Earthquakes and friction laws

Christopher H. Scholz

Earthquakes have long been recognized as resulting from a stick–slip frictional instability. The development of a full constitutive law for rock friction now shows that the gamut of earthquake phenomena—seismogenesis and seismic coupling, pre- and post-seismic phenomena, and the insensitivity of earthquakes to stress transients—all appear as manifestations of the richness of this friction law.

The traditional view of tectonics is that the lithosphere comprises a strong brittle layer overlying a weak ductile layer, which gives rise to two forms of deformation: brittle fracture, accompanied by earthquakes, in the upper layer, and aseismic ductile flow in the layer beneath. Although this view is not incorrect, it is imprecise, and in ways that can lead to serious misunderstandings. The term ductility, for example, can apply equally to two common rock deformation mechanisms: crystal plasticity, which occurs in rock above a critical temperature, and cataclastic flow, a type of granular deformation which can occur in poorly consolidated sediments. Although both exhibit ductility, these two deformation mechanisms have very different rheologies. Earthquakes, in turn, are associated with strength and brittleness—associations that are likewise sufficiently imprecise that, if taken much beyond the generality implied in the opening sentence, they can lead to serious misinterpretations about earthquake mechanics.

Lately, a newer, much more precise and predictive model for the earthquake mechanism has emerged, which has its roots in the observation that tectonic earthquakes seldom if ever occur by the sudden appearance and propagation of a new shear crack (or 'fault'). Instead, they occur by sudden slippage along a pre-existing fault or plate interface. They are therefore a frictional, rather than fracture, phenomenon, with brittle fracture playing a secondary role in the lengthening of faults[1] and frictional wear[2]. This distinction was noted by several early workers[3], but it was not until 1966 that Brace and Byerlee[4] pointed out that earthquakes must be the result of a stick–slip frictional instability. Thus, the earthquake is the 'slip', and the 'stick' is the interseismic period of elastic strain accumulation. Subsequently, a complete constitutive law for rock friction has been developed based on laboratory studies. A surprising result is that a great many other aspects of earthquake phenomena also now seem to result from the nature of the friction on faults. The properties traditionally thought to control these processes—strength, brittleness and ductility—are subsumed within the overarching concept of frictional stability regimes.

Constitutive law of rock friction

In the standard model of stick–slip friction it is assumed that sliding begins when the ratio of shear to normal stress on the surface reaches a value μ_s, the static friction coefficient. Once sliding initiates, frictional resistance falls to a lower dynamic friction coefficient, μ_d, and this weakening of sliding resistance may, depending on the stiffness of the system, result in a dynamic instability. Following the suggestion of Brace and Byerlee, a great deal of attention was focused on the physics of rock friction, from which it was found that most of the statements in the standard model had to be revised. First, it was found that μ_s depends on the history of the sliding surface. If the surfaces are in static contact under load for time t, then μ_s increases slowly as log t (ref. 5). Second, the dynamic friction, when measured in the steady-state sliding regime, depends[6] on the sliding velocity, V. This dependence, which goes as log V, may be either positive or negative, depending on the rock type and certain other parameters such as temperature[7]. Last, if subjected to a sudden change in sliding velocity, friction is found to evolve to its new steady-state value over a characteristic slip distance $\mathcal{L}$ (refs 8, 9).

The ageing of μ_s and the velocity dependence of μ_d are related behaviours[10] which result from creep of the surface contact and a consequent increase in real contact area with time of contact[11,12]. The critical slip distance $\mathcal{L}$ is interpreted as a memory distance over which the contact population changes[9,13,14]. All of the experimental results are well described by an empirical, heuristic model known as the rate- and state-variable constitutive law, outlined in Box 1. This

Box 1 Rate- and state-variable friction law

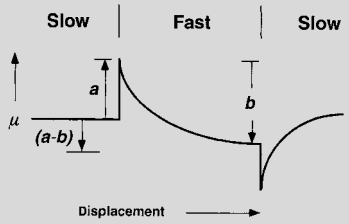
There are several forms of rate/state-variable constitutive law that have been used to model laboratory observations of rock friction. The version currently in best agreement with experimental data[20], known as the Dieterich–Ruina or 'slowness' law, is expressed as

$$\tau = \left[\mu_0 + a \ln\left(\frac{V}{V_0}\right) + b \ln\left(\frac{V_0\theta}{\mathcal{L}}\right)\right]\bar{\sigma} \quad (1)$$

where τ is shear stress and $\bar{\sigma}$ is effective normal stress (applied normal stress minus pore pressure). In the bracketed friction term, V is slip velocity, V_0 a reference velocity, μ_0 the steady-state friction at $V = V_0$, and a and b are material properties. $\mathcal{L}$ is the critical slip distance and the state variable, θ, evolves according to:

$$\frac{d\theta}{dt} = 1 - \frac{\theta V}{\mathcal{L}} \quad (2)$$

The significance of these various terms is illustrated in the diagram below, which shows schematically but faithfully the experimentally observed frictional response to a suddenly imposed e-fold increase and then decrease in sliding velocity.



On initial application of the rate increase there is an increase a in friction, known as the direct velocity effect. This is followed by an evolutionary effect involving a decrease in friction, of magnitude b.

