CHAPTER ONE

Brittle fracture of rock

Under the low-temperature and pressure conditions of Earth's upper lithosphere, silicate rock responds to large strains by brittle fracture. The mechanism of brittle behavior is by the propagation of cracks, which may occur on all scales. We begin by studying this form of deformation, which is fundamental to the topics that follow.

1.1 THEORETICAL CONCEPTS

1.1.1 Historical

Understanding the basic strength properties of rock has been a practical pursuit since ancient times, both because of the importance of mining and because rock was the principal building material. The crafting of stone tools required an intuitive grasp of crack propagation, and mining, quarrying, and sculpture are trades that require an intimate knowledge of the mechanical properties of rock. The layout and excavation of quarries, for example, is a centuries-old art that relies on the recognition and exploitation of preferred splitting directions in order to maximize efficiency and yield. One of the principal properties of brittle solids is that their strength in tension is much less than their strength in compression. This led, in architecture, to the development of fully compressional structures through the use of arches, domes, and flying buttresses.

Rock was one of the first materials for which strength was studied with scientific scrutiny because of its early importance as an engineering material and in mining. By the end of the nineteenth century the macroscopic phenomenology of rock fracture had been put on a scientific basis. Experimentation had been conducted over a variety of conditions up to moderate confining pressures. The Coulomb criterion and the Mohr circle analysis had been developed and applied to rock fracture with sufficient success that they remain the principal tools used to describe this process for many engineering and geological applications.

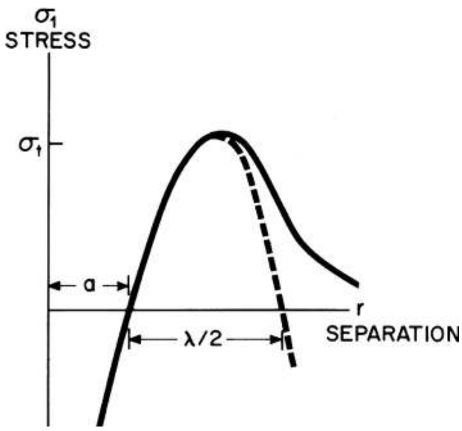

Fig. 1.1. Sketch of an anharmonic model of interatomic forces, showing the relationship between stress and atomic separation (solid curve) and a sinusoidal approximation (dashed curve).

The modern theory of brittle fracture arose as a solution to a crisis in understanding the strength of materials, brought about by the atomic theory of matter. In simplest terms, strength can be viewed as the maximum stress that a material can support under given conditions. Fracture (or flow) must involve the breaking of atomic bonds. An estimate of the *theoretical strength* of a solid is therefore the stress required to break the bonds across a lattice plane.

Consider a simple anharmonic model for the forces between atoms in a solid, as in Figure 1.1, in which an applied tension σ produces an increase in atomic separation r from an equilibrium spacing a (Orowan, 1949). Because we need only consider the prepeak region, we can approximate the stress–displacement relationship with a sinusoid,

$$\sigma = \sigma_t \sin\left[\frac{2\pi(r-a)}{\lambda}\right] \tag{1.1}$$

For small displacements, when $r \approx a$, then

$$\frac{d\sigma}{d(r-a)} = \frac{E}{a} = \frac{2\pi}{\lambda}\sigma_t \cos\left[\frac{2\pi(r-a)}{\lambda}\right] \tag{1.2}$$

but because $(r-a)/\lambda \ll 1$, the cosine is equal to 1, and

$$\sigma_t = \frac{E\lambda}{2\pi a} \tag{1.3}$$

where E is Young's modulus. When $r = 3a/2$, the atoms are midway between two equilibrium positions, so by symmetry, $\sigma = 0$ there and $a \approx \lambda$. The theoretical strength is thus about $E/2\pi$. The work done in separating the planes by $\lambda/2$ is the specific surface energy γ, the energy per unit area required to break the bonds, so

$$2\gamma = \int_0^{\lambda/2} \sigma_t \sin\left[\frac{2\pi(r-a)}{\lambda}\right] d(r-a) = \frac{\lambda\sigma_t}{\pi} \tag{1.4}$$

which, with $\sigma_t \approx E/2\pi$, yields the estimate $\gamma \approx Ea/4\pi^2$.

