Strength and Energetics of Active Fault Zones

James N. Brune
University of Nevada, Reno, USA
Wayne Thatcher
US Geological Survey, Menlo Park, USA

1. Introduction

The strength of active fault zones, i.e., the shear stress level required to cause fault slip, is fundamental to understanding the physics of earthquakes and to assessing earthquake hazard. Although many researchers have concluded that fault zones are weak (shear stresses 10 MPa or less averaged between 0 and $\sim$20 km depth), others maintain that faults are strong ($\sim$100 MPa average of an approximate linear increase with depth). Thus, despite 30 y of dedicated research, relevant data remain inconclusive and fault strength remains uncertain by an order of magnitude. In part, this is because the main source of energy release in earthquakes is at depths greater than 5 km, inaccessible to direct instrumental observation. Very large earthquakes rupture to the Earth's surface where direct observation of the shallow rupture process is possible. However, the rupture characteristics at shallow depth may differ from those at seismogenic depths. To date no great ($M \geq 7.75$) earthquakes have occurred within a network of modern strong-motion instruments, but the large, well-recorded 1999 earthquakes in Turkey and Taiwan, both $M = 7.6$, show that this data gap is rapidly being filled. Furthermore, shear stress changes at the earthquake source (and the resulting seismic waves) are nearly linear perturbations of the absolute stress field. Thus, an unknown absolute background stress does not greatly affect the basic characteristics of the observed low frequency seismic waves and the observed geodetic deformation. Other, less-direct data must then be used to infer the physical state and ambient stress levels on active faults.

This chapter discusses available evidence and current ideas about fault zone strength and energetics. In our review we begin by outlining the general physical conditions prevailing in the Earth around active faults and summarize the generally agreed upon observational features of faulting and earthquake occurrence (Section 2). In doing so we make as few assumptions as possible, because making uncertain assumptions may lead to logical inconsistencies and apparent paradoxes. For example, the absence of a measurable, frictionally generated heat flow anomaly near active faults, the so-called stress-heat flow paradox, rests upon assumptions that must somehow be incorrect. We wish to avoid such inconsistencies and to begin we introduce the observations and briefly state what they imply about fault zones. Subsequently, in Sections 3–8, we place these observations in an interpretative context and show how they have been used to infer various measures of fault shear stress. Section 9 summarizes our assessment of the average shear stress state near faults, argues for the general importance of stress heterogeneity in faulting processes, and discusses its implications.

2. General Physical Problem and Observational Constraints

Active faults are stressed by forces applied in the adjacent lithosphere, forces applied at or near plate boundaries due to the motions of the plates as well as those caused by lateral density contrasts within the lithosphere. Parts of some active faults, particularly those at major plate boundaries, slip aseismically, keeping fault zone shear stresses at about the same levels. Others slip primarily in earthquakes, with the fault shear stresses increasing between events and decreasing abruptly when sudden seismic slip occurs. During earthquakes, energy released at depth in the Earth propagates as elastic waves that cause ground shaking when they reach the Earth's surface. We have no direct access to the depths at which the major energy is released. Thus the physical processes must be inferred indirectly from evidence obtained at the surface (proposals for deep drilling into active crustal fault zones are pending, and

would provide the first direct evidence of the physical state at seismogenic depths). What we know about active faults and earthquakes thus comes from measurements made at or near the Earth's surface, and relies on seismology, structural geology, geodesy, and other geophysical data to infer the processes that are thus far inaccessible to direct observation. Studies of these data have established the following general features of stress, faulting, and earthquake occurrence. Detailed discussion follows in the section indicated.

1. Lateral density gradients in the lithosphere and variations in surface and sea-floor topography generate differential stresses of ~10–100 MPa available to drive slip on faults (Section 3).
2. Stresses measured *in situ*, in mines to depths of ~3 km and in boreholes to as deep as 9 km, are consistent with an approximately linear increase in shearing stress with depth of ~5–15 MPa km^{-1}. However, these measurements, particularly those below a few km depths, come from tectonically inactive plate interiors far from large active faults (Section 5).
3. Studies of inactive fault zones exhumed by erosion from seismogenic depths show them to be generally planar features, often with zones of crushed and comminuted rock (fault gouge, microbreccia, and cataclasite) up to 1 km in width. Faulting involves frictional slip to depths corresponding to temperatures of ~350°C; at greater temperatures, deformation is primarily ductile, though still confined to relatively narrow shear zones (see Chapter 29 by Sibson). The fault zones show evidence of having been fluid saturated throughout their depth range (Section 5).
4. Heat flow measurements near active faults show no evidence for the heat generation expected if there were significant frictional resistance to fault slip. Furthermore, the orientations of principal compressive stresses near some major active faults are nearly normal to their strike. These two observations have been used to suggest faults are weaker (support less shearing stress) than the surrounding blocks, with an upper bound on average fault zone shear stress of ~20 MPa or less (Section 6).
5. Data from seismology, geodesy, and geological mapping of earthquake faulting are all consistent with average earthquake stress drops of ~0.1–10 MPa. These same data show, however, that slippage on the fault surface is very spatially heterogeneous, indicating local stress drops up to an order of magnitude greater than the average values (Section 7).
6. Measurements of strong ground motions made near earthquakes as large as $M7$ reveal accelerations as large as $1g$ and velocities up to $2\,\text{m}\,\text{sec}^{-1}$, indicating dynamic stress changes of about 30 MPa on the fault (Section 8).

In what follows we amplify on these observations and explain the inferences derived from them. In each case we point out the assumptions implicit in the inference, our own assessment of their reliability, and the implications of the results for fault zone stress and energetics.

3. External Stresses Available to Drive Faulting

The energy that drives active faulting ultimately comes from the stresses imposed on the lithosphere by the forces that drive and resist motions of the major plates and by stresses due to lateral density gradients in the lithosphere. Although at least eight different plate forces are potentially important, only some of these are believed to be decisive in determining the force balance (Forsyth and Uyeda, 1975) and hence the intraplate stress field (Zoback *et al.*, 1989). Here we shall describe only those considered in recent work. The forces are illustrated in Figure 1, which should be referred to in the subsequent discussion of each force.

The best understood force is the "ridge push," F_{rp}, the mean excess pressure (differential stress) ΔP exerted on the lithosphere due to the elevation of the mid-ocean ridges above the surrounding sea floor. It is given by $\Delta P = 0.5g\ (\rho_a - \rho_w)e$, where g is the acceleration of gravity, ρ_a is the asthenospheric density, ρ_w is the density of sea water and e is the elevation of the ridge above the sea floor. All of these parameters are known rather well, and so the resulting differential stress estimate of 30 MPa averaged over an oceanic lithospheric thickness of ~70 km is probably accurate. Several recent studies suggest that the ridge push force is the most important determinant of intraplate stress in the North American plate (e.g., Richardson and Reding, 1991; Zoback, 1992). If so, the field has a particularly simple, predictable form and its magnitude is
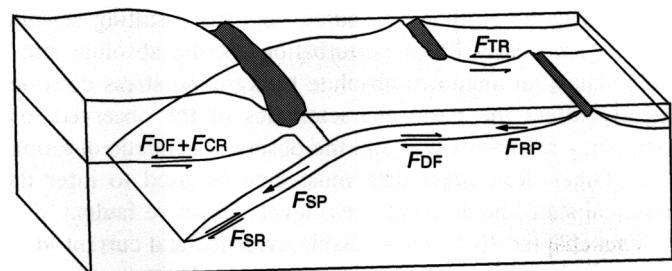
FIGURE 1 Plate driving and resisting forces (modified from Forsyth and Uyeda, 1975). $F_{\text{TR}} = $ transform fault resistance; $F_{\text{DF}} = $ plate drag force; $F_{\text{CR}} = $ plate drag under continents; $F_{\text{RP}} = $ ridge push; $F_{\text{SP}} = $ slab pull; $F_{\text{SR}} = $ slab resistance.

bounded rather well. Since the load-bearing thickness of the continental lithosphere may well be considerably less than 70 km in active regions, differential driving stresses from ridge push could be as much as ~100 MPa in these areas.

The "slab pull" force, F_{sp}, is caused by the excess density of the cold subducting oceanic slab as it sinks into the hotter, more buoyant asthenosphere. Its magnitude can be estimated from the thermal structure of the descending plate (McKenzie, 1969) and it is likely to be the largest of the driving forces. The corresponding force per unit area acting on the lithosphere is about 180 MPa, but this is an upper bound estimate of the differential driving stresses imposed on the lithosphere by subduction. Slab descent into the mantle is almost certainly resisted by poorly known forces of comparable magnitude, F_{sr} in the asthenosphere, and F_{cr} on the intraplate interface of the subduction thrust. As a result, the net drive or resistance from subduction is not well known, although it is likely that this net force importantly influences differential stresses in some plates.

Transform fault resistance, F_{tf}, opposes the strike-slip motion of the plates on both oceanic transforms and continental transcurrent faults like the San Andreas system. Although these frictional and ductile resisting shear stresses could be significant, their magnitudes are poorly known. We discuss them further below in connection with lithospheric rheology. Quasi-static frictional shear resistance on the seismogenic upper crustal faults could average 60 MPa across strike-slip faults and as much as 400 MPa across the deeper portions of subduction thrusts (F_{cr} of Fig. 1). However, as discussed below, heat flow data in both transform and subduction settings place much lower thresholds on the magnitudes of these resisting shear stresses.

Finally, shear stresses imposed on the base of the lithosphere, either as driving or resisting forces, F_{df}, are potentially important but poorly known. F_{df} is a driving force if imposed by general convective flow of the mantle, or resistive if caused by the drag of the plates over a passive asthenosphere. Driving or resisting shear stresses of only a few bars integrated over the large basal area of the plates would have important influences on the intraplate stress field.

Whatever the magnitudes of these largely unknown forces their net effect is to impose a long-wavelength stress field on the interiors of the plates. It is this field, often modified by local perturbations, that supplies the "tectonic" shear stresses that drive active faulting.

Buoyancy forces caused by lateral density contrasts in crust or mantle are the most important local perturbations. Just as the elevated topography at midocean ridges leads to large intraplate stresses, a similar process occurs when elevated topography is caused by thickened continental crust. By considering the force balance for two columns of continental crust, one with thickness y_{cco} and the other thickened so that it has an additional elevation h, and assuming the two columns

are in isostatic balance, Turcotte (1982) showed the net horizontal force F_r is

$$F_r = \rho_c \, gh \left[y_{cco} + \frac{\rho_m h}{\rho_m - \rho_c} \right]$$

where ρ_c and ρ_m are the densities of crust and mantle respectively. If $y_{cco} = 35$ km, $\rho_m = 3300$ kg m^{-3}, $\rho_c = 2750$ kg m^{-3} and we assume F_r is supported over an elastic crust 50 km thick, then we can obtain the resulting compressive stress as a function of elevated topographic height h. If $h > 3$ km then differential stresses exceed 100 MPa. In the absence of any other applied forces we expect these topographic effects will lead to extensional stresses within the elevated region and compressive stresses in the adjacent thinner crust. The same general principles that we have discussed for the crust apply to lateral density gradients in the mantle (see Fleitout and Froidevaux, 1982 for a general formulation), and both lead to stresses that can drive faulting.

3.1 Assessment

Our assessment is that intraplate differential stresses due to plate motion forces and lateral density gradients within the lithosphere can be estimated approximately and lie in the range ~10–100 MPa averaged over the entire lithosphere. They thus provide rough bounds on the magnitude of the long wavelength, far field shear stress that causes faulting.

4. Internal Fault Zone Stress and Energy Balance for Fault Slip

The tectonic stresses discussed above, supported by the lithosphere over long time intervals, provide the boundary stresses which ultimately supply energy for earthquakes and fault slip. Earthquake occurrence modulates the local stress field in an intermittent fashion while steady-state aseismic slip keeps shear stress at about the same level at all times. The fault zone (including gouge) represents an internal boundary on which we seek to infer the stresses. Relatively rapid earthquake fault slip will decrease stress on the fault and its surroundings, radiating seismic waves. In the process, work is done against frictional stresses that resist fault motions. Elastic stress accumulation subsequently restores the stress slowly, over hundreds or thousands of years, to an ambient prefailure level that represents the actual "strength" of the fault. The earthquake stress drop and the absolute ambient pre-earthquake stress, along with the frictional resisting forces active during sliding, determine the partitioning of stored elastic energy released during faulting (E) into seismic waves (E_s) and work done against resisting stresses (E_r). In what follows we outline how this partitioning provides a framework for considering

constraints on the magnitudes of stresses acting before, during and immediately after earthquake fault slip. This will in turn provide a bounding framework for evaluating various models and inferences of stress state.

