Size effect on structural strength: a review

Z. P. Bažant

Summary  The article attempts a broad review of the problem of size effect or scaling of failure, which has recently come to the forefront of attention because of its importance for concrete and geotechnical engineering, geomechanics, arctic ice engineering, as well as for designing large load-bearing parts made of advanced ceramics and composites, e.g. for aircraft or ships. First, the main results of Weibull statistical theory of random strength are briefly summarized, and its applicability and limitations described. In this theory as well as plasticity, elasticity with a strength limit, and linear elastic fracture mechanics (LEFM), the size effect is a simple power law, because no characteristic size or length is present. Attention is then focused on the deterministic size effect in quasibrittle materials which, because of the existence of a non-negligible material length characterizing the size of the fracture process zone, represents the bridging between the simple power-law size effects of plasticity and of LEFM. The energetic theory of quasibrittle size effect in the bridging region is explained, and then a host of recent refinements, extensions and ramifications are discussed. Comments on other types of size effect, including that which might be associated with the fractal geometry of fracture, are also made. The historical development of the size-effect theories is outlined, and the recent trends of research are emphasized.

Key words  Scaling, size effect, fracture mechanics, quasibrittle materials, asymptotic methods

1  Introduction

The size effect is a problem of scaling, which is central to every physical theory. In fluid mechanics research, the problem of scaling continuously played a prominent role for over a hundred years. In solid mechanics research, though, the attention to scaling had many interruptions and became intense only during the last decade.

Not surprisingly, the modern studies of nonclassical size effect, begun in the 1970's, were stimulated by the problems of concrete structures, for which there inevitably is a large gap between the scales of large structures (e.g. dams, reactor containments, bridges) and of laboratory tests. This gap involves in such structures about one order of magnitude; even in the rare cases when a full-scale test is carried out, it is impossible to acquire a sufficient statistical basis on the full scale.

The question of size effect recently became a crucial consideration in the efforts to use advanced fiber composites and sandwiches for large ship hulls, bulkheads, decks, stacks and masts, as well as for large load-bearing fuselage panels. The scaling problems are even greater in geotechnical engineering, arctic engineering, and geomechanics. In analyzing the safety of an excavation wall or a tunnel, the risk of a mountain slide, the risk of slip of a fault in the earth crust or the force exerted on an oil platform in the Arctic by a moving mile-size ice floe, the scale jump from the laboratory spans many orders of magnitude.
In most of mechanical and aerospace engineering, on the other hand, the problem of scaling has been less pressing because the structures or structural components can usually be tested at full size. It must be recognized, however, that even in that case the scaling implied by the theory must be correct. Scaling is the most fundamental characteristic of any physical theory. If the scaling properties of a theory are incorrect, the theory itself is incorrect.

The size effect in solid mechanics is understood as the effect of the characteristic structure size (dimension) \( D \) on the nominal strength \( \sigma_N \) of structure when geometrically similar structures are compared. The nominal stress (or strength, in case of maximum load) is defined as

\[
\sigma_N = \frac{c_N P}{bD} \quad \text{or} \quad \frac{c_N P}{D^c}
\]

for two- or three-dimensional similarity, respectively, where \( P \) is the load (or load parameter), \( b \) structure thickness and \( c_N \) arbitrary coefficient chosen for convenience (normally, \( c_N = 1 \)). So \( \sigma_N \) is not a real stress but a load parameter having the dimension of stress. The definition of \( D \) can be arbitrary (e.g. the beam depth or half-depth, the beam span, the diagonal dimension, etc.) because it does not matter for comparing geometrically similar structures.

The basic scaling laws in physics are power laws in terms of \( D \), for which no characteristic size (or length) exists. The classical Weibull [107] theory of statistical size effect caused by randomness of material strength is of this type. During the 1970’s it was found that a major deterministic size effect, overwhelming the statistical size effect, can be caused by stress redistributions caused by stable propagation of fracture or damage and the inherent energy release. The law of the deterministic size effect provides a way of bridging two different power laws applicable in two adjacent size ranges. The structure size at which this bridging transition occurs represents a characteristic size.

The material for which this new kind of size effect was identified first, and studied in the greatest depth and with the largest experimental effort by far, is concrete. In general, a size effect that bridges the small-scale power law for nonbrittle (plastic, ductile) behavior and the large-scale power law for brittle behavior signals the presence of a certain nonnegligible characteristic length of the material. This length, which represents the quintessential property of quasibrittle materials, characterizes the typical size of material inhomogeneities or the fracture process zone (FPZ). Aside from concrete, other quasibrittle materials include rocks, cement mortars, ice (especially sea ice), consolidated snow, tough fiber composites and particulate composites, toughened ceramics, fiber-reinforced concretes, dental cements, bone and cartilage, biological shells, stiff clays, cemented sands, grouted soils, coal, paper, wood, wood particle board, various refractories and filled elastomers as well as some special tough metal alloys. Keen interest in the size effect and scaling is now emerging for various ‘high-tech’ applications of these materials.

Quasibrittle behavior can be attained by creating or enhancing material inhomogeneities. Such behavior is desirable because it endows the structure made from a material incapable of plastic yielding with a significant energy absorption capability. Long ago, civil engineers subconsciously but cleverly engineered concrete structures to achieve and enhance quasibrittle characteristics. Most modern ‘high-tech’ materials achieve quasibrittle characteristics in much the same way – by means of inclusions, embedded reinforcement, and intentional microcracking (as in transformation toughening of ceramics, analogous to shrinkage microcracking of concrete). In effect, they emulate concrete.

In materials science, an inverse size effect spanning several orders of magnitude must be tackled in passing from normal laboratory tests of material strength to microelectronic components and micromechanisms. A material that follows linear elastic fracture mechanics (LEFM) on the scale of laboratory specimens of sizes from 1 to 10 cm may exhibit quasibrittle or even ductile (plastic) failure on the scale of 0.1 or 100 microns.

The purpose of this article is to present a brief review of the basic results and their history. For an in-depth review with a thousand of literature references, the recent article [16] may be consulted. A full exposition of most of the material reviewed here is found in the recent book [28]. The problem of scale bridging in the micromechanics of materials, e.g. the relation of dislocation theory to continuum plasticity, is beyond the scope of this review.
2 History of size effect up to Weibull

Speculations about the size effect can be traced back to Leonardo da Vinci (1500's) [112]. He observed that “among cords of equal thickness the longest is the least strong,” and proposed that “a cord is so much stronger . . . as it is shorter,” implying inverse proportionality. A century later, Galileo Galilei (1638) [58], the inventor of the concept of stress, argued that Leonardo’s size effect cannot be true. He further discussed the effect of the size of an animal on the shape of its bones, observing that bulkiness of bones is the weakness of the giants.

A major idea was spawned by Mariotte (1686) [76]. Based on his extensive experiments, he observed that “a long rope and a short one always support the same weight unless that in a long rope there may happen to be some faulty place in which it will break sooner than in a shorter”, and proposed the principle of “the inequality of matter whose absolute resistance is less in one place than another.” In other words, the larger the structure, the greater is the probability of encountering in it an element of low strength. This is the basic idea of the statistical theory of size effect.

Despite no lack of attention, not much progress was achieved for two and half centuries, until the remarkable work of Griffith (1921) [60], the founder of fracture mechanics. He showed experimentally that the nominal strength of glass fibers was raised from 42,300 psi to 491,000 psi when the diameter decreased from 0.0042 in. to 0.00013 in., and concluded that “the weakness of isotropic solids. . . is due to the presence of discontinuities or flaws. . . The effective strength of technical materials could be increased 10 or 20 times at least if these flaws could be eliminated.” In Griffith’s view, however, the flaws or cracks at the moment of failure were still only microscopic; their random distribution controlled the macroscopic strength of the material but did not invalidate the concept of strength. Thus, Griffith discovered the physical basis of Mariotte’s statistical idea but not a new kind of size effect.

