



# 5 Seismology and Plate Tectonics

*The acceptance of continental drift has transformed the earth sciences from a group of rather unimaginative studies based on pedestrian interpretations of natural phenomena into a unified science that holds the promise of great intellectual and practical advances.*

J. Tuzo Wilson, *Continents Adrift and Continents Aground*, 1976

## 5.1 Introduction

Two of the major advances in the earth sciences since the 1960s have been the growth of global seismology and the development of our understanding of global plate tectonics. The two are closely intertwined because seismological advances provided some of the crucial data that make plate tectonics the conceptual framework used to think about large-scale processes in the solid earth.

The theory of plate tectonics grew out of the earlier theory of continental drift, proposed in its modern form by Alfred Wegener in 1915. The idea that continents drifted apart was an old one, rooted in the remarkable fit of the coasts of South America and Africa. Still, without compelling evidence for motion between continents, the idea that such motions were physically impossible prevented most geologists from accepting Wegener's ideas. By the 1970s the story was very different. Geologists accepted continental drift in large part because paleomagnetic measurements, based on the geometry and history of the earth's magnetic field, showed that continents had in fact moved over millions of years. Combination of these observations with results from seismology and marine geology and geophysics led to the realization that all parts of the earth's outer shell, not just the continents, were moving.

Plate tectonics is conceptually simple: it treats the earth's outer shell as made up of about 15 rigid plates, about 100 km thick, which move relative to each other at speeds of a few cm per year.<sup>1</sup> The plates are rigid in the sense that little (ideally no) deformation occurs within them, so deformation occurs at their boundaries, giving rise to earthquakes, mountain building, volcanism, and other spectacular phenomena. These strong plates form the earth's *lithosphere*, and move over the

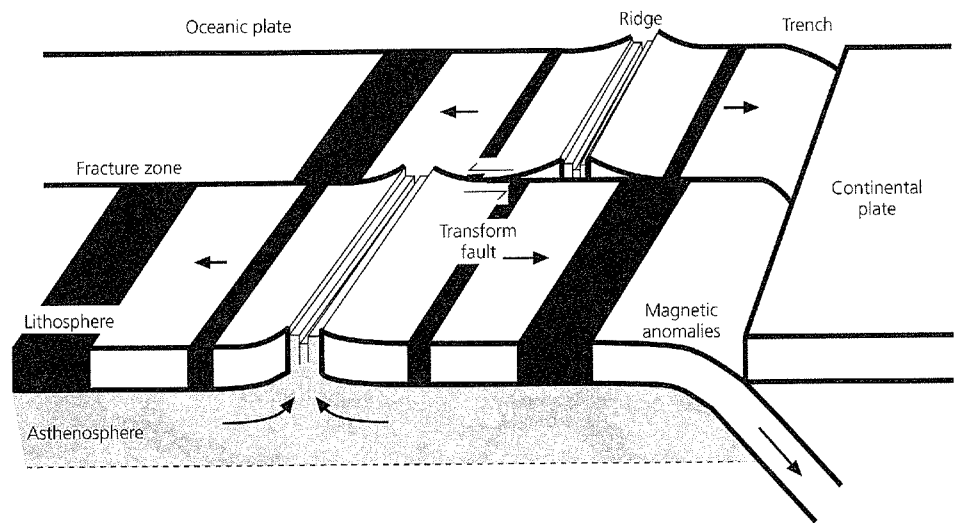
weaker *asthenosphere* below. The lithosphere and asthenosphere are mechanical units defined by their strength and the way they deform. The lithosphere includes both the crust and part of the upper mantle.

Figure 5.1-1 shows the three basic types of plate boundaries. Warm mantle material upwells at *spreading centers*, also known as mid-ocean ridges, and then cools. Because the strength of rock decreases with temperature (Section 5.7.3), the cooling material forms strong plates of new oceanic lithosphere. The cooling oceanic lithosphere moves away from the ridges, and eventually reaches *subduction zones*, or trenches,<sup>2</sup> where it descends in *downgoing slabs* back into the mantle, reheating as it goes. The direction of the relative motion between two plates at a point on their common boundary determines the nature of the boundary. At spreading centers both plates move away from the boundary, whereas at subduction zones the subducting plate moves toward the boundary. At the third boundary type, *transform faults*, relative plate motion is parallel to the boundary.

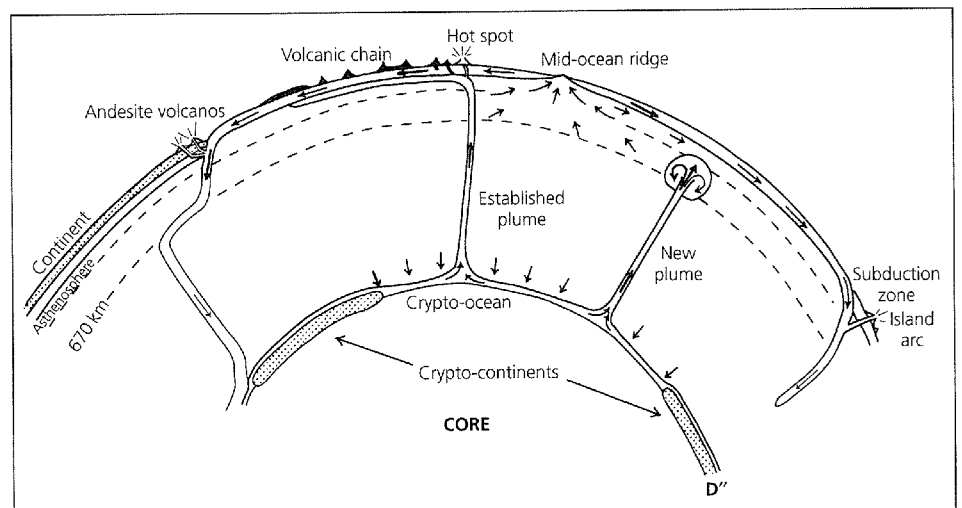
As discussed in Section 3.8, seismology shows that the structure of the mantle and the core varies with depth, due to changes in temperature, pressure, mineralogy, and composition. Plate tectonics describes the behavior of the lithosphere, the strong outer shell of the mantle, which is the cold outer boundary layer of the thermal convection system involving the mantle and the core that removes heat from the earth's interior. Although much remains to be learned about this convective system, especially in the lower mantle and the core (Fig. 5.1-2), there is general agreement that at shallow depths the warm,

<sup>1</sup> This is about the speed at which fingernails grow.

<sup>2</sup> Boundaries are described either as mid-ocean ridges and trenches, emphasizing their morphology, or as spreading centers and subduction zones, emphasizing the plate motion there. The latter nomenclature is more precise, because there are elevated features in the ocean basins that are not spreading ridges, and spreading centers like the East African rift exist within continents.



**Fig. 5.1-1** Plate tectonics at its simplest. Oceanic lithosphere is formed at ridges and subducted at trenches. At transform faults, plate motion is parallel to the boundaries. Each boundary type has typical earthquakes.



**Fig. 5.1-2** Schematic diagram showing ideas about mantle convection. Ridges reflect upper mantle upwelling. Slabs penetrate into the lower mantle, causing heterogeneity there, and in some cases descend to the base of the mantle. Mantle (hot spot) plumes reflect lower mantle upwelling. Many features shown are controversial and subject to change without notice. (Modified from Stacey, 1992.)

and hence less dense, material rising below spreading centers forms upwelling limbs, whereas the relatively cold, and hence dense, subducting slabs form downwelling limbs. Although the lithosphere is a very thin layer compared to the rest of the mantle (100 km is 1/29 of the mantle's radius), it is where the greatest temperature change occurs, from about 1300° to 1400°C at a depth of 100 km to about 0°C at the surface. For this reason, the lithosphere is called a thermal boundary layer. Because of this temperature change, the lithosphere is much stronger than the underlying rock, and so is also a mechanical boundary layer. This strong boundary layer is thought to be a primary reason why plate tectonics is much more complicated than expected from simple convection models. Moreover, the lithosphere, which contains the crust, is also a chemical boundary layer distinct from the remainder of the mantle. Continental lithosphere is especially distinct: although individual plates can contain both oceanic and continental lithosphere, the latter is made of less dense rock than the former (recall the

differences between granitic and basaltic rocks discussed in Section 3.2), and so does not subduct. The oceanic lithosphere is continuously subducted and reformed at ridges, and so never gets older than about 200 Myr. The continental lithosphere, however, can be billions of years old.

Put another way, plate tectonics is the primary surface manifestation of the heat engine whose nature and history govern the planet's thermal, mechanical, and chemical evolution.<sup>3</sup> Earth's heat engine is characterized by the balance between three modes of heat transfer from the interior: the plate tectonic cycle involving the cooling of oceanic lithosphere; mantle plumes, which are thought to be a secondary feature of mantle convection; and heat conduction through continents that are not subducted and hence do not participate directly in the oceanic plate tectonic cycle. Based on estimates from sea floor topography and heat flow, discussed shortly, terrestrial heat

<sup>3</sup> It has been said that heat is the geological lifeblood of planets.

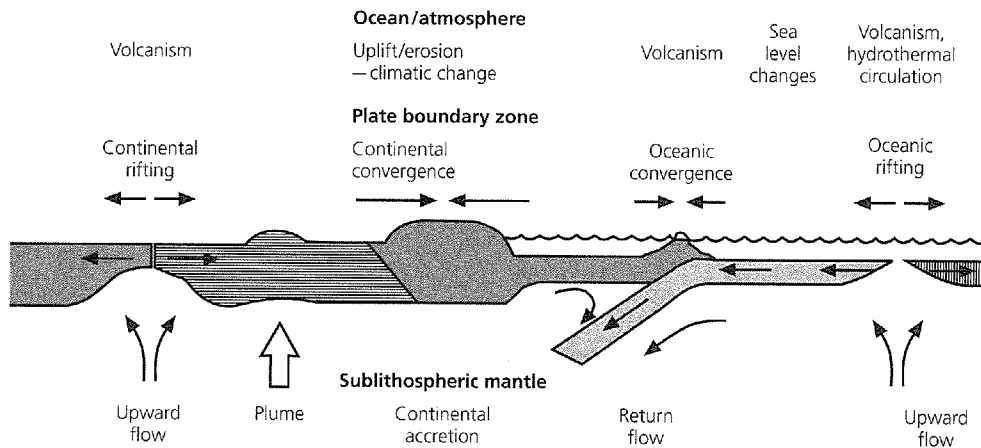


Fig. 5.1-3 Cartoon summarizing some of the primary modes of interaction between the solid earth's interior and the fluid ocean and atmosphere system. (Stein *et al.*, 1995. *Seafloor Hydrothermal Systems*, 425–45, copyright by the American Geophysical Union.)

loss seems to occur primarily (about 70%) via plate tectonics, with about 5% via hot spots (mantle plumes). By contrast, Earth's grossly similar sister planets, Mars and Venus, seem to function quite differently, because large-scale plate tectonics appears absent, at least at present.

Plate tectonics is also crucial for the evolution of Earth's ocean and atmosphere, because it involves many of the primary means (including volcanism, hydrothermal circulation through cooling oceanic lithosphere, and the cycle of uplift and erosion) by which the solid earth interacts with the ocean and the atmosphere (Fig. 5.1-3). The chemistry of the oceans and the atmosphere depends in large part on plate tectonic processes, and many long-term features of climate are influenced by mountains that are uplifted by plate convergence and the positions of continents that control ocean circulation. In fact, the presence of plate tectonics may explain how life evolved on earth (at mid-ocean ridge hot springs) and be crucial for its survival (the atmosphere is maintained by plate boundary volcanism, and plate tectonics raises the continents above sea level).

As a result, plate tectonics is heavily studied by earth scientists. Our goal in this chapter is to introduce some of the ways in which seismology contributes to these studies. Some sources for more general and more detailed treatments of these topics are listed at the end of the chapter.

Seismology plays several key roles in our studies of plate tectonics. The distribution of earthquakes provides strong evidence for the idea of essentially rigid plates, with deformation concentrated on their boundaries. Figure 5.1-4 shows maps of global seismicity covering the time period 1964–97. Such maps did not become available until the early 1960s, when the World Wide Standardized Seismographic Network (WWSSN) allowed accurate locations for earthquakes of magnitude 5 or greater anywhere in the world. The map shows several remarkable patterns.

The mid-ocean ridge system, where the oceanic lithosphere is created, is beautifully outlined by the earthquake locations. For example, the Mid-Atlantic ridge and East Pacific rise can be followed using epicenters for thousands of kilometers. The locations of the trenches, where oceanic lithosphere is subducted,

are even more apparent in the lower panel showing earthquakes with focal depths greater than 100 km, because mid-ocean ridge earthquakes are shallow and thus do not appear.

It is especially impressive to plot the locations of earthquakes on cross-sections across trenches (Fig. 5.1-5). Inclined zones of seismicity delineate the subducting oceanic plates, which travel time and attenuation studies show to be colder and stronger than the surrounding mantle. These zones, identified before their plate tectonic significance became clear, are known as *Wadati-Benioff zones* after their discoverers.<sup>4</sup>

The *interplate* earthquakes both delineate plate boundaries and show the motion occurring there. We will see that the direction of faulting reflects the spreading at mid-ocean ridges and subduction at trenches. The earthquake locations and mechanisms also show that plate boundaries in continents are often complicated and diffuse, rather than the simple narrow boundaries assumed in the rigid plate model that are a good approximation to what we see in the oceans. For example, seismicity shows that the collision of the Indian and Eurasian plates creates a deformation zone which includes the Himalayas but extends far into China. Similarly, the northward motion of the Pacific plate with respect to North America creates a broad seismic zone, indicating that the plate boundary zone spans much of the western USA and Canada.

In addition, *intraplate* earthquakes occur within plate interiors, far from boundary zones. For example, Fig. 5.1-4 shows earthquakes in eastern Canada and central Australia. Such earthquakes are much rarer than plate boundary zone earthquakes, but are common enough to indicate that plate interiors are not perfectly rigid. In some cases these earthquakes are associated with intraplate volcanism, as in Hawaii. Intraplate earthquakes are studied to provide data about where and how the plate tectonic model does not fully describe tectonic processes.

<sup>4</sup> Kiyoo Wadati (1902–95) discovered the existence of deep seismicity and its geometry under Japan; Hugo Benioff (1899–1968), also known for important contributions to seismological instrumentation, discussed the global nature of deep earthquakes and their relation to surface features (Fig. 1.1-10).

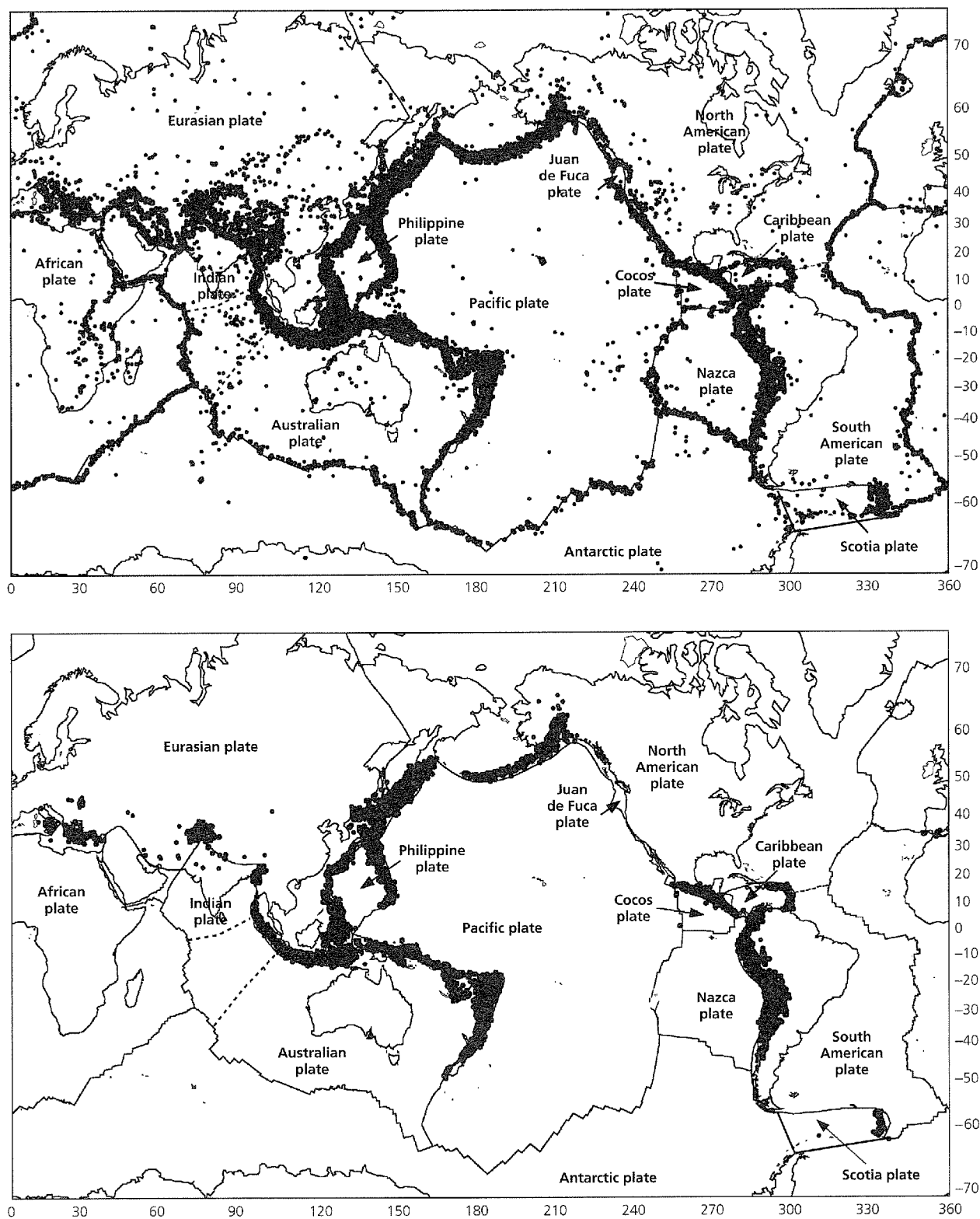


Fig. 5.1-4 Global seismicity (1964–97). *Top*: Earthquakes ( $m_b \geq 5$ , all depths) clearly delineate most plate boundaries, and show that some (e.g., India-Eurasia) are diffuse. Many intraplate earthquakes show internal plate deformation. *Bottom*: The locations of seismicity (of all magnitudes) below 100 km indicate the subduction zones.



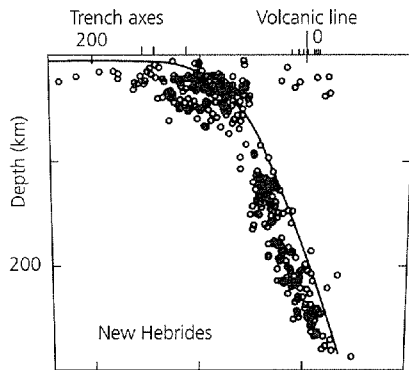


Fig. 5.1-5 Seismicity cross-section perpendicular to the New Hebrides trench showing the Wadati–Benioff zone. This dipping plane of earthquakes indicates the position of the subducting plate. (Isacks and Barazangi, 1977. *Island Arcs, Deep Sea Trenches and Back Arc Basins*, 99–114, copyright by the American Geophysical Union.)

In summary, seismology provides crucial information about both *plate kinematics*, the directions and rates of plate motions, and *plate dynamics*, the forces causing plate motions. As we will see, seismicity is one of the major tools used to identify and delineate plate boundary zones, and earthquake mechanisms are among the primary data used to determine the motion within plate boundary zones. The mechanisms also provide information about the stresses acting at plate boundaries and within plates, which, together with earthquake depths and seismic velocity structure, are important in developing ideas about the forces involved and the physical processes by which rocks deform and cause earthquakes. Conversely, plate motion data are used to draw inferences about the locations and times of future earthquakes and their societal risks. Thus it is often hard, and sometimes pointless, to decide where seismology ends and plate tectonics begins, or vice versa.

## 5.2 Plate kinematics

Understanding the distribution and types of earthquakes requires an understanding of the geometry of plate motions, or plate kinematics. In this section we sketch some basic results, of which we assume most readers have some knowledge. As full exploration of this topic is beyond our scope, readers are encouraged to delve into the suggested literature.

### 5.2.1 Relative plate motions

A basic principle of plate tectonics is that the relative motion between any two plates can be described as a rotation about an *Euler pole*<sup>1</sup> (Fig. 5.2-1). This condition controls the types of boundaries and the focal mechanisms of earthquakes resulting from relative motions, as discussed later. Specifically, at any

<sup>1</sup> This term comes from Euler's theorem, which states that the displacement of any rigid body (in this case, a plate) with one point (in this case, the center of the earth) fixed is a rotation about an axis.

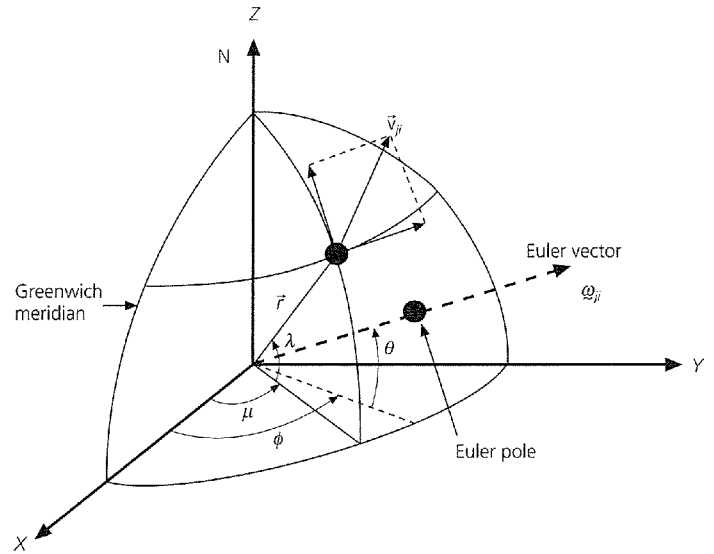


Fig. 5.2-1 Geometry of plate motions. Linear velocity at point  $\mathbf{r}$  is given by  $\mathbf{v}_{ji} = \boldsymbol{\omega}_{ji} \times \mathbf{r}$ . The Euler pole is the intersection of the Euler vector with the earth's surface. Note that west longitudes and south latitudes are negative.

point  $\mathbf{r}$  along the boundary between plate  $i$  and plate  $j$ , with latitude  $\lambda$  and longitude  $\mu$ , the *linear velocity* of plate  $j$  with respect to plate  $i$  is

$$\mathbf{v}_{ji} = \boldsymbol{\omega}_{ji} \times \mathbf{r}. \quad (1)$$

This is the usual formulation for rigid body rotations in mechanics.  $\mathbf{r}$  is the position vector to the point on the boundary, and  $\boldsymbol{\omega}_{ji}$  is the angular velocity vector, or *Euler vector*. Both vectors are defined from an origin at the center of the earth.

The direction of relative motion at any point on the boundary is a small circle, a parallel of latitude *about the Euler pole* (not a geographic parallel about the North Pole!). For example, in Fig. 5.2-2 (*top*) the pole shown is for the motion of plate 2 with respect to plate 1. The convention used is that the first named plate ( $j = 2$ ) moves counterclockwise (in a right-handed sense) about the pole with respect to the second named plate ( $i = 1$ ). The segments of the boundary where relative motion is parallel to the boundary are transform faults. Thus transforms are small circles about the pole, and earthquakes occurring on them should have pure strike-slip mechanisms. Other segments have relative motion away from the boundary, and are thus spreading centers. Figure 5.2-2 (*bottom*) shows an alternative case. The pole here is for plate 1 ( $j = 1$ ) with respect to plate 2 ( $i = 2$ ), so plate 1 moves toward some segments of the boundary, which are subduction zones.

The magnitude, or rate, of relative motion increases with distance from the pole because

$$|\mathbf{v}_{ji}| = |\boldsymbol{\omega}_{ji}| |\mathbf{r}| \sin \gamma, \quad (2)$$

where  $\gamma$  is the angle between the Euler pole and the site (corresponding to a colatitude about the pole). All points on a plate

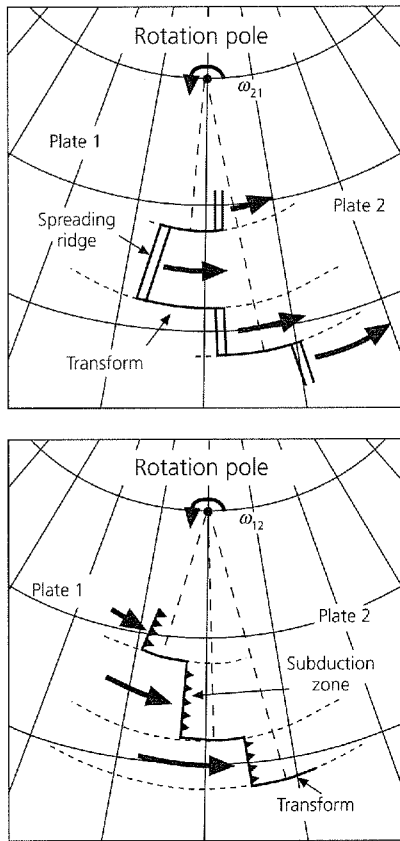


Fig. 5.2-2 Relationship of motions on plate boundaries to the Euler pole. Relative motions occur along small circles about the Euler pole (short dashed lines) at a rate that increases with distance from the pole. Note the difference the sense of rotation makes:  $\omega_i$  is the Euler vector corresponding to the rotation of plate  $j$  counterclockwise with respect to  $i$ .

boundary have the same angular velocity, but the magnitude of the linear velocity varies from zero at the pole to a maximum  $90^\circ$  away.

The components of the vectors can be written in Cartesian  $(x, y, z)$  coordinates (Fig. 5.2-1). The position vector is

$$\mathbf{r} = (a \cos \lambda \cos \mu, a \cos \lambda \sin \mu, a \sin \lambda), \quad (3)$$

where  $a$  is the earth's radius. Similarly, if the Euler pole is at latitude  $\theta$  and longitude  $\phi$ , the Euler vector is written (neglecting the  $ij$  subscripts for simplicity) as

$$\boldsymbol{\omega} = (|\boldsymbol{\omega}| \cos \theta \cos \phi, |\boldsymbol{\omega}| \cos \theta \sin \phi, |\boldsymbol{\omega}| \sin \theta), \quad (4)$$

where the magnitude,  $|\boldsymbol{\omega}|$ , is the scalar angular velocity or rotation rate. To find the Cartesian components of the linear velocity  $\mathbf{v}$ , we evaluate the cross product (Eqn 1) using its definition (Eqn A.3.28), and find

$$\mathbf{v} = (v_x, v_y, v_z),$$

$$v_x = a |\boldsymbol{\omega}| (\cos \theta \sin \phi \sin \lambda - \sin \theta \cos \lambda \sin \mu)$$

$$\begin{aligned} v_y &= a |\boldsymbol{\omega}| (\sin \theta \cos \lambda \cos \mu - \cos \theta \cos \phi \sin \lambda) \\ v_z &= a |\boldsymbol{\omega}| \cos \theta \cos \lambda \sin (\mu - \phi). \end{aligned} \quad (5)$$

At the point  $\mathbf{r}$ , the north-south and east-west unit vectors can be written in terms of their Cartesian components using Eqn A.7.4,

$$\begin{aligned} \hat{\mathbf{e}}^{\text{NS}} &= (-\sin \lambda \cos \mu, -\sin \lambda \sin \mu, \cos \lambda), \\ \hat{\mathbf{e}}^{\text{EW}} &= (-\sin \mu, \cos \mu, 0), \end{aligned} \quad (6)$$

so we find the north-south and east-west components of  $\mathbf{v}$  by taking dot products of its Cartesian components (Eqns 5) with the unit vectors (Eqns 6), and obtain

$$\begin{aligned} v^{\text{NS}} &= a |\boldsymbol{\omega}| \cos \theta \sin (\mu - \phi), \\ v^{\text{EW}} &= a |\boldsymbol{\omega}| [\sin \theta \cos \lambda - \cos \theta \sin \lambda \cos (\mu - \phi)]. \end{aligned} \quad (7)$$

We can then find the rate and direction of plate motion,

$$\begin{aligned} \text{rate} &= |\mathbf{v}| = \sqrt{(v^{\text{NS}})^2 + (v^{\text{EW}})^2} \\ \text{azimuth} &= 90^\circ - \tan^{-1} [(v^{\text{NS}})/(v^{\text{EW}})], \end{aligned} \quad (8)$$

such that azimuth is measured in the usual convention, degrees clockwise from North.

In evaluating these expressions, it is important to be careful with dimensions. Although rotation rates are typically reported in degrees per million years, they should be converted to radians per year. The resulting linear velocity will have the same dimensions as Earth's radius. By serendipity, converting radius in km to mm and Myr to years cancel out, so only the degrees to radians ( $\times \pi/180^\circ$ ) conversion actually needs to be done to obtain a linear velocity in mm/yr. Plate motions are often quoted as mm/yr, because a year is a comfortable unit of time for humans and 1 mm/yr corresponds to 1 km/Myr, making it easy to visualize what seemingly slow plate motion accomplishes over geologic time.

To see how this works, consider Fig. 5.2-3, which shows the North America-Pacific boundary zone. The map is drawn in a projection about the Euler pole, so the expected relative motion is parallel to small circles like the one shown. By analogy to Fig. 5.2-2, this geometry predicts NW-SE-oriented spreading along ridge segments in the Gulf of California, which are rifting Baja California away from the rest of Mexico. Further north, the San Andreas fault system is essentially parallel to the relative motion, so is largely a transform fault. In Alaska, the eastern Aleutian arc is perpendicular to the plate motion, so the Pacific plate subducts beneath North America. Thus this plate boundary contains ridge, transform, and trench portions, depending on the geometry of the boundary.<sup>2</sup> In addition, the

<sup>2</sup> A good way to visualize the plate motion is to photocopy Fig. 5.2-3, cut along the boundary of the Pacific plate, and then photocopy the "Pacific" onto another piece of paper. Putting the "Pacific" beneath "North America" and rotating around a thumbtack through the pole shows the ridge, transform, and trench motions both forward and backward in time.

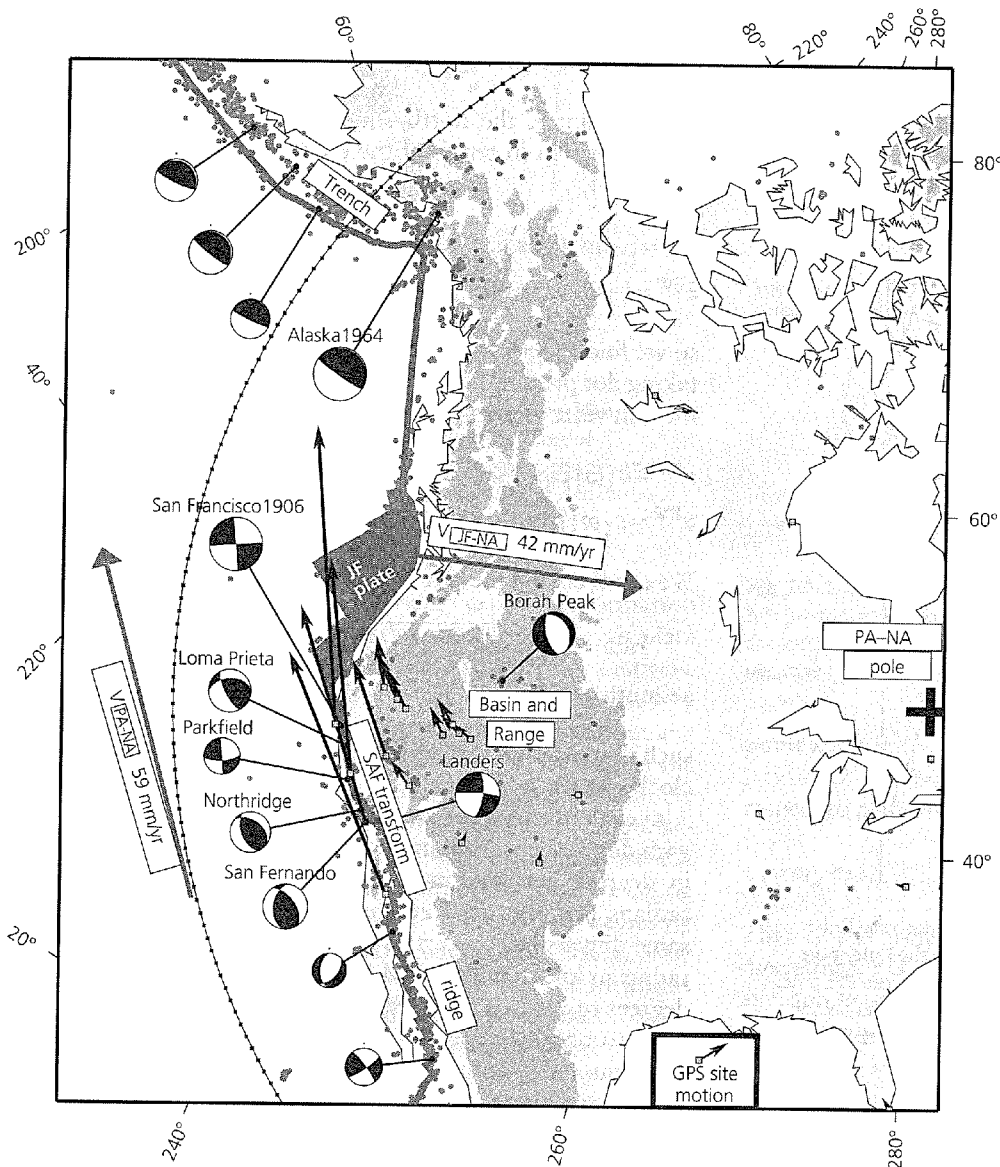


Fig. 5.2-3 Geometry and focal mechanisms for a portion of the North America-Pacific boundary zone that also includes the small Juan de Fuca (JF) plate. The map projection is about the Pacific-North America Euler pole, so the line with dots shows a small circle and thus the direction of plate motion. This small circle is further from the pole than the San Andreas fault, so the rate of motion on it is larger. The variation in the boundary type along its length from extension, to transform, to convergence, is shown by the focal mechanisms. The diffuse nature of the boundary zone is shown by seismicity (small dots), focal mechanisms, topography (elevation above 1000 m is shaded), and vectors showing the motion of GPS and VLBI sites (squares) (Bennett *et al.*, 1999) with respect to the stable interior of North America. The velocity scale is shown by the plate motion arrows; some site motion vectors are too small to be seen. (Stein and Klosko, 2002. From *The Encyclopedia of Physical Science and Technology*, ed. R. A. Meyers, copyright 2002 by Academic Press, reproduced by permission of the publisher.)

boundary zone contains the small Juan de Fuca plate, which subducts beneath the Pacific Northwest at the Cascadia subduction zone.

Equation 8 lets us find how the motion varies. The predicted motion of the Pacific plate with respect to the North American plate at a point on the San Andreas fault ( $36^{\circ}\text{N}$ ,  $239^{\circ}\text{E}$ ) has a rate of 46 mm/yr at an azimuth of  $\text{N}36^{\circ}\text{W}$ . The predicted direction agrees reasonably well with the average trend of the San Andreas fault,  $\text{N}41^{\circ}\text{W}$ . Thus, to first order, the San Andreas is a Pacific-North America transform plate boundary with right-lateral motion. However, there are some deviations from pure transform behavior. As we will see, the rate on the San Andreas fault is less than the total plate motion because some of the motion occurs elsewhere within the broad plate boundary zone. In addition, in some places the San Andreas trend differs enough from the plate motion direction that dip-

slip faulting occurs. Hence we think of the San Andreas as the primary feature of the essentially strike-slip portion of the plate boundary zone.

Similarly, at a point on the Aleutian trench near the site of the great 1964 Alaska earthquake (Fig. 4.3-15) ( $62^{\circ}\text{N}$ ,  $212^{\circ}\text{E}$ ), we predict Pacific motion of 53 mm/yr at  $\text{N}14^{\circ}\text{W}$  with respect to North America. This motion is into the trench, which is a Pacific-North America subduction zone. It is worth noting that for a given convergent relative motion either plate can be subducting. However, the relative direction is important, so the plates cannot be interchanged: if  $\text{N}14^{\circ}\text{W}$  were the direction of motion of North America with respect to the Pacific, the motion would be away from the boundary, which would then be a spreading center with the same rate. As for the San Andreas, the actual boundary zone shown by earthquakes and other deformation is wider and more complicated than the ideal.

Earthquake focal mechanisms within the boundary zone are consistent with the overall plate motions and illustrate some of their complexities. In the Gulf of California we see both strike-slip faulting along oceanic transforms and normal faulting on ridge segments. The San Andreas fault system, composed of the main fault and some others, has both pure strike-slip earthquakes (Parkfield) and earthquakes with some dip-slip motion (Northridge (Section 4.5.3), San Fernando, and Loma Prieta) when it deviates from pure transform behavior. The seismicity also shows that the plate boundary zone is quite broad. Although the San Andreas fault system is the locus of most of the plate motion (Fig. 4.5-13) and hence large earthquakes, seismicity extends as far eastward as the Rocky Mountains. For example, the Landers earthquake shows strike-slip motion east of the San Andreas, and the Borah Peak earthquake illustrates the extensional faulting that occurs in the Basin and Range. These focal mechanisms are consistent with the motions shown by space-based geodetic measurements, discussed shortly, and with geologic studies.

### 5.2.2 Global plate motions

The relative plate motions show how the plate boundary geometry is evolving and has evolved. The Juan de Fuca plate is subducting under North America faster than new lithosphere is being added to it by sea floor spreading at its boundary with the Pacific plate, so this plate was larger in the past and is shrinking. Rotating the Pacific plate backwards with respect to North America shows that 10 million years ago the Gulf of California had not yet begun to open by sea floor spreading. These changes are part of the evolution of the plate boundary in western North America, in which the large oceanic Farallon plate that used to be between the Pacific and North American plates began subducting under North America at about 40 Ma,<sup>3</sup> leaving the Juan de Fuca plate as a remnant and forming the San Andreas fault.

At this point you may be wondering how Euler poles are found. Until recently, this was done by combining three different types of data from different boundaries. The rates of spreading are found from sea floor magnetic anomalies, which form as the hot rock at ridges cools and acquires magnetization parallel to the earth's magnetic field. Because the history of reversals of the earth's magnetic field is known, the anomalies can be dated, so their distance from the ridge where they formed shows how fast the sea floor moved away from the ridge. The directions of motion are found from the orientations of transform faults and the slip vectors of earthquakes on transforms and at subduction zones. Euler vectors are found from the relative motion data, using geometrical conditions we have discussed. The process is easy to visualize. Because slip vectors and transform faults lie on small circles about the pole, the pole must lie on a great circle at right angles to them (Fig. 5.2-2). Similarly, the rate of plate motion increases with the sine of

the distance from the pole (Eqn 2). These constraints make it possible to locate the poles. Determination of Euler vectors for all the plates can thus be treated as an overdetermined least squares problem whose solution (Section 7.5) gives a *global relative plate motion model*. Because these models use spreading rates determined from magnetic anomaly data that span several million years, they describe plate motions averaged over the past few million years.<sup>4</sup>

Table 5.2-1 gives such a model, known as NUVEL-1A,<sup>5</sup> which specifies the motions of plates (Fig. 5.2-4) with respect to North America. The vectors follow the convention that each named plate moves counterclockwise relative to North America. Although the table lists only Euler vectors with respect to North America, the motion of plates with respect to other plates is easily found using vector arithmetic. For example,

$$\omega_{ij} = -\omega_{ji}, \quad (9)$$

so we reverse the plate pair using the negative of the Euler vector. The pole for the new plate pair is the antipole, with latitude of opposite sign and longitude increased by 180°. The magnitude (rotation rate) stays the same. We can also reverse the plate pair by keeping the same pole and making the rotation rate negative (clockwise rather than counterclockwise). Although we usually use positive rotation rates, negative ones sometimes help us visualize the motion. For example, the table shows the Pacific–North America pole at about 49°N, 102°E, so the North America–Pacific pole is at about 49°N, (102 + 180 = 282)°E, which is in southeastern Canada. Thus, about this pole, North America rotates counterclockwise with respect to the Pacific, or the Pacific rotates clockwise with respect to North America, as shown in Fig. 5.2-3.

For other plate pairs we assume that the plates are rigid, so all motion occurs at their boundaries. We can then add Euler vectors,

$$\omega_{jk} = \omega_{ji} + \omega_{ik} \quad (10)$$

because the motion of plate  $j$  with respect to plate  $k$  equals the sum of the motion of plate  $j$  with respect to plate  $i$  and the motion of plate  $i$  with respect to plate  $k$ . Thus if we start with a set of vectors all with respect to one plate, e.g.,  $i$ , we use

$$\omega_{jk} = \omega_{ji} - \omega_{ki} \quad (11)$$

to form any Euler vector needed. These operations are easily done using the Cartesian components (Eqn 4), as shown in this chapter's problems. We can also perform the analogous operations on linear velocity vectors at a specific site.

<sup>4</sup> The most recent magnetic reversal occurred about 780,000 years ago, so any plate model based on paleomagnetic data must average at least over that interval.

<sup>5</sup> NUVEL-1 (Northwestern University VELOCITY) was developed as a new ("nouvelle") model (DeMets *et al.*, 1990). The multiyear development prompted the suggestion that "OLDVEL" might be a better name. Due to changes in the paleomagnetic time scale the model was revised to NUVEL-1A (DeMets *et al.*, 1994). This change caused a slight difference in the rates of relative motion, but not in the poles and hence directions of relative motion.

<sup>3</sup> "Ma" is often used to denote millions of years before the present.

Table 5.2-1 Euler vectors with respect to North America (NA).

Plate	Pole latitude (°N)	Longitude (°E)	$ \omega $ (°/Myr)
Pacific (PA)	-48.709	101.833	0.7486
Africa (AF)	78.807	38.279	0.2380
Antarctica (AN)	60.511	119.619	0.2540
Arabia (AR)	44.132	25.586	0.5688
Australia (AU)	29.112	49.006	0.7579
Caribbean (CA)	74.346	153.892	0.1031
Cocos (CO)	27.883	-120.679	1.3572
Eurasia (EU)	62.408	135.831	0.2137
India (IN)	43.281	29.570	0.5803
Nazca (NZ)	61.544	-109.781	0.6362
South America (SA)	-16.290	121.876	0.1465
Juan de Fuca (JF)	-22.417	67.203	0.8297
Philippine (PH)	-43.986	-19.814	0.8389
Rivera (RI)	22.821	-109.407	1.8032
Scotia (SC)	-43.459	123.120	0.0925
NNR*	2.429	93.965	0.2064

Source: After DeMets *et al.* 1994.

\*No net rotation, defined in Section 5.2.4.

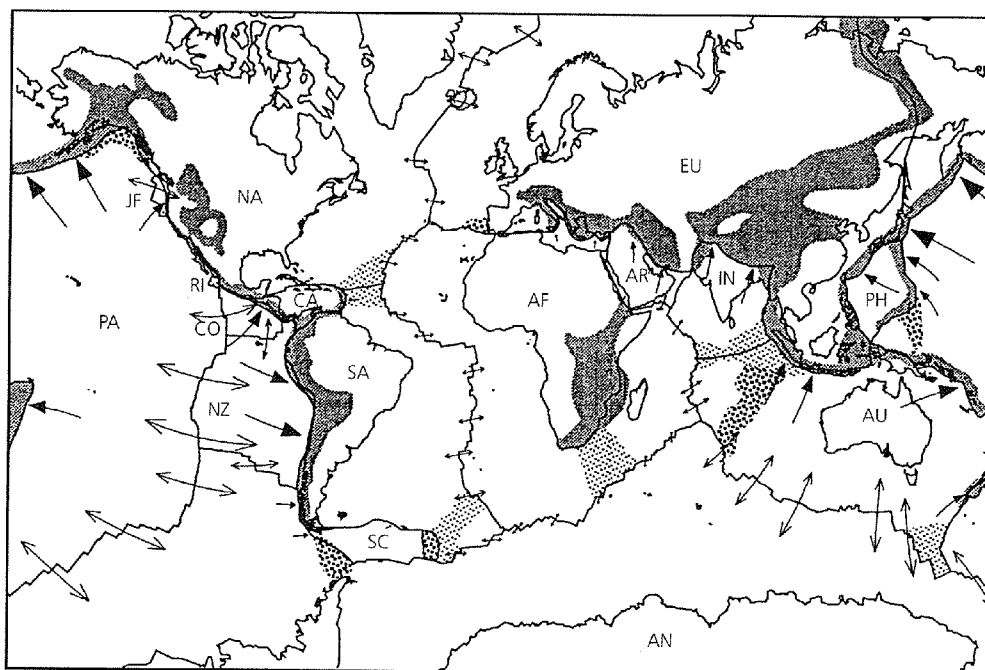


Fig. 5.2-4 Relative plate motions for the NUVEL-1 global plate motion model. Arrow lengths are proportional to the displacement if plates maintain their present relative velocity for 25 Myr. Divergence across mid-ocean ridges is shown by diverging arrows. Convergence is shown by single arrows on the underthrust plate. Plate boundaries are shown as diffuse zones implied by seismicity, topography, or other evidence of faulting. Fine stipple shows mainly subaerial regions where the deformation has been inferred from seismicity, topography, other evidence of faulting, or some combination of these. Medium stipple shows mainly submarine regions where the nonclosure of plate circuits indicates measurable deformation; in most cases these zones are also marked by earthquakes. Coarse stipple shows mainly submarine regions where the deformation is inferred mostly from the presence of earthquakes. The geometry of these zones, and in some cases their existence, is under investigation. (Gordon and Stein, 1992. *Science*, 256, 333–42, copyright 1992 American Association for the Advancement of Science.)

Such vector addition is important because we only have certain types of data for individual boundaries (Fig. 5.2-5). Although spreading centers provide rates from the magnetic anomalies and azimuths from both transform faults and slip

vectors, only the direction of motion is directly known at subduction zones. As a result, convergence rates at subduction zones are estimated by global closure, combining data from all plate boundaries (Section 7.5). Thus the predicted rate at which

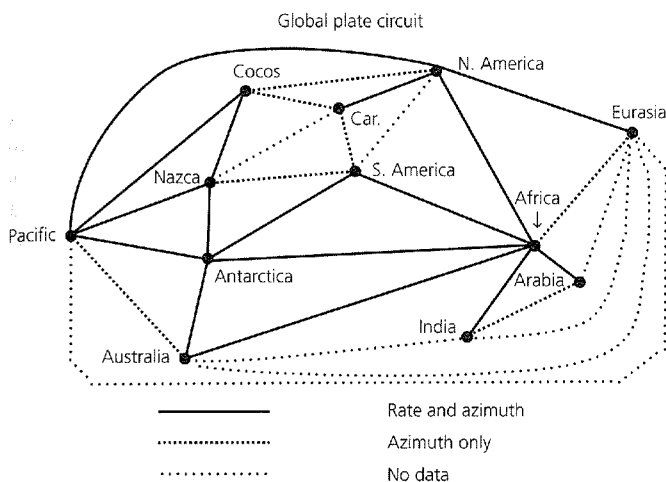


Fig. 5.2-5 Global plate circuit geometry for the NUVEL-1 plate motion model. Relative motion data are used on the boundaries indicated. (De Mets *et al.*, 1990. *Geophys. J. Int.*, 101, 425–78.)

the Cocos plate subducts beneath North America, causing large earthquakes in Mexico, depends on the measured rates of Cocos–Pacific spreading on the East Pacific rise and Pacific–North America spreading in the Gulf of California. In some cases, such as relative motion between North and South America, no direct data were used because the boundary location and geometry are unclear, so the relative motion is inferred entirely from closure. Not surprisingly, the motions of plate pairs based on both rate and azimuth data appear to be better known.

Figure 5.2-4 shows the predicted relative motions at plate boundaries around the world. As shown for the Pacific–North America boundary in Fig. 5.2-3 and discussed in general terms in later sections, the predicted motions correspond to the earthquake mechanisms. Moreover, we can use the plate motions to make inferences about future earthquakes. For example, even though we do not have seismological observations of large earthquakes along the boundary between the Juan de Fuca and North American plates, the plate motions predict that such earthquakes could result from the subduction of the Juan de Fuca plate beneath North America. Evidence for this subduction is given by the presence of the Cascade volcanoes (such as Mount Saint Helens and Mount Rainer) and paleoseismic records (Section 1.2.5) that are interpreted as evidence of large past earthquakes.

Figure 5.2-4 also illustrates that boundaries between plates are often diffuse. Seismicity, active faulting, and elevated topography often indicate a broad zone of deformation between plate interiors. This effect is evident in continental lithosphere, such as the India–Eurasia collision zone in Asia or the Pacific–North America boundary zone in the western USA, but can also sometimes be seen in oceanic lithosphere, as in the Central Indian Ocean. Plate boundary zones cover about 15% of the earth's surface, and about 40% of the earth's population lives within them.

Earthquakes are among the best tools for investigating plate boundary zones and other deviations from plate rigidity. They provide one of the best indicators of the location of boundary zones, so new earthquakes often change our views. We also use plate motion data, many of which are earthquake slip vectors. For example, Fig. 5.2-4 shows zones of seismicity in the Central Indian Ocean (Section 5.5.2) as boundaries between distinct Indian and Australian plates, rather than as within a single Indo-Australian plate, because spreading rates along the Central Indian Ocean ridge are better fit by a two-plate model. A similar argument justifies the assumption of a small Rivera plate distinct from the Cocos plate. Another approach is to use the global plate circuit closures (Fig. 5.2-5). Recall that forming a Euler vector from two others (Eqn 10) assumes that all three plates are rigid. Hence this assumption can be used to test for deviations from rigidity. To do this, we form a *best-fitting vector* for a plate pair, using only data from that pair of plates' boundary, and a *closure fitting vector* from data elsewhere in the world. If the plates were rigid, the two vectors would be the same. However, a significant difference between the two indicates a deviation from rigidity, or another problem with the plate motion model. For example, such analysis shows systematic deviations along some subduction zones, suggesting that the slip vectors of the trench earthquakes do not exactly reflect plate motions because a sliver of forearc material in the overriding plate moves separately from the remainder of the overriding plate (Section 5.4.3).

A variant of this approach is to examine the Euler vectors for three plates that meet at a *triple junction*, compute best-fitting Euler vectors for each of the three plate pairs, and sum them. For rigid plates, Eqn 10 shows that the sum should be zero. However, when this was done for the junction in the Central Indian Ocean, assuming that it was where the African, Indo-Australian, and Antarctic plates met, the Euler vector sum differed significantly from zero, indicating deviations from plate rigidity. As plate motion data improve, it seems that what was treated as a three-plate system may include as many as six resolvable plates (Antarctica, distinct Nubia (West Africa) and Somalia (East Africa), India, Australia, and Capricorn (between India and Arabia)). Hence models of plate boundaries and motions improve with time (Fig. 1.1-9). For example, although the model in Fig. 5.2-4 has a single African plate, recent models seek to resolve the motion between Nubia and Somalia (Fig. 5.6-2).

### 5.2.3 Space-based geodesy

New plate motion data have become available in recent years due to the rapidly evolving techniques of space-based geodesy. Using space-based measurements to determine plate motions was suggested by Alfred Wegener when he proposed the theory of continental drift in 1915. Wegener realized that proving continents moved apart was a formidable challenge. Although geodesy — the science of measuring the shape of, and distances on, the earth — was well established, standard surveying

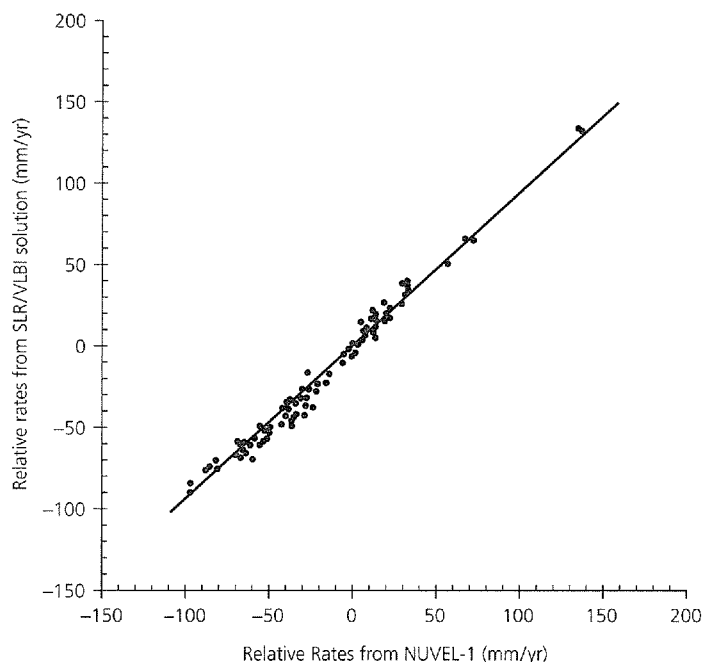


Fig. 5.2-6 Comparison of rates determined by space geodesy with those predicted by the NUVEL-1 global plate motion model. The space geodetic rates are determined from sites located away from plate boundaries to reduce the effects of deformation near the boundaries. The slope of the line is 0.94, indicating that plate motions over a decade are very similar to those predicted by a model averaging over 3 million years. (Robbins *et al.*, 1993. *Contributions of Space Geodesy to Geodynamics*, 21–36, copyright by the American Geophysical Union.)

methods offered no hope of measuring slow motions between continents far apart. Wegener thus decided to measure the distance between continents using astronomical observations.<sup>6</sup> However, because measuring continental drift called for measurement accuracies far greater than ever before to show small changes in positions over a few years, Wegener's attempts failed, and the idea of continental drift was largely rejected.

By the 1970s the story was very different. Geologists accepted continental drift, in large part because paleomagnetic measurements showed that continents had in fact moved over millions of years. It thus seemed natural to see if modern space-based technology could accomplish Wegener's dream of measuring continental motions over a few years. Three basic approaches were attempted. Each faced formidable technical challenges — and all succeeded. Hence, using the techniques discussed in Section 4.5.1, plate motions can now be measured to a precision of a few mm/yr or better, using a few years of data from systems including Very Long Baseline Interferometry (VLBI), Satellite Laser Ranging (SLR), and the Global Positioning System (GPS).

Space geodesy measures both the rate and the azimuth of the motions between sites, and can thus be used to compute rela-

tive plate motions. One of the most important results of space geodesy for seismology is that plate motions have remained generally steady over the past few million years. This is shown by the striking agreement between motions measured over a few years by space geodesy and the predictions of global plate motion models that average over the past three million years (Fig. 5.2-6). The general agreement is consistent with the idea that although motion at plate boundaries can be episodic, as in large earthquakes, the viscous asthenosphere damps out the transient motions (much like the damping element in a seismometer, Section 6.6) and causes steady motion between plate interiors. This steadiness implies that plate motion models can be used for comparison with earthquake data.

Space geodesy surmounts a major difficulty faced by models like NUVEL-1A: namely, that the data used (spreading rates, transform azimuths, and slip vectors) are at plate boundaries, so the model provides only the net motion across a boundary. By contrast, space geodesy can also measure the motion of sites within plate boundary zones. For example, Fig. 5.2-3 shows the motions of GPS and VLBI sites within the North America–Pacific boundary zone. Sites in eastern North America move so slowly — less than 2 mm/yr — with respect to each other that their motion vectors cannot be seen on this scale. These sites thus define a rigid reference frame for the stable interior of the North American plate. Sites west of the San Andreas fault move at essentially the rate and direction predicted for the Pacific plate by the global plate motion model. The site vectors show that most of the plate motion occurs along the San Andreas fault system, but significant motions occur for some distance eastward. The geodetic motions are consistent with the focal mechanisms and geological data. Thus, as discussed further in Section 5.6, the different data types are used together to study how the seismic and aseismic portions of the deformation vary in space and time in the diffuse deformation zones that characterize many plate boundaries. This is done both on large scales, as shown here, and for studies of smaller areas and individual earthquakes (Section 4.5).

Space geodesy is also used to study the relatively rare, but sometimes large, earthquakes within plates. Global plate motion models give no idea where or how often intraplate earthquakes should occur, beyond the trivial prediction that they should not occur because there is no deformation within ideal rigid plates. Space geodesy is being combined with earthquake locations, focal mechanisms, and other geological and geophysical data to investigate the motions and stresses within plates and how they give rise to intraplate earthquakes (Section 5.6.3).

#### 5.2.4 Absolute plate motions

So far, we have discussed the relative motions between plates, which have traditionally been of greatest interest to seismologists because most earthquakes reflect these motions. However, in some applications it is important to consider *absolute* plate motions, those with respect to the deep mantle.

In general, both plates and plate boundaries move with respect to the deep mantle. To see this, assume that the African

<sup>6</sup> Using an extraterrestrial reference has a long history; in about 230 BC Eratosthenes found the Earth's size from observations of the sun's position at different sites, and navigators have found their positions by observing the sun and stars.