The value of the theoretical strength from this estimate is 5–10 GPa, several orders of magnitude greater than the strength of real materials. This discrepancy was explained by the postulation and later recognition that all real materials contain defects. Two types of defects are important: cracks, which are surface defects; and dislocations, which are line defects. Both types of defects may propagate in response to an applied stress and produce yielding in the material. This will occur at applied stresses much lower than the theoretical strength, because both mechanisms require that the theoretical strength be achieved only locally within a *stress concentration* deriving from the defect. The two mechanisms result in grossly different macroscopic behavior. When cracks are the active defect, material failure occurs by its separation into parts, often catastrophically: this is brittle behavior. Plastic flow results from dislocation propagation, which produces permanent deformation without destruction of the lattice integrity.

These two processes tend to be mutually inhibiting, but not exclusive, so that the behavior of crystalline solids usually can be classed as brittle or ductile, although mixed behavior, known as semibrittle, may be more prevalent than commonly supposed. Because the lithosphere consists of two parts with markedly different rheological properties, one brittle and the other ductile, it is convenient to introduce two new terms to describe them. These are *schizosphere* (literally, the broken part) for the brittle region, and *plastosphere* (literally, the moldable part) for the ductile region. In this book we will assume, for the most part, that we are dealing with purely brittle processes, so that we will be concerned principally with the behavior of the schizosphere.

1.1.2 Griffith theory

All modern theories of strength recognize, either implicitly or explicitly, that real materials contain imperfections that, because of the stress concentrations they produce within the body, result in failure at much lower stresses than the theoretical strength. A simple example, Figure 1.2(a), is a hole within a plate loaded with a uniform tensile stress σ_∞. It can be shown

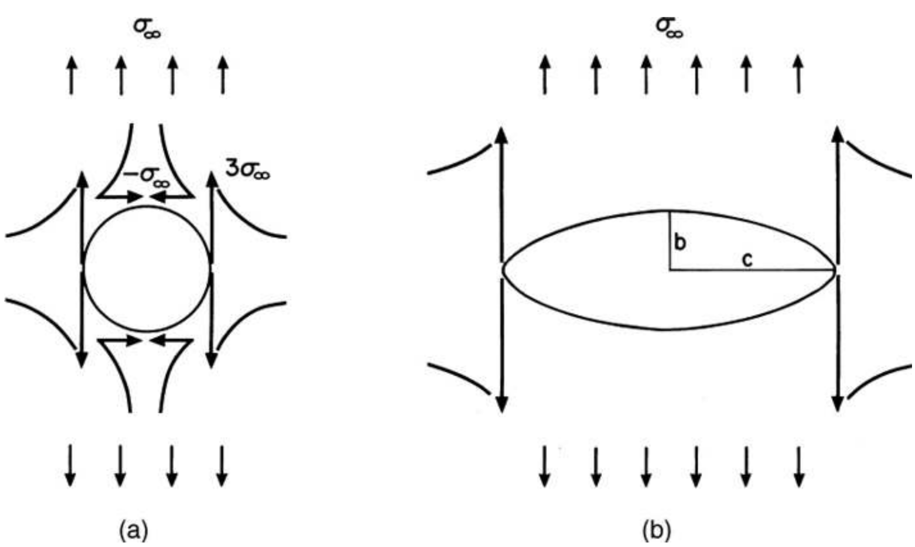

Fig. 1.2. Stress concentration around (a) a circular hole, and (b) an elliptical hole in a plate subjected to a uniform tension σ_∞.

