

1

Earthquakes and fault motion

1.1. The origin of earthquakes

Humankind’s experience of earthquakes has always given rise to shock and confusion. The ground beneath us, our basis of stability, is suddenly subject to shaking, bringing ruin to buildings and causing high numbers of casualties. Investigating the causes of these phenomena has interested human beings since antiquity. The first rational explanations, beyond mythical stories, were presented by Greek natural philosophers beginning in the sixth century BC. The most lasting theory of the Greek philosophers was presented by Aristotle (fourth century BC) in his book *Meteorologica*. By the word *meteors* ancient Greeks referred to a variety of phenomena which were believed to take place somewhere above the Earth’s surface, such as rain, wind, thunder, lightning, comets and the Milky Way, or inside the Earth, such as earthquakes and volcanoes. Aristotle refuted the ideas proposed by previous authors, as Anaxagoras, Empedocles and Democritus, and proposed that the cause of earthquakes consisted in the shaking of the Earth by dry heated underground exhalations or winds trapped in cavities of its interior, as they attempt to escape toward the exterior. This explanation was part of his general theory for “meteors”, which, he proposed, were all caused by various types of exhalation of dry or humid, cold or hot, vapors or winds, which extended from the interior of the Earth to the Lunar orbit. This was a very popular theory, which was accepted with only minor changes until the seventeenth century. Roman authors such as Pliny and Seneca presented it in their encyclopedic works and medieval authors such as Albert the Great and Thomas Aquinas wrote long commentaries on it.

The general reaction against Aristotelian physics, and the birth of modern science in the seventeenth century, gave rise to a new theory to explain the origin of earthquakes. Martin Lister in England and Nicolas Lemery in France proposed that earthquakes were caused by spontaneous explosions of flammable material accumulated underground in certain regions. The increasing use of mines in mining and military operations suggested this idea. The explosive theory was supported by Newton and Buffon and helped to define the concept of the focus of an earthquake as the location where the explosion took place.

The great Lisbon earthquake of 1755 created in Europe a widespread interest in the study of these phenomena and can be considered as marking the beginning of modern seismology. In 1760 John Michell, for the first time, related the ground-shaking due to earthquakes with the propagation through the Earth of elastic waves generated at the focus. Thomas Young, Robert Mallet and John Milne, among others, further developed this idea, laying the foundations of seismology. Mallet, in his study of the Naples earthquake of

1857, developed the theory of the seismic focus, from which elastic waves spread out in all directions. He connected the occurrence of earthquakes with changes in the Earth's crust, which often result in dislocations and fractures, though he did not abandon the explosive theory. Charles Lyell, in his foundational work on geology, related earthquakes to tectonic processes and volcanic activity. Edward Suess, Ferdinand Montessus de Ballore and Alfred Sieberg, among others, continued these studies, relating the occurrence of earthquakes to orogenic and tectonic processes. Systematic studies of field observations after the occurrence of earthquakes in the nineteenth century, such as those of George K. Gilbert (Owen, California, 1872), Bunjiro Koto (Nobi, Japan, 1891) and Richard D. Oldham (Assan, India, 1897), began relating earthquakes to motion along faults. With the increase of these field observations and precision in locating epicenters, the relation between earthquakes and faults became more evident. The cause of earthquakes began to be understood as the release by fracture of the tectonic stresses accumulated in the Earth's crust, as we now explain.