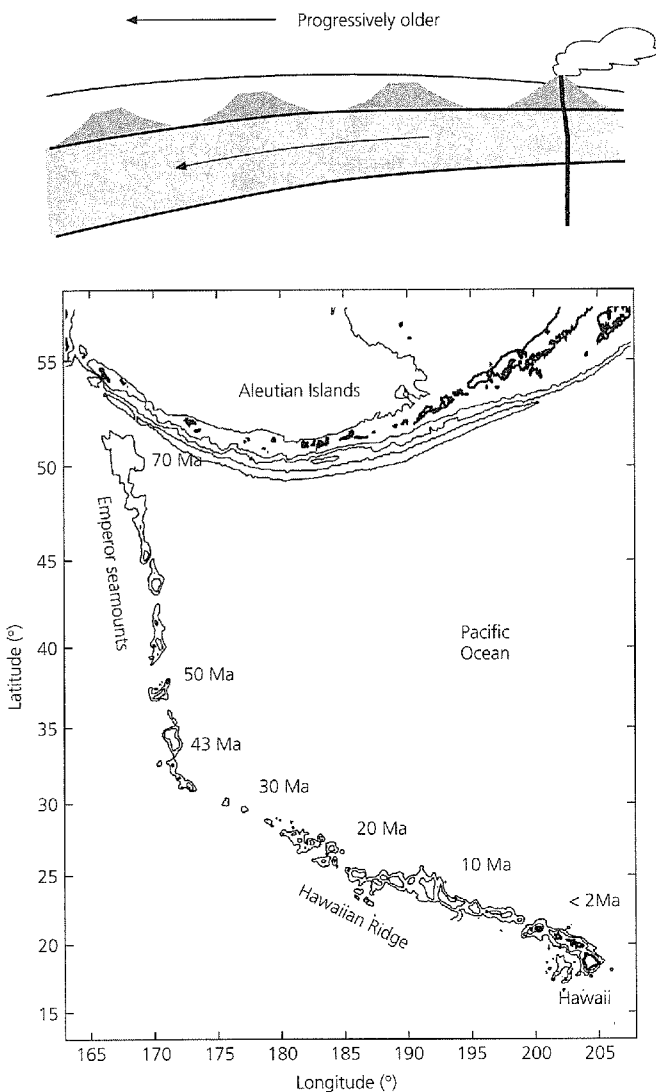


Fig. 5.2-7 Top: Illustration of the formation of a volcanic island chain by plate motion over a fixed hot spot. Bottom: Ages, in millions of years, of volcanoes in the Hawaiian-Emperor chain.

plate were not moving with respect to the deep mantle. In this case, as lithosphere was added to the plate by sea floor spreading at the Mid-Atlantic ridge (Fig. 5.2-4), both the ridge and the South American plate would move westward with respect to the mantle. Conversely, as the African plate lost area by subduction beneath the Eurasian plate in the Mediterranean, the trench would “roll backward,” causing both it and Eurasia to move southward relative to the mantle. Such motions can have important consequences for processes at plate boundaries (e.g. Fig. 5.3-10).

Absolute plate motions cannot be measured directly. Hence we infer these motions in two ways. One uses the *hot spot* hypothesis, in which certain linear volcanic trends result from the motion of a plate over a hot spot, or fixed source of volcanism, which causes melting in the overriding plate (Fig. 5.2-7). If

the overriding plate is oceanic, its motion causes a progression from active volcanism that builds the islands, to older islands, to underwater seamounts as the sea floor moves away from the hot spot, cools, and subsides. This process leaves a broad, shallow, topographic swell around the hot spot and a characteristic volcanic age progression away from it, as shown for the Hawaiian-Emperor seamount chain. The ages of volcanism range from present, on the currently active island of Hawaii, to a few million years on the other Hawaiian islands,<sup>7</sup> to about 28 Ma at Midway island, and about 70 Ma where the seamount chain vanishes into the Aleutian trench. Thus the direction and age of the volcanic chain give the motion of the plate with respect to the hot spot. For example, the bend in the Hawaiian-Emperor seamount chain has been interpreted as indicating that the Pacific plate changed direction about 40 million years ago. Hence using hot spot tracks beneath different plates, and assuming that the hot spots are fixed with respect to the deep mantle (or move relative to each other more slowly than plates), yields a hot spot reference frame.

It is often further assumed that hot spots result from plumes of hot material rising from great depth, perhaps even the core-mantle boundary (Fig. 5.1-2). The concepts of hot spots and plumes are attractive and widely used, but the relation between the persistent volcanism and possible deep mantle plumes remains a subject of active investigation because there are many deviations from what would be expected. Some hot spots move significantly, some chains show no clear age progression, evidence for plate motion changes associated with bends like that in Fig. 5.2-7 is weak, and oceanic heat flow data show little or no thermal anomalies at the swells. Seismological studies find low-velocity anomalies, but assessing their depth extent and relation to possible plumes is challenging. However, the hot spot reference frame is similar to one obtained by assuming there is no net rotation (NNR) of the lithosphere as a whole, and hence that the sum of the absolute motion of all plates weighted by their area is zero. Thus despite unresolved questions about the nature and existence of hot spots and plumes, NNR reference frames are often used to infer absolute motions.

To compute absolute motions, we recognize that motions in an absolute reference frame correspond to adding a rotation to all the plates. Thus we use the Euler vector formulation and treat the absolute reference frame as mathematically equivalent to another plate. We define  $\Omega_i$  as the Euler vector of plate  $i$  in an absolute reference frame. For example, Table 5.2-1 gives the NNR Euler vector relative to the North American plate ( $\omega_{NNR-NA}$ ), so its negative ( $\omega_{NA-NNR}$ ) is the absolute Euler vector  $\Omega_{NA}$  for North America in the NNR reference frame. The linear velocity at a point  $\mathbf{r}$  is found by analogy to Eqn 1:

$$\mathbf{v}_i = \Omega_i \times \mathbf{r}. \quad (12)$$

Thus we find the motion of North America with respect to the hot spot thought to be producing the volcanism and earthquakes in Yellowstone National Park ( $44^\circ, -110^\circ$ ) to be

<sup>7</sup> This age progression was recognized by native Hawaiians, who attributed it to the order in which the volcano goddess Pele plucked the islands from the sea.



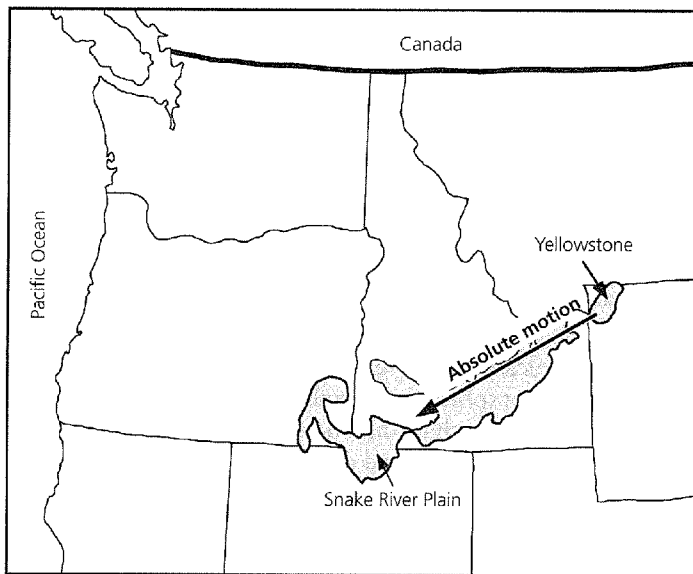


Fig. 5.2-8 Comparison of the predicted absolute motion of North America to the Snake River Plain basalts, which are thought to be the track of a hot spot now producing volcanism in Yellowstone National Park. (After Smith and Braile, 1994. *J. Volcan. Geotherm. Res.*, 61, 121–87, with permission from Elsevier Science.)

18 mm/yr directed N239°E. This motion is along the trend connecting the present volcanism in Yellowstone to the Snake River Plain basalts (Fig. 5.2-8), which are thought to be its track, a continental analogy to the Hawaiian–Emperor seamount chain.

Relative and absolute Euler vectors are simply related because

$$\omega_{ij} = \Omega_i - \Omega_j, \quad (13)$$

the relative Euler vector for two plates, is the difference between their absolute Euler vectors. Thus, if we know one plate's absolute motion, we can find all the others from the relative motions. For example, the absolute motion of the Pacific plate can be found from Table 5.2-1, which gives its vector relative to North America, using

$$\Omega_{PA} = \omega_{PA-NA} + \Omega_{NA}. \quad (14)$$

Absolute motions are important in several seismological applications. Seismology is used to study hot spots and their effects, including the resulting intraplate earthquakes like those associated with the volcanism in Hawaii. For example, Fig. 2.8-5 illustrated the use of surface wave dispersion to study the velocity structure under the Walvis ridge, which is thought to be the track produced by a hot spot under the Mid-Atlantic ridge. A second application involves seismic anisotropy in the mantle (Section 3.6), which is thought to reflect flow of olivine-rich material in a direction that is often consistent with the predicted absolute plate motions. Thus seismic anisotropy, seismic velocities, and absolute motions are being combined to model mantle flow.

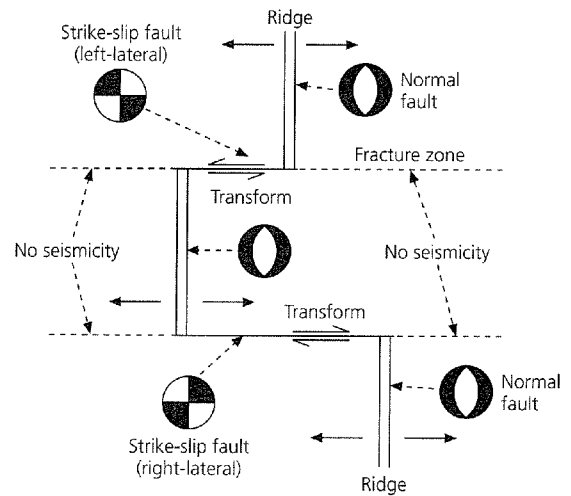


Fig. 5.3-1 Possible tectonic settings of earthquakes at an oceanic spreading center. Most events occur on the active segment of the transform and have strike-slip mechanisms consistent with transform faulting. On a slow-spreading ridge, like the Mid-Atlantic, normal fault earthquakes also occur.

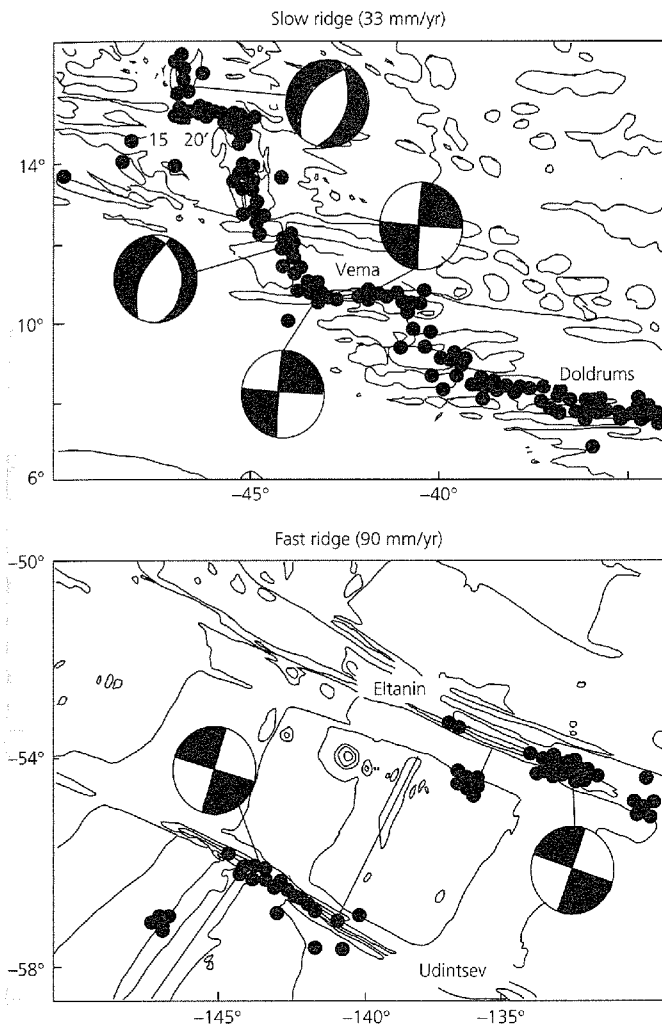
### 5.3 Spreading centers

Because the lithosphere forms at spreading centers, we begin with an overview of such systems and the earthquakes within them. We will see that seismological observations both demonstrate and reflect the basic kinematic model for ridges and transforms. Moreover, they provide key evidence for the thermal-mechanical processes that control the formation and evolution of the oceanic lithosphere.

#### 5.3.1 Geometry of ridges and transforms

Mid-ocean ridges are marked by earthquakes, which provide important information about the sea floor spreading process. Figure 5.3-1 is a schematic diagram of a portion of a spreading ridge offset by transform faults. Because new lithosphere forms at ridges and then moves away, transform faults are segments of the boundaries between plates, across which lithosphere moves in opposite directions. A given pair of plates can have either right- or left-lateral motion, depending on the direction in which a transform offsets the ridge; both reflect the same direction of relative plate motion. This motion across the transform is not what produced the offset of the ridge crest. In fact, in the usual situation such that spreading is approximately symmetric (equal rates on either side), the length of the transform will not change with time. This is a very different geometry from a transcurrent fault, where the offset between ridge segments is produced by motion on the fault and increases with time.

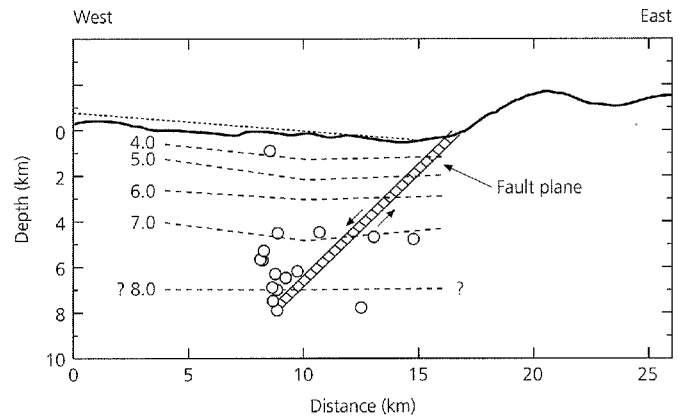
The focal mechanisms illustrate these ideas. Figure 5.3-2 (top) shows a portion of the Mid-Atlantic ridge composed of north–south-trending ridge segments that are offset by transform faults such as the Vema transform that trend approxim-



**Fig. 5.3-2** Maps contrasting faulting on slow- and fast-spreading centers. *Top:* The slow Mid-Atlantic ridge has earthquakes on both the active transform and the ridge segments. Strike-slip faulting on a plane parallel to the transform azimuth is characteristic. On the ridge segments, normal faulting with nodal planes parallel to the ridge trend is seen. *Bottom:* The fast East Pacific rise has only strike-slip earthquakes on the transforms. (Stein and Woods, 1989.)

ately east-west. Both the ridge crest and the transforms are seismically active. The mechanisms show that the relative motion along the transform is right-lateral. Sea floor spreading must be occurring on the ridge segments to produce the observed relative motion. For this reason, earthquakes occur almost exclusively on the active segment of the transform fault between the two ridge segments, although an inactive extension known as a fracture zone extends to either side. Although no relative plate motion occurs on the fracture zone,<sup>1</sup> it is often marked by a topographic feature due to the contrast in lithospheric ages across it.

<sup>1</sup> Unfortunately, some transform faults named before this distinction became clear are known as "fracture zones" along their entire length.



**Fig. 5.3-3** Cross-section through the Mid-Atlantic ridge. The fault plane inferred from the focal mechanisms of large earthquakes is consistent with the locations of microearthquakes (dots) determined using ocean bottom seismometers. Dashed lines show P-wave velocity structure. (Toomey *et al.*, 1988. *J. Geophys. Res.*, 93, 9093–112, copyright by the American Geophysical Union.)

Earthquakes also occur on the spreading segments. Their focal mechanisms show normal faulting, with nodal planes trending approximately along the ridge axis. These normal fault earthquakes are thought to be associated with the formation of the axial valley. For example, Fig. 5.3-3 shows a cross-section through the Mid-Atlantic ridge. The fault planes inferred from teleseismic focal mechanisms and the locations of microearthquakes determined using ocean bottom seismometers are consistent with normal faulting along the east side of the valley. Slip on this fault over 10,000 years would be enough to produce the observed geometry, including the eastward tilt of the valley floor.

The seismicity differs along the East Pacific rise. Here (Fig. 5.3-2, *bottom*) earthquakes occur on the transform faults with the expected strike-slip mechanisms, but few earthquakes occur on the ridge crest. This is probably because the East Pacific rise has an axial high, rather than the axial valley that occurs at the Mid-Atlantic ridge.<sup>2</sup> This difference appears to reflect the spreading rates: ridges spreading at less than about 60 mm/yr usually have axial valleys, whereas faster-spreading ridges have axial highs and thus do not have ridge crest normal faulting.

These examples show the spreading process at its simplest, but there can be complexities. Spreading can be asymmetric (one flank faster than the other) or oblique, such that the spreading is not perpendicular to the ridge axis. In addition, the geometry of a ridge system can change with time, as discussed in Section 5.3.3.

### 5.3.2 Evolution of the oceanic lithosphere

To understand the difference between fast- and slow-spreading ridges, and the nature of the earthquakes associated with them, it is important to understand the evolution of the oceanic

<sup>2</sup> This is often shown incorrectly on older maps.

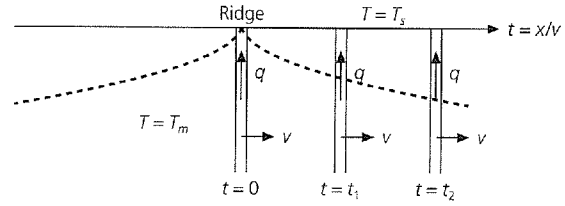
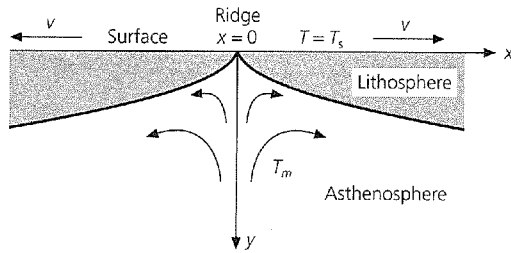


Fig. 5.3-4 Model for the cooling of an oceanic plate as it moves away from the ridge axis (*left*). Because a column moves away from the ridge faster than heat is conducted in the horizontal direction (*right*), the cooling in the vertical direction can be treated as a one-dimensional problem. (After Turcotte and Schubert, 1982.)

lithosphere. This process can be described using a simple, but powerful, model for the formation of the lithosphere by hot material at the ridge, which cools as the plate moves away.

In this model, material at the ridge at a mantle temperature  $T_m$  (1300–1400 °C) is brought to the ocean floor, which has a temperature  $T_s$ . The material then moves away at a velocity  $v$ , while its upper surface remains at  $T_s$  (Fig. 5.3-4). Because the plate moves away from the ridge faster than heat is conducted horizontally, we can consider only vertical heat conduction. Mathematically, this is the same as the cooling of a halfspace originally at temperature  $T = T_m$ , whose surface is suddenly cooled to  $T_s$  at time  $t = 0$ .

The temperature as a function of depth and time is given by the one-dimensional heat flow equation, which relates the temperature change with time in a piece of material to the rate at which heat is conducted out of it,

$$\frac{\partial T(z, t)}{\partial t} = \frac{k}{\rho C_p} \frac{\partial^2 T(z, t)}{\partial z^2} = \kappa \frac{\partial^2 T(z, t)}{\partial z^2}. \quad (1)$$

$\kappa$ , known as the *thermal diffusivity*, is a property of the material that measures the rate at which heat is conducted. It has units of distance squared divided by time, and is defined as  $\kappa = k/\rho C_p$ , where  $k$  is the thermal conductivity,  $\rho$  is the density, and  $C_p$  is the specific heat at constant pressure.

The well known solution to Eqn 1 is

$$T(z, t) = T_s + (T_m - T_s) \operatorname{erf} \left( \frac{z}{2\sqrt{\kappa t}} \right), \quad (2)$$

where

$$\operatorname{erf}(s) = \frac{2}{\sqrt{\pi}} \int_0^s e^{-\sigma^2} d\sigma \quad (3)$$

is known as the error function. Figure 5.3-5 (*right*) shows how this function varies between  $\operatorname{erf}(0) = 0$  and  $\operatorname{erf}(3) \approx 1$ . Thus cooling starts at the surface and deepens with time (Fig. 5.3-5, *left*).

Assuming that any column of oceanic lithosphere cools this way, and that the sea floor temperature is  $T_s = 0$  °C, then

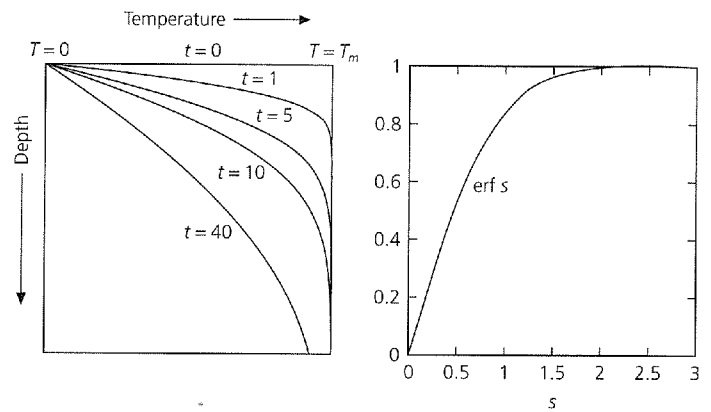


Fig. 5.3-5 *Left*: Cooling of a halfspace as described by the one-dimensional heat flow equation. The surface is cooled at time zero, and then the interior cools with time. *Right*: The error function, which controls the cooling solution shown.

$$T(z, t) = T_m \operatorname{erf} \left( \frac{z}{2\sqrt{\kappa t}} \right) \quad (4)$$

gives the temperature at a depth  $z$  for material of age  $t$ . The lithosphere moves away from the ridge at half the total spreading rate, so the age of the lithosphere is  $t = x/v$ , its distance from the ridge divided by the half-spreading rate  $v$ . Thus the temperature (Eqn 4) as a function of distance and depth is

$$T(x, z) = T_m \operatorname{erf} \left( \frac{z}{2\sqrt{\kappa x/v}} \right). \quad (5)$$

It is useful to think of *isotherms*, lines of constant temperature, in the plate. An isotherm is a curve on which the argument of the error function is constant,

$$\frac{z_c}{2\sqrt{\kappa t}} = c, \quad \text{or} \quad z_c = 2c\sqrt{\kappa t}, \quad (6)$$

so that the depth to a given temperature increases as the square root of the lithospheric age.

This is an example of a general feature of heat conduction problems: setting  $c = 1$  and examining Fig. 5.3-5 for  $\operatorname{erf}(1)$

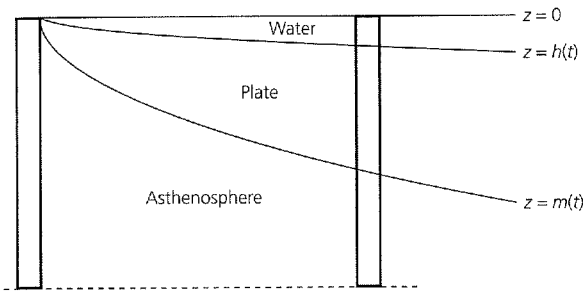


Fig. 5.3-6 The increase in ocean depth with lithospheric age due to the cooling of the lithosphere can be modeled using isostasy, the assumption that the mass in a vertical column is the same for all ages.

shows that most of the temperature change has propagated a distance  $2\sqrt{\kappa t}$  in a time  $t$ . For example, after a lava flow erupts, it cools as the square root of time. Such square root of time behavior occurs for any process described by a diffusion equation, of which the heat equation is an example.

The concept that the lithosphere cools with time such that isotherms deepen with the square root of age has many observable consequences. The simplest is that ocean depth should vary with age, which makes sense, because spreading centers are ridges precisely because the ocean deepens on either side. To model this effect, we consider the mass in two columns, one at the ridge and one at age  $t$ , and invoke the idea of isostasy, which means that the masses in the two columns balance (Fig. 5.3-6).<sup>3</sup>

Assume that the lithosphere, defined by the  $T = T_m$  isotherm, has thickness zero at the ridge and  $z = m(t)$  at age  $t$ , where the water depth is  $h(t)$ . Similarly, we assume that the asthenosphere is at temperature  $T_m$  and has density  $\rho_m$ . However, the temperature and thus density in the cooling lithosphere vary, such that at the point  $(z, t)$  the temperature is  $T(z, t)$  and the corresponding density is

$$\rho(z, t) = \rho_m + \frac{\partial \rho}{\partial T} [T(z, t) - T_m] = \rho_m + \rho'(z, t). \quad (7)$$

The change in density due to temperature, at constant pressure, is given by the coefficient of thermal expansion,

$$\alpha = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_P = -\frac{1}{\rho} \left( \frac{\partial \rho}{\partial T} \right)_P \quad (8)$$

(the minus sign is because  $\partial \rho / \partial T$  is negative). Thus the density perturbation for the halfspace cooling model is

$$\rho'(z, t) = \alpha \rho_m [T_m - T(z, t)] = \alpha \rho_m T_m \left[ 1 - \operatorname{erf} \left( \frac{z}{2\sqrt{\kappa t}} \right) \right]. \quad (9)$$

If the density of water is  $\rho_w$ , equal mass in the two columns requires that

$$\rho_m m(t) = \rho_w h(t) + \int_{h(t)}^{m(t)} [\rho_m + \rho'(z, t)] dz, \quad (10)$$

which gives the isostatic condition for ocean depth,

$$h(t) = \frac{1}{(\rho_m - \rho_w)} \int_{h(t)}^{m(t)} \rho'(z, t) dz. \quad (11)$$

Because temperature and density in the plate are defined for all values of  $z$  (the thickness of the plate is defined as some chosen isotherm), let  $z' = z - h(t)$  and  $m(t) \rightarrow \infty$ . Then

$$h(t) = \frac{\alpha \rho_m T_m}{(\rho_m - \rho_w)} \int_0^{\infty} \left[ 1 - \operatorname{erf} \left( \frac{z'}{2\sqrt{\kappa t}} \right) \right] dz'. \quad (12)$$

To evaluate the integral, substitute  $s = z'/2\sqrt{\kappa t}$  and integrate by parts (try it!) to show that

$$\int_0^{\infty} [1 - \operatorname{erf}(s)] ds = 1/\sqrt{\pi}. \quad (13)$$

Thus ocean depth should increase as the square root of plate age,

$$h(t) = 2\sqrt{\frac{\kappa t}{\pi}} \frac{\alpha \rho_m T_m}{(\rho_m - \rho_w)}. \quad (14)$$

The cooling of the lithosphere should also cause heat flow at the sea floor to vary with age. By Fourier's law of heat conduction, the heat flow at the sea floor is the product

$$q = k \frac{dT}{dz} \quad \text{at } z=0 \quad (15)$$

of the temperature gradient at the sea floor and the thermal conductivity  $k$ .<sup>4</sup> An easy approximation to see how heat flow varies with age is to consider the  $T_m$  isotherm as the base of the lithosphere, so that the thickness of the lithosphere increases

<sup>3</sup> Isostasy is the general idea that topography results from equal masses in different columns. Here we consider thermal isostasy, in which density changes produced by temperature variations cause topographic differences. Another common model, Airy isostasy, is used to explain the relation between crustal thickness variations and topography, such as crustal roots under mountains.

<sup>4</sup> Normally, this equation requires a minus sign because heat flows from hot objects to cold ones. Without this sign, hot objects would get hotter. There is none here because of our customary but inconsistent definitions: heat flow is measured upward whereas depth is measured downward.

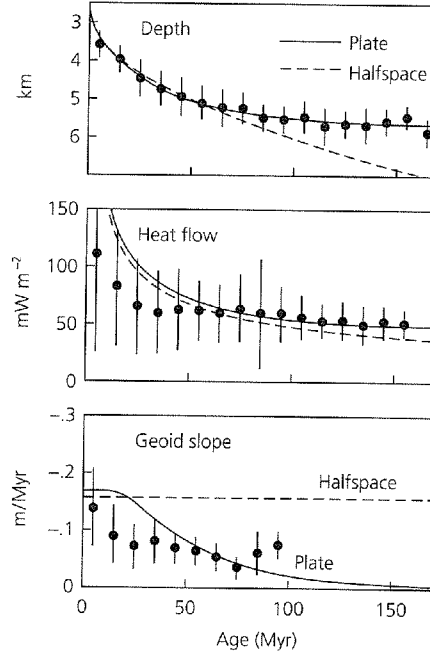
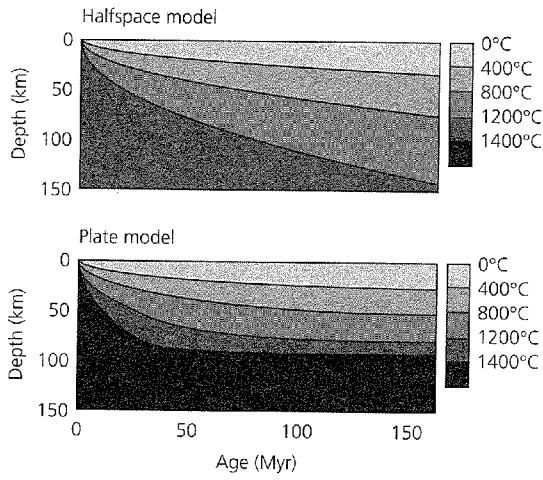


Fig. 5.3-7 Models and data for thermal evolution of the oceanic lithosphere. *Left*: Isotherms for thermal models. The lithosphere continues cooling for all ages in a halfspace model, but equilibrates for ~70 Ma lithosphere in a plate model with a 95 km-thick thermal lithosphere. The plate model shown has a higher basal temperature than the halfspace model. *Right*: Comparison of thermal model predictions to different data. All show a lithospheric cooling signal, and are better (but far from perfectly) fit by the predictions of a plate model (solid lines) than by those of a halfspace model (dashed lines). (Richardson *et al.*, 1995. *Geophys. Res. Lett.*, 22, 1913–16, copyright by the American Geophysical Union.)

with the square root of age. Approximating the gradient at the surface by the average gradient through the lithosphere,

$$q(t) \approx k \frac{\Delta T}{\Delta z} \approx \frac{k T_m}{\sqrt{\kappa t}} \quad (16)$$

predicts that the heat flow decreases as the square root of age. The same result can be obtained by differentiation of the temperature structure (Eqn 4) using

$$\frac{d}{dz} \operatorname{erf}(s) = \frac{d}{dz} \frac{2}{\sqrt{\pi}} \int_0^s e^{-\sigma^2} d\sigma = \frac{2}{\sqrt{\pi}} e^{-s^2} \frac{ds}{dz}, \quad (17)$$

which gives

$$q(t) = k \left. \frac{dT}{dz} \right|_{z=0} = k \frac{2T_m}{\sqrt{\pi}} e^{-\frac{z^2}{4\kappa t}} \frac{1}{2\sqrt{\kappa t}} \bigg|_{z=0} = \frac{kT_m}{\sqrt{\pi\kappa t}}. \quad (18)$$

This model, which predicts that lithospheric thickness, heat flow, and ocean depth vary as the square root of age for all ages is called a halfspace model (Fig. 5.3-7, *upper left*). In it, the lithosphere is the upper layer of a halfspace that continues cooling for all time. (In reality, oceanic lithosphere never gets older than 200 million years old because it gets subducted.) The model does a good job of describing the average variation in ocean depth and heat flow with lithospheric age.

However, because ocean depth seems to “flatten” at about 70 Myr, we often use a modification called a plate model (Fig. 5.3-7, *lower left*), which assumes that the lithosphere evolves toward a finite plate thickness  $L$  with a fixed basal temperature  $T_m$ . In this model,

$$T(x, z) = T_m \left[ \frac{z}{L} + \sum_{n=1}^{\infty} c_n \exp \left( -\frac{\beta_n x}{L} \right) \sin \left( \frac{n\pi z}{L} \right) \right], \quad (19)$$

where  $c_n = 2/(n\pi)$ ,  $\beta_n = (R^2 + n^2\pi^2)^{1/2} - R$ ,  $R = \nu L/(2\kappa)$ . The constant  $R$ , known as the thermal Reynolds number, relates the rates at which heat is transported horizontally by plate motion and conducted vertically. In this model isotherms initially deepen as the square root of age, but eventually level out. The flattening reflects the fact that heat is being added from below, which the model approximates by having old lithosphere reach a steady-state thermal structure that is simply a linear geotherm (Fig. 5.3-8, *top*). As a result, the predicted sea floor depth and heat flow also behave for young ages like in the halfspace model, but evolve asymptotically toward constant values for old ages. Both have simple interpretations: the heat flow is proportional to the geotherm, and thus  $T_m/L$ , whereas the depth is proportional to the thermal subsidence and hence heat lost since the plate formed at the ridge, and thus the product  $T_m L$ . The model parameters can be estimated by an inverse problem, finding those that best fit a set of depth and heat flow data versus age (Fig. 5.3-8, *bottom*).

Comparison with data shows that the plate thermal model is a good, but not perfect, fit to the average data because processes other than this simple cooling are also occurring. For example, ocean depth is also affected by uplift associated with hot spots (Section 5.2.4). Water flow in the crust transports some of the heat for ages less than about 50 Ma, making the observed heat flow lower than the model's predictions, which assume that all heat is transferred by conduction. Some topographic effects, including the spectacular volcanic oceanic plateaus, result from crustal thickness variations. Because these and other effects vary from place to place, the data vary about their average values for a given age.

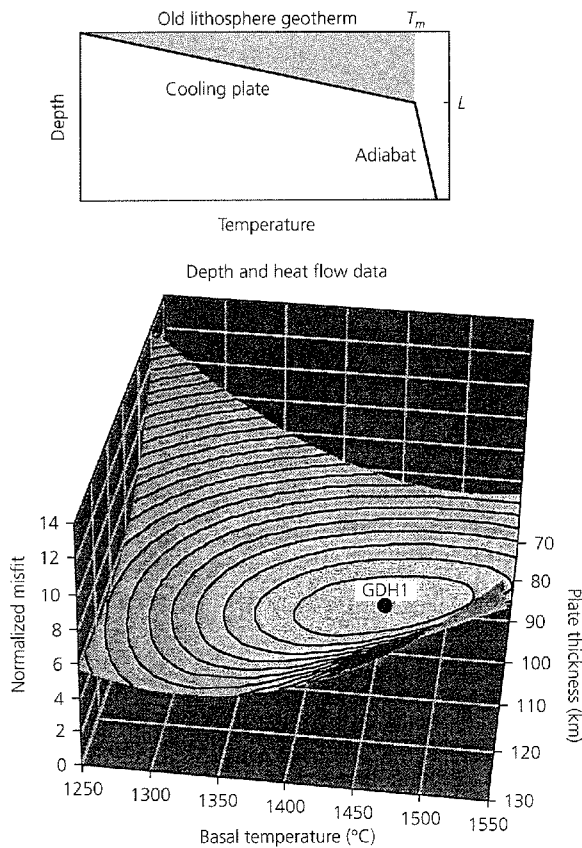


Fig. 5.3-8 *Top*: Asymptotic thermal structure for old lithosphere in a plate model. The sea floor subsidence from the ridge, and thus ocean depth, is proportional to the shaded area between the geotherm and  $T = T_m$ , whereas heat flow is proportional to the geotherm. A schematic adiabatic temperature gradient (Section 5.4.1) is shown beneath the plate. (Stein and Stein, 1992. Reproduced with permission from *Nature*.) *Bottom*: Fitting process used for thermal model parameters. The misfit to a set of depth and heat flow data has a minimum at the point labeled GDH1, a plate thermal thickness of  $95 \pm 15$  km and basal temperature of  $1450 \pm 250^\circ\text{C}$ . (Stein and Stein, 1996. *Subduction*, 1–17, copyright by the American Geophysical Union.)

We can view ocean depth, heat flow, and several other properties of the oceanic lithosphere as observable measures of the temperature in the cooling lithosphere. Because the observables depend on different combinations of parameters (Table 5.3-1), they can be used together to constrain individual parameters (a halfspace model corresponds to an infinitely thick plate). The depth depends on the integral of the temperature (Eqn 11), whereas the heat flow depends on its derivative at the sea floor (Eqn 15). Similarly, the slope of the geoid, a function of the gravity field depending on a weighted integral of the density, also varies with age in general agreement with the plate model's prediction (Fig. 5.3-7).

In addition, the elastic thickness of the lithosphere inferred from the deflection caused by loads such as seamounts (Fig. 5.3-9a), the maximum depth of intraplate earthquakes within the oceanic lithosphere (Fig. 5.3-9b), and the depth to

Table 5.3-1 Constraints on thermal models  $T(z, t)$ .

Observable	Proportional to	Reflects
Young ocean depth	$\int T(z, t) dz$	$k^{1/2} \alpha T_m$
Old ocean depth	$\int T(z, t) dz$	$\alpha T_m L$
Old ocean heat flow	$\left. \frac{\partial T(z, t)}{\partial z} \right _{z=0}$	$k T_m / L$
Geoid slope	$\frac{\partial}{\partial t} \int z T(z, t) dz$	$k \alpha T_m \exp(-kt/L^2)$

Source: Stein and Stein (1996).

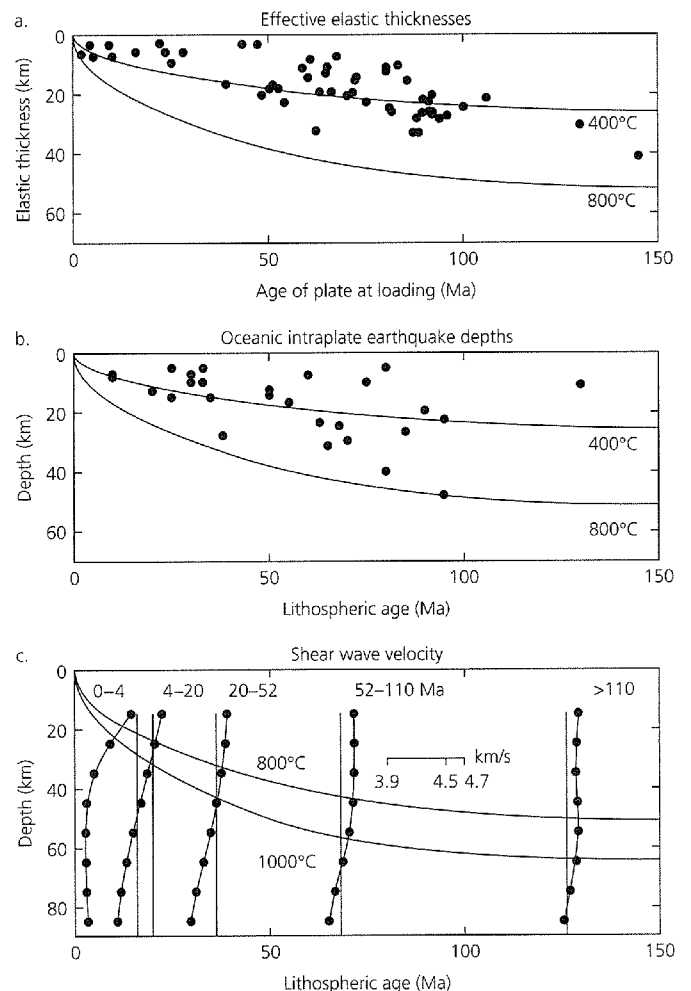
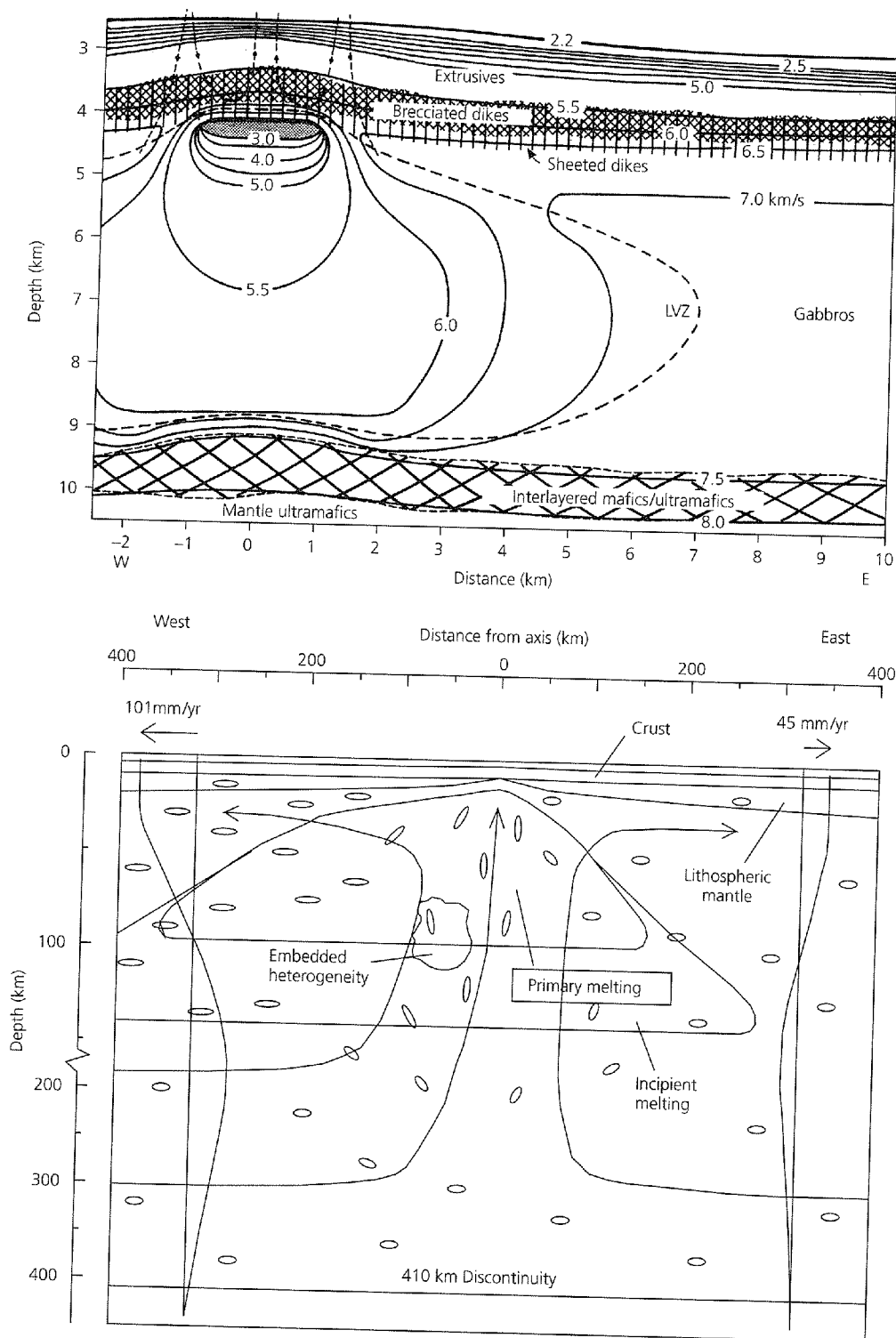


Fig. 5.3-9 Comparison of isotherms as functions of age for a plate model to three datasets whose variation with age is consistent with cooling of the lithosphere. The effective elastic thickness (a), deepest intraplate seismicity (b), and depth to the low-velocity zone, shown by velocity profiles at different ages (c), all increase with age. (After Stein and Stein, 1992. Reproduced with permission from *Nature*.)

the low-velocity zone determined from surface wave dispersion (Figs. 5.3-9c and 2.8-7), all increase with age. Hence the cooling of oceanic lithosphere causes the expected increase in strength and seismic velocity. Moreover, as discussed in Section 5.5, the resulting density increase is thought to provide a major force driving plate motions.

Because various properties vary with age, the oceanic lithosphere can be defined in various ways, so terms like "seismic lithosphere," "elastic lithosphere," and "thermal lithosphere" are often used. Interestingly, these thicknesses differ. It looks as if the deepest earthquakes are bounded by about 600–800 °C, such that hotter material cannot support seismic failure. The



**Fig. 5.3-10** *Top*: Geological interpretation of a multichannel seismic velocity study on the East Pacific rise. A low-velocity region under the axis is interpreted as a hot region of melting, capped by a magma lens. Dashed lines are possible paths of water circulation. (Vera *et al.*, 1990. *J. Geophys. Res.*, 95, 15,529–56, copyright by the American Geophysical Union.) *Bottom*: Schematic cross-section across the East Pacific rise. The broad region of low velocities is interpreted as the primary melting region. Small ellipses are directions of preferred olivine alignment inferred from anisotropy. Lines with arrows indicate inferred mantle flow, causing the distortion shown of an initially vertical line. Absolute velocities of the two plates (Pacific on left, Nazca on right) are given by small horizontal arrows. (Forsyth *et al.*, 1998. *Science*, 280, 1215–18, copyright 1998 American Association for the Advancement of Science.)

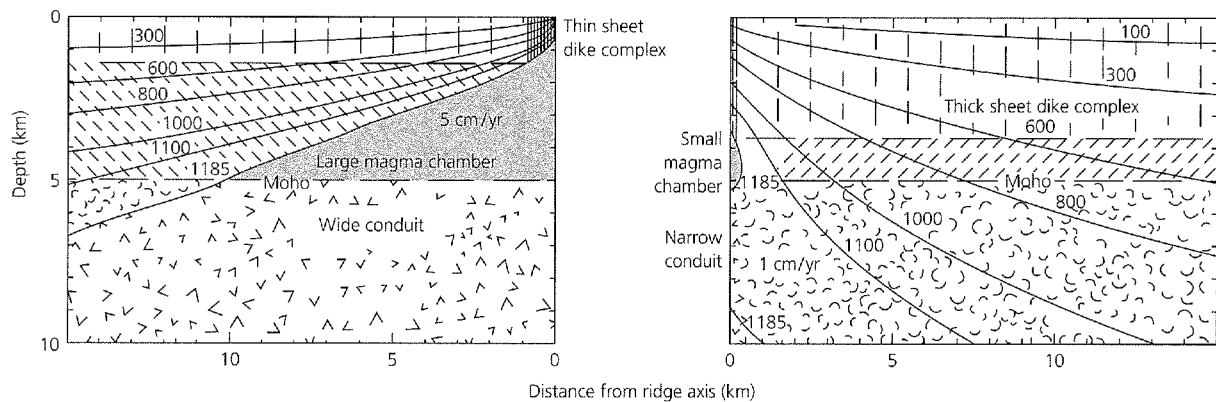


Fig. 5.3-11 Thermal and petrological model for the difference between fast-spreading (*left*) and slow-spreading (*right*) ridges. (Sleep and Rosendahl, 1979. *J. Geophys. Res.*, 84, 6831–9, copyright by the American Geophysical Union.)

elastic thickness corresponds approximately to the 400 °C isotherm, whereas the low-velocity zone begins approximately below the 1000 °C isotherm (Fig. 5.3-9c). These differences, discussed in Section 5.7, likely result from rock being stronger for more rapid deformation. All of these thicknesses, however, only approximate what we would like to know but cannot directly measure: the depth of the base of the moving plate, which is likely to be a gradational rather than a distinct boundary.

### 5.3.3 Ridge and transform earthquakes and processes

Seismology makes important contributions to understanding the properties and behavior of spreading centers. Ocean bottom seismometers yield locations of microearthquakes and data for travel time and waveform studies. Larger earthquakes are also studied using teleseismic body and surface waves. The seismological results are being integrated with marine geophysical and petrological data to develop better models. For example, Fig. 5.3-10 (*top*) shows a geological interpretation of a multichannel seismic study (Section 3.3) that used air gun and explosive sources to image velocity structure under the East Pacific rise to a depth of about 10 km. A low-velocity region under the axis is interpreted as a hot melting region capped by a magma lens. Other studies using ocean bottom seismometers and distant earthquake sources map the structure to greater depth, including inferring flow directions under the ridge axis using anisotropy (Fig. 5.3-10, *bottom*). Such studies are finding interesting features of the spreading process. For example, the broad region of low velocity presumed to be the primary melting area extends further west than east of the axis. This asymmetry may occur because the westward absolute motion of the Pacific plate is much faster than the eastward absolute motion of the Nazca plate, causing the ridge to migrate westward relative to the deep mantle. Thus the spreading process, which depends on the relative plate motion (spreading rate), also seems affected by the absolute motion.

Some effects of the spreading rate are illustrated by a model

shown in Fig. 5.3-11. At a given distance from the ridge, faster spreading produces younger lithosphere and isotherms closer to the surface than does slow spreading. If the region beneath the 1185 °C isotherm and above the Moho depth of 5 km is considered to be a magma chamber, a fast ridge has a larger magma chamber. Hence crust moving away from a fast-spreading ridge is more easily replaced than that moving away from a slow ridge. Thus, in contrast to the axial valley and normal faulting earthquakes on a slow ridge, a fast ridge has an axial high and an absence of earthquakes. Similarly, both the depths and the maximum seismic moments<sup>5</sup> of ridge crest normal faulting earthquakes decrease with spreading rate (Fig. 5.3-12). These observations are consistent with the fault area decreasing on faster-spreading and hotter ridges, because faulting requires that rock be below a limiting temperature, above which it flows (Section 5.7). The idea that the faulting depends on temperature is also implied by the increase in the maximum depth of oceanic intraplate earthquakes with age (Fig. 5.3-9b).

Transform fault earthquakes also depend on thermal structure. The temperatures along a transform fault should be essentially the average of the expected temperature on the two sides; coolest at the transform midpoint and hottest at either end (Fig. 5.3-13). As expected from the area available for faulting, the maximum seismic moment for transform earthquakes decreases with spreading rate (Fig. 5.3-14), consistent with the idea of faulting limited to a zone bounded by the isotherms.

An interesting question is how the seismic moments of transform earthquakes relate to the plate motion. The average slip rate from earthquakes can be inferred from the total seismic moment released on a transform, assuming that

$$\text{seismic slip rate} = \frac{\text{total seismic moment}}{(\text{fault area})(\text{rigidity})(\text{time period})}. \quad (20)$$

<sup>5</sup> Recall (Section 4.6) that the seismic moment is the product of the rigidity, the slip in the earthquake, and the fault area.



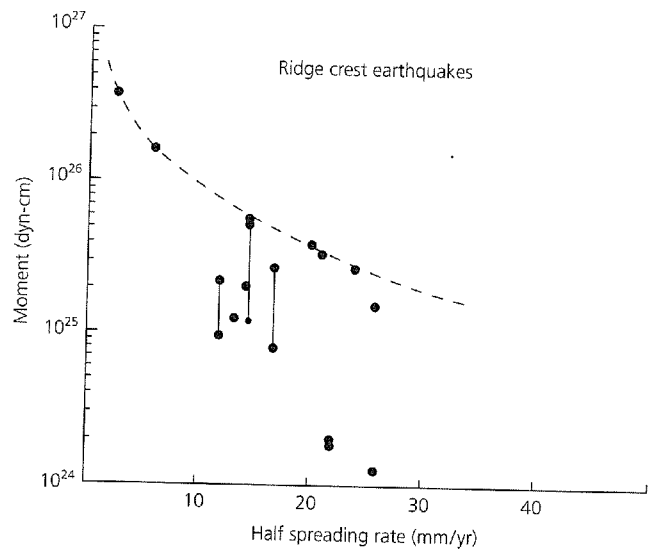
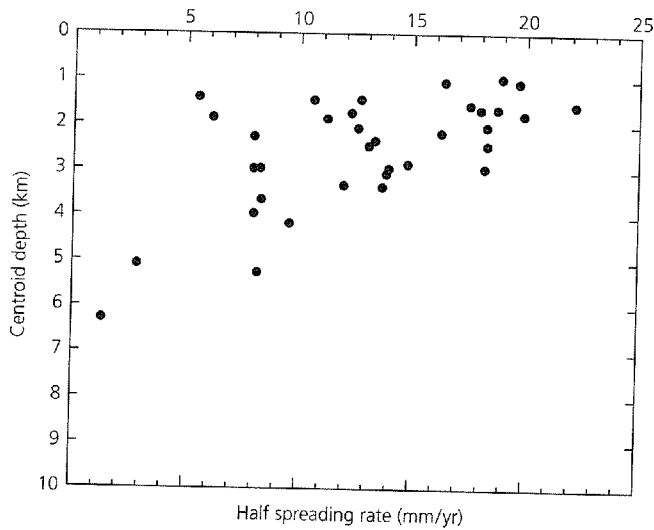


Fig. 5.3-12 Left: Shallowing of focal depth for ridge crest normal fault earthquakes with half-spreading rate. (After Huang and Solomon, 1988. *J. Geophys. Res.*, 93, 13, 445–77, copyright by the American Geophysical Union.) Right: Corresponding decrease in maximum seismic moment. (After Solomon and Burr, 1979. *Tectonophysics*, 55, 107–26, with permission from Elsevier Science.)

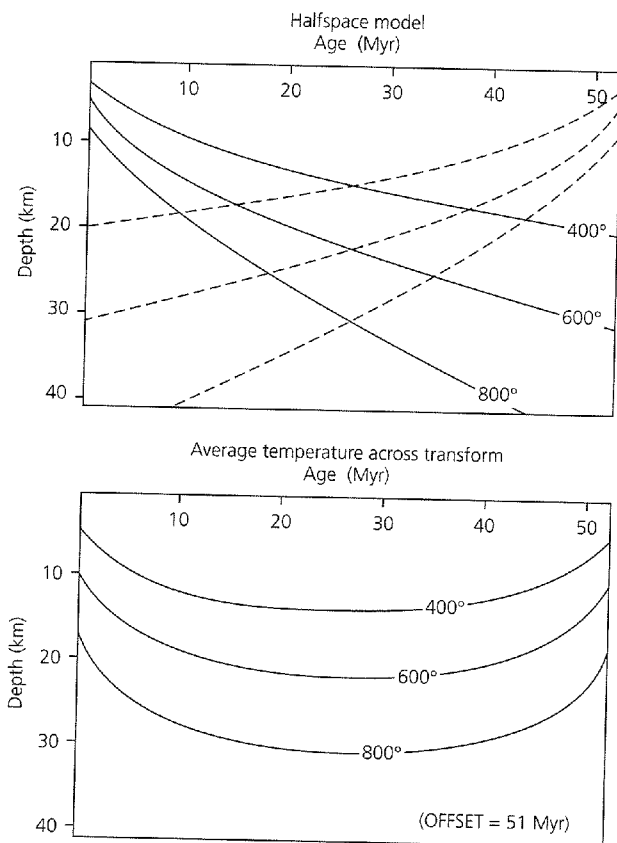


Fig. 5.3-13 Thermal model of the Romanche Transform. Top: Temperatures on either flank predicted by the cooling halfspace model. Bottom: Average temperature distribution along the transform. (After Engeln *et al.*, 1986. *J. Geophys. Res.*, 91, 548–77, copyright by the American Geophysical Union.)

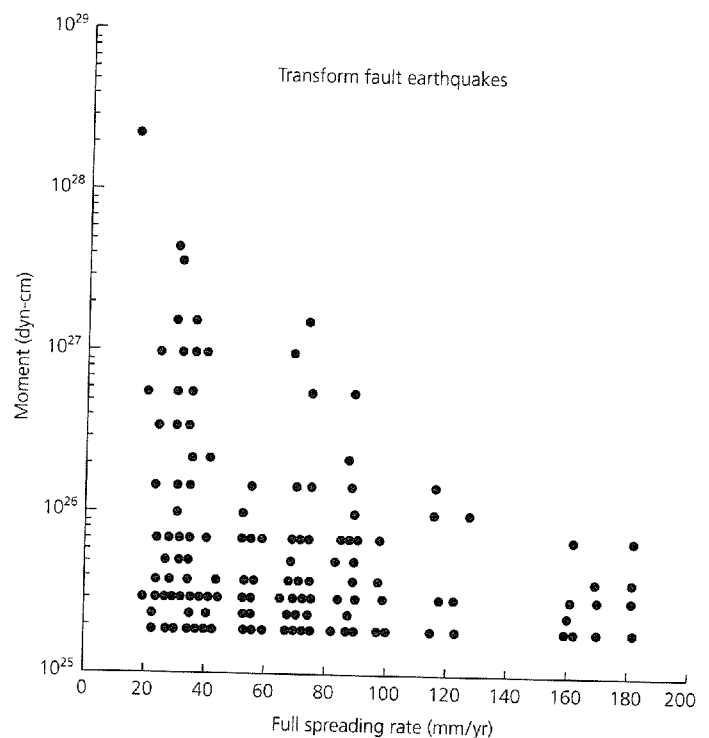


Fig. 5.3-14 Seismic moment versus spreading rate for oceanic transforms. The maximum moment decreases with spreading rate, as expected from thermal considerations. (After Solomon and Burr, 1979. *Tectonophysics*, 55, 107–26, with permission of Elsevier Science.)

