Heating and weakening of faults during earthquake slip

James R. Rice

Received 21 August 2005; accepted 23 January 2006; published 24 May 2006.

[1] Field observations of mature crustal faults suggest that slip in individual events occurs primarily within a thin shear zone, <1–5 mm, within a finely granulated, ultracataclastic fault core. Relevant weakening processes in large crustal events are therefore suggested to be thermal, and to involve the following: (1) thermal pressurization of pore fluid within and adjacent to the deforming fault core, which reduces the effective normal stress and hence also the shear strength for a given friction coefficient and (2) flash heating at highly stressed frictional microcontacts during rapid slip, which reduces the friction coefficient. (Macroscopic melting, or possibly gel formation in silica-rich lithologies, may become important too at large enough slip.) Theoretical modeling of mechanisms 1 and 2 is constrained with lab-determined hydrologic and poroelastic properties of fault core materials and lab friction studies at high slip rates. Predictions are that strength drop should often be nearly complete at large slip and that the onset of melting should be precluded over much (and, for small enough slip, all) of the seismogenic zone. A testable prediction is of the shear fracture energies that would be implied if actual earthquake ruptures were controlled by those thermal mechanisms. Seismic data have been compiled on the fracture energy of crustal events, including its variation with slip in an event. It is plausibly described by theoretical predictions based on the above mechanisms, within a considerable range of uncertainty of parameter choices, thus allowing the possibility that such thermal weakening prevails in the Earth.


1. Introduction

[2] Earthquakes occur because fault strength weakens with increasing slip or slip rate. What physical processes determine how that weakening occurs? The question is addressed here for mature crustal faults with highly granulated cores, that are capable of producing large earthquakes. Recent field observations suggest that slip in individual events may then be extremely localized, and may occur primarily within a thin shear zone, <1–5 mm thick. That localized shear zone lies within a finely granulated (ultracataclastic) fault core of typically tens to hundreds millimeter thickness, that core itself fitting within a much broader damage zone of granulated or incompletely cracked rock. Evidence for that morphology has emerged from studies of the exhumed, and now inactive, North Branch San Gabriel fault [Chester et al., 1993] and Punchbowl [Chester and Chester, 1998; Chester et al., 2004] fault of the San Andreas system in southern California, the Median Tectonic Line fault in Japan [Wibberley and Shimamoto, 2003], of the Hanaore fault in southwest Japan [Noda and Shimamoto, 2005], the Nojima fault [Lockner et al., 2000] which ruptured in the 1995 Kobe, Japan, earthquake and has been penetrated by drill holes, and from other observations summarized by Sibson [2003], Ben-Zion and Sammis [2003], Biegel and Sammis [2004], and Rice and Cocco [2006].

[3] As perhaps the best characterized case, a thin “principal slip surface” (PSS), Figure 1, was identified along an exposure of the Punchbowl fault which is exhumed from 2 to 4 km depth and which has accommodated 44 km of slip. The PSS was argued to have accommodated “several km” of slip [Chester and Chester, 1998; Chester et al., 2004]. Subsequent studies by J. S. Chester et al. (Extreme localization of slip and implications for dynamic weakening of faults, manuscript in preparation, 2005, hereinafter referred to as Chester et al., manuscript in preparation, 2005) [see also Chester et al., 2003; Chester and Goldsby, 2003] of a thin section sample (Figure 1b) showed that the nominal thickness of that shear zone, as revealed by uniform birefringence in crossed polarizers due to preferred orientation of the sheared minerals within it, varies from 0.6 to 1.1 mm at different locations along the specimen. However, within this ~1 mm thick “nominal” shear zone, most of the shearing seems to have been accommodated within a zone (dark in Figure 1b) of extreme shear localization having an apparent thickness of 100–300 μm. Such thicknesses on the order of 200 μm are small compared to hydraulic and thermal boundary layers developed near faults during significant earthquakes. Thus it will be appropriate for some purposes to address the thermal and hydraulic state near a fault as if the slip occurred in a zone of zero thickness, i.e., as slip on a plane, as considered
by Mase and Smith [1987] and Lee and Delaney [1987]. That model is the main focus of this paper, and some new analytical results are derived for it. A subsequent manuscript (A. W. Rempel and J. R. Rice, Thermal pressurization and onset of melting in fault zones, submitted to Journal of Geophysical Research, 2006, hereinafter referred to as Rempel and Rice, submitted manuscript, 2006) builds on this work and provides numerical estimates for zones of small but finite thickness, addressed also by Lachenbruch [1980] and Mase and Smith [1985, 1987].

4 Also, a thin “central slip zone” was identified by Wibberley and Shimamoto [2003, Figure 3] along the Median Tectonic Line fault; subsequent measurements (C. A. Wibberley, private communication, 2003) showed that it has a nominal thickness of $\sim 24\text{ mm}$, although there has been no study reported of possible fine scale features within it as for the case just discussed.