To begin we discuss a relatively small rupture surface over which it may be assumed that the stress is uniform. We assume a confined planar fault surface and uniform stresses and stress changes over the fault. We define u as the total slip averaged over the fault area A. τ_i and τ_f are the initial and final shear stresses on the fault, and the average resisting shearing stress during slip is τ_r. Following Brune (1976) and Lachenbruch and Sass (1980), Figure 2 shows three possibilities for the relative magnitudes of the final stress and average resisting stress ($\tau_r > \tau_f$, overshoot; $\tau_r = \tau_f$, null; and $\tau_r < \tau_f$, locking). The total elastic energy release is the area under the straight line joining τ_i and τ_f in Figure 2,

$$E = \tfrac{1}{2}(\tau_i + \tau_f)uA \qquad (1)$$

Then the work done against resisting stresses is the area under the line $\tau = \tau_r$ in Figure 2,

$$E_r = \tau_r uA \qquad (2)$$

The seismically radiated energy is the difference between Eq. (1) and Eq. (2), the area between the elastic energy release line and the resisting stress line τ_r,

$$E_s = E - E_r = \left[\tfrac{1}{2}(\tau_i + \tau_f) - \tau_r\right]uA \qquad (3)$$

Figure 2 graphically shows how the energy partitioning depends on the relative magnitudes of the initial, final, and
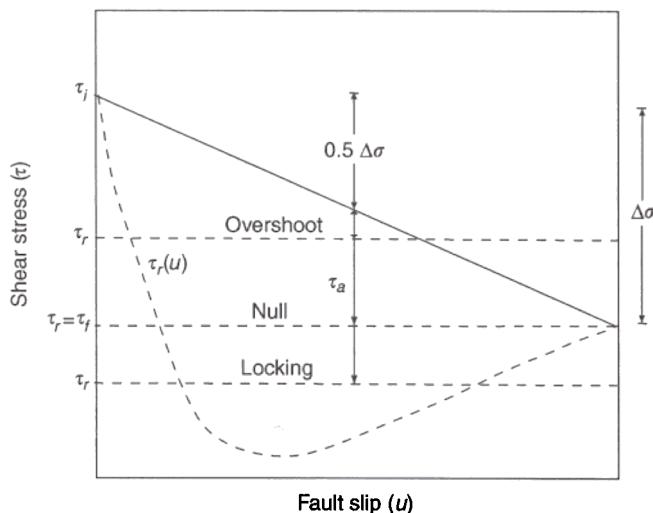
Fault slip (*u*)

FIGURE 2 Average stress levels during fault slip. Initial stress $= \tau_i$, final stress $= \tau_f$, resisting stress $= \tau_r$, apparent stress $= \tau_a$, stress drop $= \Delta\sigma$. Dashed line shows case of variable dynamic friction during slip. Three possibilities for the relative magnitudes of the final stress and average resisting stress are shown ($\tau_r > \tau_f$, overshoot; $\tau_r = \tau_f$, null; and ($\tau_r < \tau_f$, locking).

resisting stresses on the fault. The detailed dynamics of earthquake faulting certainly depend on dynamic shear resistance, as shown by the dashed line in Figure 2, but average values remain useful for intuitive understanding.

We may relate several spatially averaged measures of stress by using Eqs. (1–3) and defining some commonly used stress parameters as follows

Average loading stress:	$\tau_e = \tfrac{1}{2}(\tau_i + \tau_f)$	(4)
Apparent stress:	$\tau_a = \tfrac{1}{2}(\tau_i + \tau_f) - \tau_r$	(5)
Stress drop:	$\Delta\sigma = \tau_i - \tau_f$	(6)

Then since $E = E_s + E_r$ we have

$$\tau_e = \tau_a + \tau_r \qquad (7)$$

Using Eqs. (5) and (6) we finally obtain

$$\tau_i = \tau_a + \tfrac{1}{2}\Delta\sigma + \tau_r \qquad (8)$$

Equation (8) is useful because it relates stress measures that we can estimate with varying degrees of precision and uncertainty from observational data, permitting inter-comparisons and checks on consistency. In what follows we will critically assess the state of knowledge of each stress measure in order to constrain the true strength and energy balance in active fault zones. Our general objective is to understand what physical factors control stresses, to evaluate observations that constrain their magnitudes, and to assess whether these average stresses are relatively high ($\sim$100 MPa or greater) or rather low ($\sim$20 MPa or less). As we discussed previously, rough bounds can be placed on the initial stress τ_i based on intraplate stress estimates. The apparent stress τ_a can be obtained for individual earthquakes by seismically measuring the moment and radiated energy release. Stress drops ($\Delta\sigma$) for individual earthquakes can be estimated from geodetic measurements, as well as seismically. Stresses τ_r that resist fault motions generate heat, and thermal measurements near active faults constrain the long-term (millions of years) average resisting stress across crustal faults.

5. Quasi-static Lithospheric Rheology

The rheology of the lithosphere, the constitutive laws that determine quasi-static deformation for given applied stresses, may play a decisive role in governing the magnitude and distribution of ambient stress. It is the shear stress component of the ambient field resolved onto the fault that is identified with the initial stress τ_i in the quasi-static model for rupture discussed above. The upper 10–30 km of the crust, where ruptures typically nucleate and propagate, deforms elastically

or through brittle failure. Below these depths, aseismic fault slip and distributed bulk ductile deformation occur in the deeper portions of a lithospheric column that may be as much as ~100 km thick (Brace and Kohlstedt, 1980).

The vertical stratification and lateral heterogeneity of lithospheric rheology determine the coupling between the far-field tectonic stress, discussed above, and the local deformation. This variability in rheology affects the ways in which stresses are transferred into the "earthquake machine" in the seismogenic upper crust. Local changes in bulk lithospheric rheology, for example near major faults, could profoundly influence both the stress field and the patterns of active faulting.

The rheological behavior of fault zones, either during earthquake slip or quasi-static slip, is not yet well understood, and in what follows we use extrapolations from laboratory rock mechanics experiments to derive idealized static strength versus depth profiles for the lithosphere. It is not yet certain how far they can be applied to conditions existing in and near active faults. Indeed, as we shall see, observations from the San Andreas fault and elsewhere violate the simplest expectations based on these derived static strength profiles, and so additional complexities are needed to rationalize the field measurements within the rock mechanics framework. In addition, the stress field during faulting in the seismogenic upper crust may be dominated by dynamical effects that are not closely linked to the quasi-static, prefailure stress levels. Thus, although significant uncertainties also exist in understanding the rheology of the deeper, ductile lithosphere, the quasi-static parameters based upon laboratory experiments may more nearly apply there.

Whatever the actual stress state in the Earth, the conventional rock mechanics framework developed here is a useful standard against which to evaluate observations and models of fault zone strength. In what follows we separately consider the rheology of the seismogenic crust and the region that lies beneath it.

5.1 Brittle Upper Crust

Let us define strength as the maximum shear stress that can be supported by rocks. Within the conventional framework, in the upper 10–30 km of the Earth this strength is determined by the stresses required to cause frictional sliding on preexisting faults or fractures. For intact rock masses the brittle fracture strength is generally higher but is more difficult to estimate because it varies widely with rock type, temperature, and strain rate and it is not considered further here.

The general form of the failure condition for frictional slip is given by

$$\tau = \mu (\sigma_n - P) + \tau_0 \qquad (9)$$

where τ and σ_n are the shear and normal stresses across a planar fault surface, μ is the coefficient of static friction on the surface,

P is the ambient fluid pressure, and τ_0 is a cohesion term which we subsequently drop for simplicity. Note that high fluid pressure P could considerably reduce the tectonic shear stress τ required to cause frictional failure, a point we will return to in Section 9.

Laboratory measurements of frictional sliding on rock surfaces under the temperature and pressure conditions of the crust provide potentially strong constraints on frictional slip. They lead to the important generalization that the frictional resistance to sliding is (with a few exceptions) independent of rock type and depends only upon confining pressure (Byerlee, 1978). All of the data except those for a few clay-rich minerals can be fit well assuming

$$\mu = 0.75$$

5.1.1 Triaxial Stress State

If we consider a triaxial stress state with effective principal compressive stresses $(\sigma_1 - P) > (\sigma_2 - P) > (\sigma_3 - P)$ and the fault plane lying at angle θ to the σ_1 axis (Fig. 3), we may rewrite Eq. (9) in terms of the principal effective stresses (see Jaeger and Cook, 1976, p.14) $(\sigma_1' - \sigma_3')\sin 2\theta = \mu[(\sigma_1' + \sigma_3') - (\sigma_1' - \sigma_3')\cos 2\theta]$, where we have replaced all $(\sigma_i - P)$ by σ_i'. After a little algebra we obtain

$$\sigma_1'/\sigma_3' = (1 + \mu \cot \theta)/(1 - \mu \tan \theta)$$

This ratio has a minimum value for fault planes perpendicular to the $\sigma_1 - \sigma_3$ plane intersecting the σ_2 axis and oriented at angles of θ_0 to σ_1. This angle is given by

$$\theta_0 = 45° - \tfrac{1}{2}\tan^{-1} \mu = \tfrac{1}{2}\tan^{-1}(1/\mu)$$

Faults at this optimal orientation thus slip at the minimum value of σ_1'/σ_3' and it is often assumed that such faults exist in the crust when calculating its frictional strength. However, it is not uncommon for pre-existing faults not optimally oriented in this way to be reactivated in the current (different) tectonic stress field, in which case the stress ratio must be greater. Frictional lock-up occurs for faults oriented more than $2\theta_0$ from the σ_1-axis (Fig. 3c), which defines the asymptotic limits for θ (Sibson, 1985). If all available orientations for pre-existing faults lie close to the $2\theta_0$ limit, new faults of optimum orientation may be formed at lower stresses by fracture of previously unfaulted rock.

In principle, failure could occur for $\theta > 2\theta_0$ but this would require negative values of the stress ratio σ_1'/σ_3'. This implies the effective least principal stress $\sigma_3' < 0$, i.e., $P > \sigma_3$, and the rock mass surrounding the fault may show evidence of tensional failure by hydraulic fracturing (see Chapter 29 by Sibson). This condition may apply without pervasive hydrofracture if the region of high pore pressure is confined near the fault zone, permitting slip at low effective stresses on very misoriented faults in a stress field characterized by locally
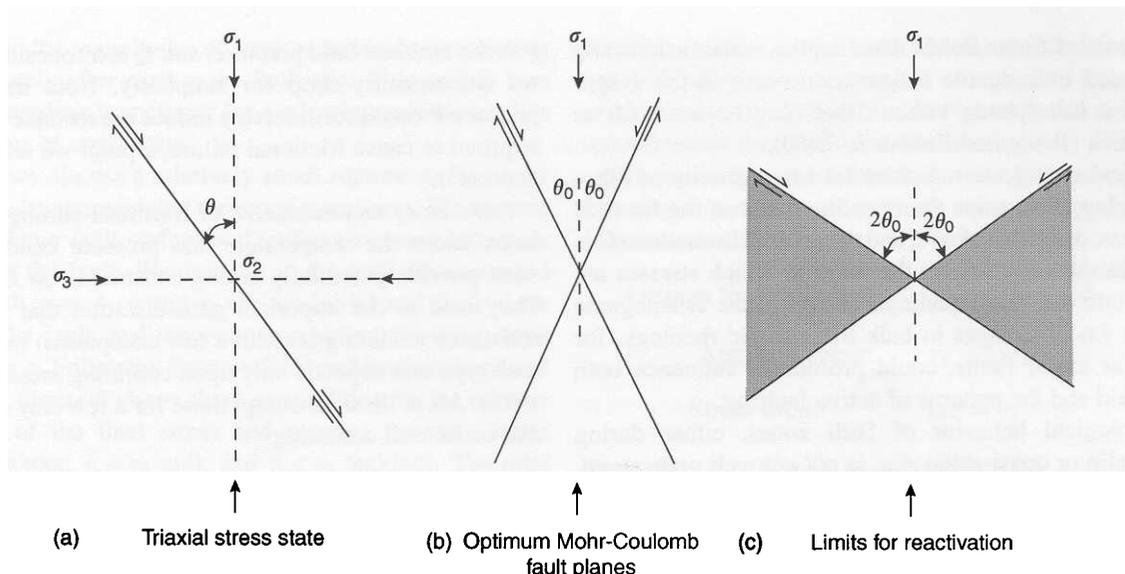
FIGURE 3 2D Mohr–Coulomb faulting theory. (a) Possible fault planes under triaxial stress state ($\sigma_1 > \sigma_2 > \sigma_3$). (b) Optimum failure planes as given by Eq. (12) in the text. (c) Frictional limits on fault plane orientation. Arrows show orientations of maximum (σ_1) and minimum (σ_3) principal stresses. The intermediate principal stress axis (σ_2) lies on intersection of other two axes and is perpendicular to them. Conjugate planes with opposite senses of slip are shown in each case.

elevated principal stress magnitudes (Rice, 1992). It is also likely that if a fluid phase were present, hydrofracture would keep σ_3' from becoming negative, thereby limiting the ambient stress and promoting lockup.

