2917-CH06.pdf 11/20/03 5:10 PM Page 113

PART B

# BRITTLE STRUCTURES

CHAPTER SIX

# Brittle Deformation



## 6.1 INTRODUCTION

Drop a glass on a tile floor and watch as it breaks into dozens of pieces; you have just witnessed an example of brittle deformation! Because you've probably broken a glass or two (or a plate or a vase) and have seen cracked buildings, sidewalks, or roads, you already have an intuitive feel for what brittle deformation is all about. In the upper crust of the Earth, roughly 10 km in depth, rocks primarily undergo brittle deformation, creating a myriad of geologic structures. To understand why these structures exist, how they form in rocks of the crust, and what they tell us about the conditions of deformation, we must first learn why and how brittle deformation takes place in materials in general. Our purpose in this chapter is to introduce the basic terminology used to describe brittle deformation, to explain the processes by which brittle deformation takes place, and to describe the physical conditions that lead to brittle deformation. This chapter, therefore, provides a basis for the discussion of brittle structures that we present in Chapters 7 and 8.

## 6.2 VOCABULARY OF BRITTLE DEFORMATION

Research in the last few decades has changed the way geologists think about brittle deformation and, as a consequence, the vocabulary of brittle deformation has evolved. To avoid misunderstanding, therefore, we begin the chapter on brittle deformation by defining our terminology. Table 6.1 summarizes definitions of terms that we use in discussing brittle deformation, and includes additional terms that you may come across when reading articles on this topic. Remember that some of the brittle deformation vocabulary is

**TABLE 6.1** | **TERMINOLOGY OF BRITTLE DEFORMATION**

| | |
|---|---|
| Brittle deformation | The permanent change that occurs in a solid material due to the growth of fractures and/or due to sliding on fractures. Brittle deformation only occurs when stresses exceed a critical value, and thus only after a rock has already undergone some elastic and/or plastic behavior. |
| Brittle fault zone | A band of finite width in which slip is distributed among many smaller discrete brittle faults, and/or in which the fault surface is bordered by pervasively fractured rock. |
| Brittle fault | A single surface on which movement occurs specifically by brittle deformation mechanisms. |
| Cataclasis | A deformation process that involves distributed fracturing, crushing, and frictional sliding of grains or of rock fragments. |
| Crack | *Verb:* to break or snap apart. *Noun:* a fracture whose displacement does not involve shear displacement (i.e., a joint or microjoint). |
| Fault | *Broad sense:* a surface or zone across which there has been measurable sliding parallel to the surface. *Narrow sense:* a brittle fault. The narrow definition emphasizes the distinctions between faults, fault zones, and shear zones. |
| Fracture zone | A band in which there are many parallel or subparallel fractures. If the fractures are wavy, they may anastomose with one another. *Note:* The term has a somewhat different meaning in the context of ocean-floor tectonics. |
| Fracture | A general term for a surface in a material across which there has been loss of continuity and, therefore, strength. Fractures range in size from grain-scale to continent-scale. |
| Healed microcrack | A microcrack that has cemented back together. Under a microscope, it is defined by a plane containing many fluid inclusions. (Fluid inclusions are tiny bubbles of gas or fluid embedded in a solid). |
| Joint | A natural fracture which forms by tensile loading, i.e., the walls of the fracture move apart very slightly as the joint develops. *Note:* A minority of geologists argue that joints can form due to shear loading. |
| Microfracture | A very small fracture of any type. Microfractures range in size from the dimensions of a single grain to the dimensions of a thin section. |
| Microjoint | A microscopic joint; microjoints range in size from the dimensions of a single grain to the dimensions of a hand-specimen. Synonymous with *microcrack.* |
| Shear fracture | A macroscopic fracture that grows in association with a component of shear parallel to the fracture. Shear fracturing involves coalescence of microcracks. |
| Shear joint | A surface that originated as a joint but later became a surface of sliding. *Note:* A minority of geologists consider a shear joint to be a joint that initially formed in response to shear loading. |
| Shear rupture | A shear fracture. |
| Shear zone | A region of finite width in which ductile shear strain is significantly greater than in the surrounding rock. Movement in shear zones is a consequence of ductile deformation mechanisms (cataclasis, crystal plasticity, diffusion). |
| Vein | A fracture filled with minerals precipitated from a water solution. |

**FIGURE 6.1** A geologist measuring fractures in an outcrop, near Tuross Point on the southeastern coast of Australia. The more intensely fractured rock is a fine-grained mafic intrusive and the less fractured rock is a coarse-grained felsic intrusive.

controversial, and you will find that not all geologists agree on the definitions that we provide.

**Brittle deformation** is simply the permanent change that occurs in a solid material due to the growth of fractures and/or due to sliding on fractures once they have formed. By this definition, a **fracture** is any surface of discontinuity, meaning a surface across which the material is no longer bonded (Figure 6.1). If a fracture fills with minerals precipitated out of a hydrous solution, it is a **vein,** and if it fills with (igneous or sedimentary) rock originating from elsewhere, it is a **dike.** A **joint** is a natural fracture in rock across which there is no measurable shear displacement. Because of the lack of shear involved in joint formation, joints can also be called **cracks** or **tensile fractures. Shear fractures,** in contrast, are mesoscopic fractures across which there has been displacement. Sometimes geologists use the term "shear fracture" instead of "fault" when they wish to imply that the amount of shear displacement on the fracture is relatively small, and that the shear displacement accompanied the formation of the fracture in once intact rock. In Chapter 7 we show several examples of fractures and veins in natural rocks.

In a broad sense, a **fault** is a surface or zone on which there has been measurable displacement. In a narrower sense, geologists restrict use of the term fault to a fracture surface on which there has been sliding. When using this narrow definition of fault, we apply the term **fault zone** to refer either to a band of finite width across which the displacement is partitioned among many smaller faults, or to the zone of rock bordering the fault that has fractured during faulting. Chapter 8 shows many examples of natural faults. We apply the term **shear zone** to a band of finite width in which the ductile shear strain is significantly greater than in the surrounding rock. Movement in shear zones can be the consequence of **cataclasis** (distributed fracturing, crushing, and frictional sliding of grains of rock or rock fragments), crystal-plastic deformation mechanisms (dislocation glide, dislocation climb), and diffusion. We'll mention cataclastic shear zones in this chapter, but other types of shear zones are discussed in Chapter 12, after we have introduced ductile deformation (Chapter 9).

Regardless of type, a fracture does not extend infinitely in all directions (Figure 6.2). Some fractures intersect the surface of a body of rock, whereas others terminate within the body. The line representing the intersection of the fracture with the surface of a rock body is the **fracture trace,** and the line that separates the region of the rock which has fractured from the

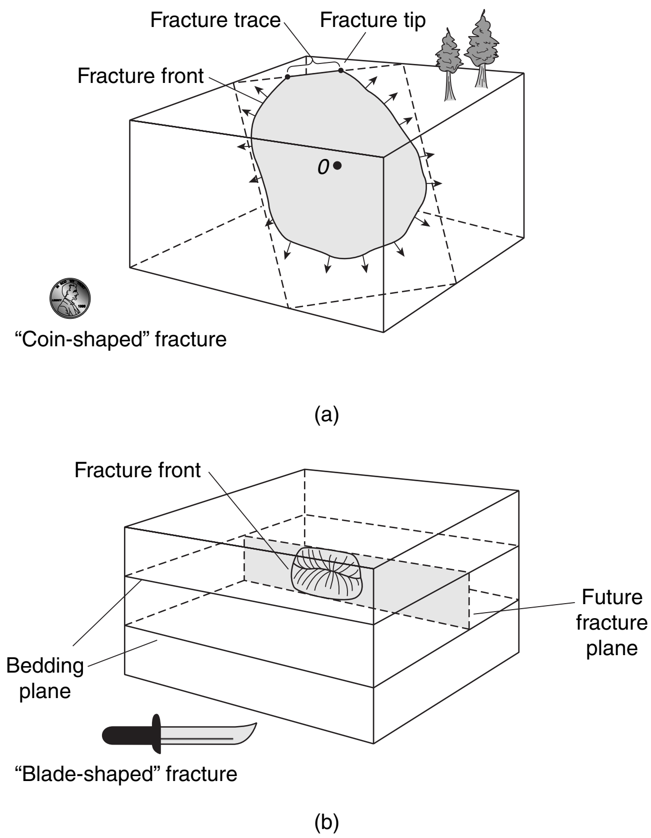


**FIGURE 6.2** (a) A block diagram illustrating that a fracture surface terminates within the limits of a rock body. The top surface of the block is the ground surface; erosion exposes the fracture trace. Note that the trace of the fracture on the ground surface is a line of finite length (with a fracture tip at each end). The arrows indicate that this particular fracture grew radially outward from an origin, labeled *O,* the dot in the center of the fracture plane. (b) A blade fracture which has propagated in a sedimentary layer and terminates at the bedding plane.

region that has not fractured is the **fracture front.** The point at which the fracture trace terminates on the surface of the rock is the **fracture tip.** In three dimensions, some fractures have irregular surfaces whereas others have geometries that roughly resemble coins or blades.

## 6.3 WHAT IS BRITTLE DEFORMATION?

To understand brittle deformation we need to look at the atomic structure of materials. Solids are composed of atoms or ions that are connected to one another by chemical bonds that can be thought of as tiny springs. Each chemical bond has an equilibrium length, and the angle between any two chemical bonds connected to the same atom has an equilibrium value (Figure 6.3a). During elastic strain, the bonds holding the atoms together within the solid stretch, shorten, and/or bend, but they do not break (Figure 6.3b)! When the stress is removed, the bonds return to their equilibrium conditions and the elastic strain disappears. In other words, elastic strain is fully *recoverable.* Recall from Chapter 5 that this elastic property of solids explains the propagation of earthquake waves though Earth.

Rocks cannot accumulate large elastic strains; you certainly cannot stretch a rock to twice its original length and expect it to spring back to its original shape!

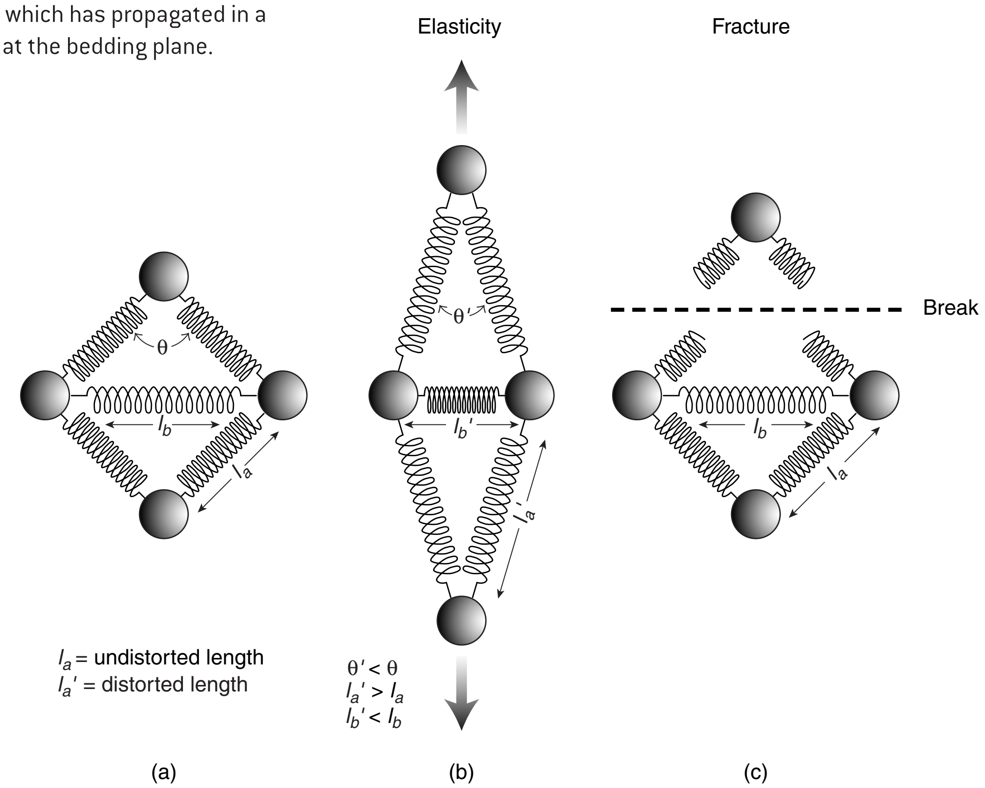


**FIGURE 6.3** A sketch illustrating what is meant by stretching and breaking of atomic bonds. The chemical bonds are represented by springs and the atoms by spheres. (a) Four atoms arranged in a lattice at equilibrium. (b) As a consequence of stretching of the lattice, some bonds stretch and some shorten, and the angle between pairs of bonds changes. (c) If the bonds are stretched too far, they break, and elastic strain is released.

| TABLE 6.2 | CATEGORIES OF BRITTLE DEFORMATION PROCESSES |
|---|---|
| Cataclastic flow | This type of brittle deformation refers to macroscopic ductile flow as a result of grain-scale fracturing and frictional sliding distributed over a band of finite width. |
| Frictional sliding | This process refers to the occurrence of sliding on a preexisting fracture surface, without the significant involvement of plastic deformation mechanisms. |
| Shear rupture | This type of brittle deformation results in the initiation of a macroscopic shear fracture at an acute angle to the maximum principal stress when a rock is subjected to a triaxial compressive stress. Shear rupturing involves growth and linkage of microcracks. |
| Tensile cracking | This type of brittle deformation involves propagation of cracks into previously unfractured material when a rock is subjected to a tensile stress. If the stress field is homogeneous, tensile cracks propagate in their own plane and are perpendicular to the least principal stress ($\sigma_3$). |

At most, a rock can develop a few percent strain by elastic distortion. If the stress applied to a rock is greater than the stress that the rock can accommodate elastically, then one of two changes can occur: the rock deforms in a ductile manner (strains without breaking), or the rock deforms in a brittle manner (that is, it breaks).