The statistical theory of size effect began to emerge after Peirce (1926) [86], formulated the weakest-link model for a chain and introduced the extreme value statistics which was originated by Tippett (1925) [101], Fischer and Tippett (1928) [52], and Fréchet (1927) [51], and refined by von Mises (1936) [102] and others (see also [50, 56, 57, 97]). The capstone of the statistical theory was laid by Weibull (1939) [107] (also [108–110]). On a heuristic and experimental basis, he concluded that the tail distribution of low strength values with an extremely small probability could not be adequately represented by any of the previously known distributions. He introduced what came to be known as the Weibull distribution, which gives the probability of a small material element as a power law of the strength difference from a finite or zero threshold. Others (e.g. [56, 97]) later offered a theoretical justification by means of a statistical distribution of microscopic flaws or microcracks. Refinement of applications to metals and ceramics (fatigue embrittlement, cleavage toughness of steels at low and brittle-ductile transition temperatures, evaluation of scatter of fracture toughness data) has continued until today (e.g. [32, 35, 50, 71]).

Most subsequent studies of the statistical theory of size effect dealt basically with refinements and applications of Weibull’s theory to fatigue embrittled metals and to ceramics (e.g. [69, 70]). Applications to concrete, where the size effect was of the greatest concern, have been studied in [36, 37, 78, 79, 82, 115, 116, 117] and elsewhere.

Until about 1985, most mechanicians paid almost no attention to the possibility of a deterministic size effect. Whenever a size effect was detected in tests, it was automatically assumed to be statistical, and thus its study was supposed to belong to statisticians rather than mechanicians. The reason probably was that no size effect is exhibited by the classical continuum mechanics in which the failure criterion is written in terms of stresses and strains: elasticity with strength limit, plasticity and viscoplasticity, as well as fracture mechanics of bodies containing only microscopic cracks or flaws [10]. The size effect was not even mentioned by S.P. Timoshenko in 1953 in his monumental History of the Strength of Materials.

The attitude, however, changed drastically in the 1980’s. In consequence of the well-funded research in concrete structures for nuclear power plants, theories exhibiting a deterministic size effect have been developed. We will discuss it later.

3 Power scaling and the case of no size effect

It is proper to explain first the simple scaling applicable to all physical systems that involve no characteristic length. Let us consider geometrically similar systems, for example the beams shown in Fig. 1a, and seek to deduce the response $Y$ (e.g. the maximum stress or the maximum
deflection) as a function of the characteristic size (dimension) $D$ of the structure; $Y = Y_0 f(D)$ where $u$ is the chosen unit of measurement (e.g. 1 m, 1 mm). We imagine three structure sizes $1, D_1, D_0$ (Fig. 1a). If we take size 1 as the reference sizes. The responses for sizes $D$ and $D_0$ are $Y = f(D)$ and $Y' = f(D')$. However, since there is no characteristic length, we can also take size $D$ as the reference size. Consequently, the equation

\[ \frac{f(D')}{f(D)} = f\left(\frac{D'}{D}\right) \]

must hold [10, 16] (for fluid mechanics, see [4, 96]). This is a functional equation for the unknown scaling law $f(D)$. It has one and only one solution, namely the power law:

\[ f(D) = \left(\frac{D}{c_1}\right)^s \]

where $s = \text{const.}$ and $c_1$ is a constant which is always implied as a unit of length measurement. Note that $c_1$ cancels out of Eq. (1) when the power function (2) is substituted.

On the other hand, when for instance $f(D) = \log(D/c_1)$, Eq. (1) is not satisfied and the unit of measurement, $c_1$, does not cancel out. So, the logarithmic scaling could be possible only if the system possessed a characteristic length related to $c_1$.

The power scaling must apply for every physical theory in which there is no characteristic length. In solid mechanics such failure theories include elasticity with a strength limit, elastoplasticity, viscoplasticity as well as LEFM (for which the FPZ is assumed shrunken into a point).

To determine exponent $s$, the failure criterion of the material must be taken into account. For elasticity with a strength limit (strength theory), or plasticity (or elasto-plasticity) with a yield surface expressed in terms of stresses or strains, or both, one finds that $s = 0$ when response $Y$ represents the stress or strain (for example the maximum stress, or the stress at certain homologous points, or the nominal stress at failure) [10]. Thus, if there is no characteristic dimension, all geometrically similar structures of different sizes must fail at the same nominal stress. By convention, this came to be known as the case of no size effect.

In LEFM, on the other hand, $s = -\frac{1}{2}$, provided that geometrically similar structures with geometrically similar cracks or notches are considered. This may be generally demonstrated with the help of Rice's J-integral [10].

If $\log \sigma_N$ is plotted versus $\log D$, the power law is a straight line, Fig. 1. For plasticity, or elasticity with a strength limit, the exponent of the power law vanishes, i.e. the slope of this line is zero. For LEFM, the slope is $-\frac{1}{2}$. A recently emerged ‘hot’ subject is the quasibrittle materials and structures, for which the size effect bridges these two power laws.
Weibull statistical size effect

The classical theory of size effect has been statistical. Three-dimensional continuous generalization of the weakest link model for the failure of a chain of links of random strength, Fig. 2 left, leads to the distribution

\[ P_f = 1 - \exp \left\{ - \int_V c(\sigma(x)) \, dV(x) \right\}, \tag{3} \]

which represents the failure probability of a structure that fails as soon as a macroscopic fracture initiates from a microcrack (or some flaw) somewhere in the structure; \( \sigma \) is the stress tensor field just before failure, \( x \) the coordinate vector, \( V \) the volume of structure and \( c(\sigma) \) a function giving the spatial concentration of failure probability of the material (=\( V^{-1} \times \) failure probability of material representative volume \( V_r \)) \[56],

\[ c(\sigma) \approx \frac{\sum \sigma_i \, P_1(\sigma_i)}{V_0} \]

where \( \sigma_i \) are the principal stresses \((i = 1, 2, 3)\) and \( P_1(\sigma) \) the failure probability (cumulative) of the smallest possible test specimen of volume \( V_0 \) (or representative volume of the material) subjected to uniaxial tensile stress \( \sigma \),

\[ P_1(\sigma) = \left( \frac{\sigma \sigma_u}{s_0} \right)^m, \tag{4} \]

[107], where \( m, s_0, \sigma_1 \) are material constants \((m \) denotes the Weibull modulus, usually between 5 and 50, \( s_0 \) is a scale parameter, \( \sigma_u \) the strength threshold, which may usually be taken as zero) and \( V_0 \) is the reference volume understood as the volume of specimens on which \( c(\sigma) \) was measured. For specimens under uniform uniaxial stress (and \( \sigma_u = 0 \)), (3) and (4) lead to the following simple expressions for the mean and the coefficient of variation of nominal strength:

\[ \sigma_N = s_0 \Gamma(1 + m^{-1}) \left( \frac{V_0}{V} \right)^{\frac{1}{m}}, \quad \omega = \left[ \frac{\Gamma(1 + 2m^{-1})}{\Gamma^2(1 + m^{-1})} - 1 \right]^{\frac{1}{2}}, \tag{5} \]

where \( \Gamma \) is the gamma function. Since \( \omega \) depends only on \( m \), it is often used for determining \( m \) from the observed statistical scatter of strength of identical test specimens. The expression for \( \sigma_N \) includes the effect of volume \( V \) which depends on size \( D \). In general, for structures with nonuniform multidimensional stress, the size effect of Weibull theory (for \( \sigma_u \approx 0 \)) is of the type:

![Fig. 2. a Chain with many links of random strength b failure probability of a small element c structure with many microcracks of different probabilities to become critical](image)
\[ \sigma_N \propto D^{-\frac{n_d}{2}}, \quad (6) \]

where \( n_d = 1, 2 \) or 3 for uni-, two- or three-dimensional similarity.

In view of (5), the value \( \sigma_W = \sigma_N(V/V_0)^{1/m} \) for a uniformly stressed specimen can be adopted as a size-independent stress measure. Taking this viewpoint, Beremin [32], proposed taking into account the nonuniform stress in a large crack-tip plastic zone by the so-called Weibull stress:

\[ \sigma_W = \left( \sum_i \sigma_{Ii}^{m} \frac{V_i}{V_0} \right)^{\frac{1}{m}}, \quad (7) \]

where \( V_i \) (\( i = 1, 2, \ldots, N_W \)) are elements of the plastic zone having maximum principal stress \( \sigma_{Ii} \). The sum in (5) was replaced by an integral in [95], see also [71]. Equation (7), however, considers only the crack-tip plastic zone whose size which is almost independent of \( D \). Consequently, Eq. (7) is applicable only if the crack at the moment of failure is not yet macroscopic, still being negligible compared to structural dimensions.