from elasticity theory that at the top and bottom of the hole a compressive stress of magnitude $-\sigma_\infty$ exists and that at its left and right edges there will be tensile stresses of magnitude $3\sigma_\infty$. These stress concentrations arise from the lack of load-bearing capacity of the hole, and their magnitudes are determined solely by the geometry of the hole and not by its size. If the hole is elliptical, as in Figure 1.2(b), with semiaxes b and c, with $c > b$, the stress concentration at the ends of the ellipse increases proportionally to c/b, according to the approximate formula

$$\sigma \approx \sigma_\infty(1 + 2\ c/b)$$

or

$$\sigma \approx \sigma_\infty\left[1 + 2(c/\rho)^{1/2}\right] \approx \sigma_\infty(c/\rho)^{1/2} \tag{1.5}$$

for $c \gg b$, where ρ is the radius of curvature at that point. It is clear that for a long narrow crack the theoretical strength can be attained at the crack tip when $\sigma_\infty \ll \sigma_t$. Because Equation (1.5) indicates that the stress concentration will increase as the crack lengthens, crack growth can lead to a dynamic instability.

Griffith (1920; 1924) posed this problem at a more fundamental level, in the form of an energy balance for crack propagation. The system he considered is shown in Figure 1.3(a) and consists of an elastic body that contains a crack of length 2 c, which is loaded by forces on its external boundary. If the crack extends an increment δc, work W will be done by the external forces and there will be a change in the internal strain energy U_e. There will also be an

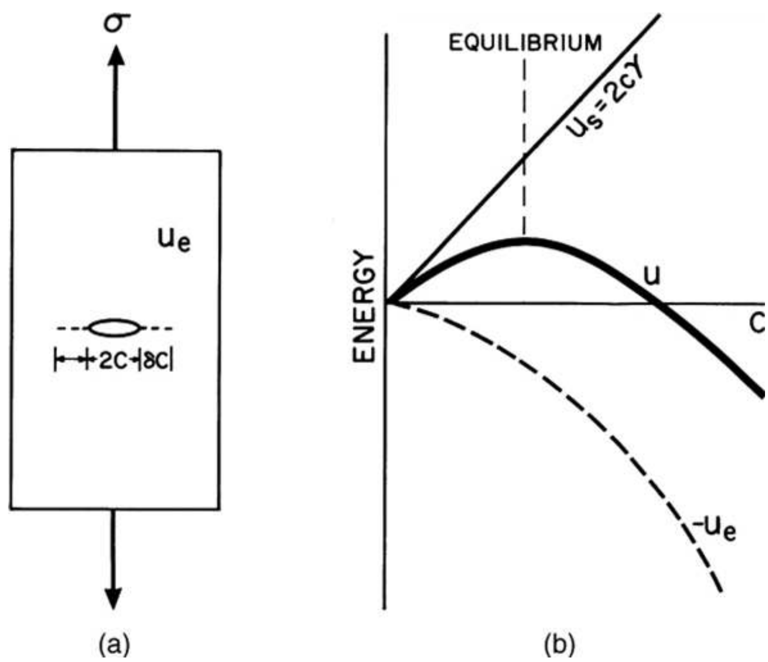

Fig. 1.3. Griffith's model for a crack propagating in a rod (a), and the energy partition for the process (b).

expenditure of energy in creating the new surfaces U_s. Thus, the total energy of the system, U, for a static crack, will be

$$U = (-W + U_e) + U_s \tag{1.6}$$

The combined term in parentheses is referred to as the mechanical energy. It is clear that, if the cohesion between the incremental extension surfaces δc were removed, the crack would accelerate outward to a new lower energy configuration. Thus, mechanical energy must decrease with crack extension. The surface energy, however, will increase with crack extension, because work must be done against the cohesion forces in creating the new surface area. There are two competing influences; for the crack to extend there must be reduction of the total energy of the system, and hence at equilibrium there is a balance between them. The condition for equilibrium is

$$dU/dc = 0 \tag{1.7}$$

Griffith analyzed the case of a rod under uniform tension. A rod of length y, modulus E, and unit cross section loaded under a uniform tension will have strain energy $U_e = y\sigma^2/2E$. If a crack of length $2\ c$ is introduced into the rod, it can be shown that the strain energy will increase an amount $\pi c^2\sigma^2/E$, so that U_e becomes