Francis Reid in 1910 presented the first mechanical model for the faulting process of earthquakes, in order to explain the observed ground fracture with horizontal displacement observed in the San Francisco earthquake of 1906. This earthquake marked a milestone in the development of earthquake source studies, as did the Lisbon earthquake for general seismology. Reid's theory, known as the *elastic rebound* theory, assumes that earthquakes take place by a fracturing of the Earth's crust with total or partial release of the elastic strain accumulated in a region owing to the stresses produced by tectonic processes. In his own words: "the difference between displacements in neighboring regions sets up elastic strains, which may become greater than the rock can endure; a rupture then takes place and the strained rock rebounds under its own elastic stresses until the strain is largely or wholly relieved" (Reid, 1911). With certain modifications Reid's basic insight, that earthquakes are caused by fractures in the Earth's crust induced by tectonic stress, still remains valid today. According to plate tectonic theory, developed between 1960 and 1970, tectonic stresses are ultimately related to the relative motion of lithospheric plates. An earthquake can be considered, then, to be produced by the rupturing of a certain part of the Earth's crust, with a relative displacement of its two sides and the release of the accumulated elastic strain, produced by the relative motion of lithospheric plates. A rupture of the crust is generally referred to as a *fault*. Owing to the confining pressure inside the Earth, faults are shear fractures; that is, motion takes place along the breaking surface.

1.2. Faults in the Earth's crust

Most earthquakes are of tectonic origin and take place on faults. Non-tectonic earthquakes, such as those of volcanic origin or induced by dams, may have different characteristics. Geological field observations show the presence of many fault traces exposed at the Earth's surface, with lengths which vary from some meters to many kilometers (Fig. 1.1). In general, observed fault traces are not continuous straight lines but have bends, bifurcations, offsets and other complexities. The non-continuous nature of observed faults has led



Figure 1.1. Surface rupture of the San Andreas fault, California. View along Carrizo Plains (US Geological Survey/Photo R. E. Wallace).

to the consideration that their nature might be fractal. The concept of a fault plane is, then, a useful simplification of a situation which involves non-planar geometry and various complexities. The depth extent at which faults break by brittle fracture generating earthquakes can be deduced from the depth of shallow earthquakes. From this evidence, it has been shown that brittle fracture extends only to about between 15 km and 20 km in depth. This shows that only the upper part of the crust has a brittle nature and can, in consequence, be fractured. This brittle upper part of the crust is called the *seismogenic* layer, that is, the layer capable of generating earthquakes. The lower part of the crust and the upper mantle, owing to the increase in temperature and pressure, behaves as a plastic material which can be deformed without breaking. This behavior of the lower crust and upper mantle has been confirmed from deep seismic reflection studies.

The limitation on the brittle depth imposes a condition on the fault-surface shape; the latter is roughly described by its aspect ratio L/W (length over width). For small earthquakes, with fracture lengths less than 20 km, there is no limitation to the propagation of fracture in all directions, so that the shape of faults is symmetrical, approximately circular or square, with aspect ratio near to unity. For large earthquakes the length is longer than the depth limit and so the shape will be rectangular or elliptical, with a longer length than width, and aspect ratio greater than unity. Very large crustal earthquakes have lengths extending for hundreds of kilometers and only a limited width, less than about 40 kilometers; their aspect ratios may be greater than 10. However, large earthquakes in subduction zones with lengths longer than 400 km may have widths of about 100 km. The shape of the fault depends, then, on the size and type of earthquake, with different shapes for small and large events.

A different problem is presented by the nature of earthquakes with foci at depths below 60 km. These earthquakes are produced in regions where tectonic processes have introduced crustal material inside the Earth's mantle. Intermediate-depth ($60 \text{ km} < h < 200 \text{ km}$)