As far as quasibrittle structures are concerned, applications of the classical Weibull theory face a number of serious objections:

1. The fact that in (6) the size effect is a power law implies the absence of any characteristic length. But this cannot be true if the material contains sizable inhomogeneities.
2. The energy release due to stress redistributions caused by macroscopic FPZ, or stable crack growth before \( P_{\text{max}} \), gives rise to a deterministic size effect which is ignored. Thus the Weibull theory is valid only if the structure fails as soon as a microscopic crack becomes macroscopic.
3. Every structure is mathematically equivalent to a uniaxially stressed bar (or chain, Fig. 2), which means that no information on the structural geometry and failure mechanism is taken into account.
4. The size-effect differences between two- and three-dimensional similarities (\( n_d = 2 \) or 3) are predicted much too large.
5. Many tests of quasibrittle materials (e.g. diagonal shear failure of reinforced concrete beams) show a much stronger size effect than predicted by Weibull theory, [28] and the review in [11]).
6. The classical theory neglects the spatial correlations of material failure probabilities of neighboring elements caused by nonlocal properties of damage evolution (while generalizations based on some phenomenological load-sharing hypotheses have been divorced from mechanics).
7. When Eq. (5) is fit to the test data on statistical scatter for specimens of one size (\( V = \text{const.} \)), and when Eq. (6) is fit to the mean test data on the effect of size (of unnotched plain concrete specimens), the optimum values of Weibull exponent \( m \) are very different, namely \( m = 12 \) and \( m = 24 \), respectively. If the theory were applicable, these value would have to coincide.

In view of these limitations, among concrete structures Weibull theory appears applicable to some extremely thick plain (unreinforced) structures, e.g. the flexure of an arch dam acting as a horizontal beam (but not for vertical bending or arch dams nor gravity dams, because large vertical compressive stresses cause long cracks to grow stably before the maximum load). Most other plain concrete structures are not thick enough to prevent the deterministic size effect from dominating. Steel or fiber reinforcement prevents it as well.

5 Quasibrittle size effect bridging plasticity and LEFM, and its history

Quasibrittle materials are those that obey on a small scale the theory of plasticity (or strength theory), characterized by material strength or yield limit \( \sigma_0 \), and on a large scale the LEFM, characterized by fracture energy \( G_f \). While plasticity alone, as well as LEFM alone, possesses no characteristic length, the combination of both, which must be considered for the bridging of plasticity and LEFM, does. Combination of \( \sigma_0 \) and \( G_f \) yields Irwin’s (1958) characteristic length (material length)

\[ \ell_0 = \frac{EG_f}{\sigma_0^2}, \quad (8) \]
which approximately characterizes the length of the FPZ; $E$ is Young’s elastic modulus. So the key to the deterministic quasibrittle size effect is a combination of the concept of strength or yield with fracture mechanics. In dynamics, this further implies the existence of a characteristic time (material time):

$$\tau_0 = l_0/v,$$

representing the time a wave of velocity $v$ takes to propagate the distance $l_0$.

After LEFM was first applied to concrete [66], it was found to disagree with test results [68, 72, 105, 106]. In [72], Leicester tested geometrically similar notched beams of different sizes, fit the results by a power law, $\sigma_N \propto D^{-n}$, and observed that the optimum $n$ was less than 1/2, the value required by LEFM. The power law with a reduced exponent, of course, fits the test data in the central part of the transitional size range well but does not provide the bridging of the ductile and LEFM size effects. It was tried to explain the reduced exponent value by notches of a finite angle, which however is objectionable for two reasons: (i) notches of a finite angle cannot propagate (rather, a crack must emanate from the notch tip), (ii) the singular stress field of finite-angle notches gives a zero flux of energy into the notch tip. Same as the Weibull theory, Leicester’s power law also implied nonexistence of a characteristic length (see [16], Eqs. (1)–(3)); which cannot be the case for concrete due to the large size of its inhomogeneities. More extensive tests of notched geometrically similar concrete beams of different sizes were carried out by Walsh [105, 106]. Although he did not attempt a mathematical formulation, he was the first to make the doubly logarithmic plot of nominal strength versus size and observe that it was transitional between plasticity and LEFM.

An important advance was made in [62] and also in [87]. Inspired by the softening and plastic FPZ models of Barenblatt [2, 3] and Dugdale [49], Hillerborg et al. formulated the cohesive (or fictitious) crack model characterized by a softening stress-displacement law for the crack opening and showed by finite element calculations that the failures of unnotched plain concrete beams in bending exhibit a deterministic size effect, in agreement with tests of the modulus of rupture.

Analyzing distributed (smeared) cracking damage, Bazant [5] demonstrated that its localization into a crack band engenders a deterministic size effect on the postpeak deflections and energy dissipation of structures. The effect of the crack band is approximately equivalent to that of a long fracture with a sizable FPZ at the tip. Subsequently, using an approximate energy release analysis, the following approximate size effect law was derived in [6] for the quasibrittle size effect in structures failing after large stable crack growth:

$$\sigma_N = B\sigma_0 \left(1 + \frac{D}{D_0}\right)^{-\frac{1}{2}} \sigma_R,$$

or more generally:

$$\sigma_N = B\sigma_0 \left[1 + \left(\frac{D}{D_0}\right)^r\right]^{-\frac{1}{r}} \sigma_R,$$

in which $r, B$ are positive dimensionless constants, $D_0 = \text{const.}$ represents the transitional size at which the power laws of plasticity and LEFM intersect; $D_0$ and $B$ characterize the structure geometry. Usually, constant $\sigma_R = 0$, except when there is a residual crack-bridging stress $\sigma_R$ outside the FPZ, as in fiber composites, or when at large sizes some plastic mechanism acting in parallel emerges and becomes dominant, as in the Brazilian split-cylinder test.

Equation (10) was shown to be closely followed by the numerical results for the crack band model [5, 26], as well as for the nonlocal continuum damage models, which are capable of realistically simulating the localization of strain-softening damage and avoiding spurious mesh sensitivity.

Beginning in the mid 1980’s, the interest in the quasibrittle size effect of concrete structures surged enormously and many researchers made noteworthy contributions; to name but a few: [36, 87, 88–90]. The size effect has recently become a major theme at conferences on concrete fracture [8, 30, 80, 81, 114].
Measurements of the size effect of $P_{\text{max}}$ were shown to offer a simple and reliable way to determine the fracture characteristics of quasibrittle materials, including the fracture energy, effective FPZ length, the critical crack tip opening displacement, and geometry-dependent $R$-curve.

6 Size-effect mechanism: stress redistribution and energy release

Let us now describe the gist of the deterministic quasibrittle size effect. LEFM applies when the FPZ is negligibly small compared to structural dimension $D$ and can be considered as a point. Thus the LEFM solutions can be obtained by methods of elasticity. The salient characteristic of quasibrittle materials is that there exists a sizable FPZ with distributed cracking or other softening damage that is not negligibly small compared to structural dimension $D$. This makes the problem nonlinear, although approximately equivalent LEFM solutions can be applied, unless FPZ reaches near the structure boundaries.

The existence of a large FPZ means that the distance between the tip of the actual traction-free crack and the tip of the equivalent LEFM crack at $P_{\text{max}}$ is equal to a certain characteristic length $c_f$ (roughly one half of the FPZ size) that is not negligible compared to $D$. This causes a nonnegligible macroscopic stress redistribution with energy release from the structure.

With respect to the fracture length $a_0$ (distance from the mouth of notch or crack to the beginning of the FPZ), two basic cases may now be distinguished: (i) $a_0 = 0$, which means that $P_{\text{max}}$ occurs at the initiation of macroscopic fracture propagation, and (ii) $a_0$ is finite and not negligible compared to $D$, which means that $P_{\text{max}}$ occurs after large stable fracture growth.