The friction at steady state is:

$$\tau = \left[\mu_0 + (a - b)\ln(V/V_0)\right]\bar{\sigma} \quad (3)$$

from which equation (1) in the main text arises. There is a continuum of changing friction values, but if dynamic friction, μ_d, is defined as steady-state friction at velocity V, then $d\mu_d/d(\ln V) = a - b$. Similarly, if static friction μ_s is defined as the starting friction following a period of time t in stationary contact, then for long t, $d\mu_s/d(\ln t) = b$. The name 'slowness law' arises because at steady state, the state variable is proportional to slowness, $\theta_{ss} = \mathcal{L}/V$. The critical slip distance $\mathcal{L}$ is often interpreted as the sliding distance required to renew the contact population. In this view θ_{ss} represents an average contact lifetime.

form of friction does not seem to be very material dependent: it also applies to some metals[8] and to paper, wood and some plastics[15,16]. The former distinction between μ_s and μ_d disappears in this model. The base friction μ_o has a value nearly independent of rock type and temperature[7,17]. It is modified by second-order effects involving a dependence on sliding velocity and a state variable θ, and it is these second-order effects that result in the interesting modes of behaviour discussed here. The base friction, which determines the frictional strength of the fault, does not concern us in this discussion. The fault strength is not involved in the seismogenic behaviour of the fault, which is solely determined by its frictional stability, not its strength. Fault strength does play a role in frictional heating of faults, which can produce several interesting effects[10,18] that will not be discussed here.

Friction stability regimes and seismogenesis

Frictional stability depends on two friction parameters, $\mathcal{L}$ and the combined parameter $(a-b)$, defined as the velocity dependence of steady-state friction (a and b are defined in Box 1):

$$a-b=\frac{\partial\mu^{ss}}{\partial[\ln(V)]} \quad (1)$$

The frictional stability regimes are described in Box 2. If $(a-b)\geqslant 0$, the material is said to be velocity strengthening, and will always be stable. In the velocity-weakening field, $(a-b)<0$, there is a Hopf bifurcation between an unstable regime and a conditionally stable one. Considering a simple spring–slider model with fixed stiffness k, the bifurcation occurs at a critical value of effective normal stress, $\bar{\sigma}_c$, given by:

$$\bar{\sigma}_c=\frac{k\mathcal{L}}{-(a-b)} \quad (2)$$

If $\bar{\sigma}>\bar{\sigma}_c$, sliding is unstable under quasistatic loading. In the conditionally stable regime, $\bar{\sigma}<\bar{\sigma}_c$, sliding is stable under quasistatic loading but can become unstable under dynamic loading if subjected to a velocity jump exceeding ΔV, as shown in Box 2. In a narrow region at the bifurcation, sliding occurs by a self-sustaining oscillatory motion[6,15,19] (shaded region, in the second diagram in Box 2). Although the friction law described in Box 1 can be written in several ways which differ in detail[20], those details do not influence the above definitions of the stability states, which control the seismic behaviour of faults discussed here.

The three stability regimes have the following consequences for earthquakes. Earthquakes can nucleate only in those regions of a fault that lie within the unstable regime. They may propagate indefinitely into conditionally stable regions, provided that their dynamic stresses continue to produce a large enough velocity jump. If earthquakes propagate into a stable region, on the other hand, a negative stress drop will occur, resulting in a large energy sink that will rapidly stop the propagation of the earthquake.

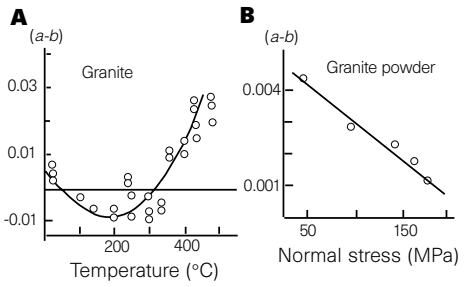


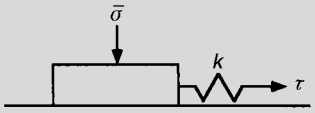
Figure 1 Systematics of the friction parameter $(a-b)$. **A**, Dependence of $(a-b)$ on temperature for granite (from refs 7, 21). **B**, Dependence of $(a-b)$ on pressure for granulated granite (ref. 24). This effect, due to lithification, should be augmented with temperature.

The primary parameter that determines stability, $(a-b)$, is a material property (see Box 1). The main systematics of this parameter that concern us are summarized in Fig. 1. Figure 1A shows the dependence of $(a-b)$ on temperature for granite[7,21]. It is negative at low temperatures and becomes positive for temperatures above about 300 °C. This transition temperature corresponds to the onset of crystal plasticity of quartz, the most ductile of the major minerals in granite[22]. It may be a general statement that for low-porosity crystalline rocks a transition from negative to positive $(a-b)$ corresponds to a change from elastic-brittle deformation to crystal plasticity in the micro-mechanics of friction. As another example, halite (rock salt), a much more ductile mineral, undergoes the same two corresponding transitions at 25 °C and a pressure of about 70 MPa (ref. 23). These observations indicate that for faults in granite, the representative rock of the continental crust, we should not expect earthquakes to occur below a depth at which the temperature is 300 °C, and faults in salt should be aseismic under almost all conditions.

Faults are not simply frictional contacts of bare rock surfaces: they are usually lined with wear detritus, called cataclasite or fault gouge. The shearing of such granular material involves an additional hardening mechanism (involving dilatancy), which tends to make $(a-b)$ more positive[24]. For such materials, $(a-b)$ is positive when

Box 2 Stability regimes

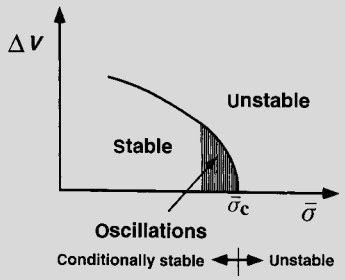
Consider a simple spring-slider model with spring stiffness k, as shown in the diagram below, in which the slider obeys the rate/state-variable friction law.



