corresponds to a continuous smearing of the discrete spectrum, with infinitely many, infinitely closely spaced, retardation times. Its advantage is that it represents a Laplace transform. The continuous retardation spectrum, representing \( \hat{E} \) as a function of \( \log \tau \), is unique and, based on the theory of Laplace transform, can even be determined, by explicit formulae, from the compliance function of a viscoelastic material. For numerical computations, the spectrum is represented by discrete values for times \( \tau_k \). The choice of these times is obviously arbitrary, the only restriction being that they must be spaced closely enough for the desired accuracy (and to prevent bumps on the compliance curve).

An analogous spectral approach, with the following features, may be taken for the size effect theory:

1. The plot of \( G_{FK} \) versus \( \log H_k \) (\( k = 1, 2, 3, \ldots \)) may be regarded as a discrete spectrum of fracture energies (Fig. 10, right). This spectrum is geometry dependent for small sizes (for the large size end of the spectrum, its geometry dependence is a moot point since only \( \sum_k G_{FK} \) matters, the FPZ being embedded in the geometry-independent LEFM near-tip field). The continuum limit for \( H_k \) values infinitely closely spaced gives a continuous fracture energy spectrum (dashed in Fig. 10, right). For numerical purposes, this spectrum may be represented, equally well, by discrete values \( H_k \) chosen in various ways but not too far apart.
2. The values of \( H_k \) cannot not be identified from size effect tests, because the identification would be an ill-conditioned problem. Rather, they must be judiciously chosen. Then the corresponding \( G_{FK} \) values may be obtained by fitting test data. The most reasonable appears to be the choice in Eq. (26), \( H_k = \lambda^k D_0 \), which corresponds to uniform spacing of \( \log H_k \) in the logarithmic scale.
3. A unique way to characterize the broad-range size effect (without any arbitrariness such as that in \( H_k \) values) would be by a continuous spectrum of fracture energies.

**Appendix B. Reinterpretation of past structural catastrophes in light of size effect**

Because of very large safety factors as well as a size effect hidden in the safety factor for dead load (Appendix C), a single inadequacy of the current design methods and codes can hardly bring down a structure. Thus it is not surprising that in most structural failures, several causes are combined. Practically all the famous catastrophes of structures have in the past been plausibly explained without revoking the size effect. However, it now transpires that many of these explanations have been incomplete. In the light of the latest research, the fracture size effect should, for example, be added as a significant contributing factor [36] to the explanations of many catastrophes.

(1) While the direct cause of the tragic failure of the Malpasset Arch Dam in French Maritime Alps in 1959 was an excessive movement of the rock abutment, it now transpires that the maximum tolerable movement must have been only about 45% of the value deduced at that time from the standard strength tests of concrete.
(2) The same, but with a reduction to 40%, applies to the Saint-Francis Dam near Los Angeles (failed in 1928).
(3) Whereas the direct cause of the failure of the Schoharie Bridge on New York Thruway in 1987 was the scouring of the foundation in a flood, the nominal bending strength of the foundation
plinth that broke was only about 54% of the value deduced from the laboratory tensile strength of concrete.

Based on what is today known about the fracture size effect, nominal strength reductions ranging from 30% to 50% compared to code-based design must have also occurred in the following catastrophes:

(4) the sinking of the Sleipner Oil Platform in Norway in 1991, caused by shear fracture of a tri-cell (an incorrect placement of reinforcement and an error in finite element analysis due to incorrect meshing were the originally cited causes);
(5) the columns of Hanshin Viaduct, Kobe (failed in 1995 earthquake);
(6) the columns of Cypress Viaduct, Oakland (failed in 1989 earthquake); and
(7) the bridge columns in Los Angeles earthquake (1994) (in all of which insufficient confining reinforce-
ment was the originally cited cause).

Appendix C. Size effect hidden in excessive safety factor for dead load

The dead load factors currently used in concrete design codes have recently been criticized by structural engineering statisticians as unjustifiably large. Proposals for reducing these factors drastically have been made. However, without simultaneously incorporating into the code provisions the fracture size effect, such a reduction would be dangerous. The point is important and deserves to be reviewed in this appendix, based on a detailed study in Ref. [27].

The larger the structure, the higher is the percentage of the own weight contribution \( \hat{D} \) to the ultimate load \( U \). So, if the load factor for the own weight is excessive, structures of large size are overdesigned from the viewpoint of strength theory or plastic limit state design—the theory underlying the current building codes. However, such an overdesign helps to counteract the neglect of size effect in the current codes, which is inherent to plastic limit analysis concepts [27]. Doubtless it is the reason why the number of structural collapses in which the size effect was a contributing factor has not been much larger than we have seen so far.

Denote by \( \hat{L} \) and \( \hat{D} \) the internal forces caused by the live load and the dead load, and by \( U \) the internal force caused by ultimate loads, i.e., the loads magnified by the load factors. Using the load factors currently prescribed by the building code (ACI Standard 318, 1999), one has

\[
U = 1.4\hat{D} + 1.7\hat{L}
\]  

(C.1)

Take it now for granted that the dead load factor 1.4 is excessive and that a realistic value, justified by statistics of dead load, should be \( \mu_D \). Then the ratio of the required ultimate design value of the internal force to the realistic ultimate value, which may be called the overdesign ratio [27], is

\[
R = \frac{U_{\text{design}}}{U_{\text{real}}} = \frac{1.4\hat{D} + 1.7\hat{L}}{\mu_D\hat{D} + 1.7\hat{L}}
\]  

(C.2)

Consideration will now be limited to dead loads caused by the own weight of structures, which for example dominate the design of large span bridges. For a bridge of very large span, the dead load may represent 90% of the total load, and the live load 10%. In that case, the overdesign ratio is

\[
R = \frac{1.4 \times 0.9 + 1.7 \times 0.1}{\mu_D \times 0.9 + 1.7 \times 0.1}
\]  

(C.3)
For the small scale tests which were used to calibrate the present code specifications, the own weight may be assumed to represent less that 2% of the total load. In that case, the overdesign ratio is

\[
R_0 = \frac{1.4 \times 0.02 + 1.7 \times 0.98}{\mu_D \times 0.02 + 1.7 \times 0.98}
\]

(A.4)

A precise value, though, may be debatable and should be determined by extensive statistics, it nevertheless seems reasonable to assume that the own weight of a very large structure cannot be underestimated by more than 5%. This means that \( \mu_D = 1.05 \). So,

\[
R \approx 1.28, \quad R_0 \approx 1.00
\]

(A.5)

It follows that, compared to the reduced scale laboratory tests used to calibrate the code, a structure of a very large span is overdesigned, according to the current theory, by about 28% [27]. Such overdesign compensates for a size effect in the ratio of about 1.28. This is approximately the size effect for very large spans that is unintentionally hidden in the current code specifications.

Further it should be noted that a hidden size effect also exists in various proposals for reliability-based codes. This is due to the fact that the reliability implied in the code increases with the contribution of the dead load to the overall gravity load effect (in detail, see Ref. [27]).

References