6.1 Scaling for failure at crack initiation

An example of the first case is the modulus of rupture test, which consists in the bending of a simply supported beam of span $L$ with a rectangular cross section of depth $D$ and width $b$, subjected to concentrated load $P$, the maximum load is not decided by the stress $\sigma_1 = 3PL/2bD^2$ at the tensile face, but by the stress value $\bar{\sigma}$ roughly at distance $c_f/2$ from the tensile face, which is at the middle of FPZ. Because $\bar{\sigma} = \sigma_1 - \sigma'_1 c_f/2$, where $\sigma'_1$ is the stress gradient equal to $2\sigma_1/D$, and also because $\bar{\sigma} = \sigma_0$ is the intrinsic tensile strength of the material, the failure condition $\bar{\sigma} = \sigma_0$ yields $P/bD = \sigma_N = \sigma_0/(1 - D_b/D)$, where $D_b = (3L/2D)c_f$, which is a constant because for geometrically similar beams $L/D = \text{const}$. This expression, however, is unacceptable for $D \leq D_b$. But since the derivation is valid only for small enough $c_f/D$, one may replace it by the following asymptotically equivalent size effect formula:

$$\sigma_N = \sigma_0 \left(1 + \frac{rD_b}{D}\right)^{1/2},$$

which happens to be acceptable for the whole range of $D$ ($D \leq \text{material inhomogeneity size}$ is excluded, of course); $r$ is any positive constant. The values $r = 1$ or 2 have been used for concrete [7], while $r \approx 1.45$ is optimum according to the latest analysis of test data at Northwestern University (yet unpublished).

6.2 Scaling for failures with a long crack or notch

Let us now give a simple explanation of the second case of structures failing only after stable formation of large cracks, or notched fracture specimens. Failures of this type, exhibiting a strong size effect [18, 28, 59, 63, 77, 98, 104] are typical of reinforced concrete structures or fiber composites [17, 113], and are also exhibited by some unreinforced structures (e.g. dams, due to the effect of vertical compression, or floating ice plates in the Arctic). Consider the rectangular panel in Fig. 3, which is initially under a uniform stress equal to $\sigma_N$. Introduction of a crack of length $a$ with a FPZ of a certain length and width $h$ may be approximately imagined to relieve the stress, and thus release the strain energy, from the shaded triangles on the flanks of the crack band shown in Fig. 3. The slope $k$ of the effective boundary of the stress-relief zone need not be determined; what is important is that $k$ is independent of the size $D$.

For the usual ranges of interest, the length of the crack at maximum load may normally be assumed approximately proportional to the structure size $D$, while the size $h$ of the FPZ is essentially a constant, related to the inhomogeneity size in the material. This has been verified.
for many cases by experiments (showing similar failure modes for small and large specimens) and finite element solutions based on the crack band, cohesive or nonlocal models.

The stress reduction in the triangular zones of areas $ka^2/2$, Fig. 3, causes (for the case $b = 1$) the energy release $U_a = 2 \times (ka^2/2)\sigma_N^2/2E$. The stress drop within the crack band of width $h$ causes further energy release $U_b = ha\sigma_N^2/E$. The total energy dissipated by the fracture is $W = aG_f$, where $G_f$ is the fracture energy, a material property representing the energy dissipated per unit area of the fracture surface. Energy balance during static failure requires that $\partial(U_a + U_b)/\partial a = dW/da$. Setting $a = D(a/D)$, where $a/D$ is approximately a constant if the failures for different structure sizes are geometrically similar, the solution of the last equation for $\sigma_N$ yields $[6]$ the approximate size effect law in (10) with $\sigma_R = 0$, Fig. 1 bottom.

More rigorous derivations of this law, applicable to arbitrary structure geometry, have been given in terms of asymptotic analysis based on equivalent LEFM [12] or on Rice's path-independent J-integral [28]. This law has also been verified by nonlocal finite element analysis, by the cohesive crack model, and by random particle (or discrete element) models. The experimental verifications, among which the earliest is found in the famous Walsh’s tests [105, 106] of notched concrete beams, have by now become abundant, e.g. Fig. 4.

For very large sizes ($D \gg D_0$), the size effect law in (10) reduces to the power law $\sigma_N \propto D^{-1/2}$, which represents the size effect of LEFM for geometrically similar large cracks and corresponds to the inclined asymptote of slope $-1/2$ in Fig. 1 (bottom). For very small sizes ($D \ll D_0$), this law reduces to $\sigma_N = \text{const}$, which corresponds to the horizontal asymptote and means that there is no size effect, as in plastic limit analysis.

The ratio $\beta = D/D_0$ is called the brittleness number of a structure. For $\beta \to \infty$, the structure is perfectly brittle, i.e. follows LEFM, in which case the size effect is the strongest possible, while for $\beta \to 0$ the structure is nonbrittle (or ductile, plastic), in which case there is no size effect. Quasibrittle structures are those for which $0.1 \leq \beta \leq 10$, in which case the size effect represents a smooth transition (or interpolation) that bridges the power law size effects for the two asymptotic cases. The law (10) has the character of asymptotic matching and serves to provide the bridging of scales. In the quasibrittle range, the stress analysis is of course nonlinear, calling for the cohesive crack model or the crack band model (which are mutually almost equivalent), or some of the nonlocal damage models.

The meaning of the term quasibrittle is relative. If the size of a quasibrittle structure becomes sufficiently large compared to material inhomogeneities, the structure becomes perfectly brittle (for concrete structures, only the global fracture of a large dam is describable by LEFM), and if the size becomes sufficiently small, the structure becomes nonbrittle (plastic, ductile), because the FPZ extends over the whole cross section of the structure. Thus, a micromachine or a miniature electronic device made of silicone or fine-grained ceramic may be quasibrittle or nonbrittle.

6.3 Size effect on postpeak softening and ductility

The plots of nominal stress versus the relative structure deflection (normalized so as to make the initial slope size-independent) have, for small and large structures, the shapes indicated in Fig. 5. Apart from the size effect on $P_{\text{max}}$, there is also a size effect on the shape of the postpeak descending load-deflection curve. For small structures the postpeak curves descend slowly, for larger structures steeper, and for sufficiently large structures they may exhibit a snapback, that is, a change of slope from negative to positive.

If a structure is loaded under displacement control through an elastic device with spring constant $C_s$, it loses stability and fails at the point where the load-deflection diagram first attains
the slope \(-C_p\) (if ever), Fig. 5. The ratio of the deflection at these points to the elastic deflection characterizes the ductility of the structure. As apparent from the figure, small quasibrittle structures have a large ductility while large quasibrittle structures have small ductility.

The areas under the load-deflection curves in Fig. 5 characterize the energy absorption. The capability of a quasibrittle structure to absorb energy decreases, in relative terms, as the structure size increases. The size effect on energy absorption capability is important for blast loads and impact.
The progressive steepening of the postpeak curves in Fig. 5 with increasing size and the development of a snapback can be most simply described by the series coupling model, which assumes that the response of a structure may be at least approximately modeled by the series coupling of the cohesive crack or damage zone with a spring characterizing the elastic unloading of the rest of the structure ([15], Sec. 13.2).