$$U_e = \sigma^2(y + 2\pi c^2)/2E \tag{1.8}$$

The rod becomes more compliant with the crack, with an effective modulus $\underline{E} = yE/(y + 2\pi c^2)$. The work done in introducing the crack is

$$W = \sigma y(\sigma/\underline{E} - \sigma/E) = 2\pi^2 c^2/E \tag{1.9}$$

and the surface energy change is

$$U_s = 4cy \tag{1.10}$$

Substituting Equations (1.8)–(1.10) into Equation (1.6) gives

$$U = -\pi c^2\sigma^2/E + 4cy \tag{1.11}$$

and applying the condition for equilibrium (Equation (1.7)), we obtain an expression for the critical stress at which a suitably oriented crack will be at equilibrium,

$$\sigma_f = (2E\gamma/\pi c)^{1/2} \tag{1.12}$$

The energies of the system are shown in Figure 1.3(b), from which it can be seen that Equation (1.12) defines a position of unstable equilibrium: when this condition is met the crack will propagate without limit, causing macroscopic failure of the body.

Griffith experimentally tested his theory by measuring the breaking strength of glass rods that had been notched to various depths. He obtained an experimental result with the form of Equation (1.12) from which he was able to extract an estimate of γ. He obtained an independent estimate of γ by measuring the work necessary to pull the rods apart by necking at elevated temperatures. By extrapolating this result to room temperature, he obtained a value that was within reasonable agreement with that derived from the strength tests.

Griffith's result stems strictly from a consideration of thermodynamic equilibrium. Returning to our original argument, we may ask if the theoretical strength is reached at the crack tip when the Griffith condition is met: that is, is the stress actually high enough to break the bonds? This question was posed by Orowan (1949), who considered the stress at the tip of an atomically narrow crack, as described before. Combining Equations (1.3) and (1.4), we obtain

$$\sigma_t = (E\gamma/a)^{1/2} \tag{1.13}$$

This stress will exist at the ends of a crack of length 2 c when the macroscopic applied stress σ_f is (Equation (1.5))

$$\sigma_t = 2\sigma_f(c/a)^{1/2} \tag{1.14}$$

so that

$$\sigma_f = (E\gamma/4c)^{1/2} \tag{1.15}$$

which is very close to Equation (1.12). The close correspondence of these two results demonstrates both necessary and sufficient conditions for crack propagation. Griffith's thermodynamic treatment shows the condition for which the crack is energetically favored to propagate, while Orowan's calculation shows the condition in which the crack-tip stresses are sufficient to break atomic bonds. For a typical value of $\gamma \approx Ea/30$ (Equation (1.4)), commonly observed values of strength of $E/500$ can be explained by the presence of cracks of length $c \approx 1$ μm. Prior to the advent of the electron microscope, the ubiquitous presence of such microscopic cracks was hypothetical, and this status was conferred upon them with the use of the term *Griffith crack.*

Griffith's formulation has an implicit instability as a consequence of the constant stress boundary condition. In contrast, the experiment of Obriemoff (1930) leads to a stable crack configuration. Obriemoff measured the cleavage strength of mica by driving a wedge into a mica book using the configuration shown in Figure 1.4(a). In this experiment the boundary condition is one of constant displacement. Because the wedge can be considered to be rigid, the bending force F undergoes no displacement and the external work done on the system is simply

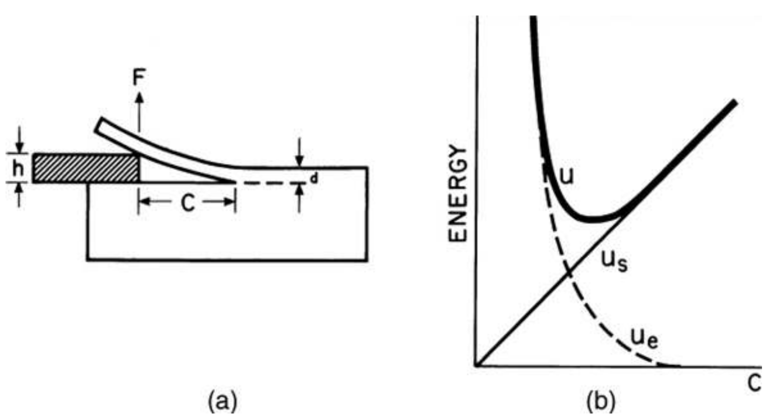

Fig. 1.4. The configuration of Obriemoff's mica cleaving experiment (a), and the energy partition for this process (b).