What actually happens during brittle deformation? Basically, if the stress becomes large enough to stretch or bend chemical bonds so much that the atoms are too far apart to attract one another, then the bonds break, resulting in either formation of a fracture (Figure 6.3c) or slip on a preexisting fracture. In contrast to elastic strain, brittle deformation is *nonrecoverable,* meaning that the distortion remains when the stress is removed. Again we will use earthquakes as an example; in this case, elastic strain is exceeded at the focus, resulting in failure and displacement. The pattern of breakage during brittle deformation depends on stress conditions and on material properties of the rock, so brittle deformation does not involve just one process. For purposes of our discussion, we divide brittle deformation processes into four categories that are listed in Table 6.2 and illustrated in Figure 6.4.

## 6.4 TENSILE CRACKING

### 6.4.1 Stress Concentration and Griffith Cracks

In Figure 6.5 we illustrate a crack in rock on the atomic scale. One way to create such a crack would be for all the chemical bonds across the crack surface to break at once. In this case, the tensile stress necessary for this to occur is equal to the strength of each chemical bond multiplied by all the bonds that had once crossed the area of the crack. If you know the strength of a single chemical bond, then you can calculate the stress necessary to break all the bonds simply by multiplying the bond strength by the number of bonds. Using realistic values for the elasticity (Young's modulus, $E$) and small strain (<10%), Equation 5.3 in Chapter 5 results in a theoretical strength of rock that is thousands of megapascals. Measurement of rock strength in the Earth's crust shows that tensile cracking occurs at crack-normal tensile stresses of less than about 10 MPa, when the confining pressure is low,[1] a value that is hundreds of times less than the theoretical strength of rock. Keeping the concept of theoretical strength in mind, we therefore face a paradox: How can natural rocks fracture at stresses that are so much lower than their theoretical strength?

The first step toward resolving the **strength paradox** came when engineers studying the theory of elasticity realized that the **remote stress** (stress due to a load applied at a distance from a region of interest) gets concentrated at the sides of flaws (e.g., holes) inside a material. For example, in the case of a circular hole in a vertical elastic sheet subjected to tensile stress at its ends (Figure 6.6a), the **local stress** (i.e., stress at the point of interest) tangent to the sides of the hole is three times the remote stress magnitude ($\sigma_r$). The magnitude of the local tangential stress at the top and bottom of the hole equals the magnitude of the remote stress, but is opposite in sign (i.e., it is compressive). If the hole has the shape of an ellipse instead

[1]The failure strength of rocks under tension is much less than the failure strength of rocks under compression; failure strength under compression depends on the confining pressure.

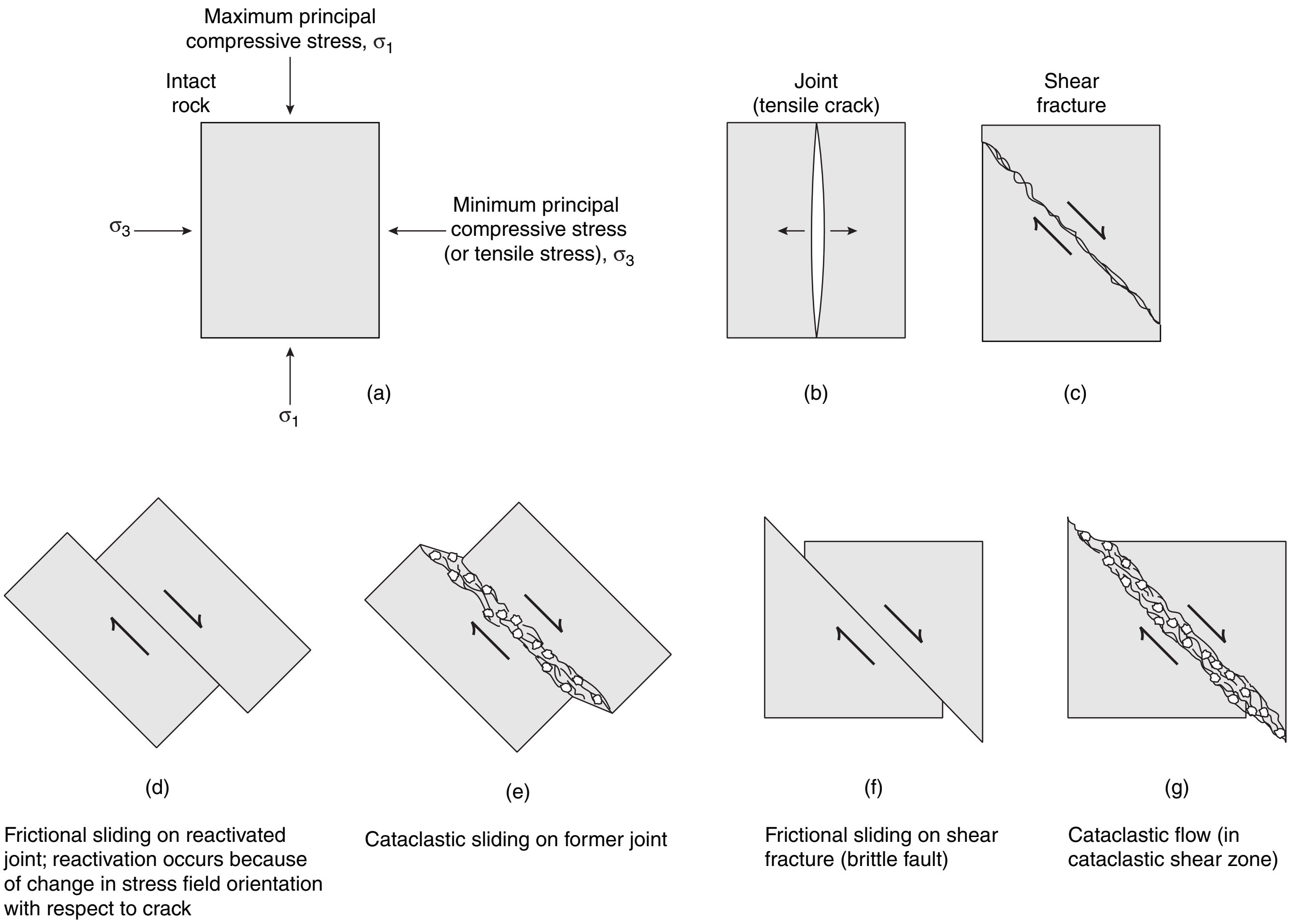


**FIGURE 6.4** Types of brittle deformation. (a) Orientation of the remote principal stress directions with respect to an intact rock body. (b) A tensile crack, forming parallel to $\sigma_1$ and perpendicular to $\sigma_3$ (which may be tensile). (c) A shear fracture, forming at an angle of about 30° to the $\sigma_1$ direction. (d) A tensile crack that has been reoriented with respect to the remote stresses and becomes a fault by undergoing frictional sliding. (e) A tensile crack which has been reactivated as a cataclastic shear zone. (f) A shear fracture that has evolved into a fault. (g) A shear fracture that has evolved into a cataclastic shear zone.

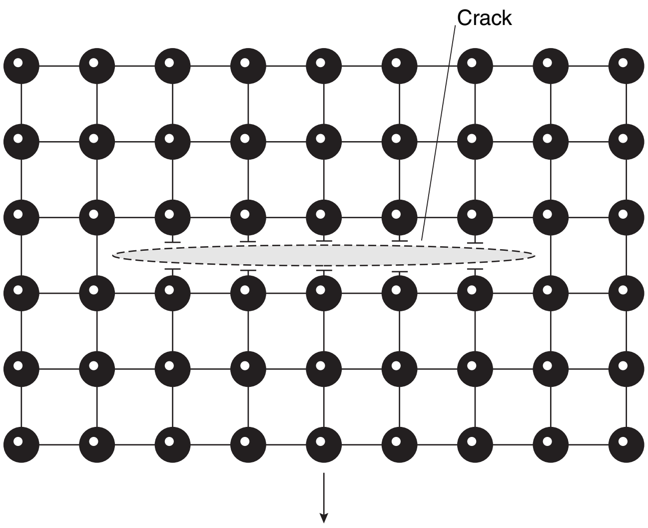


**FIGURE 6.5** A cross-sectional sketch of a crystal lattice (balls are atoms and sticks are bonds) in which there is a crack. The crack is a plane of finite extent across which all atomic bonds are broken.

of a circle (Figure 6.6b), the amount of stress concentration, $C$, is equal to $2b/a + 1$, where $a$ and $b$ are the short and long axes of the ellipse, respectively. Thus, values for stress concentration at the ends of an elliptical hole depend on the axial ratio of the hole: the larger the axial ratio, the greater the stress concentration. For example, at the ends of an elliptical hole with an axial ratio of 8:1, stress is concentrated by a factor of 17, and in a 1 µm × 0.02 µm crack the stress is magnified by a factor of ~100!

With this understanding in mind, A. W. Griffith, in the 1920s, took the next step toward resolving the strength paradox when he applied the concept of stress concentration at the ends of elliptical holes to fracture development. Griffith suggested that all materials contain preexisting microcracks or flaws at which stress concentrations naturally develop, and that because of the stress concentrations that develop at the tips of these cracks, they propagate and become larger cracks

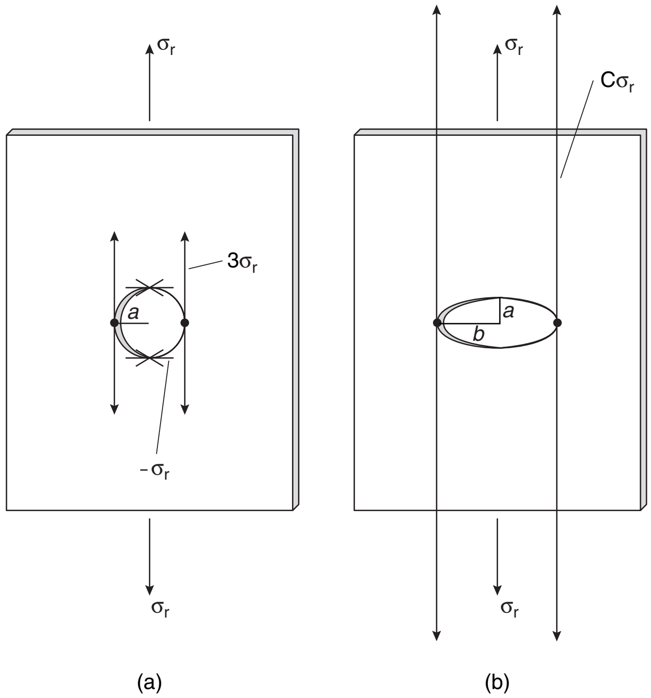


**FIGURE 6.6** Stress concentration adjacent to a hole in an elastic sheet. If the sheet is subjected to a remote tensile stress at its ends ($\sigma_r$), then stress magnitudes at the sides of the holes are equal to $C\sigma_r$, where $C$, the stress concentration factor, is $(2b/a) + 1$. (a) For a circular hole, $C = 3$. (b) For an elliptical hole, $C > 3$.

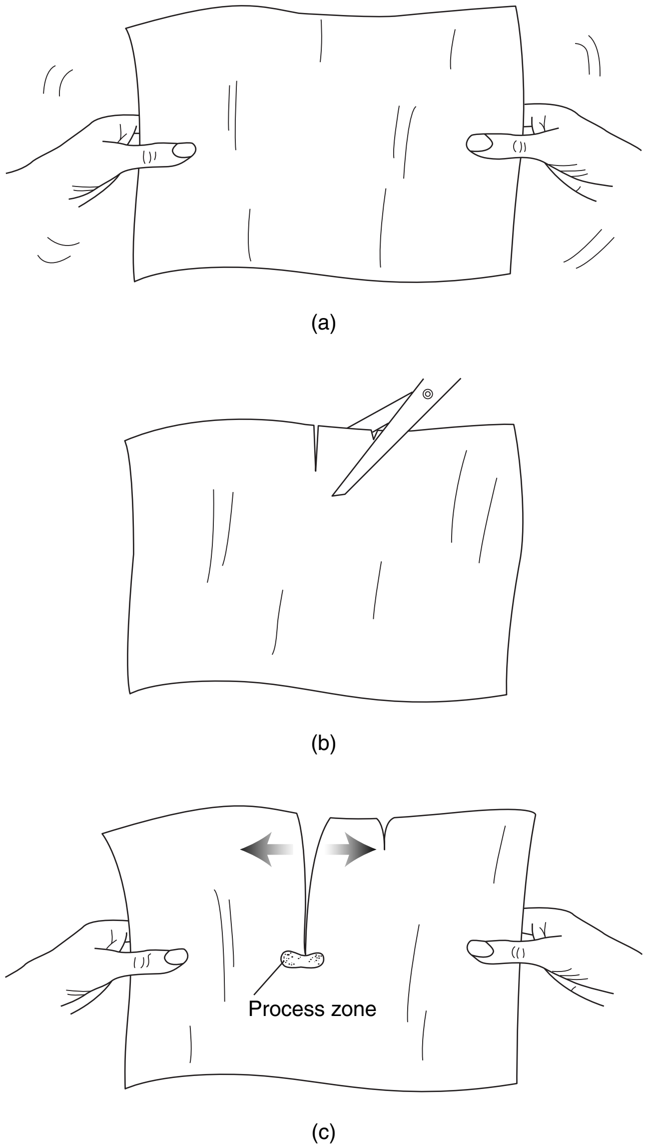


**FIGURE 6.7** Illustration of a home experiment to observe the importance of preexisting cracks in creating stress concentrations. (a) An intact piece of paper is difficult to pull apart. (b) Two cuts, a large one and a small one, are made in the paper. (c) The larger preexisting cut propagates. In the shaded area, a region called the process zone, the plastic strength of the material is exceeded and deforms.

even when the host rock is subjected to relatively low remote stresses. He discovered that in a material with cracks of different axial ratios, the crack with the largest axial ratio will most likely propagate first. In other words, stress at the tips of preexisting cracks can become sufficiently large to rupture the chemical bonds holding the minerals together at the tip and cause the crack to grow, even if the remote stress is relatively small. Preexisting microcracks and flaws in a rock, which include grain-scale fractures, pores, and grain boundaries, are now called **Griffith cracks** in his honor. Thus we resolve the strength paradox by learning that rocks in the crust are relatively weak because they contain Griffith cracks.