### 6.4 Asymptotic analysis of size effect by equivalent LEFM

To obtain simple approximate size-effect formulae that give a complete prediction of the failure load, including the effect of geometrical shape of the structure, equivalent LEFM may be used. In this approach the tip of the equivalent LEFM sharp crack is assumed to lie approximately a distance $c_f$ ahead of the tip of the traction-free crack or notch, $c_f$ being a constant representing roughly one half of the length of the FPZ ahead of the tip. Two cases are relatively simple:

(i) If a large crack grows stably prior to $P_{\text{max}}$ or if there is a long notch

$$\sigma_N = \frac{\sqrt{EG_f} + \sigma_0 \sqrt{\gamma'(x_0) c_f + \gamma(x_0) D}}{\sqrt{g'(x_0)c_f + g(x_0)D}}$$

(ii) If $P_{\text{max}}$ occurs at fracture initiation from a smooth surface

$$\sigma_N = \frac{\sqrt{EG_f} + \sigma_0 \sqrt{\gamma'(0) c_f + \gamma''(0) \frac{c_f^2}{D}}}{\sqrt{g'(0)c_f + g''(0)\frac{c_f^2}{D}}}$$

[7, 12], where the primes denote derivatives; $g(x_0) = K_{IP}^2/\sigma_N^2 D$ and $\gamma(x_0) = K_{IN}^2/\sigma_N^2 D$ are dimensionless energy-release functions of LEFM of $x = a_0/D$, where $a_0$ is the length of notch or crack up to the beginning of the FPZ; $K_{IP}, K_{IN}$ are stress intensity factors for load $P$ and for loading by uniform residual crack-bridging stress $\sigma_Y$, respectively; $\sigma_Y > 0$ for tensile fracture, but $\sigma_Y \neq 0$ in the cases of compression fracture in concrete, kink band propagation in fiber composites, and tensile fracture of composites reinforced by fibers short enough to undergo frictional pullout rather than breakage. The asymptotic behavior of (13) for $D \to \infty$ is of the LEFM type, $\sigma_N - \sigma_Y \propto D^{-1/2}$. Formula (14) approaches for $D \to \infty$ a finite asymptotic value; so does formula (13), if $\sigma_Y > 0$. 

Fig. 5. Load-deflection curves of quasibrittle structures of different sizes, scaled to the same initial slope
6.5 Size-effect method for measuring material constants and the R-curve

Comparison of (13) with (10) yields the relations

\[ D_0 = c_f \frac{g'(x_0)}{g(x_0)}, \quad B \sigma_0 = \sigma_0 \sqrt{\frac{E G_f}{c_f g'(x_0)}}. \]  

Therefore, by fitting formula (10) with \( \sigma_R = 0 \) to the values of \( \sigma_N \) measured on test specimens of different sizes with a sufficiently broad range of brittleness numbers \( \beta = D/D_0 \), the values of \( G_f \) and \( c_f \) can be identified \([19, 28]\). The fitting can best be done by using the Levenberg–Marquardt nonlinear optimization algorithm, but it can also be accomplished by a properly weighted linear regression of \( \sigma_N^2 \) versus \( D \). The specimens do not have to be geometrically similar, although, when they are, the evaluation is simpler and the error smaller. The lower the scatter of test results, the narrower is the minimum necessary range of \( \beta \) (for concrete and fiber composites, the size range 1:4 is the minimum).

The size-effect method of measuring fracture characteristics has been adopted for an international standard recommendation for concrete ([28, 93] Sec. 6.3), and has also been verified and used for various rocks, ceramics, orthotropic fiber-polymer composites, sea ice, wood, tough metals and other quasibrittle materials. The advantage of the size-effect method is that the tests, requiring only the maximum loads, are foolproof and easy to carry out. With regard to the cohesive crack model, note that the size-effect method gives the energy value corresponding to the area under the initial tangent of the softening stress-displacement curve, rather than the total area under the curve.

The size-effect method also permits determining the R-curve (resistance curve) of the quasibrittle material, a curve that represents the apparent variation of fracture energy with crack extension for which LEFM becomes approximately equivalent to the actual material with a large FPZ. The R-curve, which depends on the specimen geometry, in contrast to the classical R-curve definition, can be obtained as the envelope of the curves of the energy release rate at \( P = P_{\text{max}} \) (for each size) versus the crack extension for specimens of various sizes (or various brittleness numbers). In general, this can easily be done numerically, and if the size-effect law has the form in (10) with \( \sigma_R = 0 \), a parametric analytical expression for the R-curve exists ([19, 28], Sec. 6.4).

The fracture model implied by the size-effect law in (10) or (13) has one independent characteristic length, \( c_f \), representing about one half of the FPZ length. Because of (15), the value of \( \ell_0 \) is implied by \( c_f \) if \( \sigma_0 \) is known. The value of \( c_f \) controls the size \( D_0 \) at the center of the bridging region, intersection of the power-law asymptotes in Fig. 1 (bottom), and \( \sigma_0 \) or \( G_f \) control a vertical shift of the size-effect curve at constant \( D_0 \). The location of the large-size asymptote depends only on \( K_c \) and geometry, and the location of the small-size asymptote depends only on \( \sigma_0 \) and geometry.

6.6 Critical crack-tip opening displacement, \( \delta_{\text{CTOD}} \)

The quasibrittle size effect, bridging plasticity and LEFM, can also be simulated by the fracture models characterized by the critical stress intensity factor \( K_c \) (fracture toughness) and \( \delta_{\text{CTOD}} \); for metals see [45, 111], and for concrete [64]. Jenq and Shah’s model [64], called the two-parameter fracture model, has been shown to give essentially the same results as the R-curve derived from the size-effect law in (10) with \( \sigma_R = 0 \). The models are in practice equivalent because

\[ K_c = \sqrt{E G_f}, \quad \delta_{\text{CTOD}} = \frac{1}{\pi} \sqrt{\frac{B G_f c_f}{E}}. \]  

[18]. Using these formulae, the values of \( K_c \) and \( \delta_{\text{CTOD}} \) can be easily identified by fitting the size effect law (10) to measured \( P_{\text{max}} \) value.

Like the size-effect law in (10) with \( \sigma_R = 0 \), the two-parameter model has only one independent characteristic length, \( \ell_0 = K_c^2/\sigma_0^2 \). If \( \sigma_0 \) is known, then \( \delta_{\text{CTOD}} \) is not an independent length because \( c_f \) is implied by \( \ell_0 \), and \( \delta_{\text{CTOD}} \) then follows from (16).
Extensions, ramifications and applications

7.1 Size effects in compression fracture
Loading by high compressive stress without sufficient lateral confining stresses leads to damage in the form axial splitting microcracks engendered by pores, inclusions or inclined slip planes. This damage localizes into a band that propagates either axially or laterally.

For axial propagation, the energy release from the band drives the formation of the axial splitting fracture, and since this energy release is proportional to the length of the band, there is no size effect. For lateral propagation, the stress in the zones on the sides of the damage band gets reduced, which causes an energy release that grows in proportion to $D^2$, while the energy consumed and dissipated in the band grows in proportion to $D$. The mismatch of energy-release rates inevitably engenders a deterministic size effect of the quasibrittle type, analogous to the size effect associated with tensile fracture. In consequence of the size effect, failure by lateral propagation must prevail over the failure by axial propagation if a certain critical size is exceeded.

The size effect can again be approximately described by the equivalent LEFM. This leads to Eq. (13), in which $\sigma_Y$ is determined by analysis of the microbuckling in the laterally propagating band of axial splitting cracks. The spacing $s$ of these cracks is in (13) assumed to be dictated by material inhomogeneities. However, if the spacing is not dictated, and is such that it minimizes $\sigma_N$, then the size effect gets modified as

$$\sigma_N = CD^{1/2} + \sigma_\infty,$$  \hfill (17)

([28], Sec. 10.5.11) where $C, \sigma_\infty$ are constants, the approximate values of which have been calculated for the breakout of boreholes in rock.

7.2 Fracturing truss model for concrete and boreholes in rock
Propagation of compression fracture is what appears to control the maximum load in diagonal shear failure of reinforced concrete beams with or without stirrups, for which a very strong size effect has been demonstrated experimentally, e.g. [11, 19, 63, 65, 85, 92, 98, 103–104]. A long diagonal tension crack grows stably under shear loading until the concrete near its tip gets crushed. A simplified formula for the size effect can be obtained by energetic modification of the classical truss model (strut-and-tie model) [11].

The explosive breakout of boreholes (or mining stopes) in rock under very high pressures is known to also exhibit size effect, as revealed by the tests in [42, 43, 61, 84]. An approximate analytical solution can be obtained by exploiting Eshelby’s theorem for eigenstresses in elliptical inclusions [25].

7.3 Kink bands in fiber composites
A kink band, in which axial shear-splitting cracks develop between fibers which undergo microbuckling, is one typical mode of compression failure of composites or laminates with uniaxial fiber reinforcement. This failure mode, whose theory was begun in [1, 94], was until recently treated by the theory of plasticity, which implies no size effect. Recent experimental and theoretical studies [35], however, revealed that the kink band propagates side-way like a crack, and the stress on the flanks of the band gets reduced to a certain residual value, which is here denoted as $\sigma_Y$ and can be estimated by the classical plasticity approach of Budianski [34].