$$W = 0 \tag{1.16}$$

From elementary beam theory, the strain energy in the bent flake is

$$U_e = Ed^3 h^2/8c^3 \tag{1.17}$$

and, using $U_s = 2c\gamma$ and the condition $dU/dc = 0$, we obtain the equilibrium crack length

$$c = \left(3Ed^3 h^2/16\gamma\right)^{1/4} \tag{1.18}$$

The energies involved in this system are shown in Figure 1.4(b). It is clear that in this case the crack is in a state of stable equilibrium; it advances the same distance that the wedge is advanced. This example shows that the stability is controlled by the system response, rather than being a material property, a point that will be taken up in greater detail in the discussion of frictional instabilities in Section 2.3. In this case the loading system may be said to be infinitely stiff, and crack growth is controlled and stable. Griffith's experiment, on the other hand, had a system of zero stiffness and the crack was unstable. Most real systems, however, involve loading systems with finite stiffness so that the stability has to be evaluated by balancing the rate at which work is done by the loading system against the energy absorbed by crack propagation.

Obriemoff noticed that the cracks in his experiment did not achieve their equilibrium length instantly, but that on insertion of the wedge they jumped forward and then gradually crept to their final length. When he conducted the experiment in vacuum, however, he did not observe this transient effect. Furthermore, the surface energy that he measured in vacuum was about 10 times the surface energy measured in ambient atmosphere. He was thus the first to observe the important effect of the chemical environment on the weakening of brittle solids and the *subcritical crack growth* that results from this effect. This effect is very important in brittle processes in rock and will be discussed in more detail in Section 1.2.4.

1.1.3 Fracture mechanics

Linear elastic fracture mechanics is an approach that has its roots in the Griffith energy balance, but that lends itself more readily to the solution of general crack problems. It is a continuum mechanics approach in which the crack is idealized as a mathematically flat and narrow slit in a linear elastic medium. It consists of analyzing the stress field around the crack and then formulating a fracture criterion based on certain critical parameters of the stress field. The macroscopic strength is thus related to the intrinsic strength of the material through the relationship between the applied stresses and the crack-tip stresses. Because the crack is treated as residing in a continuum, the details of the deformation and fracturing processes at the crack tip are ignored.

The displacement field of cracks can be categorized into three modes (Figure 1.5). Mode I is the tensile, or opening, mode, in which the crack wall displacements are normal to the crack. There are two shear modes: in-plane shear, Mode II, in which the displacements are in the plane of the crack and normal to the crack edge; and antiplane shear, Mode III, in which the displacements are in the plane of the crack and parallel to the edge. The latter are analogous to edge and screw dislocations, respectively.

If the crack is assumed to be planar and perfectly sharp, with no cohesion between the crack walls, then the near-field approximations to the crack-tip stress and displacement fields may be reduced to the simple analytic expressions:

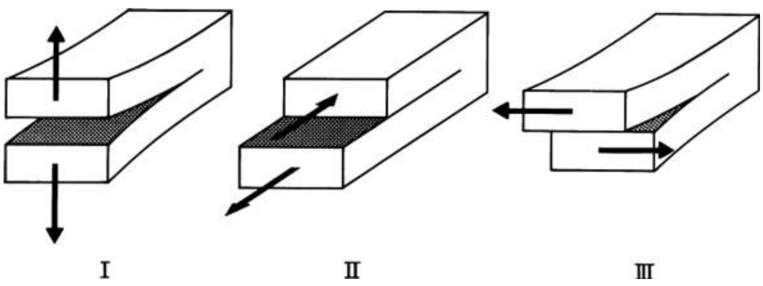

Fig. 1.5. The three crack propagation modes.