Griffith's concept provided useful insight into the nature of cracking, but his theory did not adequately show how factors such as crack shape, crack length, and crack orientation affect the cracking process. In subsequent years, engineers developed a new approach to studying the problem. In this approach, called **linear elastic fracture mechanics,** we assume that cracks in a material have nearly infinite axial ratio (defined as long axis/short axis), meaning that all cracks are very sharp. Linear elastic fracture mechanics theory predicts that, if factors like shape and orientation are equal, a longer crack will propagate before a shorter crack. We'll see why later in this chapter, when we discuss failure criteria.

We can examine how preexisting cracks affect the magnitude of stress necessary for tensile cracking in a simple experiment. Take a sheet of paper (Figure 6.7) and pull at both ends. You have to pull quite hard in order for the paper to tear. Now make two cuts, one that is ~0.5 cm long and one that is ~2 cm long, in the edge of the sheet near its center, and pull again. The pull that you apply gets concentrated at the tip of the preexisting cuts, and at this tip the strength of the paper is exceeded. You will find that it takes much less force to tear the paper, and that it tears apart by growth of the longer preexisting cut. The reason that sharp cracks do not propagate under extremely small stresses is that the tips of real cracks are blunted by a crack-tip **process zone,** in which the material deforms plastically (Figure 6.7c).

It is implicit in our description of crack propagation that the total area of a crack does not form instantaneously, but rather a crack initiates at a small flaw and then grows outward. If you have ever walked out on thin ice covering a pond, you are well aware of this fact. As you move away from shore, you suddenly hear a sound like the echo of a gunshot; this is the sound of a fracture forming in the ice due to the stress applied by your weight. If you have the presence of mind under such precarious circumstances to watch how the crack forms, you will notice that the crack initiates under your boot, and propagates outwards into intact ice at a finite velocity. This means that at any instant, only chemical bonds at the crack tip are breaking. In other words, not all the bonds cut by a fracture are broken at once, and thus the basis we used for calculating theoretical strength in the first place does not represent reality.

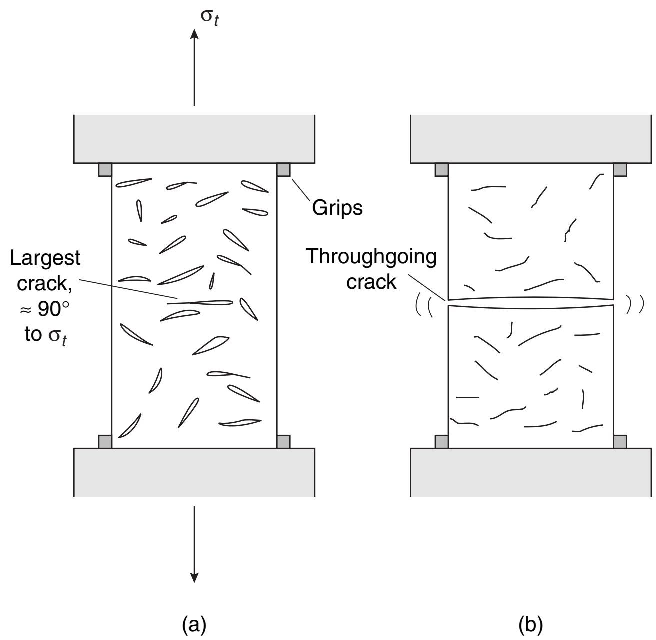


**FIGURE 6.8** Development of a throughgoing crack in a block under tension. (a) When tensile stress ($\sigma_t$) is applied, Griffith cracks open up. (b) The largest, properly oriented cracks propagate to form a throughgoing crack.

## 6.4.2 Exploring Tensile Crack Development

Let's consider what happens during a laboratory experiment in which we stretch a rock cylinder along its axis under a relatively low confining pressure (Figure 6.8a), a process called **axial stretching.** As soon as the remote tensile stress is applied, preexisting microcracks in the sample open slightly, and the remote stress is magnified to create larger local stresses at the crack tips. Eventually, the stress at the tip of a crack exceeds the strength of the rock and the crack begins to grow. If the remote tensile stress stays the same after the crack begins to propagate, then the crack continues to grow, and may eventually reach the sample's margins. When this happens, the sample fails, meaning it separates into two pieces that are no longer connected (Figure 6.8b).

We can also induce tensile fracturing by subjecting a rock cylinder to axial compression, under conditions of low confining pressure. Under such stress conditions, mesoscopic tensile fractures develop parallel to the cylinder axis (Figure 6.9a), a process known as **longitudinal splitting.** Longitudinal splitting is similar to tensile cracking except that, in uniaxial compression, the cracks that are not parallel to the $\sigma_1$ direction are closed, whereas cracks that are parallel to the compression direction can open up. To picture

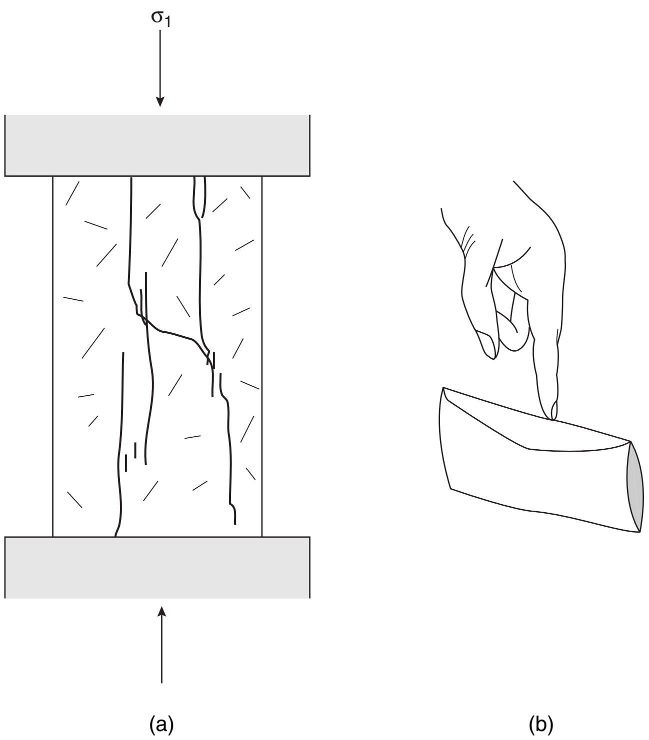


**FIGURE 6.9** (a) A cross section showing a rock cylinder with mesoscopic cracks formed by the process of longitudinal splitting. (b) An "envelope" model of longitudinal splitting. If you push down on the top of an envelope (whose ends have been cut off), the sides of the envelope will move apart.

this, imagine an envelope standing on its edge. If you push down on the top edge of the envelope, the sides of the envelope pull apart, even if they were not subjected to a remote tensile stress (Figure 6.9b). In rocks, as the compressive stress increases, the tensile stress at the tips of cracks exceeds the strength of the rock, and the crack propagates parallel to the compressive stress direction.

In the compressive stress environment illustrated in Figure 6.9, the confining pressure required is very small; but tensile cracks can also be generated in a rock cylinder when the remote stress is compressive under higher confining pressure when adding fluid pressure in pores and cracks of the sample (i.e., the pore pressure; Figure 6.10). The uniform, outward push of a fluid in a microcrack can have the effect of creating a local tensile stress at crack tips, and thus can cause a crack to propagate. We call this process **hydraulic fracturing.** As soon as the crack begins to grow, the volume of the crack increases, so if no additional fluid enters the crack, the fluid pressure decreases. Crack propagation ceases when pore pressure drops below the value necessary to create a sufficiently large tensile stress at the crack tip, and does not begin again until the pore pressure builds up sufficiently. Therefore, tensile cracking driven by an increase in pore pressure typically occurs in pulses. We'll return to the role of hydraulic fracturing later in this chapter and again in Chapter 8.

### 6.4.3 Modes of Crack-Surface Displacement

Before leaving the subject of Griffith cracks, we need to address one more critical issue, namely, the direction in which an individual crack grows when it is loaded. So far we have limited ourselves to cracks that are perpendicular to a remote stress. But what about cracks in other orientations with respect to stress, and how do they propagate? Materials scientists identify three configurations of crack loading. These configurations result in three different modes of crack-surface displacement (Figure 6.11). Note that the "displacement" we are referring to when describing crack propagation is only the infinitesimal movement initiating propagation of the crack tip and is not measurable mesoscopic displacement as in faults.

During **Mode I displacement,** a crack opens very slightly in the direction perpendicular to the crack surface, so Mode I cracks are tensile cracks. They form parallel to the principal plane of stress that is perpendicular to the $\sigma_3$ direction, and can grow in their plane without changing orientation. During **Mode II displacement** (the sliding mode), rock on one side of the crack surface moves very slightly in the direction parallel to the fracture surface and perpendicular to the fracture front. During **Mode III displacement** (the tearing mode), rock on one side of the crack slides very slightly parallel to the crack but in a direction parallel to the fracture front.

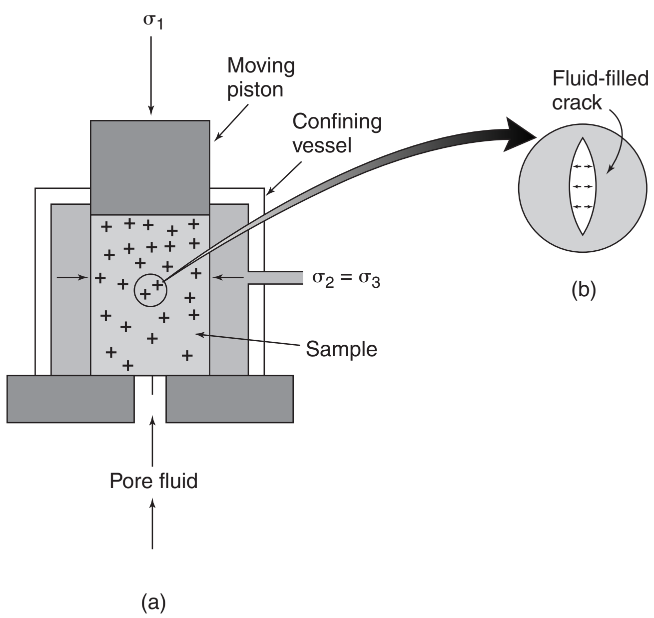


**FIGURE 6.10** (a) Cross-sectional sketch illustrating a rock cylinder in a triaxial loading experiment. Fluid has access to the rock cylinder and fills the cracks. (b) A fluid-filled crack that is being pushed apart from within by pore-fluid pressure.

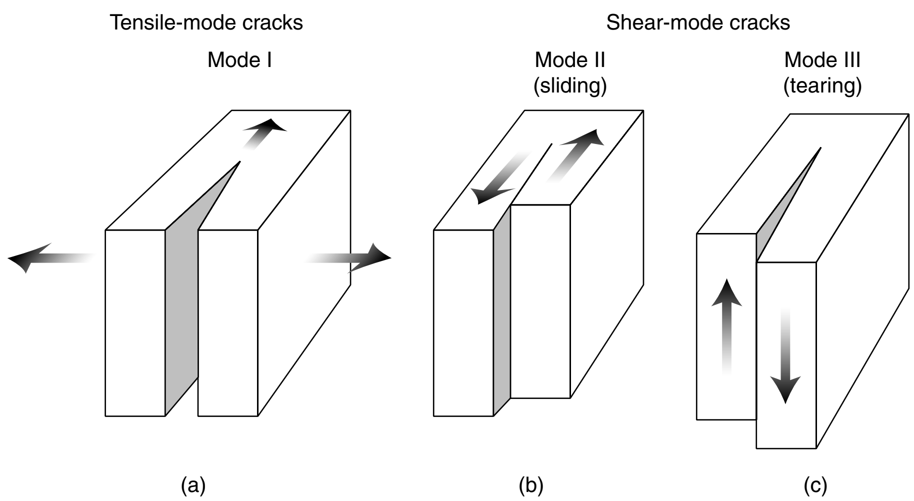


**FIGURE 6.11** Block diagrams illustrating the three modes of crack surface displacement: (a) Mode I, (b) Mode II, (c) Mode III. Mode I is a tensile crack, and Mode II and Mode III are shear-mode cracks.

Although shear-mode cracks appear similar to mesoscopic faults, Mode II (and Mode III) cracks are not simply microscopic equivalents of faults. We realize the difference when we examine the propagation of shear-mode cracks. As they start growing, shear-mode cracks immediately curve into the orientation of tensile or Mode I cracks, meaning that shear-mode cracks do *not* grow in their plane. Propagating shear-mode cracks spawn new tensile cracks called **wing cracks,** a process illustrated in Figure 6.12.

## 6.5 PROCESSES OF BRITTLE FAULTING

A *brittle fault* is a surface on which measurable slip has developed without much contribution by plastic deformation mechanisms. Brittle faulting happens in response to the application of a differential stress, because slip occurs in response to a shear stress parallel to the fault plane. In other words, for faulting to take place, $\sigma_1$ cannot equal $\sigma_3$, and the fault surface cannot parallel a principal plane of stress. Faulting causes a change in shape of the overall rock body that contains the fault. Hence, faulting contributes to development of regional strain; however, because a brittle fault is, by definition, a discontinuity in a rock body, the occurrence of faulting does not require development of measurable ductile strain in the surrounding rock.

There are two basic ways to create a brittle fault (Figure 6.4). The first is by shear rupturing of a previously intact body of rock. The second is by shear reactivation of a previously formed weak surface (for example, a joint, a bedding surface, or a preexisting fault) in a body of rock. A preexisting weak surface may slip before the differential stress magnitude reaches the failure strength for shear rupture of intact rock. Once formed, movement on brittle faults takes place either by frictional sliding, by the growth of fault-parallel veins, or by cataclastic flow. We'll focus mostly on fault formation and friction, the central processes of brittle faulting, and only briefly mention the other processes here.