The crack-like behavior implies a size effect, which is demonstrated by the latest laboratory tests of notched carbon-PEEK specimens [20], Fig. 4; these tests also demonstrated the possibility of a stable growth of a long kink band, which was achieved by rotational restraint at the ends.

There are again two types of size effect, depending on whether $P_{max}$ is reached (i) when the FPZ of the kink band is attached to a smooth surface or (ii) or when there exists either a notch or a long segment of the kink band in which the stress has been reduced to $\sigma_Y$. Formulae (13) and (14), respectively, approximately describe the size effects for these two basic cases; in this case, $G_f$ plays the role of fracture energy of the kink band (area below the stress-contraction
curve of the kink band and above the $\sigma_f$ value), and $\epsilon_f$ - the role of the FPZ of the kink band, which is assumed to be approximately constant, governed by material properties.

The aforementioned carbon-PEEK tests also confirm that case (ii), in which a long kink band grows stably prior to $P_{\text{max}}$, is possible; in those tests, stability is achieved by virtue of a lateral shift of compression resultant in wide notched prismatic specimens with ends restrained against rotation.

7.4 Size effects in sea ice

Normal laboratory specimens of sea ice exhibit no notch sensitivity. Therefore, failure of sea ice has been thought to be well-described by plastic limit analysis, which exhibits no size effect, e.g. [66, 100]. This perception, however, changed drastically after Dempsey carried out in 1993 on the Arctic Ocean size-effect tests of floating notched square specimens with an unprecedented, record-breaking size range (with square sides ranging from 0.5 m to 80 ml), see [46, 47, 83].

It is now clear that floating sea ice plates are quasibrittle and their size effect on the scale of 100 m approaches that of LEFM. Among other things, Dempsey’s major experimental result (Fig. 4 middle) explains why the measured forces exerted by moving ice on a fixed oil platform are one to two orders of magnitude smaller than the predictions of plastic limit analysis based on the laboratory strength of ice. The size effect law in (10) with $\sigma_R = 0$, or in (13) with $\sigma_Y = 0$, agrees with these results well, permitting the values of $G_f$ and $c_f$ of sea ice to be extracted by linear regression of the $P_{\text{max}}$ data. The value of $c_f$ is in the order of meters, which can be explained by inhomogeneities such as brine pockets and channels, as well as preexisting thermal cracks, bottom roughness of the plate, warm and cold spots due to alternating snow drifts, etc..

Information on the size effect in sea ice can also be extracted from acoustic measurements [74]. Rapid cooling in the Arctic can in the floating plate produce bending moments large enough to cause fracture. According to plasticity or elasticity with a strength limit, the critical temperature difference $\Delta T_{cr}$ across the plate would have to be independent of plate thickness $D$. Fracture analysis, however, indicated a quasibrittle size effect. Curiously, its asymptotic form is not $\Delta T_{cr} \propto D^{-3/2}$ but

$$\Delta T_{cr} \propto D^{-1/2},$$

[9]. The reason is that $D$ is not a characteristic dimension in the plane of the boundary value problem of plate bending; rather it is the flexural wavelength of a plate on elastic foundation, which is proportional to $D^{4/3}$ rather than $D$. It seems that (18) may explain why long cracks of length 10 to 100 km, which suddenly form in the fall in the Arctic ice cover, often run through thick ice floes and do not follow the thinly refrozen water leads around the floes.

In analyzing the vertical penetration of floating ice plate (load capacity for heavy objects on ice, or the maximum force $P$ required for penetration from below), one must take into account that bending cracks reach only through part of the thickness, their ligaments transmitting compressive forces which produces a dome effect. Because at maximum load the part-through bending crack (of variable depth profile) is growing vertically, the asymptotic size effect is not $P/D^2 = \sigma_N \propto D^{-3/8}$ [99], but $\sigma_n \propto D^{-1/2}$. This was determined by a simplified analytical solution (with a uniform crack depth) in [48], and confirmed by a detailed numerical solution with a variable crack depth profile [21]. The latter also led to an approximate prediction formula for the entire practical range of $D$, which is of the type of (10) with $\sigma_N = 0$. This formula was shown to agree with the existing field tests [53, 54, 75].

7.5 Influence of crack separation rate, creep and viscosity

There are two mechanisms in which the loading rate affects fracture growth: (i) creep of the material outside the FPZ, and (ii) rate dependence of the severance of material bonds in the FPZ. The latter may be modeled as a rate process controlled by activation energy, with Arrhenius-type temperature dependence. This leads to a dependence of the softening stress-separation relation of the cohesive crack model on the rate of opening displacement. In an equivalent LEFM approach, the latter is modeled by considering the crack-extension rate to be a power function of the ratio of the stress intensity factor to its critical $R$-curve value.

For quasibrittle materials exhibiting creep (e.g. concretes and polymer composites, but not rocks or ceramics), the consequence of mechanism 1 (creep) is that a decrease of loading rate,
or an increase of duration of a sustained load, causes a decrease of the effective length of the FPZ. This in turn means an increase of the brittleness number manifested by a leftward rigid-body shift of the size-effect curve in the plot of log $\sigma_N$ versus log $D$, i.e., a decrease of effective $D_0$. For slow or long-time loading, quasibrittle structures become more brittle and exhibit a stronger size effect [24].

Mechanism 2 (rate-dependence of separation) causes that an increase of loading rate, or a decrease of sustained load duration, leads to an upward vertical shift of the size effect curve for log $\sigma_N$ but has no effect on $D_0$ and thus on brittleness. This mechanism also explains an interesting recently discovered phenomenon — a reversal of softening to hardening after a sudden increase of the loading rate, which cannot be explained by creep.

So far all our discussions dealt with statics. In dynamic problems, any type of viscosity $\eta$ of the material (present in models for creep, viscoelasticity or viscoplasticity) implies a characteristic length. Indeed, since $\eta$ means stress/strain rate $= kg/ms$, and the Young’s modulus $E$ and mass density $\rho$ have dimensions $E = kg/ms^2$ and $\rho = kg/m^3$, the material length associated with viscosity is given by

$$\ell_v = \frac{\eta}{\sqrt{\rho E}}, \quad v = \sqrt{\frac{E}{\rho}}, \quad (19)$$

where $v$ is the wave velocity. Consequently, any rate-dependence in the constitutive law implies a size effect (and a nonlocal behavior as well).

There is, however, an important difference. Unlike the size effect associated with $\ell_0$ or $c_f$, the viscosity-induced size effect (as well as the width of damage localization zones) is not time-independent. It varies with the rates of loading and deformation of the structure, and vanishes as the rates drop to zero. For this reason, an artificial viscosity or rate effect can approximate the nonviscous size effect and localization only within a narrow range of time delays and rates, but not generally.

7.6 Size effect in fatigue crack growth

Cracks slowly grow under fatigue (repeated) loading. This is for metals and ceramics described by the Paris (or Paris-Erdogan) law, which states that the plot of the logarithm of the crack length increment per cycle versus the amplitude of the stress intensity factor is a rising straight line. For quasibrittle material it turns out that a size increase causes this straight line to shift to the right, the shift being derivable from the size effect law in (10) ([28], Sec. 11.7).

7.7 Size effect for cohesive crack model and crack band model

The cohesive (or fictitious) crack model (called by Hillerborg et al., [62], and Petersson, [87], the fictitious crack model) is more accurate yet less simple than the equivalent LEFM. It is based on the hypothesis that the three-dimensional behavior of a finite FPZ can be approximated by a unique one-dimensional decreasing function $v = g(\sigma)$ relating the crack-opening displacement $w$ (separation of crack faces) to the crack bridging stress $\sigma$ in the FPZ. The obvious way to determine the size effect is to solve $P_{\text{max}}$ by numerical integration for step-by-step loading [87].