$$\sigma_{ij} = K_n(2\pi r)^{-1/2} f_{ij}(\theta) \tag{1.19}$$

and

$$u_i = (K_n/2E)(r/2\pi)^{1/2} f_i(\theta) \tag{1.20}$$

where r is the distance from the crack tip and θ is the angle measured from the crack plane, as shown in Figure 1.6. K_n is called the *stress-intensity factor* and depends on mode, that is, K_{I}, K_{II}, and K_{III}, refer to the three corresponding crack modes. The functions $f_{ij}(\theta)$ and $f_i(\theta)$ can be found in standard references (e.g. Lawn, 2010), and are illustrated in Figure 1.6. The stress–intensity factors depend on the geometry and magnitudes of the applied loads and determine the intensity of the crack–tip stress field. They also can be found tabulated, for common geometries, in standard references (e.g. Tada, Paris, and Irwin, 1973). The other terms describe only the distribution of the fields.

In order to relate this to the Griffith energy balance it is convenient to define an *energy release rate*, or *crack extension force*,

$$\mathcal{G} = -d(-W + U_e)/dc \tag{1.21}$$

which can be related to K by

$$\mathcal{G} = K^2/E \tag{1.22}$$

(Lawn 2010, p. 29) for plane stress or

$$\mathcal{G} = K^2(1 - \nu^2)/E \tag{1.23}$$

for plane strain (ν is Poisson's ratio). In Mode III, the right–hand sides of the corresponding expressions must be multiplied by $(1 + \nu)$ for plane stress and divided by $(1 - \nu)$ for plane strain, respectively. From Equations (1.6) and (1.7), it is clear that the condition for crack propagation will be met when

$$\mathcal{G}_c = K_c^2/E = 2\gamma \tag{1.24}$$

for plane stress, with a corresponding expression for plane strain. Thus $\mathcal{K}_c$, the *critical stress-intensity factor*, and $\mathcal{G}_c$ are material properties that, because they can be related to the applied