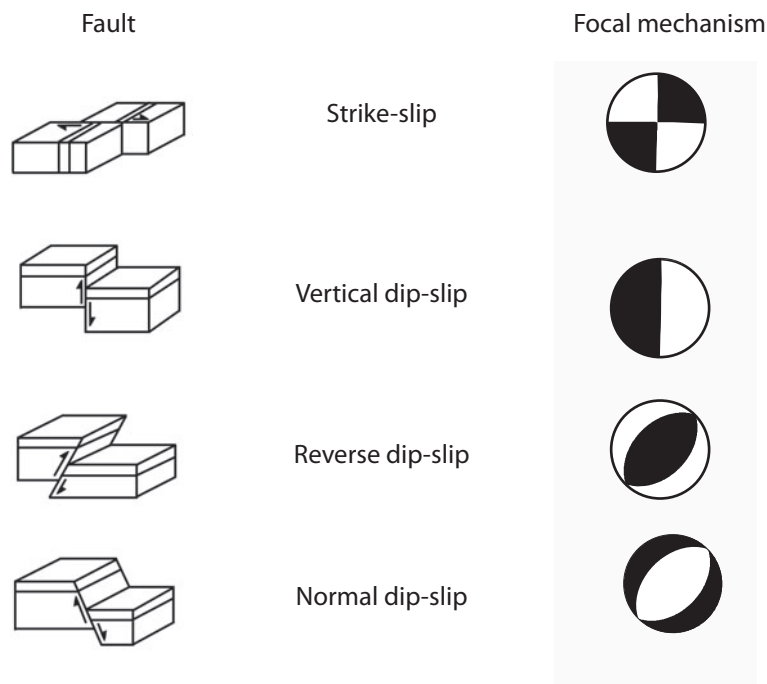


Figure 1.2. Types of fault and their focal mechanism diagrams.

and deep ($200\text{ km} < h < 700\text{ km}$) earthquakes cannot be produced by brittle fracture, owing to the high temperature and pressure existing at those depths. Different processes, such as sudden phase transitions, have been proposed for them; however, they do occur on weakness surfaces similar to those of faults.

Faults can be classified according to their type of motion into *strike-slip* faults, where the motion is horizontal and parallel to the strike, and *dip-slip* faults, where in cross-section the motion is in a vertical plane (see Fig. 1.2). Strike-slip faults are divided into right-lateral and left-lateral, according to the direction in which one side of the fault moves as seen from the other side. Dip-slip faults are divided into *normal* faults, with the hanging wall sliding downward with respect to the foot wall, and *reverse* faults, with motion in the opposite direction; the hanging wall moves upward with respect to the foot wall. Normal faults are also called *gravity* faults, because their motion results from a process driven by gravity. Reverse faults with a small dip angle are called *thrust* faults (Fig. 1.2). Combinations of horizontal and vertical motion result in *oblique* faults. Normal faults are caused by horizontal tension and reverse faults by horizontal pressure. In strike-slip faults the tension and pressure are both horizontal.

In general, earthquakes occur on pre-existing fault planes although, naturally, at some time faults involve newly broken material. The material between the two moving surfaces of a fault is usually filled with crushed and highly deformed rocks, called *fault gouge*, in a zone that may be up to several meters wide in some places. This material is formed by the repeated differential movement of the two sides of the fault over a long time; this breaks and