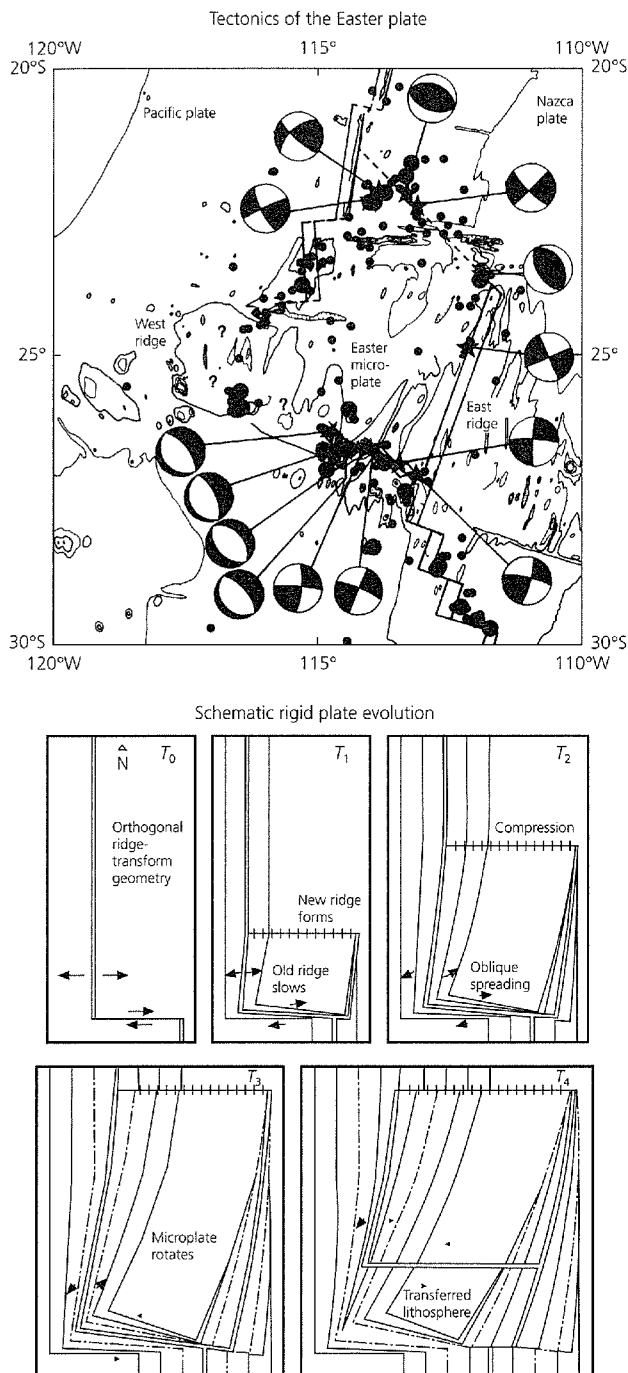


Fig. 5.3-15 The Easter microplate on the East Pacific rise. *Top*: Seismicity (dots) and focal mechanisms in the microplate region. Note the normal faulting on the southern boundary. (After Engeln and Stein, 1984.) *Bottom*: Schematic model for the evolution of a rigid microplate between two major plates by rift propagation. Successive isochrons illustrate the northward propagation of the east ridge, slowing of spreading on the west ridge, the rotation of the microplate, the reorientation of the two ridges, and the conversion of the initial transform into a slow and obliquely spreading ridge. (Engeln *et al.*, 1988. *J. Geophys. Res.*, 93, 2839–56, copyright by the American Geophysical Union.)

Using this relation requires inferring the fault area, which depends on both the transform length and the depth to which faulting occurs. Assuming the area above the 600–700 °C isotherms fails seismically, the seismic slip rate for major Atlantic transforms is generally less than predicted by the plate motion. Thus, if the time period sampled is long enough to be representative — a major question — some of the plate motion occurs aseismically. The issue of how much slip occurs seismically remains unresolved, as we will see when we discuss subduction zones (Section 5.4.3) and intraplate deformation zones (Section 5.6.2).

In addition, seismology helps study how ridge-transform systems evolve. For example, the East Pacific rise near Easter Island contains two approximately parallel sections (Fig. 5.3-15, *top*). Earthquakes occur on these ridges, but not between them, suggesting that the area in between is an essentially rigid microplate. The normal fault earthquakes on the microplate's southern boundary are surprising because the East Pacific rise here is a very fast-spreading (15 cm/yr) ridge, which should not have normal fault earthquakes (Fig. 5.3-12). Magnetic anomalies show that the east ridge segment is propagating northward and taking over from the old (west) ridge segment. Figure 5.3-15 (*bottom*) shows a simplified model of this process. Because finite time is required for the new ridge to transfer spreading from the old ridge, both ridges are active at the same time, and the spreading rate on the new ridge is very slow at its northern tip and increases southward. As a result, the microplate rotates, causing compression (thrust faulting) and extension (normal faulting) at its north and south boundaries, respectively. Ultimately the old ridge will die, transferring lithosphere originally on the Nazca plate to the Pacific plate, and leaving inactive fossil ridges on the sea floor. Both V-shaped magnetic anomalies characteristic of ridge propagation and fossil ridges are widely found in the ocean basins, showing that this is a common way that ridges reorganize. Even for smaller (a few km) propagating ridge systems, studies of the associated earthquakes can yield useful information about the propagation process.

## 5.4 Subduction zones

We have seen that earthquakes at spreading centers, which at shallow depths are upwelling limbs of the mantle convection system, reflect the processes forming oceanic lithosphere there. In a similar way, earthquakes at subduction zones, downwelling limbs of the convection system, reflect the processes by which oceanic lithosphere reenters the mantle. Plate convergence takes different forms, depending on the plates involved. Figure 5.4-1 shows the basic model for a situation where oceanic lithosphere of one plate subducts beneath oceanic lithosphere of the overriding plate. Typically, a volcanic island arc forms, and sea floor spreading occurs behind the arc, forming a back-arc basin or marginal sea. Earthquakes occur both at the trench and to great depth, forming a dipping

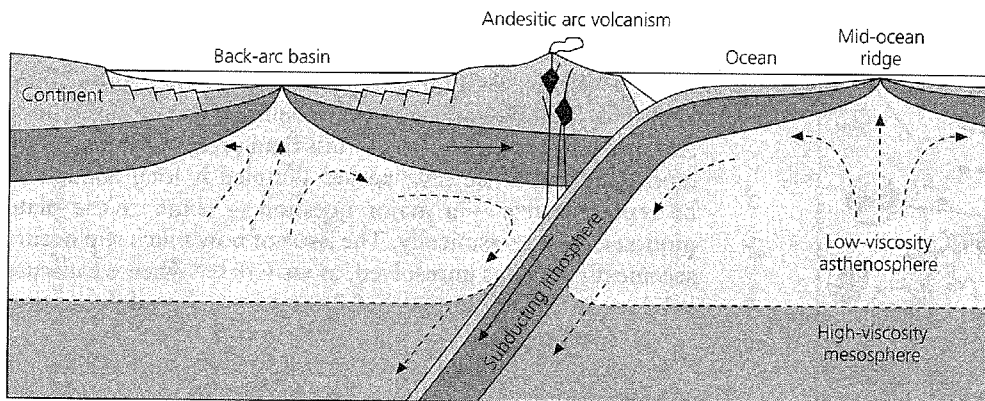


Fig. 5.4-1 Schematic diagram of processes associated with the subduction of one oceanic plate beneath another.

Wadati–Benioff zone. By contrast, when oceanic lithosphere subducts beneath a continent, a mountain chain like the Andes forms on the continent, and the oceanic lithosphere forms a Wadati–Benioff zone. Finally, because continental crust cannot subduct, convergence between two continental plates, as in the Himalayas, causes crustal thickening, mountain building, and shallow earthquakes but does not create a Wadati–Benioff zone.

Subduction zones have a wide variety of earthquakes with different focal mechanisms and depths. There are shallow (less than 70 km deep), intermediate (70–300 km deep), and deep (more than 300 km deep) focus earthquakes.<sup>1</sup> These earthquakes occur in different tectonic environments. The intermediate and deep earthquakes forming the Wadati–Benioff zone occur in the cold interiors of downgoing slabs. The shallow earthquakes are associated with the interaction between the two plates. The largest and most common of these shallow earthquakes occur at the interface between the plates, and release the plate motion that has been locked at the plate interface. In addition, shallow earthquakes can occur within both the overriding and the subducting plates. Figure 5.4-2 shows some features of seismicity observed in subduction zones. Not all features have been observed at all places. For example, the dips and shapes of subduction zones vary substantially. Some show double planes of intermediate or deep seismicity, whereas others do not.

In discussing subduction zones, we follow an approach similar to that used in the last section for ridges. We introduce thermal models for subduction, then use them to gain insight into earthquake and seismic velocity observations. We will see that seismological observations, thermal models, and calculations of the behavior of materials at high temperature and pressure are combined to investigate these complicated regions. In general, the seismological observations are fairly clear, but they can be interpreted in terms of a variety of models. As a result, subduction zone studies remain active, fruitful, and exciting.

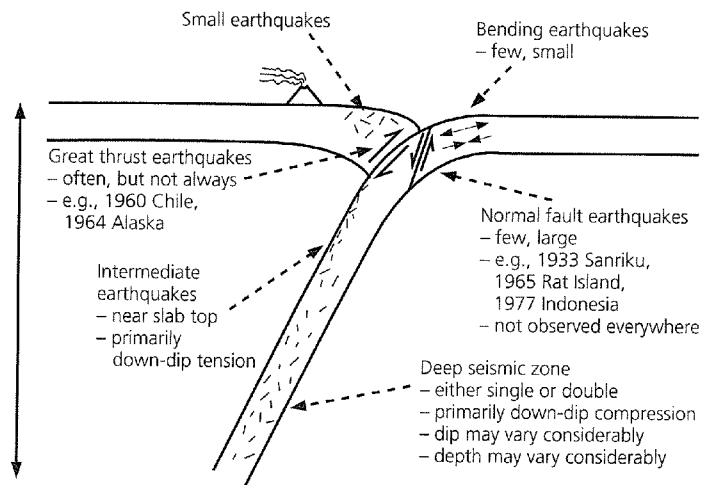


Fig. 5.4-2 Composite subduction zone showing some earthquake types. Not all are observed at all subduction zones.

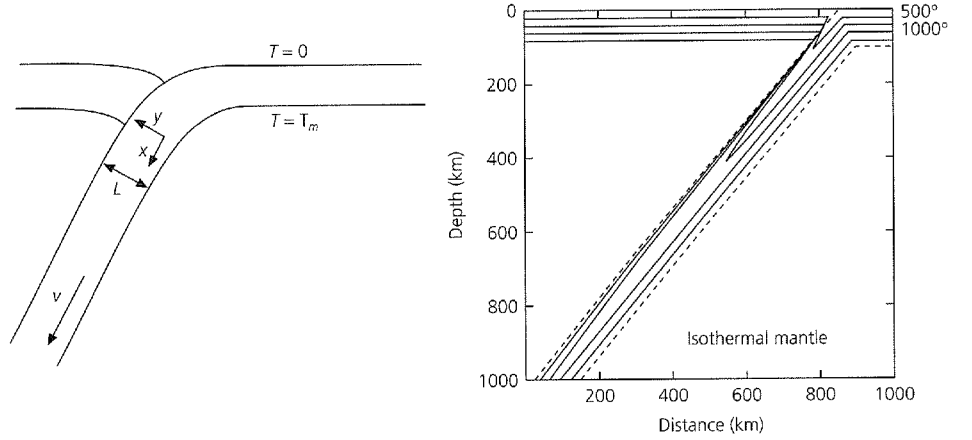
#### 5.4.1 Thermal models of subduction

The essence of subduction is the penetration and slow heating of a cold slab of lithosphere as it descends into the warmer mantle. As we will see, slabs subduct rapidly compared to the time needed for heat conducted from the surrounding mantle to warm them up. Thus they remain colder, denser, and mechanically stronger than the surrounding mantle. Consequently, slabs transmit seismic waves faster and with less attenuation than the surrounding mantle, making it possible to map slabs and to show that deep earthquakes occur within them. Moreover, the negative thermal buoyancy of cold slabs appears to be the primary force driving plate motions and provides a major source of stress within them that causes deep earthquakes.

To explore the thermal evolution of slabs, we use two approaches. First, we discuss a simplified analytic thermal model that allows insights into the physics. We then discuss numerical models that incorporate additional effects in the hope of providing a more realistic description. We highlight some significant points, and more complete information can be found in the references.

<sup>1</sup> Slightly different definitions have been used for these depth ranges; for example, 325 km has also been used as the upper limit for deep earthquakes.

Fig. 5.4-3 An analytic model for temperatures in a subducting plate. *Left*: model geometry. *Right*: Results, showing the cold slab heating up as it descends through the hotter surrounding mantle.



The analytic model (Fig. 5.4-3) considers a semi-infinite slab of thickness  $L$  subducting at rate  $v$ . The surrounding mantle is at temperature  $T_m$ , and the plate enters the trench with a linear temperature gradient from  $T=0$  at its top to  $T_m$  at its base. We define the  $x$  axis down the dip of the slab, and the  $y$  axis across the slab. The evolution of the region is given by a slightly more complicated version of the heat equation (Eqn 5.3.1) used to model the cooling of the lithosphere as it moves away from the ridge. This version,

$$\rho C_p \left( \frac{\partial T}{\partial t} + v \nabla T \right) = \nabla \cdot (k \nabla T) + \varepsilon, \quad (1)$$

describes the evolution of the temperature field,  $T(x, y, t)$ , as a function of time and the two space coordinates. In addition to the heat conduction term  $\nabla \cdot (k \nabla T)$ , Eqn 1 includes a  $v \nabla T$  term describing the transfer (or advection) of heat by movement of material, and the  $\varepsilon$  term representing additional sources or sinks of heat such as radioactivity and phase changes. This form allows key parameters such as the density  $\rho$ , specific heat  $C_p$ , thermal conductivity  $k$ , and heat sources or sinks  $\varepsilon$  to vary with position. For a simple analytic solution, we assume that the problem is steady state ( $\partial T / \partial t = 0$ ) and neglect heat sources and sinks ( $\varepsilon = 0$ ). We further assume that the physical properties of the material ( $\rho$ ,  $C_p$ ,  $k$ , and hence the thermal diffusivity  $\kappa = k / \rho C_p$ ) are independent of position.

With these simplifications, Eqn 1 becomes

$$\rho C_p v \frac{\partial T}{\partial x} = k \left( \frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} \right), \quad (2)$$

which has a series solution

$$T(x, y) = T_m \left[ 1 + 2 \sum_{n=1}^{\infty} c_n \exp(-\beta_n x / L) \sin(n\pi y / L) \right], \quad (3)$$

with

$$c_n = (-1)^n / (n\pi), \quad \beta_n = (R^2 + n^2 \pi^2)^{1/2} - R, \quad R = vL / (2\kappa).$$

$R$ , the dimensionless thermal Reynolds number, is the ratio of the rate at which cold material is subducted to that at which it heats up by conduction. This solution resembles the temperature field in the plate model of cooling lithosphere (Eqn 5.3.19), because both models describe the thermal evolution of a plate of finite thickness with temperature boundary conditions at the top, bottom, and one end. In the previous case the plate cools, whereas in this case it heats up.

To find how far along the slab a given isotherm penetrates, we approximate the series by its first term and use the fact that  $R \gg \pi$ , so

$$T(x, y) \approx T_m \left[ 1 - (2/\pi) \exp(-\pi^2 x / (2RL)) \sin(\pi y / L) \right]. \quad (4)$$

Solving for the point where  $\partial T / \partial y = 0$  yields  $y = L/2$ , the middle of the slab. In fact, taking additional terms shows that this point is actually closer to the colder top (Fig. 5.4-3). Using the first-term approximation, a temperature  $T_0$  goes furthest into the subduction zone at

$$T_0(x_0, L/2) = T_m \left[ 1 - (2/\pi) \exp(-\pi^2 x_0 / (2RL)) \right], \quad (5)$$

and reaches a maximum down-dip distance

$$x_0 = -vL^2 / (\pi^2 \kappa) \ln [\pi(T_m - T_0) / (2T_m)]. \quad (6)$$

To convert this distance to depth in the mantle, we multiply by  $\sin \delta$ , where  $\delta$  is the slab dip. This correction converts the subduction rate  $v$  to the slab's vertical descent rate  $v \sin \delta$ . Thus an isotherm's maximum depth is proportional to the subduction rate and the square of the plate thickness, so faster subduction or a thicker slab allows material to go deeper before heating up. If we assume that the square of the plate thickness is proportional to its age, the maximum depth to an isotherm in the downgoing slab is proportional to the vertical descent rate times the age,  $t$ , of the subducting lithosphere.

This idea can be tested by assuming, as we did for spreading center earthquakes, that the maximum depth of earthquakes is temperature-controlled, so earthquakes should cease once material reaches a temperature that is too high. To compare

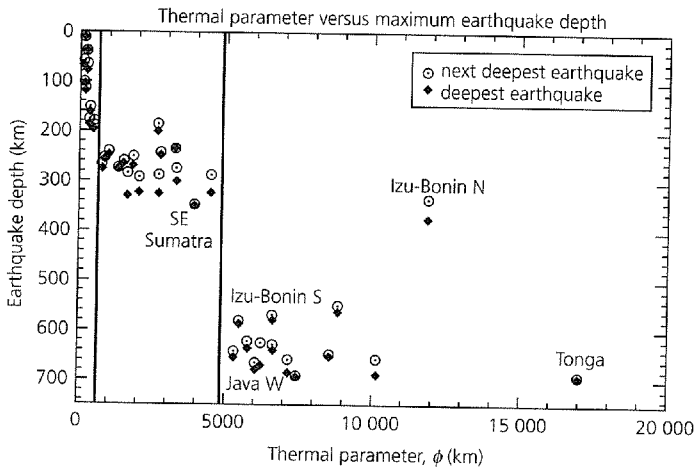


Fig. 5.4-4 Maximum earthquake depths for different subduction zones as a function of thermal parameter, the product of vertical descent rate and lithospheric age. If earthquakes are limited by temperature, this observation is consistent with the simple thermal model's prediction that the maximum depth to an isotherm should vary with the thermal parameter. (After Kirby *et al.*, 1996b. *Rev. Geophys.*, 34, 261–306, copyright by the American Geophysical Union.)

various subduction zones, we examine the maximum depth of earthquakes as a function of their *thermal parameter*

$$\phi = tv \sin \delta. \quad (7)$$

Figure 5.4-4 shows that the maximum depth increases with thermal parameter, and deep earthquakes below 300 km occur only for slabs with a thermal parameter greater than about 5000 km.

However, the fact that the earthquakes stop does not mean that the slab has equilibrated with the surrounding mantle. Figure 5.4-5 shows the predicted minimum temperature within a slab as a function of time since subduction, assuming it maintains its simple planar geometry and does not buckle or thicken. The coldest portion reaches only about half the mantle temperature in about 10 Myr, which is about the time required for the slab to reach 660 km. Thus the restriction of seismicity to depths shallower than 660 km does not indicate that the slab is no longer a discrete thermal and mechanical entity. From a thermal standpoint, there is no reason for slabs not to penetrate into the lower mantle, an issue we discuss shortly. If a slab descended through the lower mantle at the same rate (in fact, it would probably slow down due to the more viscous lower mantle), it would retain a significant thermal anomaly at the core–mantle boundary, consistent with some models of that region (Section 3.8.4).<sup>2</sup>

The thermal model can be improved with simple modifications. Although we assumed that the slab subducts into an isothermal mantle, temperature should increase with depth,

<sup>2</sup> The oceanic lithosphere takes about 70 Myr to cool to equilibrium with the mantle below, and so takes about half that time to heat up again from both sides after it subducts.

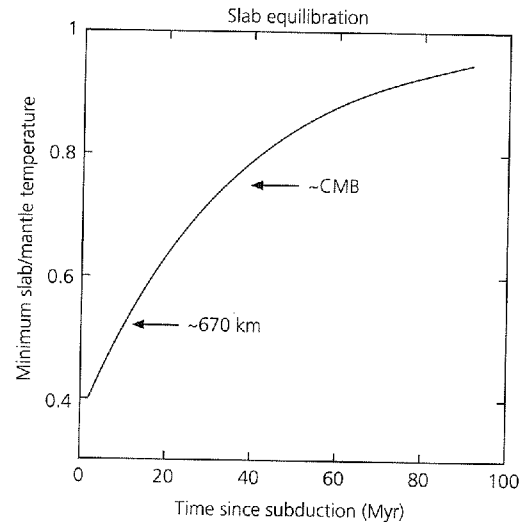


Fig. 5.4-5 Minimum temperature within a slab as a fraction of the mantle temperature, as a function of the time since subduction, computed using the analytic thermal model (Fig. 5.4-3). The coldest portion reaches half the mantle temperature in about 10 Myr, by which time a typical slab is approximately at 670 km depth, and 80% of it in 40 Myr, by which time a slab that continued descending at the same rate would reach the core–mantle boundary. Slabs can thus remain thermally distinct for long periods of time. (Stein and Stein, 1996. *Subduction*, 1–17, copyright by the American Geophysical Union.)

as the material is compressed due to increasing pressure from the overlying rock. Because the mantle below the lithosphere is thought to be convecting, it is often assumed that self-compression occurs adiabatically, such that material moving vertically neither loses nor gains heat. In this case, equilibrium thermodynamics requires that the effects of temperature and pressure changes exactly offset each other,

$$dS = \frac{C_p}{T} dT - \frac{\alpha}{\rho} dP = 0, \quad (8)$$

so that the entropy  $S$  does not change. This condition gives the adiabatic temperature gradient, or *adiabat*, as

$$\left( \frac{dT}{dP} \right)_s = \frac{\alpha}{\rho C_p} T, \quad (9)$$

where  $\alpha$  is the coefficient of thermal expansion. Because pressure increases with depth as  $dP/dz = \rho g$ , temperature increases with depth as

$$\left( \frac{dT}{dz} \right)_s = \frac{\alpha g}{C_p} T. \quad (10)$$

We can thus correct the temperatures for the isothermal mantle case to include adiabatic heating. Using the entropies requires using absolute (Kelvin) temperatures, equal to the Celsius temperature plus 273.15°. Thus if the absolute temperature at

depth  $z_0$ , the base of the plate, is  $T_0^K$ , we integrate Eqn 10 to find the absolute temperature at depth  $z$ ,

$$T^K(z) = T_0^K \exp [(\alpha g / C_p)(z - z_0)]. \quad (11)$$

Another possibly important effect is that of heat sources and sinks. For example, the olivine to spinel transition, which gives rise to the 410 km discontinuity outside the slab, should release heat as it occurs in the slab. Heat might also be generated by friction at the top of the downgoing slab. The heat produced is the product of the subduction rate and the shear stress on the slab interface. The magnitude of this effect is difficult to estimate. It should not be significant unless the shear stress is greater than a few kilobars. As discussed later (Section 5.7.5), the stress on faults is unknown. A further complexity results from the fact that the viscosity of the mantle, which controls the stress, decreases exponentially with temperature. Thus, if frictional heating raises the temperature at the slab interface, viscosity, and hence stress, would decrease, tending to counteract the effect.

To address these complexities, we use numerical models to solve the heat equation at every point in the slab. These models allow parameters such as density to vary with position. In addition, heat sources and sinks such as radioactive heating, phase changes, and frictional heating can be incorporated. The results of such calculations are similar to those of the analytic model and are used to explore how temperatures should vary between subduction zones. For example, Fig. 5.4-6 compares models for a relatively younger and slower-subducting slab (thermal parameter about 2500 km), approximating the Aleutian arc, and an older, faster-subducting slab (thermal parameter approximately 17,000 km), approximating the Tonga arc. As expected, the slab with the higher thermal parameter warms up more slowly, and is thus colder. This prediction is consistent with the observation that Tonga has deep earthquakes, whereas the Aleutians do not (Fig. 5.4-4).

Although we can compute such thermal models, a question is whether they make sense. We test them using two seismological datasets: earthquake locations and seismic velocities. Travel time tomography (Section 7.3) across subduction zones shows high-velocity slabs (Fig. 5.4-7). These results are compared to the velocities predicted using a thermal model of the subducting slab and laboratory values for the variation in velocity with temperature. The model predicts coldest temperatures in the slab interior where the earthquakes occur. Because the tomographic inversion finds the velocity within rectangular cells, the model is converted to that grid and then “blurred” because the seismic rays do not uniformly sample the slab. As shown by the hit count, the number of rays sampling each cell, most rays go down the high-velocity slab, yielding a somewhat distorted image. The fact that this image and the tomographic result are similar suggests that the model is a reasonable description of the actual slab. A similar conclusion emerges from the observation that the tomographic result also resembles parts of the model image that are artifacts, velocity anomalies

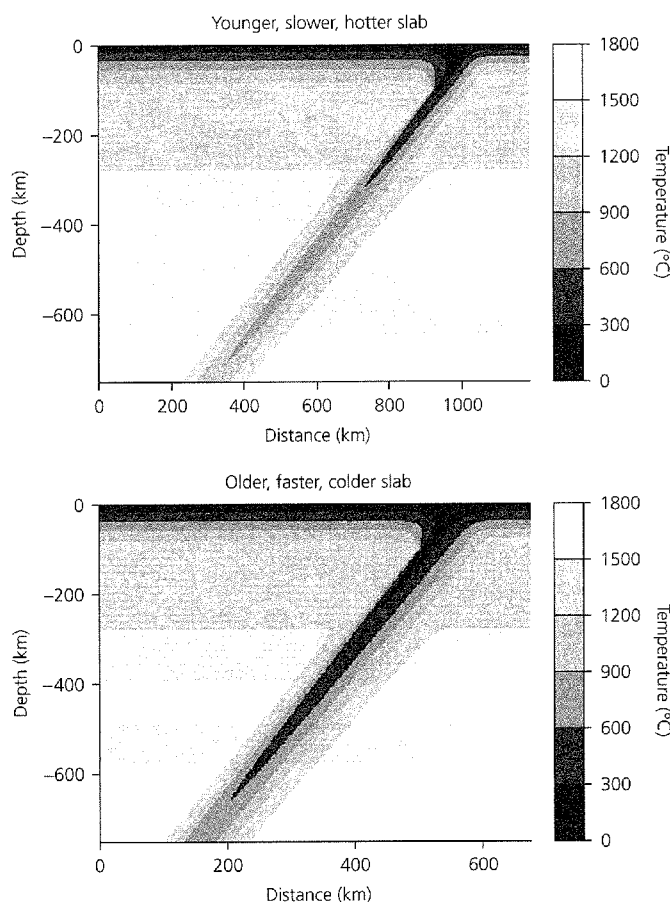


Fig. 5.4-6 Comparison of thermal structure for a relatively younger, slower-subducting slab (50 Myr-old lithosphere subducting at 70 mm/yr; thermal parameter about 2500 km), which approximates the Aleutian arc, and an older, faster-subducting slab (140 Myr-old lithosphere subducting at 140 mm/yr; thermal parameter about 17,000 km) which approximates the Tonga arc. (Stein and Stein, 1996. *Subduction*, 1-17, copyright by the American Geophysical Union.)

that are not present in the original model. These artifacts, generally of low amplitude, cause the slab to appear to broaden, shallow in dip, or flatten out. Hence, although slab thermal models are simplifications of complicated real slabs, and many key parameters are not well known, it seems likely that the models are reasonable approximations (perhaps accurate to a few hundred degrees) to the temperatures within actual slabs.

Seismology provides other tools to study the contrast between the cold, rigid, downgoing plate and the hotter, less rigid material around it. Figure 3.7-20 showed that a cold slab transmits seismic energy with less attenuation than its surroundings. Figure 5.4-8 shows some of the earliest data for this effect: seismograms from a deep earthquake are contrasted at stations NIU, to which waves travel through the downgoing plate; and VUN, to which waves arrive through the surrounding mantle. The VUN record shows much more long-period energy, especially for S waves, than that at NIU. Thus the

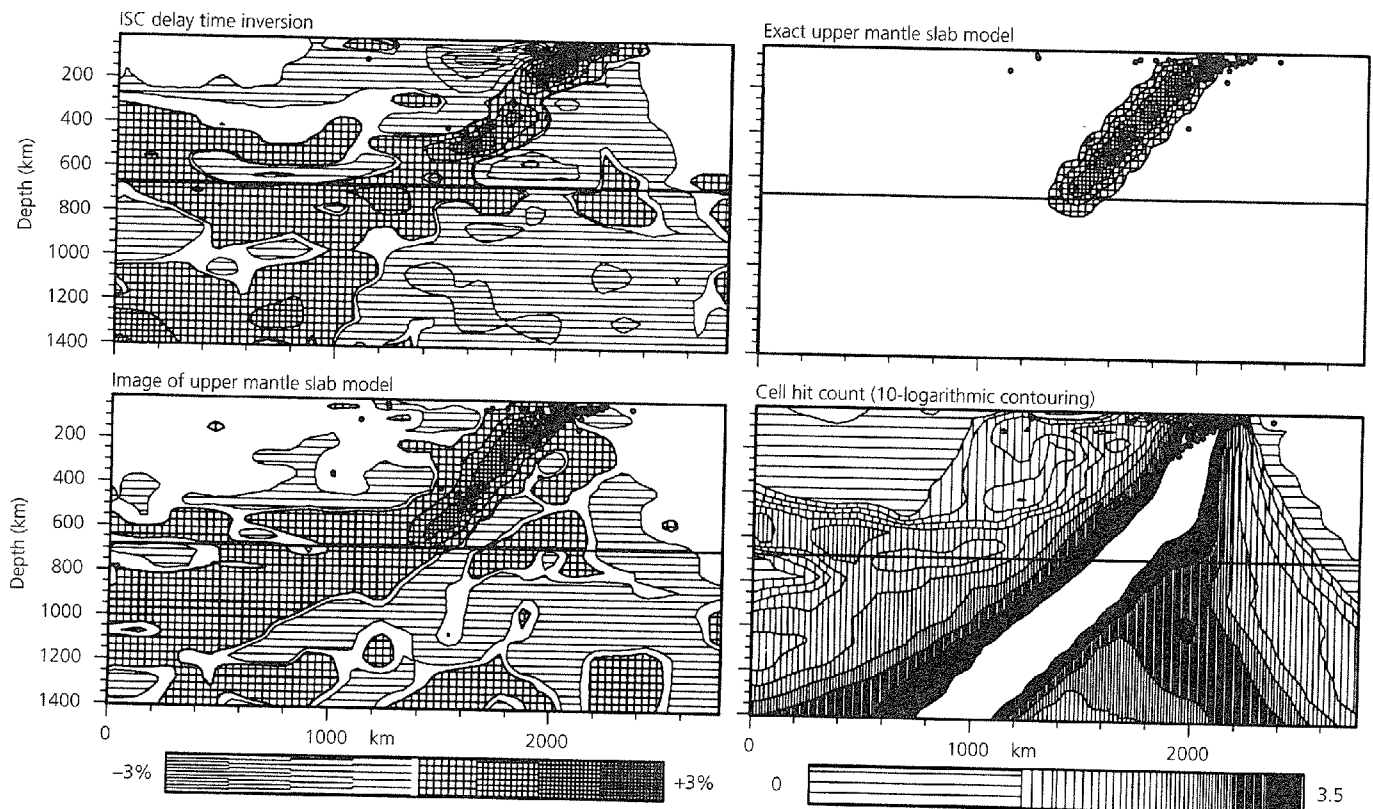


Fig. 5.4-7 Comparison of a seismic tomographic image of a subducting slab, indicated by the velocity anomaly and earthquake hypocenters (dots) (*upper left*) to the image (*lower left*) predicted for a slab thermal model. The seismic velocity anomaly predicted by the thermal model (*upper right*) is imaged by a simulated tomographic study using the same seismic ray path sampling as the data. The hit count (*lower right*) shows the number of rays sampling each cell used in the inversion. As a result of ray geometry and noise, the slab model gives a somewhat distorted image (*lower left*), showing how the model would appear in such a tomographic study. The similarity of the image of the model and the tomographic result suggests that the model generally describes the major features of the actual slab. Left scale bar gives velocity perturbations in percent, with positive values representing fast material. Right scale bar is for hit count, showing values as logarithm to base 10; the white region in the hit count plot is densely sampled and off scale. (Spakman *et al.*, 1989. *Geophys. Res. Lett.*, 16, 1097–110, copyright by the American Geophysical Union.)

high-frequency components were more absorbed on the path to VUN due to higher attenuation (lower  $Q$ ) than on the more rigid slab path to NIU. In addition, the sharp contrast in seismic velocity at the top of slabs can be detected using reflected and converted seismic waves (Fig. 2.6-15).

#### 5.4.2 Earthquakes in subducting slabs

The deep and intermediate earthquakes forming the Wadati-Benioff zone extend in some places to depths of almost 700 km (Fig. 5.4-9). These are the deepest earthquakes that occur: away from subduction zones, earthquakes below about 40 km are rare. The Wadati-Benioff zone earthquakes illustrate that material cold enough to fail seismically (rather than flow) is being subducted, and give our best information about the geometry and mechanics of slabs.

The number of earthquakes as a function of depth illustrates why we distinguish intermediate and deep earthquakes; seismicity decreases to a minimum near about 300 km, and then increases again. Deep earthquakes, those below about 300 km,

are thus generally treated as distinct from intermediate earthquakes. Deep earthquakes peak at about 600 km, and then decline to a minimum before 700 km. The focal mechanisms also vary with depth; those shallower than 300 km show generally down-dip tension, whereas those below 300 km show generally down-dip compression (Fig. 5.4-10).

Various explanations for this distribution of earthquakes and focal mechanisms are under consideration. One is that near the surface the slab is extended by its own weight, whereas at depth it encounters stronger lower mantle material, causing down-dip compression. Another possible factor may be mineral phase changes that occur at different depths in the cold slab than in the surrounding mantle.

It is generally assumed that the most crucial effect is the negative buoyancy (sinking) of the cold and dense slabs. The thermal model gives the force driving the subduction due to the integrated negative buoyancy of a slab resulting from the density contrast between it and the warmer and less dense material at the same depth outside. Because the slab does not have a discrete lower end in the analytic model, the net force is

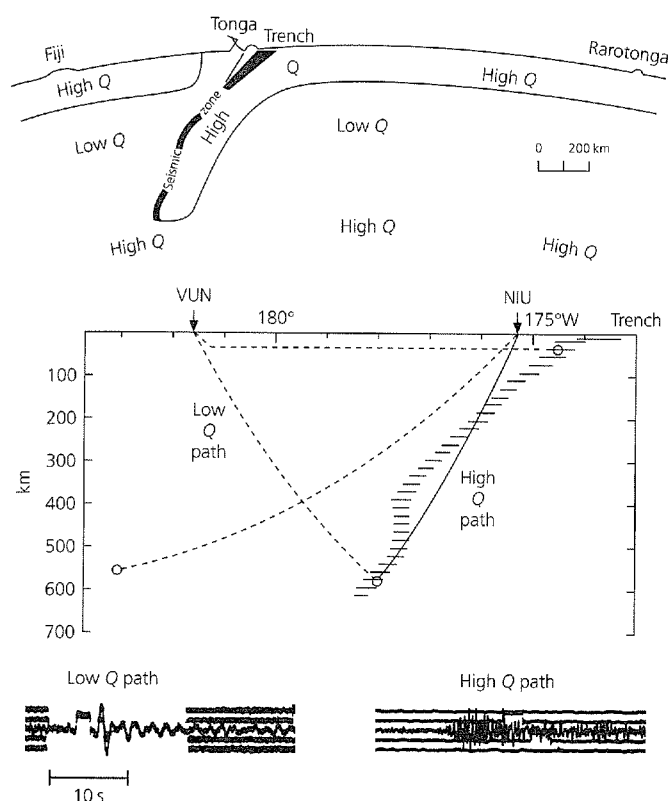


Fig. 5.4-8 Seismological observations showing the difference between the cold slab and hotter ambient mantle. Comparison of the seismograms at NIU and VUN shows that high frequencies are transmitted better by the slab, so the slab is a less attenuating, or higher  $Q$  path. (Oliver and Isacks, 1967. *J. Geophys. Res.*, 72, 4259–75, copyright by the American Geophysical Union.)

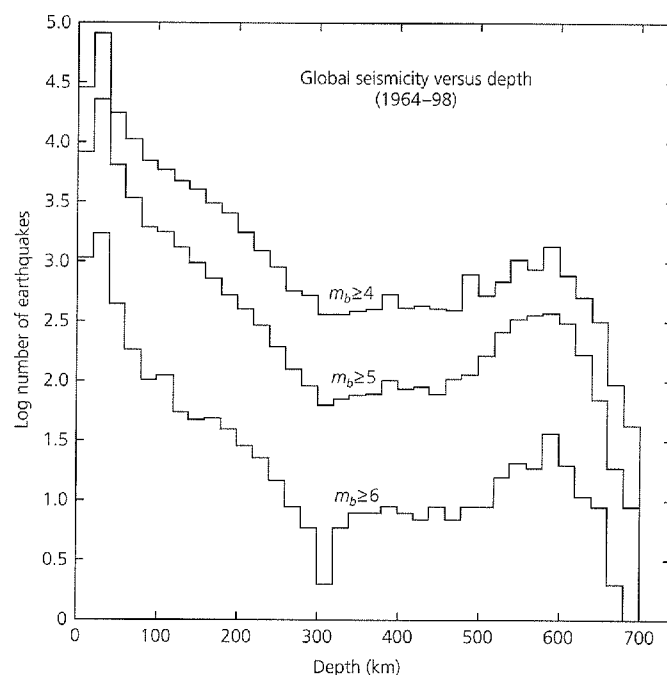


Fig. 5.4-9 Distribution of seismicity with depth.

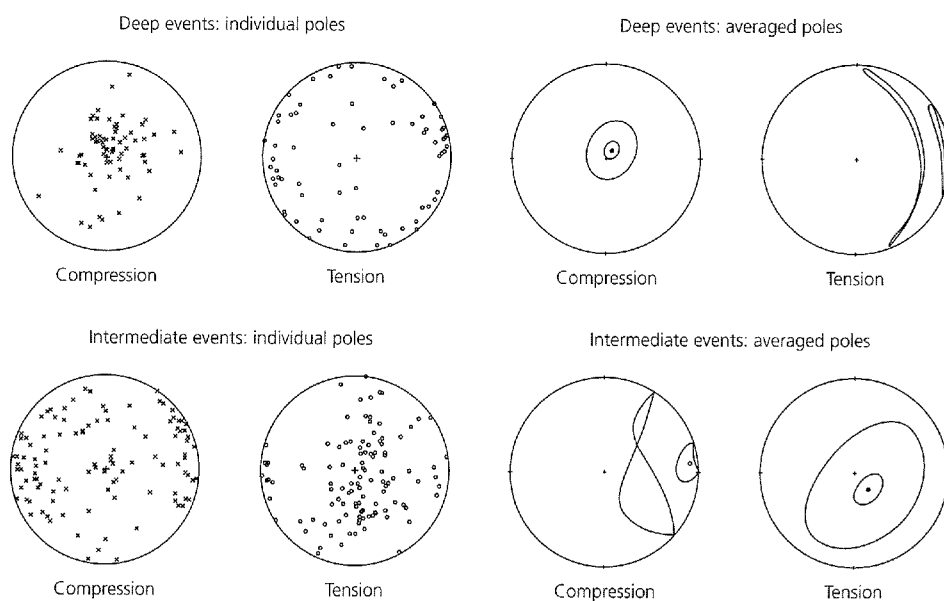


Fig. 5.4-10 Stress orientations inferred from focal mechanisms of subduction zone earthquakes. The P and T axes are rotated so that the down-dip direction is at the center of each plot, and their distributions are contoured. *Top*: Events below 300 km are dominated by down-dip compression. *Bottom*: Events from 70–300 km are dominated by down-dip tension. (After Vassiliou, 1984. *Earth Planet. Sci. Lett.*, 69, 195–201, with permission from Elsevier Science.)



$$F = \int_0^L \int_0^\infty g[\rho(x, y) - \rho_m] dx dy. \quad (12)$$

If material outside the slab is at temperature  $T_m$  and density  $\rho_m$ , material in the slab at the point  $(x, y)$  has density

$$\rho(x, y) \approx \rho_m + \frac{\partial \rho}{\partial T} [T(x, y) - T_m] = \rho_m + \rho'(x, y). \quad (13)$$

As for the cooling plate (Eqn 5.3.9), the density perturbation is

$$\rho'(x, y) = \alpha \rho_m [T_m - T(x, y)], \quad (14)$$

so for the analytic temperature model (Eqn 3) the integral over the slab yields a force

$$F = \frac{g \alpha \rho_m T_m \nu L^3}{24 \kappa}. \quad (15)$$

This force, known as “slab pull,” is the plate driving force due to subduction. Specifically, it is the negative buoyancy associated with a cold downgoing limb of the convection pattern. Its significance for stresses in the downgoing plate and for driving plate motions depends on its size relative to the resisting forces at the subduction zone. There are several such forces. As the slab sinks into the viscous mantle, the material displaced causes a force depending on the viscosity of the mantle and the subduction rate. The slab is also subject to drag forces on its sides and to resistance at the interface between the overriding and downgoing plates, which is often manifested as earthquakes.

To gain insight into the relative size of the negative buoyancy (“slab pull”) and resistive forces, we consider the stress in the downgoing slab and the resulting focal mechanisms. Figure 5.4-11 shows a simple analogy, the stress due to the weight of a vertical column of length  $L$  of material with density  $\rho$ . Using the equilibrium equation (Eqn 2.3.49), we equate the stress gradient to the body force,

$$\frac{\partial \sigma_{zz}(z)}{\partial z} = -\rho g, \quad (16)$$

so the stress as a function of depth is found by integration,

$$\sigma_{zz}(z) = -\rho g z + C, \quad (17)$$

where  $C$  is a constant of integration. To determine  $C$ , and thus the stress in the column, the boundary conditions must be known.

First, suppose the stress is zero at the top,  $z = 0$ . In this case  $C = 0$  and

$$\sigma_{zz}(z) = -\rho g z, \quad (18)$$

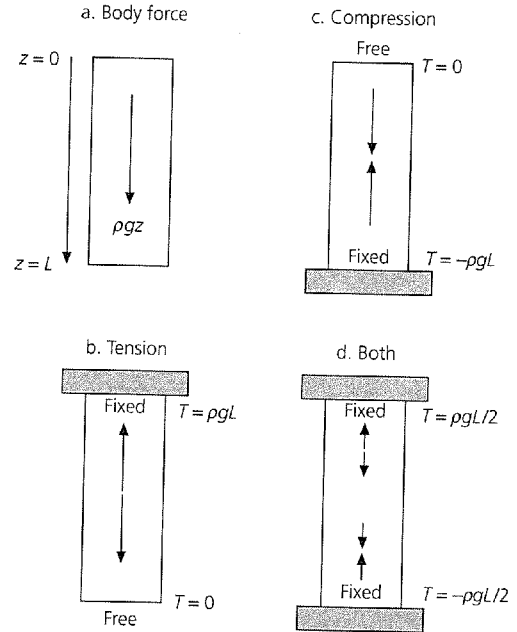


Fig. 5.4-11 Stress within a vertical column of material under its own weight, a simple analogy to stress within a downgoing slab. For the same body force, different stress distributions result from different boundary conditions. If the load is supported at the bottom, the column is under compression; if the support is at the top, the column is under tension. A combination of the two produces a transition.

which is negative, corresponding to compression everywhere. The forces required at the top and the bottom to maintain equilibrium are given by the relation between the traction, stresses, and outward normal vector on a surface (Eqn 2.3.8),

$$T_z = \sigma_{zz} n_z. \quad (19)$$

At the top  $T_z(0) = 0$ , whereas at the bottom a force

$$T_z(L) = -\rho g L \quad (20)$$

holds the column up. This situation is like a column of material sitting on the earth's surface, under compression everywhere.

Alternatively, suppose the stress is zero at the bottom. In this case the constant is chosen so that

$$\sigma_{zz}(z) = \rho g (L - z) \quad (21)$$

and the column is in extension ( $\sigma_{zz}$  positive) everywhere. The force at the bottom is zero, and the force at the top,

$$T_z(0) = \rho g L, \quad (22)$$

supports the column, because  $n_z$  points in the  $-z$  direction. This situation corresponds to the material hanging under its own weight.

If the column is supported equally at both ends, the forces at either end are equal, so we find the stress from the condition

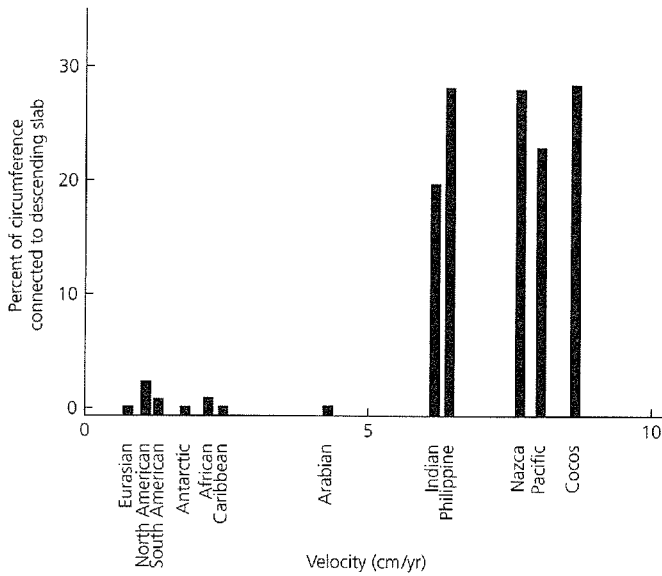


Fig. 5.4-12 The absolute velocity of lithospheric plates increases with the fraction of the plate's boundary formed by subducting slabs, suggesting that slabs provide a major driving force for plate motions. (Forsyth and Uyeda, 1975.)

$$T_z(0) = -T_z(L), \quad (23)$$

which gives

$$\sigma_{zz}(z) = \rho g(L/2 - z). \quad (24)$$

Thus the column is in extension in its upper half,  $z < L/2$ , and in compression below this point.

The stress in the column shows how the body force due to gravity is balanced by forces on the boundaries. By analogy, if the downgoing slab were in tension, the negative buoyancy force must exceed the resistive forces at the subduction zone, and the slab would be "pulling" on and supported by the remainder of the plate outside the subduction zone. In fact, most earthquakes in the deeper portions of the slab show down-dip compression, whereas the intermediate earthquakes show down-dip tension (Fig. 5.4-10). This situation is like the column supported at both ends.

These ideas about the forces within subduction zones are consistent with two important pieces of data. First, the average absolute velocity of plates increases with the fraction of their area attached to downgoing slabs (Fig. 5.4-12), suggesting that slabs are a major determinant of plate velocities. Second, as discussed in Section 5.5.2, earthquakes in old oceanic lithosphere have thrust mechanisms, demonstrating deviatoric compression. Thus the net effect of the subduction zone on the remainder of the plate is not a "pull," so the term "slab pull" is misleading. Instead, as implied by the slab stress models, the "slab pull" force is balanced by local resistive forces, a combination of the effects of the viscous mantle and the interface

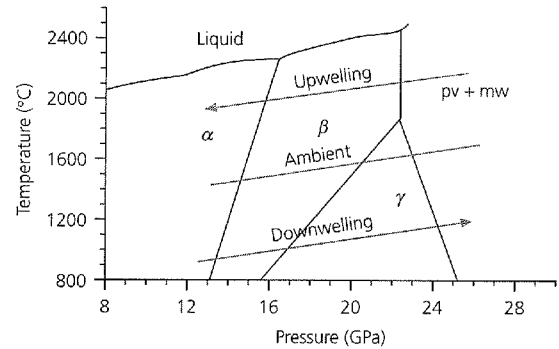


Fig. 5.4-13 Phase diagram for transitions in olivine with increasing depth. The phase boundaries as functions of temperature and pressure are known as Clapeyron curves. The downwelling and upwelling lines contrast conditions in slabs and plumes, respectively, to those in the ambient mantle. A reaction with a positive slope, such as the olivine ( $\alpha$  phase) to spinel ( $\beta$  phase) change thought to give rise to the 410 km discontinuity outside the slab, is displaced upward (to lower pressure) within the cold slab. By contrast, the  $\gamma$  spinel to perovskite plus magnesiowustite (pv + mw) transition has a negative slope, so the 660 km discontinuity should be deeper in slabs than outside. (After Bina and Liu, 1995. *Geophys. Res. Lett.*, 22, 2565–8, copyright by the American Geophysical Union.)

between plates. This situation is like an object dropped in a viscous fluid, which is accelerated by its negative buoyancy until it reaches a terminal velocity determined by its density and shape and the viscosity and density of the fluid.

An interesting possible complication is that slabs are not just thermally different from their surroundings; they are probably also mineralogically different. Slabs extend through the mantle transition zone, where mineral phase changes are thought to occur (Section 3.8). However, because a downgoing slab is colder than material at that depth elsewhere, phase changes within the slab are displaced relative to their normal depth. The displacement can be calculated using the thermodynamic relation, known as the Clapeyron equation, for the boundary between two phases as a function of pressure and temperature. If  $\Delta H$  and  $\Delta V$  are the heat and volume changes resulting from the phase change, then a change  $dT$  in temperature moves the phase change by a pressure  $dP$  given by the Clapeyron slope (the reciprocal of Eqn 9),

$$\gamma = \frac{dP}{dT} = \frac{\Delta H}{T\Delta V}. \quad (25)$$

For example, the 410 km discontinuity is attributed to the phase change with increased pressure from olivine to a denser spinel structure (the  $\beta$  phase, wadsleyite) described by a phase diagram like that in Fig. 5.4-13. Because the spinel phase is denser,  $\Delta V$  is less than zero. This reaction is exothermic (gives off heat), so  $\Delta H$  is also negative, causing a positive Clapeyron slope. If we know the depth (pressure) and temperature at which a phase change occurs in the mantle, the Clapeyron equation gives its position in the slab. The slab is colder than

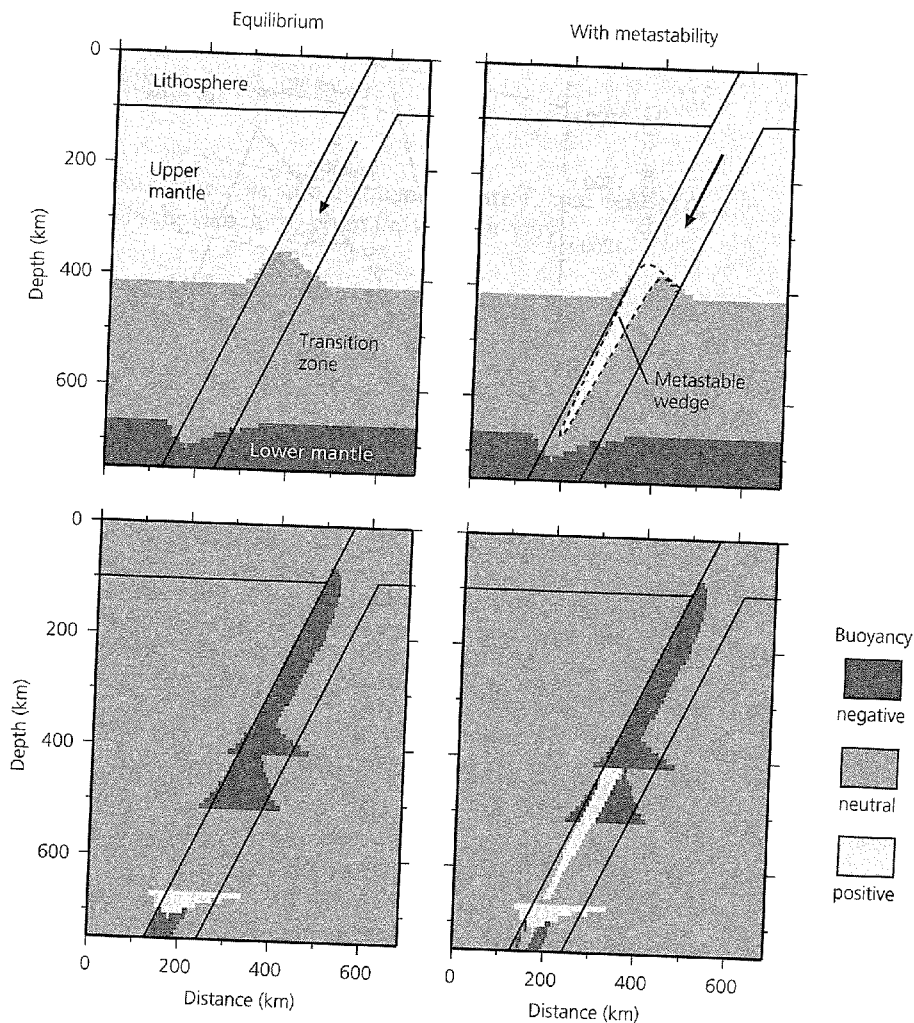


Fig. 5.4-14 Predicted mineral phase boundaries and resulting buoyancy forces in a downgoing slab without (*left panels*) and with (*right panels*) a metastable olivine wedge. Assuming equilibrium mineralogy the cold slab has negative thermal buoyancy, negative compositional buoyancy associated with the elevated 410 km discontinuity, and positive compositional buoyancy associated with the depressed 660 km discontinuity. A metastable wedge gives positive compositional buoyancy and hence decreases the force driving subduction. Negative buoyancy favors subduction, whereas positive buoyancy opposes it. (Stein and Rubie, 1999. *Science*, 286, 909–10, copyright 1999 American Association for the Advancement of Science.)

the ambient mantle ( $dT < 0$ ), so this phase change occurs at a lower pressure ( $dP < 0$ ), corresponding to a shallower depth. Converting the pressure change to depth, the vertical displacement of this phase change is

$$\frac{dz}{dT} = \frac{\gamma}{\rho g} \quad (26)$$

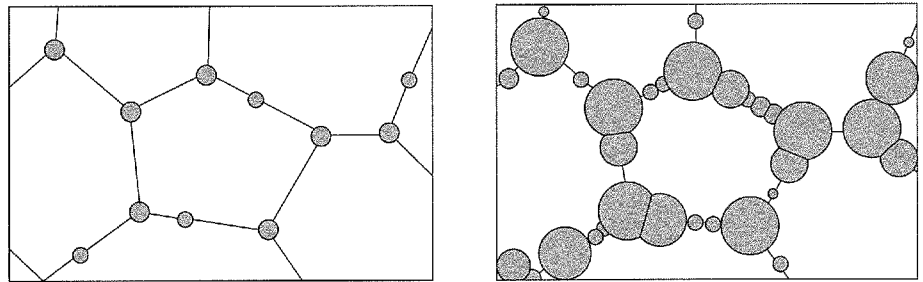
By contrast, the ringwoodite ( $\gamma$  spinel phase) to perovskite plus magnesiowustite transition, thought to give rise to the 660 km discontinuity, is endothermic (absorbs heat), so  $\Delta H$  is positive. Because this is a transformation to denser phases ( $\Delta V$  less than zero), the Clapeyron slope is negative, and the 660 km discontinuity should be deeper in slabs than outside. These opposite effects — upward deflection of the 410 km and downward deflection of the 660 km discontinuities (Fig. 5.4-14) — have been observed in travel time studies. An interesting way to think about these is to note that the negative buoyancy associated with the elevated 410 km discontinuity helps the subduction, whereas the positive buoyancy associated with the depressed 660 km discontinuity opposes the subduction. The reverse effect should not occur at the 660 km discontinuity for

upcoming plumes, however, because the phase diagram shows that at these higher temperatures the Clapeyron curve for the perovskite plus magnesiowustite transition is vertical, so the transition is not displaced (Fig. 5.4-13).

The position of the olivine–spinel phase change may be further affected. The Clapeyron slope predicts what happens if a phase change occurs at equilibrium. However, the phase change actually occurs by a process in which grains of the high-pressure phase nucleate on the boundaries between grains of the lower-pressure phase and then grow with time (Fig. 5.4-15). Studies of mineral nucleation and growth rates suggest that in the coldest slabs the phase transformation cannot keep pace with the rate of subduction, causing a wedge of olivine in the cold slab core to persist metastably<sup>3</sup> to greater depths (Fig. 5.4-14).

<sup>3</sup> *Metastability* describes the situation where a mineral phase survives outside its equilibrium stability field in temperature–pressure space. Such metastable persistence is expected because the relatively colder temperatures in slabs should inhibit reaction rates. This effect explains why diamonds, which are unstable at the low pressures of earth's surface, survive metastably rather than transform to graphite. The situation in slabs is similar to that of supercooled water, which persists as a liquid at temperatures below its equilibrium freezing point.

Fig. 5.4-15 Diagram showing the early stages of a phase transformation. Grains of the new phase (shaded) nucleate on grain boundaries and grow by consuming the original phase until none remains. (Kirby *et al.*, 1996b. *Rev. Geophys.*, 34, 261–306, copyright by the American Geophysical Union.)



The deflections of the phase boundaries have several possible consequences. First, phase changes affect the thermal structure of the slab due to the heat of the phase change. Thus the exothermic olivine–spinel change should add heat to slabs. This effect is simulated in thermal models by increasing the temperature at the phase change. Second, the phase boundaries are probably important for the buoyancy and stresses within slabs. We have already discussed the idea that the cold slabs are denser than their surroundings, causing negative thermal buoyancy, which favors sinking. The phase boundaries cause additional mineralogical buoyancy. For example, if the olivine–spinel boundary is uplifted in the slab, the presence of slab material denser than at that depth outside causes additional negative buoyancy. However, if a wedge of metastable olivine exists, it would be less dense than material at

that depth outside and produce positive buoyancy (Fig. 5.4-14) in addition to that caused by the downward deflection of the 660 km discontinuity. Although the net buoyancy must be negative because slabs subduct, the details of the buoyancy can be important. For example, metastable olivine may help regulate subduction rates. Faster subduction would cause a larger wedge of low-density metastable olivine, reducing the driving force and slowing the slab.