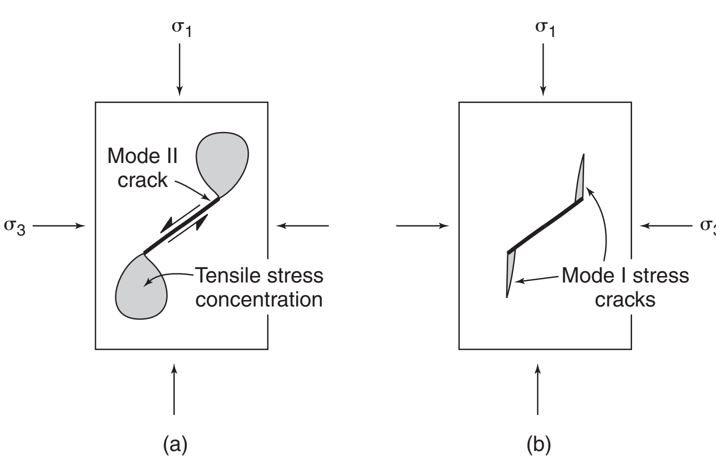


**FIGURE 6.12** Propagating shear-mode crack and the formation of wing cracks. (a) A tensile stress concentration occurs at the ends of a Mode II crack that is being loaded. (b) Mode I wing cracks form in the zones of tensile-crack concentration.

### 6.5.1 Slip by Growth of Fault-Parallel Veins

Not all faults that undergo displacement in the brittle field move by frictional sliding. On some faults, the opposing surfaces are separated from one another by mineral crystals (e.g., quartz or calcite) precipitated out of water solutions present along the fault during movement (Figure 6.13). Such **fault-surface veins** may be composed of mineral fibers (needle-like crystals), of blocky crystals, or of both.

The process by which fault-surface veins form is not well understood. In some cases, they may form when high fluid pressures cause a crack to develop along a weak fault surface. Immediately after cracking, one side of the fault moves slightly with respect to the other; the crack then seals by precipitation of vein material in the crack (see also Chapter 7). In other cases, vein formation may reflect gradual dissolution of steps (or asperities) on the fault surface, followed by the transfer of ions through fluid films to sites of lower stress, where mineral precipitation takes place. This second process may occur without the formation of an actual discontinuity, across which the rock loses cohesion. Whether syn-slip veining or frictional sliding takes place on a fault surface probably depends on strain rate and on the presence of water. Fault-surface veining is probably more common when water is present along the fault, and when movement occurs slowly.

### 6.5.2 Cataclasis and Cataclastic Flow

Cataclasis refers to movement on a fault by a combination of microcracking, frictional sliding of fragments past one another, and rotation and transport of grains. To picture the process of cataclasis, imagine what happens to corn passing between two old-fashioned mill stones. The millstones slide past one another and, in the process, transform the corn into cornmeal. Cataclasis, if affecting a relatively broad band of rock, results in mesoscopic ductile strain, in which case it is also called **cataclastic flow** (Chapter 9),

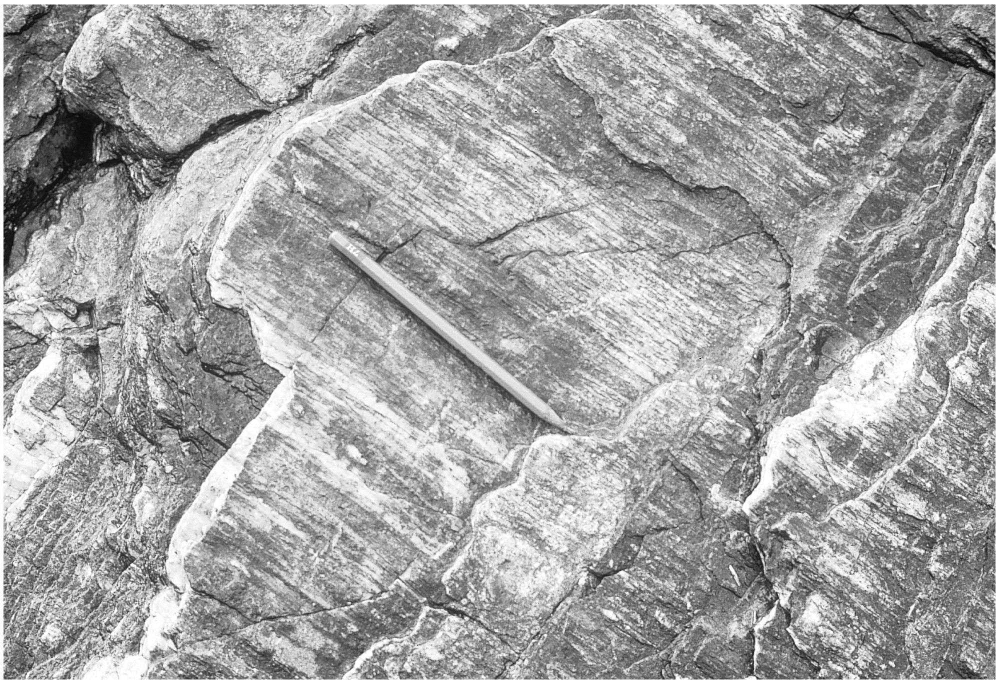

**FIGURE 6.13** Photograph of stepped calcite slip fibers on a fault surface; pencil indicates displacement direction.

because the rock over the width of the band effectively flows. To picture cataclastic flow, think of how the cornmeal that we just produced, when poured from one container to another, behaves much like a fluid, even though the individual grains certainly are solid. Movement on a fault that involves development of a zone in which cataclastic flow occurs is often referred to as a **cataclastic shear zone.**

## 6.6 FORMATION OF SHEAR FRACTURES

Shear fractures differ markedly from tensile cracks. A shear fracture is a surface across which a rock loses continuity when the shear stress parallel to the surface is sufficiently large. Shear fractures are initiated in laboratory rock cylinders at a typical angle of about 30° to $\sigma_1$ under conditions of confining pressure ($\sigma_1 > \sigma_2 = \sigma_3$). Because there is a component of normal stress acting on the fracture in addition to shear stress, friction resists sliding on the fracture during its formation. If the shear stress acting on the fracture continues to exceed the frictional resistance to sliding, the fracture grows and displacement accumulates. Shear fractures (or faults) are therefore not simply large shear-mode cracks, because, as we have seen, shear-mode cracks cannot grow in their own plane. This conceptual difference is very important.

So how do shear fractures form? We can gain insight into the process of shear-fracture formation by generating shear ruptures during a laboratory triaxial loading experiment, using a rock cylinder under confining pressure. So to begin our search for an answer to this question, we first describe such an experiment.

In a confined-compression triaxial-loading experiment, we take a cylinder of rock, jacket it in copper or rubber, surround it with a confining fluid in a pressure chamber, and squeeze it between two hydraulic pistons. In the experiment shown in Figure 6.14, the rock itself stays dry. During the experiment, we apply a confining pressure ($\sigma_2 = \sigma_3$) to the sides of the cylinder by increasing the pressure in the surrounding fluid, and an axial load ($\sigma_1$) to the ends of the cylinder by moving the pistons together at a constant rate. By keeping the value of $\sigma_3$ constant while $\sigma_1$ gradually increases, we increase the differential stress ($\sigma_d = \sigma_1 - \sigma_3$). In this experiment we measure the magnitude of $\sigma_d$, the change in length of the cylinder (which is the axial strain, $\mathbf{e}_a$), and the change in volume ($\Delta$) of the cylinder.

A graph of $\sigma_d$ versus $\mathbf{e}_a$ (Figure 6.14a) shows that the experiment has four stages. In Stage I, we find that as $\sigma_d$ increases, $\mathbf{e}_a$ also increases and that the relationship between these two quantities is a concave-up curve. In Stage II of the experiment, the relationship between $\sigma_d$ and $\mathbf{e}_a$ is a straight line with a positive slope. During Stage I and most of Stage II, the volume of the sample decreases slightly. In Stage III of the experiment, the slope of the line showing the relation between $\sigma_d$ and $\mathbf{e}_a$ decreases. The stress at which the curve changes slope is called the **yield strength.** During the latter part of Stage II and all of Stage III, we observe a slight increase in volume, a phenomenon known as **dilatancy,** and if we had a very sensitive microphone attached to the sample during this time, we would hear lots of popping sounds that reflect the formation and growth of microcracks. Suddenly, when $\sigma_d$ equals the failure stress ($\sigma_f$), a shear rupture surface develops at an angle of about 30° to the cylinder axis,

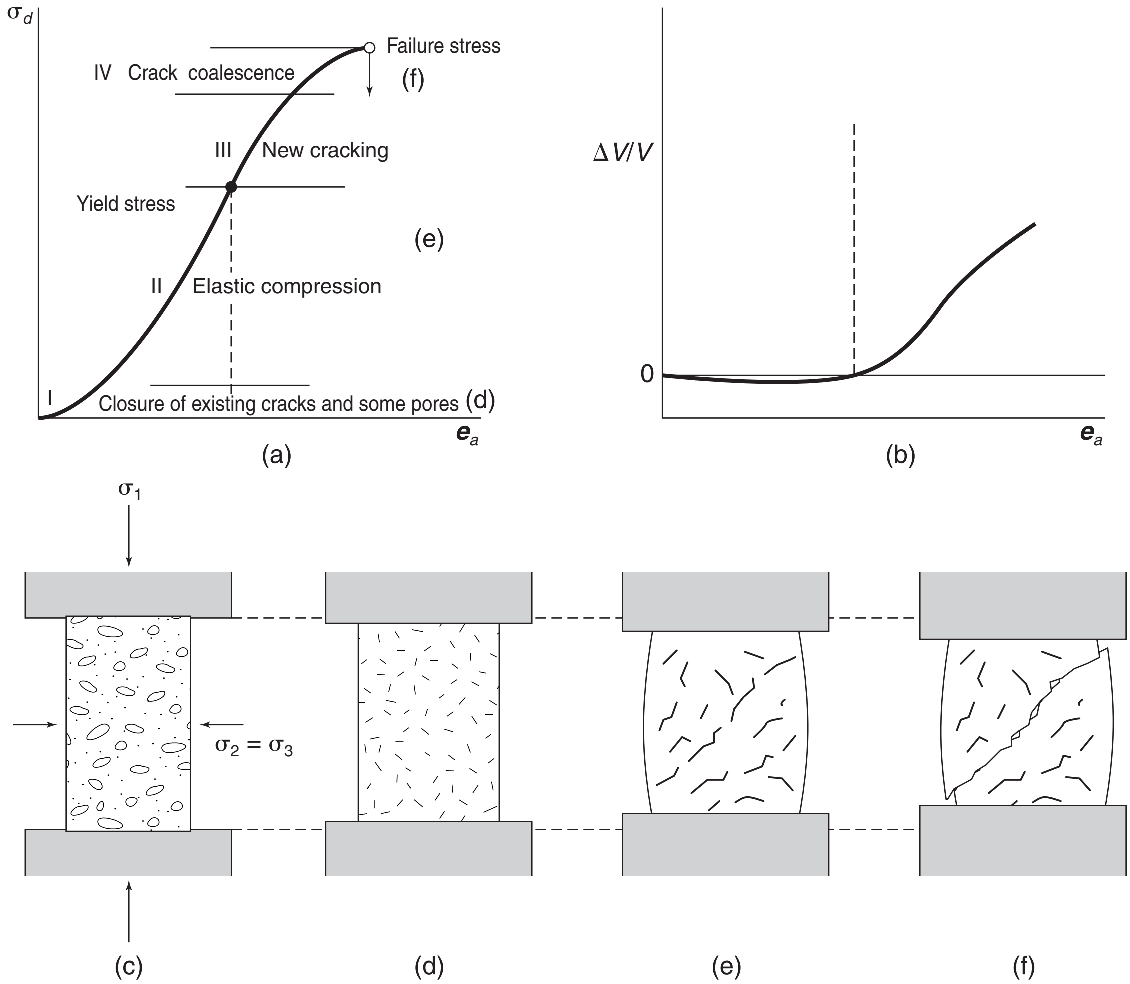


**FIGURE 6.14** Fracture formation. (a) Stress–strain plot (differential stress versus axial shortening) showing the stages (I–IV) in a confined compression experiment. The labels indicate the process that accounts for the slope of the curve. (b) The changes in volume accompanying the axial shortening illustrate the phenomenon of dilatancy; left of the dashed line, the sample volume decreases, whereas to the right of the dashed line the sample volume increases. (c–f) Schematic cross sections showing the behavior of rock cylinders during the successive stages of a confined compression experiment and accompanying stress–strain plot, emphasizing the behavior of Griffith cracks (cracks shown are much larger than real dimensions). (c) Pre-deformation state, showing open Griffith cracks. (d) Compression begins and volume decreases due to crack closure. (e) Crack propagation and dilatancy (volume increase). (f) Merging of cracks along the future throughgoing shear fracture, followed by loss of cohesion of the sample (mesoscopic failure).

and there is a stress drop. A stress drop in this context means that the axial stress supported by the specimen suddenly decreases and large strain accumulates at a lower stress. To picture this stress drop, imagine that you're pushing a car that is stuck in a ditch. You have to push hard until the tires come out of the ditch, at which time you have to stop pushing so hard, or you will fall down as the car rolls away. The value of $\sigma_d$ at the instant that the shear rupture forms and the stress drops is called the **failure strength for shear rupture.** Once failure has occurred, the sample is no longer intact and frictional resistance to sliding on the fracture surface determines its further behavior.

What physically happened during this experiment? During Stage I, preexisting open microcracks underwent closure. During Stage II, the sample underwent elastic shortening parallel to the axis, and because of the Poisson effect expanded slightly in the direction perpendicular to the axis (Figure 6.14d). The Poisson effect refers to the phenomenon in which a rock that is undergoing elastic shortening in one direction extends in the direction at right angles to the shortening direction. The ratio between the amount of shortening and the amount of extension is called Poisson's ratio, $\nu$.[2] At the start of Stage III, tensile microcracks begin to grow throughout the sample, and wing cracks grow at the tips of shear-mode cracks. The initiation and growth of these cracks causes the observed slight increase in volume, and accounts for the popping noises (Figure 6.14e). During Stage III, the tensile cracking intensifies along a narrow band that cuts across the sample at an angle of about 30° to the axial stress. Failure occurs in Stage IV when the cracks self-organize to form a throughgoing surface along which the sample loses continuity, so that the rock on one side can frictionally slide relative to the rock on the other side (Figure 6.14f). As a consequence, the cylinders move together more easily and stress abruptly drops.