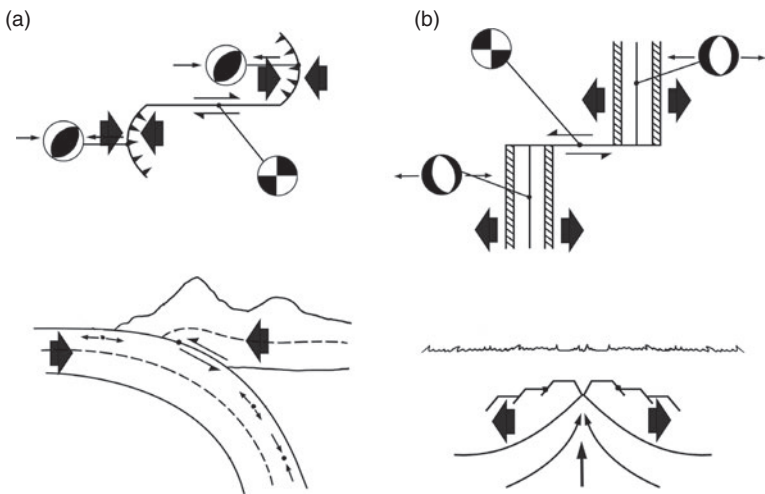


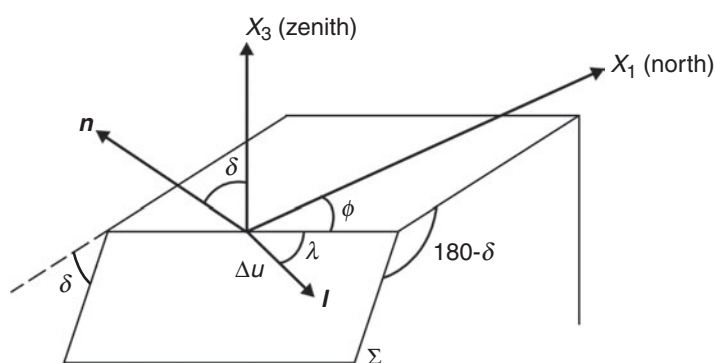
Figure 1.3. Plate boundaries and focal mechanisms: horizontal plans and vertical cross-sections. (a) Subduction: reverse and transform (strike-slip) faults. (b) Rift: normal and transform (strike-slip) faults. The large arrows show the stress directions. The directions of motion are represented by the small half-arrows.

shears the rock into a fine granular and powdery form, which can be altered by the addition of water into clays and sandy silts. Earthquakes do not happen at present on all observed geological faults; that is, not all faults observed today in the field are seismically active. On a given active fault, earthquakes of different size break partial sections of its surface. On a long fault several large earthquakes may be needed to break its whole extent. Faults can also move slowly without producing an earthquake; this type of motion is called *fault creep*.

The active factor responsible for movement along faults is the tectonic stress resulting from the relative motion of *lithospheric plates*, which are driven by the convective currents in the mantle. According to plate tectonics the Earth surface is divided into several plates (the main plates are the Eurasia, Africa, America, Pacific, Indo-Australia, Antarctic, Nazca and Arabia plates), which move relative to each other with velocities ranging from about 1 to 8 cm/year. The plate boundaries are, then, the places where most earthquakes occur (*interplate* earthquakes), though earthquakes also occur in the interior of plates (*intraplate* earthquakes). The types of boundary between plates can be reduced to three and at each of these a corresponding type of fault is predominant: where plates are separating from each other (at *extension* zones), faults are normal; where plates are colliding (at *collision* and *subduction* zones) faults are reverse; and, where plates slide horizontally over each other, faults are strike-slip (Fig. 1.3).

1.3. Geometry of a fault

A fault is defined as a rupture in the Earth’s crust with a relative displacement of its two sides. The size of the fault is given by its area S , which for a rectangular fault is given by $S = LW$, where L is the length and W the width, and for a circular fault is given by $S = \pi a^2$, where a is the radius. These two approximations to the shapes of faults are very often used.



Geometry of a shear fault. The fault orientation is given by either angles ϕ (azimuth), δ (dip) and λ (the rake or slip angle) or unit vectors \mathbf{n} (normal to the fault plane) and \mathbf{l} (in the slip direction). The fault slip is Δu and the fault surface is Σ .

The values of δ and λ correspond to different types of fault. We have mentioned above (and see Fig 1.2) strike-slip faults, for which $\delta = 90^\circ$ and $\lambda = 0^\circ$, dip-slip vertical faults, for which $\delta = 90^\circ$ and $\lambda = 90^\circ$, and dip-slip faults on an inclined plane, for which $0^\circ < \delta < 90^\circ$ and $\lambda = 90^\circ$. According to the value of λ , on an inclined fault we can have horizontal motion ($\lambda = 0^\circ, 180^\circ, -180^\circ$), vertical motion ($\lambda = 90^\circ$) or inclined motion with horizontal and vertical components ($-180^\circ < \lambda < 180^\circ$). If λ is negative then the fault is a normal fault and if positive it is a reverse fault. The angles ϕ , δ and λ are shown on a stereographic projection of the fault plane in Fig. 1.5.