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

$$D \int_{x_0}^{x} C^{\sigma}(\xi, \xi') v(\xi) d\xi' = -g'(\sigma(\xi)) v(\xi), \quad P_{\text{max}} = \frac{\int_{x_0}^{x} v(\xi) d\xi}{D \int_{x_0}^{x} C^{\sigma P}(\xi) v(\xi) d\xi}, \quad (20)$$

where the first represents an eigenvalue problem for a homogeneous Fredholm integral equation, with $D$ as the eigenvalue and $v(\xi)$ as the eigenfunction; $\xi = x/D, x$ is the coordinate along the crack, Fig. 6; $x = a/D, x_0 = a_0/D, a, a_0$ are the total and the traction-free crack lengths (or notch lengths); and $C^{\sigma}(\xi, \xi'), C^{\sigma P}(\xi)$ denote the compliance functions of structure for crack surface force and given load $P$. Choosing a sequence of $x$-values, one obtains for each, from (20), the corresponding values of $D$ and $P_{\text{max}}$. These results have also been generalised to obtain directly the load and displacement corresponding, on the load-deflection curve, to a point with any given tangential stiffness, including the displacement at the snapback point which characterizes the ductility of the structure.
The cohesive crack model possesses at least one, but for concrete typically two, independent characteristic lengths: \( \ell_0 = \frac{EG_f}{\sigma_0^2} \) and \( \ell_1 = \frac{EG_F}{\sigma_0^2} \), where \( G_F \) is the area under the entire softening stress–displacement curve \( \sigma = f(w) \), and \( G_f \) is the area under the initial tangent to this curve, which is equal to \( G_F \) only if the curve is simplified as linear (typically \( G_F \approx 2G_f \)). The bilinear stress-displacement law used for concrete involves further parameters of the length dimension – the opening displacement \( w_f \) when the stress is reduced to zero, and the displacement at the change of slope; but their values are implied by \( G_f, G_F, \sigma_0 \) and the stress at the change of slope. The minimum admissible spacing \( h \) of parallel cohesive cracks is also a necessary parameter.

The scatter of size-effect measurements within a practicable size range (up to 1:30) normally does not permit identifying more than one characteristic length (measurements of postpeak behavior are used for that purpose). Vice versa, when only the maximum loads of structures in the bridging region between plasticity and LEFM are of interest, hardly more than one characteristic length (namely \( e_y \)) is needed.

The crack band model, which is easier to implement and is used in commercial codes (e.g. DIANA, SBETA [44]), is, for localized cracking or fracture, nearly equivalent to the cohesive crack model [28, 91], provided that the effective (average) transverse strain in the crack band is taken as \( \bar{\epsilon}_y = w/h \), where \( h \) is the width of the band. All that has been said about the cohesive crack model also applies to the crack band model. Width \( h \), of course, represents an additional characteristic length, \( \ell_4 = h \). It matters only when the cracking is not localized but distributed (e.g. due to the effect of dense and strong enough reinforcement), and it governs the minimum spacings of parallel cracks. Their spacing cannot be unambiguously captured by the standard cohesive crack model.

### 7.8 Size effect via nonlocal, gradient or discrete element models

The hypostatic feature of any model capable of bridging the power law size effects of plasticity and LEFM is the presence of some characteristic length, \( \ell \). In the equivalent LEFM associated with the size effect law in (10), \( e_y \) serves as a characteristic length of the material; although this length can equivalently be identified with \( \delta_{C_TOD} \) in Wells-Cottrell or Jenq-Shah models, or with the crack opening \( w_f \) at which the stress in the cohesive crack model (or crack band model) is reduced to zero (for size effect analysis with the cohesive crack model, see [23, 28]).

In the integral-type nonlocal continuum damage models, \( \ell \) represents the effective size of the representative volume of the material, which in turn plays the role of the effective size of the averaging domain in nonlocal material models. In the second-gradient nonlocal damage models, which may be derived as an approximation of the nonlocal damage models, a material length is involved in the relation of the strain to its Laplacian. In damage simulation by the discrete element (or random particle) models, the material length is represented by the statistical average of particle size.

The existence of \( \ell \) in these models engenders a quasibrittle size effect that bridges the power-law size effects of plasticity and LEFM and follows closely Eq. (10) with \( \sigma_N = 0 \), as documented by numerous finite element simulations. It also poses a lower bound on the energy dissipation during failure, prevents spurious excessive localization of softening continuum damage, and eliminates spurious mesh sensitivity ([28], Ch. 13).

These important subjects will not be discussed here any further because there exists a recent extensive review [14].
7.9 Nonlocal statistical generalization of Weibull theory

Two cases need to be distinguished: (a) the front of the fracture that causes failure can lie at only one place in the structure, or (b) the front can lie, with different probabilities, at many different places. The former case occurs when a long crack, whose path is dictated by fracture mechanics, grows before attaining the maximum load, or if a notch is cut in a test specimen. The latter case occurs when the maximum load is achieved at the initiation of fracture growth.

In both cases, the existence of a large FPZ calls for a modification of Weibull concept: the failure probability \( P \) at a given point of the continuous structure depends not on the local stress at that point, but on the nonlocal strain, which is calculated as the average of the local strains within the neighborhood of the point constituting the representative volume of the material. The nonlocal approach broadens the applicability of Weibull concept to the case of notches or long cracks, for which the existence of crack-tip singularity causes the classical Weibull probability integral to diverge at realistic \( m \)-values. In cleavage fracture of metals, the problem of crack singularity has been circumvented differently – by dividing the crack-tip plastic zone into small elements and superposing their Weibull contributions [71].

Using the nonlocal Weibull theory, one can show that the proper statistical generalizations of (10), with \( \sigma_B = 0 \) and Eq. (12) having the correct asymptotic forms for \( D \to \infty, D \to 0 \) and \( m \to \infty \), are:

**Case (a):** \[ \sigma_N = B\sigma_0 (\beta^{\frac{m_D}{m_B}} + \beta^m)^{-\frac{1}{m}}, \quad \beta = \frac{D}{D_0}, \quad (21) \]

**Case (b):** \[ \sigma_N = \sigma_0 \xi \pi (1 + \xi \frac{n_D}{n_B})^\frac{1}{2}, \quad \xi = \frac{D_b}{D}, \quad (22) \]

(Fig. 7) where it is assumed that \( n_D < m \), which is normally the case.

The first formula, which was obtained for \( r = 1 \) in [31] and refined for \( r \neq 1 \) by Planas, has the property that the statistical influence on the size effect disappears asymptotically for large \( D \). The reason is that, for long cracks or notches with stress singularity, a significant contribution to the Weibull probability integral comes only from the FPZ, whose size does not vary much with \( D \). The second formula has the property that the statistical influence asymptotically disappears for small sizes. The reason is that the FPZ occupies much of the structure volume.

Numerical analyses of test data for concrete show that the size ranges in which the statistical influence on the size effect in case (a) as well as (b) would be significant do not lie within the range of practical interest except for unreinforced beams many meters thick (arch dams, foundation plinths). Thus, the deterministic size effect dominates and its statistical correction in (21) and

![Fig. 7. Scaling laws according to nonlocal generalization of Weibull theory for failures after long stable crack growth (a) or at crack initiation (b)](image-url)
(22) may be ignored for normal concrete structures, with the exception of the rare situations where the deterministic size effect vanishes (e.g. for centric tension of an unreinforced bar).

8 Other size effects

8.1 Hypothesis of fractal origin of size effect

The partly fractal nature of crack surfaces and of the distribution of microcracks in concrete has recently been advanced as the physical origin of the size effects observed on concrete structures. A possible role of fractality in size effects in sea ice was discussed in [33]. In [38–40] the so-called 'multi-fractal scaling law' (MFSL) was proposed for failures occurring at fracture initiation from a smooth surface; it reads

$$\sigma_N = \sqrt{A_1 + \frac{A_2}{D}},$$

where $A_1, A_2 = \text{const}$. There are, however, four objections to the fractal theory [13]:

(i) a mechanical analysis (of either invasive or lacunar fractals) predicts a different size effect trend than (23), disagreeing with experimental observations;

(ii) the fractality of the final fracture surface should not matter, because typically about 99% of energy is dissipated by microcracks and frictional slips on the sides of this surface;

(iii) the fractal theory does not predict how $A_1$ and $A_2$ should depend on the geometry of the structure, which makes the MFSL not too useful for design application;

(iv) the MFSL is a special case of the second formula in (12) for $r = 2$, which logically follows from fracture mechanics:

$$A_1 = \frac{EG_f}{c_f g(0)}, \quad A_2 = -\frac{EG_f g''(0)}{2c_f [g'(0)]^3},$$

[7]. Unlike fractality, the fracture explanation of (23) has the advantage that, by virtue of these formulae, the geometry dependence of the size effect coefficients can be determined.