The fracture development scenario described above shows that the failure strength for shear fracture is not a definition of the stress state at which a single crack propagates, but rather it is the stress state at which a multitude of small cracks coalesce to form a throughgoing rupture. Also note that two ruptures form in some experiments, both at ~30° to the axial stress. The angle between these **conjugate fractures** is ~60°, and the acute bisector is parallel to the maximum principal stress. With continued displacement, however, it is impossible for both fractures to remain active, because displacement on one fracture will offset the other. Thus, typically only one fracture will evolve into a throughgoing fault (see next section).

## 6.7 PREDICTING INITIATION OF BRITTLE DEFORMATION

We have seen that brittle structures develop when rock is subjected to stress, but so far we have been rather vague about defining the stress states in which brittle deformation occurs. Clearly, an understanding of the stress state at which brittle deformation begins is valuable, not only to geologists, who want to know when, where, and why brittle geologic structures (joints, faults, veins, and dikes) develop in the Earth, but also to engineers, who must be able to estimate the magnitude of stress that a building or bridge can sustain before it collapses. When we discuss brittle deformation, we are really talking about three phenomena: tensile crack growth, shear fracture development, and frictional sliding. In this section, we examine the stress conditions necessary for each of these phenomena to occur.

### 6.7.1 Tensile Cracking Criteria

A **tensile cracking criterion** is a mathematical statement that predicts the stress state in which a crack begins to propagate. All tensile cracking criteria are based on the assumption that macroscopic cracks grow from preexisting flaws (Griffith cracks) in the rock, because preexisting flaws cause stress concentrations. Griffith was one of the first researchers to propose a tensile cracking criterion. He did so by looking at how energy was utilized during cracking. Griffith envisioned that a material in which a crack forms can be modeled as a thermodynamic system consisting of an elastic plate containing a preexisting elliptical crack. If a load is applied to the ends of the sheet so that it stretches and the crack propagates, the total energy of this system can be defined by the following equation (known as the **Griffith energy balance**):[3]

$$dU_T = dU_s - dW_r + dU_E \qquad \text{Eq. 6.1}$$

where $dU_T$ is the change in total energy of the system, $dU_s$ is the change in surface energy due to growth of the crack (this term arises because crack formation breaks bonds, and energy stored in a broken bond is greater than the energy stored in a satisfied bond), $dW_r$ is the work done by the load in deforming the plate, and $dU_E$ is the change in the strain energy stored in the plate (strain energy is the energy stored by chemical bonds that have been stretched out of their equilibrium length or angle). Griffith pointed out that because the system starts with the load already in place, formation of the crack does not change the total energy of the system. Thus, the change in total energy for an increment of crack growth ($dc$) equals 0 (in equation form, $dU_T/dc = 0$ for an equilibrium condition). With this point in mind, Equation 6.1 can be rewritten as

$$dW_r = dU_s + dU_E \qquad \text{Eq. 6.2}$$

In words, Equation 6.2 means that the work done on a system is divided between creation of new elastic strain energy in the sheet and creation of new crack surface by propagation of the crack. Using this con-

[2]Since Poisson's ratio has the units length/length, it is dimensionless; a typical value of $\nu$ for rocks is 0.25.

[3]This is a plane stress criterion; for plane strain conditions, $E$ is replaced by $E/(1 - \nu)^2$, where $\nu$ is Poisson's ratio.

cept, along with several theorems from elasticity theory, Griffith devised a tensile cracking criterion, which we present without derivation:

$$\sigma_t = (2E\gamma/\pi c)^{1/2} \qquad \text{Eq. 6.3}$$

where $\sigma_t$ = critical remote tensile stress (tensile stress at which the weakest Griffith crack begins to grow), $E$ = Young's modulus, $\gamma$ = energy used to create new crack surface, and $c$ = half-length of the preexisting crack. Reading this equation, we see that the critical remote tensile stress for a rock sample is proportional to material properties of the sample and the length of the crack. Note that as crack length increases, the value of critical remote tensile stress ($\sigma_t$) decreases.

Subsequently, researchers have utilized concepts from the engineering study of linear elastic fracture mechanics to develop tensile cracking criteria. According to this work, the following equation defines conditions at which Mode I cracks propagate:

$$K_{\mathrm{I}} = \sigma_t\, Y(\pi c)^{1/2} \qquad \text{Eq. 6.4}$$

where $K_{\mathrm{I}}$ (pronounced "kay one," where the "one" represents a Mode I crack) is the **stress intensity factor,** $\sigma_t$ is the remote tensile stress, $Y$ is a dimensionless number that takes into account the geometry of the crack (e.g., whether it is penny-shaped, blade-shaped, or tunnel-shaped), and $c$ is half of the crack's length. In this analysis, all cracks in the body are assumed to have very large ellipticity; that is, cracks are assumed to be very sharp at their tips. Note that cracks with smaller ellipticity require higher stresses to propagate.

Equation 6.4 says that the value of $K_{\mathrm{I}}$ increases when $\sigma$ increases. A crack in the sample begins to grow when $K_{\mathrm{I}}$ attains a value of $K_{\mathrm{I}c}$, which is the **critical stress intensity factor** or the **fracture toughness** (a measure of tensile strength). The fracture toughness is constant for a given material. When $K_{\mathrm{I}}$ reaches $K_{\mathrm{I}c}$, the value of $\sigma$ reaches $\sigma_t$, where $\sigma_t$ is the **critical remote tensile stress** at the instant the crack starts to grow. We can rewrite Equation 6.4 to create an equation that more directly defines the value of $\sigma_t$ at the instant the crack grows:

$$\sigma_t = K_{\mathrm{I}c}/[Y(\pi c)^{1/2}] \qquad \text{Eq. 6.5}$$

Note that, in this equation, the remote stress necessary for cracking depends on the fracture toughness, the crack shape, and the length of the crack. If other factors are equal, a longer crack generally propagates before a shorter crack. Similarly, if other factors are equal, crack shape determines which crack propagates first. Because $c$ increases as the crack grows, crack propagation typically leads to sample failure. This relationship has another interesting consequence for natural rocks. Because crack length depends on the grain size of a sample, fine-grained rocks should be stronger than coarse-grained rocks.

Equations like Equations 6.4 and 6.5 can also be written for Mode II and Mode III cracks. By comparing equations for the three different modes of cracking, you will find that, other factors being equal, a Mode I crack (i.e., a crack perpendicular to $\sigma_3$) propagates before a Mode II or Mode III crack. However, since other factors like crack shape and length come into play, Mode II or Mode III cracks sometimes propagate before Mode I cracks in a real material. Remember that the instant they propagate, Mode II and III cracks either bend and become Mode I cracks, or they develop wing cracks at their tips; shear cracks cannot propagate significantly in their own plane.

In summary, we see that the stress necessary to initiate the propagation of a crack depends on the ellipticity, the length, the shape, and the orientation of a preexisting crack. Study of crack-propagation criteria is a very active research area, and further details concerning this complex subject (such as subcritical crack growth under long-term loading or corrosive conditions) are beyond the scope of this book (see the reading list).

### 6.7.2 Shear-Fracture Criteria and Failure Envelopes

A **shear-fracture criterion** is an expression that describes the stress state at which a shear rupture forms and separates a sample into two pieces. Because shear-fracture initiation in a laboratory sample inevitably leads to failure of the sample, meaning that after rupture the sample can no longer support a load that exceeds the frictional resistance to sliding on the fracture surface, shear-rupture criteria are also commonly known as **shear-failure criteria.**

Charles Coulomb[4] was one of the first to propose a shear-fracture criterion. He suggested that if all the principal stresses are compressive, as is the case in a confined compression experiment, a material fails by the formation of a shear fracture, and that the shear stress parallel to the fracture surface, at the instant of failure, is related to the normal stress by the equation

$$\sigma_s = C + \mu\sigma_n \qquad \text{Eq. 6.6}$$

[4]An eighteenth-century French naturalist.

where $\sigma_s$ is the shear stress parallel to the fracture surface at failure; $C$ is the **cohesion** of the rock, a constant that specifies the shear stress necessary to cause failure if the normal stress across the potential fracture plane equals zero (note that this $C$ is not the same as the $c$ in Equations 6.3–6.5); $\sigma_n$ is the normal stress across the shear fracture at the instant of failure; and $\mu$ is a constant traditionally known as the **coefficient of internal friction.** The name for $\mu$ originally came from studies of friction between grains in unconsolidated sand and of the control that such friction has on slope angles of sand piles, so the name is essentially meaningless in the context of shear failure of a solid rock; $\mu$ should be viewed simply as a constant of proportionality. Equation 6.6, also known as **Coulomb's failure criterion,** basically states that the shear stress necessary to initiate a shear fracture is proportional to the normal stress across the fracture surface.

The Coulomb criterion plots as a straight line on a Mohr diagram[5] (Figure 6.15). To see this, let's plot the results of four triaxial loading experiments in which we increase the axial load on a confined granite cylinder until it ruptures. In the first experiment, we set the confining pressure ($\sigma_2 = \sigma_3$) at a relatively low value, increase the axial load ($\sigma_1$) until the sample fails, and then plot the Mohr circle representing this **critical stress state,** meaning the stress state at the instant of failure, on the Mohr diagram. When we repeat the experiment, using a new cylinder, and starting at a higher confining pressure, we find that as $\sigma_3$ increases, the differential stress ($\sigma_1 - \sigma_3$) at the instant of failure also increases. Thus, the Mohr circle representing the second experiment has a larger diameter and lies to the right of the first circle. When we repeat the experiment two more times and plot the four circles on the diagram, we find that they are all tangent to a straight line with a slope of $\mu$ (i.e., tan $\phi$) and a $y$-intercept of $C$, and this straight line is the Coulomb criterion. Note that we can also draw a straight line representing the criterion in the region of the Mohr diagram below the $\sigma_n$-axis.

A line drawn from the center of a Mohr circle to the point of its tangency with the Coulomb criterion defines $2\theta$, where $\theta$ is the angle between the $\sigma_3$ direction and the plane of shear fracture (typically about 30°). Because the Coulomb criterion is a straight line, this angle is constant for the range of confining pressures for which the criterion is valid. The reason for the 30° angle becomes evident in a graph plotting normal stress magnitude and shear stress magnitude as a function of the angle between the plane and the $\sigma_1$ direction (Figure 6.16). Notice that the minimum normal stress does not occur in the same plane as the maximum shear stress. Shear stress is at its highest on a potential failure plane oriented at 45° to $\sigma_1$, but the normal stress across this potential plane is still too large to permit shear fracturing in planes of this orientation. The shear stress is a bit lower across a plane oriented at 30° to $\sigma_1$, but is still fairly high. However, the normal stress across the 30° plane is substantially lower, favoring shear-fracture formation.

Coulomb's criterion is an empirical relation, meaning that it is based on experimental observation alone, not on theoretical principles or knowledge of atomic-scale or crystal-scale mechanisms. This failure criterion does not relate the stress state at failure to physical parameters, as does the Griffith criterion, nor does it define the state of stress in which the microcracks, which eventually coalesce to form the shear rupture, begin to propagate. The Coulomb criterion does not predict whether the fractures that form will dip to the right or to the left with respect to the axis of the rock cylinder in a triaxial loading experiment. In fact, as

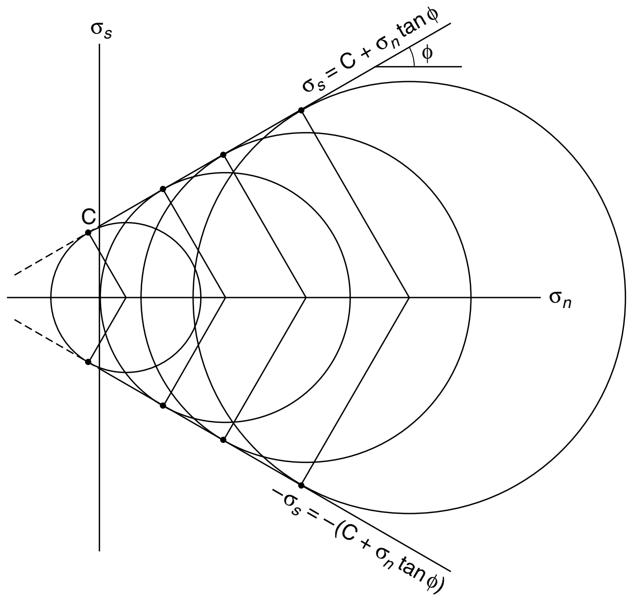


**FIGURE 6.15** Mohr diagram showing a Coulomb failure envelope based on a set of experiments with increasing differential stress. The circles represent differential stress states at the instant of shear failure. The envelope is represented by two straight lines, on which the dots represent failure planes.

[5]Recall that on the Mohr diagram, normal stresses ($\sigma_n$) are plotted on the $x$-axis and shear stresses ($\sigma_s$) are plotted on the $y$-axis. A Mohr circle represents the stress state by indicating the values of $\sigma_n$ and $\sigma_s$ acting on a plane oriented at $\theta$° to the $\sigma_3$ direction. The circle intersects the $x$-axis at $\sigma_1$ and $\sigma_3$ (both of which are normal stresses, because they are principal stresses), and the angle between the $x$-axis and a radius from the center of the circle to a point on the circle defines the angle $2\theta$.