If we assume the frictional faults are optimally oriented according to Eq. (12) and assume a value for the friction coefficient μ of 0.75 we may use Eq. (11) to relate the magnitudes of the maximum and minimum principal effective stresses. At frictional equilibrium we obtain

$$\sigma_1' = 4.0\,\sigma_3'$$ (13)

5.1.2 Andersonian Faulting Types and Strength Versus Depth

In the Earth one of the principal stresses is usually assumed to be vertical. This condition must apply at the surface, where shear stresses vanish, and experimental data from mines and boreholes indicate it is a good generalization to at least 3 km depth (McGarr and Gay, 1978). Then, in the 2D faulting theory of Anderson (1951) the style of faulting depends upon which of the principal effective stresses is vertical. If the maximum effective principal stress σ_1' is vertical, the minimum σ_3' is horizontal and the failure planes are normal faults. When the minimum principal stress σ_3' is vertical, thrust faulting results. If the intermediate principal stress σ_2' is vertical, σ_1' and σ_3' are horizontal and faulting is strike slip.

Given this geometry of the principal stresses and assuming a value for the coefficient of friction, we can determine the

orientations of thrust, normal and strike-slip faults that will slip at the lowest differential stress ($\sigma_1'-\sigma_3'$). Assuming $\mu = 0.75$, Eq. (12) can be used to obtain $\theta_0 = 27°$. Thus, optimally oriented thrusts will dip at 27°, normal faults at 63°, and strike-slip faults will be oriented 27° to the direction of maximum compression.

The vertical effective principal stress σ_v' can be estimated independently, permitting the other principal effective stress to be calculated from Eq. (13). σ_v' is simply given by the weight of the overburden reduced by the ambient pore pressure,

$$\sigma_v' = \rho g z - P$$ (14)

where ρ is the density of the overburden and z is depth. In a fluid-saturated crust it is often assumed that pores or cracks are interconnected and a part of the rock column is supported by the fluid pressure $P = \rho_w g z$, where $\rho_w = 1000\,\text{kg m}^{-3}$ is the fluid density. For $\rho = 2700\,\text{kg m}^{-3}$, the gradient of σ_v' is $26.5\,\text{MPa km}^{-1}$ under dry conditions and $16.7\,\text{MPa km}^{-1}$ if pore pressure is hydrostatic.

For each of the main faulting types we can thus use Eqs. (13) and (14) to compute the gradient in resisting shear stress $\tau = 0.5\,(\sigma_1'-\sigma_3')\sin 2\theta_0$ on optimally oriented faults under hydrostatic and dry conditions.

Normal:

$$\sigma_1' = \sigma_v \qquad \text{Hydrostatic, dry gradients} = 5.0,\ 8.0\,\text{MPa km}^{-1}$$

(15a)

Strike slip:

$$\sigma_2' = \sigma_v' \quad \text{Hydrostatic, dry gradients} = 8.0, 12.7\,\text{MPa}\,\text{km}^{-1}$$
$$(\text{strike-slip case assumes}\; \sigma_v' = 0.5(\sigma_1' + \sigma_3'))$$
$$\text{(15b)}$$

Thrust:

$$\sigma_3' = \sigma_v' \quad \text{Hydrostatic, dry gradients} = 20.0, 31.8\,\text{MPa}\,\text{km}^{-1}$$
$$\text{(15c)}$$

For normal and strike-slip faults extending from the surface to 15 km depth, the average resisting stresses under hydrostatic conditions are thus 38 and 60 MPa. For a thrust fault extending to 30 km, a common earthquake nucleation depth in subduction zones, the corresponding average value is 300 MPa.

Recalling our previous discussion on plate driving and resisting forces, we see that if resisting stresses of this magnitude actually occur on plate boundary faults they will play an important role in determining the force balance of the plates and the intraplate "tectonic" stress field of the lithosphere. However, such high resisting stresses would lead to significant frictional heat generation. Section 6 shows how heat flow measurements sharply constrain the amounts of possible frictionally generated heat along both the San Andreas fault in California and on the Cascadia subduction megathrust offshore of Washington, Oregon, and British Columbia.

5.2 Lower Crust and Lithospheric Mantle Rheology

Above temperatures ranging from 300° to 450°C both laboratory results and field observations show that the rheological behavior of rocks alters drastically. Pressure-sensitive fracture strength and frictional fault slip yield to bulk ductile flow controlled by dislocation creep, a very temperature- and strain rate-sensitive process that is also very dependent on rock type. However, strength in the ductile field is largely independent of both lithostatic pressure and pore pressure. For this deformation mechanism, the strain rate ε and differential stress $(\sigma_1 - \sigma_3)$ are related by an expression of the form

$$\varepsilon = A\,(\sigma_1 - \sigma_3)^n\,\exp[-Q/RT] \qquad (16)$$

or equivalently

$$(\sigma_1 - \sigma_3) = (\varepsilon/A)^{1/n}\,\exp[Q/nRT] \qquad (16a)$$

where A, n and Q are constants that depend on rock type, R is the universal gas constant $(8.316\,\text{J}\,\text{mol}^{-1}\,\text{K}^{-1})$, and T is absolute temperature. For many rock types $n{\sim}3$, implying $(\sigma_1{-}\sigma_3)$ increases by a factor of 2 for each eightfold increase in strain rate. Since temperature normally increases with depth, Eq. (16a) shows that for each rock type $(\sigma_1{-}\sigma_3)$ is greatest at

the shallowest depths for which the flow law applies and thereafter decreases exponentially with increasing depth.

The exponential dependence of differential stress on rock type and temperature demonstrates the great influence of these factors on ductile strength. For example, laboratory experiments on ductile flow of dry quartzite, representing quartz-rich crustal rocks, yield $Q = 190\,\text{KJ}\,\text{mol}^{-1}$, whereas similar results for olivine, an upper mantle constituent, give $Q = 520\,\text{KJ}\,\text{mol}^{-1}$. Thus at the same temperatures and strain rates the ductile strength of these two rock types will differ by many orders of magnitude and there may be a significant strength contrast between the lower crust and uppermost mantle in many regions. Similarly, a change of just 100° in temperature of the lower crust or upper mantle will change ductile strength by a factor of 10.

5.3 Composite Quasi-static Strength Profile

If we consider lithospheric rheology to be controlled only by the frictional and ductile processes discussed above, the strength envelope for the crust is then determined by the lowest shear stress at a given depth that satisfies Eq. (9) or Eq. (16a). The depth at which both are satisfied is called the brittle/ductile transition. This image of lithospheric strength is certainly oversimplified, because there is a significant depth range over which a transitional, semibrittle behavior is likely to be important (Kohlstedt *et al.*, 1995) and frictional (i.e., pressure-sensitive) stable sliding may well occur below the seismogenic zone. Recent work suggests that the sharp peak in strength predicted at the brittle–ductile transition (see Fig. 4) may be blunted and modestly decreased by semibrittle processes (e.g., Kohlstedt *et al.*, 1995, Fig. 7, 9). Although these complexities may be important in the earthquake nucleation process, they are likely to be confined to a small fraction of the crustal column and the simple partition into frictional and ductile domains remains useful and intuitively instructive (see also Chapter 29 by Sibson).

Figure 4 shows strength versus depth, plotted as $(\sigma_h-\sigma_v)$ for thrust and normal faulting. Dry and hydrostatic pore pressure gradients are shown for each frictional sliding case. Pore pressure is parameterized by λ the ratio of the pore pressure P to the total vertical stress σ_v. Ductile strength curves are shown for both dry quartzite and olivine, a strain rate of $10^{-15}\,\text{sec}^{-1}$ and a geothermal gradient of $15°\,\text{km}^{-1}$ (appropriate for old oceanic lithosphere or stable continental interiors) are assumed.

What general conclusions can be drawn from composite strength profiles like Figure 4? Provided our assumptions that (1) upper crustal rocks are fractured and frictional resistance to slip on optimally oriented faults limits ambient stresses there, and (2) temperature-sensitive creep properties of rocks rich in quartz and olivine determine limiting stresses at greater depths, then lithospheric stresses will lie within the bounds shown in Figure 4. A region of high strength is then expected
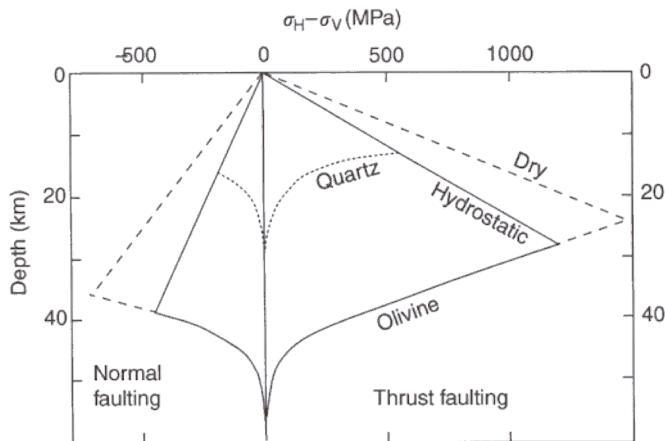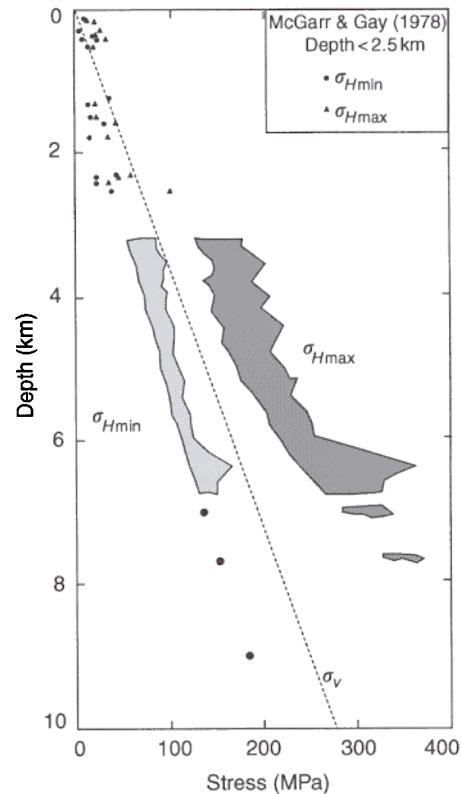
FIGURE 4 Composite strength profile for the continental cratonic lithosphere. Difference between maximum and minimum horizontal and vertical stress versus depth. If λ = ratio of pore pressure to vertical (overburden) stress, dry corresponds to $\lambda = 0$ and hydrostatic to $\lambda = 0.37$. Both quartz and olivine rheologies are used for ductile lithosphere and assumed thermal gradient is $15°C \, km^{-1}$. A strain rate of $10^{-15} \, s^{-1}$ is assumed.

FIGURE 5 Observations of maximum and minimum horizontal *in situ* stress versus depth. Determinations are from mines in southern Africa at depths shallower than 2.5 km (McGarr and Gay, 1978) and from below 3 km in the KTB scientific borehole in southern Germany (Brudy *et al.*, 1997), both in stable continental interiors. The range of possible values from KTB are shown shaded.

in the midcrust, near the transition between the frictional and ductile fields, and possibly also in the uppermost mantle, where a compositional change from quartzo-feldspathic rocks to more basic, olivine-rich compositions is expected. Depending on many factors this region of high strength lies in the depth range 15–40 km, with much lower ambient values above and below this depth interval. Depending on composition, temperature, and to a lesser extent strain rate, it is also possible that the ductile lower crust has a lower strength than either the overlying crust or the underlying upper mantle. For example, the strength curve for quartz (QTZ) in Figure 4 shows a sharp decrease in the 15–20 km depth range that is caused by the normal increase in temperature with depth and the exponential dependence of differential stress on temperature in Eq. (16a). Therefore, in some regions ductile flow may be concentrated in this low strength layer. However, it should be noted that quartz content generally diminishes in the lower continental crust.

5.4 Observational Constraints on Quasi-static Strength Profile

Direct measurements of stress in active fault zones at seismogenic depths have not yet been made. For constraints we must presently rely largely on measurements in plate interiors and on indirect estimates based on the effects of fault zone stress. *In situ* horizontal stresses measured in deep level mines to depths of 3 km (McGarr and Gay, 1978) and from the KTB deep scientific borehole drilled to 8 km depth in Germany (Brudy *et al.*, 1997) are plotted versus depth in Figure 5. The results from the KTB hole in particular support the expectations

based on the frictional strength profiles, and the gradient of horizontal differential stress of 14–$19 \, MPa \, km^{-1}$ (not shown) lies between the strike slip and thrust gradients for hydrostatic pore pressure.

The results of Figure 5 support the "rock mechanics" strength model for frictional slip in the regions sampled, but several indirect observations suggest major faults are considerably weaker. We discuss heat flow constraints in detail in the next section but first consider an independent measure of relative fault strength.

This indicator of fault strength, first suggested by Mount and Suppe (1987) and Zoback *et al.* (1987), is the orientation of maximum horizontal compressive stress direction inferred from earthquake fault plane solutions and borehole elongations in the blocks adjacent to major fault zones. As shown in Figure 3a and Eq. (12), the angle, θ, between fault strike and the σ_1 axis should be about 27° for an expected fault friction coefficient $\mu \sim 0.75$. However, near both major strike-slip faults and subduction zones $\theta \sim 50°$–$90°$, suggesting that the shear strength of these faults is much less than that of the adjacent crust (if the fault were a free surface, supporting no

shear stress, two principal stresses would lie in its plane and the third would be perpendicular to the fault). Detailed profiles of θ versus distance from the San Andreas fault recently obtained by Hardebeck and Hauksson (1999) from fault plane solutions show a minimum of $\theta \sim 45°-70°$ at the fault and generally support previous observations from strike slip faults and subduction zones. However, the results of Hardebeck and Hauksson also reveal considerable along-strike and strike-normal variability in southern California.