The stability of this system depends entirely on $\bar{\sigma}$, τ, k, the friction parameters $(a-b)$ and $\mathcal{L}$, and is independent of base friction μ_0. The following are the conditions defining the stability regimes.

$(a-b)>0$. This is velocity-strengthening behaviour, which is intrinsically **stable**. No earthquake can nucleate in this field, and any earthquake propagating into this field will produce there a negative stress drop, which will rapidly terminate propagation.

$(a-b)<0$. The stability diagram for the system exhibiting velocity weakening is shown in the diagram below.



This diagram shows the velocity jump, ΔV, necessary to destabilize the system as a function of the applied normal stress $\bar{\sigma}$. If $\bar{\sigma}\geqslant\bar{\sigma}_c$, as defined by equation (2) in the main text, the system is unstable with respect to a vanishing velocity perturbation ΔV–that is, to quasistatic loading. This is the **unstable** field. If the effective normal stress is less than the critical value, it requires a finite velocity 'kick' to become unstable. Thus, in this **conditionally stable** field, the system is stable under quasistatic loading but may become unstable under sufficiently strong dynamic loading. Earthquakes may nucleate only in the unstable field, but may propagate into the conditionally stable field. At the border of the stability transition there is a narrow region in which self-sustaining oscillatory motion occurs, as indicated by the shaded region.

the material is poorly consolidated, but decreases at elevated pressure and temperature as the material becomes lithified (Fig. 1B). Therefore faults may also have a stable region near the surface, owing to the presence of such loosely consolidated material[25].

These considerations allow the construction of synoptic models for the two primary sites of tectonic earthquakes, crustal faults and subduction zone interfaces, as shown in Fig. 2. In the centre of this figure is drawn the expected variation of the friction stability parameter $\zeta = (a - b)\bar{\sigma}$. This is positive at shallow depths because of the presence of unconsolidated granular material and at large depths because of the onset of plasticity at a critical temperature—hence, regions above and below these stability transitions are stable (shown blue). The regions in which the stability parameter exceeds the threshold defined in equation (2) are unstable, indicated by red, and yellow indicates the regions of conditional stability. Thus, the red regions define the seismogenic zone, the depth range over which earthquakes may nucleate, as indicated by their hypocentral depths, an example of which is given on the right of Fig. 2.

For crustal faults, the upper transition depth is typically observed to be at 3–4 km, but may be absent at faults on which there has been little slip and hence little or no gouge developed[25]. The lower transition occurs at 15–20 km, corresponding to the onset of plasticity of quartz at about 300 °C. The depth at which this occurs depends on the local thermal gradient[26]. For subduction zones the upper transition occurs at the base of the accretionary prism of scraped-off oceanic sediments, where it encounters a 'backstop' of competent rock[27]. Because the thickness of the sedimentary wedge is quite variable, so is the depth of this transition—it may be as deep as 10 km. The lower transition occurs at depths as great as 45 km at subduction zones. This greater depth is a result of lower thermal gradients, owing to the subduction of the cold oceanic plate, although this may vary widely because of variations in the age of the subducting plate, which strongly affects the thermal regime[28]. The transition is also deeper because the basalt of the oceanic plate contains no quartz—the most ductile mineral in basalt is feldspar, which becomes plastic at about 450 °C (ref. 22). Because the seismogenic zone is much wider than for crustal faults—up to 150 km—and because they tend to be more continuous along strike, subduction zones produce by far the largest earthquakes in the world.

If a large earthquake occurs on a crustal fault, it will often have enough energy to propagate through the narrow shallow stable region and breach the surface. It may also often propagate a short distance into the ductile stable region at depth, for which there is geological evidence[29,30]. However, for subduction zones with a wide accretionary prism, large earthquakes will often not breach the surface. Whether they do or not is thought to be important in determining how efficient they are in generating tsunamis[31].

Seismic coupling and seismic styles

The linear measure of earthquake size is seismic moment, $M_0 = GuA$, where u is the mean slip in the earthquake, A the rupture area and G the shear modulus. The moment release rate of a fault or plate boundary is thus $\dot{M}_0 = G\nu A$ where ν is the long-term slip velocity and A is now the total fault area. We define the seismic coupling coefficient χ as the ratio of the moment release rate determined from summing earthquakes to the total rate obtained by determining ν from a plate-tectonic model or geological data. The parameter χ is a good measure of the overall stability state of a fault. If the fault is entirely in the unstable field, $\chi = 1$, and if entirely in the stable field $\chi = 0$; otherwise, χ will be somewhere in between.

For most crustal faults, χ is indistinguishable from 1; that is, all of the fault slip occurs during earthquakes and these faults are said to be fully seismically coupled. An important exception is the so-called 'creeping' section of the San Andreas fault, a 170-km-long stretch in central California where the fault slips aseismically. Much of this aseismic slip occurs as 'creep episodes' (Fig. 3), which appear to be the same as the oscillatory behaviour observed at the stability boundary—*prima facie* evidence that this part of the fault is in the conditionally stable regime close to the bifurcation of equation (2). It is sufficiently far from the stability boundary, though, to prevent earthquakes on neighbouring sections of the fault from propagating very far into this region. The most likely mechanism for the anomalous behaviour of this section of the fault is the presence there of unusually high pore pressures in the fault zone[32]. I note that the effective normal stress is $\bar{\sigma} = (\sigma - p)$, where σ is the applied normal stress and p the pore pressure. If p approaches σ, the stability parameter ζ may be reduced so that the entire depth range of the fault normally in the unstable (red) field in Fig. 2 is shifted to the yellow field.