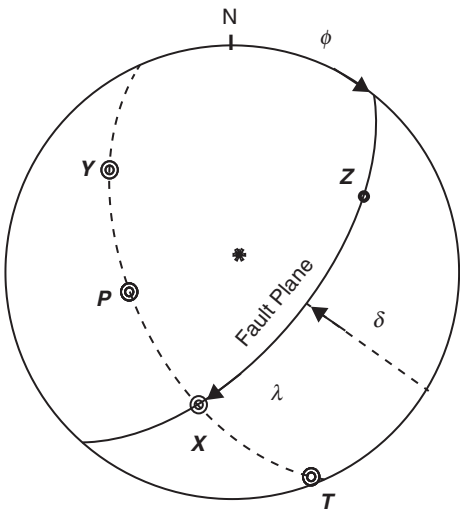


Figure 1.5. Stereographic projection of the fault plane: the angles ϕ , δ , λ , shown by arrows; the double-couple axes X , Y , Z (couples along X and Y) and the axes of the principal stresses P , T , Z (P , pressure; T , tension).

1.4. Elastic rebound and the earthquake cycle

As we mentioned in the introduction, the most accepted model of the origin of earthquakes is the earthquake rebound model proposed by Francis Reid in 1910 (Reid, 1911). Following the earthquake that devastated the city of San Francisco in 1906, the American government charged a commission led by Reid, a professor at Johns Hopkins University in Baltimore, to evaluate the damage produced by the earthquake and to try to understand its origin. Using geodetic data that had been collected before and after the earthquake of 1906, he proposed a model that has become the basis of modern earthquake source studies.

Let us consider the simplified model of a crustal block, shown in Fig. 1.6, and assume that an earthquake has just occurred on the main fault, which runs across the middle of the block. During a certain period of time many aftershocks will occur on the fault itself and its immediate vicinity and a certain amount of post-seismic deformation will occur. After a certain time the activity will cease; we choose this time as the zero instant, $t = 0$, of our earthquake cycle. Then a slow continuous process of stress accumulation will start around the main fault from time $t = 0$. This stage in the fault progress is shown in Fig. 1.6a. As time passes, the slow movement of the plates, of the order of a few centimeters per year, will produce a pre-seismic deformation of the rocks that surround the fault. This deformation, shown in Fig. 1.6b, produces a generalized increase of the stress level around the fault.

An essential point with Reid's model is that plate motion does not produce deformation that is uniformly distributed over the plates. Reid and his colleagues proposed that the deformation around the fault was actually concentrated in a relatively narrow band some 40 km in width. This localization of deformation is an essential part of the elastic rebound theory. Without a mechanism that localizes deformation, there would be no earthquakes;