A third possibility is that a phase change causes deep earthquakes. Although this idea is a natural consequence of the observation that deep earthquakes occur at transition zone depths, it was not given serious consideration for a long time because deep earthquake focal mechanisms show slip on a fault, rather than isotropic implosions (Section 4.4.6). However, laboratory studies now suggest that an instability

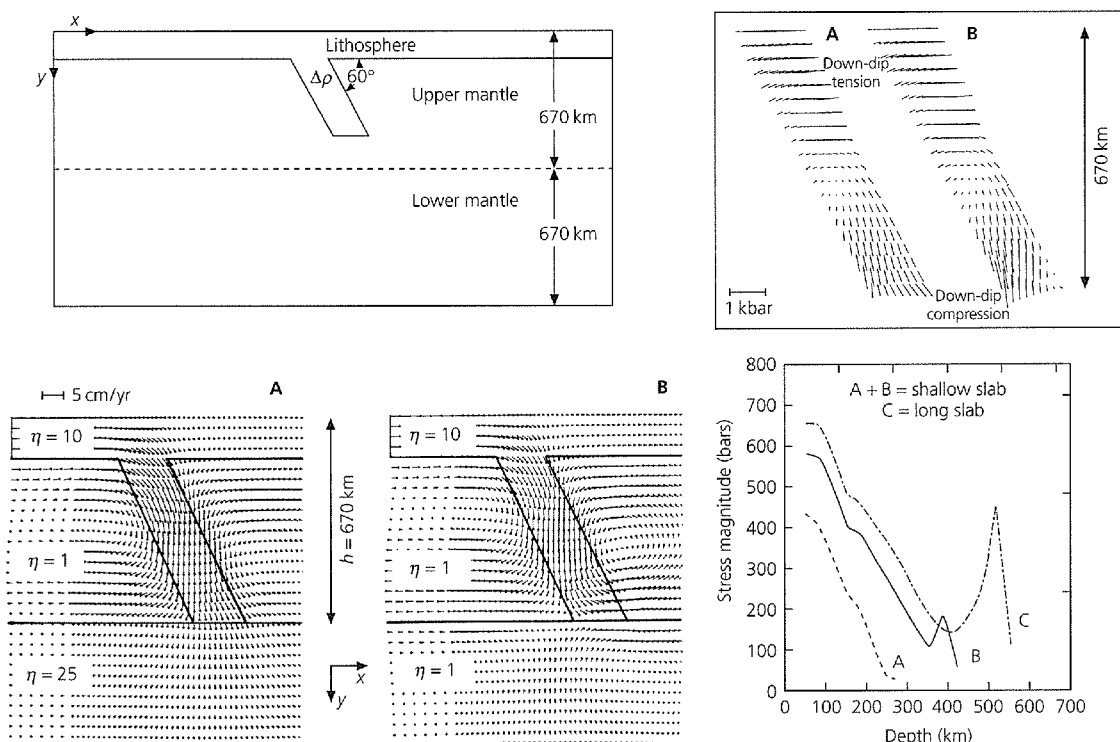


Fig. 5.4-16 Numerical model of mantle flow fields (lower left) and resulting stresses (upper right) within a downgoing slab for the cases of a slab that (A) encounters higher-viscosity material below 670 km and (B) cannot penetrate below this depth.  $\eta$  values show relative viscosities. Both predict down-dip tension in the upper portion of the slab and down-dip compression in the lower portion. The calculated stresses are highest near the bottom of the slab. (Vassiliou *et al.*, 1984. *J. Geodynam.*, 1, 11–28, with permission from Elsevier Science.)

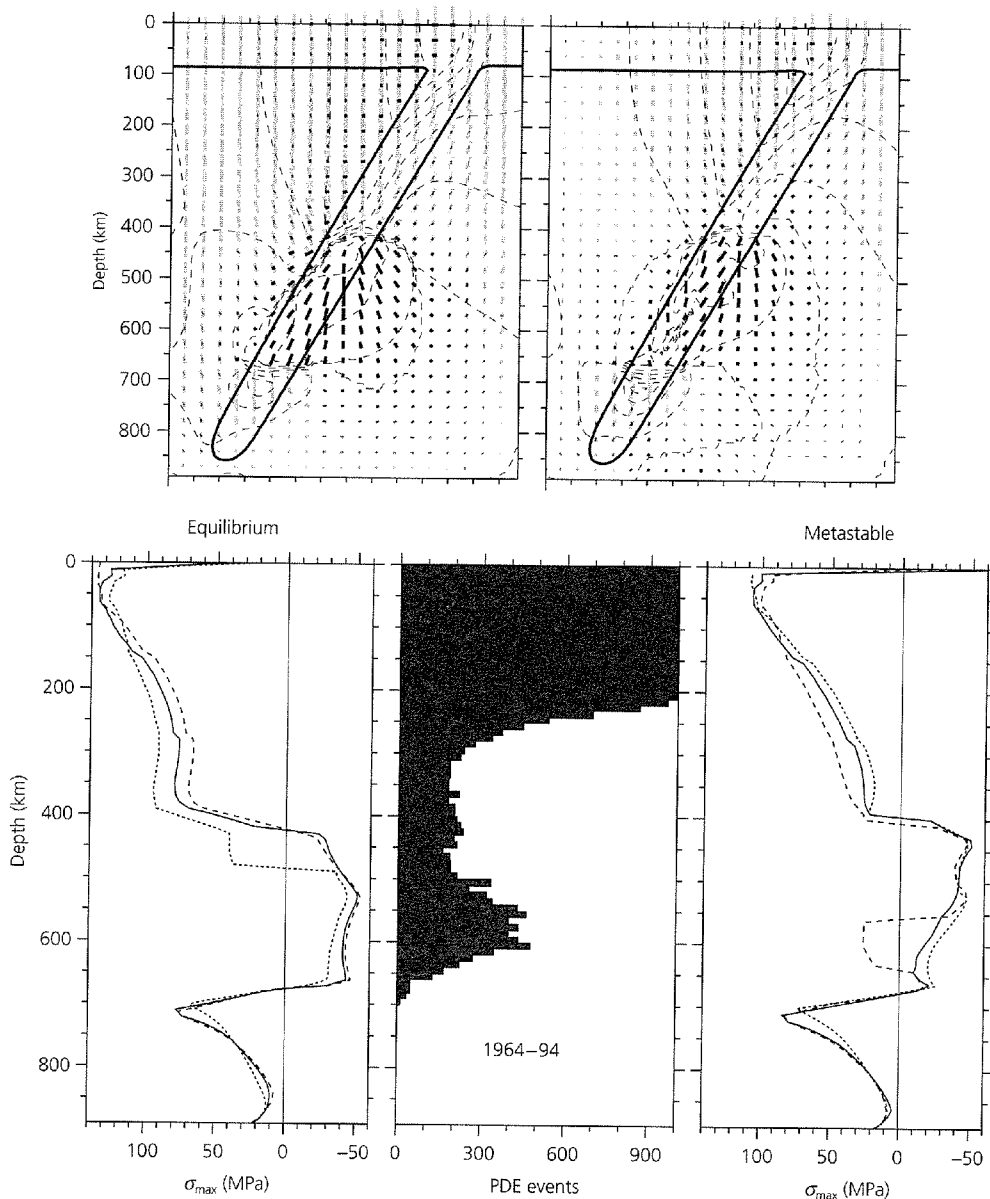
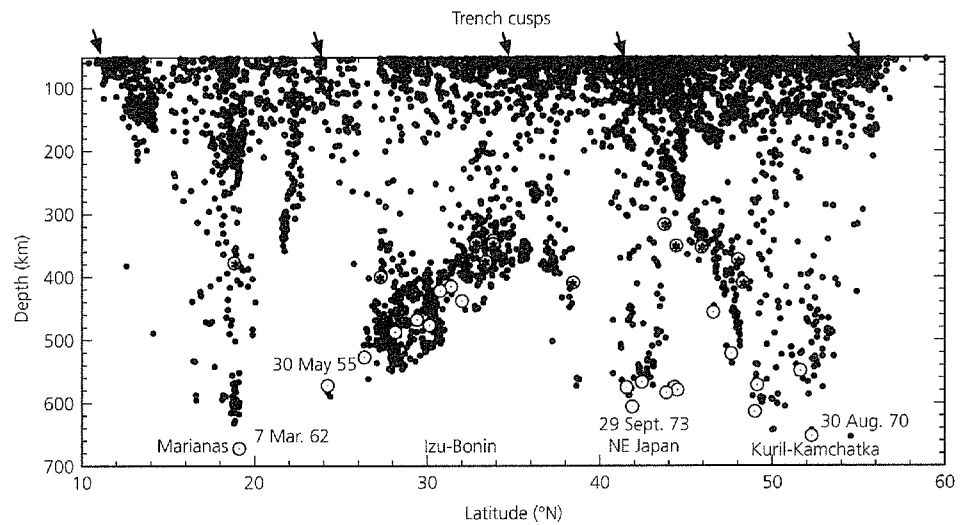


Fig. 5.4-17 Numerical models of stresses within a downgoing slab assuming the density distribution corresponding to equilibrium mineralogy (*left panels*) and with metastable olivine (*right panels*). Upper panels show stress orientations, and lower panels show stress magnitudes, with compression as negative, compared to the distribution of seismicity (*lower center*). (Bina, 1997. *Geophys. Res. Lett.*, 24, 3301–4, copyright by the American Geophysical Union.)

called transformational faulting can cause slip along thin shear zones where metastable olivine transforms to denser spinel. Such faulting can occur for the exothermic olivine to spinel transition, but not for the endothermic spinel to perovskite plus magnesiowustite transition, so deep earthquakes would occur only in the transition zone. Because the metastable wedge's lower boundaries are essentially isotherms, this model offers a physical mechanism for the observation (Fig. 5.4-4) that the depth of earthquakes increases with thermal parameter. This idea is attractive, but to date seismological studies show no evidence for a metastable wedge, and large deep earthquakes occur on fault planes that appear to extend beyond the boundaries of the expected metastable wedge. If such wedges exist, earthquakes may nucleate by transformational faulting, but then propagate outside the wedge via another failure mechanism.

Together these ideas offer several possible explanations for features of slab earthquakes. One key feature is the depth variation in seismicity and focal mechanisms. The first explanation is that the depth distribution and stresses are largely due to the negative thermal buoyancy of slabs and their encountering either a region of much higher viscosity or a barrier to their motion at the 660 km discontinuity. Numerical models (Fig. 5.4-16) predict stress orientations similar to those implied by the focal mechanisms. Moreover, the magnitude of the stress varies with depth in a fashion similar to the depth distribution of seismicity — a minimum at 300–410 km and an increase from 500 to 700 km. Alternatively, numerical models including the buoyancy effects of the phase changes (Fig. 5.4-14) also predict a similar variation in stress magnitude and orientation with depth (Fig. 5.4-17), without invoking a barrier or higher viscosity in the lower mantle. Thus, in such

Fig. 5.4-18 North-south cross-section showing seismicity of subduction zones of the Northwest Pacific. Seismicity shallows near the cusps where arcs meet, making individual Wadati-Benioff zones tongue-shaped. Large deep earthquakes ( $M_0$  greater than  $10^{26}$  dyn-cm), shown by open circles, tend to be at the edges or bottoms of deep seismicity, or isolated from the main Wadati-Benioff zones. (Kirby *et al.*, 1996b. *Rev. Geophys.*, 34, 261–306, copyright by the American Geophysical Union.)



models, deep earthquakes need not be physically different from intermediate ones, because the minimum in seismicity reflects a stress minimum.

A second key issue is how deep earthquakes can occur at all. As discussed in Section 5.7, the strength of rock that must be exceeded for fracture increases with pressure. The pressures deep in a subducting slab should be high enough to prevent fracture. One possibility is that the slabs become hot enough that water released by decomposition of hydrous minerals lubricates (reduces the effective stress on) faults. Another possibility, mentioned earlier, is transformational faulting in metastable olivine. It is also possible that the earthquakes occur by very rapid creep, possibly associated with weakening due to unusually small spinel grains formed in the coldest slabs.

The different explanations offered by these models all have attractive features and may be true in part. However, although such simple models based on idealized slabs explain some gross features of deep earthquakes, none fully explains the complexity of deep earthquakes. As shown by Fig. 5.4-18, a cross-section along the subduction zones of the Northwest Pacific, deep seismicity is “patchy” and variable. For example, it shallows dramatically at the cusps between the Marianas, Izu-Bonin, NE Japan, and Kuril-Kamchatka arcs. Moreover, the largest earthquakes occur at the edges of the regions of deep seismicity, as especially evident at the northern edge of the Izu-Bonin seismicity. These sites may reflect tears in the down-going lithosphere at the junctions between arcs, where hot mantle material penetrates slabs. A further complexity is that some deep earthquakes occur in unusual locations off the down-dip extension of the main Wadati-Benioff zones and have focal mechanisms differing from those of the deepest earthquakes in the main zone (Fig. 5.4-19). Some other deep earthquakes are isolated from actively subducting slabs. Such unusual earthquakes may occur in slab fragments where metastable olivine survives, and thus have mechanisms related to local stresses rather than those expected for continuous slabs.

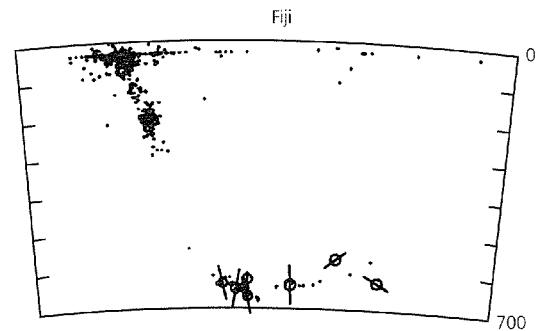


Fig. 5.4-19 Seismicity cross-section for the Fiji subduction zone, showing “outlier” deep earthquakes. Lines through symbols show P axes, which often differ from those for the main Wadati-Benioff zone. (Lundgren and Giardini, 1994. *J. Geophys. Res.*, 99, 15, 833–42, copyright by the American Geophysical Union.)

Another interesting observation from precise earthquake locations in some subduction zones (Fig. 5.4-20) shows that the Wadati-Benioff zone is made up of two distinct planes, separated by 30–40 km. The upper plane seems to coincide with the conversion plane for  $ScSp$  (Fig. 2.6-15), a sharp velocity contrast that is presumably near the slab top. Focal mechanisms suggest that the upper plane is in down-dip compression and the lower one in down-dip extension. A variety of models have been proposed. One is that the double plane results from “unbending” of the slab — the release of the bending stresses produced when the slab began to subduct. Another model is that the slab “sags” under its own weight, because at depth it runs into a more viscous mesosphere, while at intermediate depths it encounters a less viscous asthenosphere. Explaining the phenomenon is complicated by the observation that only some subduction zones have double zones.

The nature of deep earthquakes, especially the mechanism restricting them to the transition zone, has implications for

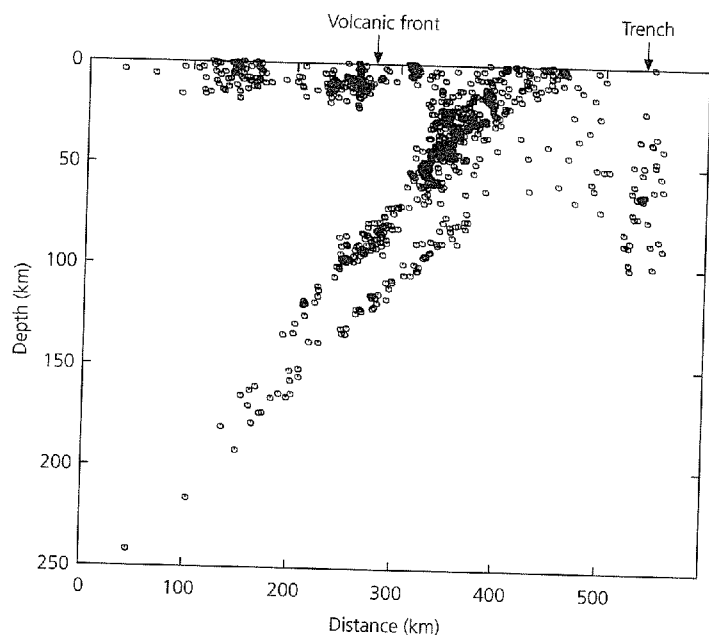


Fig. 5.4-20 Double seismic zone beneath Tohoku, Japan. (Hasegawa *et al.*, 1978. *Tectonophysics*, 47, 43–58, with permission from Elsevier Science.)

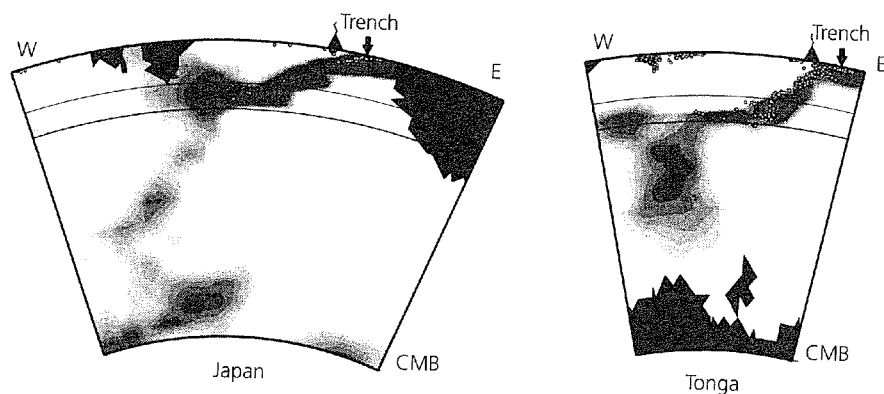


Fig. 5.4-21 Tomographic images across Pacific subduction zones with deep earthquakes. Horizontal lines are at 410 and 660 km depth. White dots are earthquake hypocenters. The Wadati-Benioff zone seismicity generally coincides with the high-velocity anomaly (dark regions) due to the cold subducting slab. Slabs are deflected at the base of the transition zone before penetrating into the lower mantle. (van der Hilst *et al.*, 1998. *The Core-Mantle Boundary Region*, 5–20, copyright by the American Geophysical Union.)

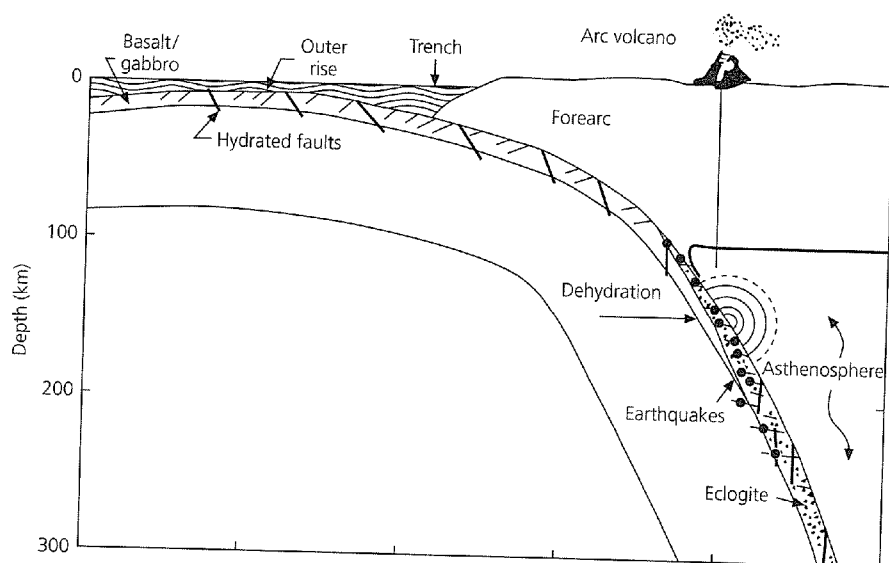


Fig. 5.4-22 Schematic model for intermediate depth earthquakes. Earthquakes are assumed to occur in subducting crust and be associated with the dehydration of mineral phases and the gabbro to eclogite transition. (Kirby *et al.*, 1996a. *Subduction*, 195–214, copyright by the American Geophysical Union.)



schematic model in which the earthquakes are presumed to occur in subducting oceanic crust, rather than throughout the subducting mantle that makes up most of the slabs, because detailed location studies show that the earthquakes are close to the top of the subducting slabs. The crust should undergo two important mineralogical transitions as it subducts. Hydrous (water-bearing) minerals formed at fractures and faults should warm up and dehydrate. Eventually, the gabbro transforms to eclogite, a rock of the same chemical composition composed of denser minerals.<sup>4</sup> Under equilibrium conditions, eclogite should form by the time slab material reaches about 70 km depth. However, travel time studies in some slabs find a low-velocity waveguide interpreted as subducting crust extending to deeper depths. Hence it has been suggested that the eclogite-forming reaction is slowed in cold downgoing slabs, allowing gabbro to persist metastably. Once dehydration occurs, the freed water weakens the faults, favoring earthquakes and promoting the eclogite-forming reactions. In this model the intermediate earthquakes occur by slip on faults, but the phase changes favor faulting. The extensional focal mechanisms may also reflect the phase change, which would produce extension in the subducting crust. Support for this model comes from the fact that the intermediate earthquakes occur below the island arc volcanoes, which are thought to result when water released from the subducting slab causes partial melting in the overlying asthenosphere.

The fact that various explanations are under discussion illustrates the difficulty in understanding the complex thermal structure, mineralogy, rheology, and geometry of real slabs. We can think of the deep subduction process as a chemical reactor that brings cold shallow minerals into the temperature and pressure conditions of the mantle transition zone, where these phases are no longer thermodynamically stable (Fig. 5.4-23). Because we have no direct way of studying what is happening and what comes out, we seek to understand this system by studying earthquakes that somehow reflect what is happening. This is a major challenge, and we have a long way to go.

#### 5.4.3 Interplate trench earthquakes

Much of what is known about the geometry and mechanics of the interaction between plates at subduction zones comes from the distribution and focal mechanisms of shallow earthquakes at the interface between the plates. These include the largest earthquakes that occur, as illustrated by Fig. 5.4-24, showing the largest earthquakes (surface wave magnitude greater than 8.0) during 1904–76. Among these are the two largest earthquakes ever recorded seismologically: the 1960 Chilean ( $M_0$   $2 \times 10^{30}$  dyn-cm,  $M_s$  8.3) and 1964 Alaska ( $M_0$   $5 \times 10^{29}$  dyn-cm,  $M_s$  8.4) earthquakes. Figure 5.4-25 shows the geometry of the Chilean earthquake: 21 meters of slip occurred on a fault

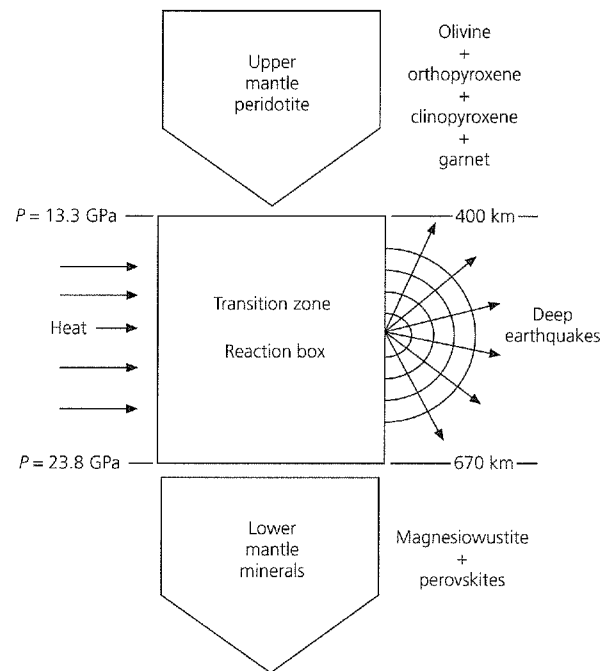


Fig. 5.4-23 Cartoon of subducting slabs in the transition zone as a chemical reactor. (Kirby *et al.*, 1996b. *Rev. Geophys.*, 34, 261–306, copyright by the American Geophysical Union.)

800 km long along strike, and 200 km wide down-dip. The mechanism shows thrusting of the South American plate over the subducting oceanic lithosphere of the Nazca plate. The aftershock zone was 800 km long, and the surface deformation was dramatic, reaching 6 meters of uplift in places. Thrust earthquakes of this type, although smaller, make up most of the large, shallow events at subduction zones. Such *interplate* earthquakes release the plate motion that has been locked at the plate interface. As we saw in Section 4.6.1, these can be much bigger than the largest earthquakes at transform fault boundaries like the San Andreas. For example, even the 1906 San Francisco earthquake was tiny (100 times smaller seismic moment) compared to the 1964 Alaska earthquake, although both occurred along different segments of the same plate boundary. The difference reflects the fact that faulting occurs only when rock is cooler than a limiting temperature. Thus a vertically dipping transform like the San Andreas has a much shorter cold down-dip extent than the shallow-dipping thrust interfaces (sometimes called megathrusts) at subduction zones.

Major thrust earthquakes at the interface between subducting and overriding plates directly indicate the nature of subduction. In most cases, their focal mechanisms show slip toward the trench, approximately in the convergence direction predicted by global plate motion models or space-based geodesy (Section 5.2) (Fig. 5.2-3). However, in some cases when the plate motion is oblique to the trench, a forearc sliver moves separately from the overriding plate (Fig. 5.4-26). This effect,

<sup>4</sup> Most of the oceanic crust consists of gabbro, the intrusive version of the extrusive basalt seen at mid-ocean ridges (Section 3.2.5). With increasing pressure, gabbro becomes eclogite as feldspar and pyroxene transform to garnet.



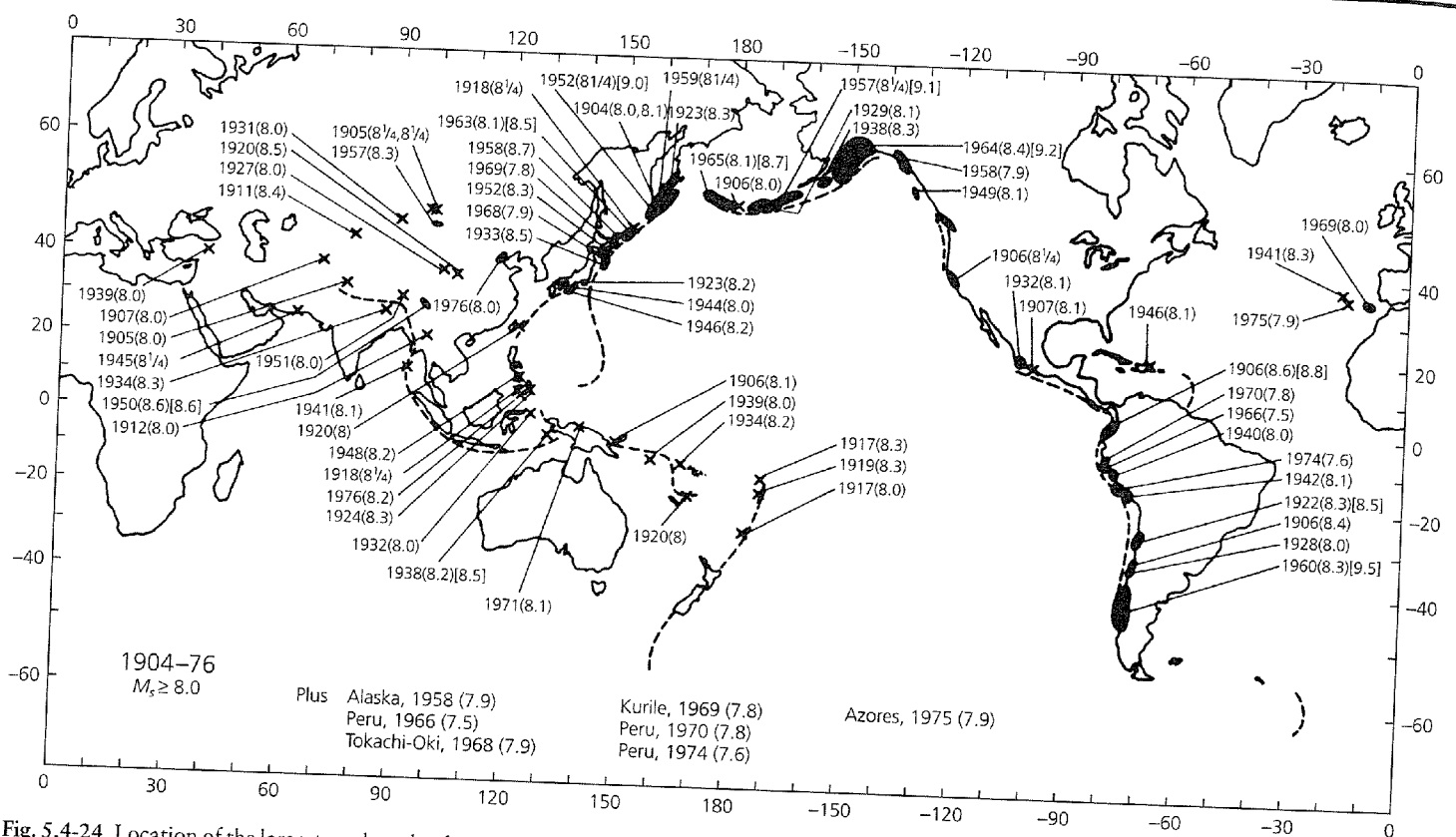


Fig. 5.4-24 Location of the largest earthquakes between 1904 and 1976.  $M_s$  values are in parentheses and  $M_w$  values in square brackets. Most are at subduction zones and result from thrust faulting at the interface between the two plates. (Kanamori, 1978. Reproduced with permission from *Nature*.)

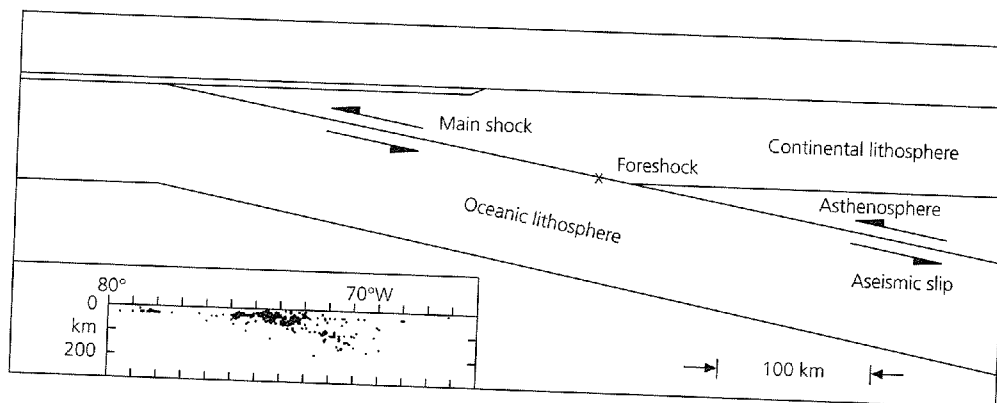


Fig. 5.4-25 Fault geometry and aftershock distribution (insert) for the 1960 Chilean earthquake. (Kanamori and Cipar, 1974. *Phys. Earth Planet. Inter.*, 9, 128-36, with permission from Elsevier Science.)

called *slip partitioning*, makes earthquake slip vectors at the trench trend between the trench-normal direction and the predicted convergence direction, and causes strike-slip motion between the forearc and the stable interior of the overriding plate. This effect can be seen in plate motion studies and with GPS data, and can cause misclosure of plate circuits. In the limiting case of pure slip partitioning, pure thrust faulting would occur at the trench, and all the oblique motion would be accommodated by trench-parallel strike-slip.

How the thrust earthquakes release the accumulated plate motion is both interesting scientifically and important for assessing earthquake hazards. In many subduction zones,

thrust earthquakes have characteristic patterns in space and time. For example, large earthquakes have occurred in the Nankai trough area of southern Japan approximately every 125 years since 1498 with similar fault areas (Fig. 5.4-27).<sup>5</sup> In some cases the entire region seems to have slipped at once; in others, slip was divided into several events over a few years.

<sup>5</sup> Due to its location where between four and six plates (North America, Pacific, Philippine, Eurasia, and perhaps Okhotsk and Amuria) interact, Japan has a high level of seismicity, which was originally attributed to the motion of the namazu, a giant underground catfish. As a result, Japan has an outstanding tradition of seismology and some of the best data in the world for studying subduction-related earthquakes.

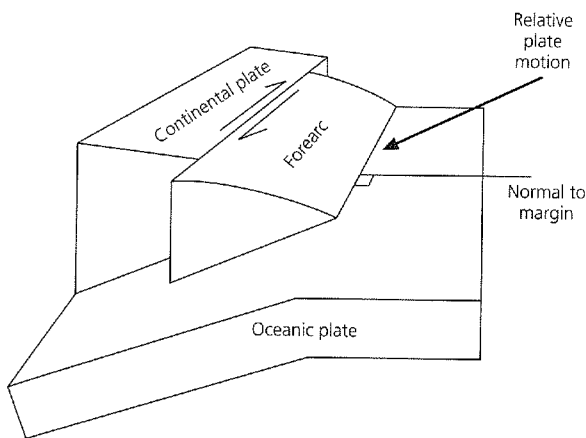


Fig. 5.4-26 Schematic illustration of forearc sliver motion when convergence is oblique. (Courtesy of D. Davis.)

Given such repeatability, it seems likely that a segment of a subduction zone that has not slipped for some time constitutes a seismic gap and is due for an earthquake. For example, the Tokai area (segment D) may be such a case and is the focus of extensive earthquake prediction studies. However, despite the intuitive appeal of the gap idea, efforts to predict the location of future earthquakes using it have not generally been successful (Sections 1.2.5, 4.7.3).

One difficulty is that not all of the plate motion occurs seismically. Figure 5.4-28 shows that during 1952–73 a large segment of the Kuril trench slipped in a series of six major earthquakes with similar thrust fault mechanism. Seismic moment studies show that the average slip was 2–3 meters. Since the previous major earthquake sequence in the area occurred about 100 years earlier, the average seismic slip rate is 2–3 cm/yr, about one-third of the plate motion predicted from relative motion models. The remaining two-thirds of the slip occurs aseismically, as postseismic or interseismic motion. Similar studies around the world find that the fraction of plate motion that occurs as seismic slip, sometimes called the *seismic coupling* factor, is generally much less than 1, implying that much of the plate motion occurs aseismically if the time interval sampled is adequate.

The Chilean subduction zone shows the other extreme. The seismic slip rate, estimated from the slip in the great 1960 earthquake and historical records indicating that major earthquakes occurred about every 130 years during the past 400 years, exceeds the convergence rate predicted by plate motion models (Fig. 5.4-29). Because the convergence rate is an upper bound on the seismic slip rate, the two estimates are inconsistent. One possibility is that the seismic slip is overestimated: either the earlier earthquakes were significantly smaller than the 1960 event or their frequency in the past 400 years is higher than the long-term average.

More generally, these examples illustrate the difficulty in inferring seismic slip from historical seismicity, owing to pro-

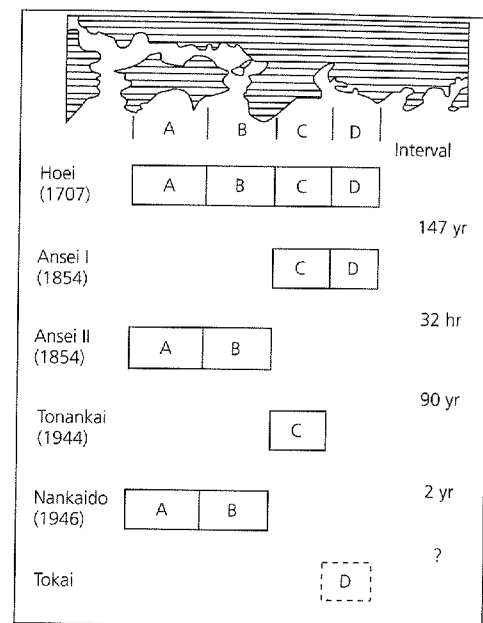


Fig. 5.4-27 Time sequence of large subduction zone earthquakes along the Nankai trough, suggesting both some space and time periodicity and some variability. (Ando, 1975. *Tectonophysics*, 27, 119–40, with permission from Elsevier Science.)

blems including the variability of earthquakes on a given plate boundary, the issue of whether the time sample is long enough, and the difficulty in estimating source parameters for earthquakes that pre-dated instrumental seismology. Given the uncertainties in estimating the slip in an earthquake even with seismological data (Section 4.6), doing so without such data is particularly challenging. An alternative approach to estimating plate coupling, discussed in Sections 4.5.4 and 5.6.2, uses GPS geodesy to measure the deflection of the overriding plate, which will be released in future large earthquakes. This deflection depends on the mechanical coupling at the interface, so directly measures what we infer indirectly from the earthquake history. However, the GPS data sample only the present earthquake cycle, which may not be representative of long-term behavior.

Perhaps for similar reasons, efforts to interpret the seismic slip fraction in terms of the physical processes of subduction have not yet been successful. Although the term “seismic coupling” implies a relation between the seismic slip fraction with properties such as the mechanical coupling between the subducting and overriding lithospheres, this has been hard to establish. This relation was originally posed in terms of two end members: coupled Chilean-type zones with large earthquakes and uncoupled Mariana-style zones with largely aseismic subduction. The largest subduction zone earthquakes appear to occur where young lithosphere subducts rapidly (Fig. 5.4-30, *top*), where we might expect the minimum “slab pull” effects and hence the strongest coupling. However,

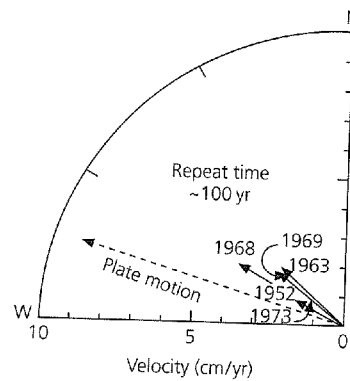
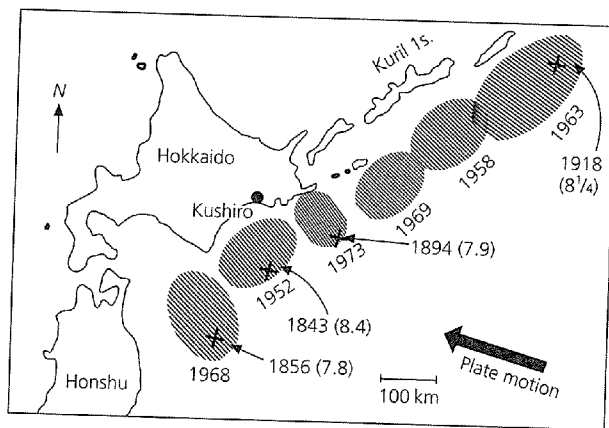


Fig. 5.4-28 Rupture areas for a sequence of large subduction zone earthquakes along the Kuril trench. Different segments of the boundary slip seismically over time. Arrows show the direction and rate of seismic slip and plate motion. If such sequences occur about every 100 years and this time sample is representative, the seismic slip is only about one-third of the plate motion. (Kanamori, 1977b. *Island Arcs, Deep Sea Trenches and Back Arc Basins*, 163-74, copyright by the American Geophysical Union.)

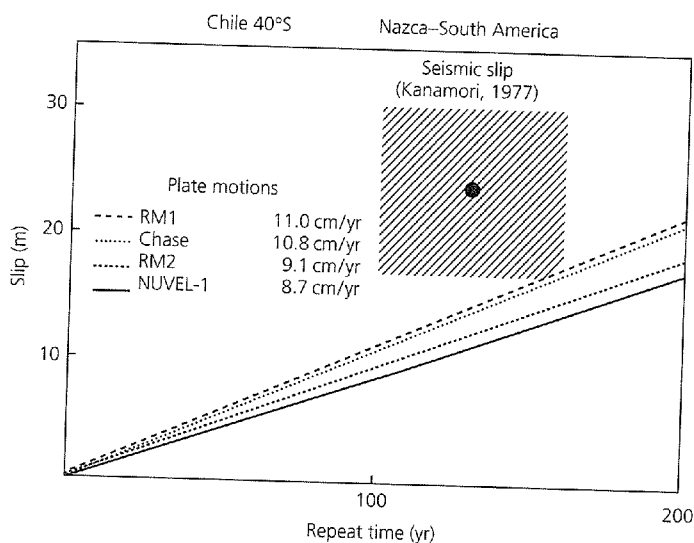


Fig. 5.4-29 Comparison of seismic slip rate and plate motions for the area of the great 1960 Chilean earthquake. Shaded region gives slip rate estimated from slip in the 1960 event and recurrence of large trench earthquakes in the last 400 years. The estimated slip rate exceeds that predicted by any of the four plate motion models shown. (Stein *et al.*, 1986. *Geophys. Res. Lett.*, 13, 713-16, copyright by the American Geophysical Union.)

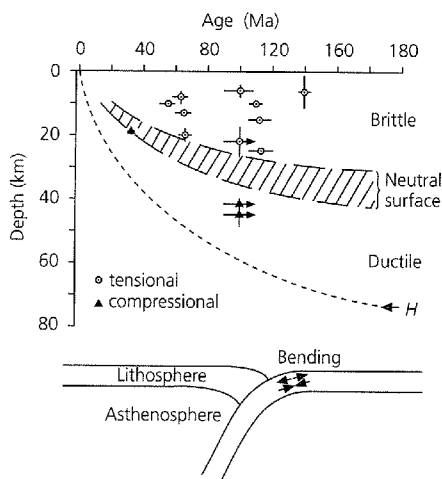
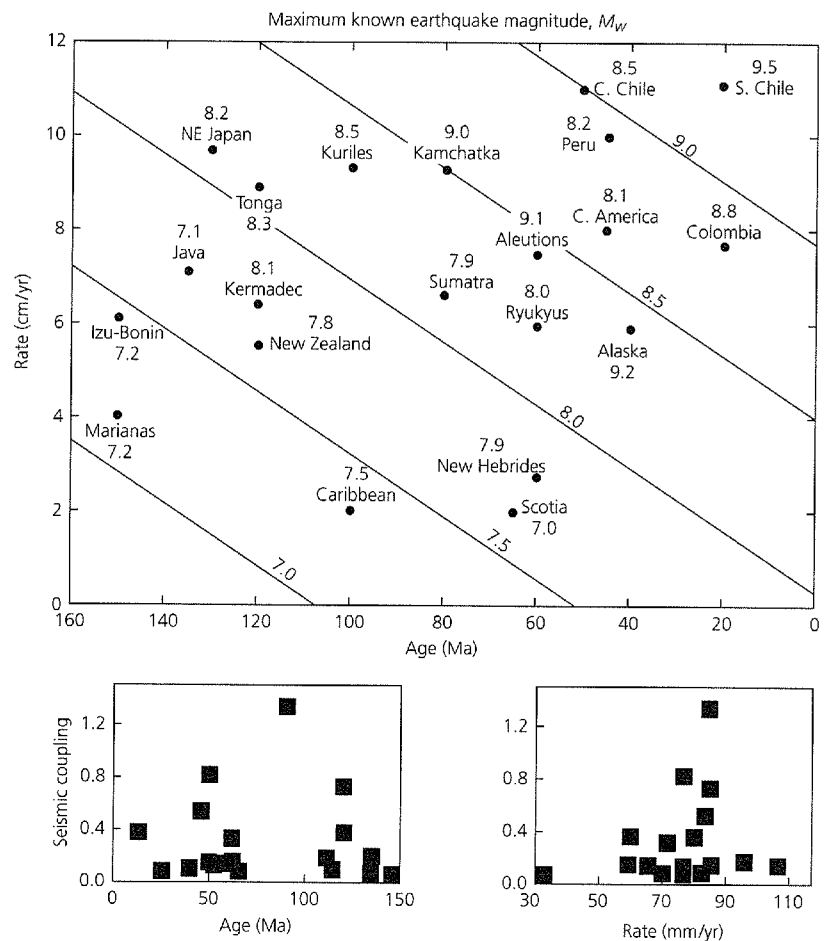
efforts to correlate the seismic slip fraction with subduction zone properties such as convergence rate or plate age find no clear pattern (Fig. 5.4-30, bottom). It has also been suggested that seismic coupling may be lowest for sedimented trenches and where normal stress on the plate interface is low, although these plausible ideas have yet to be demonstrated. Thus, although seismic coupling can be defined from the seismic slip fraction, its relation to the mechanics of plate coupling is still unclear. It appears that most subduction zones have significant components of aseismic slip, as do oceanic transforms and many continental plate boundaries (Section 5.6.2). Hence, even given the considerable uncertainties in such estimates, it appears common for a significant fraction of plate motion to occur aseismically.

The difficulty in estimating seismic coupling and understanding the process of aseismic plate motion has consequences for estimating the recurrence of earthquakes on a plate boundary and the seismic gap concept. It may be difficult to distinguish between gaps and areas where much of the slip is aseismic. For example, we would not want to say both that areas with recent major seismicity have high seismic hazard and that areas with little recent seismicity are gaps with high seismic hazard.<sup>6</sup> Moreover, as discussed in Sections 1.2 and 4.7.3, the process of earthquake faulting may be sufficiently random that it is hard to use the plate motion rate and seismic history to usefully predict how long it will be until the next large earthquake.

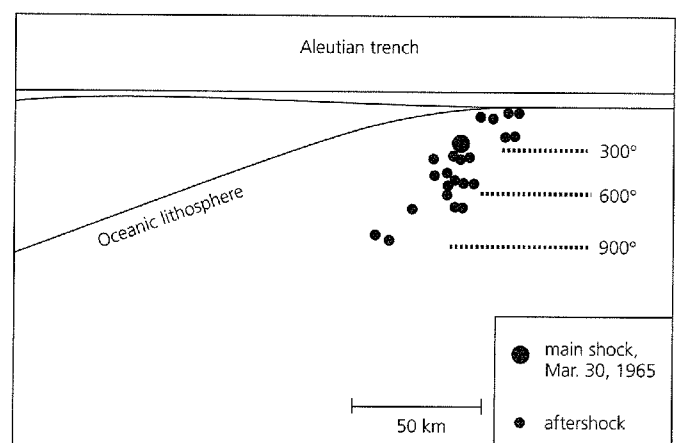
Although most shallow subduction zone seismicity is at the plate interface, some earthquakes occur within either plate. Some appear to result from flexural bending of the downgoing plate as it enters the trench (Fig. 5.4-31). Focal depth studies show a pattern of normal faulting in the upper part of the plate to a depth of 25 km, and thrusting in its lower part, between 40 and 50 km. These observations constrain the position of the neutral surface dividing the mechanically strong lithosphere (Section 5.7.4) into upper extensional and lower compressional zones. In some cases the normal fault earthquakes are so large that they may be "decoupling" events due to "slab pull" that rupture the entire downgoing plate (Fig. 5.4-32). After-shock distributions and studies of the rupture process indicate that faulting extended through a major portion, and perhaps all, of the lithosphere. Rupture through the entire lithosphere favors the decoupling model. If only a portion of the lithosphere breaks, the interpretation is more complicated. Rupture may have been restricted to one side of the neutral surface (in the flexural model) or reflect the material below being too hot and weak for seismic rupture. In the latter case, the entire lithosphere could have failed, with the deeper rupture being aseismic.

<sup>6</sup> The observation that more recent grizzly bear attacks have occurred in Montana than in Illinois might indicate either a perilous "gap" in Illinois or a greater intrinsic hazard in Montana.

**Fig. 5.4-30** *Top:* Variation in the magnitude ( $M_w$ ) of the largest known subduction thrust fault earthquake between subduction zones as a function of the convergence rate and age of the subducting lithosphere. (Ruff and Kanamori, 1980. *Phys. Earth Planet. Inter.*, 23, 240–52, with permission from Elsevier Science.) *Bottom:* Seismic coupling fraction estimated from historical seismicity at various subduction zones. Although most subduction zones show considerable aseismic slip, there is no obvious correlation with either age of the subducting lithosphere (*left*) or subduction rate (*right*). (Pacheco *et al.*, 1993. *J. Geophys. Res.*, 98, 14, 133–59, copyright by the American Geophysical Union.)



**Fig. 5.4-31** Focal depths of flexural earthquakes due to the bending of subducting plates as they enter the trench. Tensional events occur above the neutral surface, and compressional events occur below it. The plate mechanical thickness,  $H$ , increases with age, as expected from thermal models. (After Bodine *et al.*, 1981. *J. Geophys. Res.*, 86, 3695–707, copyright by the American Geophysical Union.)



**Fig. 5.4-32** Large normal faulting earthquakes at trenches, such as the 1965  $M_s$  7.5 Rat Island earthquake, may be due to flexure or failure of the lithosphere under its own weight. The extent of aftershocks, which appear not to cut the entire lithosphere, may reflect the extent of rupture or be a temperature effect. (Wiens and Stein, 1985. *Tectonophysics*, 116, 143–62, with permission from Elsevier Science.)

## 5.5 Oceanic intraplate earthquakes and tectonics

The vast majority of earthquakes — especially when measured in terms of seismic moment release — occur on plate boundaries and reflect the relative plate motions there. However, *intraplate* earthquakes, those within plates, also provide important tectonic information. We discuss intraplate earthquakes that occur in oceanic lithosphere in this section, and then discuss their counterparts in continental lithosphere in the next.

### 5.5.1 Locations of oceanic intraplate seismicity

Figure 5.5-1 illustrates the distribution of earthquakes in the Atlantic Ocean, excluding those along the Mid-Atlantic ridge. Although these earthquakes are rarer than those along the

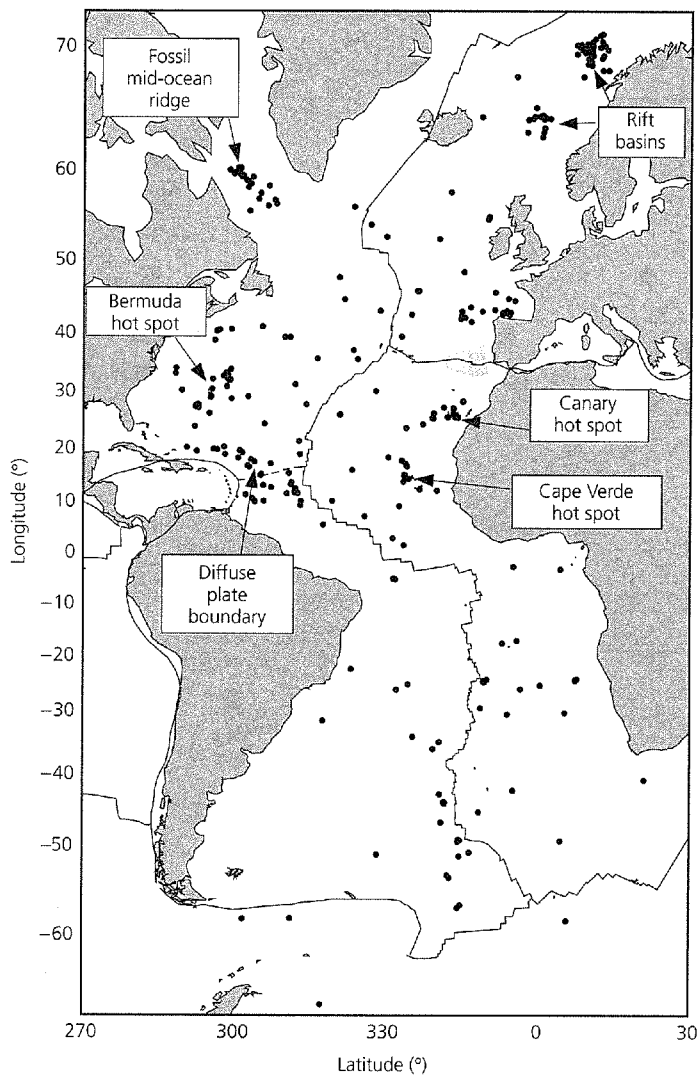


Fig. 5.5-1 Distribution of earthquakes in the Atlantic Ocean other than those on ridge and transform segments of the Mid-Atlantic ridge system. (Wyssession *et al.*, 1995. © Seismological Society of America. All rights reserved.)

ridges and transforms making up the Mid-Atlantic ridge plate boundaries, there are enough to justify interest. They nicely illustrate that plates deviate from the ideal case of perfect rigidity without internal deformation, such that all motion occurs at narrow boundaries. Instead, as noted in Section 5.2, real plates are complicated entities that have both internal deformation and diffuse boundary zones.

One way to think about these earthquakes is to consider a hierarchy, from slow-moving plate boundaries, to recognizable weak structures, and then to apparently isolated earthquakes. For example, the Atlantic portion of the boundary between the Eurasian and African plates, which stretches from Gibraltar to the Azores, is poorly defined by topography and seismicity compared to the Mid-Atlantic ridge. However, the focal mechanisms (Fig. 5.5-2, *top*) show a transition from extension at the Terceira Rift near the Azores, to strike-slip along a segment that includes the mapped Gloria transform fault, to compression near Gibraltar, and then into the Mediterranean. This transition reflects the fact that the Euler pole is close enough that the relative motions are small and change rapidly with distance (Fig. 5.5-2, *bottom*). For example, near the triple junction the NUVEL-1A model (Table 5.2-1) predicts 4 mm/yr of extension resulting from the small difference between Eurasia–North America (23 mm/yr at N97°E) and Africa–North America (20 mm/yr at N104°E) spreading across the Mid-Atlantic ridge. Even in the western Mediterranean, the motions are too slow to generate a well-developed subduction zone like those of the Pacific, but instead cause a broad convergent zone indicated by large earthquakes like the 1980  $M_s$  7.3 El Asnam, Algeria, earthquake.

Even slower motion appears to be why sea floor topography shows no clear evidence for the boundary between the North American and South America plates shown by the dashed line in Fig. 5.5-1, despite a diffuse zone of seismicity in this area. This zone is considered to be a plate boundary, based on detailed studies of plate motions. These studies invert plate motion data (spreading rates, transform fault directions, and earthquake slip vectors; Section 5.2.2) to find Euler vectors under two different assumptions: either there is a single American plate, or there are two. The Euler vectors derived by assuming there are two plates fit the data better, which would be expected, because a model with more parameters always fits data better. However, statistical tests (Section 7.5.2) show that the fit to the data improves more than expected purely by chance due to the additional parameters, implying that the two plates are distinct.

The North America–South America Euler vector that results from inverting the data is not well constrained, because it is not derived directly from data recording the motion between North America and South America, but is estimated from closure of the plate circuit (Fig. 5.2-5). Thus the estimate of motion results from the difference between North America–Africa and South America–Africa motions, which are quite similar (if they were not, the data would clearly show two distinct American plates). The predicted motion along the North

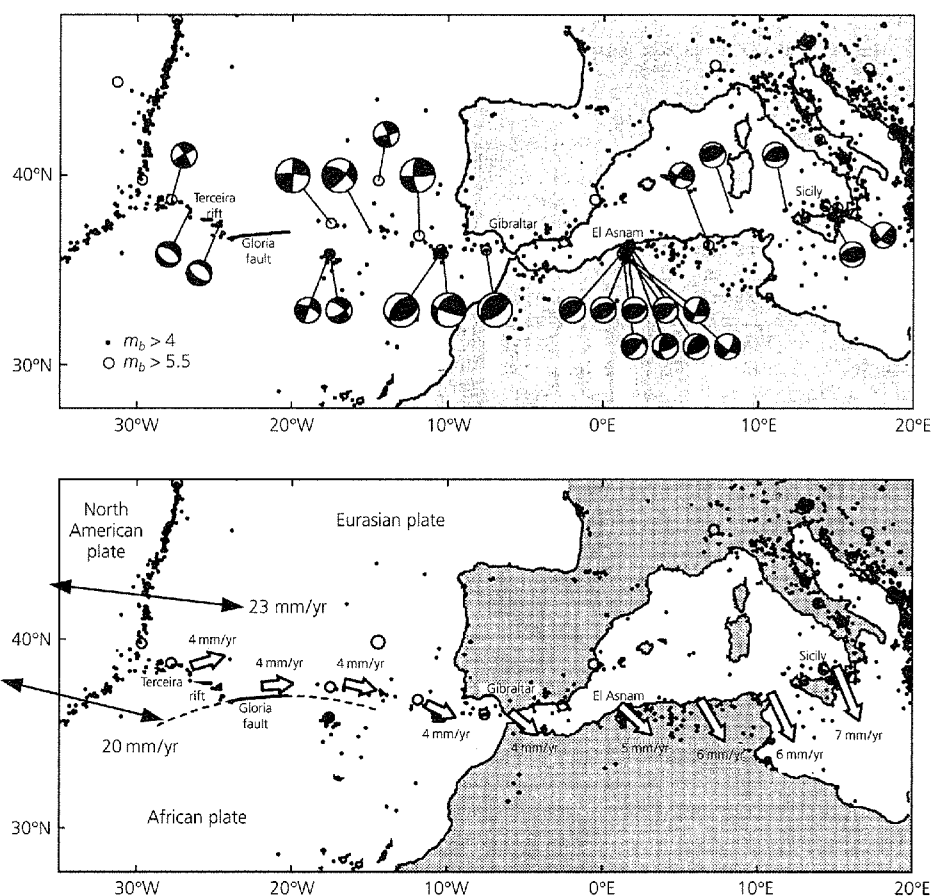


Fig. 5.5-2 *Top*: Focal mechanisms along the western section of the Eurasia–Africa plate boundary. Note the transition from extension near the Azores, to strike-slip (the Gloria fault is a transform), to compression near Gibraltar and into the Mediterranean. *Bottom*: Motions with respect to Africa along the boundary predicted by an Euler pole slightly south of the mapped area, near 20°N, 20°W. The dashed line is a small circle about this pole. (Argus *et al.*, 1989. *J. Geophys. Res.*, 94, 5585–5602, copyright by the American Geophysical Union.)

America–South America boundary is only about 1 mm/yr — much slower than the approximately 20 mm/yr along the Mid-Atlantic ridge. The North America–South America boundary is thus considered a diffuse, slow-moving boundary zone, although its location and motion are not well constrained. Another reason for treating this as a boundary zone is that paleomagnetic reconstructions find that over the past 70 Myr the two plates have moved relative to each other as the Atlantic Ocean opened.