Although such seismic decoupling seems to be rare for crustal faults, it is not rare for subduction zones, which vary from being fully coupled to almost entirely decoupled[33]. The difference seems to be due to the stress state. Stress measurements in deep boreholes[34] in continents typically show that deviatoric stresses increase with depth so that at all depths the stresses are just below that necessary to cause sliding on a favourably oriented fault with a friction coefficient consistent with laboratory values (about 0.6). The pore pressures observed in deep boreholes usually increase with the hydrostatic gradient, and the vertical stress usually increases with the weight of the overlying rock. The most-studied fault, the San Andreas of California, seems to be exceptional, in that it seems to be

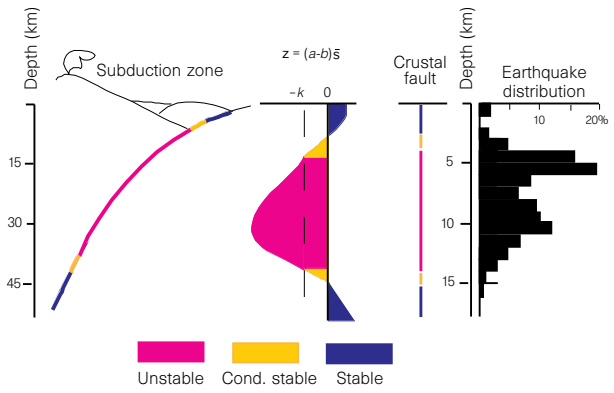


Figure 2 A synoptic model for stability as a function of depth for crustal faults and subduction zones. The central panel and the crustal fault model are taken from ref. 22; the subduction zone model (left) is from refs 27, 28; the histogram of the depth distribution of earthquakes (right) is for a section of the San Andreas fault near Parkfield, California (data from ref. 25).

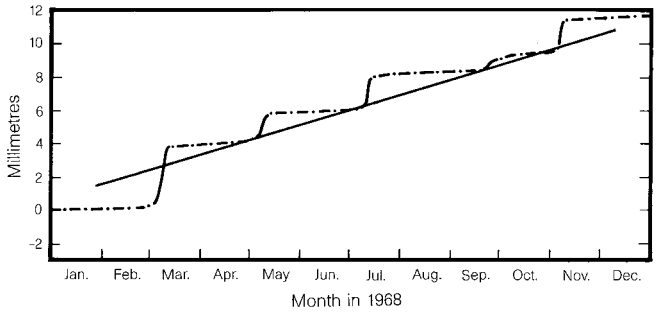


Figure 3 Oscillatory motion (creep episodes) of the creeping section of the San Andreas fault in central California (from ref. 10). The straight line is for reference.

sliding under very low shear stresses[35]. This can be consistent with the friction law only if the pore pressure within the fault is near the lithostatic load (weight of overburden), which is very problematical[36]. The important point is that crustal faults are subjected to remotely applied loads, and the effective normal stress on them is mainly determined by the lithostatic load minus the pore pressure, which is in the more usual case the hydrostatic head. For subduction zones, on the other hand, the forces that drive the plates are local to the subduction zone and may vary widely, which results in great variation of the effective normal stress supported by the plate interface. An analysis of the reduction of normal force (relative to a standard state) applied across subduction interfaces[37], calculated from the plate-tectonic driving forces, is shown in Fig. 4 for most of the world's subduction zones. The seismic coupling coefficient χ, determined from seismicity data, decreases from high to low values at a critical value corresponding to $\bar{\sigma}_c$, which was determined independently of the data shown in the figure. Owing to the shortness of the seismic record, these values of χ are not very well determined[38], but they are good enough to allow one to distinguish the coupled from the decoupled zones. Thus coupled and decoupled subduction zones are on either side of the stability transition boundary. On a local scale, irregularities caused, for example, by the subduction of seamounts can produce local increases in normal stress with the result that otherwise decoupled subduction zones may become locally coupled[39].

The three stability states result in three distinctive seismic styles. Regions with the stable field, such as the outer parts of accretionary prisms and faults in salt, are totally aseismic[27,40]. Faults in the unstable field are characterized by infrequent large earthquakes separated by long interseismic periods of quiescence. On the other hand, faults in the conditionally stable regime, such as the creeping section of the San Andreas and the decoupled subduction zones, are characterized by high steady rates of small-event activity and no large events ('large' events are earthquakes that rupture the entire seismogenic thickness). These small events together contribute very little to the total moment release, which is primarily aseismic[39,41]. Small events are found to occur repeatedly with a high repetition rate at the same spots[42]. These spots may mark small geometric irregularities[43] where the normal stress is higher, causing a transition to the unstable field.

Stages in the seismic cycle

As noted above, seismically coupled faults are typified by infrequent large events, separated by long quiescent interseismic periods in which the stresses relaxed by the preceding earthquake are restored. A frictional model of the seismic cycle of a strike–slip fault[44,45], purposely reminiscent of the San Andreas fault, is shown in Fig. 5. In this two-dimensional model, the fault is driven remotely at a constant velocity; the figure shows the slip on the fault as a function of depth at different times during the seismic cycle. The only assumption in the model is that it obeys the friction law of Box 1 with the $(a - b)$ parameter varying as in Fig. 1A and $\bar{\sigma}$ increasing with depth in a way consistent with the borehole data summarized above. The depth of the transition from unstable to stable regimes is at 11 km, in accordance with the geothermal gradient typical of the San Andreas fault.