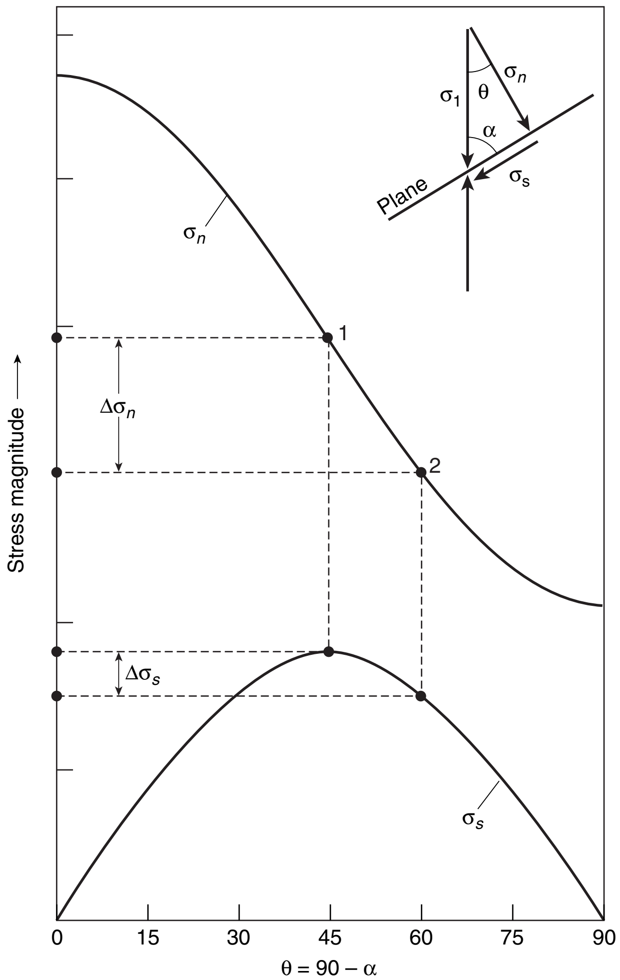


**FIGURE 6.16** The change in magnitudes of the normal and shear components of stress acting on a plane as a function of the angle $\alpha$ between the plane and the $\sigma_1$ direction; the angle $\theta = 90 - \alpha$ is plotted for comparison with other diagrams. At point 1 ($\alpha = \theta = 45°$), shear stress is a maximum, but the normal stress across the plane is quite large. At point 2 ($\theta = 60°$, $\alpha = 30°$), the shear stress is still quite high, but the normal stress is much lower.

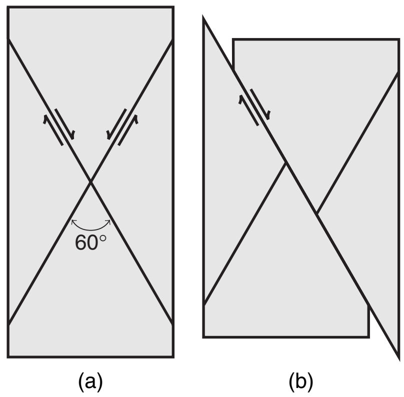


**FIGURE 6.17** Cross-sectional sketch showing how only one of a pair of conjugate shear fractures (a) evolves into a fault with measurable displacement (b).

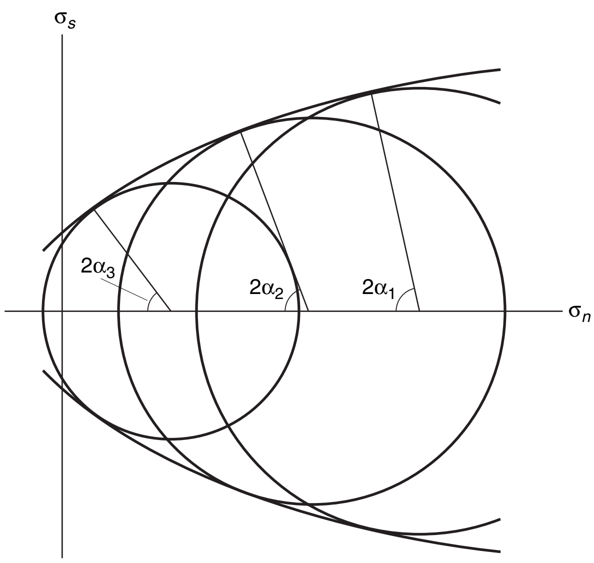


**FIGURE 6.18** Mohr failure envelope. Note that the slope of the envelope steepens toward the $\sigma_s$-axis. Therefore, the value of $\alpha$ (the angle between fault and $\sigma_1$) is not constant (compare $2\alpha_1$, $2\alpha_2$, and $2\alpha_3$).

mentioned earlier, **conjugate shear fractures,** one with a right-lateral shear sense and one with a left-lateral shear sense, may develop (Figure 6.17). The two fractures, typically separated by an angle of ~60°, correspond to the tangency points of the circle representing the stress state at failure with the Coulomb failure envelope.

The German engineer Otto Mohr conducted further studies of shear-fracture criteria and found that Coulomb's straight-line relationship only works for a limited range of confining pressures. He noted that at lower confining pressure, the line representing the stress state at failure curved with a steeper slope, and that at higher confining pressure, the line curved with a shallower slope (Figure 6.18). Mohr concluded that over a range of confining pressure, the failure criteria for shear rupture resembles a portion of a parabola lying on its side, and this curve represents the **Mohr-Coulomb criterion** for shear fracturing. Notice that this criterion is also empirical. Unlike Coulomb's straight-line relation, the change in slope of the Mohr-Coulomb failure envelope indicates that the angle between the shear fracture plane and $\sigma_1$ actually does depend on the stress state. At lower confining pressures, the angle is smaller, and at high confining pressures, the angle is steeper.

The Mohr-Coulomb criterion (both for positive and negative values of $\sigma_s$) defines a failure envelope on the Mohr diagram. A **failure envelope** separates the field on the diagram in which stress states are "stable" from

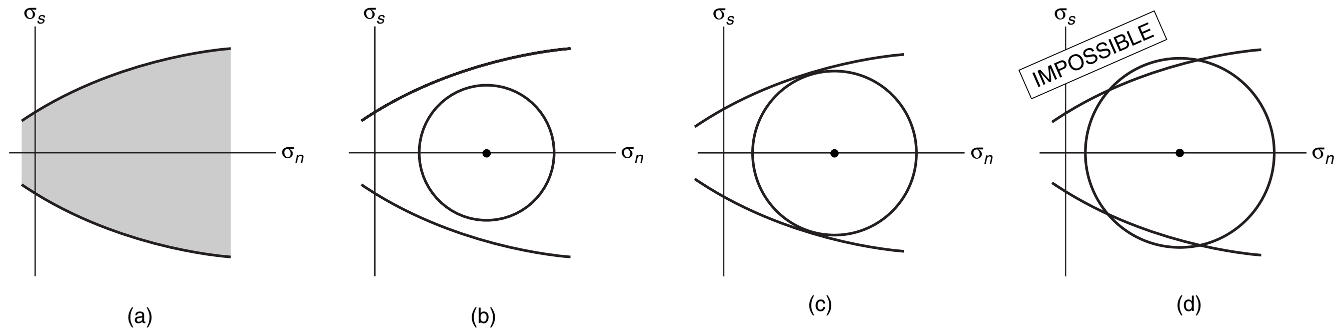


**FIGURE 6.19** (a) A brittle failure envelope as depicted on a Mohr diagram. Within the envelope (shaded area), stress states are stable, but outside the envelope, stress states are unstable. (b) A stress state that is stable, because the Mohr circle, which passes through values for $\sigma_1$ and $\sigma_3$ and defines the stress state, falls entirely inside the envelope. (c) A stress state at the instant of failure. The Mohr circle touches the envelope. (d) A stress state that is impossible.

the field in which stress states are "unstable" (Figure 6.19). By this definition, a **stable stress state** is one that a sample can withstand without undergoing brittle failure. An **unstable stress state** is an impossible condition to achieve, for the sample will have failed by fracturing before such a stress state is reached (Figure 6.19). In other words, a stress state represented by a Mohr circle that lies entirely within the envelope is stable, and will not cause the sample to develop a shear rupture. A circle that is tangent to the envelope specifies the stress state at which brittle failure occurs. Stress states defined by circles that extend beyond the envelope are unstable, and are therefore impossible within the particular rock being studied.

Can we define a failure envelope representing the critical stress at failure for very high confining pressures, very low confining pressures, or for conditions where one of the principal stresses is tensile? The answer to this question is controversial. We'll look at each of these conditions separately.

At high confining pressures, samples may begin to deform plastically. Under such conditions, we are no longer really talking about brittle deformation, so the concept of a "failure" envelope no longer really applies. However, we can approximately represent the "yield" envelope, meaning the stress state at which the sample begins to yield plastically, on a Mohr diagram by a pair of lines that parallel the $\sigma_n$-axis (Figure 6.20a). This yield criterion, known as **Von Mises criterion,** indicates that plastic yielding is effectively independent of the differential stress, once the yield stress has been achieved.

If the tensile stress is large enough, the sample fails by developing a throughgoing tensile crack. The tensile stress necessary to induce tensile failure may be represented by a point, $T_0$, the **tensile strength,** along the $\sigma_n$-axis to the left of the $\sigma_s$-axis (Figure 6.20b). As we have seen, however, the position of this point depends on the size of the flaws in the sample. Thus, even for the same rock type, experiments show that the tensile strength is very variable and that it is best represented by a range of points along the $\sigma_n$-axis.

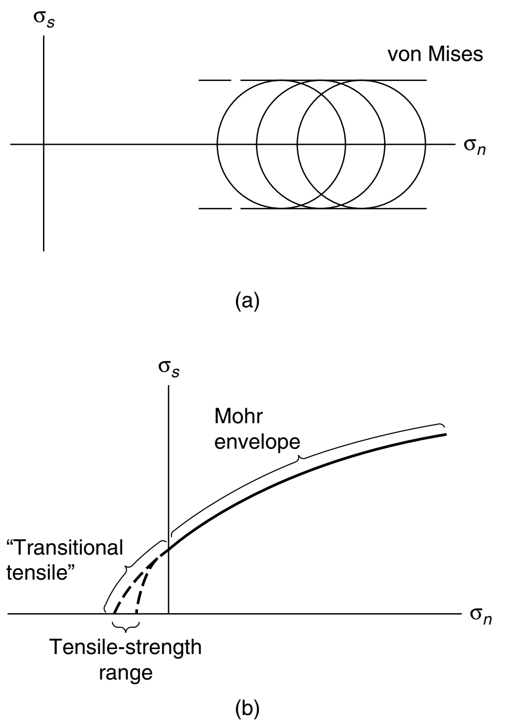


**FIGURE 6.20** (a) Mohr diagram illustrating the Von Mises yield criterion. Note that the criterion is represented by two lines that parallel the $\sigma_n$-axis. (b) The extrapolation of a Mohr envelope to its intercept with the $\sigma_n$-axis, illustrating the "transitional-tensile" regime, and the tensile strength ($T_0$). Note that the tensile strength has a range of values, because it depends on the dimensions of preexisting flaws in the deforming sample.

There are competing views as to the nature of failure for rocks subjected to tensile stresses that are less than the tensile strength. Some geologists have suggested that failure occurs under such conditions by the formation of fractures that are a hybrid between tensile cracks and shear ruptures, and have called these fractures **transitional-tensile fractures** or **hybrid shear fractures.** The failure envelope representing the conditions for initiating transitional-tensile fractures is the steeply sloping portion of the parabolic failure envelope (Figure 6.20b). Most fracture specialists, however, claim that transitional-tensile fractures do not occur in nature, and point out that no experiments have yet clearly produced transitional-tensile fractures in the lab. We'll discuss this issue further in Chapter 8.

Taking all of the above empirical criteria into account, we can construct a **composite failure envelope** that represents the boundary between stable and unstable stress states for a wide range of confining pressures and for conditions for which one of the principal stresses is tensile (Figure 6.21). The envelope roughly resembles a cross section of a cup lying on its side. The various parts of the curve are labeled. Starting at the right side of the diagram, we have Von Mises criteria, represented by horizontal lines. (Remember that the Von Mises portion of the envelope is really a plastic yield criterion, not a brittle failure criterion.) The portion of the curve where the lines begin to slope effectively represents the brittle–plastic transition. To the left of the brittle–plastic transition, the envelope consists of two straight sloping lines, representing Coulomb's criterion for shear rupturing. For failure associated with the Coulomb criterion, remember that the angle between the shear rupture and the $\sigma_1$ direction is independent of the confining pressure. Closer to the $\sigma_s$-axis, the slope of the envelope steepens, and the envelope resembles a portion of a parabola. This parabolic part of the curve represents Mohr's criterion, and for failure in this region, the decrease in the angle between the fracture and the $\sigma_1$ direction depends on how far to the left the Mohr circle touches the failure curve. The part of the parabolic envelope with steep slopes specifies failure criteria for supposed transitional-tensile fractures formed at a very small angle to $\sigma_1$, but as we discussed, the existence of such fractures remains controversial. The point where the envelope crosses the $\sigma_n$-axis represents the failure criterion for tensile cracking, but as we have discussed, this criterion really shouldn't be specified by a point, for the tensile strength of a material depends on the dimensions of the flaws it contains. Note that for a circle tangent to the composite envelope at $T_0$, $2\theta = 180$ (or, $\alpha = 0$), so

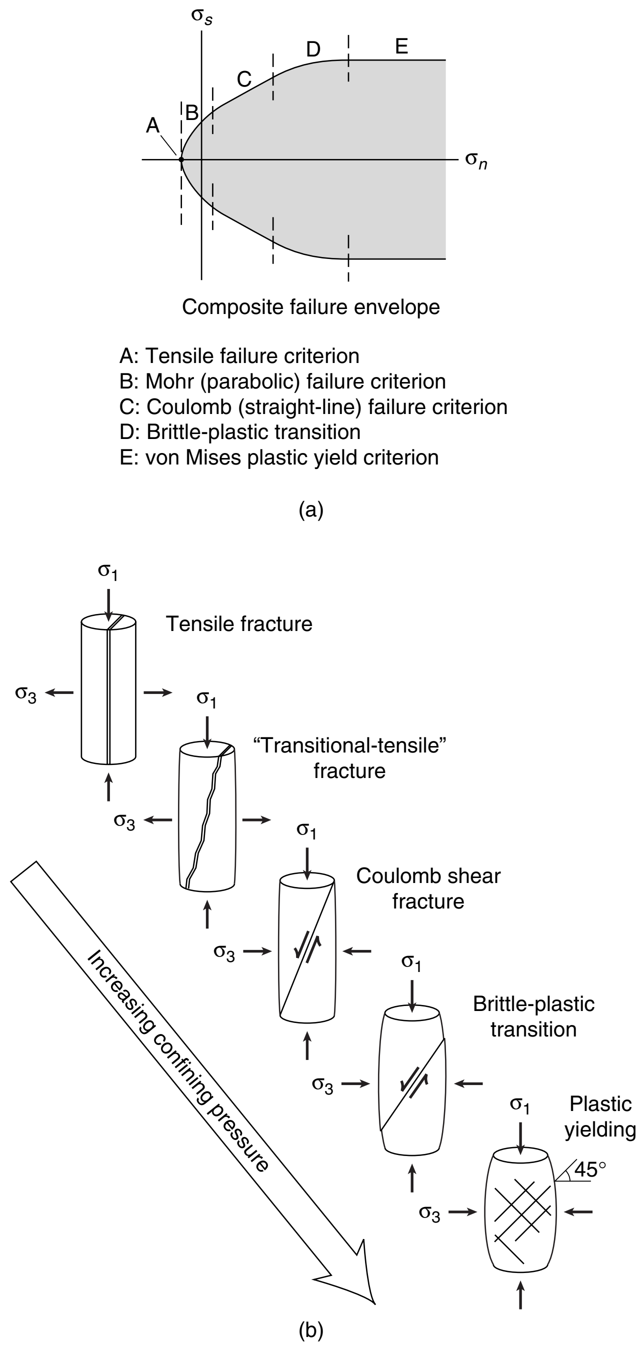


**FIGURE 6.21** (a) A representative composite failure envelope on a Mohr diagram. The different parts of the envelope are labeled, and are discussed in the text. (b) Sketches of the fracture geometries that form during failure. Note that the geometry depends on the part of the failure envelope that represents failure conditions, because the slope of the envelope is not constant.

the fracture that forms is parallel to $\sigma_1$! Also, note that there is no unique value of differential stress needed to cause tensile failure, as long as the magnitude of the differential stress (the diameter of the Mohr circle) is less than about $4T_0$, for this is the circle whose curvature is the same as that of the apex of the parabola.