In general, 1–2 mm/yr is an approximate lower limit for plate boundary deformation. Regions with motions faster than this are generally viewed as plate boundaries, and slower deformation is generally treated as intraplate. However, there is no generally accepted criterion, and evidence from seismicity and topography is also considered. Put another way, in many cases one can regard a region as either a slow-moving plate boundary zone or a zone of intraplate deformation, and “intraplate” earthquakes are often just ones not on an obvious plate boundary.

The Atlantic example (Fig. 5.2-1) shows that in addition to the North America–South America boundary zone, some intraplate seismicity is concentrated in other areas associated with tectonic features. For example, seismicity between Green-

land and North America is likely related to the former spreading ridge that opened this part of the Atlantic (the Labrador Sea). Although this spreading stopped about 43 Myr ago, the fossil ridge appears to remain a weak zone along which intraplate stresses cause some motion. Intraplate seismicity is often associated with such fossil structures. Concentrations of seismicity are also associated with the Bermuda (32°N, 65°W), Cape Verde (17°N, 25°W), and Canary (26°N, 17°W) hot spots. Focal mechanism studies are consistent with the earthquakes reflecting heating of the lithosphere by the hot spots.

Hawaii, the most impressive hot spot trace in the oceans (Fig. 5.2-7),<sup>1</sup> provides the best example of intraplate earthquakes associated with hot spot processes (Fig. 5.5-3). Small earthquakes are associated with magma upwelling in the rift zones. Larger earthquakes, which occur on a time scale of tens of years, reflect sliding of the volcanic edifice on subhorizontal faults that are thought to be a layer of weak sediments at the top of the old oceanic crust on which the volcanic island formed. These earthquakes can be quite large — the 1975

<sup>1</sup> Numerical models that infer the amount of upwelling mantle material from how elevated the sea floor is relative to the normal depth–age curves estimate that Hawaii has a buoyancy flux 5–10 times greater than that of Bermuda (Sleep, 1990).

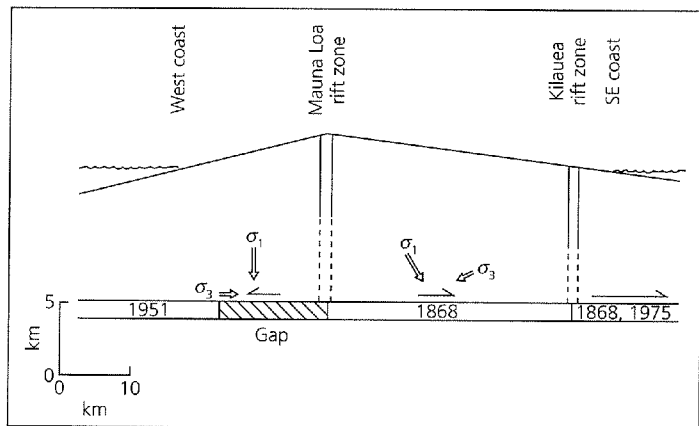


Fig. 5.5-3 Schematic model for large intraplate earthquakes below the island of Hawaii. Small earthquakes are associated with magma upwelling in the rift zones. Larger earthquakes, at dates shown, reflect sliding of the volcanic edifice on subhorizontal faults. The portion of the basal fault that has not ruptured in historic time may be a seismic gap. (Wyss and Koyanagi, 1992. © Seismological Society of America. All rights reserved.)

Kalapana earthquake had  $M_s$  7.2, caused a tsunami that killed two campers on the seashore, and did considerable property damage. The earthquake was followed by a small volcanic eruption near the summit of Kilauea, perhaps because the ground shaking triggered an eruption of shallow magma. Curiously, some earthquakes occur to considerable depths under Hawaii, including a magnitude 6.2 earthquake at 48 km depth.

Although many oceanic intraplate earthquakes are associated with tectonic features, some appear to occur far from plate boundaries, hot spots, or major bathymetric features. Thus the stresses generated by plate driving forces and other sources, including mantle flow near hot spots, appear to reactivate weak zones in the plate resulting from small-scale structure acquired during the lithosphere's evolution.<sup>2</sup>

These earthquakes can be dramatic. For example, the enormous ( $M_w$  8.2) intraplate earthquake that occurred near the Balleny Islands in an oceanic part of the Antarctic plate (63°S, 149°E) in March 1998 was the largest earthquake that had occurred on earth for several years. The fault inferred from waveform modeling (Section 4.3) followed no observable lineaments and cut straight across existing fracture zones. Moreover, in the previous hundred years, no other earthquakes had been located in this region. It is not clear what caused the earthquake or whether this area has any special properties or stress acting there. Although the earthquake occurred south of a puzzling hypothesized deformation zone in the extreme southeast corner of the Australian plate (Fig. 5.2-4), its fault plane solution is inconsistent with its being on the boundaries of a microplate. It is thus unclear whether this area is now any

more prone to future earthquakes than other areas, and what the recurrence time of such earthquakes might be. Similar issues arise in considering the intraplate seismicity and associated seismic hazard in the more structurally complex continents.

Oceanic intraplate seismicity often occurs in swarms. Regions without previously known seismicity sometimes become active for several years, with hundreds of teleseismically located earthquakes.<sup>3</sup> The seismicity then dies out, and seems not to recur. For example, during 1981–3, an intraplate earthquake swarm occurred near the Gilbert Islands in Micronesia. A total of 225 earthquakes were detected, mostly over a 15 month period, with 87 above  $m_b$  5. No major tectonic features are known in this area, and a ship survey found no bathymetric anomalies. Before and after the swarm, no other earthquakes have been recorded in this region. The swarms thus differ from plate boundary seismicity, which occurs on features that remain active for long periods even if there are intervening quiet intervals. Moreover, the intraplate swarms often appear not to have a single well-developed fault, and no event is significantly bigger than the others. By contrast, plate boundary earthquakes usually have one or two main ruptures and many aftershocks, perhaps reflecting local adjustments to the stress field after the mainshock has ruptured the entire fault.

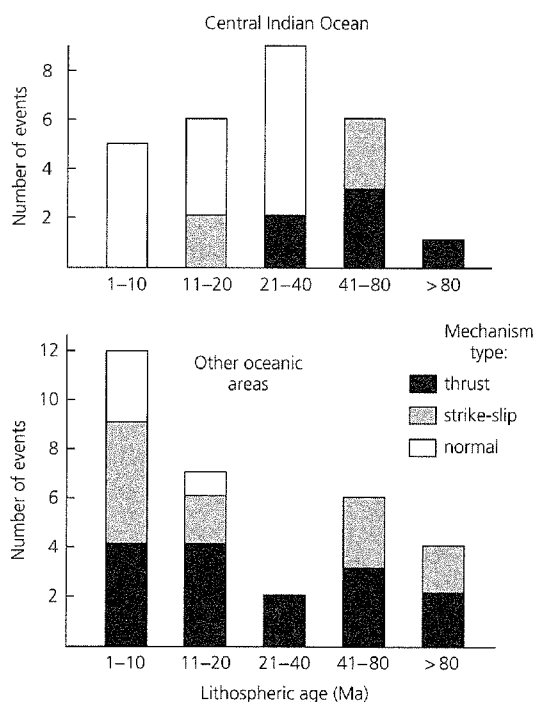
These swarms raise an interesting issue. We can assume that these areas are analogous to plate boundaries in having special, if not yet understood, tectonic significance. If so, they are likely to be the sites of future swarms. Alternatively, perhaps all areas of oceanic lithosphere are equally susceptible to such swarms. In this case, over time, swarms will occur in many places, and future swarms are no more likely in one place than another. We will see that similar issues surface in trying to estimate seismic hazards due to intraplate earthquakes within continents.

### 5.5.2 Forces and stresses in the oceanic lithosphere

In addition to using oceanic intraplate seismicity to investigate the specific processes acting at individual sites, we study the seismicity to learn about plate-wide processes. For example, Fig. 5.5-4 shows the variation of mechanism type with lithospheric age. Most of the oceanic lithosphere seems to be in horizontal deviatoric compression, as shown by thrust and strike-slip mechanisms. This compression is in approximately the spreading direction, and is thought to be related to “ridge push”: the plate driving force due to lithospheric cooling and subsidence. The major exceptions are the extensional events occurring in the central Indian Ocean. Although originally regarded as intraplate, these earthquakes now appear to be in a diffuse plate boundary zone (Section 5.2.2). In the model shown, the focal mechanisms (Fig. 5.5-5) reflect counterclockwise rotation of Australia with respect to India, causing normal fault earthquakes in the young lithosphere near the Euler pole

<sup>2</sup> This situation is analogous to timbers creaking as a wooden boat rocks in the waves.

<sup>3</sup> There may be many more smaller earthquakes associated with these swarms, but because the swarms often occur in remote regions, only the larger events are detected.



**Fig. 5.5-4** Focal mechanism type as a function of lithospheric age for oceanic intraplate earthquakes. Older oceanic lithosphere is in compression, whereas younger lithosphere has both extensional and compressional mechanisms. Extensional events are located primarily in the central Indian Ocean. (Wiens and Stein, 1984. *J. Geophys. Res.*, 89, 11, 442–64, copyright by the American Geophysical Union.)

and thrust and strike-slip earthquakes to the east. These earthquakes reach magnitude 7 on the Ninetyeast ridge.<sup>4</sup>

The general trend of compressive mechanisms in the oceanic plates is consistent with the plate driving force due to the cooling of the oceanic lithosphere. Consider a plate, defined as the area above the  $m(t)$  isotherm, out to age  $t$ , where the water depth is  $h(t)$  (Fig. 5.5-6). The plate is cooler, and thus denser, than material below. The thermal model we used for ocean depth and heat flow also predicts the resulting force.

The total horizontal force on the base of the lithosphere,  $F_1$ , equals the integrated horizontal pressure force of the asthenosphere at the ridge, because the material is in hydrostatic equilibrium:

$$F_1 = \int_0^{m(t)} \rho_m g z dz = \rho_m g (m(t))^2 / 2. \quad (1)$$

Similarly,  $F_2$ , the horizontal force due to water pressure on the plate, equals the integrated horizontal pressure force of the water,

<sup>4</sup> Although hot spot tracks like the Ninetyeast and Chagos-Laccadive ridges have been termed “aseismic” ridges, to distinguish them from spreading ridges, these two are more seismically active in terms of moment release than many spreading ridges.

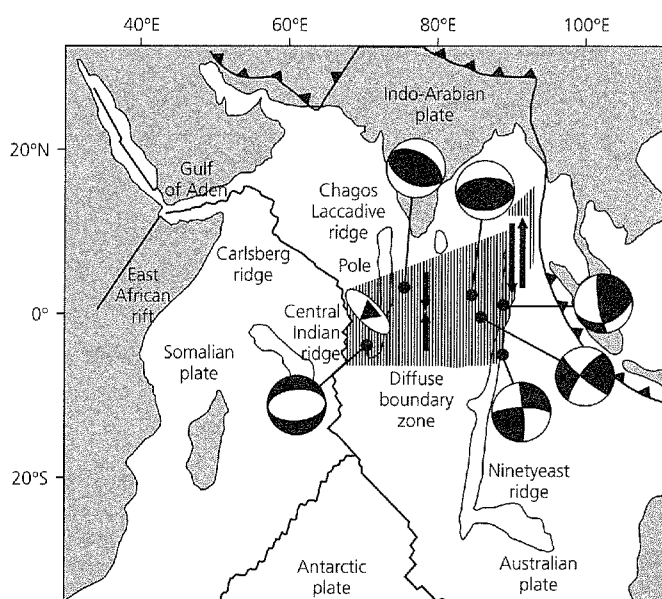


Fig. 5.5-5 Schematic map of earthquake mechanisms in the central Indian Ocean, shown here as a diffuse boundary zone (shaded) between the Indian and Australian plates. Later studies have refined the location and geometry of the boundary zone (Fig. 5.2-4) and pole (triangle) (Wiens *et al.*, 1985. *Geophys. Res. Lett.*, 12, 429–32, copyright by the American Geophysical Union.)

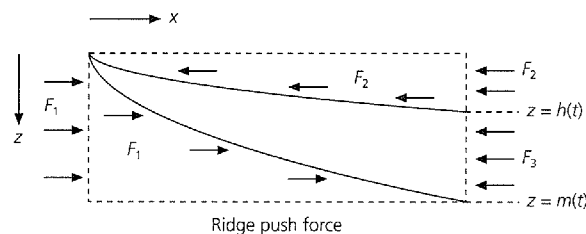


Fig. 5.5-6 Derivation of the "ridge push" force.

$$F_2 = \int_0^{b(t)} \rho_w g z dz = \rho_w g (h(t))^2 / 2. \quad (2)$$

$F_3$  is the remaining horizontal force due to lithospheric pressure  $P(z, t)$ ,

$$F_3 = \int_{ht(t)}^{m(t)} P(z, t) g z d z, \quad (3)$$

where the pressure depends on the density perturbation due to lithospheric cooling (Eqn 5.3.7),

$$P(z, t) = \rho_w g b(t) + g \int_{b(t)}^z [\rho_m + \rho'(z', t)] dz'. \quad (4)$$



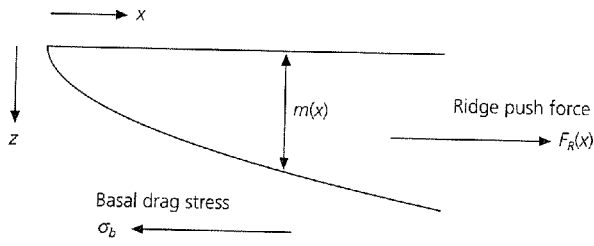


Fig. 5.5-7 Geometry for a simple model of intraplate stresses.

If the plate is not accelerating, the force difference is balanced by a net horizontal force

$$F_R = F_1 - F_2 - F_3. \quad (5)$$

For the cooling halfspace temperature structure (Eqn 5.3.2), this force is

$$F_R = g\alpha\rho_m T_m \kappa t, \quad (6)$$

whereas for a plate model it approaches a constant value for old lithosphere. The convention of calling this force “ridge push” is confusing because it is zero at the ridge and increases linearly with plate age. It results not from force at the ridge but from the total force due to the density anomaly within the cooling plate out to any given age.

The expression for the “ridge push” force is similar to that for the “slab pull” force (Eqn 5.4.15) because both are thermal buoyancy forces due to the density contrast resulting from the temperature difference between the plate and its surroundings. The two depend in the same way on the  $g\alpha\rho_m T_m$  term that describes the force due to the density contrast, but differently on  $\kappa$  because faster cooling increases ridge push whereas faster heating decreases slab pull. Although it is useful to think of the forces separately, both are net buoyancy forces due to the mantle convection system of which the plates are a part.<sup>5</sup>

To discuss the stresses within the oceanic lithosphere, we compare the ridge push force to the other forces applied at the boundaries of the plate. These include forces at the plate base and forces at the subduction zone. As for the downgoing slab, earthquake focal mechanisms constrain the relative size of the forces. Here, we use the observation (Fig. 5.5-4) that stress in the spreading direction is typically compressive at all ages.

Consider a simple model of stress in the oceanic lithosphere, using the geometry of Fig. 5.5-7. Using the stress equilibrium equation (Eqn 2.3.49) in the spreading ( $x$ ) direction, we relate the deviatoric stresses to the body force  $f(x, z)$ , which is the contribution to ridge push from the material at  $(x, z)$ ,

$$\frac{\partial \sigma_{xx}(x, z)}{\partial x} + \frac{\partial \sigma_{xz}(x, z)}{\partial z} + f(x, z) = 0. \quad (7)$$

<sup>5</sup> Verhoogen (1980) offers the analogy that rain occurs because of the negative buoyancy of the drops relative to the surrounding air, as part of the process by which solar heat evaporates water which rises as vapor due to positive buoyancy and is transported by wind to the point where it cools, condenses into drops, and then falls.

Integrating first with respect to  $x$  and then with respect to  $z$  from  $z = 0$  to the base of the lithosphere  $m(x)$  yields the force balance

$$\bar{\sigma}_{xx}(x) = \frac{\sigma_b x - F_R(x)}{m(x)} + \sigma_r. \quad (8)$$

Here the stress in the spreading direction is given by its vertical average  $\bar{\sigma}_{xx}(x)$ ;  $\sigma_r = \bar{\sigma}_{xx}(0)$  characterizes the strength of the ridge; the drag force at the base of the plate is given by the basal shear stress  $\sigma_b$ ; and  $F_R(x)$  is the net ridge push force

$$F_R(x) = \int_0^{m(x)} \int_0^x f(x, z) dx dz. \quad (9)$$

Written in terms of plate age,  $t$ ,

$$\bar{\sigma}_{xx}(t) = \frac{\sigma_b vt - F_R(t)}{m(t)} + \sigma_r, \quad (10)$$

where  $v$  is a half spreading rate, assumed constant. A useful form for comparing different plates comes from the usual assumption that the basal drag force equals the product of absolute velocity  $u$  and drag coefficient  $C$  ( $\sigma_b = Cu$ ),

$$\bar{\sigma}_{xx}(t) = \frac{Cuvt - F_R(t)}{m(t)} + \sigma_r \quad (11)$$

Thus a drag depending on absolute velocity is applied over an area proportional to the spreading rate. For simplicity, we assume that  $v = u$ , spreading rate equals absolute velocity (the ridge is fixed with respect to the mantle), so the net drag force is proportional to velocity squared.

A subduction zone would provide a boundary condition on the oldest lithosphere. For example, if focal mechanisms in the lithosphere near trenches were extensional, an extensional condition could be imposed. Because such mechanisms are not seen, it is often assumed that the negative buoyancy of slabs (slab pull) is balanced by local resistive forces (Section 5.4.2). Thus, although the ridge push force is probably smaller than the slab pull forces, the thrust fault mechanisms suggest that it is more crucial for determining stress in oceanic lithosphere.

Although this stress model is schematic and does not describe any individual plate, it lets us use focal mechanism observations to estimate several important quantities. Figure 5.5-8 shows the predicted intraplate stress as a function of plate age and drag coefficient. For zero drag the stress is purely compressive ( $\bar{\sigma}_{xx} < 0$ ) and varies as  $\sqrt{t}$ , because the force increases linearly with age, whereas the plate thickens as its square root. For larger drag coefficients,  $\bar{\sigma}_{xx}$  follows  $\sqrt{t}$  curves corresponding to less and less compression, until the lithosphere is in extension for all ages. All lithospheric plates

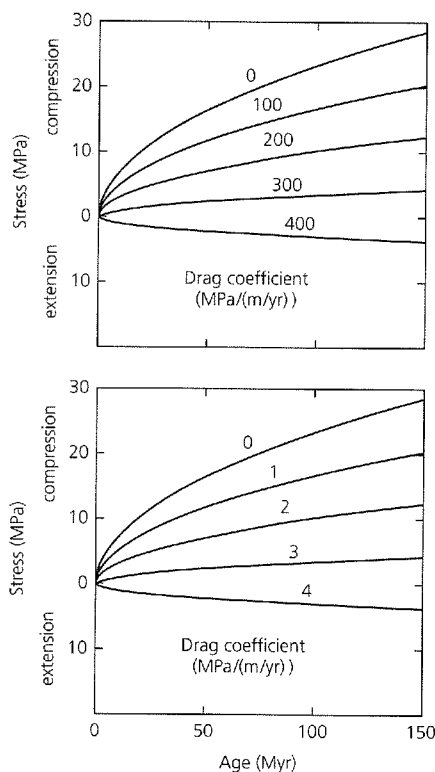


Fig. 5.5-8 Intraplate stress in the spreading direction as a function of lithospheric age and assumed basal drag coefficient for slow-moving (1 cm/yr, *top*) and fast-moving (10 cm/yr, *bottom*) plates. The compressional stresses in old oceanic lithosphere place an upper bound on the drag coefficient of 4 MPa/(m/yr). (Wiens and Stein, 1985, *Tectonophysics*, 116, 143–62, with permission from Elsevier Science.)

appear to be in compression, so a rapidly moving plate (such as the Pacific, which moves at about 10 cm/yr) constrains the drag coefficient to less than about 4 MPa/(m/yr). Similar results emerge for a cooling plate model.

This model assumes a zero stress boundary condition at the ridge axis, so the axis has no tensile strength. The predicted stress in young lithosphere, especially the location of a possible transition from compression to extension in the direction of spreading, would be sensitive to the strength of the ridge (Fig. 5.5-9). Models with substantial strength at the axis predict a wide band of extension in the spreading direction. Since such a zone of normal-faulting earthquakes is not observed, the axis seems weak.

Although this simple model describes only a hypothetical average plate, more sophisticated models use realistic plate geometries to calculate the stresses expected from ridge push, slab pull, and basal drag forces. These models' predictions can be compared to earthquake focal mechanisms and other data for specific areas. For example, Fig. 5.5-10 shows stresses predicted for the Indian Ocean region. Although the model was calculated assuming a single Indo-Australian plate, it predicts stresses in the region now considered a diffuse boundary

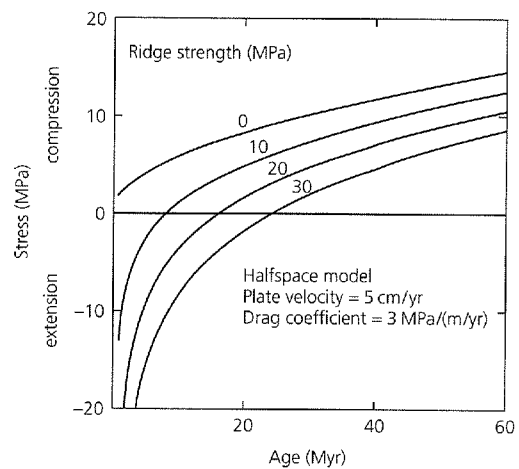


Fig. 5.5-9 Intraplate stress in the spreading direction as a function of lithospheric age computed for several values of ridge strength. The age of the transition from ridge-normal extension to compression increases with the strength of the ridge. (Wiens and Stein, 1984, *J. Geophys. Res.*, 89, 11, 442–64, copyright by the American Geophysical Union.)

zone (Fig. 5.5-5) that are generally consistent with the focal mechanisms and the folding seen in gravity and seismic reflection data.

### 5.5.3 Constraints on mantle viscosity

The last section's analysis relating earthquake mechanisms to drag at the base of the lithosphere also gives insight into the viscosity of the mantle. The viscosity,<sup>6</sup> the proportionality constant between shear stress and the strain rate (or velocity gradient), controls how the mantle flows in response to applied stress, and is thus crucial for mantle convection. If the drag on the base of a plate is due to motion over the viscous mantle, compressive earthquake mechanisms in old lithosphere constrain the viscosity.

Consider a simple two-dimensional geometry where mass flux due to the moving plate is balanced by a return flow at depth (Fig. 5.5-11, *top*). The drag coefficient is proportional to the viscosity and inversely proportional to the flow depth. Figure 5.5-12 shows that the basal drag constraint from the focal mechanism data,  $C \leq 4$  MPa/(m/yr), requires an average mantle viscosity less than  $2 \times 10^{20}$  poise if flow occurs to a depth of 700 km in the upper mantle, or  $10^{21}$  poise if flow occurs in the entire mantle. These values are lower than the  $1\text{--}5 \times 10^{22}$  poise typically estimated from glacial rebound, earth rotation, and satellite orbits.

This discrepancy can be reconciled by assuming that the plate is underlain by a thin, low-viscosity asthenosphere (Fig. 5.5-11, *bottom*). The low-viscosity layer, in which only a fraction of the return flow occurs, decouples the plates from the underlying

<sup>6</sup> Viscosity, defined in Section 5.7, is given in cgs units as poise (dyn-s/cm<sup>2</sup>) or in SI units as Pascal-seconds (1 poise = 0.1 Pa-s).

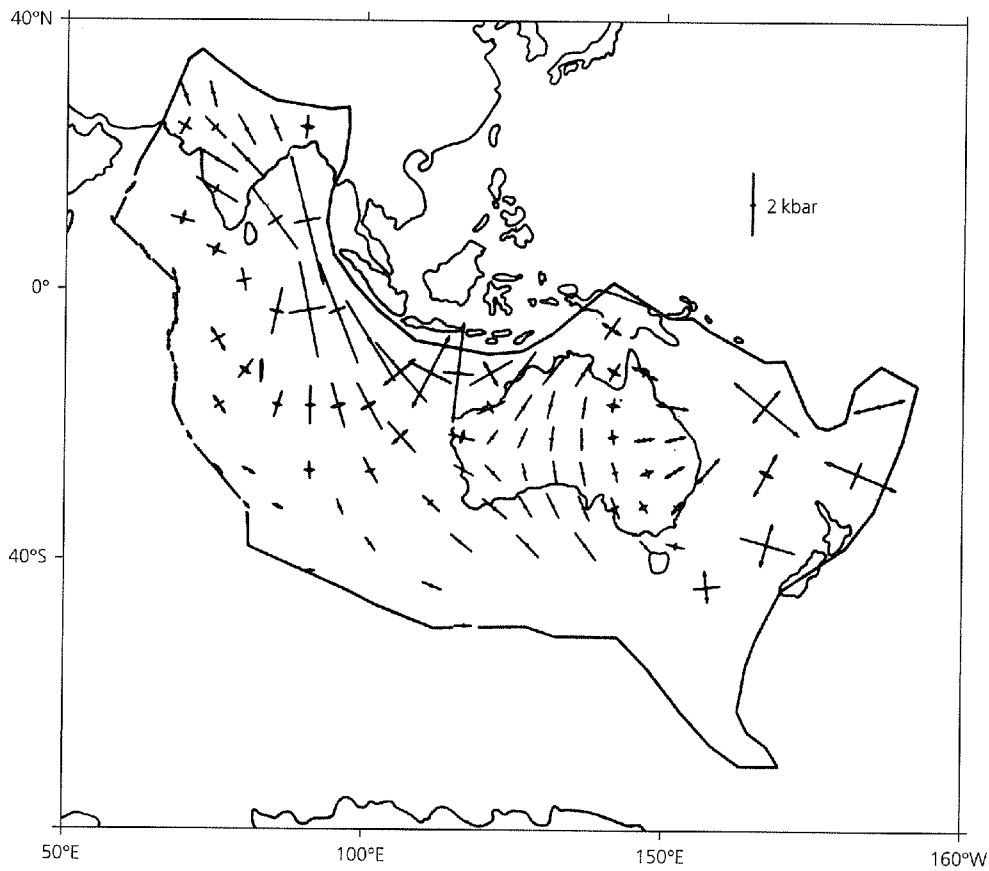


Fig. 5.5-10 Intraplate stress predicted by a force model for the Indo-Australian plate. The bars show the principal horizontal deviatoric stresses, with arrowheads marking tension. The location and orientation of the highest stresses, such as the transition between compression and tension, are generally consistent with earthquake mechanisms in the region now regarded as a diffuse plate boundary (Fig. 5.5-5). (Cloetingh and Wortel, 1985. *Geophys. Res. Lett.*, 12, 77-80, copyright by the American Geophysical Union.)

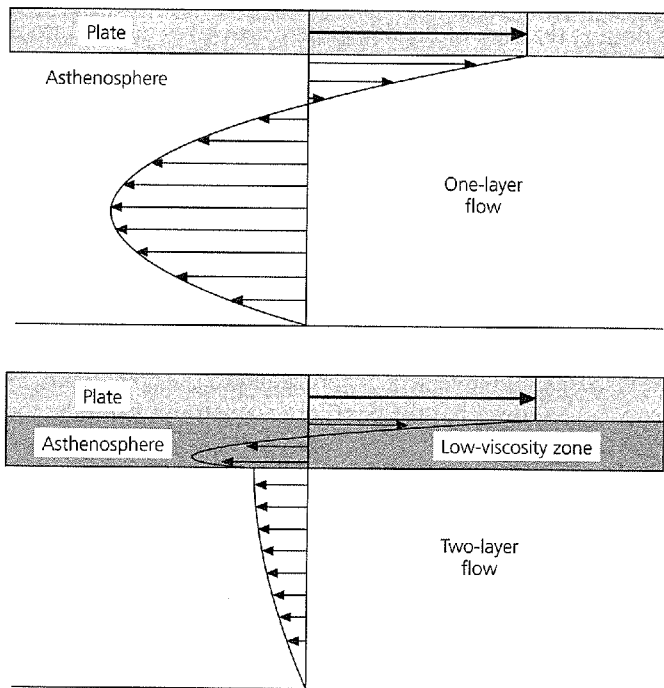


Fig. 5.5-11 *Top*: Velocity profile associated with a return flow of uniform-viscosity asthenosphere that balances the mass flux due to plate motions. *Bottom*: Velocity profile associated with a return flow of two layers of different viscosity. The upper, low-viscosity layer decouples the plates from the underlying mantle. (McKenzie and Richter, 1978.)

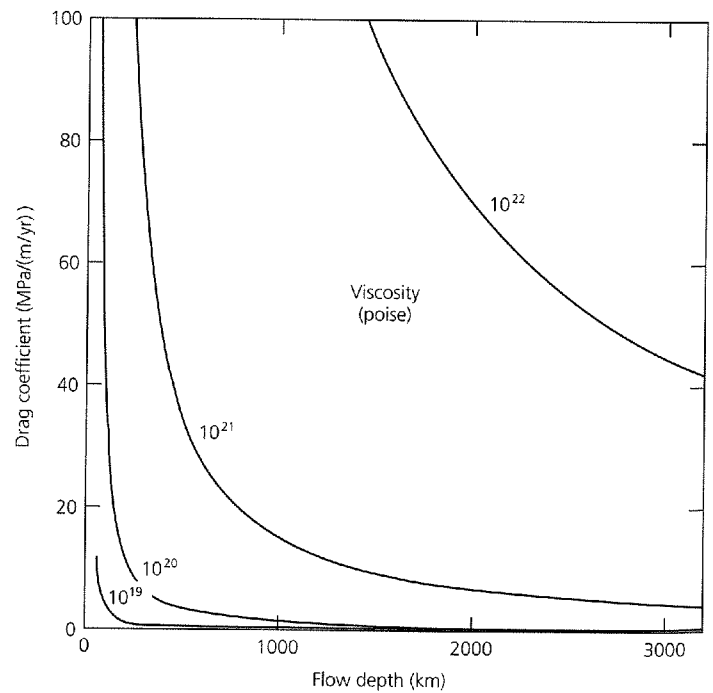
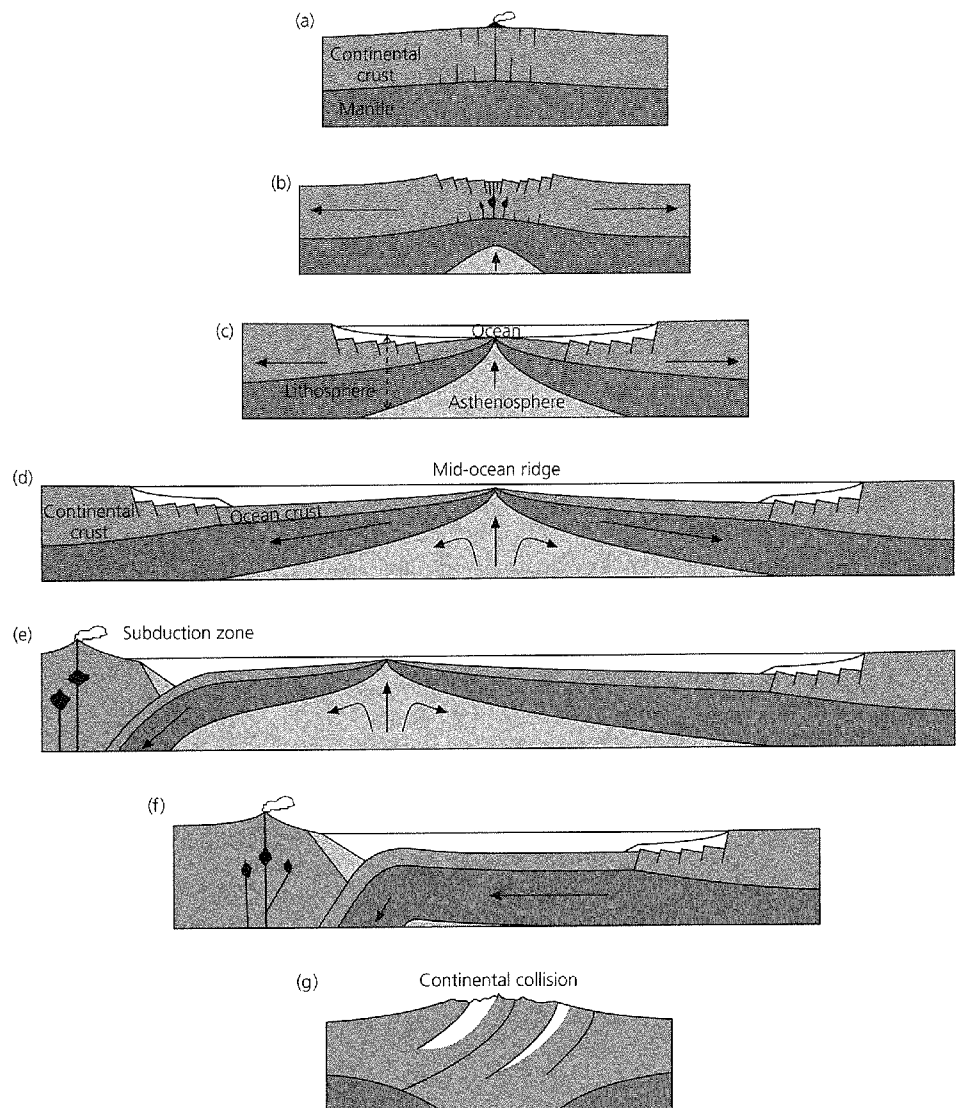


Fig. 5.5-12 Basal drag coefficients as a function of the mantle viscosity and flow depth, assuming single-layer flow. (Wiens and Stein, 1985. *Tectonophysics*, 116, 143-62, with permission from Elsevier Science.)



**Fig. 5.6-1** Schematic illustration of the Wilson cycle, the fundamental geological process controlling the evolution of the continents. (a)–(b): A continent rifts, such that the crust stretches, faults, and subsides. (c): Sea floor spreading begins, forming a new ocean basin. (d): The ocean widens and is flanked by sedimented passive margins. (e): Subduction of oceanic lithosphere begins on one of the passive margins, closing the ocean basin (f) and starting continental mountain building. (g): The ocean basin is destroyed by a continental collision, which completes the mountain building process. At some later time, continental rifting begins again.

mantle. Viscosity values that satisfy the focal mechanisms are consistent with constraints from gravity and glacial isostasy, and such decoupling is consistent with the lack of correlation between oceanic plate area and absolute velocity (Fig. 5.4-12).

## 5.6 Continental earthquakes and tectonics

Although the basic relationships between plate boundaries, plate interiors, and earthquakes apply to continental as well as oceanic lithosphere, the continents are more complicated. The continental crust is much thicker, less dense, and has different mechanical properties from the oceanic crust. As a result, plate boundaries in continental lithosphere are generally broader and more complicated than in the oceanic lithosphere (Fig. 5.2-4).

Studies of continental plate boundaries, which rely heavily on seismology, provide important insights into the funda-

mental geological processes controlling the evolution of the continents. The basic process, known as the *Wilson cycle*,<sup>1</sup> is illustrated in Fig. 5.6-1. A continental region undergoes extension, such that the crust is stretched, faulted, and subsides, yielding a rift valley like the present East African rift. Because the uppermost mantle participates in the stretching, hotter mantle material upwells, causing partial melting and basaltic volcanism. Sometimes the extension stops after only a few tens of kilometers, leaving a failed or fossil rift such as the 1.2 billion-year-old mid-continent rift in the central USA. In other cases the extension continues, so the continental rift evolves into an oceanic spreading center (identifiable from sea floor magnetic anomalies), which forms a new ocean basin like

<sup>1</sup> Named after J. Tuzo Wilson (1908–93), whose key role in developing plate tectonic theory included introducing the ideas of transform faults, hot spots, and that the Atlantic had closed and then reopened.

the Gulf of Aden or the Red Sea. With time, the ocean widens and deepens due to thermal subsidence of oceanic lithosphere (Section 5.3.2), and thick sediments accumulate on the continental margins, such as those on either side of the Atlantic. These margins are not plate boundaries — the oceanic and continental crust on the two sides are on the same plate — and are called *passive margins*, to distinguish them from active continental margins, which are plate boundaries. Subduction often begins along one of the passive margins, and the ocean basin closes, such that magmatism and mountain building occur, as along the west coast of South America today. Continental collision like that currently in the Himalayas occurs eventually, and the mountain building process reaches its climax. If the continental materials on either side cease to move relative to each other, this process leaves a mountain belt within the interior of a single plate. At some future time, however, a new rifting phase can begin, often near the site of the earlier rifting, and a new ocean will start to grow. Thus the Appalachian Mountains record a continental collision that closed an earlier Atlantic Ocean about 270 million years ago, and remain despite the opening of the present Atlantic Ocean during the past 200 Myr.

As a result, continental and oceanic crust have very different life cycles. Because the relatively less dense continental crust is not subducted, the continents have accreted over a much longer time than the 200-million-year age of the oldest oceanic crust. Hence the continents preserve a complex set of geologic structures, many of which can be sites of deformation, including earthquake faulting. Thus both plate boundary and intraplate deformation zones within continents are more complex than their oceanic counterparts.

Earth scientists seek to understand the continental evolution process for both intellectual and practical reasons. The process is fundamental to how the planet works, but also provides information about geologic hazards (earthquakes, volcanism, uplift, and erosion) and mineral resources. In addition, the large mountain belts have major impacts on earth's climate. Seismology contributes to these studies by providing data about earthquakes and velocity structure in regions where different parts of the evolutionary cycle occur today or occurred in the past. These data are combined with other geophysical and geological data to form an integrated picture of the complicated continental evolution processes. Hence, although the processes are not fully understood, important progress continues to be made.

### 5.6.1 Continental plate boundary zones

As for oceanic boundaries, we seek to first describe the motion (kinematics) within boundary zones, and then to combine the kinematics with other data to investigate their mechanics (dynamics). One example is the East African rift (Fig. 5.6-2), a spreading center between the Nubian (West Africa) and Somalian (East Africa) plates. The extension rate is so slow, less than 10 mm/yr, that it is hard to resolve in plate motion models, and the two plates are often treated as one (Fig. 5.2-4).

However, the rift topography, normal faulting, and seismicity distribution show the presence of an extensional boundary zone broader, more diffuse, and more complex than at a mid-ocean ridge. For example, the seismicity ends in southern Africa and has no clear connection to the southwest Indian ridge, where the plate boundary must go. A recent estimate is that the northern East Africa rift opens at about 6 mm/yr, whereas the southern part opens at about half that, because the Euler pole is to the south. Some of the complexity of such continental extensional zones results from the fact that, unlike a mid-ocean ridge, the lithosphere starts off with reasonable thickness and then is stretched and thinned in the extending zone. The rifting process can eventually progress far enough that a new oceanic spreading center forms. This has already occurred in the Gulf of Aden and the Red Sea, which are newly formed (and hence narrow) oceans separating the Arabian plate from Somalia and Nubia at rates of about 22 and 16 mm/yr, respectively. Whether the East African rift will evolve this far is still unclear, because the geologic record shows many rifts that, although active for some time, failed to develop into oceanic spreading centers and simply died. As we will see, these fossil rifts can be loci for intraplate earthquakes.

The earthquakes also indicate that the thermal and mechanical structure of continental rifts is more complicated than on mid-ocean ridges. Normal-faulting earthquakes extend to depths of 25–30 km, considerably deeper than at mid-ocean ridges. Hence the lower crust appears to be surprisingly stronger and colder than might be expected in an active rift.

Continental transforms are also more complicated than their oceanic counterparts. As we saw in Section 5.2, the transform portion of the Pacific–North America plate boundary in western North America is an active seismic zone hundreds of kilometers wide (Fig. 5.2-3), in contrast to widths of less than 10 km for oceanic transforms. Thus the focal mechanisms show primarily strike-slip motion on the San Andreas fault itself and demonstrate complexities including thrust faulting for events like the 1971 San Fernando and 1994 Northridge earthquakes and normal faulting due to the regional extension in the Basin and Range province. The earthquakes and space-geodetic data show that although most of the motion occurs along the San Andreas (Fig. 4.5-13) and nearby faults, a reasonable fraction of the motion occurs elsewhere (Figs. 5.6-3 and 5.2-3). The boundary zone is further complicated by volcanism in areas including the Long Valley caldera in eastern California and the Yellowstone hot spot, which also have associated seismicity. Hence, we think of a boundary zone in which the overall steady motion between the plate interiors is distributed in both space and time (Fig. 5.6-4). Although much of the motion occurs in occasional large earthquakes or steady creep on the main boundary segment, some deformation occurs elsewhere in the zone.

The breadth of continental plate boundary zones has important implications for seismic hazards within them. Because ground shaking decays rapidly with distance (Fig. 1.2-5), nearby smaller earthquakes within a boundary zone, but not

Fig. 5.6-2 Seismicity and focal mechanisms ( $T$  axes shown by black arrows) for the East African rift system, with relative plate motions (white arrows) from Chu and Gordon (1998, 1999).

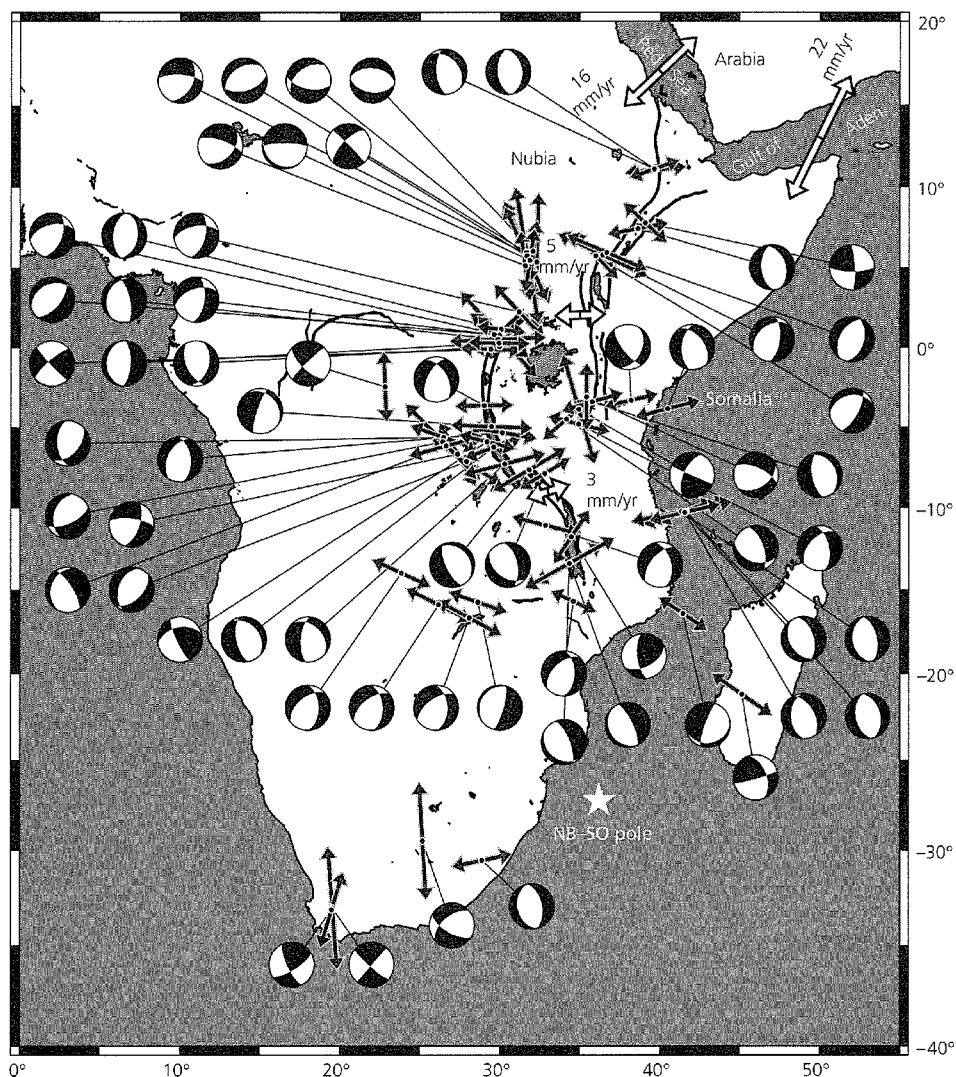
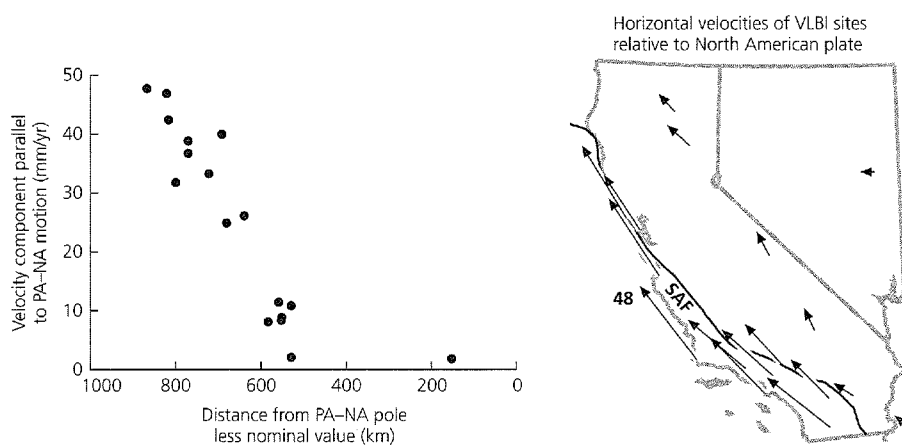


Fig. 5.6-3 Variation in motion of space-geodetic sites across part of the Pacific–North America boundary zone. *Right:* Horizontal velocities of sites in California, Nevada, and Arizona relative to stable North America. The velocity of the southwesternmost site nearly equals the predicted 48 mm/yr velocity of the Pacific plate relative to the North American plate. *Left:* Component of motion tangent to small circles centered on the Pacific–North America Euler pole versus angular distance from that pole. Velocities increase with distance from the Euler pole, with a discontinuity due to the approximately 35 mm/yr of time-averaged slip across the San Andreas fault. (Gordon and Stein, 1992. *Science*, 256, 333–42, copyright 1992 American Association for the Advancement of Science.)





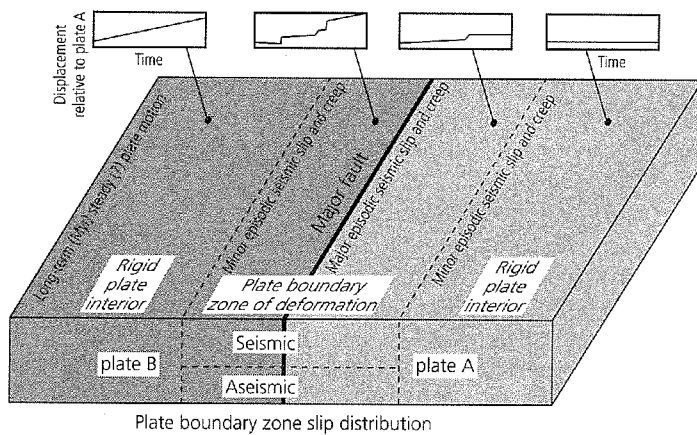


Fig. 5.6-4 Schematic illustration of the distribution of motion in space and time for a strike-slip boundary zone between two major plates. (Stein, 1993. *Contributions of Space Geodesy to Geodynamics*, 5–20, copyright by the American Geophysical Union.)

on the main boundary fault, can be more damaging than larger but more distant ones on the main fault. Hence the Los Angeles area is vulnerable to both nearby earthquakes like the 1994 Northridge ( $M_w$  6.7) or 1971 San Fernando ( $M_s$  6.6) earthquakes and larger ones on the more distant San Andreas Fault, such as a recurrence of the 1857 Fort Tejon earthquake which is estimated to have had  $M_w$  about 8. Similarly, the earthquake hazard in the Seattle area involves both great earthquakes at the subduction interface and smaller, but closer, earthquakes in the subducting Juan de Fuca plate (like the 2001  $M_w$  6.7

Nisqually earthquake) or at shallow depth in the North American plate.

Of the three boundary types, continental convergence zones may be the most complicated compared to their oceanic counterparts. One primary difference is that because continental crust is much less dense than the upper mantle, it is not subducted, and a Wadati–Benioff zone is not formed. As a result, continental convergence zones in general do not have intermediate and deep focus earthquakes. However, the plate boundary tectonics occur over a broader and more complex region than in the oceanic case.

A spectacular example is the collision between the Indian and Eurasian plates. This area is the present type example of mountain building by continental collision, which has produced a boundary zone extending thousands of km northward from the nominal plate boundary at the Himalayan front (Fig. 5.6-5). The total plate convergence is taken up in several ways. About half of the convergence occurs across the locked Himalayan frontal faults such as the Main Central Thrust (Fig. 5.6-6), and gives rise to large destructive earthquakes. These faults are part of the interface associated with the underthrusting Indian continental crust, which thickens the crust under the high Himalayas. However, the earthquakes also show normal faulting behind the convergent zone, in the Tibetan plateau, presumably because the uplifted and thickened crust spreads under its own weight. GPS data (Fig. 5.6-5) show that this extension is part of a large-scale process of crustal “escape,” or “extrusion,” in which large fragments of continental crust are displaced eastward by the collision along

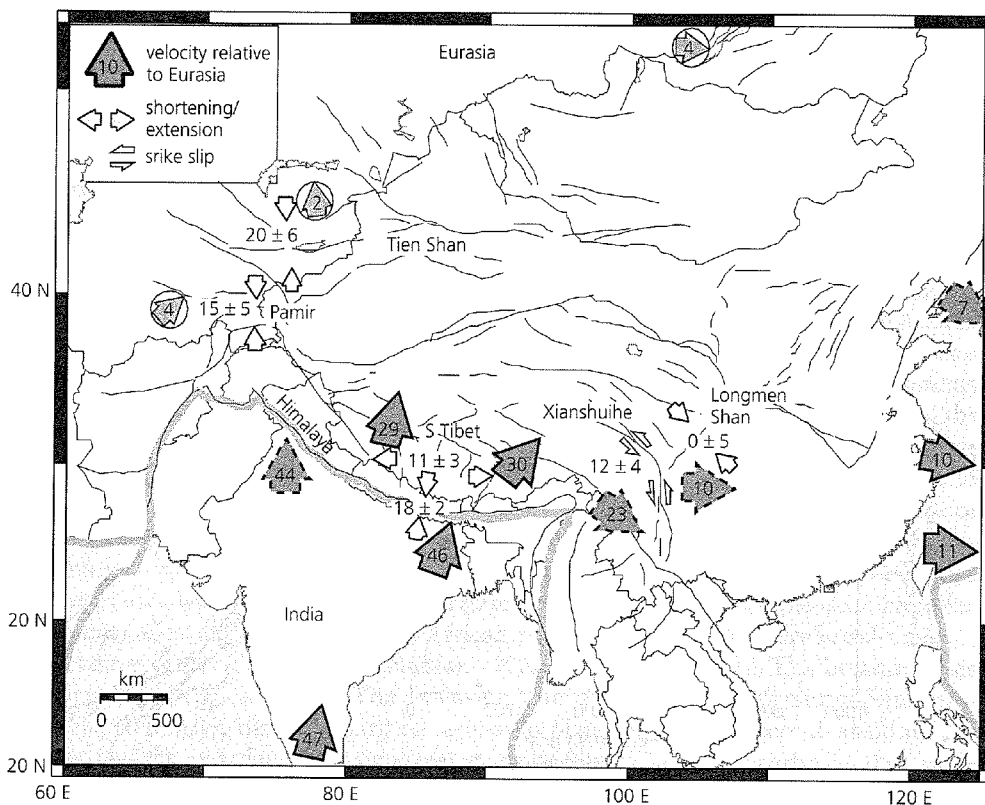


Fig. 5.6-5 Summary of crustal motions determined using space geodesy in the India–Eurasia plate collision zone. Large arrows indicate velocities relative to Eurasia. Arrows in circles show velocities with no significant motion with respect to Eurasia. Small arrows show local relative deformation. (Larson *et al.*, 1999. *J. Geophys. Res.*, 104, 1077–94, copyright by the American Geophysical Union.)

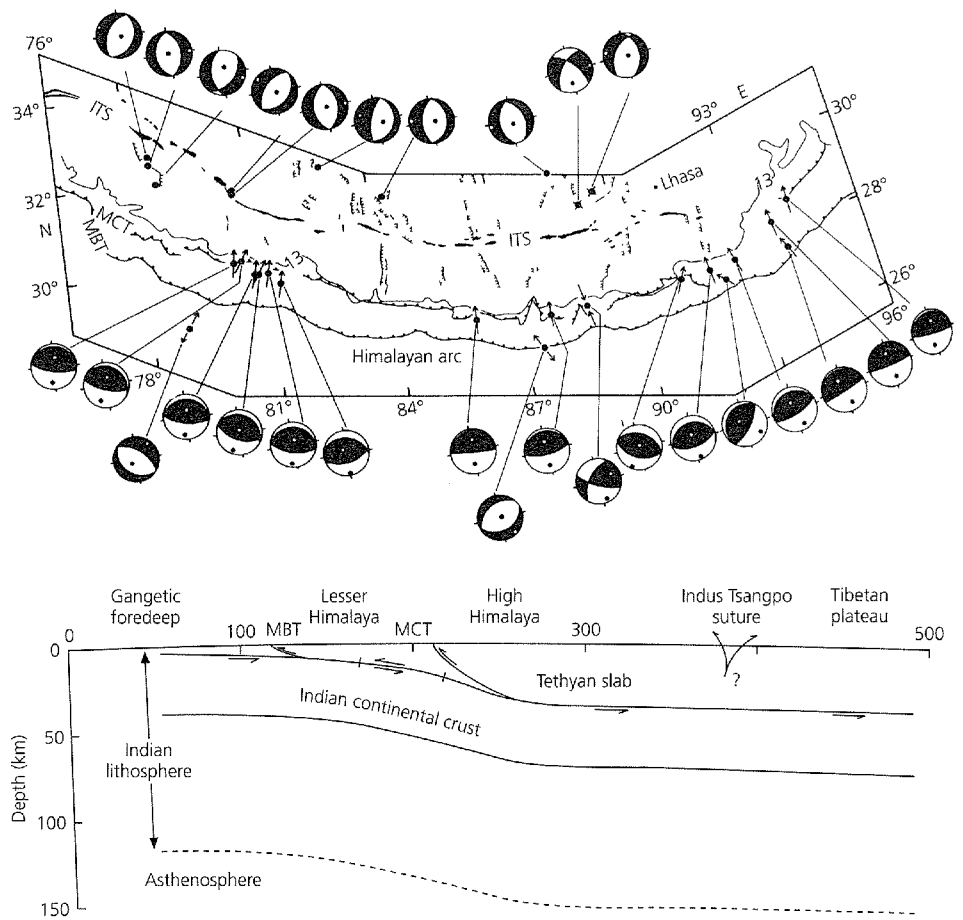


Fig. 5.6-6 Focal mechanisms and tectonic interpretation for the Himalayan continental convergence zone. MCT and MBT are the Main Central and Main Boundary thrust faults. (After Ni and Barazangi, 1984. *J. Geophys. Res.*, 89, 1147–64, copyright by the American Geophysical Union.)

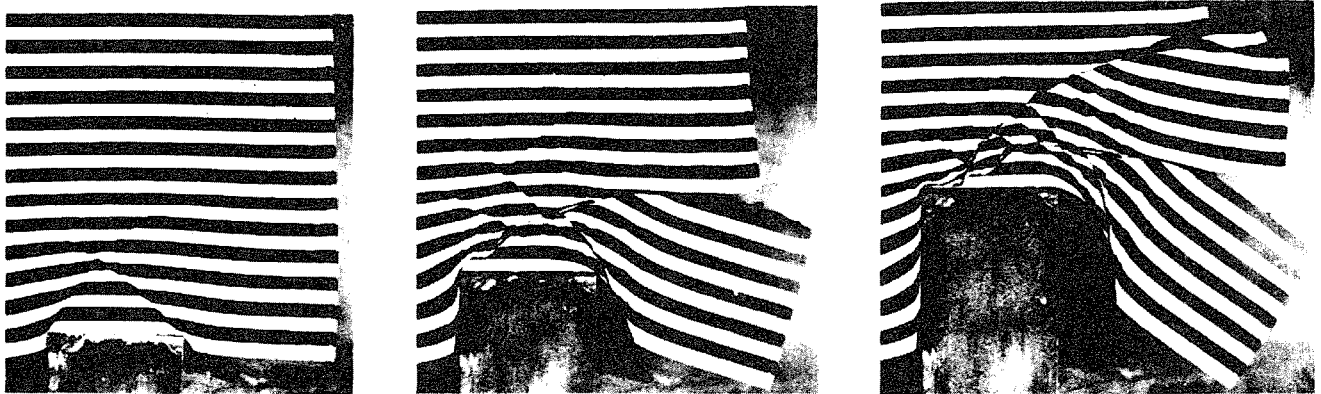


Fig. 5.6-7 Demonstration of the deformation of Asia, modeled by a striped block of plasticine, as the result of a collision with a rigid block simulating the Indian subcontinent. The plasticine is constrained on the left side, so the impact forces blocks to be extruded to the right, analogous to the eastward motion of blocks in Indochina and China. (Tapponnier *et al.*, 1982. *Geology*, 10, 611–16, with permission of the publisher, the Geological Society of America, Boulder, Co. © 1982 Geological Society of America.)

major strike-slip faults. This extrusion has been modeled assuming that India acts as a rigid block indenting a semi-infinite plastic medium (Asia), giving rise to a complicated faulting and slip pattern (Fig. 5.6-7). The extent of the collision is illustrated by GPS data and focal mechanisms showing that the Tien Shan

intracontinental mountain belt, 1000–2000 km north of the Himalayas, accommodates almost half the net plate convergence in the western part of the zone.