In a related observation, Thatcher and Hill (1991, 1995) have shown that $M\,5.8-7.3$ normal faulting earthquakes have dips that are strongly clustered near 45° (see Chapter 29 by Sibson, Fig. 6). Anderson's faulting theory and Eq. (12) suggest values should be centered near 60°, and the lower value suggests low friction coefficient μ or high pore pressure (or both) in these normal fault zones.

5.4.1 Assessment

The results in Figure 5 indicate that for inactive plate interiors the rock mechanics strength profile applies, and the brittle upper crust supports substantial shear stress. However, horizontal compressive stress orientations show that major fault zones support considerably lower stress, and the clustering of normal fault dips near 45° suggests these faults may also be weak.

6. Heat Flow and Frictional Stresses Resisting Fault Motions at Major Plate Boundaries

As we discussed in Section 4, stresses that resist fault motions generate heat, and measurements of surface heat flux near active faults can thus be used to infer bounds on the magnitudes of these resisting stresses. Using Eq. (2) the rate of work Q done against resisting stress τ_r for average fault slip rate v is $\tau_r v$ per unit fault area, or

$$Q = \tau_r v \tag{17}$$

No significant heat sinks are known (see Lachenbruch and Sass, 1980, p. 6186 for one justification), so the thermal effects of dissipative heating could be substantial. For example if $\tau_r = 100\,\text{MPa}$ and $v = 30\,\text{mm y}^{-1}$, then $Q = 96\,\text{m Wm}^{-2}$. We will discuss below how such a thermal flux at depth is conducted to the Earth's surface, but clearly sources of this magnitude would contribute significantly to observed heat flow and might violate observational constraints. For example, typical surface heat flow values above subduction thrusts average only about $40\,\text{m Wm}^{-2}$ (Hyndman and Wang, 1993), and in a 50–100 km wide region near the San Andreas fault are only 30–50 m Wm^{-2} above background levels (Lachenbruch and Sass, 1980).

6.1 Resisting Stresses on the San Andreas Fault

The vertical strike-slip geometry of the San Andreas fault makes it particularly straightforward to detect any surface heat flux due to dissipative heating at depth on the fault. About 100 heat flux measurements have been made by the US Geological Survey in the vicinity of the fault, and they provide a uniquely detailed picture of the thermal regime near a major active fault. The data possess considerable scatter. Nonetheless, they show a zone of high heat flux averaging $\sim 80\,\text{m Wm}^{-2}$, called the Coast Range Anomaly (CRA) by Lachenbruch and Sass (1980). It is spread over a region 50–100 km wide spanning the San Andreas fault zone, and heat flux is 40–50 m Wm^{-2} above background values to the east.

An important feature of the data is that it applies to all of the San Andreas fault system from its inception at the Mendocino triple junction on the north to the "big bend" of the fault 700 km to the south. It thus includes both the seismogenic, predominantly locked segments of the northern San Andreas fault system, which has several active strands, and the central $\sim$160-km long creeping segment of the San Andreas, where nearly all of the motion currently occurs as aseismic slip on a single fault. In neither region is there any obvious local anomaly of the kind expected from dissipative heating on the upper crustal portions of the fault.

To understand the form of this expected thermal anomaly, consider the simple 2D model of dissipative heat generation on the fault shown in Figure 6a. For resisting stress τ_r and slip velocity v acting over depth range d, the rate of heat generation q per unit length of fault is given by

$$q = \tau_r v d \tag{18}$$

Provided the depth of this source, a, is large compared to its width d, following Brune et al. (1969) we may write the surface heat flux $Q_0(x, t)$ due to conductive heat transfer as a function of distance x from the fault and time since initiation of slip,

$$Q_0(x,t) = q/\pi[a/(a^2 + x^2)]\exp[-(a^2 + x^2)/4\kappa t] \tag{19}$$

where κ is the thermal diffusivity ($10^{-6}\,\text{m}^2\text{s}^{-1}$).

Assume for the moment that the resisting stress increases roughly linearly with depth, as suggested by Figures 4 and 5, and is greatest near earthquake nucleation depths, about 10 km on the San Andreas fault. The time-dependent exponential term in Eq. (19) will then be negligible after a few million years of slip. The steady-state surface heat flux will have a maximum value $q/\pi a$ at the fault, decrease to half this value at $x = a$, and thereafter decrease rapidly with increasing distance. Therefore if we take $v = 30\,\text{mm y}^{-1}$ and $\tau_r = 80\,\text{MPa}$ acting over a depth range of 5 km centered at depth $a = 10\,\text{km}$, we easily find that the maximum surface heat flow anomaly (i.e., excess over background) is $40\,\text{m Wm}^{-2}$ and it decreases

(a)

(b) Distance from San Andreas fault (km)
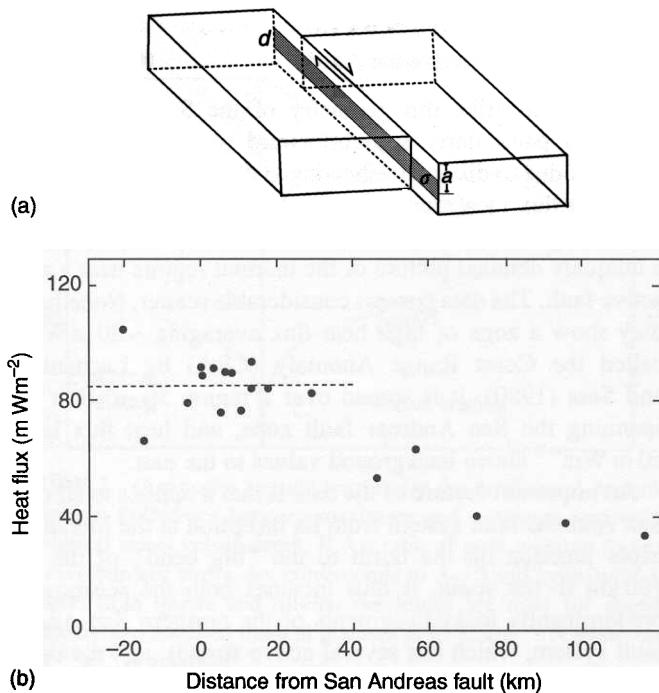
FIGURE 6 (a) Model geometry for heat generation on a vertical strike-slip fault. Resisting stress σ acts on a fault strip of width d centered at depth a below the ground surface. (b) Heat flux versus distance from central San Andreas fault (data from Region 3 of Lachenbruch and Sass, 1980). Mean heat flux within about 40 km of the fault, $86\,\mathrm{m\,Wm^{-2}}$, is shown by dashed line. This "Coast Range Anomaly" (CRA) lies $40{-}50\,\mathrm{m\,Wm^{-2}}$ above background values $\sim$100 km east of the San Andreas fault.

rapidly (with scale length a) away from the fault. No such local heat flow anomaly is observed in the data (Fig. 6b). Indeed, individual heat flow profiles across the fault, like that in Figure 6b, fail to show any anomaly larger than the noise in the data, $5{-}10\,\mathrm{m\,Wm^{-2}}$, so an upper bound on frictional resistance of about 10 MPa is indicated.

Assuming uniform or linearly increasing distributions of resisting stress with depth (Brune *et al.*, 1969; Lachenbruch and Sass, 1980) makes the analysis more precise. However, unless the dissipative heating is all concentrated below $\sim$10 km depth, the resolvable features of the computed surface heat flow profile do not change and the main conclusion is unaltered. Average resisting stress on the seismogenic portion of the San Andreas fault is thus no greater than about 10 MPa, about a factor of five less than suggested by the extrapolation of laboratory-based estimates of frictional fault strength (Fig. 4).

If frictionally generated heat were concentrated at depths below 15 km, the anomaly at the surface would be considerably broadened, and less obvious. A possibility (Thatcher and England, 1998) is that the CRA (coast range anomally), the regionally high heat flux in Figure 6b, is caused by

distributed shear heating due to ductile flow beneath the depths where frictional processes dominate. As the strength profiles of Figure 4 suggest, and Thatcher and England (1998) show, ductile shear heating is concentrated near the ductile strength maximum at the brittle–ductile transition. The resulting heat source thus behaves like a line source concentrated at this depth and its effects can be approximated well by Eq. (19). The surface heat flux distribution predicted by this model is relatively broad and decreases away from the fault with a scale length equal to the depth of the brittle–ductile transition. The ductile shear zone model can generally account for the main features of the CRA (see Thatcher and England, 1998, Figs. 14, 15, 16).

Another proposed explanation of the CRA is that it results from advective transfer of frictionally generated heat by fluid flow away from the fault (Hanks, 1977; Williams and Narasimhan, 1989; Scholz, 2000). In this hypothesis, resisting stresses on the fault could be as high as given in Eq. (15b), with the CRA being smeared out by convective flow of groundwater away from the fault.

6.2 Resisting Stresses on Subduction Thrusts

At subduction zones the geometry of faulting and several potential sources and sinks of heat influence the surface heat flux and affect our ability to distinguish effects of dissipative heating on the fault. These features have been considered in detail with sophisticated numerical models of heat conduction, but the essential features can be intuitively understood and resisting stresses bounded by a simple analytic model of the thermal regime presented by Molnar and England (1990).

Recall from frictional faulting theory that resisting stresses on thrusts could be *very* high (the shear stress gradient for hydrostatic pore pressure is $\sim$20 MPa km^{-1}, implying 600 MPa at 30 km depth!). If shear stress on subduction thrusts was this high it would strongly resist plate motions. Furthermore, the thermal effects of such dissipative heating would be enormous. For example, if $v = 40\,\mathrm{mm\,y^{-1}}$ and resisting stress was even just half the value suggested by Andersonian faulting theory, then $Q \sim 400\,\mathrm{m\,Wm^{-2}}$ of fault surface! We will discuss below how such a thermal flux at depth would be conducted to the Earth's surface, but clearly sources of this magnitude would significantly violate observed heat flow constraints. (Note that the heating required to generate arc magmas occurs 100–200 km farther landward and at depths of $\sim$80 km in the slab.)

The geometry of the model is shown in Figure 7a. A plate with dip δ is subducting at velocity v. Its motion is resisted by shear stress τ_r, which may be a function of the depth to the fault plane Z_f. Three sources of heat are important in determining the temperature above the slab and conduction to the Earth's surface: (1) resistive shear heating, as given by Eq. (17), acts on the subduction thrust; (2) the heat Q_0 from the descending slab; and (3) heat generated by radiogenic decay in
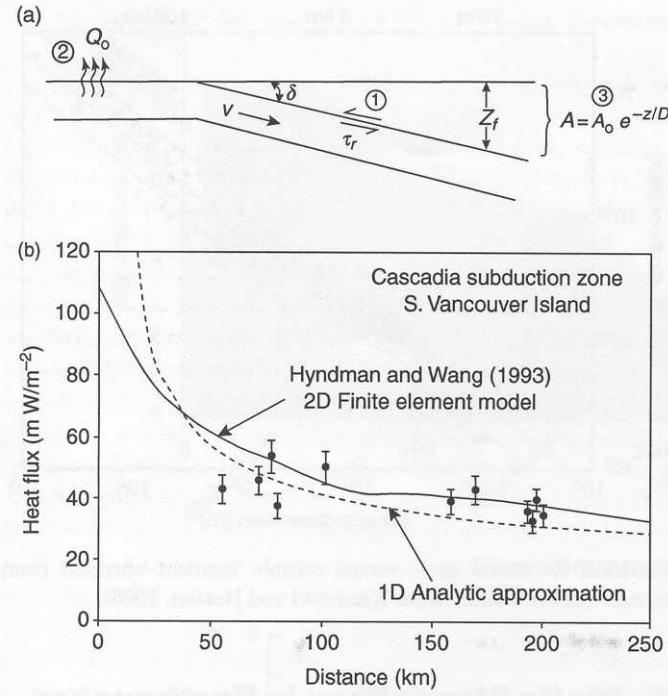
FIGURE 7 (a) Geometry and heat sources of subduction thrust fault model. 1, Shear heating on interplate thrust; 2, heat from subducting oceanic slab; 3, heat from upper plate radiogenic decay. See text for definition of symbols. (b) Cascadia heat flux versus distance from trench axis compared with 1D model discussed in text and with 2D heat finite element model of Hyndman and Wang (1993).

the overlying plate. The advective downward transport of heat by the descending slab must be accounted for in deriving the effect of these three heat sources on the surface heat flux. However, Molnar and England have shown that for gentle slab dips ($\delta < 30°$) it is possible to correct for this advection, ignore the effect of lateral heat conduction in the overlying plate, and obtain 1D analytic expressions for the temperature field. The heat flux at any point on the surface a distance Z_f above the slab is then given by

$$k\frac{dT}{dz} = \frac{\tau_r \nu}{S} + \frac{Q_o}{S} + A_o D[1 - e^{-Z_f/D}] \qquad (20)$$

where the advective correction term, S, is

$$S = 1 + \sqrt{\frac{Z_f \nu \sin\delta}{\kappa}} \qquad (21)$$

and κ is the thermal diffusivity. The last term in Eq. (20) represents the contribution to surface heat flux of radioactive heat production A_o that decreases exponentially with depth on a scale length D.