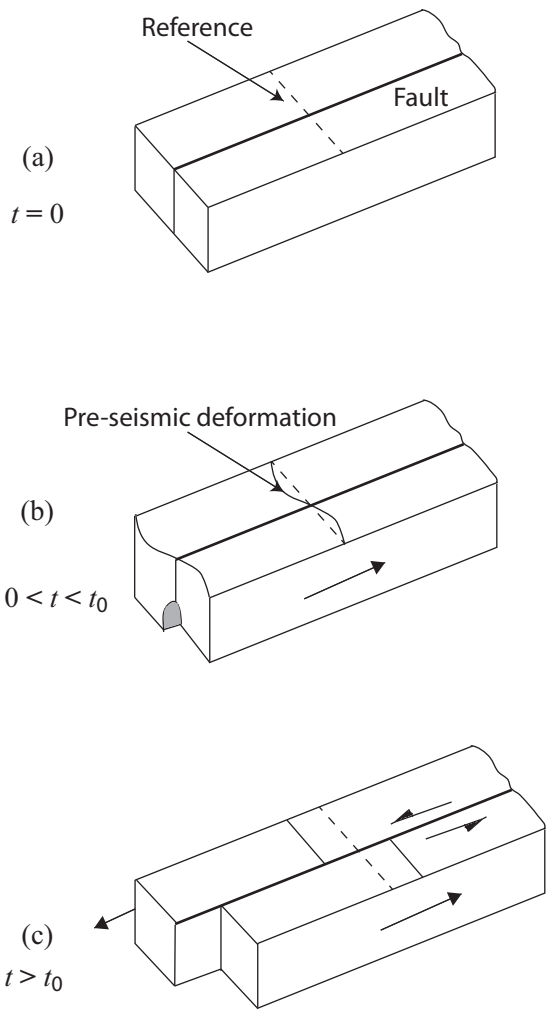


Figure 1.6. Schematic representation of the elastic rebound theory: (a) initial situation, at $t = 0$; (b) pre-seismic deformation due to applied stress; (c) displacement after earthquake occurrence at time t_0 .

the Earth would deform continuously as the plates move. In Reid's time the internal structure of fault zones was practically unknown, but since then geodetic observations have improved significantly. We now know that the localization of deformation does occur because at depths greater than 15 or 20 km the crust reaches temperatures close to about 30% of the melting temperatures of the rocks that make up the crust, and at those temperatures rocks flow under stress and are no longer elastically deformed. Thus, the deeper part of the San Andreas fault, shown by the shaded area in Fig. 1.6b, undergoes continuous aseismic slip due to creep. Creep may be localized on the deep roots of the faults or may be spread over a large zone around the fault. In Fig. 1.6b we illustrate the former model.

Finally, once the deformation around the upper region near the fault is such that the stress across the fault is larger than its frictional resistance, an earthquake is triggered,

at time $t = t_0$, producing a sudden release of accumulated stresses on the fault (Fig. 1.6c). This sudden release of stresses produces a slip on the fault that is usually called *elastic rebound*. This is a way of expressing the idea that during the earthquake the brittle part of the crust slips very rapidly until the slowly accumulated stress is eliminated. In the idealized model shown at Fig. 1.6c, the two sides of the fault have slipped uniformly so that there is no stress left on the fault. In reality, of course, elastic rebound is not perfect and, once the earthquake has occurred, a period of slow post-seismic slip occurs. At the end of the current earthquake the entire cycle starts again with a slow accumulation of deformation near the surface and continuous slip at depth. The process we have just described is cyclic, although not necessarily periodic; it is what seismologists call the *earthquake cycle*, an endless repetition of slow loading and fast release of stresses on the fault zone.

Reid’s model has been extensively tested from geodetic field observation in the last 40 years, starting with the classical work by Savage and Burford (1973) on the Parkfield segment of the San Andreas fault. Observation of the pre-seismic accumulation of stresses has now been extended to subduction zones, where, although the details are more complex, again the basic mechanism of the slow accumulation of deformation at shallow depths and its eventual release by elastic rebound has been observed after many large earthquakes.

1.5. Energy, stress drop and seismic moment

Motion on a fault happens when the applied shear stress on the fault plane overcomes the strength of the material or the friction which holds its two sides locked. The strength under the given conditions can be defined as the maximum stress that the material can support without suffering a permanent deformation or failure. The material has a plastic behavior if it suffers a permanent deformation without breaking and a brittle fracture behavior if it suffers permanent deformation by failure, with a break and a relative displacement of the two sides of the fault. We will consider always brittle fracture either of unbroken material or against friction on a pre-existing fault.

Let us now consider a fault kept locked by friction. An earthquake happens when shear stress overcomes the static friction and displaces the two sides of the fault. After the earthquake the fault is locked again by friction forces. This simple model of the source can be used to define fundamental concepts such as energy dissipation, stress drop, fault slip, seismic moment and magnitude (Fig. 1.7).