In addition to providing data about a collision region's kinematics, seismological studies provide insight into its mech-





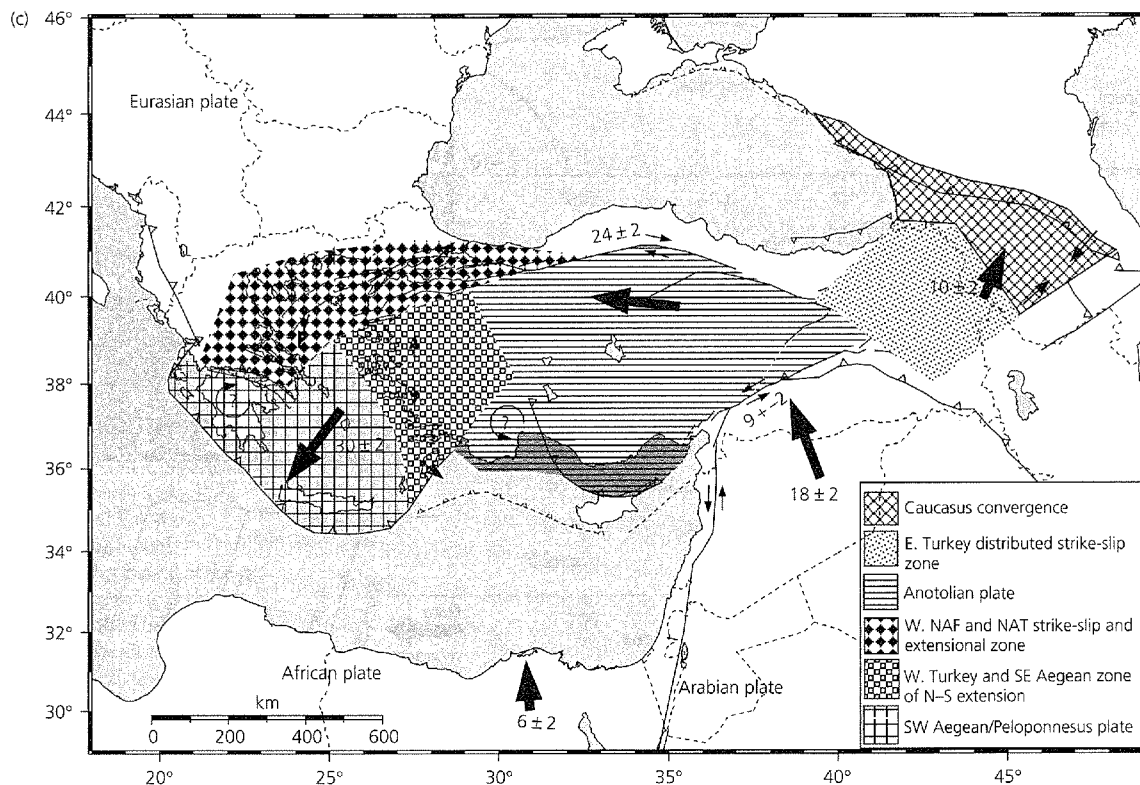


Fig. 5.6-8 GPS observations of motions relative to Eurasia (a), focal mechanisms (b), and tectonic interpretation (c) for a portion of the Africa–Arabia–Eurasia plate collision zone. Note strike-slip along the North Anatolian fault, extension in western Anatolia and the Aegean region, and compression in the Caucasus mountains. Rates are in mm/yr. (McClusky *et al.*, 2000. *J. Geophys. Res.*, 105, 5695–5719, copyright by the American Geophysical Union.)

anics. The collision process is thought to involve a complex interplay between forces due directly to the collision, gravitational forces due to the resulting uplift and crustal thickening, and forces from the resulting mantle flow. Earthquake depths and studies of seismic velocity, attenuation, and anisotropy are providing data on crustal thicknesses, thermal and mechanical structures, and mantle flow. For example, *P*-wave travel time tomography shows high velocity under the presumably cold Himalayas, which contrasts with low velocity under Tibet. These and other seismological data are consistent with the idea that Tibet deforms easily during the collision.

An equally complicated situation occurs in the eastern Mediterranean collision zone involving the African, Arabian, and Eurasian plates. Combining GPS and focal mechanism data shows the complex motions. Figure 5.6-8 (a) shows the motions of sites in the western Mediterranean relative to Eurasia. Northern portions of Arabia move approximately N40°W, consistent with global plate motion models. Western Turkey rotates as the Anatolian plate about a pole near the Sinai peninsula. Anatolia is thus “squeezed” westward between Eurasia and northward-moving Arabia (Fig. 5.6-8, c).<sup>2</sup> The motion across the North Anatolian fault, about 25 mm/yr, gives rise to

large right-lateral strike-slip earthquakes (Fig. 5.6-8, b) such as the 1999  $M_s$  7.4 Izmit earthquake, which occurred about 100 km east of Istanbul and caused more than 30,000 deaths. To the west, the data show interesting deviations from a rigid Anatolian plate. The increasing velocities toward the Hellenic trench, where the Africa plate subducts below Crete and Greece, show that western Anatolia and the Aegean region are under extension, consistent with the normal fault mechanisms. This region may be being “pulled” toward the arc, perhaps by an extensional process similar to oceanic back-arc spreading, as the trench “rolls back” (Section 5.2.4). By contrast, eastern Turkey is being driven north-ward into Eurasia, causing compression that appears as the thrust fault earthquakes in the Caucasus mountains. The Dead Sea transform separates Arabia from the region to the west, sometimes viewed as the Sinai microplate. Strike-slip motion along this fault gives rise to the earthquakes mentioned in the Bible that repeatedly destroyed famous cities like Jericho.

#### 5.6.2 Seismic, aseismic, transient, and permanent deformation

The examples in the previous section illustrate that earthquakes give powerful insights into the crustal deformation shaping the

<sup>2</sup> Consider a melon seed squeezed between a thumb and a forefinger.

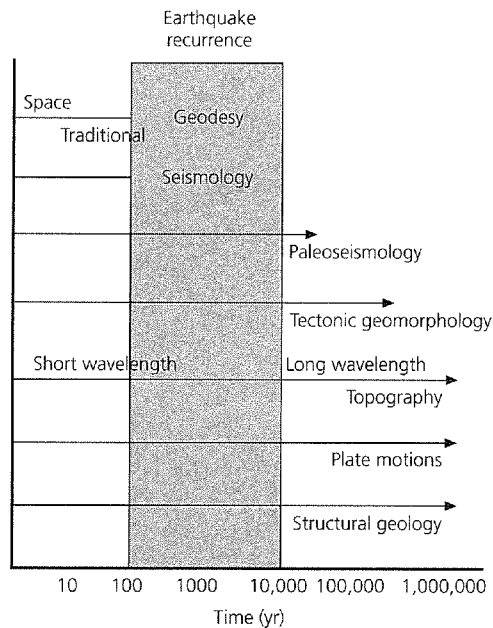


Fig. 5.6-9 Schematic illustration of how crustal deformation on various time scales is observed by different techniques.

continents. Other approaches to studying this deformation, including various geodetic and geological means, sample the deformation in different ways on various time scales (Fig. 5.6-9). Hence, considerable attention goes into understanding how what we see with these different techniques are related. For example, as discussed earlier (Sections 4.5.4, 5.4.3), in many places only part of the plate motion seems to occur as earthquakes, and the rest takes place as aseismic slip. A related question is how the deformation shown by earthquakes, which has a time scale of a few years, is related to the longer-term deformation that is recorded by topography and the geologic record.

To explore these ideas, consider the distribution of motion within the boundary zone extending from the stable interior of the oceanic Nazca plate, across the Peru–Chile trench to the coastal forearc, across the high Altiplano and foreland thrust belt, and into the stable interior of the South American continent. Figure 5.6-10 shows GPS site velocities relative to stable South America, which would be zero if the South American plate were rigid and all motion occurred at the trench plate boundary. However, the site velocities are highest near the coast and decrease relatively smoothly from the interior of the Nazca plate to the interior of South America.

Figure 5.6-10 (*bottom*) shows an interpretation of these data. In this model, about half of the plate convergence (approximately 35 mm/yr) is locked at the subduction interface, causing elastic strain of the overriding plate that will be released in large interplate thrust earthquakes (Section 4.5.4) like those whose focal mechanisms are shown. Thus the locked fraction of the plate motion corresponds to the seismic slip rate, perhaps via a process in which only a fraction of the interface is locked at any time. Approximately 20 mm/yr of the plate motion occurs by stable sliding at the trench, which does not

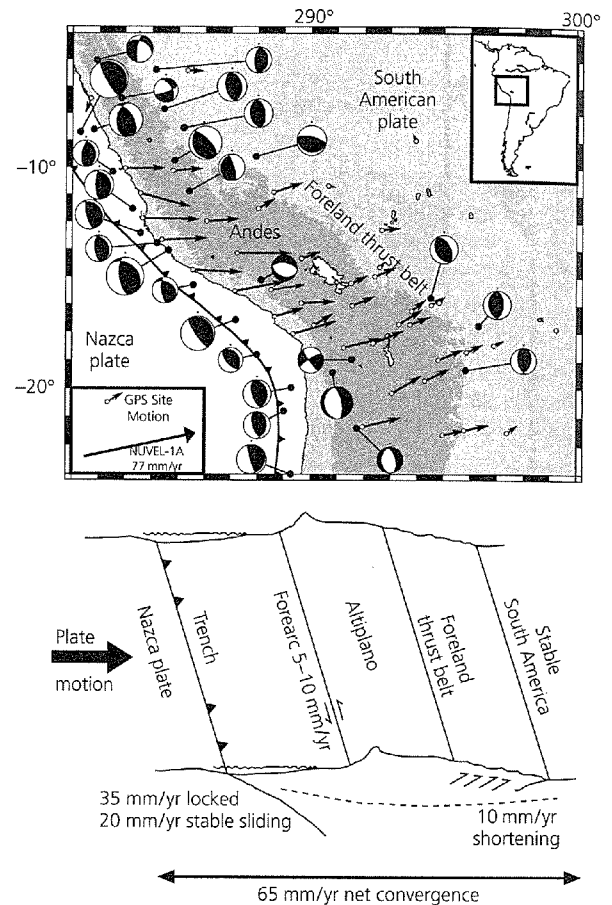


Fig. 5.6-10 *Top*: GPS site velocities relative to stable South America (Norabuena *et al.*, 1998. *Science*, 279, 358–62, copyright 1998 American Association for the Advancement of Science), and selected earthquake mechanisms in the boundary zone. Rate scale is given by the NUVEL-1A vector. *Bottom*: Cross-section showing approximate velocity distribution inferred from GPS data. (Stein and Klosko, 2002. From *The Encyclopedia of Physical Science and Technology*, ed. R. A. Meyers, copyright 2002 by Academic Press, reproduced by permission of the publisher.)

deform the overriding plate. This portion of the plate motion corresponds to aseismic slip. The rest occurs across the sub-Andean foreland fold-and-thrust belt, causing permanent shortening and mountain building, as shown by the inland thrust fault mechanisms. This portion of the plate motion would be considered aseismic slip if we considered only the fraction of the plate motion that appears in the trench seismic moment release, whereas in reality it occurs as inland deformation. These interpretations come from analyzing the GPS data in the convergence direction relative to the stable interior of South America (Fig. 5.6-11). If all the convergence were locked on the interplate thrust fault, the predicted rates would exceed those observed within about 200 km of the trench. However, if only about half of the predicted convergence goes into locking the fault, the predicted rates near the trench are less, because only the portion of the slip locked at the interface deforms the overriding plate. Similarly, the data farther than about 300 km

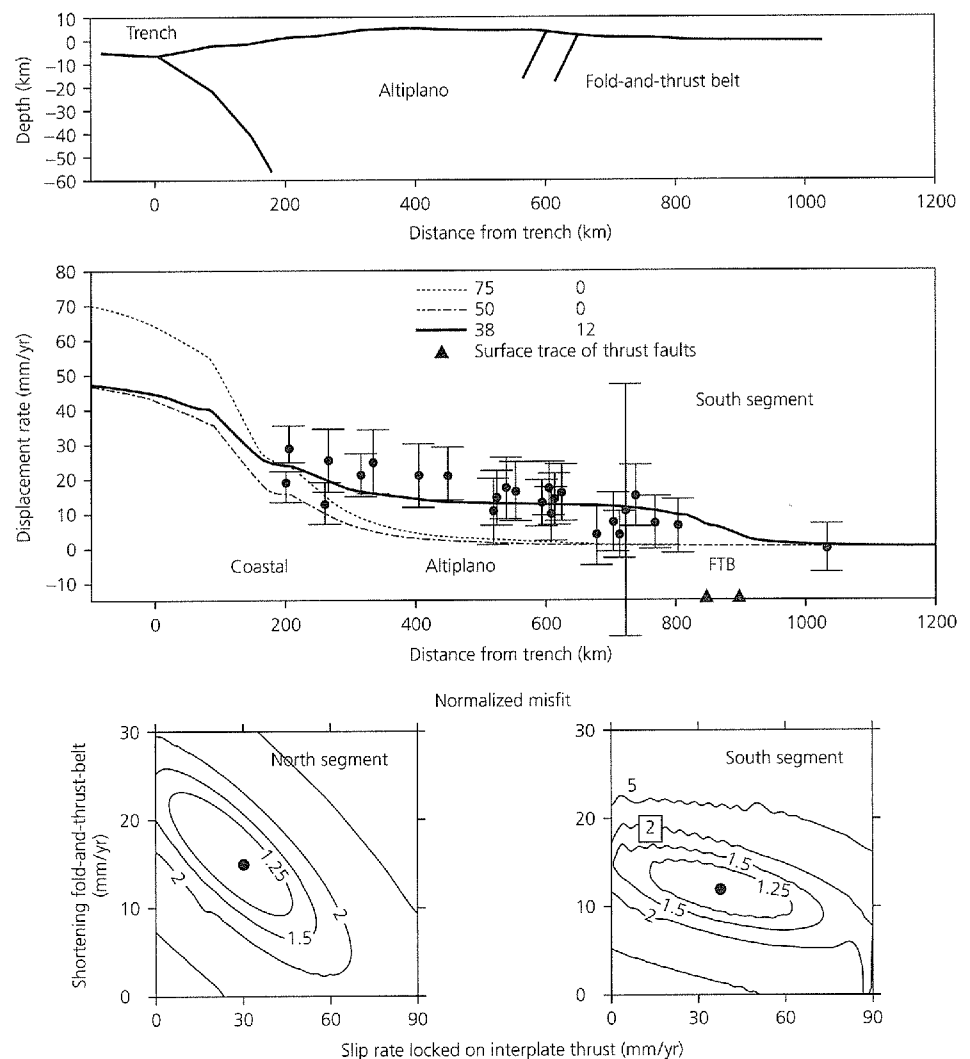


Fig. 5.6-11 Derivation of the model in Fig. 5.6-10 (bottom). Top: Model geometry, assuming partial slip locked at the plate boundary and shortening in the eastern Andes. Center: GPS site velocities in the convergence direction and various models, given by the rates of locked slip and shortening. Solid line shows predictions of best-fitting model, including both partial slip locked at the plate boundary and shortening in the eastern Andes. Short dashed line shows predictions of model with all slip locked on the plate boundary and no shortening. Long-short dashed line shows predictions of model with no shortening and partial slip locked on the plate boundary equal to the sum of best-fitting slip and shortening. Bottom: Contour plot showing misfit to the data as a function of the slip rate locked on the plate boundary and shortening rate in the eastern Andes. The best fits (dots) occur for about 30–40 mm/yr of locking and about 10–20 mm/yr shortening. (Norabuena *et al.*, 1998. *Science*, 279, 358–62, copyright 1998 American Association for the Advancement of Science.)

from the trench are better fit by assuming that about 10 mm/yr motion is locked on thrust faults in the eastern Andes. The locking and shortening rates are the best-fit parameters for this simple model, which does not include other possible complexities such as deformation in the Altiplano.

The idea that about 40% of the plate motion at the trench occurs by aseismic slip seems plausible, because studies using the history of large earthquakes at trenches often estimate that only about half the slip occurs seismically (Fig. 5.4-30). Given the problems of estimating source parameters of earthquakes from historical data, it is encouraging that the geodetic answer seems similar.

The relation between the shortening rate in the thrust belt inferred from GPS data and that implied by the earthquakes can also be studied. Assessing the seismic slip rate is a little more complicated than for transform faults (Section 5.3.3) or subduction zone thrust faulting (Section 5.4.3), because in continental deformation zones earthquakes occur over a dis-

tributed volume, rather than on a single fault, and have diverse focal mechanisms. Thus we sum the earthquakes' moment tensors (Section 4.4) to estimate a seismic strain rate tensor<sup>3</sup> using

$$\dot{\epsilon}_{ij} = \sum M_{ij} / (2\mu Vt), \quad (1)$$

where  $t$  is the time interval, and  $\mu$  is the rigidity.  $V$ , the assumed seismic source volume, the product of the length and width of the zone of seismicity and the depth to which seismicity extends. For example, the thrust belt can be assumed to be approximately 2000 km long, 250 km wide, and faulting extends to about 40 km depth. We can then diagonalize the result and consider the eigenvalue associated with the P axis. Scaling this value by the assumed zone width gives an estimate of the shortening rate. The resulting value, less than 2 mm/yr, is significantly less than the approximately 10 mm/yr indicated by the

<sup>3</sup> Strain rates are often written using a dot to indicate the time derivative.

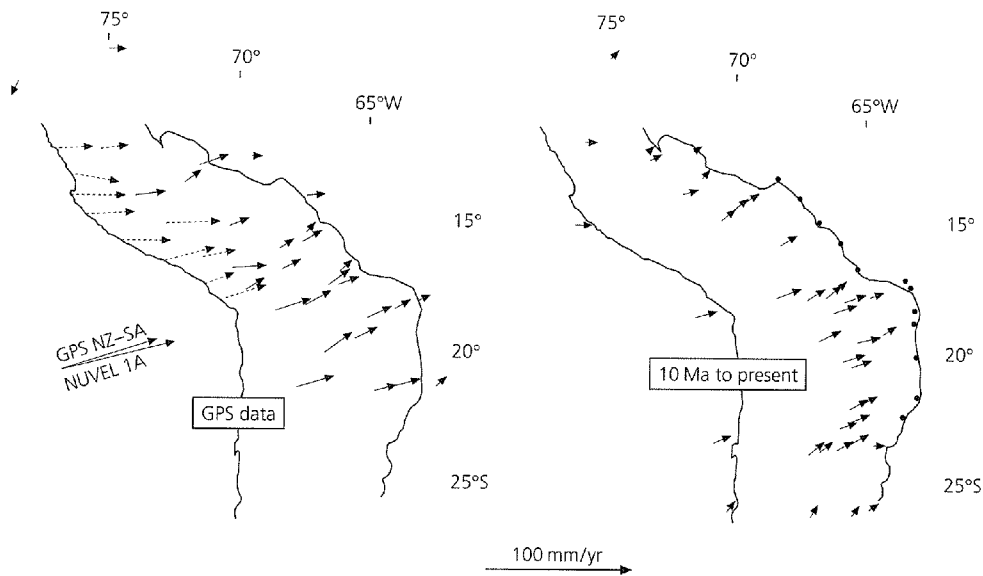


Fig. 5.6-12 Comparison of shortening across the Andes with respect to stable South America from GPS data (*left*) and geological studies (*right*). The dashed GPS vectors reflect elastic strain due to the earthquake cycle at the trench, and are not directly comparable to the permanent shortening in the geological data. Motion decreases toward the eastern extent of the mountain range, shown by the solid line. The geological vectors are largest at about 18°S and decrease to the north and south, showing how the variation in shortening that built the Andes bent them and made them widest about this point. (Hindle *et al.*, 2002.)

GPS data. Thus, even given the usual problem that the seismic history is short and may have missed the largest earthquakes, an effect one can attempt to correct for using earthquake frequency–magnitude data (Section 4.7.1), it looks like much of the shortening occurs aseismically.

An interesting question is how what we see today with earthquakes and GPS data relates to what occurs over geologic time. Figure 5.6-12 shows the results of geological studies, in which the arrows indicate the deformation that occurred over the past 10 Myr as the Andes formed. The directions and rates are similar to what are seen today, suggesting that the mountain building process has occurred relatively uniformly, although there have been some rate changes.<sup>4</sup>

Putting all this together gives some ideas about how the different measures of crustal deformation are related in this area. The first issue involves the relative amounts of seismic and aseismic deformation. It appears that about half of the plate motion at the trench occurs seismically. Similar fractions are also seen in other subduction zones (Fig. 5.4-30), implying that stable sliding at trenches is relatively common. Moreover, only about 10–20% of the shortening in the foreland thrust belt appears to occur seismically. Thus aseismic, and presumably permanent, deformation of rocks in the thrust belt seems like a major phenomenon. Similar results have also been observed for other continental deformation zones (Fig. 5.6-13). The next issue is that of permanent versus transient deformation. In the model of Fig. 5.6-11, the deformation of the South American plate due to the locked slip at the trench is transient, and will be released in the upcoming large trench earthquake. However, it seems likely that the deformation of the foreland thrust belt is permanent, and goes into faulting and folding rocks. Over

time, this permanent displacement adds up (Fig. 5.6-12) to build the mountains.

Similar studies are going on around the world, and should lead to an improved understanding of the partitioning between seismic, aseismic, transient, and permanent deformation. Models are being developed to explore these issues (Section 5.7), which are important both for understanding continental evolution and for earthquake hazard assessment, because an apparent seismic moment deficit could indicate either overdue earthquakes or aseismic deformation.

### 5.6.3 Continental intraplate earthquakes

Another important application of earthquake studies deals with the internal deformation of the continental portions of the major plates. Although idealized plates would be purely rigid, intraplate earthquakes reflect the important and poorly understood tectonic processes of intraplate deformation. As in the oceans (Section 5.5.1), there appears to be a hierarchy of places that have such earthquakes. There are areas like the East African rift that can be thought of as either slow-moving plate boundaries or intraplate deformation, less active zones associated with either fossil structures or other processes like hot spots, and then intraplate earthquakes that are not easily correlated with any particular structure or cause.

One example is the New Madrid area in the central USA, which had large earthquakes in 1811–12 and has small earthquakes today. Other continental interiors, including Australia, western Europe, and India, have also had significant intraplate earthquakes. Because motion in these zones is at most a few mm/yr, compared to the generally much more rapid plate boundary motions, seismicity is much lower (Fig. 5.6-14) and thus harder to study. This difficulty is compounded by the fact that, unlike at plate boundaries, where plate motions give insight into why and how often earthquakes occur, we have little

<sup>4</sup> The similarity of the focal mechanism, GPS, and geological data illustrates the principle of *uniformitarianism*, that studying present processes gives insight into the past, a tenet of geology since Lyell and Hutton's seminal work almost two centuries ago.

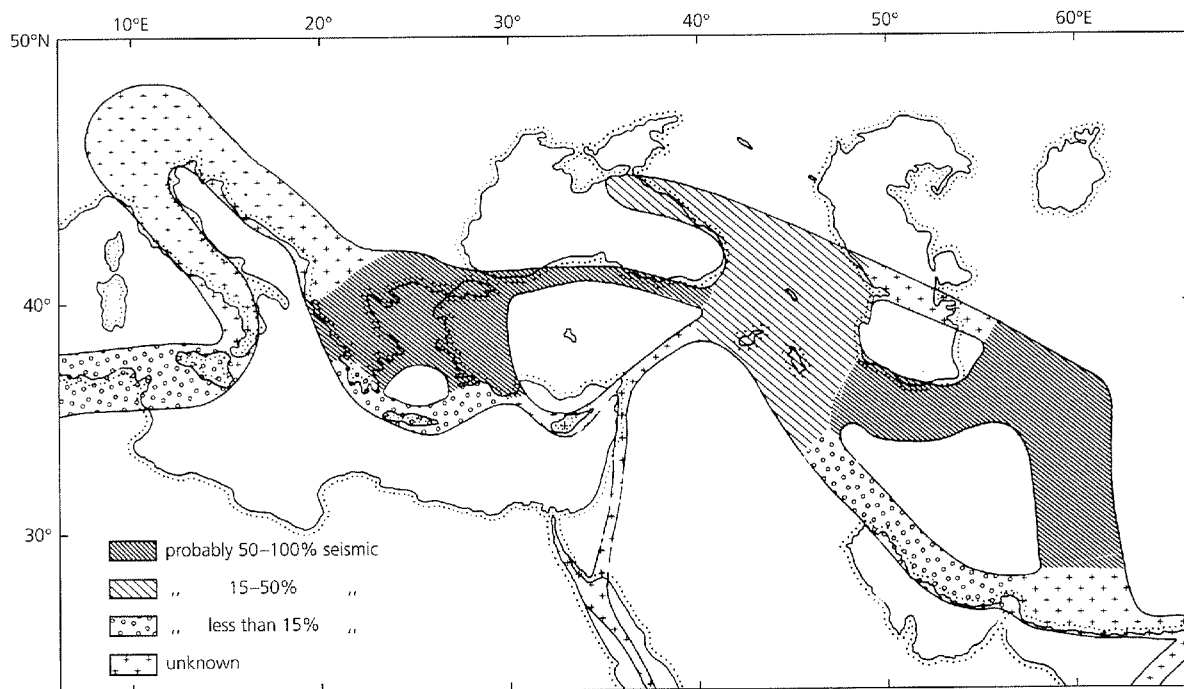


Fig. 5.6-13 Estimates of seismic deformation fractions for areas in the Mediterranean and Middle East. Seismicity appears to account for most or all of the deformation in western Turkey, Iran, and the Aegean, much of the deformation in the Caucasus and eastern Turkey, and little of the deformation in the Zagros and the Hellenic trench. (Jackson and McKenzie, 1988.)

idea of what causes intraplate earthquakes, and no direct way to estimate how often they should occur. As a result, progress in understanding these earthquakes is much slower than for earthquakes on plate boundaries, and key issues may not be resolved for a very long time.

Geodetic data illustrate the challenge. For example, comparison of the absolute velocities of GPS sites in North America east of the Rocky Mountains to velocities predicted by modeling these sites as being on a single rigid plate shows that the interior of the North American plate is rigid at least to the level of the average velocity residual, less than 1 mm/yr (Fig. 5.6-15). Similar results emerge from studies across the New Madrid zone itself and for the interiors of other major plates, showing that plates thought to have been rigid on geological time scales are quite rigid on decadal scales. For example, 1 mm/yr motion spread over 100 or 1000 km distance corresponds to strain rates of  $10^{-8}$  and  $10^{-9}$  yr $^{-1}$  ( $3 \times 10^{-16}$  and  $3 \times 10^{-17}$  s $^{-1}$ ), respectively. Because the geodetic data include measurement errors due to effects including instabilities of the geodetic markers, it seems likely that the tectonic strains are even smaller. However, over long enough time, even such small motions can accumulate enough slip for large earthquakes to occur.

This idea is consistent with what is known about large intraplate earthquakes. Although there is little seismological data for such events because they are rare, insight can be obtained from combining the seismological data with geodetic, paleoseismological, and other geological and geophysical data.

For example, intensities estimated from historical accounts of the 1811–12 New Madrid earthquakes (Fig. 1.2-4) suggest magnitudes in the low 7 range. Paleoseismic studies (Section 1.2) indicate that several previous large earthquakes, presumably comparable to those of 1811–12, occurred 500–800 years apart. Thus, in 500–1000 years (Fig. 5.6-16, *top*) steady strain accumulation less than 2 mm/yr could provide up to 1–2 m of motion available for future earthquakes, suggesting that they would be about magnitude 7. A similar view comes from considering the earthquake history for the area. As discussed in Section 4.7.1, earthquakes of a given magnitude are approximately ten times less frequent than those one magnitude unit smaller. Thus, although the instrumental data contain no earthquakes with magnitude greater than 5, both these and a historical catalog in which magnitudes were estimated from intensity data can be extrapolated to imply that a magnitude 7 earthquake would occur about once every  $1400 \pm 600$  years (Fig. 5.6-16, *bottom*). Hence, as expected, major intracontinental earthquakes occur substantially less frequently than comparable plate boundary events (Fig. 5.6-17). However, because of the lower attenuation in continental interiors (Section 3.7.10), such earthquakes can cause greater shaking than ones of the same magnitude on a plate boundary (Fig. 1.2-5).

Such earthquakes are generally thought to be due to the reactivation of preexisting faults or weak zones in response to either local or intraplate stresses. The New Madrid earthquakes, for example, are thought to occur on faults associated



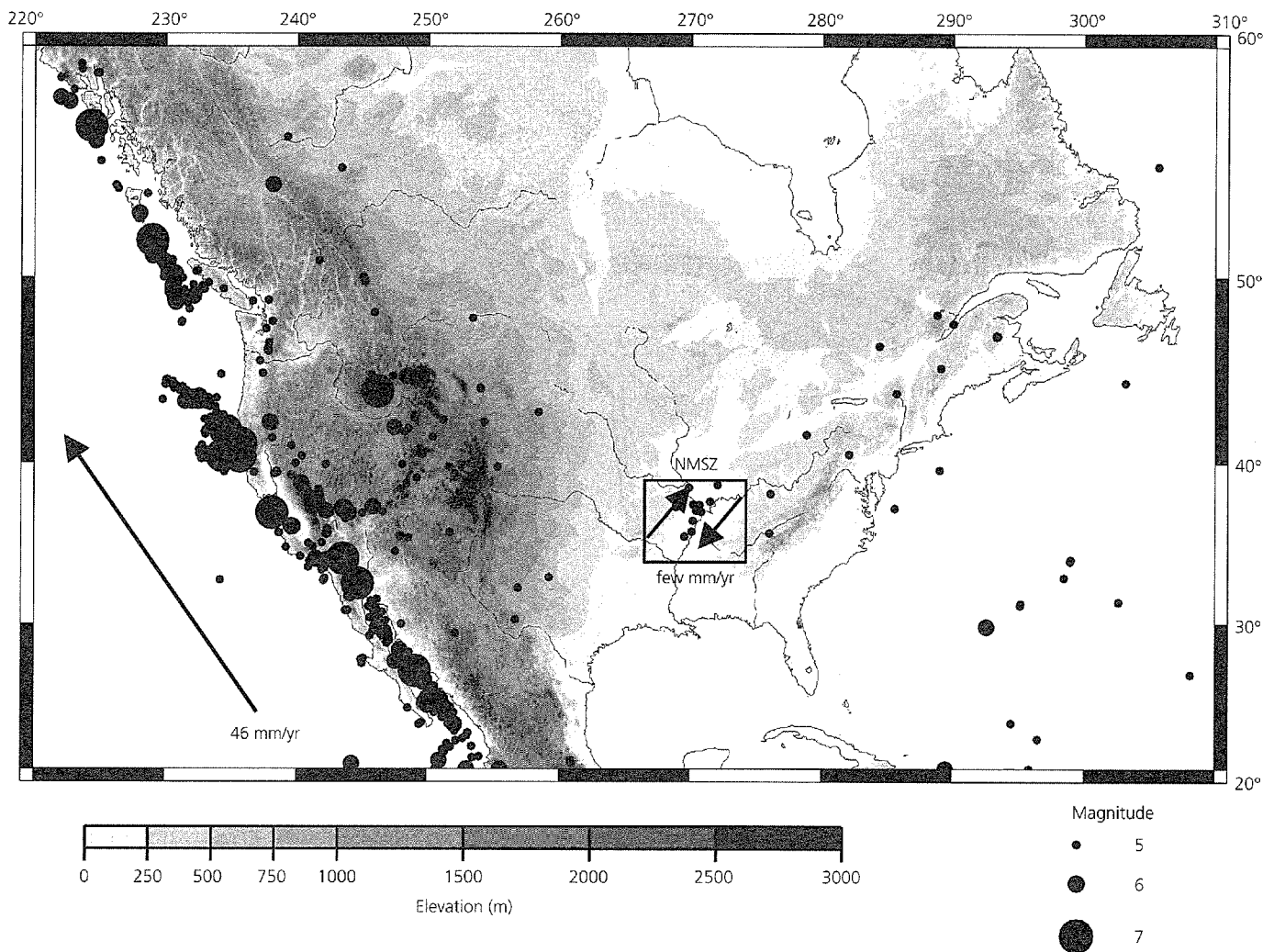


Fig. 5.6-14 Seismicity (magnitude 5 or greater since 1965) of the continental portion of the North American plate and adjacent area. Seismicity and deformation are concentrated along the Pacific–North America plate boundary zone, reflecting the relative plate motion. The remaining eastern portion of the continent, approximately that east of 260°, is much less seismically active. Within this relatively stable portion of the continent, seismicity, and thus presumably deformation, are concentrated in several zones, most notably the New Madrid seismic zone. (Weber *et al.*, 1998. *Tectonics*, 17, 250–66, copyright by the American Geophysical Union.)

with a Paleozoic failed continental rift, now buried beneath thick sediments deposited by the Mississippi river and its ancestors (Fig. 5.6-18). As a result, the faults are not exposed at the surface, so most ideas about them are based on inferences from seismology and other data. The intraplate stress field has been studied by combining focal mechanism and fault orientations with data from drill holes and *in situ* stress measurements (Fig. 5.6-19). In general, the eastern USA shows a maximum horizontal stress oriented NE–SW, consistent with the predictions of the stresses due to plate driving forces. Similar stress maps are being developed for other areas and are being used to investigate both intraplate deformation and plate driving forces. As noted in Section 3.6.5, it appears that seismic anisotropy in the lower continental crust may reflect the stress field that acted during a major tectonic event such as mountain building.

An intriguing question is why intraplate stresses cause earthquakes on particular faults, given that many weak zones could serve this purpose. Geological and paleoseismic data, together with the absence of significant fault-related topography, suggest that individual intraplate seismic zones may be active for only a few thousands of years, so intraplate seismicity migrates. This possibility is akin to that suggested for intermittent oceanic intraplate earthquake swarms. If so, there is nothing special about New Madrid or the other concentrations of intraplate seismicity we observe now — these zones will die off and be replaced by others. Moreover, there are enough tectonic structures available that (typically small) earthquakes will occur almost randomly throughout continental interiors.

A special case of this phenomenon occurs at passive continental margins, where continental and oceanic lithospheres join. Although these areas are in general tectonically inactive,

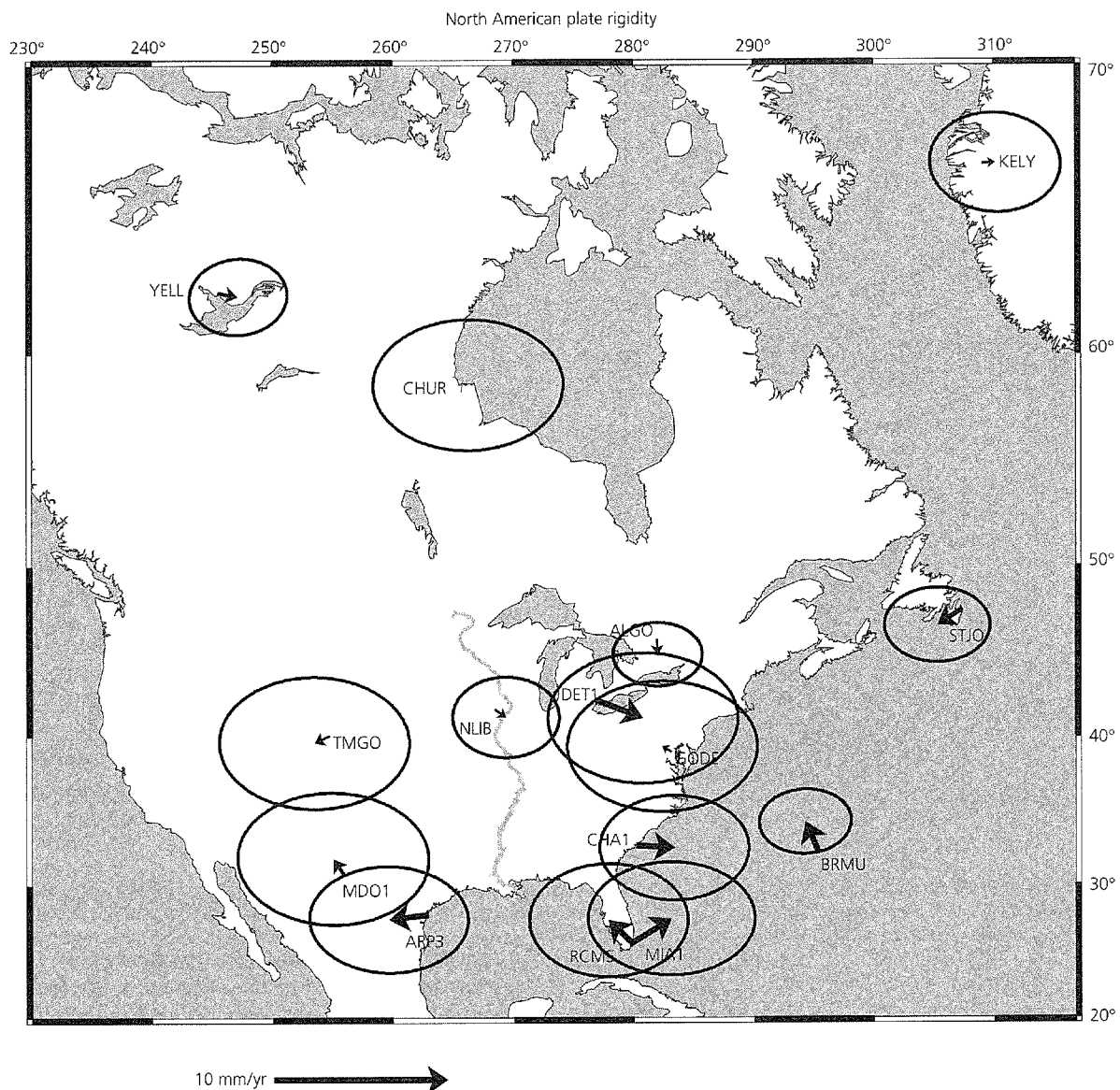


Fig. 5.6-15 Locations of continuously recording GPS sites used to estimate a Euler vector for the presumably stable portion of North America. For each, the misfit between the observed velocity and that predicted for a single plate is shown. The average misfit is less than 1 mm/yr, showing that eastern North America is quite rigid. (Newman *et al.*, 1999. *Science*, 284, 619–21, copyright 1999 American Association for the Advancement of Science.)

magnitude 7 earthquakes can occur, as on the eastern coast of North America (Fig. 5.6-20). Such earthquakes may be associated with stresses, including those due to the removal of glacial loads, which reactivate faults remaining from the original continental rifting (Fig. 5.6-1). Although such earthquakes are observed primarily on previously glaciated margins, they also occur on nonglaciated passive margins, perhaps due to sediment loading. In some cases large sediment slides occur, as was noted for the 1929  $M_s$  7.2 earthquake on the Grand Banks of Newfoundland, because the slides broke trans-Atlantic telephone cables and generated a tsunami that caused

27 fatalities.<sup>5</sup> An interesting unresolved question is whether tectonic faulting is required for such earthquakes, or whether the slump itself can account for what is seen on seismograms. Some studies find that the seismograms are best fit by a double-couple fault source, whereas others favor a single force consistent with the slump (Fig. 4.4-3). The issue is important because slumps occur in the sedimentary record along many passive

<sup>5</sup> These deaths account for all but one of Canada's known earthquake fatalities to date, although this situation could change after a large Cascadia subduction zone earthquake.



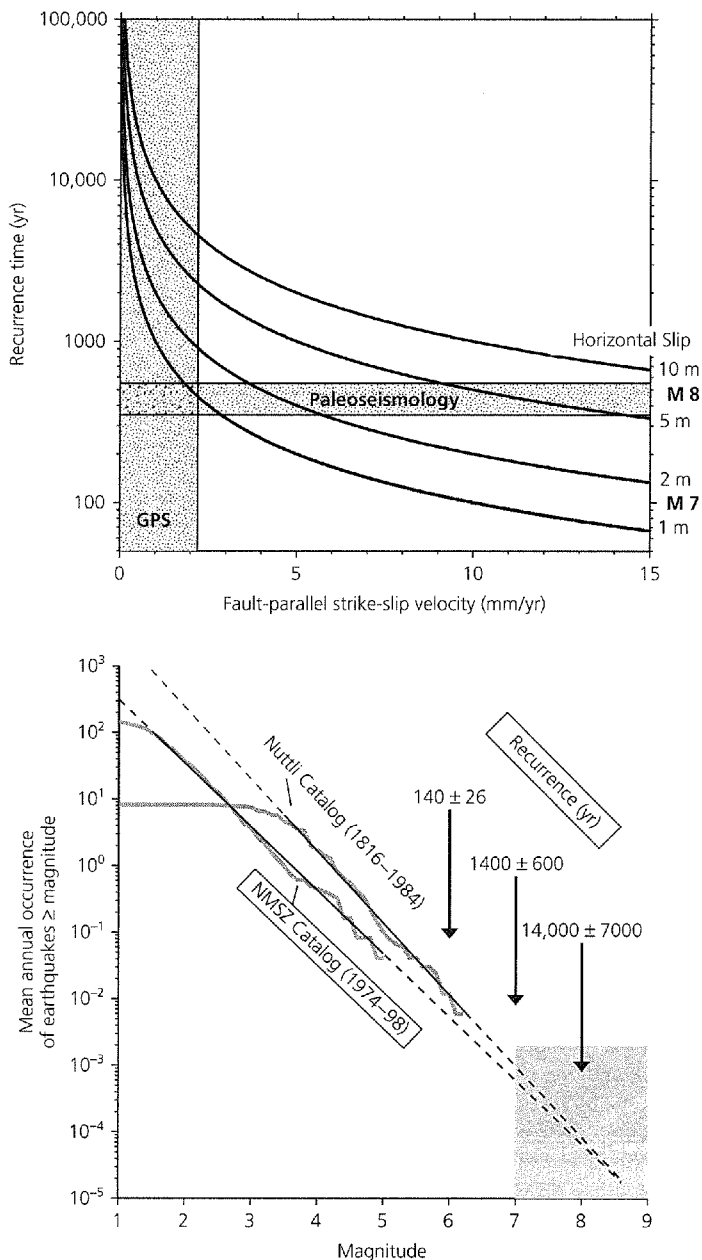


Fig. 5.6-16 Top: Relation between interseismic motion and the expected recurrence of large New Madrid earthquakes. The recurrence estimates from paleoseismic studies and geodetic data are jointly consistent with slip in the 1811–12 earthquakes of about 1 m, corresponding to a low magnitude 7 earthquake. Bottom: Earthquake frequency–magnitude data for the New Madrid zone. Both the instrumental and historic (1816–1984) data predict a recurrence interval of about 1000 yr for magnitude 7 earthquakes. (Newman *et al.*, 1999, *Science*, 284, 619–21, copyright 1999 American Association for the Advancement of Science.)

margins, even those that have not been recently deglaciated. Stresses associated with the removal of glacial loads may also play a role in causing earthquakes within continental interiors such as the northeastern USA and eastern Canada. It has also been suggested that the huge 1998 Balleny Island intraplate

earthquake (Section 5.5.1) may have been triggered by stresses due to the shrinking Antarctic ice cap.

As in the oceans, another interesting class of intraplate seismicity is associated with hot spots. The area near the Yellowstone hot spot in the western USA shows an intriguing pattern of seismicity along the margins of the Snake River plain (Fig. 5.6-21), which is the volcanic track the hot spot produced as the North American plate moved over it (Fig. 5.2-8). This seismicity, which includes the 1959  $M_s$  7.5 Hebgen Lake, Montana,<sup>6</sup> and 1983  $M_s$  7.3 Borah Peak, Idaho, earthquakes, forms a parabolic pattern extending southwestward from Yellowstone itself. It thus stands out from the regional seismicity (Fig. 5.2-3) associated with the extensional tectonics of the eastern portion of the Basin and Range province, termed the Intermountain Seismic Belt. The absence of seismicity along the track itself seems likely to be a consequence of the thermal and magmatic perturbations produced by the hot spot, although the specific mechanism is still under discussion. Seismic tomography (Fig. 5.6-21) shows a low-velocity anomaly in the crust and upper mantle under Yellowstone itself, presumably due to partial melting and hydrothermal fluids, and a deeper anomaly that persists along the track.

In summary, although continental intraplate seismicity is a minor fraction of global seismic moment release, it has both scientific and societal interest precisely because it is rare. It provides one of our few ways of studying the limits of plate rigidity and intraplate stresses, and poses the challenge of deciding the appropriate level of earthquake preparedness for rare, but potentially destructive, earthquakes.

## 5.7 Faulting and deformation in the earth

Because earthquake faulting is a spectacular manifestation of the processes that deform the solid earth, we seek to understand how earthquakes result from and reflect this deformation. Valuable insight comes from laboratory experiments and theoretical models for the behavior of solid materials. Although the experiments and models are much simpler than the complexities of the real earth, they allow us to think about key features. Seismology and geophysics thus exploit research devoted to material behavior by a range of disciplines, including engineering, materials science, and solid state physics. We touch only briefly on some basic ideas, and more information can be found in the references at the end of the chapter.

### 5.7.1 Rheology

Materials can be characterized by their *rheology*, the way they deform. In seismology we typically take a continuum

<sup>6</sup> This earthquake triggered an enormous landslide that buried a campground, causing 28 deaths and dammed the Madison River, forming Quake Lake. These dramatic effects are still visible today and make the site well worth visiting. A visitor center and parking lot are built on the slide.

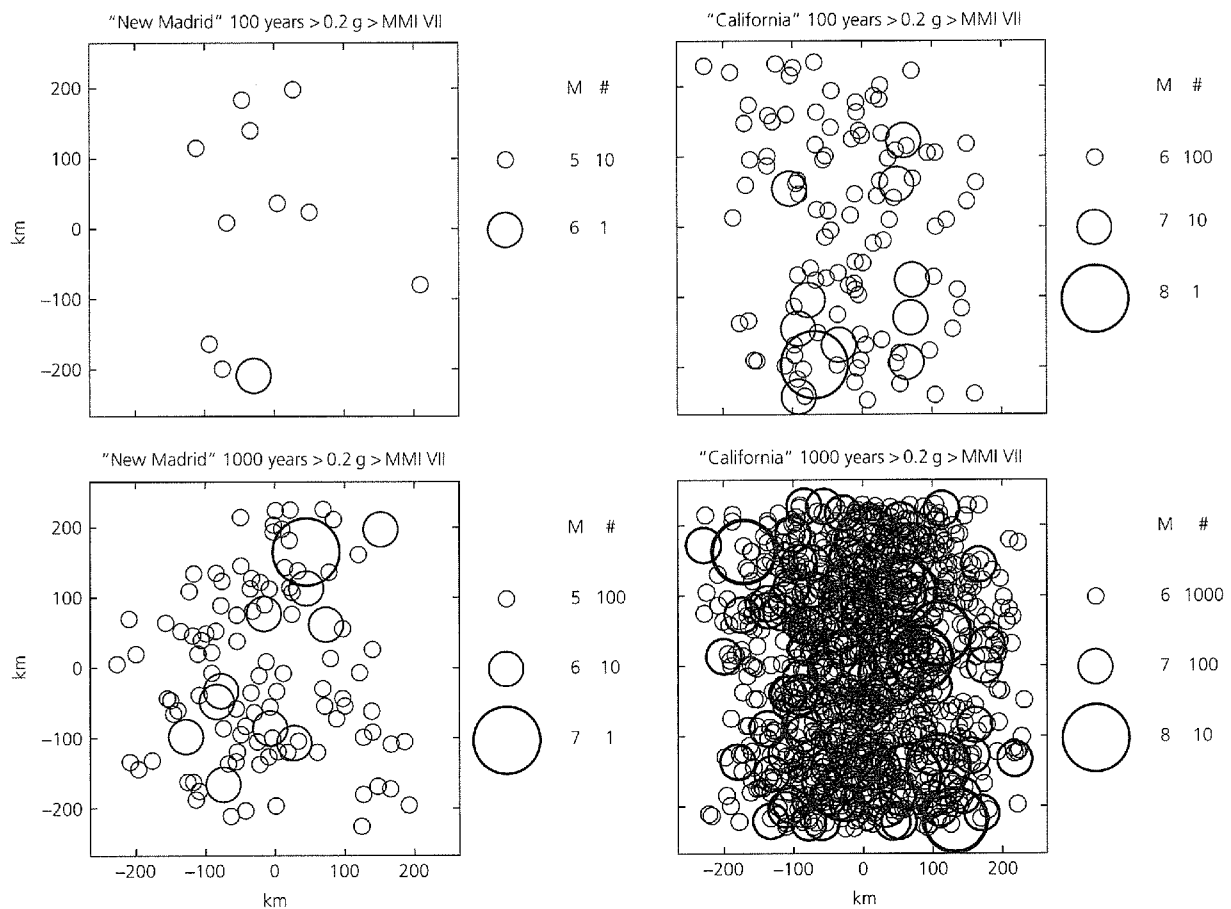


Fig. 5.6-17 Schematic illustration of the relation between the recurrence times of seismicity and resulting seismic hazard for the intraplate New Madrid seismic zone and the southern California plate boundary zone. Seismicity is assumed to be randomly distributed about an N-S line through 0, with California 100 times more active, but New Madrid earthquakes causing potentially serious damage (circles show areas with acceleration 0.2 g or greater, Table 1.2-4) over an area comparable to that for a California earthquake one magnitude unit larger.

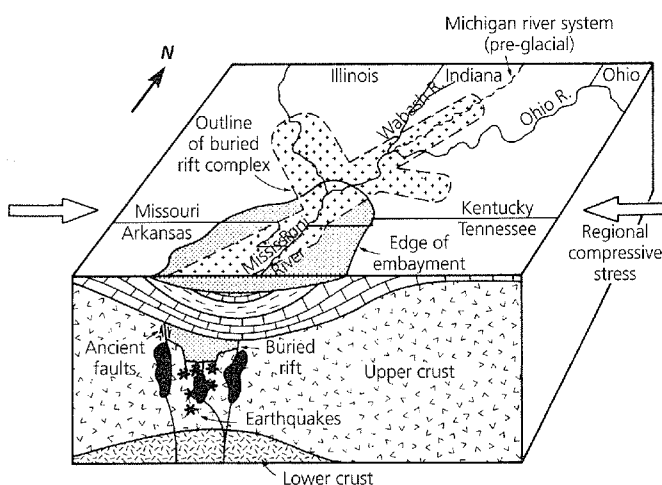


Fig. 5.6-18 Schematic tectonic model for the New Madrid earthquakes. (Braille *et al.*, 1986. *Tectonophysics*, 131, 1-21, with permission from Elsevier Science.)

approach, considering the earth to be a continuous deformable material. This means that we focus on its aggregate behavior (Section 2.3) rather than on how its behavior is determined by what happens at a microscopic scale.

To do this, consider the strain that results from compressing a rock specimen. The simplest case is shown in Fig. 5.7-1a. For small stresses, the resulting strain is proportional to the applied stress, so the material is purely *elastic*. Elastic behavior happens when seismic waves pass through rock, because the strains are small (Section 2.3.8). However, once the applied stress reaches a value  $\sigma_f$ , known as the rock's *fracture strength*, the rock suddenly breaks. Such *brittle* fracture is the simplest model for what happens when an earthquake occurs on a fault. Thus brittle fracture — a deviation from elasticity — generates elastic seismic waves.

Other materials show a change in the stress-strain curve for increasing stresses (Fig. 5.7-1b). For stresses less than the *yield stress*,  $\sigma_o$ , the material acts elastically. Thus, if the stress is released, the strain returns to zero. However, for stresses greater

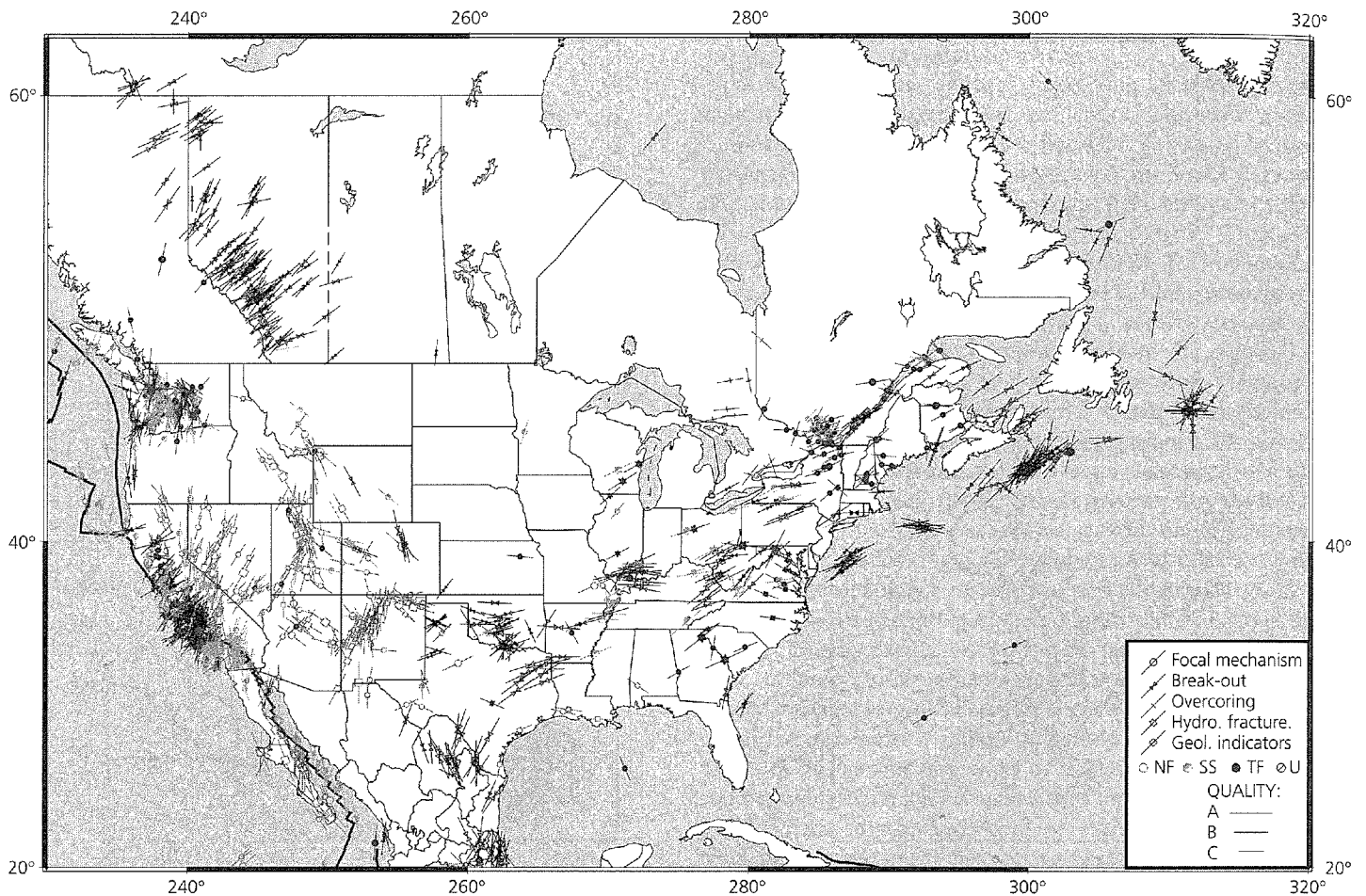


Fig. 5.6-19 Stress map for North America. (World Stress Map project, 2000. Courtesy of the US Geological Survey.)

than the yield stress, releasing the stress relieves the elastic portion of the strain, but leaves a permanent deformation (Fig. 5.7-1c). If the material is restressed, the stress-strain curve now includes the point of the permanent strain. The material behaves as though its elastic properties were unchanged, but the yield strength has increased from  $\sigma_0$  to  $\sigma'_0$ . The portion of the stress-strain curve corresponding to stress above the yield stress is called *plastic* deformation, in contrast to the elastic region where no permanent deformation occurs. Materials showing significant plasticity are called *ductile*. A common approximation is to treat ductile materials as *elastic-perfectly plastic*: stress is proportional to strain below the yield stress and constant for all strains when stress exceeds the yield stress (Fig. 5.7-2).