Using the Cascadia subduction zone as an example we can show that any contribution to surface heat flux from resistive shear heating must be quite small. The following parameters

are taken as representative for Cascadia (Hyndman and Wang, 1993):

Subduction velocity, $v = 45\,\text{mm}\,\text{y}^{-1}$
Fault dip, $\delta = 15°$
Heat flux, Oceanic flab, $Q_o = 120\,\text{m}\,\text{Wm}^{-2}$
Crustal radioactive concentration, $A_o = 0.6\,\mu\,\text{Wm}^{-3}$
Decay depth, Crustal radioactivity, $D = 10\,\text{km}$
Thermal diffusivity, $\kappa = 10^{-6}\,\text{m}^2\,\text{s}^{-1}$

Hyndman and Wang (1993) show that the surface heat flux at a distance $Z_f = 15\,\text{km}$ above the subducting Juan de Fuca slab is about $50\,\text{m}\,\text{Wm}^{-2}$. For the parameters listed above the advective correction term $S = 3.3$ and the surface flux contribution from the slab is $36\,\text{m}\,\text{Wm}^{-2}$ and from crustal radioactivity is $6\,\text{m}\,\text{Wm}^{-2}$. Thus the last two terms on the right-hand side of Eq. (20) account for most of the observed heat flux, leaving no more than about $8\,\text{m}\,\text{Wm}^{-2}$ to be allocated to shear strain heating. Clearly resisting stresses (either quasi-static or dynamic) on the Cascadia megathrust must be quite small. Figure 7b shows a sample comparison between model and data from southern Vancouver Island. For illustrative purposes we have taken a modest resisting stress gradient $\tau_r = 0.05\,\rho g\,Z_f\text{MPa}$, but this contributes no more than $10\,\text{m}\,\text{Wm}^{-2}$ to the surface heat flux. This figure shows that resisting stresses on the Cascadia megathrust are quite low and, as is the case for the San Andreas fault, any thermal contribution is not distinguishable above the noise levels of the heat flow data.

Similar results have been obtained in other subduction zones. Hyndman *et al.* (1995) use the extensive suite of heat-flow data from the Nankai Trough subduction zone of southwest Japan to demonstrate resisting stresses there are as low as in Cascadia. A synoptic review by Hyndman and Wang (1993) of heat-flow data from other circum-Pacific subduction zones suggests the same conclusion (see also Tichelaar and Ruff, 1993). In particular, in Northern Honshu, landward of a segment of the Japan Trench that is subducting predominantly aseismically (Pacheco *et al.*, 1993), no anomalous heat generation attributable to frictional resistance was observed.

6.3 Assessment

We suggest that the heat flux measurements from the San Andreas transform and the Cascadia subduction zone provide strong evidence that stresses resisting slip on major faults are relatively low, on average $\sim$20 MPa or less on subduction thrusts and $\sim$10 MPa or less on the San Andreas. Since this estimate is an average over the seismogenic fault plane, values at any point could be considerably higher (or lower) than this mean. In particular, these resisting stresses could locally be higher on the deeper portions of the fault, at the depths where most large and great earthquakes nucleate.

Despite suggestions that advective transport of frictionally generated heat away from active faults is a quantitatively

important process, the preponderance of evidence supports a conductive model of heat transfer near major faults. On the San Andreas fault the integrated heat energy represented by the CRA is about what would be generated by frictional heating on a fault stressed according to Eq. (15b). With a viable mechanism for distributing this locally generated heat across the California Coast Ranges, the CRA might be explained. Williams and Narasimhan (1989) suggest fluid flow induced by the hydraulic gradients due to high topography along parts of the San Andreas, but this mechanism would seem to apply only to restricted segments of the fault. Lachenbruch and Sass (1980, pp. 6196–6198) give qualitative arguments against vigorous hydrothermal circulation near the San Andreas, noting that the integrated output of thermal springs along the fault is orders of magnitude too small to be due to frictional heating from a fault obeying Eq. (15b). However, they do concede that heat transport at low flow rates and moderate temperatures could be significant but very difficult to detect. There is no regionally high heat flow above the upper ~40 km of subduction thrusts (see Fig. 7b). Thus, regardless of heat transfer mechanism, it seems difficult to argue for the frictional heat generation implied by Eq. (15c).

In several regions, the heat flow constraint on frictional resistance applies to major faults that currently slip aseismically. Although it is not possible to be certain that the current aseismic behavior is typical of long-term patterns, the similar behavior of seismogenic and creeping faults suggests the mechanism responsible for weakening these faults may operate for both quasi-static and dynamic slip.

The role of heat generation by ductile shearing is uncertain. Although this mechanism can account for the heat-flow patterns observed in California, the slab window model of Dickinson and Snyder (1979) is at least as successful. In this model, the Juan de Fuca slab north of the Mendocino triple junction is subducting at a shallow angle and the process of triple junction migration thus exposes hot upper mantle at the base of the crust adjacent to the newly created San Andreas fault. Lachenbruch and Sass (1980) showed that this model explains the heat flow data and also the age progression of young volcanics southeast of the Mendocino triple junction. In addition, it matches the observation that the CRA is not symmetrically centered on the San Andreas, as required by models (like ductile shearing) that localize heat sources on a single fault or its downward continuations. However, the extent and history of exposure of the slab window is uncertain, and this heating mechanism may not be generally applicable throughout California.

Ductile shearing has not yet been modeled in subduction settings. However, reference to the 1D model (Yuen et al., 1978) and relevant heat flux data (Fig. 8) suggest the existence of ductile shear heating below ~40 km depth would be difficult to prove (or disprove) from surface measurements.
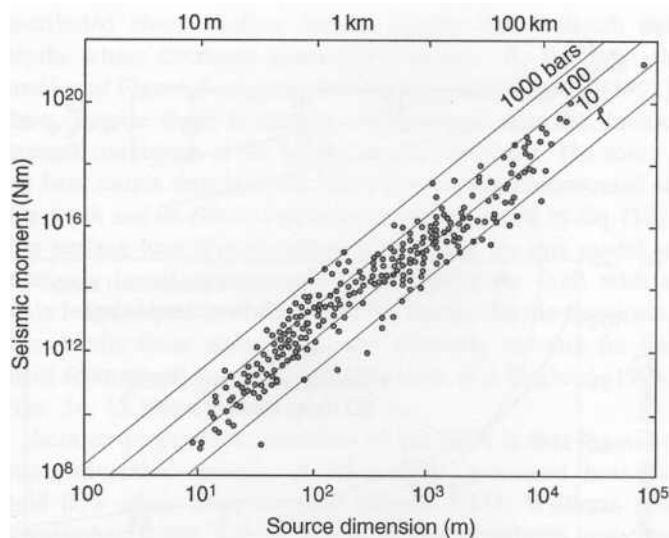

FIGURE 8 Stress drop versus seismic moment obtained from seismic data (modified from Kanamori and Heaton, 2000).

7. Static Stress Drop in Earthquakes

As mentioned in Section 3, observations of the average static stress change due to earthquake faulting supply one of the important stress measures needed to understand fault energetics. Geodetic and seismologic methods can be used to estimate this parameter and its variations with position on the earthquake rupture. The average shear stress drop is given by

$$\Delta\sigma = Gu/L \tag{22}$$

where G is an elastic rigidity modulus, u is fault slip and L is a fault dimension. For small and moderate magnitude earthquakes, $\Delta\sigma$ is measured using the amplitude spectra of seismic body waves (Brune, 1970). For larger events, $\Delta\sigma$ is obtained using independent estimates of earthquake rupture dimensions from aftershock zone size and seismic moment obtained from long period seismograms. Geodetic measurements can be used to measure the strain change (essentially u/L) near the fault rupture, which is typically about 10^{-4}. Multiplied by an elastic modulus typical for the Earth's upper crust (~30 GPa), Eq. (22) yields $\Delta\sigma \sim 3$ MPa.

Figure 8 shows average stress drop $\Delta\sigma$ over a wide range of earthquake moment M_0 and equivalent magnitude, M_W. This compilation shows that $\Delta\sigma$ ranges from about 0.1 to 10 MPa and is essentially independent of magnitude (the rather lower values at smaller magnitudes may indicate a bias related to attenuation of high frequency seismic waves; see Hanks, 1982). Figure 8 also shows that some values as high as 100 MPa are occasionally observed.

The stress drop values given in Figure 8 are averaged over the entire earthquake rupture. However, both seismological and geodetic methods can be used to infer the spatial
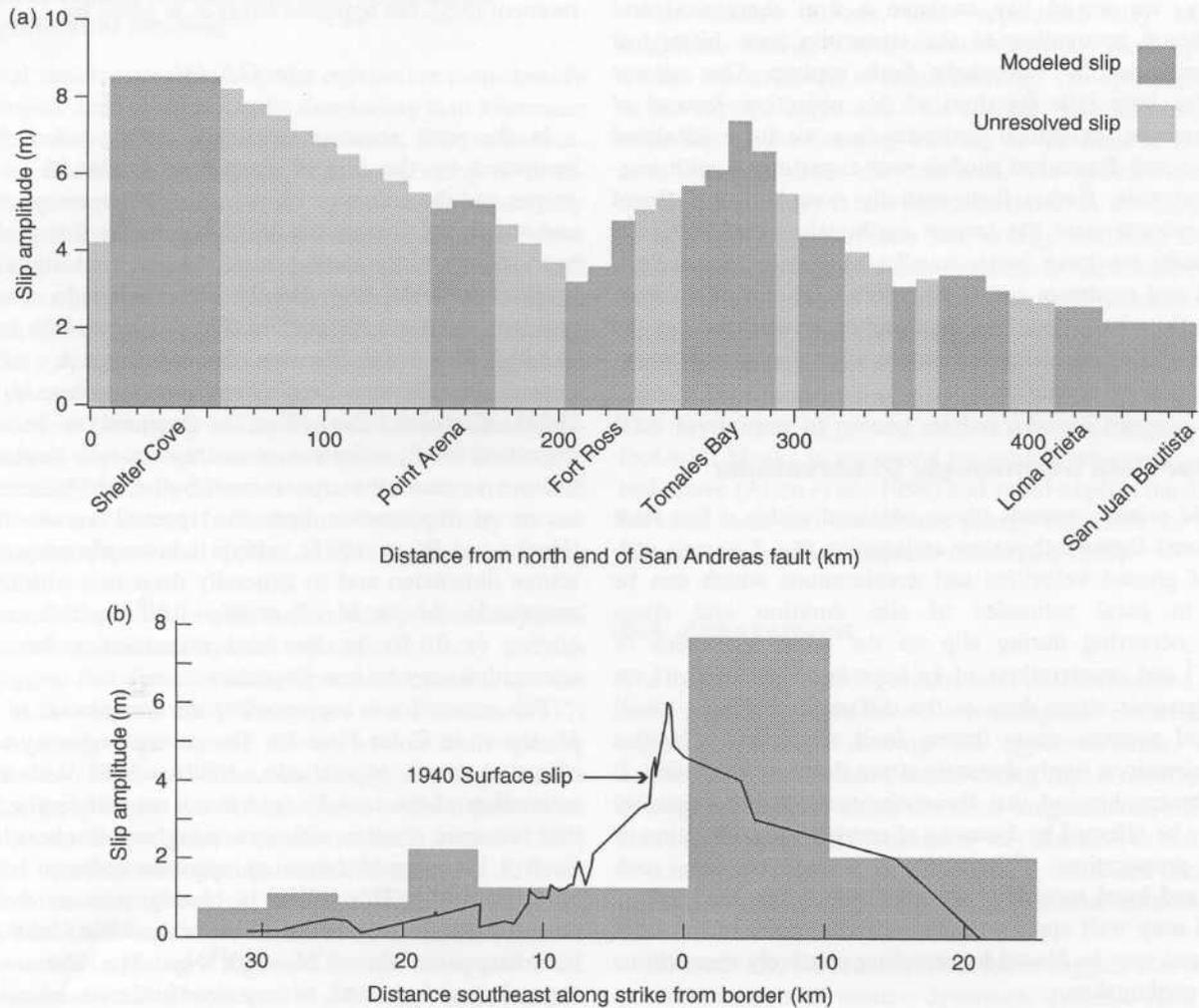
FIGURE 9 Slip distributions in (a) the 1906 $M = 7.8$ San Francisco earthquake (Thatcher *et al.*, 1997) and (b) the 1940 $M = 7.1$ Imperial Valley, California earthquake (King and Thatcher, 1998).

distribution of seismic slip and hence map static stress drop variations on the earthquake fault plane. Two examples, for the great $M = 7.8$ 1906 San Andreas earthquake and the 1940 $M = 7.1$ Imperial Valley earthquake, both strike-slip events with extensive surface faulting and measurements of surface fault offset, are plotted in Figure 9. Inevitable smoothing and nonuniqueness of the fine scale features of the derived slip distributions preclude estimation of the largest static stress changes on the fault plane. However, these results, as well as seismological determinations of slip distribution (e.g., Archuleta, 1984; Cohee and Beroza, 1994; Wald and Heaton, 1994) indicate considerable heterogeneity in stress drop on earthquake rupture planes.