8.2 Boundary layer, singularity, and diffusion

Aside from the statistical and quasibrittle size effects, there are three further types of size effect that influence nominal strength:

1. The boundary layer effect, which is due to material heterogeneity (i.e. the fact that the surface layer of heterogeneous material such as concrete has a different composition because the aggregates cannot protrude through the surface), and to Poisson effect (i.e. the fact that a plane strain state on planes parallel to the surface can exist in the core of the test specimen but not at its surface).

2. The existence of a three-dimensional stress singularity at the intersection of crack edge with a surface, which is also caused by the Poisson effect ([28], Sec. 1.3). This causes the portion of the FPZ near the surface to behave differently from that in the interior.

3. The time-dependent size effects caused by diffusion phenomena such as the transport of heat or the transport of moisture and chemical agents in porous solids (this is manifested, e.g., in the effect of size on shrinkage and drying creep, due to size dependence of the drying half-time [22] and its effect on shrinkage cracking [90].

9 Closing remarks

Substantial though the recent progress has been, the understanding of the scaling problems of solid mechanics is nevertheless far from complete. Mastering the size effect that bridges different behaviors on adjacent scales in the microstructure of material will be contingent upon the development of realistic material models that possess a material (or characteristic) length. The theory of nonlocal continuum damage will have to move beyond the present phenomenological approach based on isotropic spatial averaging, and take into account the directional and tensorial interactions between the effects causing nonlocality. A statistical description of such interactions will have to be developed. Discrete element models of the microstructure of fracturing or
damaging materials will be needed to shed more light on the mechanics of what is actually
happening inside the material and to separate the important processes from the unimportant
ones.

Appendix:

Broad-range size effect law

The simple size effect law in (10) is normally adequate for a size range up to about 1:20. For a
size range up to about 1:500, the generalized size effect law (11) (with \( \sigma_R = 10 \)) has been shown
to give excellent approximation of the numerical results obtained by Hillerborg with the co-
hesive crack model for the range 1:250 (Bažant 1985). However, only the first term of the
asymptotic expansion in terms of \( 1/D \) is correct. For a modified formula \( \sigma_N = \sigma_R (1 + R)^{-1/2} \),
the first two terms in \( 1/D \) are correct but the description of size effect in the intermediate range,
which is more important, is poor.

The following three broad-range size effect laws (proposed here), which are capable of
approximating the large-size asymptotic behavior (ensuing for example from the cohesive
crack model) up to order \( n/2 \) in \( 1/D \), are currently investigated:

\[
\sigma_N^2 = \frac{\sigma_0^2}{1 + D/D_0} \sum_{n=1}^{N} h_n (1 - e^{-D_n/D_0^2}), \quad \sum_{n=1}^{N} h_n = 1
\]

(25)

\[
\sigma_N^2 = \frac{\sigma_0^2}{1 + D/D_0} \sum_{n=1}^{N} \frac{h_n D_n^0}{D_n^0 + D_n^4}
\]

(26)

\[
\sigma_N^2 = \frac{\sigma_0^2}{1 + D/D_0} \left( \frac{1 + D/H_1}{1 + D/D_1} \right) \left( \frac{1 + D/H_2}{1 + D/D_2} \right) \cdots \left( \frac{1 + D/H_N}{1 + D/D_N} \right), \quad \sigma_0 = \sqrt{\frac{EG_f}{g'(x)G_f}}
\]

(27)

Here \( D_1, \ldots, D_N, H_1, \ldots, H_N, h_1, \ldots, h_N, D_1, \ldots, D_N \) and \( q \) are positive constants. To ensure that
each term of the chain product in (3) would not increase, it is necessary that \( H_i \geq D_i \)
\( (i = 1, 2, \ldots, N) \). It is evident, that these formulae preserve the finite limit \( \lim_{D \to 0} \sigma_N = \sigma_0 \).

Progressively increasing the number of terms in these formulae, one obtains the 'cascading'
size effect plot shown in Fig. 8, in which the individual terms of (25)–(27) terminate with
progressively shifted asymptotes of slope \(-1/2\). Each asymptote corresponds to a progressively
higher fracture energy \( G_{f,i} \). Obtaining very accurate \( \sigma_N \) values, e.g., with the cohesive crack
model, for a broad range such as 1:10000, and fitting these results with one of these three
formulae, one could thus obtain a set of \( G_{f,i} \) values, which can be regarded as a fracture energy
spectrum of the cohesive crack model defined by a certain softening stress-displacement curve.

The first fracture energy, \( G_{f,1} \), corresponds to the area under the initial tangent of the
softening curve. This is known to suffice for predicting with the cohesive crack model the
maximum loads of structures for a range of about 1:20 (which covers most practical cases). If
the size range is extended much farther, say as 1:200, then much of the area under the softening
curve of the cohesive crack model, except for a remote tail of the curve, must matter for the
maximum load predictions. This area may then be associated with the fracture energy \( G_{f,2} \),
which ought to coincide with the fracture energy $G_F$ determined by the work-of-fracture method introduced for concrete by Hillerborg. It is well known that $G_F$ is about 2 to 2.5 times larger than $G_f$. This apparent discrepancy can thus be explained in terms of the broad-range size effect law.

In analogy to what is known from the theory of retardation spectrum in linear viscoelasticity, one may expect the problem of identification of $D_i$ from a given size effect curve to be ill-conditioned, which means that these values would need to be chosen, within some constraints, just as the retardation times of a viscoelastic material with a broad retardation spectrum must be chosen. Therefore, the discrete fracture energy spectrum $G_{f1}, G_{f2}, \ldots, G_{fN}$ will be non-unique, depending on the choice of $D_0, D_1, \ldots, D_N$ in the foregoing formulae. Similar to the continuous retardation spectrum in viscoelasticity, one could consider infinitely closely spaced $D_i$ values and generalize (25) to a continuous spectrum of fracture energies, which would then be unique.

References

27. **Bažant, Z. P.; Pfeiffer, P. A.:** Determination of fracture energy from size effect and brittleness number. ACI Materials J. 84 (1987) 463–480
29. **Bažant, Z. P.; Pfeiffer, P. A.** Determination of fracture energy from size effect and brittleness number. ACI Materials Jour. 84 (1987) 463–480
38. **Carpinteri, A.:** Fractal nature of material microstructure and size effects on apparent mechanical properties. Mech. Mat. 18 (1994) 89–101
43. **Carter, B. C.; Lajtai, E. Z.; Yuan, Y.:** Tensile fracture from circular cavities loaded in compression. Int. J. Fracture 57 (1992) 221–236
45. **Cottrell, A. H.:** Iron and Steel Institute Special Report 69 (1963) 281
47. **Dempsey, J. P.; Adamson, R. M.; Mulmule, S. V.:** Scale effect on the in-situ tensile strength and failure of first-year sea ice at Resolute, NWR. Int. J. Fracture (Special Issue on Fracture Scaling, Bažant, Z. P.; Rajapakse, Y. D. S.: (eds.) 95 (1999) 325–345
50. **Evans, A. G.:** A general approach for the statistical analysis of multiaxial fracture. J. of the American Ceramic Soc. 61 (1978) 302–308
52. **Fischer, R. A.; Tippett, L. H. C.:** Limiting forms of the frequency distribution of the largest and smallest member of a sample. Proc., Cambridge Philosophical Society 24 (1928) 180–190
60. **Griffith, A. A.:** The phenomena of rupture and flow in solids. Phil. Trans. 221A (1921) 179–180
101. Tippett, L. H. C.: On the extreme individuals and the range of samples. Biometrika 17 (1925) 364
115. Wittmann, F. H.; Zaitsev, Yu., V.: Crack propagation and fracture of composite materials such as concrete. Proc. 5th Int. Conf. on Fracture (ICF5), Cannes 1981