During the interseismic period (shown blue in Fig. 5), the fault is loaded by steady slip on the deep, stable portion of the fault. Just before the earthquake a pre-seismic phase, known as nucleation, occurs (orange); in this phase, slip accelerates until the instability results in the coseismic motions (red). These penetrate below the stability boundary, reloading that region, which relaxes in a post-seismic phase of accelerated deep slip (green), at a rate that decays exponentially with time within a few years to a decade following the mainshock. Geodetic data strongly support the main features of this model[46–48]: the interseismic strain accumulation resulting from deep slip below a locking depth, above which the coseismic slip occurs, and a postseismic relaxation phase with decelerating deep slip[49]. The pre-seismic nucleation phase is sometimes associated with the occurrence of foreshocks[50], and may also be responsible for certain other precursory phenomena that have been occasionally observed[10].

A shallow relaxation phenomenon called afterslip is often observed, in which a fault slips aseismically at the surface in proportion to the logarithm of the time elapsed since an earthquake just below. Afterslip is usually observed where a thick layer of poorly consolidated sediments overlies the fault, which partially or totally impedes the earthquake from breaching the surface. Afterslip can be described by a model similar to that illustrated in Fig. 5 but with a thick stable layer at the top[51]; an example is shown in Fig. 6. This phenomenon has also been observed in partially coupled subduction zones, where an earthquake in an unstable patch drives afterslip in adjacent conditionally stable or stable regions[52]. The other typical postseismic phenomena, aftershock sequences, which obey a well defined hyperbolic decay law known as the Omori law, have also been shown to be a prediction of the rate- and state-variable friction

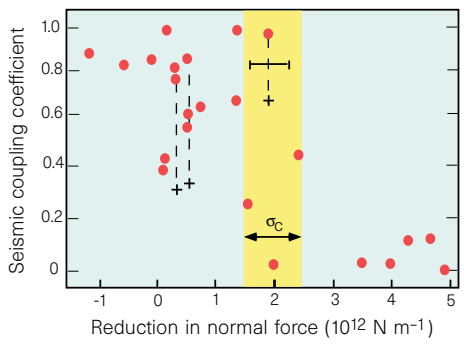


Figure 4 The observed seismic coupling coefficient χ versus the calculated reduction in normal force from a standard state for most of the Earth's subduction zones (after ref. 37). The transition point, $\bar{\sigma}_c$, was independently estimated from the Izu-Bonin/Mariana arc.

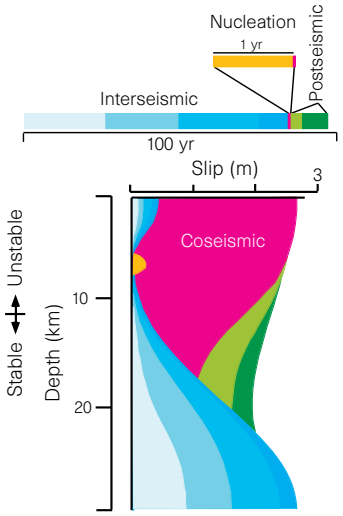


Figure 5 Slip as a function of depth over the seismic cycle of a strike–slip fault, using a frictional model containing a transition from unstable to stable friction at 11 km depth (after ref. 44).

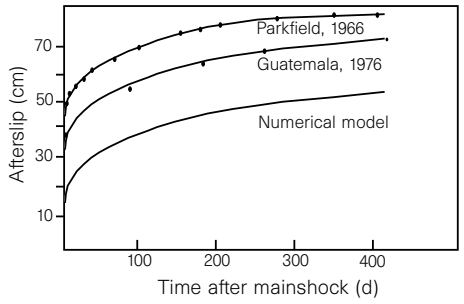


Figure 6 Afterslip observed at the surface following two large strike-slip earthquakes compared with the results of a numerical model of slip in a stable layer overlying an unstable region which slipped coseismically (from ref. 51).

law[53]. This part of the theory has also been successfully tested with observations[54].

When the instability condition for a one-dimensional spring slider, equation (2), is generalized for the two- or three-dimensional case of a slipping patch of size L, the stiffness, k, scales inversely with L according to $k = \eta G/L$, where G is the shear modulus and η is a geometrical constant of the order of unity. This implies that the instability occurs when the slipping patch reaches a critical size L_c, the nucleation length, given by:

$$L_c = \frac{G\eta\mathcal{L}}{(b-a)\bar{\sigma}} \quad (3)$$

Modelling[55] and laboratory observations[6,56] of this nucleation process indicate that stable sliding initiates at a point and then spreads out with an accelerating sliding velocity until the instability arises at L_c. Whether or not such nucleation occurs on natural faults, and whether it is large enough to be detected, is central to the problem of short-term earthquake prediction. However, the physical significance and scaling of the key parameter $\mathcal{L}$ is not known. It is very small ($\sim$10 μm) in the laboratory. Various attempts to model $\mathcal{L}$, assuming that it is a property of the surface contact topography[57] or gouge zone thickness[58], suggest that it may be much larger at the fault scale. If L_c is a constant for natural faults, it also represents a minimum earthquake dimension. In that case $L_c < 10$ m, the rupture size of the smallest observed earthquakes[59]. On the other hand, observations of the dimensions of foreshock zones and of a precursory seismic phase both indicate that nucleation length may be of the order of kilometres and scale with the size of the ensuing earthquake[60,61].