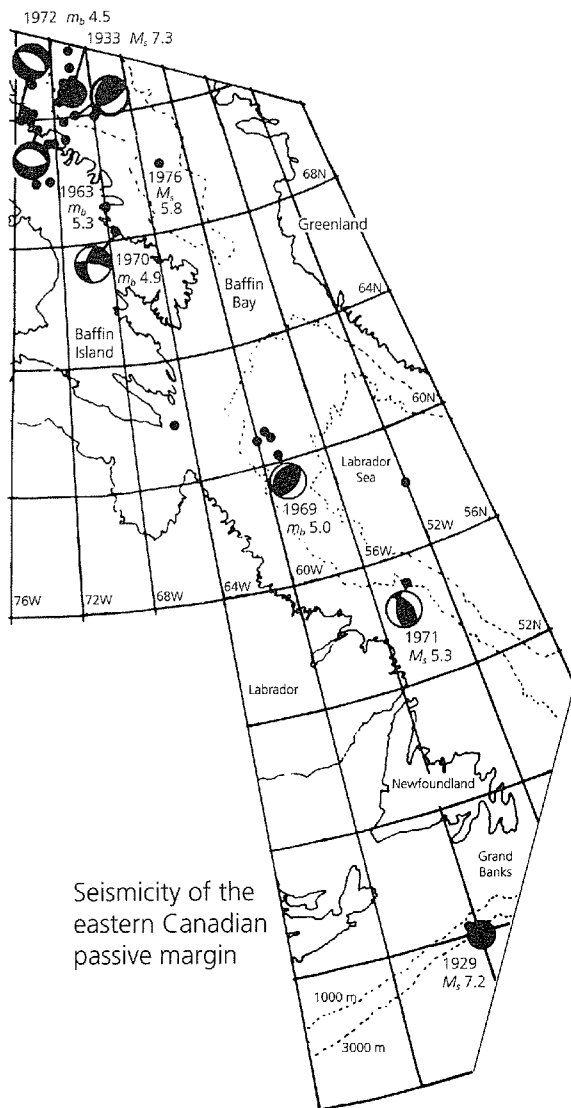
An important result of laboratory experiments is that at low pressures rocks are brittle, but at high pressures they behave ductilely, or flow. Figure 5.7-3 shows experiments where a rock is subjected to a compressive stress  $\sigma_1$  that exceeds a confining pressure  $\sigma_3$ . For confining pressures less than about 400 MPa the material behaves brittly — it reaches the yield strength, then fails. For higher confining pressures the material

flows ductilely. These pressures occur not far below the earth's surface — as discussed earlier, 3 km depth corresponds to 100 MPa pressure — so 800 MPa is reached at about 24 km. This experimental result is consistent with the idea that the strong lithosphere is underlain by the weaker asthenosphere.

A related phenomenon is that materials behave differently at different time scales. A familiar example is that although an asphalt driveway is solid if one falls on it, a car parked on it during a hot day can sink a little ways into it. On short time scales the driveway acts rigidly, but on longer time scales it starts to flow as a *viscous fluid*. This effect is crucial in the earth, because the mantle is solid on the time scale needed for seismic waves to pass through it, but flows on geological time scales.

### 5.7.2 Rock fracture and friction

The first question we address is how and when rocks break. In the brittle regime of behavior, the development of faults and the initiation of sliding on preexisting faults depend on the applied stresses.



Seismicity of the eastern Canadian passive margin

Fig. 5.6-20 Earthquakes along the passive continental margin of eastern Canada. These earthquakes may have occurred on faults remaining from continental rifting. (Stein *et al.*, 1979. *Geophys. Res. Lett.*, 6, 537–40, copyright by the American Geophysical Union.)

Given a stress field specified by a stress tensor, we use the approach of Section 2.3.3 to find the variation in normal and shear stress on faults of various orientations. For simplicity, we consider the stress in two dimensions. If the coordinate axes ( $\hat{e}_1, \hat{e}_2$ ) are oriented in the principal stress directions, the stress tensor is diagonal,

$$\sigma_{ij} = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}. \quad (1)$$

To find the stress on a plane whose normal  $\hat{e}'_1$  is at an angle of  $\theta$  from  $\hat{e}_1$ , the direction of  $\sigma_1$  (Fig. 5.7-4), we transform the stress

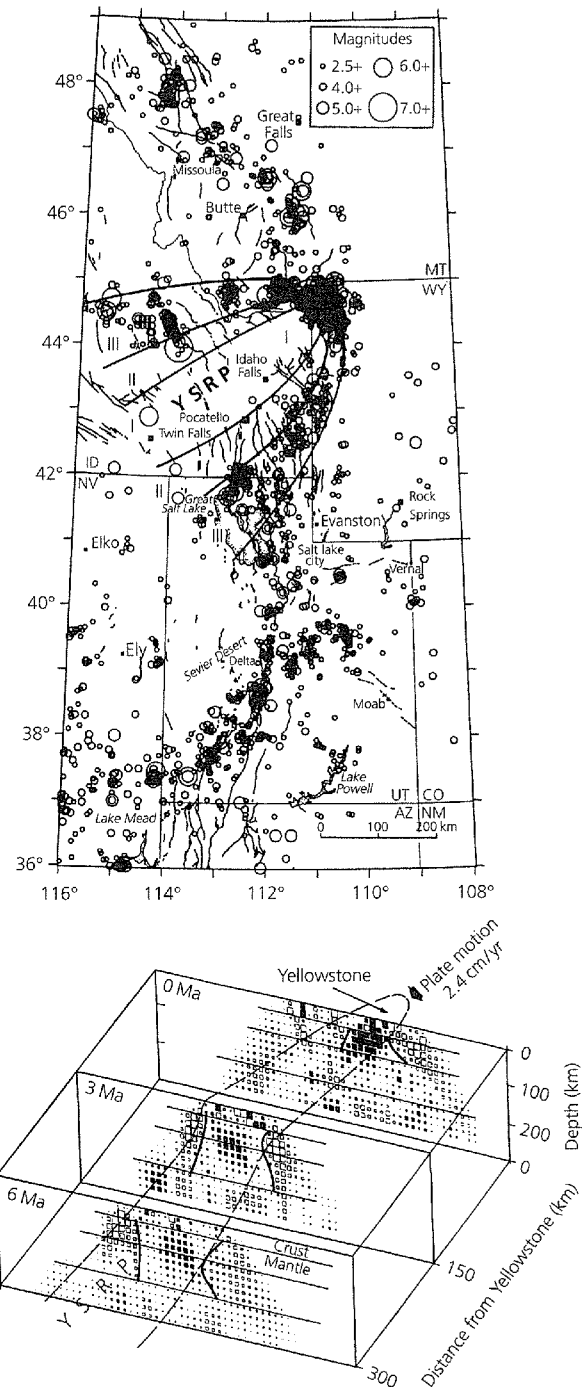


Fig. 5.6-21 Top: Seismicity (1900–85) of the Intermountain area of the western USA. Superimposed on the regional seismicity are earthquakes forming a parabola along the margins of the Yellowstone–Snake River plain (YSRP), the volcanic track of the Yellowstone hot spot. Bottom: P-wave velocities across the hot spot track, shown by squares scaled in size to the differences from a uniform-velocity model. The largest symbols are  $\pm 3\%$ , with dark and open symbols showing low and high velocities. (Smith and Braile, 1994. *J. Volcan. Geotherm. Res.*, 61, 121–87, with permission from Elsevier Science.)

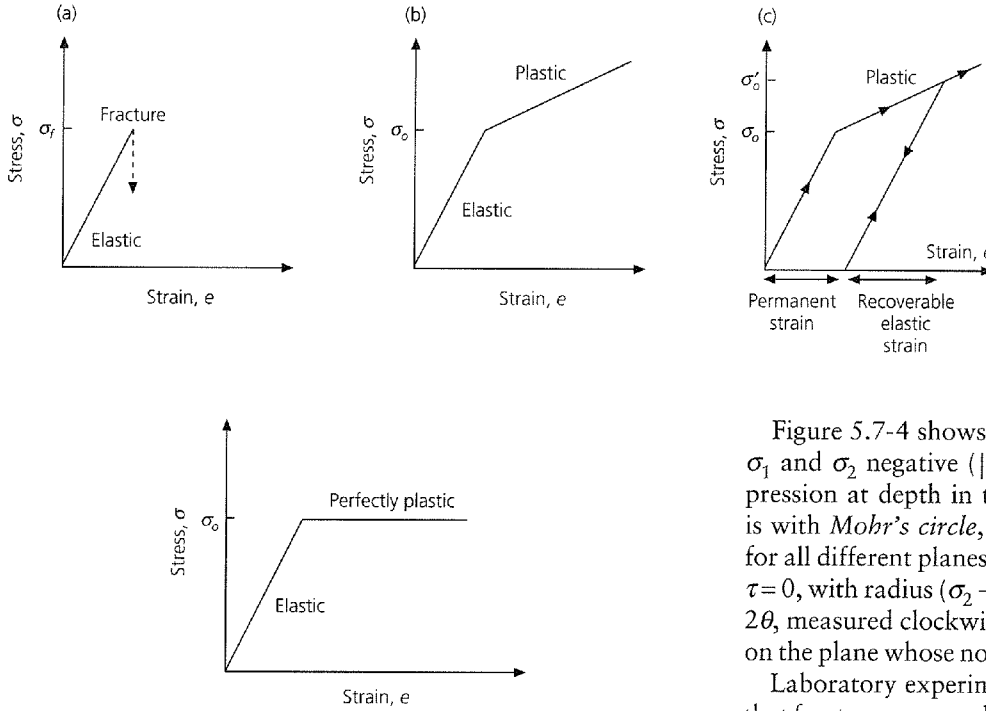


Fig. 5.7-1 (a): A material is perfectly elastic until it fractures when the applied stress reaches  $\sigma_f$ . (b): A material undergoes plastic deformation when the stress exceeds a yield stress  $\sigma_o$ . (c): A permanent strain results from plastic deformation when the stress is raised to  $\sigma_o$  and released.

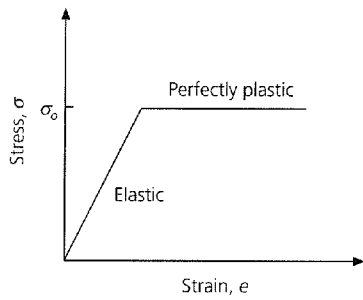


Fig. 5.7-2 An elastic-perfectly plastic rheology, which is a commonly used approximation for the behavior of ductile materials.

tensor from the principal axis coordinate system to a new coordinate system using the transformation matrix (Section 2.3.3)

$$A = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \quad (2)$$

so that the stress in the new (primed) system is

$$\begin{aligned} \sigma'_{ij} &= A \sigma A^T = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1 \cos^2 \theta + \sigma_2 \sin^2 \theta & (\sigma_2 - \sigma_1) \sin \theta \cos \theta \\ (\sigma_2 - \sigma_1) \sin \theta \cos \theta & \sigma_1 \sin^2 \theta + \sigma_2 \cos^2 \theta \end{pmatrix}. \end{aligned} \quad (3)$$

The normal and shear stresses on the plane vary, depending on the plane's orientation. The normal stress component, denoted by  $\sigma$ , is

$$\sigma = \sigma'_{11} = \sigma_1 \cos^2 \theta + \sigma_2 \sin^2 \theta = \frac{(\sigma_1 + \sigma_2)}{2} + \frac{(\sigma_1 - \sigma_2)}{2} \cos 2\theta, \quad (4a)$$

and the shear component, denoted by  $\tau$ , is

$$\tau = \sigma'_{12} = (\sigma_2 - \sigma_1) \sin \theta \cos \theta = \frac{(\sigma_2 - \sigma_1)}{2} \sin 2\theta. \quad (4b)$$

Figure 5.7-4 shows  $\sigma$  and  $\tau$  as functions of  $\theta$  for the case of  $\sigma_1$  and  $\sigma_2$  negative ( $|\sigma_1| > |\sigma_2|$ ), which corresponds to compression at depth in the earth. A graphic way to show these is with *Mohr's circle*, a plot of  $\sigma$  versus  $\tau$  (Fig. 5.7-5). Values for all different planes lie on a circle centered at  $\sigma = (\sigma_1 + \sigma_2)/2$ ,  $\tau = 0$ , with radius  $(\sigma_2 - \sigma_1)/2$ . The point on the circle with angle  $2\theta$ , measured clockwise from the  $-\sigma$  axis, gives the  $\sigma$ ,  $\tau$  values on the plane whose normal is at angle  $\theta$  to  $\sigma_1$ .<sup>1</sup>

Laboratory experiments on rocks under compression show that fracture occurs when a critical combination of the absolute value of shear stress and the normal stress is exceeded. This relation, known as the *Coulomb-Mohr failure criterion*, can be stated as

$$|\tau| = \tau_o - n\sigma, \quad (5)$$

where  $\tau_o$  and  $n$  are properties of the material known as the *cohesive strength* and *coefficient of internal friction*. (The minus sign reflects the convention that compressional stresses are negative.) The failure criterion plots as two lines in the  $\tau$ - $\sigma$  plane, with  $\tau$  axis intercepts  $\pm \tau_o$  and slope  $\pm n$  (Fig. 5.7-6). If the principal stresses are  $\sigma_1, \sigma_2$ , such that Mohr's circle does not intersect the failure lines, the material does not fracture. However, given the same  $\sigma_2$  but a higher  $\sigma'_1$ , Mohr's circle intersects the line, and the material breaks.

The failure lines show how much shear stress,  $\tau$ , can be applied to a surface subject to a normal stress  $\sigma$  before failure occurs. The cohesive strength is the minimum (absolute value) shear stress for failure. The coefficient of internal friction indicates the additional shear stress sustainable as the normal stress increases. Thus, deeper in the crust, where the pressure and hence normal stress are higher, rocks are stronger, and higher shear stress is required to break them.

The failure lines and Mohr's circle show on which plane failure occurs for a given stress state. To find  $\theta$ , the angle between the plane's normal and the maximum compressive stress ( $\sigma_1$ ) direction, we write the failure lines as

$$|\tau| = \tau_o - \sigma \tan \phi, \quad (6)$$

<sup>1</sup> Following the seismological convention of compressive stresses being negative, Mohr's circle is shown for  $\sigma < 0$ . The opposite convention is often used in rock mechanics, e.g. Figs. 5.7-3 and 5.7-10.

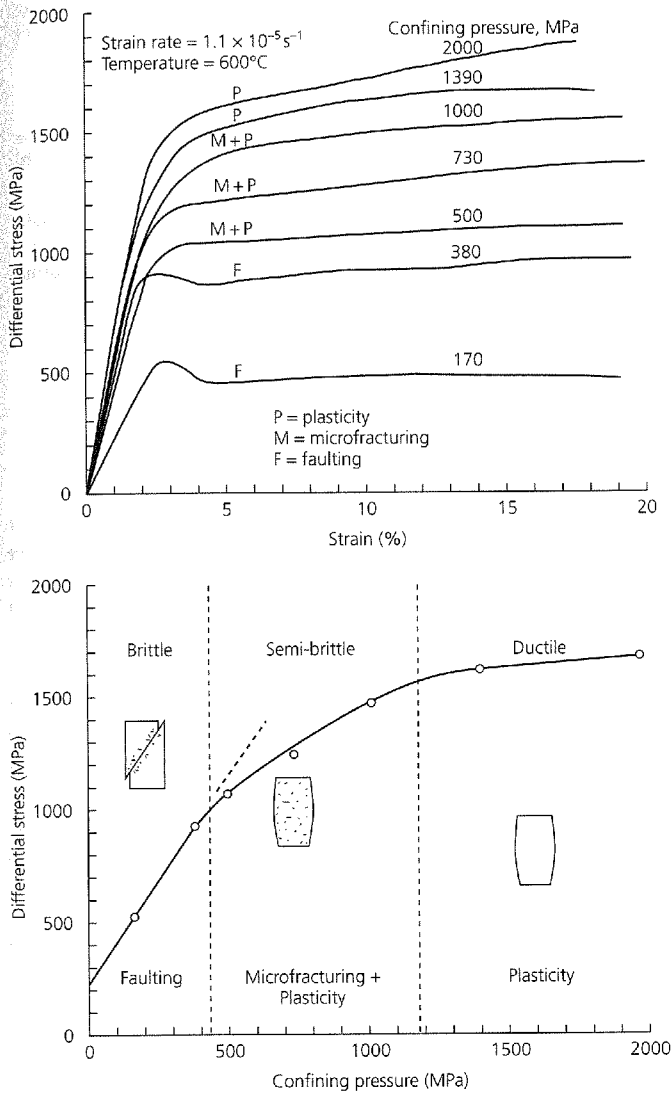


Fig. 5.7-3 Results of an experiment in which rocks are subjected to a compressive stress  $\sigma_1$  greater than the confining pressure  $\sigma_3$ . *Top*: Differential stress ( $\sigma_1 - \sigma_3$ ) versus strain (compare to Figs 5.7-1 and 2) curves for various confining pressures. *Bottom*: Ultimate strength ( $\sigma_1 - \sigma_3$  at 10% strain rate, from top) for various confining pressures. For low (< 400 MPa) confining pressures, the material fractures, and its strength increases with pressure. For higher pressures, the material is ductile, and its strength increases only slowly with pressure. A semi-brittle transition regime, in which both microfractures and crystal plasticity occur, separates the brittle and ductile regimes. (Kirby, 1980. *J. Geophys. Res.*, 85, 6353–63, copyright by the American Geophysical Union.)

where  $n = \tan \phi$ , and  $\phi$ , the *angle of internal friction*, is formed by extending the failure line to the  $\sigma$  axis (Fig. 5.7-7). Fracture occurs at point F, where the failure line is tangent to Mohr's circle. Considering the right triangle AFB, we see that

$$\phi = 2\theta - 90^\circ, \text{ so } \theta = \phi/2 + 45^\circ. \quad (7)$$

For example, in introducing the relation between fault plane solutions and crustal stresses in Section 2.3.5, we made the simplest assumption that fracture occurs at  $45^\circ$  to the principal stress axes, corresponding to the case  $\phi = 0^\circ$ ,  $n = 0$ ,  $\theta = 45^\circ$ . Physically, this means that the normal stress has no effect on the strength of the rock. However, rocks typically have  $n$  about 1, so  $\phi = 45^\circ$ ,  $\theta = 67.5^\circ$ , and the fault plane is closer ( $22.5^\circ$ ) to the maximum compression ( $\sigma_1$ ) direction (Fig. 5.7-8). This idea is important when using P and T axes of focal mechanisms to characterize stress directions.

Figure 5.7-7 also shows how to find the stresses when fracture occurs. Consider the point T on the failure line such that  $T\sigma_2$  is perpendicular to the  $\sigma$  axis. Because the angle  $AT\sigma_2$  is  $\theta$  (triangles AFT and  $A\sigma_2T$  are congruent),

$$\overline{T\sigma_2} = A\sigma_2 \cot \theta, \quad (8)$$

or, since  $\overline{A\sigma_2} = (\sigma_2 - \sigma_1)/2$ ,

$$\overline{T\sigma_2} = \frac{(\sigma_2 - \sigma_1)}{2} \cot \theta. \quad (9)$$

Similarly,

$$\overline{T\sigma_2} = \tau_o - \sigma_2 \tan \phi \quad (10)$$

(the minus sign is because  $\sigma_2$  is negative), so

$$\frac{(\sigma_2 - \sigma_1)}{2} \cot \theta = \tau_o - \sigma_2 \tan \phi. \quad (11)$$

This relation can be written in terms of the angle of the fracture plane, using Eqn 7 and trigonometric identities,

$$\tan \phi = -\cot 2\theta = \frac{-1}{\tan 2\theta} = \frac{\tan^2 \theta - 1}{2 \tan \theta}, \quad (12)$$

yielding

$$\sigma_1 = -2\tau_o \tan \theta + \sigma_2 \tan^2 \theta. \quad (13)$$

We will use this relation between the stresses when fracture occurs to estimate the maximum stresses in the crust.

Similar analyses show when the shear stress is high enough to overcome friction and cause sliding on a previously existing fault. The results are similar to those for a new fracture in unbroken rock, except that at low stress levels the preexisting fault has no cohesive strength. Thus slip on the fault occurs when  $|\tau| = -\mu\sigma$ , where  $\mu$  is the *coefficient of sliding friction*, which can be expressed by an *angle of sliding friction*

$$\tan \alpha = \mu. \quad (14)$$

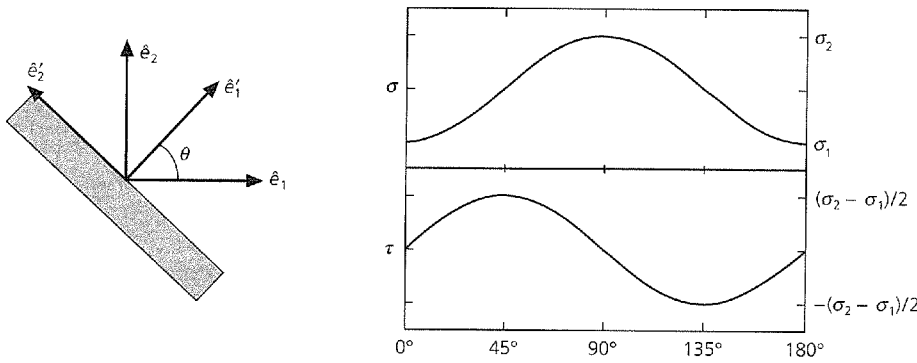


Fig. 5.7-4 Left: Geometry of a plane with normal  $\hat{e}'_1$ , oriented at an angle  $\theta$  from  $\hat{e}_1$ , the direction of the maximum compressive stress  $\sigma_1$ . Right: Normal stress,  $\sigma$ , and shear stress,  $\tau$ , as functions of the angle  $\theta$ .

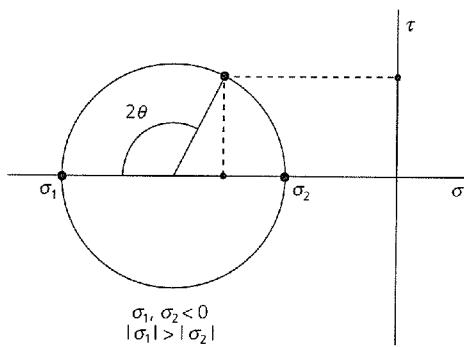


Fig. 5.7-5 Mohr's circle: Given a state of stress described by principal stresses  $\sigma_1$  and  $\sigma_2$ , the normal stress,  $\sigma$ , and the shear stress,  $\tau$ , for planes of all orientations lie on a circle with radius  $(\sigma_2 - \sigma_1)/2$ . The point on the circle with angle  $2\theta$ , measured clockwise from the  $-\sigma$  axis, gives  $\sigma$  and  $\tau$  on a plane whose normal is at an angle  $\theta$  from the direction of  $\sigma_1$ .

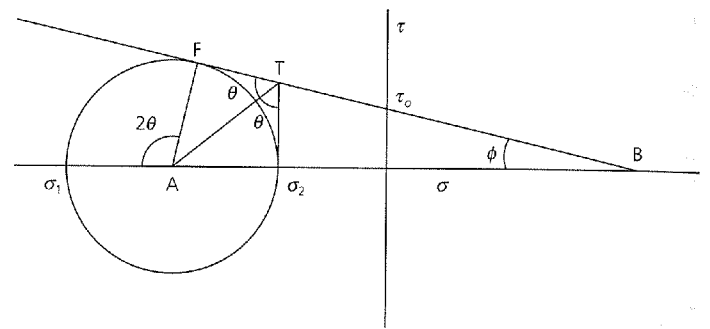


Fig. 5.7-7 Fracture occurs at point F, where a material's failure line, characterized by its cohesive strength,  $\tau_0$ , and angle of internal friction,  $\phi$ , is tangent to Mohr's circle. Hence  $\theta$  is the angle of the plane on which fracture occurs, and F gives the stresses at fracture. Point A is the center of Mohr's circle, B is where the failure line intersects the  $\sigma$  axis, and  $T\sigma_2$  is perpendicular to the  $\sigma$  axis. For simplicity, only the upper failure line for  $\tau > 0$  is shown in this and subsequent figures.

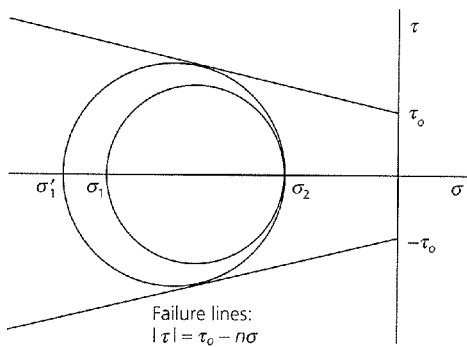


Fig. 5.7-6 The Coulomb-Mohr failure criterion assumes that a material fractures when Mohr's circle intersects the failure line.

Figure 5.7-9 shows the Mohr's circle representation of a rock with preexisting faults. In addition to the failure line, there is a frictional sliding line corresponding to

$$\tau = -\mu\sigma = -\sigma \tan \alpha. \quad (15)$$

Because the sliding line starts at the origin, it is initially below the failure line. Assume that the stresses are large enough that Mohr's circle touches the failure line at the point yielding frac-

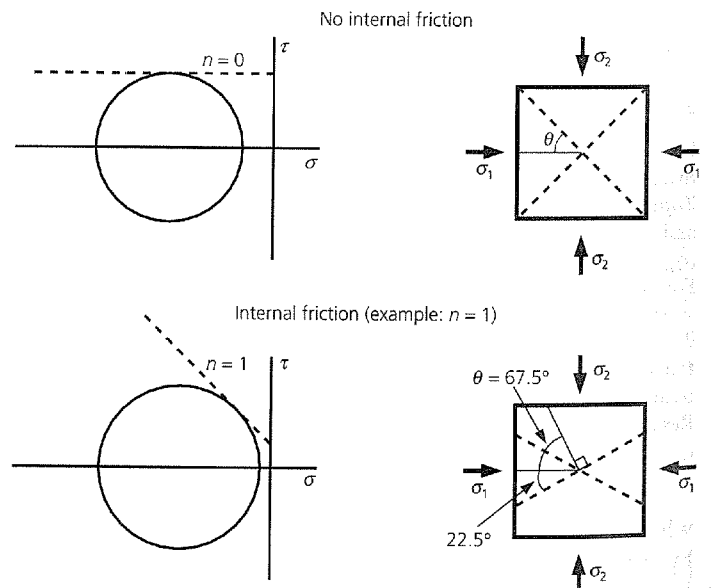


Fig. 5.7-8 With no internal friction, fracture occurs at an angle of  $45^\circ$ . For  $n = 1$ , the fracture angle is  $67.5^\circ$ , and the fault plane is closer ( $22.5^\circ$ ) to the maximum compression ( $\sigma_1$ ) direction.



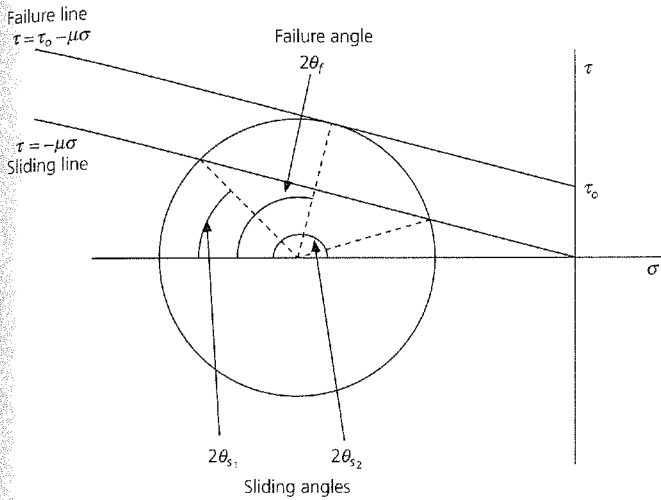


Fig. 5.7-9 Mohr's circle for sliding on a rock's preexisting faults. A new fracture would form at an angle  $\theta_f$ , given by the failure line. However, slip will occur on a preexisting fault if there are any with angles between  $\theta_{s1}$  and  $\theta_{s2}$ , given by the intersection of the circle with the frictional sliding line.

ture on a plane corresponding to an angle  $\theta_f$ . Similarly, the frictional sliding line intercepts the circle at two points, corresponding to angles  $\theta_{s1}$  and  $\theta_{s2}$ . Thus the rock can fail in several ways. If there are preexisting faults with angles  $\theta_{s1}$  or  $\theta_{s2}$ , slip on these faults may occur. Alternatively, a new fracture may form on the plane given by  $\theta_f$ . However, because this fracture occurs at higher shear stress than is needed for frictional sliding on the preexisting faults, sliding is favored over the formation of a new fracture. Thus, if the stress has gradually risen to this level, sliding on preexisting faults would probably have prevented a new fracture from forming.

This effect can have seismological consequences. The simplest way to use focal mechanisms to infer stress orientations is to assume that the earthquakes occurred on newly formed faults. However, if the rock had been initially faulted, the earthquakes may have occurred on preexisting faults. In the representation of Fig. 5.7-9, if faults exist with normals oriented between  $\theta_{s1}$  and  $\theta_{s2}$  to the maximum compressive stress, slip on these faults will occur rather than the formation of a new fracture. Thus the inferred stress direction will be somewhat inaccurate. For example, the thrust focal mechanisms along the Himalayan front (Fig. 5.6-6) or eastern Andean foreland thrust belt (Fig. 5.6-10) have fault planes that rotate as the trend of the mountains changes, suggesting that the fault planes are controlled by the existing structures, so the P axes only partially reflect the stress field. A similar pattern appears for T axes along the East African rift (Fig. 5.6-2). In general, stress axes inferred from many fault plane solutions in an area seem relatively coherent (Fig. 5.6-19). Thus we assume that the crust contains preexisting faults of all orientations, so the average stress orientation inferred from the focal mechanisms is not seriously biased.

At this point, it is worth noting other complexities. Both the failure and sliding curves may be more complicated than straight lines. These curves, known as Mohr envelopes, can be derived from experiments at various values of stress. Additional complexity comes from the fact that water and other fluids are often present in rocks, especially in the upper crust. The fluid pressure, known as the *pore pressure*, reduces the effect of the normal stress and allows sliding to take place at lower shear stresses. This effect is modeled by replacing the normal stress  $\sigma$  with  $\bar{\sigma} = \sigma - P_f$ , known as the *effective normal stress*, where  $P_f$  is the pore fluid pressure.<sup>2</sup> Because the pore pressure is defined as negative, the effective normal stress is reduced (less compressive). Similarly, effective principal stresses taking into account pore pressure,

$$\bar{\sigma}_1 = \sigma_1 - P_f \quad \text{and} \quad \bar{\sigma}_2 = \sigma_2 - P_f, \quad (16)$$

are used in the fracture theory.

The relations we have discussed can be used to estimate the maximum stresses that the crust can support. Laboratory experiments (Fig. 5.7-10) for sliding on existing faults in a variety of rock types find relations sometimes called *Byerlee's law*:

$$\begin{aligned} \tau &\approx -0.85\bar{\sigma}, \quad |\bar{\sigma}| < 200 \text{ MPa} \\ \tau &\approx 50 - 0.6\bar{\sigma}, \quad |\bar{\sigma}| > 200 \text{ MPa}. \end{aligned} \quad (17)$$

These relations, written in terms of the normal and shear stresses on a fault, can be used to infer the principal stress as a function of depth. To do so, we write the minimum compressive stress as  $\sigma_3$ , because we are in three dimensions. We assume that the crust contains faults of all orientations, and that the stresses cannot exceed the point where Mohr's circle is tangent to the frictional sliding line, or else sliding will occur (Fig. 5.7-11). At shallow depths where  $|\bar{\sigma}| < 200$  MPa, Eqn 17 shows that  $\tau_0 = 0$ . Thus Eqn 13, the relation between the stresses when fracture occurs, yields

$$\bar{\sigma}_1 = \bar{\sigma}_3 \tan^2 \theta_s. \quad (18)$$

Using Eqn 7 for the case of frictional sliding,

$$\theta_s = \alpha/2 + 45^\circ, \quad (19)$$

and the values in Eqn 17 give

$$\mu = \tan \alpha = 0.85, \quad \alpha \approx 41^\circ, \quad \theta_s \approx 66^\circ, \quad \tan^2 66^\circ \approx 5, \quad (20)$$

so the stresses are related by

$$\bar{\sigma}_1 \approx 5\bar{\sigma}_3. \quad (21)$$

At greater depths, where  $|\bar{\sigma}| > 200$  MPa,  $\alpha \approx 31^\circ$  and  $\theta_s = 60.5^\circ$ , so the stresses are related by

$$\bar{\sigma}_1 \approx -175 + 3.1\bar{\sigma}_3. \quad (22)$$

<sup>2</sup> The role of pore pressure in making sliding easier can be seen by trying to slide an object across a dry table and then wetting the table.



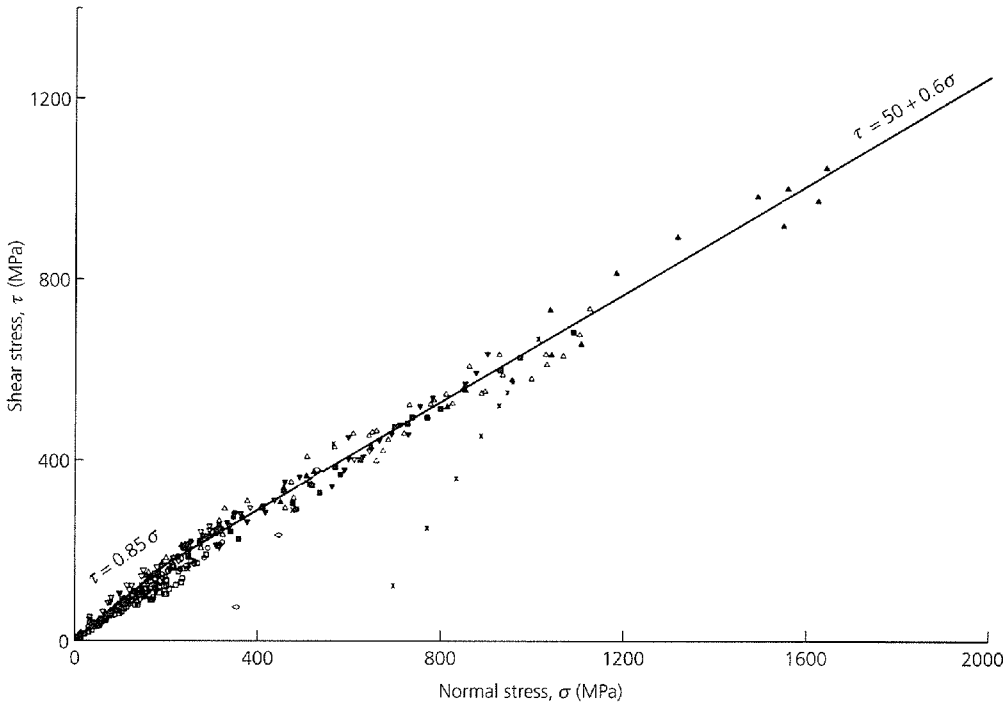


Fig. 5.7-10 Shear stress versus normal stress for frictional sliding, compiled for various rock types. Compressive stress is positive. (Byerlee, 1978. *Pure Appl. Geophys.*, 116, 615–26, reproduced with the permission of Birkhauser.)

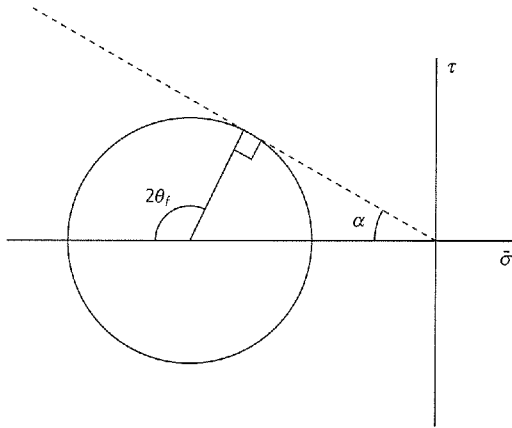


Fig. 5.7-11 Mohr's circle and sliding line for  $|\bar{\sigma}| < 200$  MPa. If the lithosphere contains fractures in all directions, the stresses cannot exceed those at the point where Mohr's circle is tangent to the sliding line, because sliding would occur.

We assume that one principal stress,  $\sigma_1$  or  $\sigma_3$ , is the vertical stress due to the lithostatic pressure as a function of depth ( $z$ ),

$$\sigma_V = -\rho_f g z. \quad (23)$$

The other principal stress, which must be horizontal, is denoted  $\sigma_H$ . The pore pressure  $P_f(z)$  is unknown. One common assumption is that the rock is dry, so  $P_f(z) = 0$ . Another is that the pore pressure is *hydrostatic*, which is equivalent to assuming that pores are connected up to the surface, so

$$P_f(z) = -\rho_f g z, \quad (24)$$

where  $\rho_f$  is the density of the fluid, which is usually water, with  $\rho_f = 1 \text{ g/cm}^3$ . Alternatively, the pore pressure can be assumed to be a fixed fraction of the lithostatic pressure (Section 2.3.6).

We now can find the *strength* of the crust, defined by the maximum difference between the horizontal and vertical stresses that the rock can support. At shallow depths where  $|\bar{\sigma}| < 200$  MPa, Eqn 21 shows that  $\bar{\sigma}_1 = 5\bar{\sigma}_3$ . There are two possibilities, depending on whether the vertical stress is the most ( $\bar{\sigma}_1$ ) or least ( $\bar{\sigma}_3$ ) compressive. If the vertical stress is the most compressive,

$$\begin{aligned} \sigma_V &= \sigma_1, & \bar{\sigma}_1 &= \sigma_V - P_f = -\rho_f g z - P_f(z) \\ \sigma_H &= \sigma_3, & \bar{\sigma}_3 &= \bar{\sigma}_1/5 = -(\rho_f g z + P_f(z))/5. \end{aligned} \quad (25)$$

Alternatively, if the vertical stress is the least compressive,

$$\begin{aligned} \sigma_V &= \sigma_3, & \bar{\sigma}_3 &= \sigma_V - P_f = -\rho_f g z - P_f(z) \\ \sigma_H &= \sigma_1, & \bar{\sigma}_1 &= 5\bar{\sigma}_3 = -5(\rho_f g z + P_f(z)). \end{aligned} \quad (26)$$

In the first case,

$$\sigma_H - \sigma_V = \sigma_3 - \sigma_1 = 0.8(\rho_f g z + P_f(z)), \quad (27)$$

corresponds to an extensional (positive) stress. In the second,

$$\sigma_H - \sigma_V = \sigma_1 - \sigma_3 = -4(\rho_f g z + P_f(z)) \quad (28)$$

corresponds to a compressive (negative) stress that is much greater in absolute value. Thus, at any depth, the crust can

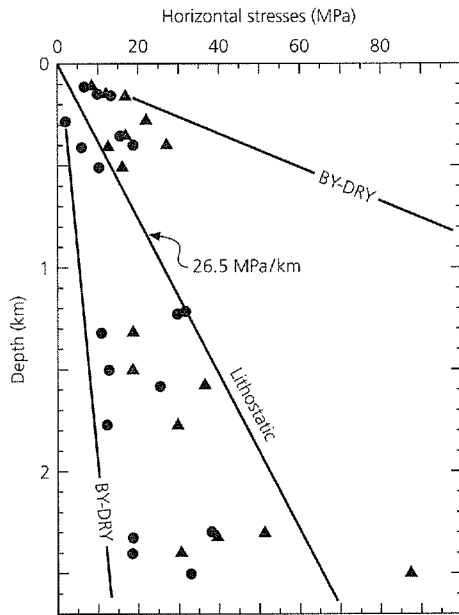


Fig. 5.7-12 Horizontal stresses measured in southern Africa. Dots are for horizontal stresses being the least compressive ( $\sigma_3$ ), and triangles are for horizontal stresses being the most compressive ( $\sigma_1$ ). The lithostatic stress gradient (26.5 MPa/km) is shown, along with Byerlee's law (BY) for zero pore pressure (DRY). The stronger line is for compression, and the weaker one is for extension. The observed stresses are within the maximum and minimum BY-DRY lines. (Brace and Kohlstedt, 1980. *J. Geophys. Res.*, 85, 6248–52, copyright by the American Geophysical Union.)

support greater compressive deviatoric stress than extensional deviatoric stress (Fig. 5.7-12).

### 5.7.3 Ductile flow

When rocks behave brittly, their behavior is not time-dependent; they either strain elastically or fail. By contrast, the deformation of ductile rock depends on time. A common model for the time-dependent behavior is a *Maxwell viscoelastic material*, which behaves like an elastic solid on short time scales and like a viscous fluid on long time scales. This model can describe the mantle because seismic waves propagate as though the mantle were solid, whereas postglacial rebound and mantle convection occur as though the mantle were fluid.

To see this difference, consider two types of deformation in one dimension. For an elastic solid subjected to elastic strain  $e_E = e_{11}$ ,

$$\sigma = E e_E, \quad (29)$$

where  $E$  is Young's modulus, and  $\sigma$  is  $\sigma_{11}$ . The simplest viscous fluid obeys

$$\sigma = 2\eta \frac{de_F}{dt}, \quad (30)$$

where  $\eta$  is the viscosity, and  $e_F$  is the fluid portion of the strain. This equation defines the viscosity, the property that measures a fluid's resistance to shear.<sup>3</sup>

We often think of an elastic material as a spring, which exerts a force proportional to distance. Thus stress and strain are proportional at any instant, and there are no time-dependent effects. By contrast, the viscous material is thought of as a dashpot, a fluid damper that exerts a force proportional to velocity. Hence the stress and strain rate are proportional, and the material's response varies with time. These effects are combined in a viscoelastic material, which can be thought of as a spring and dashpot in series (Fig. 5.7-13). The combined elastic and viscous response comes from the combined strain rate

$$\frac{de}{dt} = \frac{de_E}{dt} + \frac{de_F}{dt} = \frac{1}{E} \frac{d\sigma}{dt} + \frac{\sigma}{2\eta}. \quad (31)$$

This differential equation, the rheological law for a Maxwell substance, shows how the stress in the material evolves after a strain  $e_o$  is applied at time  $t = 0$  and then remains constant. At  $t = 0$  the derivative terms dominate, so the material behaves elastically, and has an initial stress

$$\sigma_o = E e_o. \quad (32)$$

For  $t > 0$ ,  $de/dt = 0$ , so

$$\frac{d\sigma}{dt} = -\frac{E}{2\eta} \sigma, \quad (33)$$

whose integral is

$$\sigma(t) = \sigma_o \exp[-(Et/2\eta)]. \quad (34)$$

Thus stress relaxes from its initial value as a function of time (Fig. 5.7-13). A useful parameter is the *Maxwell relaxation time*,<sup>4</sup>

$$\tau_M = \frac{2\eta}{E} \approx \frac{\eta}{\mu}, \quad (35)$$

required for the stress to decay to  $e^{-1}$  of its initial value. For times less than  $\tau_M$  the material can be considered an elastic solid, whereas for longer times it can be considered a viscous fluid.

For example, if the mantle is approximately a Poisson solid with  $\mu \approx 10^{12}$  dyn/cm<sup>2</sup> and  $\eta \approx 10^{22}$  poise, its Maxwell time is about  $10^{10}$  s or 300 years. Although the viscosity is not that well known, so estimates of the Maxwell time vary, it is clear

<sup>3</sup> In familiar terms, viscosity measures how "gooey" a fluid is. Maple syrup is somewhat more viscous than water, and the earth's mantle is about  $10^{24}$  times more viscous.

<sup>4</sup> Definitions of the Maxwell time vary, but always involve the ratio of the viscosity to an elastic constant.

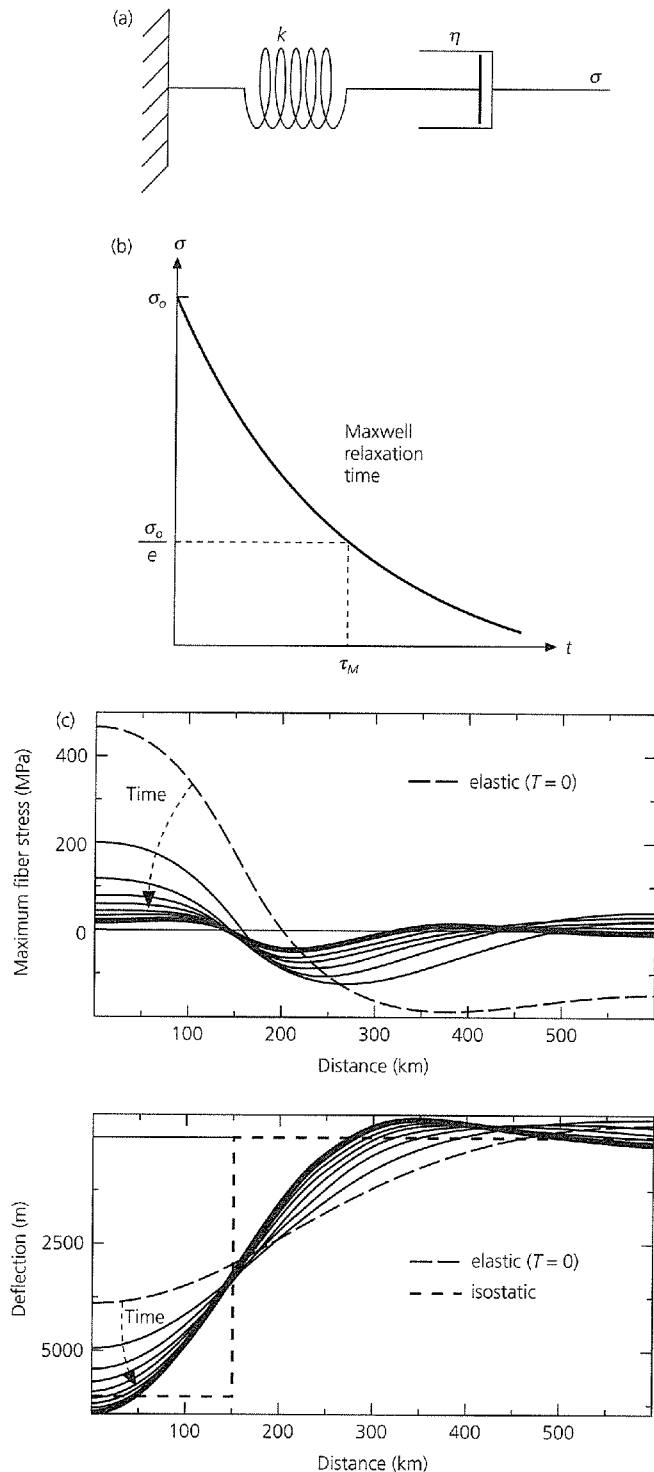


Fig. 5.7-13 (a) Model of a viscoelastic material as an elastic spring and viscous dashpot in series. (b) Stress response of a viscoelastic material to an applied strain. The Maxwell relaxation time,  $\tau_M$ , is the time the stress takes to decay to  $e^{-1}$  of its initial value. (c) Evolution of the deflection and bending stress produced by a sediment load on a viscoelastic earth. At first the earth responds elastically, as shown by the long-dashed line, but with time it flows, so the deflection beneath the load deepens and the stresses relax. (Stein *et al.*, 1989, with kind permission from Kluwer Academic Publishers.)

that we can treat the mantle as a solid for seismological purposes and as a fluid in tectonic modeling. If we model the mantle as viscoelastic, then a load applied on the surface has an effect that varies with time. Figure 5.7-13c shows the effect of a 150 km-wide sediment load, as might be expected on a passive continental margin. Initially, the earth responds elastically, causing large flexural bending stresses. With time, the mantle flows, so the deflection beneath the load deepens and the stresses relax. In the time limit, the stress goes to zero, and the deflection approaches the isostatic solution, because isostasy amounts to assuming that the lithosphere has no strength. Stress relaxation may explain why large earthquakes are rare at continental margins, except where glacial loads have been recently removed (Fig. 5.6-20). Although the large sediment loads should produce stresses much greater than other sources of intraplate stress, including the smaller and less dense ice loads, the stresses produced by sediment loading early in the margin's history may have relaxed.

Laboratory experiments indicate that the rheology of minerals in ductile flow can be described by

$$\frac{de}{dt} = \dot{\epsilon} = f(\sigma) A \exp [-(E^* + PV^*)/RT], \quad (36)$$

where  $T$  is temperature,  $R$  is the gas constant, and  $P$  is pressure.  $f(\sigma)$  is a function of the stress difference  $|\sigma_1 - \sigma_3|$ , and  $A$  is a constant. The effects of pressure and temperature are described by the *activation energy*  $E^*$  and the *activation volume*  $V^*$ . Observed values of  $f(\sigma)$  are often fit well by assuming

$$f(\sigma) = |\sigma_1 - \sigma_3|^n$$

$$\dot{\epsilon} = |\sigma_1 - \sigma_3|^n A \exp [-(E^* + PV^*)/RT]. \quad (37)$$

The rheology of such fluids is characterized by a power law. If  $n = 1$ , the material is called *Newtonian*, whereas a non-Newtonian fluid with  $n = 3$  is often used to represent the mantle. From Eqn 30 we see that for a Newtonian fluid the viscosity depends on both temperature and pressure:

$$\eta = (1/2A) \exp [(E^* + PV^*)/RT]. \quad (38)$$

Thus the viscosity decreases exponentially with temperature. This decrease is assumed to give rise to a strong lithosphere overlying a weaker asthenosphere, and the restriction of earthquakes to shallow depths.<sup>5</sup> For a non-Newtonian fluid, Eqn 30 gives the *effective viscosity*, the equivalent viscosity if the fluid were Newtonian.

We think of equations like Eqn 37 as showing the strength, or maximum stress difference  $|\sigma_1 - \sigma_3|$ , that the viscous material can support. This stress difference depends on temperature, pressure, strain rate, and rock type. The material

<sup>5</sup> Temperature-dependent viscosity is an effect familiar to automobile drivers in cold temperatures, when the engine and the transmission became noticeably sluggish.

is stronger at higher strain rates, and weakens exponentially with high temperatures. At shallow depths, the small pressure effect is often neglected, so the activation volume  $V^*$  is treated as zero. For example, a commonly used flow law for dry olivine is<sup>6</sup>

$$\begin{aligned}\dot{\epsilon} &= 7 \times 10^4 |\sigma_1 - \sigma_3|^3 \exp\left(\frac{-0.52 \text{ MJ/mol}}{RT}\right) \\ &\quad \text{for } |\sigma_1 - \sigma_3| \leq 200 \text{ MPa} \\ &= 5.7 \times 10^{11} \exp\left[\frac{-0.54 \text{ MJ/mol}}{RT} \left(1 - \frac{|\sigma_1 - \sigma_3|}{8500}\right)^2\right] \\ &\quad \text{for } |\sigma_1 - \sigma_3| > 200 \text{ MPa},\end{aligned}\quad (39)$$

where  $\dot{\epsilon}$  is in  $\text{s}^{-1}$ . Similarly, for quartz,

$$\begin{aligned}\dot{\epsilon} &= 5 \times 10^6 |\sigma_1 - \sigma_3|^3 \exp\left(\frac{-0.19 \text{ MJ/mol}}{RT}\right) \\ &\quad \text{for } |\sigma_1 - \sigma_3| < 1000 \text{ MPa}.\end{aligned}\quad (40)$$

At a given strain rate, quartz is much weaker (can sustain a smaller stress difference) than olivine. Thus the quartz-rich continental crust should be weaker than the olivine-rich oceanic crust, an effect whose tectonic consequences are discussed next.

#### 5.7.4 Strength of the lithosphere

The strength of the lithosphere as a function of depth depends upon the deformation mechanism. At shallow depths, rocks fail by either brittle fracture or frictional sliding on preexisting faults. Both processes depend in a similar way on the normal stress, with rock strength increasing with depth. However, at greater depths the ductile flow strength of rocks is less than the brittle or frictional strength, so the strength is given by the flow laws and decreases with depth as the temperatures increase. This temperature-dependent strength is the reason why the cold lithosphere forms the planet's strong outer layer.

To calculate the strength, a strain rate and a geotherm giving temperature as a function of depth are assumed. At shallow depths the strength, the maximum stress difference before frictional sliding occurs, is computed using Eqns 27 and 28. At some depth, the frictional strength exceeds the ductile strength allowed by the flow law, so for deeper depths the maximum strength is given by the flow law. Figure 5.7-14 shows a strength plot, known as a *strength envelope*, for a strain rate of  $10^{-15} \text{ s}^{-1}$  and a temperature gradient appropriate for old oceanic lithosphere or stable continental interior. In the frictional region, curves are shown for various values of  $\lambda$ , the ratio of pore pressure to lithostatic pressure. The higher

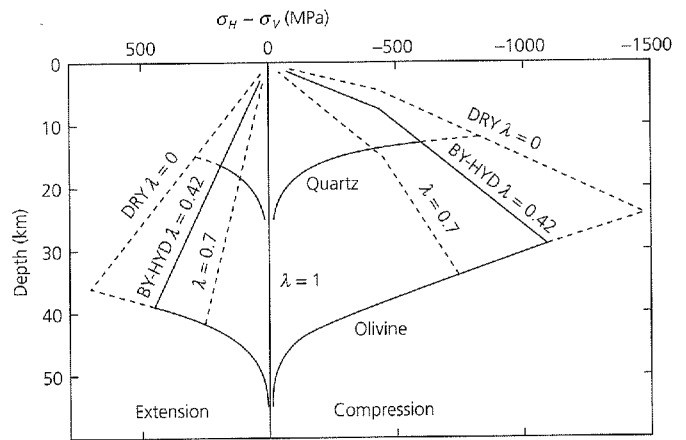


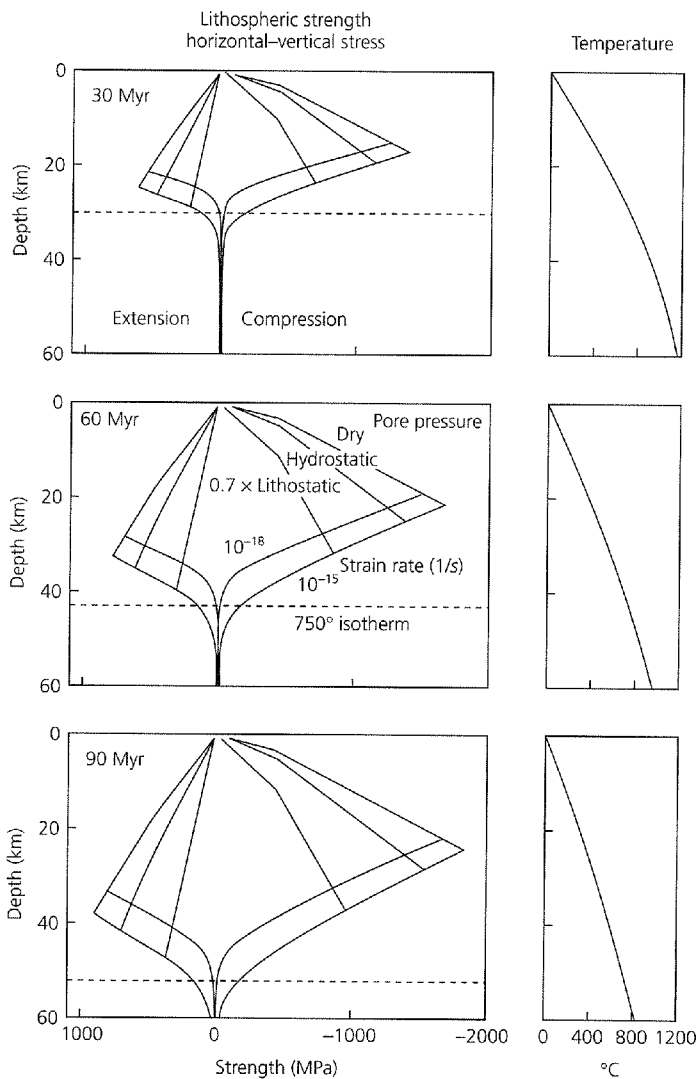
Fig. 5.7-14 Strength envelopes as a function of depth for various values of  $\lambda$ , the ratio of pore pressure to lithostatic pressure. BY-HYD lines are for Byerlee's law with hydrostatic pore pressure. At shallow depths, strength is controlled by brittle fracture; at greater depths ductile flow laws predict rapid weakening. In the ductile flow regime, quartz is weaker than olivine. In the brittle regime, the lithosphere is stronger in compression (right side) than in extension (left side). (Brace and Kohlstedt, 1980. *J. Geophys. Res.*, 85, 6248–52, copyright by the American Geophysical Union.)

pore pressures result in lower strengths. Ductile flow laws are shown for quartz and olivine, minerals often used as models for continental and oceanic rheologies. Strength increases with depth in the brittle region, due to the increasing normal stress, and then decreases with depth in the ductile region, due to increasing temperature. Hence strength is highest at the *brittle-ductile transition*. Strength decreases rapidly below this transition, so the lithosphere should have little strength at depths greater than about 25 km in the continents and 50 km in the oceans. The strength envelopes show that the lithosphere is stronger for compression than for tension in the brittle regime, but the two are symmetric in the ductile regime. Strength envelopes are often plotted using the rock mechanics convention of compression positive.

The actual distribution of strength with depth is probably more complicated, because the brittle-ductile transition occurs over a region of semi-brittle behavior that includes both brittle and plastic processes (Fig. 5.7-3). However, this simple model gives insight into various observations. In particular, we have seen that the depths of earthquakes in several tectonic environments seem to be limited by temperature. This makes sense, because for a given strain rate and rheology the exponential dependence on temperature would make a limiting strength for seismicity approximate a limiting temperature.

To see this, consider Fig. 5.7-15, which shows that as oceanic lithosphere ages and cools, the predicted strong region deepens. This result seems plausible because earthquake depths, seismic velocities, and effective elastic thicknesses imply that the strong upper part of the lithosphere thickens with age (Fig. 5.3-9). The strength envelopes are thus consistent with the observation that the maximum depth of earthquakes within

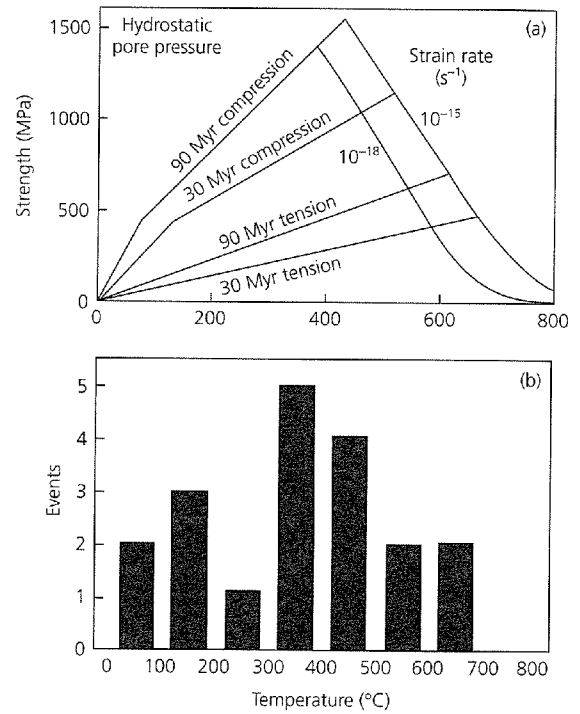
<sup>6</sup> Brace and Kohlstedt (1980).