7.1 Assessment

The results summarized in Figure 8 reliably bound average stress drop, which has a mean value of $\sim$3 MPa and ranges from 0.1 to 10 MPa. However, the along-strike and depth variations of slip indicated by slip mappings like Figure 9 show that stress release is quite heterogeneous, with stress drop varying by at least a factor of 5 on the fault.

8. Dynamics of Earthquake Faulting

The occurrence of earthquakes is the result of sudden release of elastic stress stored in blocks adjacent to the fault, and while slip is in progress inertial forces are very large, probably considerably larger than the quasi-static surface forces (stresses) which initially led to rupture nucleation and fault failure. Thus we anticipate that understanding the dynamics of earthquake slip may be crucial to understanding fault energetics, and inferences drawn from quasi-static analyses and observations may be a poor guide to dynamic behavior.

Ideally, we would like to have a firm theoretical and observational accounting of the stress-slip time history at every point on the earthquake fault rupture. Our current knowledge base falls far short of this objective. Instead of comprehensive theoretical understanding we have idealized kinematic and dynamical models and experiments with analogue materials. Rather than spatially complete, broadband seismic records near the largest earthquakes that occur on major faults we have local, usually bandlimited recordings of small and moderate events and distant seismograms from great earthquakes. Here we summarize these data and the inferences based on them and assess what they reveal about dynamic rupture processes.

8.1 Near-field Seismologic Observations

Near-field seismic records (those obtained within a few fault dimensions) from earthquakes as large as $M \sim 7$ permit estimates of ground velocities and accelerations which can be related to local estimates of slip duration and stress changes occurring during slip on the fault. Velocities of $2 \, \mathrm{m \, sec^{-1}}$ and accelerations of $1g$ have been observed. If we define dynamic stress drop as the difference between initial stress and average stress during fault slip, such velocities and accelerations imply dynamic stress drops of $\sim 20 \, \mathrm{MPa}$. It should be emphasized that these observations are relatively few, may be affected by focusing of energy in the direction of rupture propagation ("directivity"), source–receiver path effects, and local recording site conditions. The peak values obtained may well apply to only restricted parts of the fault rupture and may be biased by sampling relatively more thrust faulting earthquakes.

However, the occurrence of the 1999 $M_W = 7.6$ Chi-Chi (Taiwan) earthquake supplies a large new data set that will be very important in refining our views of earthquake dynamics. This event produced surface rupture over a 125 km-long fault with up to 8 m of reverse dip-slip motion. It was recorded by over 500 free-field strong-motion accelerographs, many of them located in the near field. Preliminary analysis of these records (Mori and Ma, 1999) shows ground velocities of $1–3 \, \mathrm{m \, sec^{-1}}$ near the fault, with clearly larger motions on the upper or hanging wall block of the fault (Ni *et al.*, 1999). These observations reinforce earlier field evidence from thrust faults showing motions larger than $1g$ and $1 \, \mathrm{m \, sec^{-1}}$ on the shallow part of the hanging wall, with associated evidence of relatively low accelerations and velocities on the adjacent foot-wall (Allen *et al.*, 1998).

8.2 Estimates of Radiated Energy and Apparent Stress

Although observations of near-field waveforms are rare, more observations exist of total seismically radiated energy (E_s). Using measures of E_s along with seismic determinations of

moment (M_o), the apparent stress τ_a is given by

$$\tau_a = G E_s / M_o \qquad (23)$$

In the past, accurate estimation of E_s and τ_a has been hampered by the limited frequency bandwidth of seismographs and the unknown effects of high frequency attenuation and scattering. During the past 15 years the first problem has been solved by the deployment of broadband digital seismographic networks both locally and worldwide. The second problem, unknown attenuation effects, may remain and appear as a high frequency saturation of corner frequency and marked dependence of stress drop on seismic moment, M_o (Hanks, 1982). However, this effect is expected to become less important for E_s estimates as earthquake size increases. This follows because the major contribution to radiated energy occurs at frequencies near the spectral corner frequency (Hanks and Wyss, 1972), which is inversely proportional to source dimension and so generally decreases with increasing magnitude. Above $M \sim 7$ or $M_o \sim 10^{19} \, \mathrm{Nm}$ this corner frequency is 0.1 Hz or less and uncertainties due to wave attenuation may be less important.

This contention is supported by the compilation of τ_a versus M_o shown in Color Plate 13. The corner frequency saturation effect at small M_o (Hanks, 1982) would lead to underestimation of the true E_s (and hence an underestimate of τ_a) that becomes smaller with increasing magnitude or moment. Such a bias would cause an apparent increase of τ_a with increasing M_o. This effect is clearly seen at the smaller moments for several of the data sets shown in Color Plate 13, but disappears above $M_o \sim 10^{18} \, \mathrm{Nm}$. The absence of any dependence of τ_a on M_o is very clear in the teleseismic data of Choy and Boatwright (1995), which show apparent stresses spanning their full range over moments varying from 10^{17} to $10^{20.5} \, \mathrm{Nm}$.

There is one observational bias that deserves mention. Some existing data suggest teleseismic estimates of radiated energy are significantly less than those based on limited near-field data, especially for thrust faults (e.g., Singh and Ordaz, 1994). More recently, analysis of the uniquely complete data for the 1999 Chi-Chi (Taiwan) earthquake by Ni *et al.* (1999) has confirmed this. They showed that teleseismic estimates E_s and τ_a determined by the method of Choy and Boatwright (1995) yielded an apparent stress value of 0.5 MPa while the near-field value was about 5.0 MPa. Because the near-field data from the 1999 earthquake records are so singular it is not possible to assess how general this bias may be. That the event occurred on a shallowly dipping thrust fault may be significant and is discussed further below. However, we do note that the apparent stress value for the 1999 earthquake lies within the range of values shown for large earthquakes in Color Plate 13. For these events τ_a is small, averaging $\sim 0.5 \, \mathrm{MPa}$ and ranging from 0.03 to 6.7 MPa.

8.3 Theoretical, Numerical and Analogue Dynamical Models

Dynamical simulations of earthquake rupture are considerably more complex and computationally demanding than kinematic modeling, and much yet remains to be learnt. Elastic dislocation theory is most commonly applied to modeling of earthquake waveforms. However, this approach is essentially kinematic and requires specification either of idealized point sources and rise times or adjustable rupture velocities and distributions of slip-time functions on the rupture surface (e.g., Archuleta, 1984; Cohee and Beroza, 1994; Wald and Heaton, 1994). Fully dynamical models are only now beginning to be applied to earthquake rupture (e.g., Boatwright and Cocco, 1996) and most commonly utilize slip-weakening rules developed for fault slip in rock mechanics laboratories (Dieterich, 1979; Ruina, 1983). Fully 3D models that treat both vertical and inclined faults and account for fault plane heterogeneity are just now being developed and applied to match observed strong motion data.

One simple theoretical result supported by laboratory analogue modeling deserves special note. Both theory and lab results suggest that dynamic fault slip is accompanied by fault opening at the rupture front. Comninou and Dundurs (1977) derived a steady-state rupture model with a fault-opening mode. Freund (1978) criticized this model, showing that the assumed singularities at the front and back of the rupture were physically unrealistic. However, more recently Adams (1999) has argued that the fault-opening mode can occur for different types of assumed singularities. In a related model, Andrews and Ben-Zion (1997) and Ben-Zion and Andrews (1998) have shown such fault opening is a general feature of rupture at the contact of two elastic half-spaces with differing elastic moduli.

Fault opening is commonly observed in laboratory physical analogue experiments. Anooshehpoor and Brune (1994) observed a steady-state propagating rupture with fault opening in a foam rubber physical model with a rough fault between both dissimilar and identical media. Anooshehpoor and Brune (1994) showed that this mode was associated with a strong reduction in the amount of frictional heat generation on the fault. The opening mode between identical media was evidently associated with asperity interactions that fed energy into the opening mode. Fault opening or nearly complete reduction in fault normal stress has also been observed in plastic models of slip between identical media (Brown, 1998; Bodin *et al.*, 1998; Bouissou *et al.*, 1998; Uenishi *et al.*, 1999).

Numerical models have also shown the existence of a fault-opening mode under some conditions. Mora and Place (1994) showed the existence of such a mode in a lattice numerical model with a rough interface. This mode disappears when the roughness is decreased to zero (Mora and Place, 1998; Shi *et al.*, 1998). Mora and Place (1999) have shown that introducing particles on the fault which are allowed to rotate dynamically might also reduce fault friction, in part by causing local fault opening.

For shallow thrusts, fault separation may be greatly enhanced. Brune (1996) observed strong fault opening in a foam rubber model of thrust faulting. The opening increased as the rupture approached the surface, resulting in a spectacular flip of the hanging wall tip of the fault at the surface as it detached from the foot-wall. The ground motions on the foot-wall were much less than those on the hanging wall ($\sim 1/5$), resulting in much less energy radiating downward (corresponding to teleseismic radiation in the Earth) compared to the energy trapped in the hanging wall (corresponding to energy typically recorded on near-field accelerograms). Shi *et al.* (1998) reproduced many of the features observed in the foam rubber model using a dynamic lattice numerical model. This asymmetry of ground motion between hanging-wall and foot-wall blocks is supported by seismic observations described above (Allen *et al.*, 1998) and could explain the difference between local and teleseismic energy estimates for the 1999 Chi-Chi (Taiwan) earthquake (Ni *et al.*, 1999).

8.4 Assessment

The dynamics of faulting are the least well understood aspect of fault mechanics. However, seismological observations are suggestive that *on average* the apparent stress, which is a measure of the dynamic stress during slip, is relatively small, averaging ~ 3 MPa, about the same magnitude as the average static stress drop. This estimate is very sensitive to the accuracy of seismic estimates of radiated energy, which are uncertain by at least a factor of 3 (H. Kanamori, personal communication, 1998).

Collectively, the studies cited above suggest that one of the possible aspects of rupture dynamics, dynamic fault opening or normal stress reduction, may play an important role in earthquake rupture and explain some of the more puzzling aspects of faulting.

9. Summary and Discussion

Although uncertainties and caveats leave room for doubts, the picture that emerges from our review is that plate interiors are strong, supporting shear stresses consistent with the frictional strength of rocks determined in the laboratory (see Townend and Zoback, 2000), whereas major fault zones along plate boundaries are much weaker, for reasons that are not yet well understood.

We may usefully summarize the average stress state near faults by recalling the analysis from Section 4 of stress and energetics at a single point on an idealized fault surface with uniform properties. We showed that the initial, pre-slip shear stress τ_i is

$$\tau_i = \tau_a + \tfrac{1}{2}\Delta\sigma + \tau_r \tag{8}$$

where τ_a is apparent stress, σ is stress drop and τ_r is stress resisting fault slip. The heat flow constraint (Section 6) gives $\tau_r < 20$ MPa; seismic and geodetic estimates (Section 7) yield $\Delta\sigma = 0.1$–10 MPa. These estimates argue that average values of τ_r and $(\tau_i - \tau_f)$ are small but [see Eq. (8)] leave open the possibility that initial stress τ_i could be large if τ_a were large. This might occur, if for example, the inertia of the accelerating fault blocks causes the resisting stress τ_r to overshoot the final stress τ_f, as sketched in Figure 2. However, such a process would lead to radiation of anomalous seismic energy (E_s) and be reflected in high values of apparent stress τ_a $(\tau_a = E_s/M_o)$. Estimates of radiated energy from seismograms (Section 8) give $\tau_a = 1$–10 MPa. These values then yield, using Eq. (8), an upper bound for initial stress τ_i of 20–40 MPa. This is consistent with an earlier upper bound estimate of 20 MPa obtained by Lachenbruch and Sass (1980) for the San Andreas fault.

This upper bound on initial stress averaged over the earthquake rupture plane has important implications for the strength of faults. At subduction thrusts this upper bound is much lower than the fault strength inferred from frictional slip experiments in the laboratory [Eq. (15c)]. It is marginally less than the same strength estimate for strike-slip faulting [Eq. (15b)]. At subduction zones, τ_i could only be as large as suggested by Eq. (15b) (~300 MPa average in the seismogenic upper crust) if τ_a were two orders of magnitude larger than observed. Therefore, the main weakening mechanism relative to the laboratory estimates, at least in the deeper parts of subduction zones, must be quasi-static. Constraints are poorer on the shallower parts of thrust faults and on strike-slip faults, leaving open the possibility that dynamic weakening could be more important there.