Earthquake insensitivity to transients

There is one last property of this friction law which clears up some long-known mysteries about earthquakes. The direct friction effect (Box 1) and the finite size and duration of nucleation prohibit earthquakes from being triggered by high-frequency stress oscillations[62,75]. Hence, as has been repeatedly verified over the past 75 years, earthquakes are not triggered by Earth tides[63]. They are also not triggered (except in magmatic systems[64]) by the seismic waves from other earthquakes[65], even though the smaller residual static stresses from earthquakes may trigger earthquakes following some time delay[66].

Outstanding problems in earthquake mechanics

There are, of course, many details yet to be worked out about the friction law, its parameters, and their scaling properties and applications to natural seismic phenomena. The scaling of $\mathcal{L}$ and its effect on nucleation is but one of these. But, lest the above discussion give the impression that most of what is worth knowing about earthquake mechanics is now well understood, I will describe several important problems that are currently at issue.

One important question is: what gives rise to the complexity of earthquakes? Earthquake populations obey a very well defined power-law size distribution, known as the Gutenberg–Richter law—such power laws being the hallmarks of systems exhibiting 'self-organized criticality' (ref. 76). The internal dynamics of earthquakes also exhibit complexity, with very broad-band velocity and acceleration spectra, while at the same time obeying well defined self-similar scaling laws in terms of the static parameters[10]. Is this complexity a result of the heterogeneity of faults, which have quasi-fractal scaling of surface topography[67], or of the nonlinearity of the friction laws, or both?

Dynamic models of arrays of spring-coupled block sliders[68–70] obeying simple versions of these friction laws have been successful in reproducing Gutenberg–Richter statistics and many other aspects of observed earthquake complexity. These models assume no built-in heterogeneity, hence these results imply that this complexity may arise solely from the nonlinearity of the friction. On the other hand, continuum models show that complexity does not arise easily from the friction alone[45,71], unless extreme values for the friction parameters are assumed[72]. But these studies are far from exploring all regimes of friction parameter space. Furthermore, as remarked on below, the friction laws are still quite simple: there may be other aspects not yet uncovered in laboratory experiments.

There is a newly discovered class of earthquakes that are not expected from the friction laws as currently formulated: these are the 'slow' earthquakes, which are characterized by moment release rates much lower than those seen in other earthquakes[73]. A slow earthquake occurring at a subduction zone may generate a much larger tsunami than expected from its moment as measured in the usual frequency band[74]. One might suppose that some additional dissipative term, not yet recognized in the laboratory, may be present in order to explain this kind of earthquake.

Although there are many interesting phenomena yet to be explored, the success of the rate- and state-variable friction laws, simple as they are, in explaining a wide range of earthquake phenomena gives confidence that they will provide the basis for many exciting future discoveries as well. □

Christopher H. Scholz is at the Lamont-Doherty Earth Observatory and Department of Earth and Environmental Sciences, Columbia University, Palisades, New York 10964, USA

1. Cowie, P. A. & Scholz, C. H. Growth of faults by the accumulation of seismic slip. *J. Geophys. Res.* **97,** 11085–11095 (1992).
2. Scholz, C. H. Wear and gouge formation in brittle faulting. *Geology* **15,** 493–495 (1987).
3. Gilbert, G. K. A theory of the earthquakes of the Great Basin, with a practical application. *Am. J. Sci.* **XXVII,** 49–54 (1884).
4. Brace, W. F. & Byerlee, J. D. Stick slip as a mechanism for earthquakes. *Science* **153,** 990–992 (1966).
5. Dieterich, J. Time-dependence of rock friction. *J. Geophys. Res.* **77,** 3690–3697 (1972).
6. Scholz, C., Molnar, P. & Johnson, T. Detailed studies of frictional sliding of granite and implications for the earthquake mechanism. *J. Geophys. Res.* **77,** 6392–6406 (1972).
7. Stesky, R. *et al.* Friction in faulted rock at high temperature and pressure. *Tectonophysics* **23,** 177–203 (1974).
8. Rabinowicz, E. The intrinsic variables affecting the stick-slip process. *Proc. Phys. Soc. (London)* **71,** 668–675 (1958).
9. Dieterich, J. Time dependent friction and the mechanics of stick slip. *Pure Appl. Geophys.* **116,** 790–806 (1978).
10. Scholz, C. H. *The Mechanics of Earthquakes and Faulting* (Cambridge Univ. Press, 1990).
11. Scholz, C. H. & Engelder, T. Role of asperity indentation and ploughing in rock friction. *Int. J. Rock Mech. Min. Sci.* **13,** 149–154 (1976).
12. Wang, W. & Scholz, C. H. Micromechanics of the velocity and normal stress dependence of rock friction. *Pure Appl. Geophys.* **143,** 303–316 (1994).
13. Dieterich, J. Modelling of rock friction: 1. Experimental results and constitutive equations. *J. Geophys. Res.* **84,** 2161–2168 (1979).
14. Ruina, A. L. Slip instability and state variable friction laws. *J. Geophys. Res.* **88,** 10359–10370 (1983).
15. Heslot, F., Baumberger, T., Perrin, B., Caroli, B. & Caroli, C. Creep, stick-slip, and dry friction dynamics: experiments and heuristic model. *Phys. Rev. E* **49,** 4973–4988 (1994).
16. Dieterich, J. H. & Kilgore, B. Direct observation of frictional contacts: new insights for state-dependent properties. *Pure Appl. Geophys.* **143,** 283–302 (1994).
17. Byerlee, J. D. Friction of rock. *Pure Appl. Geophys.* **116,** 615–626 (1978).
18. Segall, P. & Rice, J. R. Dilatancy, compaction, and slip instability of a fluid infiltrated fault. *J. Geophys. Res.* **100,** 22155–22171 (1995).
19. Gu, J. C., Rice, J. R., Ruina, A. L. & Tse, S. T. Slip motion and stability of a single degree of freedom elastic system with rate and state dependent friction. *J. Mech. Phys. Solids* **32,** 167–196 (1984).
20. Beeler, N. M., Tullis, T. E. & Weeks, J. D. The roles of time and displacement in the evolution effect in rock friction. *Geophys. Res. Lett.* **21,** 1987–1990 (1994).
21. Blanpied, M. L., Lockner, D. A. & Byerlee, J. D. Fault stability inferred from granite sliding experiments at hydrothermal conditions. *Geophys. Res. Lett.* **18,** 609–612 (1991).