## 6.8 FRICTIONAL SLIDING

Friction is the resistance to sliding on a surface. **Frictional sliding** refers to the movement on a surface that takes place when shear stress parallel to the surface exceeds the frictional resistance to sliding. The principles of frictional sliding were formulated hundreds of years ago. We can do a simple experiment that produces, at first, counterintuitive behavior (Figure 6.22). Attach a spring to a beam of wood that is placed on a table with the flat side down. Pull the spring until the beam slides. Now place the beam on its narrow side and repeat the experiment. Surprisingly, the spring extends by the same amount before beam sliding occurs, irrespective of the area of contact. Similar experiments led to what are often called Amontons's laws[6] of friction:

- Frictional force is a function of normal force.
- Frictional force is independent of the (apparent) area of contact (in his words, "the resistance caused by rubbing only increases or diminishes in proportion to greater or lesser pressure [load] and not according to the greater or lesser extent of the surfaces").
- Frictional force is mostly independent of the material used (in his words, "the resistance caused by rubbing is more or less the same for iron, lead, copper and wood in any combination if the surfaces are coated with pork fat").

Let's explore the reason for this behavior, which has important consequences for natural faulting processes and earthquake mechanics. Note that we should distinguish between **static friction,** which is associated with first motion (associated with Coulomb failure), and **dynamic friction,** which is associated with continued motion.

Friction exists because no real surface in nature, no matter how finely polished, is perfectly smooth. The bumps and irregularities which protrude from a rough surface are called **asperities** (Figure 6.23a). When two surfaces are in contact, they touch only at the asperities, and the asperities of one surface may indent or sink into the face of the opposing surface (Figure 6.23b and c). The cumulative area of the asperities that contact the opposing face is the **real area of contact** ($A_r$ in Figure 6.23). In essence, asperities act like an anchor holding a ship in place. In order for the ship to drift,

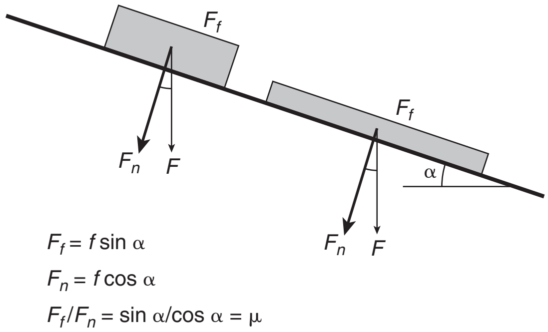


**FIGURE 6.22** Frictional sliding of objects with same mass, but with different (apparent) contact areas. The friction coefficients and, therefore, sliding forces ($F_f$) are equal for both objects, regardless of (apparent) contact area.

either the anchor chain must break, or the anchor must drag along the sea floor. Similarly, in order to initiate sliding of one rock surface past another, it is necessary either for asperities to break off, or for them to plow a furrow or groove into the opposing surface.

The stress necessary to break off an asperity or to cause it to plow depends on the real area of contact, so, as the real area of contact increases, the frictional resistance to sliding (that is, the force necessary to cause sliding) increases. Again, considering our ship analogy, it takes less wind to cause a ship with a small anchor to drift than it does to cause the same-sized ship with a large anchor to drift. Thus, the frictional resistance to sliding is proportional to the normal force component across the surface, because of the relation between real area of contact and friction. An increase in the normal force (load) pushes asperities into the opposing wall more deeply, causing an increase in the real area of contact. Returning to our earlier experiment with a sliding beam, we can now understand that the object's mass, rather than the (apparent) area of contact (the side of the beam), determines the ability to slide.

### 6.8.1 Frictional Sliding Criteria

Because of friction, a certain critical shear stress must be achieved in a rock before frictional sliding is initiated on a preexisting fracture, and a relation defining this critical stress is the **failure criterion for frictional sliding.** Experimental work shows that failure criteria for frictional sliding, just like the Coulomb failure criterion for intact rock, plot as sloping straight lines on a Mohr diagram. A compilation of friction data from a large number of experiments, using a great variety of rocks (Figure 6.24), shows that the failure criterion for frictional sliding is basically independent of rock type:

$$\sigma_s/\sigma_n = \text{constant} \qquad \text{Eq. 6.7}$$

[6]Named after the seventeenth-century French physicist Guillaume Amontons; Leonardo da Vinci earlier described similar relationships in the fifteenth century.

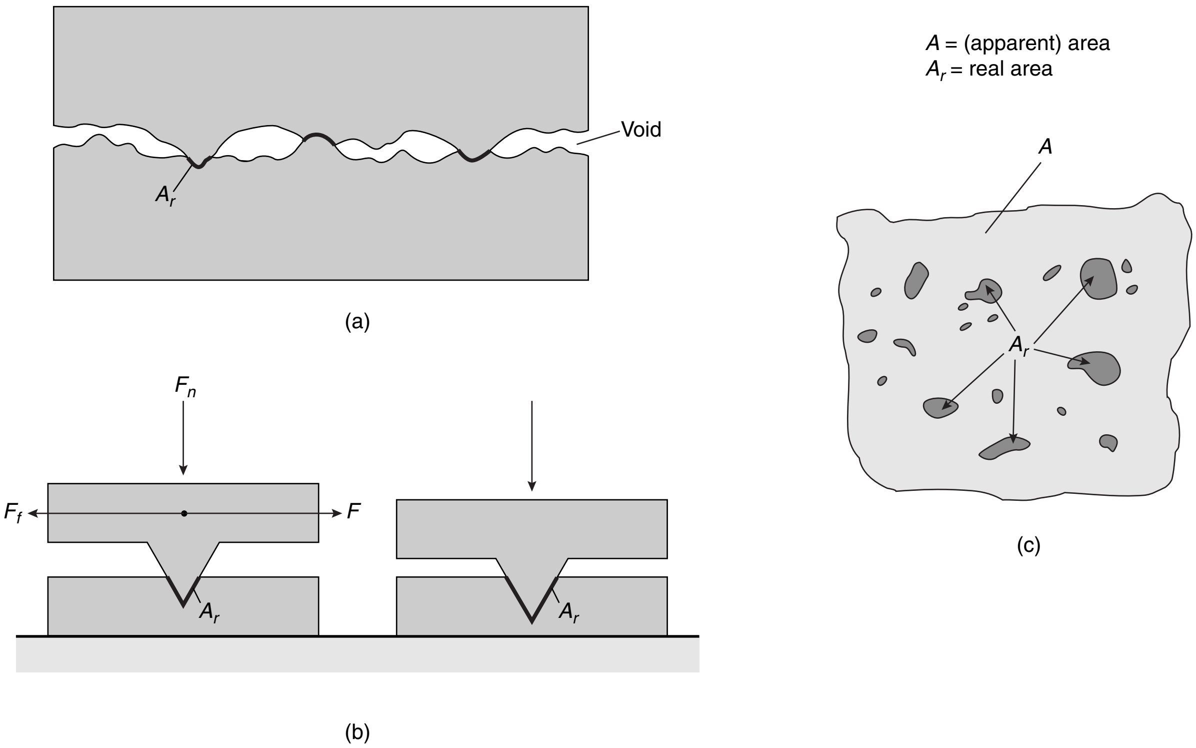


**FIGURE 6.23** Concept of asperities and the real area of contact ($A_r$). (a) Schematic cross-sectional close-up showing the irregularity of a fracture surface and the presence of voids and asperities along the surface. (b) Idealized asperity showing the consequence of changing the load (normal force) on the real area of contact. (c) Map of a fracture surface; the shaded areas are real areas of contact.

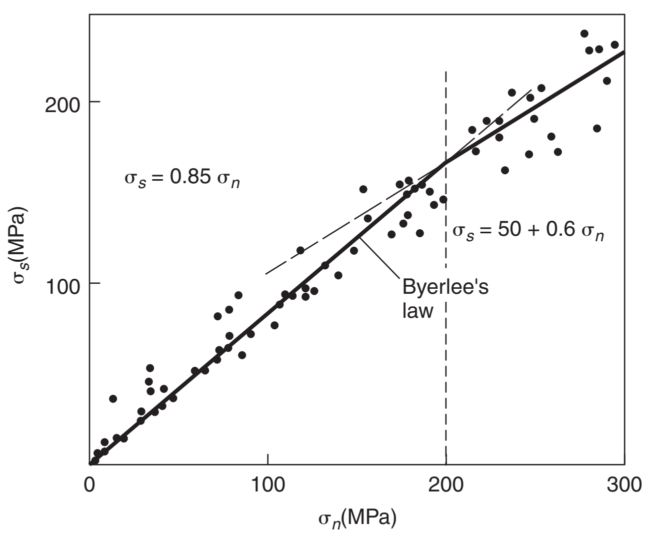


**FIGURE 6.24** Graph of shear stress and normal stress values at the initiation of sliding on preexisting fractures in a variety of rock types. The best-fit line defines Byerlee's law, which is defined for two regimes.

The empirical relationship between normal and shear stress that best fits the observations, known as **Byerlee's law,**[7] depends on the value of $\sigma_n$. For $\sigma_n < 200$ MPa, the best-fitting criterion is a line described by the relationship $\sigma_s = 0.85\sigma_n$, whereas for 200 MPa $< \sigma_n <$ 2000 MPa, the best-fitting criterion is a line described by the equation $\sigma_s = 50$ MPa $+ 0.6\sigma_n$. The proportionality between normal and shear stress is, as before, called the **friction coefficient.**

## 6.8.2 Will New Fractures Form or Will Existing Fractures Slide?

Failure envelopes allow us to quickly determine whether it is more likely that an existing shear rupture will slip in a sample, or that a new shear rupture will form (Figure 6.25). For example, look at Figure 6.25b, which shows both the Byerlee's frictional sliding envelope and the Coulomb shear fracture envelope for Blair dolomite. Note that the slope and intercept of the two envelopes

[7]After the geophysicist J. Byerlee, who first proposed the equations in the 1970s.

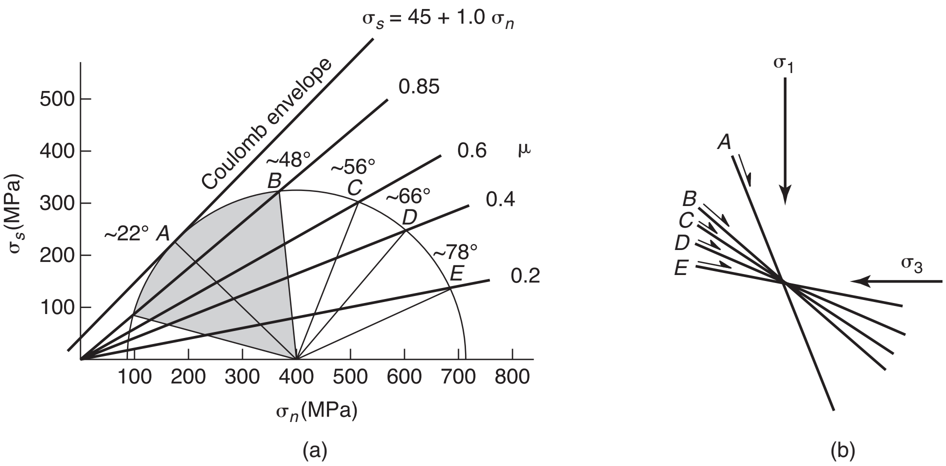


**FIGURE 6.25** (a) Mohr diagram based on experiments with Blair dolomite, showing how a single stress state (Mohr circle) would contact the frictional sliding envelope before it would contact the Coulomb envelope (heavy line). Surface *A* in (b) is the Coulomb shear fracture that would form in an intact rock. However, sliding would occur on surfaces between intersections with the friction envelope (marked by shaded area for friction envelope $\mu = 0.85$) before new fracture initiation. Preexisting surfaces *B* to *E* are surfaces that will slide with decreasing friction coefficients. Consider the geologic relevance of decreasing friction coefficients for stress state, failure, and fracture orientation.

are different, so that for a specific range of preexisting fracture orientations, the Mohr circle representing the stress state at failure touches the frictional envelope before it touches the fracture envelope, meaning the preexisting fracture slides before a new fracture forms.

However, preexisting fractures do not always slide before a new fracture is initiated. Confined compression experiments indicate that if the preexisting fracture is oriented at a high angle to the $\sigma_1$ direction (generally > 75°; plane *E* in Figure 6.25), the normal stress component across the discontinuity is so high that friction resists sliding, and it is actually easier to initiate a new shear fracture at a smaller angle to $\sigma_1$ (plane *B* in Figure 6.25). Sliding then occurs on the new fracture. If a preexisting fracture is at a very small angle to $\sigma_1$ (generally < 15°; plane *A* in Figure 6.25), the shear stress on the surface is relatively low, so again it ends up being easier to initiate a new shear fracture than to cause sliding on a preexisting weak surface. Thus, preexisting planes whose angles to $\sigma_1$ are between 15° and 75° probably will be reactivated before new fractures form.