9.1 Weakening Mechanisms

The reasons for fault weakness has been debated for over 40 years, dating from the work of Hubbert and Rubey (1959) on the mechanics of slip on low-angle overthrust faults. Though many mechanisms have since been suggested, no single one has yet been universally accepted. Here we mention only some of the candidates we judge to be strong possibilities. Interested readers are referred to additional mechanisms mentioned briefly below and listed in the references section at the end of this chapter.

As shown by Eq. (9), very high fluid pore pressure in fault zones can, formally at least, overcome the effects of overburden pressure and permit slip at arbitrarily low shear stresses. This mechanism was originally suggested by Hubbert and Rubey (1959) to explain how large thrust sheets could be transported tens of kilometers when the weight of the overlying rock imposed frictional resisting stresses of several hundred megapascals or more on the fault. Through increasing elaboration over 40 years, the Hubbert–Rubey mechanism remains one of the leading candidates for reducing fault

strength. As applied to a strike-slip setting, Lachenbruch and Sass (1980) pointed out a possible limitation of the Hubbert–Rubey model, that near-lithostatic fluid pressure would produce pervasive hydrofracture normal to the least principal stress in the blocks adjacent to the fault. Rice (1992) pointed out that this difficulty could be overcome if very high fluid pressures were confined to the fault zone only, producing a fault that slips under very low shear stresses bounded by much stronger surrounding blocks where pore pressures were nearly hydrostatic. He suggested upward fluid flow into the fault zone from the ductile roots of faults, where pore pressure may be nearly lithostatic. However, Chery et al. (2000) have recently pointed out that unless the highly pressured fault zone is very narrow (~10–100 m wide), such a model implies large, super-lithostatic vertical stresses in the fault zone, leading to plastic extrusion of fault zone rocks toward the Earth's surface. Furthermore, it seems questionable that very high pore pressures could be permanently confined within the fault zone when periodic large earthquakes create fracturing that extends from seismogenic depths to the surface.

A related class of weakening mechanism relies on transient pore pressure changes accompanying earthquake rupture. Sibson (1973) and Lachenbruch (1980) suggested fluid hydraulically confined within the fault zone during earthquake slip could be sufficiently frictionally heated to create high pore pressure via thermal expansion and a phase transition from water to steam. Sibson (1990) later proposed that fluid pressures in the bottom half of the seismic zone could transiently cycle between lithostatic and hydrostatic as a result of fracture permeability created during slip and resealing of these fractures by post-slip precipitation ("fault valve" behavior), and he cited field evidence from exhumed high angle reverse faults to support this model.

Dynamic weakening could result if frictional heat generation were sufficient to actually melt fault zone rocks during slip (McKenzie and Brune, 1972; Kanamori et al., 1998; Kanamori and Heaton, 2000). This inference is supported by restricted occurrences of unusual fault zone rocks ("pseudotachylyte") interpreted as products of melting during slip (Sibson, 1975). For temperature rises to be sufficient to produce melting, the zone of earthquake slip must be very narrow (a few centimeters or less), which may occur in some fault zones. However, the relative rarity of pseudotachylytes generally argues for lower resisting stresses and temperature rises during fault slip.

Another class of weakening models includes dynamic inertial effects. As mentioned in Section 8, theoretical models and analogue experiments show fault opening during dynamical slip, and if this occurs at seismogenic depths in the Earth it would be a weakening mechanism of major importance.

Additional weakening mechanisms include: (1) low intrinsic friction coefficient clays and related rocks (Wu, 1978; Wang et al., 1979); however, note more recent laboratory experiments showing near-normal μ at upper crustal confining

pressure (Morrow *et al.*, 1992); (2) rotation of rounded fragments of fault gouge (Brune and Anooshehpoor, 1997; Mora and Place, 1998); (3) dynamic compaction generating high pore fluid pressure (Sleep and Blanpied, 1992); (4) acoustic fluidization of fault gouge during slip (Melosh, 1996).

9.2 Inhomogeneous Fault Strength and Its Implications

Thus far we have discussed models of fault strength that are either explicitly homogeneous or rely largely on observations like heat flux and radiated energy that are integrated averages of stress effects. However, much of what is known and understood about faults indicates they are inhomogeneous in their physical properties and their effects over a wide range of scales. Surface maps show faults are often discontinuous and change in orientation along strike, and seismic and structural evidence often shows these variations persist at depth. Mappings of the distributions of both coseismic slip and moment release from large and great earthquakes show heterogeneity is the rule. Geodetic mappings of earthquake slip show slip magnitude varies by factors of 4–8 on the coseismic rupture (Fig. 9). Using Eq. (22) and a fault

zone width of 10 km, these observations indicate stress drop varying from 3 to 25 MPa along fault strike. Analysis of broadband seismograms from great earthquakes (e.g., Kikuchi and Fukao, 1987; Beck and Ruff, 1989) shows moment release rate is very nonuniform. Much of the energy radiated to teleseismic distances originates from restricted portions of the earthquake fault (see compilation of 21 such determinations in Thatcher, 1990, Fig. 4), indicating τ_a is also heterogeneous. Similar results have been reported from analysis of regional strong motion recordings of $M\sim7$ earthquakes (Trifunac and Brune, 1971; Cohee and Beroza, 1994; Wald and Heaton, 1994). These observations thus indicate that if the average fault stress is low, the observed heterogeneity in $\Delta\sigma$ and τ_a require that the ambient stress distribution be strongly heterogeneous as well.

9.3 Conceptual Model

These observations suggest a simplistic fault zone strength model shown schematically in Figure 10. Its essential features include patches of strong fault located at the deepest parts of the brittle crust, where large earthquakes nucleate and where Eq. (15) suggests the highest shear stresses occur. These strong
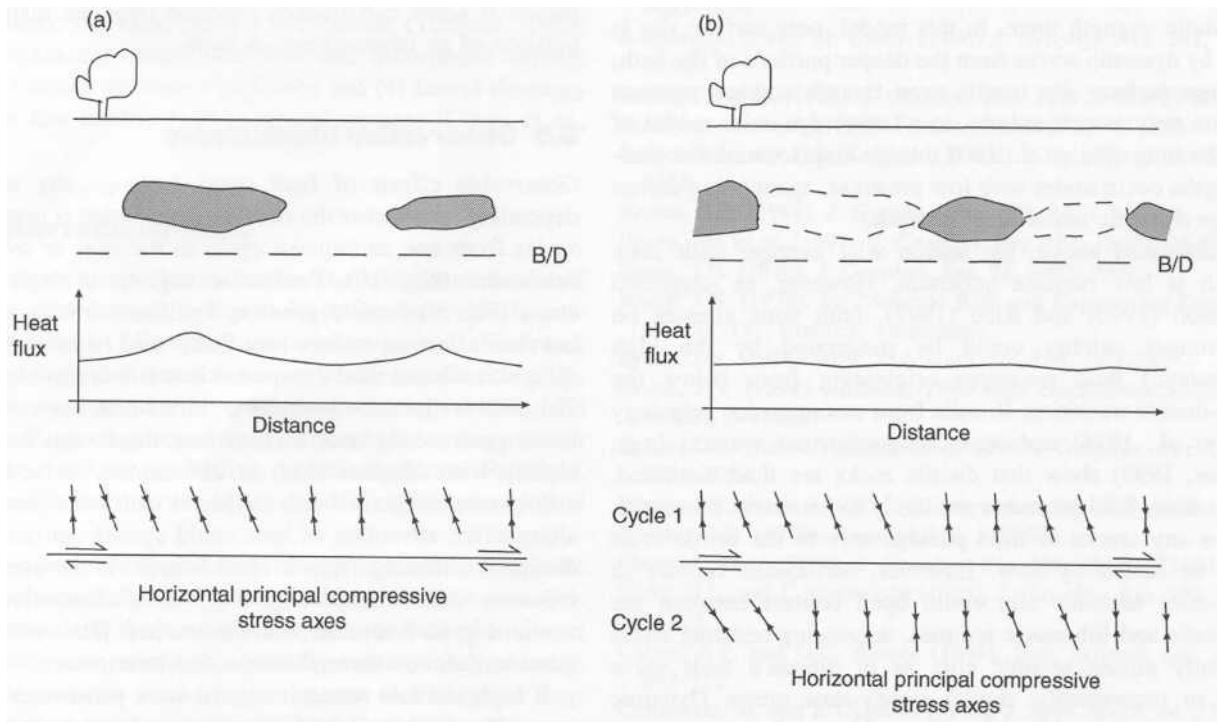
FIGURE 10 Schematic of inhomogeneous strength model: (a) strong and weak regions are permanent; (b) strength varies from cycle to cycle or over long time intervals. Top, strike-parallel longitudinal section view of vertical strike-slip fault, with strong patches shaded. In (b), the strong regions from the previous cycle are shown outlined by dashed lines. Middle, steady-state surface heat flux plotted versus distance along fault strike. Bottom, map view of horizontal principal compressive stress axes along fault strike. In (b), orientations during the preceding cycle (Cycle 1) are shown "north" of fault, and from Cycle 2 to the "south." B/D, brittle–ductile transition.

patches separate much weaker regions at shallower depths and along fault strike at depth.

The model shown in Figure 10 has some similarities with the "asperity" model of faulting inspired by ideas on the distribution of seismic and aseismic slip on plate boundary faults and by seismological mappings of moment release in large earthquakes (e.g., Wesson *et al.*, 1973; Kanamori, 1981; Lay *et al.*, 1982). In this view, the asperities are the locked patches of a fault which slip only during earthquakes, with much of the seismically radiated energy originating from these localized regions.

In the model of Figure 10, the strong patches could have a range of behaviors. They could slip locally with large dynamic stress drops during earthquakes, in which case large amounts of seismic energy will be radiated from these locations. Alternatively, stress drops here may be modest, in which case locally large amounts of energy will be dissipated by frictional heating. The strong patches could be permanent features of the fault zone or they could shift in location from cycle to cycle or over longer time intervals.

Earthquake stress release on high strength patches near the base of the seismogenic zone could be very important in dynamically driving slip at shallower depth. The dynamic model described by Brune (1999), suggested by analogue modeling in foam rubber (Brune, 1996), emphasizes the effect of large driving stresses available at depth due to the higher quasi-static strength there. In this model, near surface slip is driven by dynamic waves from the deeper portions of the fault, and large surface slip results even though ambient stresses there are very low. Similarly, in a lattice dynamics model of thrust faulting (Shi *et al.*, 1998), large displacements at shallow depths occur under very low prestress, again being driven by large dynamic stress drop at depth.

As discussed above, the reason why average fault zone strength is low remains uncertain. However, as suggested by Sibson (1990) and Rice (1992), fault zone stresses on the stronger patches could be moderated by the high (~lithostatic) fluid pressures originating from below the brittle–ductile transition. Results from metamorphic petrology (Fyfe *et al.*, 1978) and study of geothermal systems (e.g., Fournier, 1999) show that ductile rocks are fluid saturated. Furthermore, fluid pressures are likely to be nearly lithostatic, because any cracks or fluid passageways to the brittle crust would be sealed by flow. However, earthquake rupture or steady-state aseismic slip could open contact between the hydrostatic and lithostatic regimes, decreasing resisting stress transiently during seismic slip, as in Sibson's fault valve model or permanently, during steady-state creep. Dynamic weakening during earthquake slip could complement this process or independently supply the fault-averaged low strength required by our inhomogeneous model. This weakening mechanism would not, however, apply for aseismically slipping faults unless the aseismic behavior were a transient feature of the fault zone.

9.4 Role of Average Constraints on τ_r and τ_a

Although successful strength models must simultaneously satisfy the heat flux and radiated energy constraints, these constraints apply only for values of τ_r and τ_a averaged over the fault plane. In the case of heat flux, the constraint is also an average over time and applies over conductive timescales (approximately 100 ka or longer). Thus the requirement that $\tau_r < 10$ MPa on the San Andreas would still permit values consistent with Eq. (15b) over restricted portions of the upper crustal fault surface. If the most dissipative zone were not a permanent feature of the fault, the upper bound value of τ_r might more nearly represent the time-averaged strength of the entire fault zone. The heat flux constraint would appear to be much more stringent on the deeper portions of subduction thrusts, where Eq. (15c) would suggest resisting stresses of ~600 MPa under hydrostatic conditions at 30 km whereas observations permit no more than about 40 MPa at this depth.

The radiated energy constraint applies in the same averaged way as heat flux, and τ_a could locally be much higher than its mean value of 0.5 MPa. Indeed, τ_a values as high as 22 MPa are shown in Color Plate 13. In this regard it may be noteworthy that McGarr (1999) finds the upper bound values of τ_a obtained from seismology are approximately consistent with extrapolations based on laboratory rock mechanics and data from mining-induced seismicity (i.e., 15). This might be expected if some earthquakes ruptured only the high strength portions of an inhomogeneous fault.

9.5 Observation Implications

Observable effects of fault stress heterogeneity will differ depending on whether the strength distribution is permanent or varies from one earthquake cycle to the next or over longer timescales (Fig. 10). Permanent regions of high resisting stress (Fig. 10a) could produce significant frictional heating, but their effect on surface heat flux would be more subtle and difficult to detect than dissipation that is uniform along strike. This follows because laterally confined heat sources generate lower peak steady-state surface heat fluxes that decay more rapidly with distance than do 2D sources. In addition, for a fluid-saturated crust with enhanced fault zone permeability, along-strike advection of heat could spread out the effect of dissipative heating from a local source. If the strength distribution varied temporally (Fig. 10b), dissipative heating would migrate with time, and surface heat flux would become quasi-uniform on thermal conduction timescales.