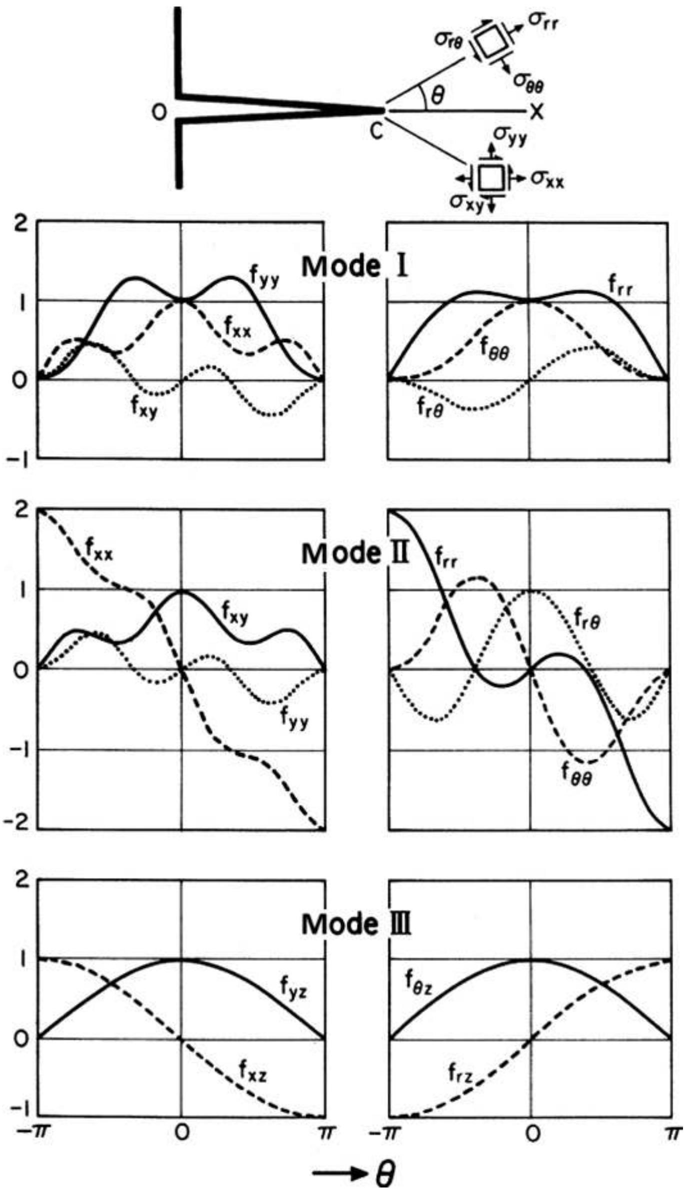

Fig. 1.6. The stress functions near the tips of the three modes of cracks, using both Cartesian and cylindrical coordinates, as shown in the geometrical key. (After Lawn, 2010.)

stresses through a stress analysis, provide powerful and general failure criteria. $\mathcal{K}_c$ is also sometimes called the *fracture toughness*, and $\mathcal{G}_c$ the *fracture energy.*

A simple and useful case is when uniform stresses σ_{ij} are applied remote from the crack, as in Figure 1.7. In this case the stress–intensity factors are given by

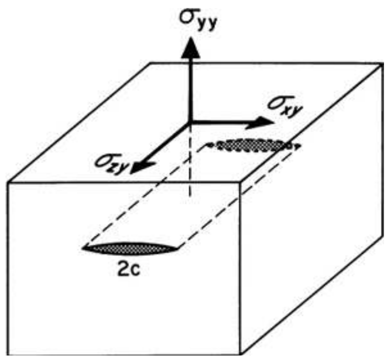

Fig. 1.7. Geometry of a crack in a uniform stress field.

$$\left.\begin{aligned} K_{\rm I} &= \sigma_{yy}(\pi c)^{1/2} \\ K_{\rm II} &= \sigma_{xy}(\pi c)^{1/2} \\ K_{\rm III} &= \sigma_{zy}(\pi c)^{1/2} \end{aligned}\right\} \tag{1.25}$$

and, using Equation (1.22), the corresponding crack extension forces, for plane stress, are

$$\left.\begin{aligned} G_{\rm I} &= (\sigma_{yy})^2 \pi c/E \\ G_{\rm II} &= (\sigma_{xy})^2 \pi c/E \\ G_{\rm III} &= (\sigma_{zy})^2 \pi c(1+\nu)/E \end{aligned}\right\} \tag{1.26}$$

In plane strain, E is replaced by $E/(1-\nu^2)$ for Modes I and II.

Equation (1.25) may be compared with the approximate expression for the stress concentration at the tip of an elliptical crack, Equation (1.5). However, inspection of Equation (1.19) indicates that there is a stress singularity at the crack tip. This results from the assumptions of perfect sharpness of the slit. This is nonphysical, both because it internally violates the assumption of linear elasticity, which implies small strains, and because no real material can support an infinite stress. There must be a region of nonlinear deformation near the crack tip that relaxes this singularity. This can be ignored in the fracture mechanics approach, because it can be shown that the strain energy in the nonlinear zone is bounded, and because the small nonlinear zone does not significantly distort the stress field at greater distances from the crack. It is, of course, of paramount importance for studies concerned with the detailed mechanics of crack advancement, but it suffices here to state that linear elastic fracture mechanics is not applicable at that scale or if there is large-scale yielding.

Within the nonlinear zone distributed cracking, plastic flow, and other dissipative processes may occur that contribute to the crack extension force. To account for these additional contributions we can rewrite Equation (1.24) as

$$\mathcal{G}_{\rm c} = 2\Gamma \tag{1.27}$$

where Γ is a lumped parameter that includes all dissipation within the crack-tip region. This failure criterion is associated with the work of Irwin (1958). The fact that we do not usually know the specific processes that contribute to G is not normally of practical significance because $\mathcal{G}$ still can be evaluated if mechanical measurements can be made suitably outside the nonlinear zone (because integration around the crack tip is path-independent [Rice, 1968]).