## 6.9 EFFECT OF ENVIRONMENTAL FACTORS IN FAILURE

The occurrence and character of brittle deformation at a given location in the earth depends on environmental conditions (confining pressure, temperature, and fluid pressure) present at that location, and on the strain rate (see Chapter 5). Conditions conducive to the occurrence of brittle deformation are more common in the upper 10–15 km of the Earth's crust. However, at slow strain rates or in particularly weak rocks, ductile deformation mechanisms can also occur in this region, as evident by the development of folds at shallow depths in the crust. Below 10–15 km, plastic deformation mechanisms dominate. However, at particularly high fluid pressures or at very rapid strain rates, brittle deformation can still occur at these depths.

In this chapter, we have described brittle deformation without considering how it is affected by environmental factors. Not surprisingly, temperature, fluid pressure, strain rate, and rock anisotropy play significant roles in the stress state at failure and/or in the orientation of the fractures that form when failure occurs. Most of these factors have already been discussed in Chapter 5, so we close this chapter on brittle deformation by focusing on the effect of fluids.

### 6.9.1 Effect of Fluids on Tensile Crack Growth

All rocks contain pores and cracks—we've already seen how important these are in the process of brittle failure. In the upper crust of the earth below the water table, these spaces, which constitute the porosity of rock, are filled with fluid. This fluid is most commonly water, though in some places it is oil or gas.

If there is a high degree of **permeability** in the rock, meaning that water can flow relatively easily from pore to pore and/or in and out of the rock layer, then the pressure in a volume of pore water at a location in the crust is roughly hydrostatic, meaning that the pressure reflects the weight of the overlying water column (Figure 6.26). **Hydrostatic** (fluid) **pressure** is defined by the relationship $P_f = \rho \cdot g \cdot h$, where $\rho$ is the density of water (1000 kg/m$^3$), $g$ is the gravitational constant (9.8 m/s$^2$), and $h$ is the depth. **Pore pressure,** which is the fluid pressure exerted by fluid within the pores of a rock, may exceed the hydrostatic pressure if permeability is restricted. For example, the fluid trapped in a

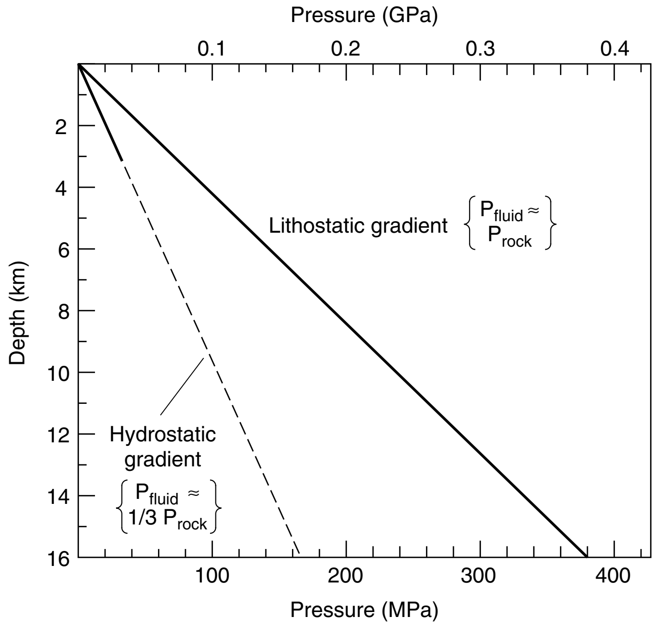


**FIGURE 6.26** Graph of lithostatic versus hydrostatic pressure as a function of depth in the Earth's crust.

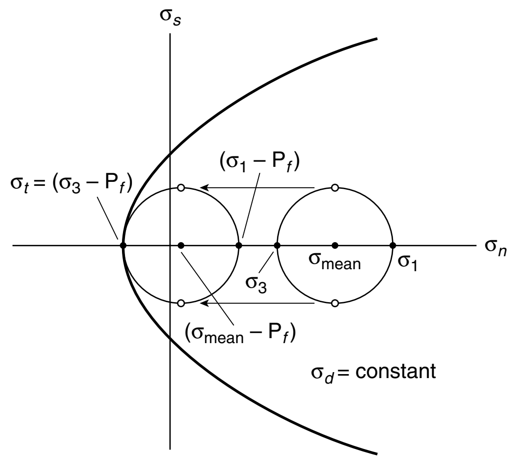


**FIGURE 6.27** A Mohr diagram showing how an increase in pore pressure moves the Mohr circle toward the origin. The increase in pore pressure decreases the mean stress ($\sigma_{mean}$), but does not change the magnitude of differential stress ($\sigma_1 - \sigma_3$). In other words, the diameter of the Mohr circle remains constant, but its center moves to the left.

sandstone lens surrounded by impermeable shale cannot escape, so the pore pressure in the sandstone can approach or even equal **lithostatic pressure** ($P_l$), meaning that the pressure approaches the weight of the overlying column of rock (i.e., $P_f = P_l = \rho \cdot g \cdot h$, where $\rho$ = 2000–3000 kg/m$^3$). Note that rock, on average, is two to three times denser than water. When the fluid pressure in pore water exceeds hydrostatic pressure, we say that the fluid is **overpressured.**

How does pore pressure affect the tensile failure strength of rock? The pore pressure is an outward push that opposes inward compression from the rock, so the fluid supports part of the applied load. If pore pressure exceeds the least compressive stress ($\sigma_3$) in the rock, tensile stresses at the tips of cracks oriented perpendicularly to the $\sigma_3$ direction become sufficient for the crack to propagate. In other words, pore pressure in a rock can cause tensile cracks to propagate, even if none of the remote stresses are tensile, because pore pressure can induce a crack-tip tensile stress that exceeds the magnitude of $\sigma_3$. This process is called **hydraulic fracturing.** On a Mohr diagram, it can be represented by movement of Mohr's circle to the left (Figure 6.27). Note that rocks do not have to be overpressured in order for natural hydraulic fracturing to occur, but $P_f$ must equal or exceed the magnitude of $\sigma_3$.

Another effect of fluids comes from the chemical reaction of the fluids with the minerals comprising a rock. Reaction with fluids may lower the tensile stress needed to cause a crack to propagate, even if the pore-fluid pressure is low. Water, for example, reacts with quartz to bring about substitution of OH molecules for O atoms in quartz lattice at a crack tip (Figure 6.28). Since the bond between adjacent OH groups is not as strong as the bond between oxygen atoms, it breaks more easily, so it takes less remote tensile stress to cause the crack to propagate. This phenomenon is called **subcritical crack growth,** because crack propagation occurs at stresses less than the critical stress necessary to cause a crack to propagate in dry rock.

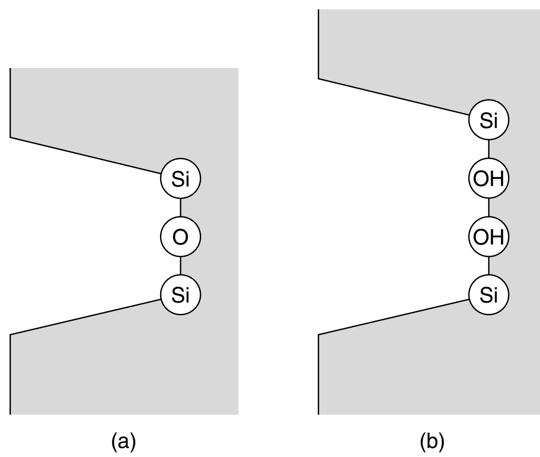


**FIGURE 6.28** Sketch of the atoms at the tip of a crack that illustrates the principle of hydrolytic weakening. (a) Si atoms are bonded to an O atom at the tip of a dry crack. (b) At the tip of a wet crack, Si atoms are bonded to two OH molecules that replace the O (same charge). The Si—O—Si bonding is stronger than the Si—OH to OH—Si bonding and the Si—O—Si bond length is less than that of Si—OH—OH—Si.

### 6.9.2 Effect of Dimensions on Tensile Strength

Rock tensile strength is *not* independent of scale, in that larger rock samples are inherently weaker than smaller rock samples. Why? Because larger samples are more likely to contain appropriately oriented and larger Griffith cracks that will begin to propagate when a stress is applied, thereby nucleating the throughgoing cracks that result in failure of the whole sample. Equation 6.2 emphasizes this point, because the tensile stress at failure is inversely proportional to the crack half-length. You can imagine that if a sample is so small that it consists only of a piece of perfect crystal lattice, it will be very strong indeed. In fact, the reason that turbine blades in modern jet engines are so strong is that they are grown as relatively flawless single crystals.

### 6.9.3 Effect of Pore Pressure on Shear Failure and Frictional Sliding

We can observe the effects of pore pressure on shear fracturing by running a confined compression experiment in which we pump fluid into the sample through a hole in one of the pistons, thereby creating a fluid pressure, $P_f$, in pores of the sample (Figure 6.10). The fluid creating the confining pressure acting on the sample is different from and is not connected to the fluid inside the sample. The magnitude of $P_f$ decreases the confining pressure ($\sigma_3$) and $\sigma_1$ by the same amount. So if the pore pressure increases in the sample, the mean stress decreases but the differential stress remains the same. This effect can be represented by the Coulomb failure criterion equation; $P_f$ decreases the magnitude of $\sigma_n$ on the right side of the equation

$$\sigma_s = C + \mu(\sigma_n - P_f) \qquad \text{Eq. 6.8}$$

The term $(\sigma_n - P_f)$ is commonly labeled $\sigma_n{}^*$, and is called the **effective stress.**

From a Mohr diagram, we can easily see the effect of increasing the $P_f$ in this experiment. When $P_f$ is increased, the whole Mohr circle moves to the left but its diameter remains unchanged (Figure 6.27), and when the circle touches the failure envelope, shear failure occurs, even if the relative values of $\sigma_1$ and $\sigma_3$ are unchanged. In other words, a differential stress that is insufficient to break a dry rock, may break a wet rock, if the fluid in the wet rock is under sufficient pressure. Thus, an increase in pore pressure effectively weakens a rock. In the case of forming a shear fracture in intact rock, pore pressure plays a role by pushing open microcracks, which coalesce to form a rupture at smaller remote stresses.

Similarly, an increase in pore pressure decreases the shear stress necessary to initiate frictional sliding on a preexisting fracture, for the pore pressure effectively decreases the normal stress across the fracture surface. Thus, as we discuss further in Chapter 8, fluids play an important role in controlling the conditions under which faulting occurs.

### 6.9.4 Effect of Intermediate Principal Stress on Shear Rupture

Fractures form parallel to $\sigma_2$, and so the value of $\sigma_2$ does not affect values of normal stress ($\sigma_n$) and shear stress ($\sigma_s$) across potential shear rupture planes. Therefore, in this chapter we have assumed that the value of $\sigma_2$ does not have a major effect on the shear failure strength of rock, and we considered failure criteria only in terms of $\sigma_1$ and $\sigma_3$. In reality, however, $\sigma_2$ does have a relatively small effect on rock strength. Specifically, rock is stronger in confined compression when the magnitude of $\sigma_2$ is closer to the magnitude of $\sigma_1$, than when the magnitude of $\sigma_2$ is closer to the magnitude of $\sigma_3$. An increase in $\sigma_2$ has the same effect as an increase in confining pressure.

## 6.10 CLOSING REMARKS

Why study brittle deformation? For starters, since rocks deform in a brittle manner under the range of pressures and temperatures found at or near the Earth's surface (down to 10–15 km), fractures pervade rocks of the upper crust. In fact every rock outcrop that you will ever see contains fractures at some scale (Figure 6.29). Because they are so widespread, fractures play a major role in determining the permeability and strength of rock, and the resistance of rock to erosion. Therefore, fractures affect the velocity and direction of toxic waste transport, the location of an ore deposit, the durability of a foundation, the stability of a slope, the suitability of a reservoir, the safety of a mine shaft, the form of a landscape, and so on. Moreover, fracturing is the underlying cause of earthquakes, and contributes to the evolution of regional tectonic features. This chapter has provided an introduction to the complex and rapidly evolving subject of brittle deformation and fracture mechanics. We have

**FIGURE 6.29** Aerial view of the San Andreas Fault, north of Landers, California.

tried to describe what fractures are, how they form, and under what conditions. In the next two chapters, we apply this information to developing an understanding of the major types of brittle structures: joints, veins, and faults.

## ADDITIONAL READING


Atkinson, B. K., ed., 1987. *Fracture mechanics of rock.* London: Academic Press.

Bowden, F. P., and Tabor, D., 1950. *The friction and lubrication of solids.* Clarendon Press: Oxford.

Brace, W. F., and Bombolakis, E. G., 1963. A note on brittle crack growth in compression. *Journal of Geophysical Research,* 68, 3709–3713.

Brace, W. F., and Byerlee, J. D., 1966. Stick-slip as a mechanism for earthquakes. *Science,* 153, 990–992.

Engelder, T., 1993. *Stress regimes in the lithosphere.* Princeton University Press.

Hancock, P. L., 1985. Brittle microtectonics: principles and practice. *Journal of Structural Geology,* 7, 437–457.

Jaeger, J. C., and Cook, N. G. W., 1979. *Fundamentals of rock mechanics* (3rd ed.). Chapman and Hall: London.

Lockner, D. A., 1991. A multiple-crack model of brittle fracture. *Journal of Geophysical Research,* 96, 19,623–19,642.

Paterson, M. S., 1978. *Experimental rock deformation—the brittle field.* Springer-Verlag: New York.

Pollard, D. D., and Aydin, A., 1988. Progress in understanding jointing over the past century. *Geological Society of America Bulletin,* 100, 1181–1204.

Price, N. J., 1966. *Fault and joint development in brittle and semi-brittle rock.* Pergamon Press: London.

Scholz, C. H., 2002. *The mechanics of earthquakes and faulting* (2nd edition). Cambridge University Press: Cambridge.

Scholz, C. H., 1998. Earthquakes and friction laws. *Nature,* 391, 37–42.

Secor, D. T., 1965. Role of fluid pressure in jointing. *American Journal of Science,* 263, 633–646.