22. Scholz, C. H. The brittle-plastic transition and the depth of seismic faulting. *Geol. Rundsch.* **77,** 319–328 (1988).
23. Shimamoto, T. A transition between frictional slip and ductile flow undergoing large shearing deformation at room temperature. *Science* **231,** 711–714 (1986).
24. Marone, C., Raleigh, C. B. & Scholz, C. Frictional behavior and constitutive modelling of simulated fault gouge. *J. Geophys. Res.* **95,** 7007–7025 (1990).
25. Marone, C. & Scholz, C. The depth of seismic faulting and the upper transition from stable to unstable slip regimes. *Geophys. Res. Lett.* **15,** 621–624 (1988).
26. Sibson, R. H. Fault zone models, heat flow, and the depth distribution of earthquakes in the continental crust of the United States. *Bull. Seismol. Soc. Am.* **72,** 151–163 (1982).
27. Byrne, D. E., Davis, D. M. & Sykes, L. R. Loci and maximum size of thrust earthquakes and the mechanics of the shallow region of subduction zones. *Tectonics* **7,** 833–857 (1988).
28. Hyndman, R. D. & Wang, K. Thermal constraints on the zone of major thrust earthquake failure: the Cascadia subduction zone. *J. Geophys. Res.* **98,** 2039–2060 (1993).
29. Sibson, R. H. Transient discontinuities in ductile shear zones. *J. Struct. Geol.* **2,** 165–171 (1980).
30. Stel, H. The effect of cyclic operation of brittle and ductile deformation on the metamorphic assemblage in cataclasites and mylonites. *Pure Appl. Geophys.* **124,** 289–307 (1986).
31. Kanamori, H. & Kikuchi, M. The 1992 Nicaragua earthquake: a slow tsunami earthquake associated with subducted sediments. *Nature* **361,** 714–716 (1993).
32. Irwin, W. P. & Barnes, I. Effects of geological structure and metamorphic fluids on seismic behavior of the San Andreas fault system in central and northern California. *Geology* **3,** 713–716 (1975).
33. Ruff, L. & Kanamori, H. Seismicity and the subduction process. *Phys. Earth Planet. Inter.* **23,** 240–252 (1980).
34. Zoback, M. L. & Zoback, M. D. Crustal stress and intraplate deformation. *Geowissenshaften* **15,** 116–112 (1997).
35. Zoback, M. D. *et al.* New evidence on the state of stress of the San Andreas fault system. *Science* **238,** 1105–1111 (1987).
36. Scholz, C. H. Faults without friction? *Nature* **381,** 556–557 (1996).
37. Scholz, C. H. & Campos, J. On the mechanism of seismic decoupling and back-arc spreading in subduciton zones. *J. Geophys. Res.* **100,** 22103–2212 (1995).
38. McCaffrey, R. Statistical significance of the seismic coupling coefficient. *Bull. Seismol. Soc. Am.* **87,** 1069–1073 (1997).
39. Scholz, C. H. & Small, C. The effect of seamount subduction on seismic coupling. *Geology* **25,** 487–490 (1997).
40. Seeber, L., Armbruster, J. G. & Quittmeyer, R. C. *Seismicity and continental subduction in the Himalayan arc* in *Geodynamics Series, V* (eds Gupta & Delany) 215–242 (Am. Geophys. Un., Washington DC, 1981).
41. Amelung, F. & King, G. C. P. Earthquake scaling laws for creeping and non-creeping faults. *Geophys. Res. Lett.* **24,** 507–510 (1997).
42. Nadeau, R., Foxall, W. & McEvilly, T. Clustering and periodic recurrence of microearthquakes on the San Andreas fault at Parkfield, California. *Science* **267,** 503–507 (1995).
43. Bakun, W. H., Stewart, R. M., Bufe, C. G. & Marks, S. M. Implication of seismicity for failure of a section of the San Andreas fault. *Bull. Seismol. Soc. Am.* **70,** 185–201 (1980).
44. Tse, S. & Rice, J. Crustal earthquake instability in relation to the depth variation of friction slip properties. *J. Geophys. Res.* **91,** 9452–9472 (1986).
45. Rice, J. R. Spacio-temporal complexity of slip on a fault. *J. Geophys. Res.* **98,** 9885–9907 (1993).
46. Thatcher, W. Nonlinear strain buildup and the earthquake cycle on the San Andreas fault. *J. Geophys. Res.* **88,** 5893–58902 (1978).
47. Savage, J. C. & Prescott, W. H. Asthenospheric readjustment and the earthquake cycle. *J. Geophys. Res.* **83,** 3369–3376 (1978).
48. Gilbert, L. E., Scholz, C. H. & Beavan, J. Strain localization along the San Andreas fault: consequences for loading mechanisms. *J. Geophys. Res.* **99,** 23975–23984 (1994).