Strain elastic energy accumulates in the region surrounding a fault, by tectonic processes; part of this energy is consumed in the production of an earthquake. The *total energy* E_T dissipated in an earthquake is given by

$$E_T = E_F + E_S + E_H \tag{1.1}$$

where E_F is the energy needed to create new fracture or to overcome the friction between the two sides of the fault, E_S is the energy dissipated into the surrounding medium in the form of elastic waves, called the *seismic energy*, and E_H is the energy dissipated by friction

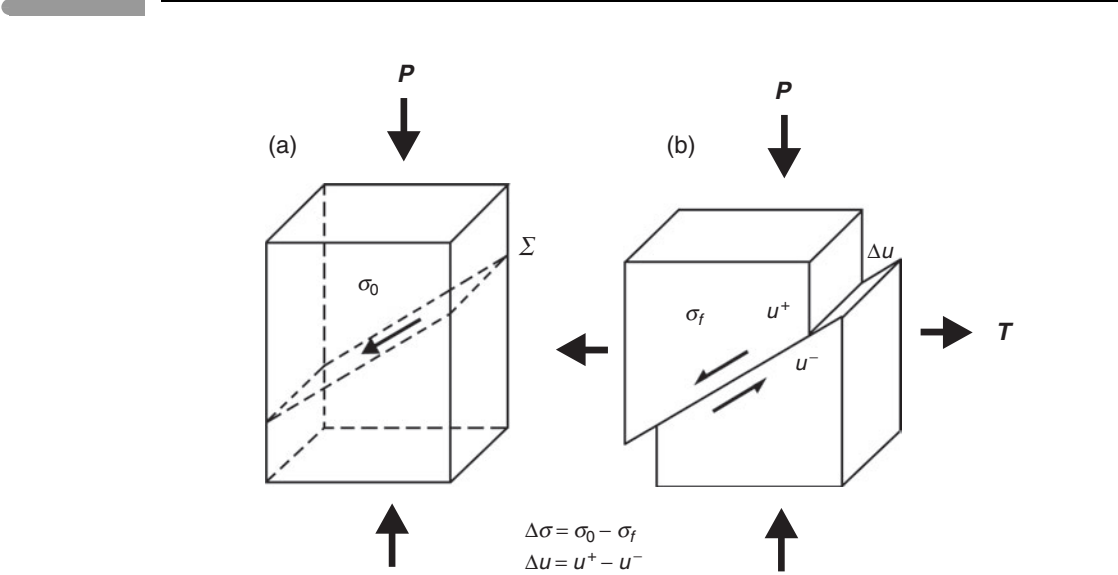


Figure 1.7. Schematic representation of acting stresses, (a) before and (b) after an occurrence on a shear normal fault with slip Δu and stress drop $\Delta\sigma$.

as heat at the fault surface or in non-elastic processes. The units used for energy are joules (SI) or ergs (cgs).

The energy E_S radiated as seismic waves, which propagate with velocity c , can be determined from the amplitudes of waves observed at a distance from the source, correcting first for the amplitude distribution at the source, the geometrical spreading of the wave front and the anelastic attenuation:

$$E_S = \int_{-\infty}^{\infty} c p \dot{u}(t)^2 dt \tag{1.2}$$

The energies E_F and E_H are dissipated during the time that the fracture process lasts, $t_0 < t < t_f$, where t_f is the final time, when motion stops; E_S is finally dissipated when all elastic waves in the Earth are attenuated by anelastic processes. Since the only part we can directly measure by analysing seismograms is the seismic energy E_S , we can express it as a fraction of the total energy, defining the *seismic efficiency coefficient* η by

$$\eta = \frac{E_S}{E_T} \tag{1.3}$$

The seismic efficiency coefficient $\eta < 1$ measures the proportion of the total energy radiated as seismic waves.

From the point of view that earthquakes are caused by shear fracture, the size of an earthquake can be measured by its *seismic moment* M_0 , introduced by Aki (1966) and given by

$$M_0 = \mu \Delta \bar{u} S \tag{1.4}$$

where μ is the shear or rigidity modulus of the material at the fault, $\Delta \bar{u}$ is the spatial average value of the final slip or displacement on the fault surface and S is the area of the fault surface;