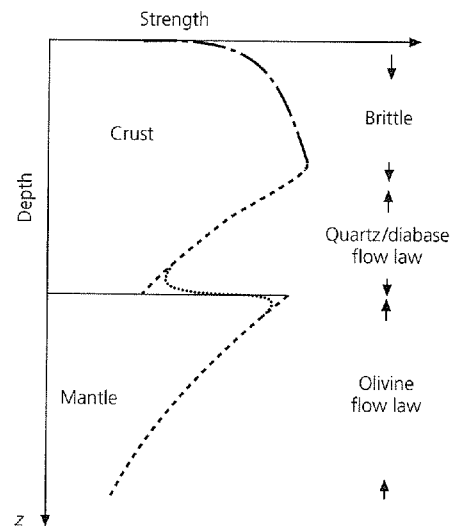
**Fig. 5.7-15** Strength envelopes showing maximum stress difference (strength) as a function of depth for an olivine rheology, for geotherms (right) corresponding to cooling oceanic lithosphere of different ages. Strength in the brittle regime is reduced by higher pore pressure; strength in the ductile regime is reduced by lower strain rate. The depth range in which the material is strong enough for faulting increases with age. (Wiens and Stein, 1983. *J. Geophys. Res.*, 88, 6455–68, copyright by the American Geophysical Union.)

the oceanic lithosphere is approximately bounded by the 750°C isotherm (Fig. 5.7-16). These envelopes are drawn for strain rates of  $10^{-15}$  and  $10^{-18} \text{ s}^{-1}$ , appropriate for slow deformation within plates. By contrast, a seismic wave with a period of 1 s, a wavelength of 10 km, and a displacement of  $10^{-6} \text{ m}$  corresponds to a strain rate of  $10^{-10} \text{ s}^{-1}$ . The successively greater effective elastic thicknesses, depth of the deepest earthquakes, and depth of the low-velocity zone are thus consistent with strength increasing with strain rate.

The strength envelopes give insight into differences between continental and oceanic lithospheres (Fig. 5.7-17). First, quartz is weaker than olivine at a given temperature (Fig. 5.7-14),



**Fig. 5.7-16** Plots of strength and seismicity versus temperature. The strength envelopes explain the observation that intraplate oceanic seismicity occurs only above the 750°C isotherm. (Wiens and Stein, 1985. *Tectonophysics*, 116, 143–62, with permission from Elsevier Science.)



**Fig. 5.7-17** Schematic strength envelope for continents. Below the ductile lower crust may be a stronger zone in the olivine-rich mantle. (Chen and Molnar, 1983. *J. Geophys. Res.*, 88, 4183–4214, copyright by the American Geophysical Union.)

consistent with the fact that the limiting temperature for continental seismicity is lower than for oceanic earthquakes (Fig. 5.7-18). Second, the strength profiles differ. The strength of oceanic lithosphere increases with depth and then decreases. However, in continental lithosphere we expect such a profile in

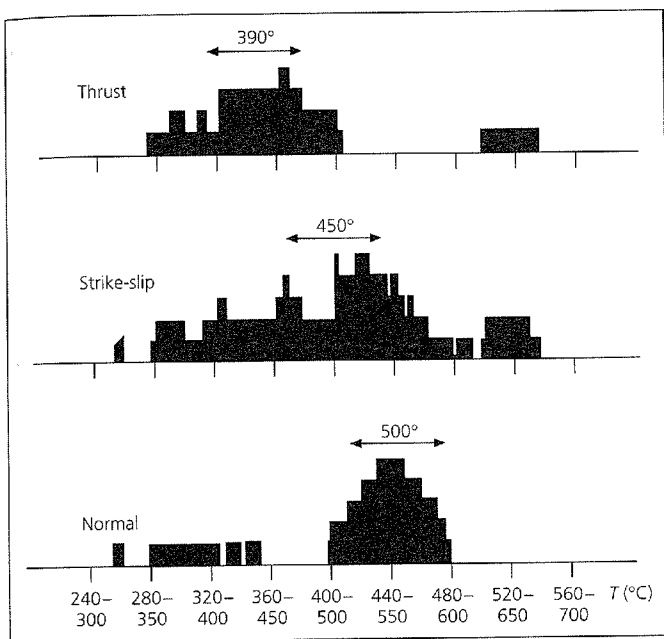


Fig. 5.7-18 Limiting temperatures for continental seismicity. These temperatures are much lower than those for oceanic lithosphere, since the quartz rheology in continents is much weaker than olivine. (Courtesy of J. Strehlau and R. Meissner.)

the quartz-rich crust, but also a second, deeper zone of strength below the Moho, due to the olivine rheology. This “jelly sandwich” profile including a weak zone may be part of the reason why continents deform differently than oceanic lithosphere. For example, some continental mountain building (Fig. 5.6-6) may involve crustal thickening in which slices of upper crust, which are too buoyant to subduct, are instead thrust atop one another. The weaker lower crust may also contribute in other ways to the general phenomenon that continental plate boundaries are broader and more complex than their oceanic counterparts (Fig. 5.2-4).

### 5.7.5 Earthquakes and rock friction

It is natural to assume that earthquakes occur when tectonic stress exceeds the rock strength, so a new fault forms or an existing one slips. Thus steady motion across a plate boundary seems likely to give rise to a cycle of successive earthquakes at regular intervals, with the same slip and stress drop (Fig. 5.7-19). However, we have seen that the earthquake process is more complicated. The time between earthquakes on plate boundaries varies (Fig. 1.2-15), although the plate motion causing the earthquakes is steady. Earthquakes sometimes rupture along the same segments of a boundary as in earlier earthquakes, and other times along a different set (Fig. 5.4-27). Moreover, many large earthquakes show a complicated rupture pattern, with some parts of the fault releasing more seismic energy than others (Fig. 4.5-10). Attempts to understand these

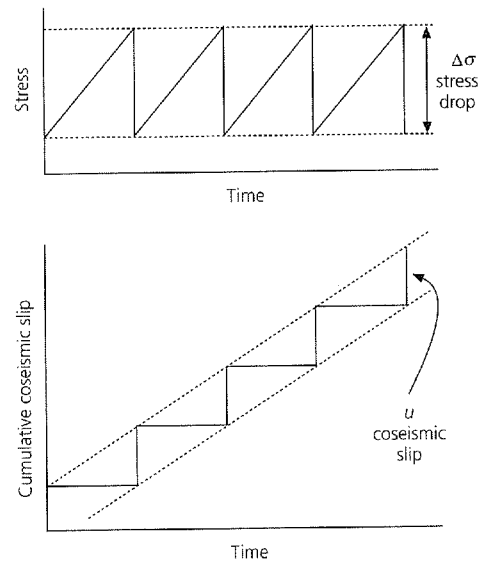


Fig. 5.7-19 Stress and slip history for an idealized earthquake cycle on a plate boundary, in which all earthquakes have the same stress drop and coseismic slip. (Shimazaki and Nakata, 1980. *Geophys. Res. Lett.*, 7, 279–82, copyright by the American Geophysical Union.)

complexities often combine two basic themes. Some of the complexity may be due to intrinsic randomness of the failure process, such that some small ruptures cascade into large earthquakes, whereas others do not (Section 1.2.6). Other aspects of the complexity may be due to features of rock friction.

Interesting insight emerges from considering an experiment in which stress is applied until a rock breaks. When the fault forms, some of the stress is released, and then motion stops. If stress is reapplied, another stress drop and motion occur once the stress reaches a certain level. So long as stress is reapplied, this pattern of jerky sliding and stress release continues (Fig. 5.7-20).

This pattern, called *stick-slip*, looks like a laboratory version of what happens in a sequence of earthquakes on a fault. By this analogy, the stress drop in an earthquake relieves only part of the total tectonic stress, and as the fault continues to be loaded by tectonic stress, occasional earthquakes occur. The analogy is strengthened by the fact that at higher temperatures (about 300° for granite), stick-slip does not occur (Fig. 5.7-20). Instead, *stable sliding* occurs on the fault, much as earthquakes do not occur at depths where the temperature exceeds a certain value. Thus, understanding stick-slip in the laboratory seems likely to give insight into the earthquake process.

Stick-slip results from a familiar phenomenon: it is harder to start an object sliding against friction than to keep it going once it is sliding. This is because the *static friction* stopping the object from sliding exceeds the *dynamic friction* that opposes motion once sliding starts.<sup>7</sup> To understand how this difference

<sup>7</sup> This effect is the basis of cross-country skiing, where loading one ski makes it grip the snow, while unloading the other lets it glide.

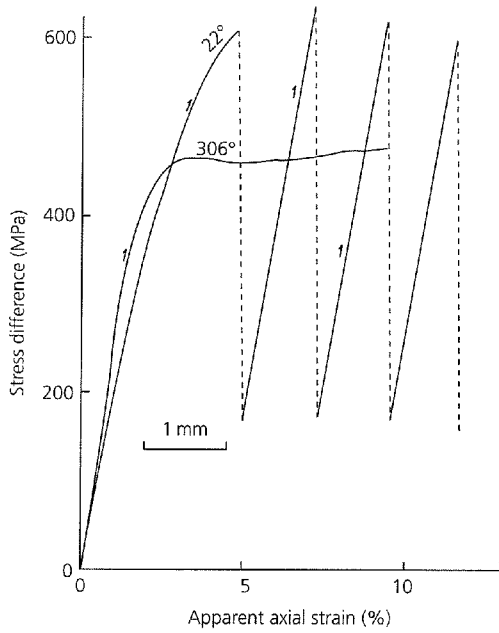


Fig. 5.7-20 Force versus slip history for a rock sample. At low temperature, so long as stress is reapplied, a stick-slip pattern of jerky sliding and stress release continues. By contrast, stable sliding occurs at high temperature. (Brace and Byerlee, 1970. *Science*, 168, 1573–5, copyright 1970 American Association for the Advancement of Science.)

causes stick-slip, and get insight into stick-slip as a model for earthquakes, consider the experiment in Fig. 5.7-21. It turns out that if an object is pulled across a table with a rubber band, jerky stick-slip motion occurs.<sup>8</sup> Thus a steady load, combined with the difference in static and dynamic friction, causes an instability and a sequence of discrete slip events.

We analyze this situation assuming that a block (sometimes called a slider) is loaded by a spring that applies a force  $f$  proportional to the spring constant (stiffness)  $k$  and the spring extension. If the loading results from the spring's far end moving at a velocity  $v$ , the spring force is

$$f = k(\zeta + vt - u), \quad (41)$$

where  $u$  is the distance the block slipped, and  $\zeta$  is the spring extension when sliding starts at  $t = 0$ . This motion is opposed by a frictional force  $|\tau| = -\mu\sigma$  equal to the product of  $\sigma$ , the compressive (negative) normal stress due to the block's weight, and the friction coefficient,  $\mu$ . By Newton's second law that force equals mass times acceleration,

$$m \frac{d^2 u}{dt^2} = f - \tau = k(\zeta + vt - u) + \mu\sigma. \quad (42)$$

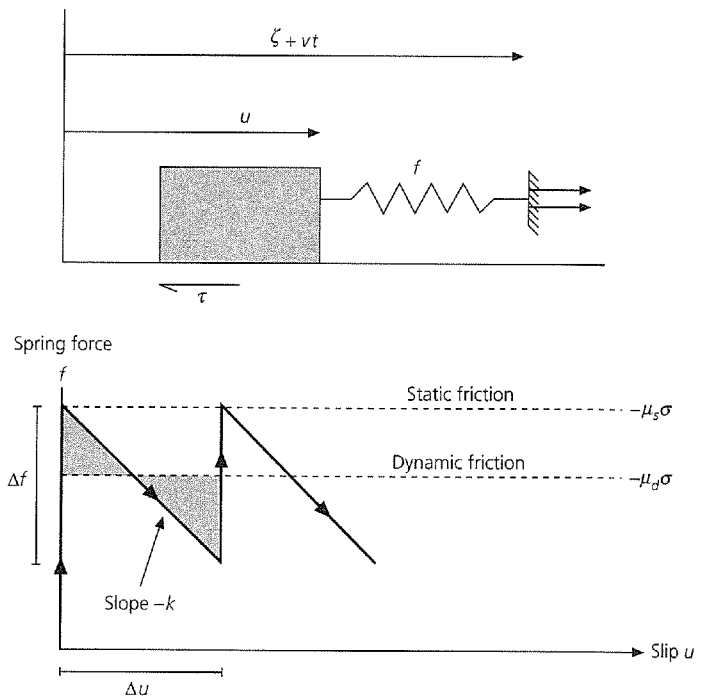


Fig. 5.7-21 A simple spring and slider block analog for stick-slip as a model for earthquakes. The slider is loaded by force  $f$  due to the spring end moving at velocity  $v$ . Before sliding, the block is retarded by a static friction force  $\tau = -\mu_s \sigma$ , but once sliding starts, the friction force decreases to  $-\mu_d \sigma$ . A series of slip events occur, each with slip  $\Delta u$  and force change (stress drop)  $\Delta f$ .

However, the block starts sliding only once the spring force exceeds the frictional force, so just before sliding starts at  $t = 0$ ,

$$0 = k\zeta + \mu_s \sigma, \quad (43)$$

where  $\mu_s$  is the static friction coefficient. For simplicity, assume that at the instant sliding starts, the friction drops to its dynamic value  $\mu_d$ , and

$$m \frac{d^2 u}{dt^2} = k(\zeta - u) + \mu_d \sigma. \quad (44)$$

Subtracting Eqn 43 from Eqn 44 gives

$$m \frac{d^2 u}{dt^2} = -ku + (\mu_d - \mu_s)\sigma = -ku + \Delta\mu\sigma, \quad (45)$$

which we can use as the equation of motion for the block's slip history  $u(t)$  if the loading rate  $v$  is slow enough to ignore during the slip event.

A solution to Eqn 45, with initial conditions  $u(0) = 0$  and  $\frac{du(0)}{dt} = 0$ , is

<sup>8</sup> We suggest trying this experiment.



$$\begin{aligned}
 u(t) &= \frac{\Delta\mu\sigma}{k} (1 - \cos \omega t) \quad (\text{slip}), \\
 \frac{du(t)}{dt} &= \frac{\Delta\mu\sigma}{\sqrt{km}} \sin \omega t \quad (\text{velocity}), \\
 \frac{du^2(t)}{dt^2} &= \frac{\Delta\mu\sigma}{m} \cos \omega t \quad (\text{acceleration}),
 \end{aligned} \tag{46}$$

where  $\omega = \sqrt{k/m}$ . As shown, the block starts slipping because the spring force exceeds the friction force. During the slip event, the spring force decreases as the spring shortens, until it becomes less than the friction force and the block slows and eventually stops. The block stops once the shaded area above the spring force line equals that below the line, or when the work done accelerating the block equals that which decelerated it. If the spring end continues to move, loading continues until the spring force again equals the static friction force and another slip event occurs.

It is interesting to think of analogies between this model of slip events and earthquakes. The slip event's duration  $t_D$ , analogous to an earthquake rise time (Section 4.3.2), satisfies

$$\frac{du(t_D)}{dt} = 0, \quad t_D = \frac{\pi}{\omega} = \pi\sqrt{m/k}. \tag{47}$$

The total slip during the event is

$$\Delta u = u(t_D) = 2\Delta\mu\sigma/k, \tag{48}$$

and the drop in the spring force, which is analogous to an earthquake stress drop (Section 4.6.3), is

$$\Delta f = 2\Delta\mu\sigma. \tag{49}$$

Thus the rise time depends on the spring constant, but not on the difference between static and dynamic friction. However, the total slip and stress drop depend upon the friction difference. None of these depend upon the loading rate, which is analogous to the rate of plate motion causing earthquakes on a plate boundary. But the loading rate determines the time between successive slip events. Thus, in the plate boundary analogy, the time between large earthquakes depends on the plate motion rate, but their slip and stress drop depend on the frictional properties of the fault and the normal stress. Hence faster-slipping boundaries would have more frequent large earthquakes, but the slip and stress drop in them would not be greater than on a slower boundary with similar frictional properties and normal stress.

Laboratory experiments show that the difference between static and dynamic friction is more complicated than the constant values assumed in this simple model. We can think of the lower dynamic friction as showing either velocity weakening, decreasing as the object moves faster, or slip weakening, decreasing as the object moves further. Frictional models called

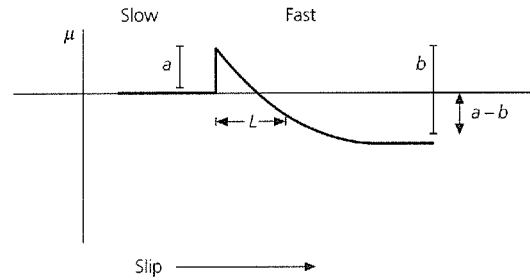


Fig. 5.7-22 Evolution of friction in a simple rate- and state-dependent model. If the slip rate increases by a factor of  $e$ , friction increases by  $a$ , and then decreases as slip progresses to a steady-state value  $a - b$ . (After Scholz, 1990. Reprinted with the permission of Cambridge University Press.)

rate- and state-dependent friction with a variable coefficient of sliding friction,  $\mu$ , are used to describe these effects. In a simple model of this sort,

$$\mu = [\mu_0 + b\psi + a \ln(v/v^*)], \tag{50}$$

where  $\mu_0$  is the coefficient of static friction. The friction depends on the slip rate  $v$ , normalized by a rate  $v^*$ , and a state variable  $\psi$  that represents the slip history

$$\frac{d\psi}{dt} = -(v/L)[\psi + \ln(v/v^*)], \tag{51}$$

where  $L$  is an experimentally determined characteristic distance. The friction also depends on  $a$  and  $b$ , which characterize the material.

Figure 5.7-22 illustrates how friction evolves. If the slip rate increases by a factor of  $e$ , the friction increases by  $a$ , and then decreases as slip progresses, reaching a new steady-state value. With time,  $\psi$  reaches a steady-state value given by Eqn 51,

$$0 = -(v/L)[\psi_{ss} + \ln(v/v^*)], \quad \psi_{ss} = -\ln(v/v^*). \tag{52}$$

The steady state friction (Eqn 50) is

$$\mu_{ss} = [\mu_0 + b\psi + a \ln(v/v^*)] = [\mu_0 + (a - b) \ln(v/v^*)], \tag{53}$$

and varies with slip rate as

$$\frac{d\mu_{ss}}{d \ln v} = (a - b), \tag{54}$$

so after the slip velocity change, the net friction change is  $(a - b)$ . If  $(a - b)$  is negative, the material shows velocity weakening, which permits earthquakes to occur by stick-slip. However, for  $(a - b)$  positive, the material shows velocity strengthening, and stable sliding is expected. Laboratory results (Fig. 5.7-20) show that  $a - b$  for granite changes sign at about  $300^\circ$ , which should be the limiting temperature for earthquakes. Thus the frictional

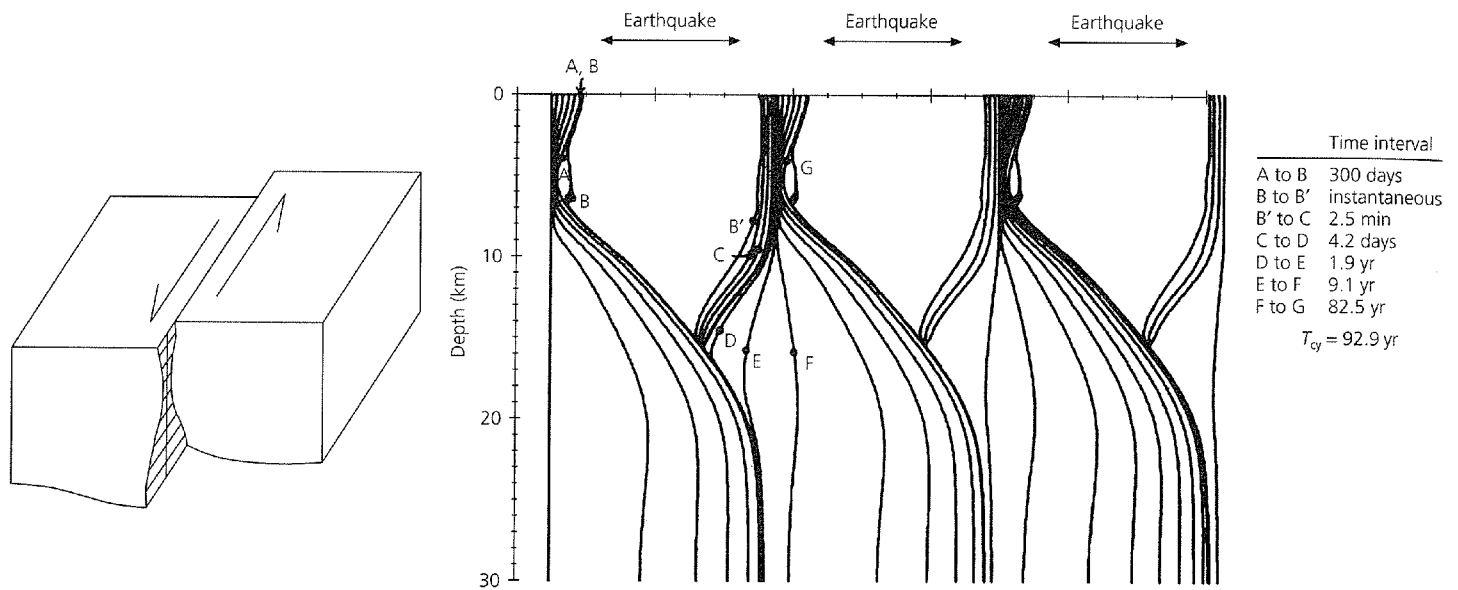


Fig. 5.7-23 Earthquake cycle for a model in which a strike-slip fault with rate- and state-dependent frictional properties is loaded by plate motion. The slip history for three cycles as a function of depth and time is shown by the lines, each of which represents a specific time. Steady motion occurs at depth, and stick-slip occurs above 11 km. (After Tse and Rice, 1986, *J. Geophys. Res.*, 91, 9452–72, copyright by the American Geophysical Union.)

model predicts a maximum depth for continental earthquakes similar to that predicted by the rock strength arguments.

These results can be used to simulate the earthquake cycle, using fault models analogous to the simple slider model (Fig. 5.7-21). Figure 5.7-23 shows the slip history as a function of depth and time for a model in which a strike-slip fault is loaded by plate motion. The fault is described by rate- and state-dependent frictional properties as a function of depth, such that stick-slip occurs above 11 km. Initially from time A to B, stable sliding occurs at depth, and a little precursory slip occurs near the surface. The earthquake causes 2.5 m of sudden slip at shallow depths, as shown by the curves for times B and B'. As a result, the faulted shallow depths "get ahead" of the material below, loading that material and causing postseismic slip from times B' to F. Once this is finished, the 93-year cycle starts again with steady stable sliding at depth.

Such models replicate many aspects of the earthquake cycle. An interesting difference, however, is that the models predict earthquakes at regular intervals, whereas earthquake histories are quite variable. Some of the variability may be due to the effects of earthquakes on other faults, or other segments of the same fault. Figure 5.7-24 shows this idea schematically for the slider model in Fig. 5.7-21. Assume that after an earthquake cycle, the compressive normal stress  $\sigma$  on the slider is reduced. This "unclamping" reduces the frictional force resisting sliding, so it takes less time for the spring force to rise again to the level needed for the next slip event. Conversely, increased compression "clamps" the slider more, and so increases the time until the next slip event. In addition, by Eqn 49, the stress drop in the slip event changes when  $\sigma$  changes.

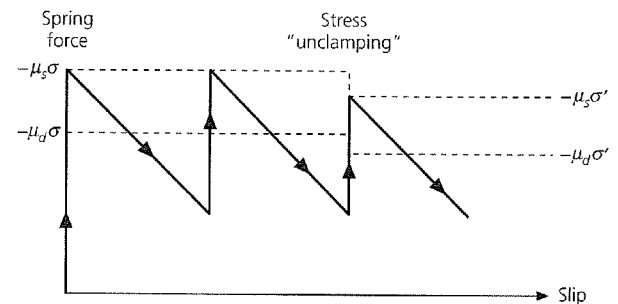


Fig. 5.7-24 Modification of a slider block model (Fig. 5.7-21) to include the effects of changes in normal stress. Reduced normal stress ( $|\sigma| < |\sigma'|$ ) reduces the frictional force, and so "unclamps" the fault and decreases the time until the next slip event.

For earthquakes, the analogy implies that earthquake occurrence on a segment of a fault may reflect changes in the stress on the fault resulting from earthquakes elsewhere. This concept is quantified using the Coulomb–Mohr criterion (Eqn 5) that sliding can occur when the shear stress exceeds that on the sliding line (Fig. 5.7-9), or  $\tau > \mu\sigma$ . We can thus define the Coulomb failure stress

$$\sigma_f = \tau + \mu\sigma \quad (55)$$

such that failure occurs when  $\sigma_f$  is greater than zero. Whether a nearby earthquake brings a fault closer to or further from failure is shown by the change in Coulomb failure stress due to the earthquake,

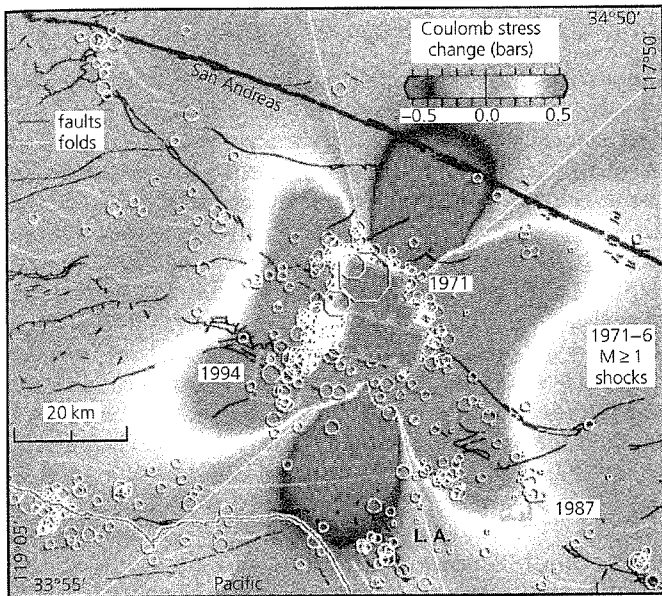


Fig. 5.7-25 Predicted changes in Coulomb failure stress due to the 1971 San Fernando earthquake. The Whittier Narrows and Northridge earthquakes subsequently occurred in regions where the 1971 earthquake increased the failure stress. (Stein *et al.*, 1994. *Science*, 265, 1432–5, copyright 1994 American Association for the Advancement of Science.)

$$\Delta\sigma_f = \Delta\tau + \mu\Delta\sigma. \quad (56)$$

Failure is favored by positive  $\Delta\sigma_f$ , which can occur either from increased shear stress  $\tau$  or a reduced normal stress (compression is negative, so  $\Delta\sigma > 0$  favors sliding).

Some earthquake observations provide support for this idea. Figure 5.7-25 shows the predicted Coulomb failure stress changes in the Los Angeles region due to the 1971 ( $M_s$  6.6) San Fernando earthquake. The stress change pattern reflects the earthquake's focal mechanism, thrust faulting on a NW–SE-striking fault (Fig. 5.2-3). Two moderate earthquakes, the 1987 Whittier Narrows ( $M_L$  5.9) and 1994 Northridge ( $M_w$  6.7) earthquakes subsequently occurred in regions where the 1971 earthquake increased the failure stress, suggesting that the stress change may have had a role in triggering the earthquakes. A similar pattern has been found after other earthquakes, and some studies have found that aftershocks are concentrated in regions where the mainshock increased the failure stress. Stress triggering may explain why successive earthquakes on a fault sometimes seem to have a coherent pattern. For example, the 1999  $M_s$  7.4 Izmit earthquake on the North Anatolian fault (Fig. 5.6-8) appears to be part of a sequence of major ( $M_s$  7) earthquakes over the past 60 years, which occurred successively further to the west, and hence closer to the metropolis of Istanbul.

An intriguing feature of such models is that the predicted stress changes are of the order of 1 bar, or only 1–10% of the typical stress drops in earthquakes (Section 4.6.3). Such small

stress changes should only trigger an earthquake if the tectonic stress is already close to failure. However, as in the slider model (Fig. 5.7-24), stress changes can affect the time until the tectonic stress is large enough to produce earthquakes. It has been argued that the 1906 San Francisco earthquake reduced the failure stress on other faults in the area, causing a “stress shadow” and increasing the expected time until the next earthquake on these faults. This is consistent with the observation that during the 75 years before the 1906 earthquake, the area had 14 earthquakes with  $M_w$  above 6, whereas only one occurred in the subsequent 75 years. Such analyses may help improve estimates of the probability that an earthquake of a certain size will occur on a given fault during some time period. To date, such estimates have large uncertainties (Section 4.7.3), in part because of the large variation in the time intervals between earthquakes. Stress loading models, some of which incorporate rate- and state-dependent friction because simple Coulomb friction does not predict large enough changes in recurrence time, may explain some of the variations and thus reduce these uncertainties.

This discussion brings out the importance of understanding the state of stress on faults. On this issue, the friction models give some insight, but major questions remain. Earthquake stress drops estimated from seismological observations are typically less than a few hundred bars (tens of MPa). Yet, the expected strength of the lithosphere (e.g., Fig. 5.7-14–16) is much higher, in the kilobar (hundreds of MPa) range. The laboratory results (Fig. 5.7-20) and frictional models (Fig. 5.7-21) suggest an explanation for this difference, because in both the stress drop during a slip event is only a fraction of the total stress.

However, the frictional models do not explain an intriguing problem called the “San Andreas” or “fault strength” paradox. As noted in Section 5.4.1, a fault under shear stress  $\tau$  slipping at rate  $v$  should generate fractional heat at a rate equal to  $\tau v$ . Thus, if the shear stresses on faults are as high (kbar or hundreds of MPa) as expected from the strength envelopes, significant heat should be produced. But little if any heat flow anomaly is found across the San Andreas fault (Fig. 5.7-26), suggesting that the fault is much weaker than expected. A similar conclusion emerges from consideration of stress orientation data. Although the Coulomb–Mohr model predicts that the maximum principal stress directions inferred from focal mechanisms, geological data, and boreholes should be about  $23^\circ$  from the San Andreas fault (Fig. 5.7-8), the observed directions are essentially perpendicular to the fault (Fig. 5.6-19), implying that the fault acts almost like a free surface. To date, there is no generally accepted explanation for these observations. The most obvious one is that the effective stress on the fault is reduced by high pore pressure, but there is discussion about whether pressures much higher than hydrostatic pressure could be maintained in the fault zone. An alternative explanation, that the fault zone is filled by low-strength clay-rich fault gouge, faces the difficulty that experiments on such material find that it has normal strength unless pore pressures are high.

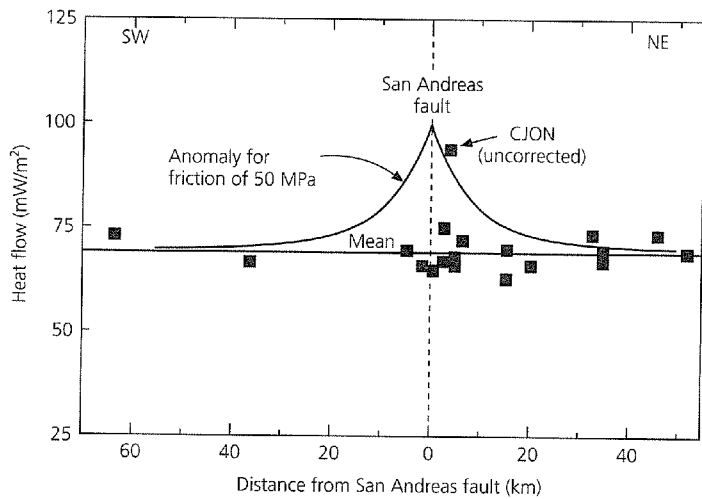
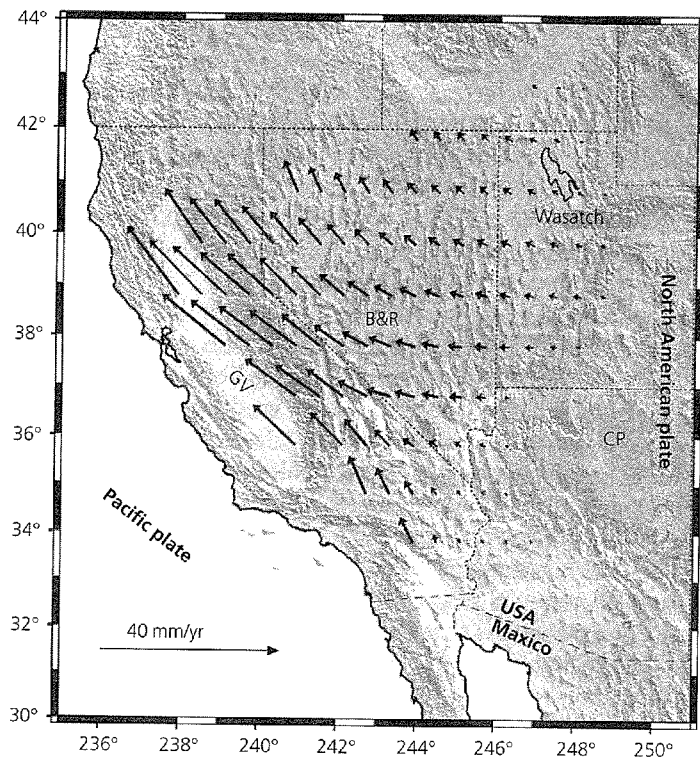


Fig. 5.7-26 Observed (squares) heat flow across the San Andreas fault. The elevated heat flow predicted by shear heating (solid line) is not observed, except for one point (CJON, Cajon pass), where alternative interpretations are possible, implying that the fault is weak. (Lachenbruch and Sass, 1988. *Geophys. Res. Lett.*, 15, 981–4, copyright by the American Geophysical Union.)

In summary, ideas based on rock friction are providing important insights into earthquake mechanics. Although many issues remain unresolved, and some attractive notions remain to be fully demonstrated, rock friction seems likely to play a growing role in addressing earthquake issues.



### 5.7.6 Earthquakes and regional deformation

The large, rapid deformation in earthquakes is often part of a slow deformation process occurring over a broader region. As discussed in Section 5.6.2, there often appear to be differences between the seismic, aseismic, transient, and permanent deformations sampled by different techniques on different time scales. Experimental and theoretical ideas about rheology and lithospheric dynamics are being used to investigate the relation between earthquakes and the regional deformations that produce them.

We have seen that earthquakes often reflect deformation distributed over a broad plate boundary zone. In this case, we can think of the lithosphere as a viscous fluid and use earthquakes as indicators of its deformation. This idea is like the physical model (Fig. 5.6-7) that used deformable plasticine as an analogy for the deformation of Asia resulting from the Himalayan collision. Figure 5.7-27 shows such an analysis for part of the Pacific–North America plate boundary zone in the western United States. The deformation is assumed to result from a combination of forces due to the transform plate boundary and forces due to the potential energy of elevated topography, which tends to spread under its own weight. To test this idea, a continuous velocity field has been interpolated from space-geodetic, fault slip, and plate motion data (Figs 5.2-3 and 5.6-3). The velocity field is treated as being due to the motion of a viscous fluid, and is converted to a strain rate tensor field. This is then compared to the magnitude of the stress tensor inferred

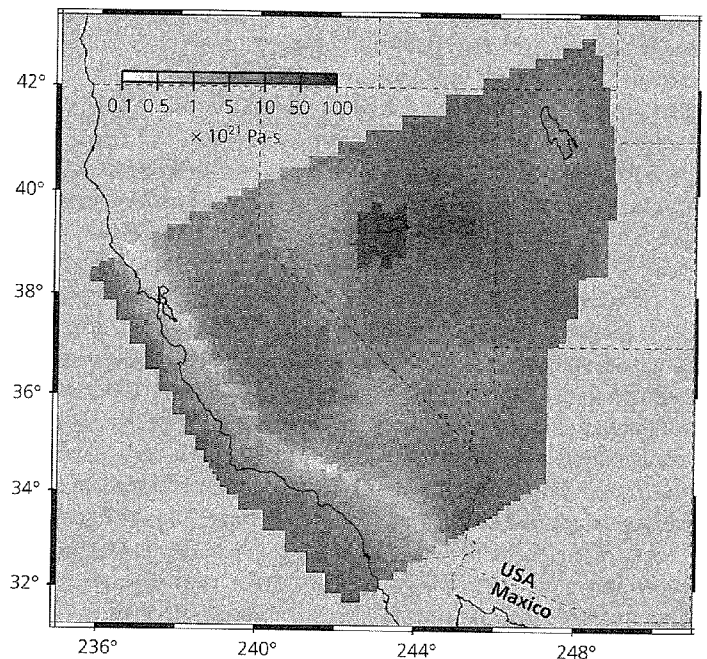


Fig. 5.7-27 *Left*: Estimated velocity field for part of the Pacific–North America plate boundary zone in the western USA. *Right*: Effective viscosity determined by dividing the magnitude of the deviatoric stress tensor by the magnitude of the strain rate tensor. (Flesch *et al.*, 2000. *Science*, 287, 834–6, copyright 2000 American Association for the Advancement of Science.)

from topography and plate boundary forces. The ratio of stress to strain rate at any point, which is the vertically averaged effective viscosity, varies significantly. Low values along the San Andreas fault and western Great Basin show that the strain rates are relatively high for the predicted stress, consistent with a weak lower crust. The Great Valley–Sierra Nevada block has little internal deformation, and thus acts relatively rigidly and appears as a high-viscosity region. Summing seismic moment tensors (Section 5.6.2) yields a seismic strain rate averaging about 60% of the inferred total strain. As discussed earlier, this discrepancy may indicate some aseismic deformation or that the 150 years of historical seismicity is too short for a reliable estimate.

Viscous fluid models can be used to study how the lithosphere deforms on different time scales. For example, as noted in Section 5.6.2, GPS data across the entire Nazca–South America plate boundary zone show faster motion than is inferred from structural geology or topographic modeling. The difference probably occurs because the GPS data record instantaneous velocities that include both permanent deformation and elastic deformation that will be recovered during future earthquakes, whereas the lower geological rates reflect only the permanent deformation. This can be modeled by representing the overriding South American plate using a simple one-dimensional system of a spring, a dashpot, and a pair of frictional plates (Fig. 5.7-28). This system approximates the behavior of the crust: the spring gives the elastic response over short periods, the dashpot gives the viscous response over geological time scales, and the frictional plates simulate the thrust faulting earthquake cycle at the trench. As plate convergence compresses the system, the stress  $\sigma(t)$  increases with time until it reaches a yield strength  $\sigma_y$ , when an earthquake occurs, stress drops to  $\sigma_b$ , and the process repeats. Displacement accumulates at a rate  $v_0$  except during earthquakes, when the displacement drops by an amount  $\Delta u$ . The topography and geologic data record the averaged long-term shortening rate  $v_c$  shown by the envelope of the sawtooth curve, whereas GPS data record the higher instantaneous velocity  $v_0$ . The instantaneous velocity thus results from the portion of the plate motion locked at the trench that deforms the overriding plate elastically (Fig. 4.5-14) and is released as seismic slip in interplate earthquakes. By contrast, the aseismic slip component at the trench has no effect because it does not contribute to locking on the interface and deformation of the overriding plate. Similar models are being explored for other regions where deformation appears to vary on different time scales.

Viscous fluid models are also used to analyze other aspects of the earthquake cycle. For example, Fig. 5.7-29 shows the strain rate near portions of the San Andreas fault compared to the time since the last great earthquake on that portion of the fault. Postseismic motion seems to continue for a period of years after an earthquake and then slowly decays, presumably due to the steady interseismic motion. A similar picture emerges from GPS and other geodetic results following large trench thrust faulting earthquakes. For a number of years, sites near the trench on the overriding plate move seaward, showing

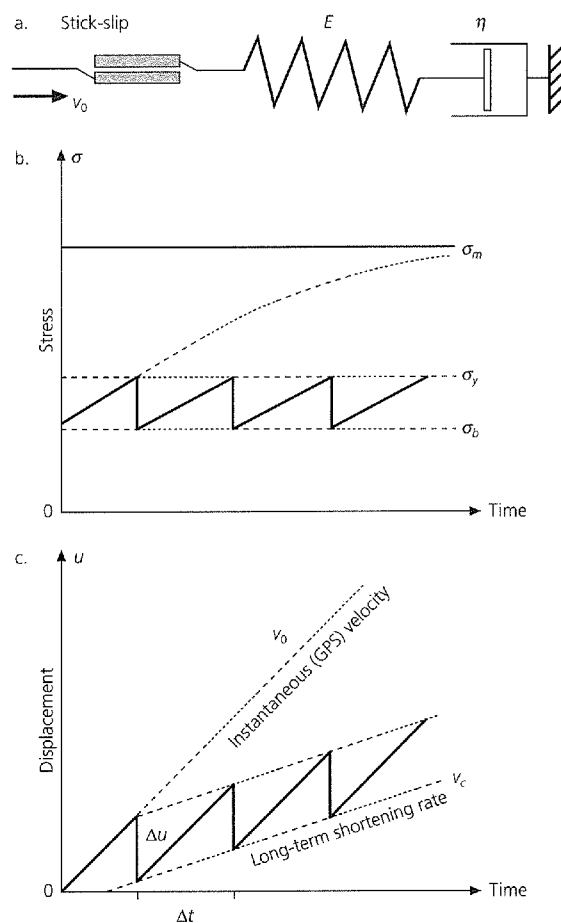


Fig. 5.7-28 a: Model for a viscoelastic–plastic crust to describe the response of the overriding South American plate to the subduction of the Nazca plate. The dashpot represents a viscous body modeling the permanent deformation, the spring represents an elastic body modeling the transient deformation, and the frictional plates represent the earthquake cycle at the trench. b: Stress evolution for the model. c: Displacement history for the model. Displacement accumulates at the instantaneous rate  $v_0$  except during earthquakes, when slip  $\Delta u$  occurs. GPS data record a gradient starting at  $v_0$  from the trench, whereas the envelope of the displacement curve  $v_c$  is the long-term shortening rate reflected in geological records and topography. (Liu *et al.*, 2000, *Geophys. Res. Lett.*, 18, 3005–8, copyright by the American Geophysical Union.)

postseismic motion consistent with the earthquake focal mechanism (Fig. 4.5-15). Eventually, however, the sites resume the landward interseismic motion usually seen at trenches (Fig. 5.6-10). Such observations are challenging to interpret, because postseismic afterslip on or near a fault can have effects at the surface similar to viscoelastic flow of the asthenosphere (Fig. 5.7-29), but offer the prospect of improving our understanding of both earthquake processes and the rheology of the lithosphere and the asthenosphere. A tantalizing possibility is that the viscous asthenosphere permits stress waves generated by large earthquakes to travel slowly for large distances and contribute to earthquake triggering.

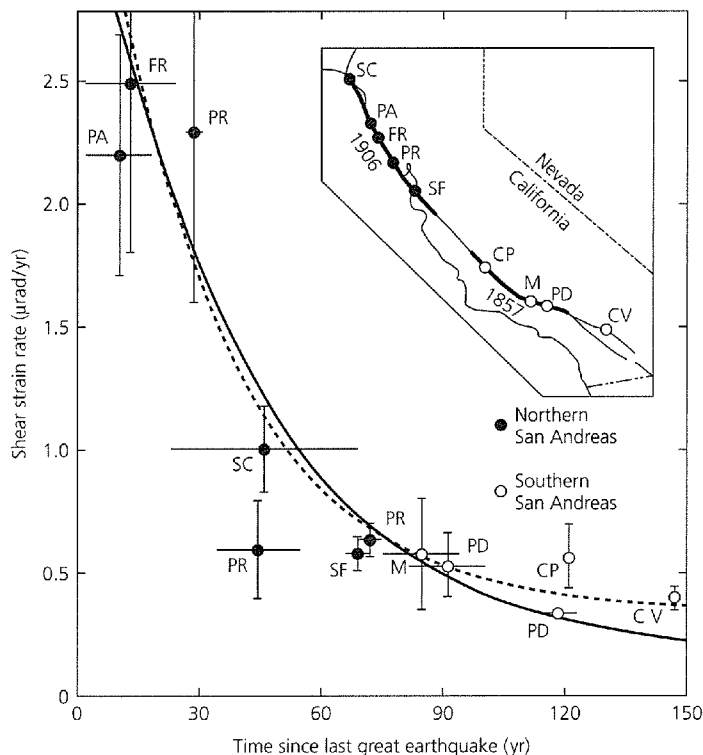


Fig. 5.7-29 Shear strain rate near portions of the San Andreas fault compared to the time since the last great earthquake. The data are similar to the predictions of two alternative models: viscoelastic stress relaxation (solid curve) and aseismic postseismic slip beneath the earthquake fault plane (dashed line). (Thatcher, 1983, *J. Geophys. Res.*, 88, 5893–902, copyright by the American Geophysical Union.)

## Further reading

Given the comparatively recent discovery of plate tectonics, its importance for most aspects of geology, and its crucial role in the earthquake process, many excellent sources, a few of which are listed here, offer more information about this chapter's topics.

The dramatic development of plate tectonics is discussed from the view of participants by Menard (1986) and in Cox's (1973) collection of classic papers. Basic ideas in plate tectonics are treated in most introductory and structural geology texts. More detailed treatments include Uyeda (1978), Fowler (1990), Kearey and Vine (1990), and Moores and Twiss (1995). Cox and Hart (1986) present the basic kinematic concepts, and global

plate motion models are discussed by Chase (1978), Minster and Jordan (1978), and DeMets *et al.* (1990).

Thermal and mechanical aspects of plate tectonics are discussed by Turcotte and Schubert (1982) and Sleep and Fujita (1997). Mid-ocean ridge tectonics and structure are discussed by Solomon and Toomey (1992) and Nicolas (1995). The thermal evolution of oceanic lithosphere is discussed by Parsons and Sclater (1977) and Stein and Stein (1992); McKenzie (1969) presents the subduction zone thermal model we follow. Papers in Bebout *et al.* (1996) cover many aspects of subduction, and Kanamori (1986) reviews subduction zone thrust earthquakes. Lay (1994) treats the nature and fate of subducting slabs, and deep earthquakes are reviewed by Frohlich (1989), Green and Houston (1995), and Kirby *et al.* (1996b). For a derivation of the ridge push force see Parsons and Richter (1980); Wiens and Stein (1985) discuss its application to oceanic intraplate stresses. Yeats *et al.* (1997) cover a wide variety of topics about the relation of earthquakes to regional geology. Rosendahl (1987) reviews continental rifting. Papers in Gregersen and Basham (1989) treat aspects of passive margin and continental interior earthquakes with emphasis on postglacial effects.

Concepts in continental deformation are treated by Molnar (1988) and England and Jackson (1989); Gordon (1998) gives an overview of plate rigidity and diffuse plate boundaries. Applications of space geodesy to tectonics are reviewed by papers in Smith and Turcotte (1993) and by Dixon (1991), Gordon and Stein (1992), and Segall and Davis (1997). Many GPS data and results, including an overview brochure, can be found on the University NAVSTAR Consortium WWW site <http://www.unavco.org>. Stress maps and their interpretations are discussed by Zoback (1992) and other papers in the same journal issue; stress maps are available at the World Stress Map project WWW site <http://www-wsm.physik.uni-karlsruhe.de>.

Mantle plumes in general are reviewed by Sleep (1992); Nataf (2000) and Foulger *et al.* (2001) discuss seismic imaging of plumes; Smith and Braile (1994) discuss the Yellowstone hot spot; and Stein and Stein (1993) discuss oceanic hot spot swells. Papers in Peltier (1989) treat many aspects of mantle convection; Silver *et al.* (1988) explore the relationship between subduction, convection, and mantle structure; and Christensen (1995) reviews the effects of phase transitions on mantle convection. The heat engine perspective on global tectonics is discussed by Stacey (1992), and Ward and Brownlee (2000) summarize the arguments advocating a crucial role for plate tectonics in the origin and survival of life on Earth.

Topics involving rock mechanics, flow, and their tectonic applications are discussed by Jaeger (1970), Weertman and Weertman (1975), Jaeger and Cook (1976), Turcotte and Schubert (1982), Kirby (1983), Kirby and Kronenberg (1987), and Ranalli (1987). Scholz (1990) and Marone (1988) cover topics dealing with the relation of rock mechanics to earthquakes, with special emphasis on rock friction. Our treatment of the slider model for faulting follows Scholz (1990). Related topics, including issues of continental deformation and fault strength, are also treated by papers in Evans and Wong (1992). Stein (1999) summarizes the concept of stress triggering of earthquakes.



## Problems

- Assume that Pacific–North America plate motion along the San Andreas fault occurs at 35 mm/yr.
  - If all this motion occurs seismically in earthquakes about 22 years apart, which is a typical recurrence interval for the Parkfield fault segment, how much slip would you expect in the earthquakes? From Fig. 4.6-7, estimate likely fault lengths and magnitudes for such earthquakes.
  - Give similar estimates if the earthquakes occur about 132 years apart, as at Palte Creek.
- Assume that all the earthquakes in the Palte Creek sequence (Fig. 1.2-15) involved 4 m of seismic slip. Using the time interval from the present to the 1857 earthquake, calculate the seismic slip rate on this portion of the San Andreas fault. Next, do so by averaging the recurrence intervals for the past two earthquakes (1857 and 1812), the past three, and so on for the entire earthquake history. What are the implications of this simple experiment for seismic slip estimates? What other sources of uncertainty should also be considered, and how might they affect this estimate?
- Use Table 5.2-1 to find the rate that the Juan de Fuca plate subducts beneath North America at 46°N, 125°W.
  - If all this motion occurs in large earthquakes, how often would you expect an earthquake if the slip in each were 5 m? How would this estimate change if the slip were 10 or 20 m?
  - How would the answers to (b) change if only 25% or 50% of the plate motion occurred by seismic slip?
  - Paleoseismic observations and historic records of a tsunami imply that this subduction zone has had very large earthquakes approximately 500 years apart. Suggest some possibilities in view of parts (a)–(c). How might you attempt to distinguish between them?
  - The crust subducting at this trench is about 10 million years old. Given the convergence rate and the observations from other trenches in Fig. 5.4-30, what might you infer about the moment magnitude of the largest earthquake expected here? Find the corresponding seismic moment and suggest a plausible fault geometry and amount of slip that would also be consistent with the paleoseismic and plate motion observations.
- For rigid plates, Eqn 5.2.10 shows that we can find the *angular* velocity vector of one plate from the sum of two others. Show that at a point we can also do this for the *linear* velocity vectors.
- The news media sometimes ask “How large would the largest possible earthquake be?” Estimate the seismic moment and moment magnitude by assuming that all the trenches in the world (48,000 km) slip at the same time, that 10 m of slip occurs, and the fault width is 250 km.
- Estimate the thermal Reynolds number  $R$  defined in Eqns 5.3.19 and 5.4.3, assuming that  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ . What does this estimate imply about the processes of plate cooling and subduction?
- Assume that oceanic lithosphere has a thermal conductivity of  $3.1 \text{ W m}^{-1} \text{ }^\circ\text{C}^{-1}$ .
  - Find the heat flow for old oceanic lithosphere, assuming a linear temperature gradient (Fig. 5.3-8), a basal temperature of 1450°C, and a plate thickness of 95 km.
  - How would this value change for a basal temperature of 1350°C and plate thickness 125 km?
  - If the lithosphere under a midplate region were thinned to 50 km while the basal temperature remained 1350°, what would the heat flow be, assuming a linear temperature gradient?
- A way to get insight into the physics of subduction is to use a classic result from fluid mechanics, called Stokes’ problem, which describes the terminal velocity  $v$  at which a sphere of radius  $a$  and density  $\rho$  sinks due to gravity in a fluid with viscosity  $\eta$  and lower density  $\rho'$ . The result is  $v = 2ga^2(\rho - \rho')/9\eta$ . Estimate the subduction velocity of a slab assuming the slab is a sphere with radius equal to half its thickness. To do this, estimate the density contrast from the thermal model (Eqn 5.4.14) and a coefficient of thermal expansion  $\alpha = 3 \times 10^{-5} \text{ }^\circ\text{C}^{-1}$ . Use a mantle viscosity from Section 5.5.3. Because this is a back-of-the-envelope calculation, there is no correct answer, but you should be able to come up with something reasonable (within an order of magnitude or two of reality).
- The result that a subducting slab that reaches the core should still be thermally distinct (Fig. 5.4-5) may seem surprising. For another estimate, use the one-dimensional cooling equation in Section 5.3.2 to estimate how long a slab should need to warm up to 90% of the ambient lowermost mantle temperatures, assuming that it were immediately transported to the base of the mantle and that  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ .
- Using the definition of the slab pull force (Eqn 5.4.15):
  - Write the force in terms of the age of the subducting plate.
  - Explain whether this force would be greater or smaller, and why, for increased values of subducting plate age, coefficient of thermal expansion, and thermal diffusivity.
- Assume that in a subducted slab the depth of the spinel–perovskite phase transition deepens from its usual 660 km outside the slab to 700 km, and that the core of the slab is 800° colder than the surrounding mantle. What is the Clapeyron slope of the phase change?
- The surface of Venus is much hotter (450°C) than that of Earth. If Venus had plate tectonics and the rocks were similar, so that the temperature gradient in old lithosphere there were the same as on Earth, how would the thickness of the “oceanic” lithosphere differ? How would the slab pull and ridge push forces differ? What other differences might you expect?
- Express the ratio of the slab pull (Eqn 5.4.15) and ridge push (Eqn 5.5.6) forces. Explain why this ratio depends on thermal diffusivity. Estimate this ratio near a trench where old oceanic lithosphere is subducting, assuming that  $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$ .
- To see if momentum can be responsible for the Indian plate’s northward motion long after its collision with Asia began, estimate the momentum of the Indian plate and that of an ocean liner, and compare the two.
- Use Mohr’s circle to show why
  - Rocks at depth do not fracture under lithostatic pressure alone.
  - The deviatoric stress needed for fracture increases at greater depth.
- Suppose that a rock is stressed close to its brittle limit. Show graphically which will make the rock fracture sooner: (a) increasing  $\sigma_1$  or (b) decreasing  $\sigma_2$  by the same amount (assume a two-dimensional case where  $\sigma_1$  and  $\sigma_2$  are both negative, and internal friction exists).
- Suppose that the fracture line for a particular rock is  $\tau = 80 - 0.5\sigma$ , where stresses are in MPa. What angle would the normal to a fracture plane make with  $\sigma_1$ ? If  $\sigma_1$  is 400 MPa at failure, what is  $\sigma_2$ ?
- For the slider block earthquake model in Section 5.7.5:
  - Derive an expression for the time between successive slip events.



- (b) Sketch the force-slip diagram for two different spring constants, and use the sketch to explain how the slip and force drop in a slip event change and why.
- (c) For the slider block model, formulate a quantity analogous to an earthquake's seismic moment, and explain why it depends on each term. What is the major difference between this quantity and the seismic moment?
- (d) Recall the observation (Fig. 4.6-11) that earthquake stress drops are similar for a wide range of earthquakes. If the slider block model is relevant, what does this imply?
- (e) What conditions might correspond to aseismic slip, which could be viewed as the limit of a continuous series of very small slip events?

### Computer problems

- C-1. (a) Write a subroutine to compute the rate and azimuth of plate motion at a point, given the location and an Euler vector in the form (pole latitude, longitude, magnitude).
- (b) Use the Euler vector in Table 5.2-1 to test your program on the San Andreas and Aleutian site examples in Section 5.2.1.
- C-2. (a) Find the rate and azimuth of Cocos-North America plate motion at 18.3°N, 102.5°W.
- (b) This location is the epicenter of a large 1985 Mexican earthquake, whose mechanism had nodal planes whose strike and dip are (127°, 81°) and (288°, 9°). Infer from the tectonics of the Middle American trench which plane was the fault plane. Using the methods of Section 4.2, determine the azimuth of slip during the earthquake. How does this compare to your predicted azimuth?
- C-3. (a) Write a subroutine to add and subtract two Euler vectors given in the form (pole latitude, longitude, magnitude). The output should be a Euler vector in the same form.
- (b) Use your program to determine the absolute Euler vector for the Pacific plate using Table 5.2-1.
- (c) Determine the rate and azimuth of absolute plate motion at Hawaii (Fig. 5.2-7). Compare the direction to the Hawaiian-Emperor seamount chain.
- C-4. Write a program to plot the temperature distribution in the oceanic lithosphere as a function of age using the cooling half-space thermal model (Eqn 5.3.4). Compute  $\text{erf}(s)$  (Eqn 5.3.3) using either available software or numerical integration as discussed in problem 4C-6.
- C-5. (a) Write a program to plot the temperature distribution in a subducting slab using the analytic thermal model (Eqn 5.4.3). Compute it for a plate subducting at 80 mm/yr at an angle of 45°. Make assumptions that seem reasonable and justify them.
- (b) Change the program to make the age of the subducting plate a parameter and generate temperature fields for different slabs, as in Fig. 5.4-6.
- (c) Using the results of (b) and Fig. 5.4-4, estimate a temperature above which deep earthquakes are not observed.