If high and low strength regions were permanent fault features (Fig. 10a), temporal changes in principal stress orientations would be small or undetectable. If on the other hand, the zones changed from one earthquake cycle to the next (perhaps as slip minima from one earthquake were followed by high slip in a subsequent event), then stress orientations could change from cycle to cycle (Fig. 10b) and also change within

each cycle as the fault strengthened by accumulation of elastic stresses.

Mount and Suppe (1987) and Zoback *et al.* (1987) originally proposed that strength differences exist between the central creeping segment of the San Andreas fault (where compressive stresses are nearly normal to the fault), and the currently locked portion of the fault system further north (where these orientations are more oblique). In principle, determinations of principal stress orientations can also be used to map out stress heterogeneities on a more local scale. Hardebeck and Hauksson (1999) used a large suite of earthquake focal mechanisms to infer local variations in compressive stress orientations both along strike and perpendicular to the San Andreas system in southern California. Although it is tempting to attribute these results as evidence for fault strength heterogeneity, facile interpretations are frustrated by the considerable complexity of the derived orientations, which show rapid spatial changes and seldom decay to stable values far from the San Andreas. While perhaps containing evidence for fault stress heterogeneity, we suspect some of the variability is also due to effects other than simply the ambient stresses on the faults themselves. These include: (1) unavoidable limitations of the inversion method (e.g., inadvertently including more than one stress regime in a single stress field determination); (2) biasing of inversions by using focal mechanisms from faults subparallel to the San Andreas; (3) local stress perturbations (Turcotte, 1982) due to significant crustal thickness and lithospheric density variations across southern California; and (4) lateral rheology variations that modify fault-generated stresses (Chery *et al.*, 2000).

9.6 Observational Tests

We do not see simple tests of either the conceptual model of fault zone stress shown in Figure 10 or those previously proposed by others. Like previous reviewers (e.g., Hickman, 1991), we support direct observation of the constitutive properties and physical state of active faults through drilling and sampling. Plans are proceeding to carry out such an experiment to a depth of ~4 km on the San Andreas fault near Parkfield, California (Hickman *et al.*, 1994) and drilling is now expected to begin during 2003.

Because seismically radiated energy is such a diagnostic measure of stress changes during faulting, more precise estimates, particularly those obtained in the near field of large and great earthquakes, would be especially valuable. In particular, such measurements would be one means of testing our suggestion that apparent stresses obtained from radiated energy estimates are low because τ_a is averaged over a rupture surface where quasi-static or dynamic strength is heterogeneous. More generally, high quality near-field recordings of strong ground motion like those now becoming available for the 1999 $M_w = 7.6$ Chi-Chi (Taiwan) earthquake will constrain the

dynamics of fault slip through mappings of the stress-slip history over the entire rupture plane.

Acknowledgments

References

Adams, G.G. (1999). *ASME J. Tribol.* **121**, 455–461.

Allen, C.R., et al. (1998). *Seismol. Res. Lett.* **69**, 524–531.

Anderson, E.M. (1951). "The Dynamics of Faulting," Oliver and Boyd, Edinburgh.

Andrews, D.J. and Y. Ben-Zion (1997). *J. Geophys. Res.* **102**, 553–571.

Anooshehpoor, A. and J.N. Brune (1994). *Pageoph* **142**, 735–747.

Archuleta, R.J. (1984). *J. Geophys. Res.* **89**, B6, 4559–4585.

Beck, S.E. and L.J. Ruff (1989). *Phys. Earth Planet Inter.*

Ben-Zion, Y. and J.D. Andrews (1998). *Bull. Seismol. Soc. Am.* **88**, 1085–1094.

Boatwright, J. and M. Cocco (1996) *J. Geophys. Res.* **101**, 13 895–13 909.

Bodin, P., et al. (1998). *J. Geophys. Res.* **103**, 29 931–29 944.

Bouissou, S., et al. (1998). *Tectonophysics* **295**, 341–350.

Brace, W.F. and D.L. Kohlstedt (1980) *J. Geophys. Res.* **85**, 6248–6252.

Brown, S.R. (1998). *J. Geophys. Res.* **103**, 7413–7420.

Brudy, M., et al. (1997). *J. Geophys. Res.* **102**, 18 453–18 475.

Brune, J.N. (1970). *J. Geophys. Res.* **75**, 4997–5009.

Brune, J.N. (1976). In: "Seismic Risk and Engineering Decisions," pp. 141–171, Elsevier, Amsterdam.

Brune, J.N. (1996). *Proc. Ind. Acad. Sci.* **105**, 197–206.

Brune, J.N. (1999). *Abstract, 1999 GSA Cordilleran Section Centennial,* 2–4 June.

Brune, J.N. and A. Anooshehpoor (1997). *Geophys. Res. Lett.* **24**, 2071–2074.

Brune, J.N., et al. (1969). *J. Geophys. Res.* **74**, 3821–3827.

Byerlee, J.D. (1978). *Pageoph* **116**, 615–626.

Chery, J., et al. (2000). *J. Geophys. Res.* **106**, 22 051–22 061.

Choy, G.L. and J. Boatwright (1995). *J. Geophys. Res.* **100**, 18 205–18 228.

Cohee, B.P. and G.C. Beroza (1994). *Bull. Seismol. Soc. Am.* **84**, 692–712.

Comninou, M. and J. Dundurs (1977). *J. Appl. Mech.* **44**, 222–226.

Dickinson, W.R. and W.S. Snyder (1979). *J. Geol.* **87**, 609–627.

Dieterich, J.H. (1979). *J. Geophys. Res.* **84**, 2161–2168.

Fleitout, L. and C. Froidevaux (1982). *Tectonics* **1**, 21–56.

Forsyth, D. and S. Uyeda (1975). *Geophys. J. R. Astron. Soc.* **43**, 163–200.

Fournier, R.O. (1999). *Econ. Geol.* submitted.

Freund, L.B. (1978). *J. Appl. Mech.* **45**, 226–227.

Fyfe, W., *et al.* (1978). "Fluids in the Earth's Crust," Elsevier.

Hanks, T.C. (1977). *Pageoph* **115**, 441–458.

Hanks, T.C. (1982). *Bull. Seismol. Soc. Am.* **72**, 1867–1879.

Hanks, T.C. and M. Wyss (1972). *Bull. Seismol. Soc. Am.* **62**, 561–589.

Hardebeck, J.L. and E. Hauksson (1999). *Science* **285**, 233–236.

Hickman, S.H. (1991). *Rev. Geophys. Suppl.* 759–775.

Hickman, S., *et al.* (1994). *EOS, Am. Geophys. Un. Trans.* **75**, 137–142.

Hubbert, M.K. and W.W. Rubey (1959). *Geol. Soc. Am. Bull.* **70**, 115–166.

Hyndman, R.D. and K. Wang (1993). *J. Geophys. Res.* **98** B2, 2039–2060.

Hyndman, R.D., *et al.* (1995). *J. Geophys. Res.* **100**, 15 373–15 392.

Jaeger, J.C. and N.G.W. Cook (1976). "Fundamentals of Rock Mechanics," Wiley.

Kanamori, H. (1981). In: "Earthquake Prediction, An International Review," Maurice Ewing Ser., vol. 4, pp. 1–19, American Geophysical Union.

Kanamori, H. and T.H. Heaton (2000). *AGU Monograph Series* **120**, 147–163.

Kanamori, H., *et al.* (1998). *Science* **279**, 839–842.

Kikuchi, M. and Y. Fukao (1987). *Tectonophysics* **144**, 231–247.

King, N.E. and W. Thatcher (1998). *J. Geophys. Res.* **103**, 18 069–18 078.

Kohlstedt, D.L., *et al.* (1995). *J. Geophys. Res.* **100**, 17 587–17 602.

Lachenbruch, A.H. (1980). *J. Geophys. Res.* **85**, 6097–6112.

Lachenbruch, A.H. and J.H. Sass (1980). *J. Geophys. Res.* **85**, 6185–6223.

Lay, T., *et al.* (1982). *Earthq. Predict. Res.* **1**, 3–71.

McGarr, A. (1999). *J. Geophys. Res.* **104**, 3003–3011.

McGarr, A. and N.C. Gay (1978). *Annu. Rev. Earth Planet,. Sci.* **6**, 405–436.

McKenzie, D.P. (1969). *Geophys. J. R. Astron. Soc.* **18**, 1–32.

McKenzie, D.P. and J.N. Brune (1972). *Geophys. J. R. Astron. Soc.* **29**, 65–78.

Melosh, J. (1996). *Nature* **379**, 601–606.

Molnar, P. and P.C. England (1990). *J. Geophys. Res.* **95**, 4833–4856.

Mora, P. and D. Place (1994). *Pageoph* **143**, 61–87.

Mora, P. and D. Place (1998). *Geophys, Res. Lett.* **103**, 21 067–21 089.

Mora, P. and D. Place (1999). *Geophys. Res. Lett.* **26**, 123–126.

Mori, J. and K.-F. Ma (1999). *EOS, Suppl.* **13**.

Morrow, C., *et al.* (1992). In: "Fault Mechanic sand Transport Properties in Rocks," pp. 69–88, Academic Press.

Mount, V.S. and J. Suppe (1987). *Geology* **15**, 1143–1146.

Ni., S.-D., *et al.* (1999). *EOS, Suppl.* **14**.

Pacheco, J., *et al.* (1993). *J. Geophys. Res.* **98**, 14 133–14 159.

Rice, J.R. (1992). In: "Fault Mechanic sand Transport Properties in Rocks," pp. 475–503, Academic Press.

Richardson, R.M. and L.M. Reding (1991). *J. Geophys. Res.* **96**, 12 201–12 223.

Ruina, A.L. (1983). *J. Geophys. Res.* **88**, 359–370.

Scholz, C.H. (2000). *Geology* **28**, 163–166.

Shi, B., *et al.* (1998). *Bull. Seismol. Soc. Am.* **88**, 1484–1494.

Sibson, R.H. (1973). *Nature* **243**, 66–68.

Sibson, R.H. (1975). *Geophys. J. R. Astron. Soc.* **43**, 775–794.

Sibson, R.H. (1985). *J. Struct. Geol.* **7**, 751–754.

Sibson, R.H. (1990). *Bull. Seismol. Soc. Am.* **80**, 1580–1604.

Singh, S.K. and M. Ordaz (1994). *Bull. Seismol. Soc. Am.* **84**, 1533–1550.

Sleep, N.H. and M. Blanpied (1992). *Nature* **359**, 687–692.

Thatcher, W. (1990). *J. Geophys. Res.* **95**, 2609–2623.

Thatcher, W. and P.C. England (1998). *J. Geophys. Res.* **103**, 891–905.

Thatcher, W. and D.P. Hill (1991). *Geology* **19**, 1116–1120.

Thatcher, W. and D.P. Hill (1995). *J. Geophys. Res.* **100**, 561–570.

Thatcher, W., *et al.* (1997). *J. Geophys. Res.* **102**, 5353–5367.

Tichelaar, B.W. and L.J. Ruff (1993). *J. Geophys. Res.* **98**, 2017–2037.

Townend, J. and M.D. Zoback (2000). *Geology* **28**, 399–402.

Trifunac, M. and J.N. Brune (1970). *Bull. Seismol. Soc. Am.* **60**, 137–160.

Turcotte, D.L. (1982). In: "Mountain Building Processes," pp. 141–146, Academic Press.

Uenishi, K., *et al.* (1999). *Bull. Seismol. Soc. Am.* **89**(5), 1296–1312.

Wald, D.J. and T.H. Heaton (1994). *Bull. Seismol. Soc. Am.* **84**, 668–691.

Wang, C.-Y. and N.-H. Mao (1979). *Geophys. Res. Lett.* **6**, 825–828.

Wesson, R.L., *et al.* (1973). *Stanford Univ. Publ. Geol. Sci.* **13**, 303–321.

Williams, C.F. and T.N. Narasimhan (1989). *Earth Planet. Sci. Lett.* **92**, 131–143.

Wu, F.T. (1978). *Pageoph* **116**, 655–689.

Yuen, D.A., *et al.* (1978). *Geophys. J. R. Astron. Soc.* **54**, 93–119.

Zoback, M.L. (1992). *J. Geophys. Res.* **97**, 11 761–11 781.

Zoback, M.L., *et al.* (1989). *Nature* **341**, 291–298.

Zoback, M.D., *et al.* (1987). *Science* **238**, 1105–1111.

Editor's Note

References with full citation are placed in BruneFullReferences.pdf on the Handbook CD under directory\35Brune. Please see also Chapter 32, Rock failure and earthquakes, by Lockner and Beeler; Chapter 33, State of stress within the Earth, by Ruff; and Chapter 34, State of stress in the Earth's lithosphere, by Zoback and Zoback.