49. Savage, J. C. & Svarc, J. L. Postseismic deformation associated with the 1992 $M_W = 7.3$ Landers earquake, southern California. *J. Geophys. Res.* **102,** 7565–7577 (1997).
50. Das, S. & Scholz, C. Theory of time-dependent rupture in the earth. *J. Geophys. Res.* **86,** 6039–6051 (1981).
51. Marone, C. J., Scholz, C. H. & Bilham, R. On the mechanics of earthquake afterslip. *J. Geophys. Res.* **96,** 8441–8452 (1991).
52. Heki, K., Miyazaki, S. & Tsuji, H. Silent fault slip following an interplate thrust earthquake at the Japan trench. *Nature* **386,** 595–598 (1997).
53. Dieterich, J. H. A constitutive law for rate of earthquake production and its application ot earthquake clustering. *J. Geophys. Res.* **99,** 2601–2618 (1994).
54. Gross, S. & Kisslinger, C. Estimating tectonic stress rate and state with Landers aftershocks. *J. Geophys. Res.* **102,** 7603–7612 (1997).
55. Dieterich, J. H. Nucleation on faults with rate and state-dependent strength. *Tectonophysics* **211,** 115–134 (1992).
56. Ohnaka, M., Kuwahana, Y., Yamamoto, K. & Hirasawa, T. in *Earthquakes Source Mechanics* (eds Das, S., Boatwright, J. & Scholz, C.) 13–24 (AGU Monogr. 37, Am. Geophys. Un., Washington DC, 1986).
57. Scholz, C. H. The critical slip distance for seismic faulting. *Nature* **336,** 761–763 (1988).
58. Marone, C. & Kilgore, B. Scaling of the critical slip distance for seismic faulting with shear strain in fault zones. *Nature* **362,** 618–622 (1993).
59. Abercrombie, R. & Leary, P. Source parameters of small earthquakes recorded ant 2.5 km depth, Cajon Pass, southern California: Implications for earthquake scaling. *Geophys. Res. Lett.* **20,** 1511–1514 (1993).
60. Dodge, D. A., Beroza, G. C. & Ellsworth, W. L. Detailed observations of California foreshock sequences: Implications for the earthquake initiation process. *J. Geophys. Res.* **101,** 22371–22392 (1996).
61. Ellsworth, W. L. & Beroza, G. C. Seismic evidence for a seismic nucleation phase. *Science* **268,** 851–855 (1995).
62. Dieterich, J. H. Nucleation and triggering of earthquake slip: effect of periodic stresses. *Tectonophysics* **144,** 127–139 (1987).
63. Heaton, T. H. Tidal triggering of earthquakes. *Bull. Seismol. Soc. Am.* **72,** 2181–2200 (1982).
64. Hill, D. P. *et al.* Seismicity in the western United States triggered by the M 7.4 Landers, California, earthquake of June 28, 1992. *Science* **260,** 1617–1623 (1993).
65. Cotton, F. & Coutant, O. Dynamic stress variations due to shear faults in a plane layered medium. *Geophys. J. Int.* **128,** 676–688 (1997).
66. King, G. C. P., Stein, R. S. & Lin, J. Static stress changes and the triggering of earthquakes. *Bull. Seismol. Soc. Am.* **84,** 935–953 (1994).
67. Power, W. L., Tullis, T. E., Brown, S., Boitnott, G. & Scholz, C. H. Roughness of natural fault surfaces. *Geophys. Res. Lett.* **14,** 29–32 (1987).
68. Carlson, J. M. & Langer, J. S. Mechanical model of an earthquake fault. *Phys. Rev. A* **40,** 6470–6484 (1989).
69. Myers, C., Shaw, B. & Langer, J. Slip complexity in a crustal plane model of an earthquake fault. *Phys. Rev. Lett.* **77,** 972–975 (1996).
70. Shaw, B. E. Frictional weakening and slip complexity on earthquake faults. *J. Geophys. Res.* **100,** 18239–18248 (1995).
71. Ben-Zion, Y. & Rice, J. R. Slip patterns and earthquake populations along different classes of faults in elastic solids. *J. Geophys. Res.* **100,** 12959–12983 (1995).
72. Cochard, A. & Madariaga, R. Complexity of seismicity due to highly rate-dependent friction. *J. Geophys. Res.* **101,** 25321–25336 (1996).
73. Ihmlé, P. F. & Jordan, T. H. Teleseismic search for slow precursors to large earthquakes. *Science* **266,** 1547–1550 (1994).
74. Ihmlé, P. F. Monte Carlo slip inversion in the frequency domain: application to the 1992 Nicaragua slow earthquake. *Geophys. Res. Lett.* **23,** 913–916 (1996).
75. Johnson, T. Time-dependent friction of granite: implications for precursory slip on faults. *J. Geophys. Res.* **86,** 6017–6028 (1981).
76. Bak, P., Tang, C. & Wiesenfeld, K. Self-organized criticality. *Phys. Rev. A* **38,** 364–374 (1988).

Acknowledgements. I thank L. Sykes and B. Shaw for comments. This work was partially supported by the US Geological Survey.

Correspondence should be addressed to the author.