1.1. Localization in Granular Systems

5 It is useful to compare the apparent thickness of the Punchbowl shear zone to other experimental estimates of the minimum thickness of a localized flow in a disordered granular material. Such shear cannot be localized to a mathematical plane because other particles get in the way and must move (and/or crack) to accommodate the motion. Thus some broader zone of particles must participate in the shear, if not instantaneously, at least when averaged over slips of order of a typical particle diameter or more. Localized shear is widely observed in geomechanics and several attempts have been made to characterize the thickness of shear bands in soils. Morgenstern and Tschalenko [1967] illustrate shear bands for a clay, kaolin, and thicknesses of order 10–100 times a clay platelet diameter ($\sim 0.5\text{ \mu m}$) seem representative. For sands and other granular materials, thickness scales principally with the “mean particle diameter” $d_{50}$, defined for nonuniform size distributions such that 50% by weight of the particles have larger size.

5 Tordesillas et al. [2004, pp. 982–983] summarize that In real sands, reported shear band widths range from around 7–8 $d_{50}$ [Oda and Kazama, 1998] up to over 30 $d_{50}$ [see Oda and Iwashita, 1999, chapter 4]. There is some evidence to suggest that the initial packing density influences the shear band width, with denser assemblies producing thinner shear bands. Harris et al. [1995] reported shear band widths of up to 17 $d_{50}$ in loose granular assemblies, compared with widths of around 10 $d_{50}$ reported, for example, by Roscoe [1970], Scarpelli and Muir Wood [1982] and Tatsuoka et al. [1990] for dense assemblies. [Citation format changed and some wording replaced with $d_{50}$ notation.]

Similarly, Muir Wood [2002, Table 1] shows values of thickness ranging from 7 to 19 $d_{50}$; that includes some results from above and also, e.g., the Muhlhaus and Vardoulakis [1987] thickness of $\sim 16 d_{50}$ from plane strain compression tests.

7 If these apply to the ultracataclasite of concern here the shear zone thickness should be of order 10–20 $d_{50}$. Chester et al. [2005] [see also Chester et al., 2004] show the size distribution for the Punchbowl ultracataclasite hosting the PSS from optical and TEM studies, based on binning particles by effective diameter $d$ into size ranges that are each a factor of two smaller than the next larger bin, starting at 100 $\mu$m size (a bin centered on $d$ has $2d/3 < \text{size} < 4d/3$). The number of particles per unit of macroscopic area sampled is accurately described as $\propto 1/d^2$ for bin centers $d$ between 50 nm and 50 $\mu$m. However, the number of particles in the 100 $\mu$m bin has only $\sim 0.002$ times as many...
particles as predicted by that scaling, so that the population is, statistically, essentially devoid of particles above 67 μm. Thus the areal distribution \( n_4(d) \) is \( \propto 1/d^2 \), where \( n_4(L)dL \) is the number per unit area with size \( L < d < L + dL \). The distribution \( n_4(d) \) per unit volume contains an extra \( 1/d \) [e.g., Sammis et al., 1987] and thus is \( \propto 1/d^2 \) over a range which should be assumed to extend from as small as \( 10 – 30 \) nm to an upper limit of \( \approx 67 \) μm, and then cut off at both ends. That is extraordinarily broad, and leads to \( d_{50} \approx 1 \) μm. Thus the thickness of a shear band in the ultracataclasite is predicted to be \( 10 – 20 \) μm by the conventional guidelines, much smaller than the 100–300 μm range inferred from the thin section (Figure 1b), although the prediction is paradoxical because 20–25% of the ultracataclasite mass has size \( > 10 \) μm. A possibility is that \( d_{50} \) does not provide a relevant scale for very broad size distributions; essentially all the thickness studies have been done on sands with a comparatively narrow distribution [Muir Wood, 2002] (e.g., \( d_{10}/d_{50} < 5 \) typically, whereas \( d_{10}/d_{50} \approx 100 \) for the ultracataclasite). Also, interparticle cohesion would be much more important for the small particles of the ultracataclasite, and clumping of smaller particles could raise the effective \( d_{50} \) for purposes of localization.

For the MTL, Japan, ultracataclasite containing the PSS, Figure 5 of Wibberley and Shimamoto [2003] shows the cumulative weight percent with size \( > d \), versus \( d \), but only for \( d > 10 \) μm. The results imply that \( d_{50} < 10 \) μm in that case too. Also, extrapolating a \( d = 10 \) μm corresponds to \( d_{20} \) to \( d_{10} \) in that case, very consistent with the correspondence of \( d = 10 \) μm for the above Punchbowl distribution with \( d_{20} \) to \( d_{50} \). So again, a very fine shear localization is expected based on the conventional guidelines.

The ultracataclasite is sheared naturally under vastly larger normal stresses than for the clay and sand experiments mentioned, so that particle cracking might normally be thought to be an issue. However, it can be argued [Sammis et al., 1987; Biegel et al., 1989] that the broad size distribution is the result of prior constrained comminution, which has evolved the gouge to a state at which it is protected against further particle cracking, by assuring that all particles, down to sizes that are too small to crack, have contact with abundant smaller neighbors and thus experience no exceptionally large individual contact forces.

While granular materials often show localized shear, even under slow quasi-static and isothermal deformation, the thermal weakening mechanisms addressed in this paper are of themselves localization promoting, so that it is appropriate as a first approximation to treat them as if slip were occurring on a plane. Nevertheless, the difference between zero and, say, 200 μm for the width of the shearing zone is not negligible for all purposes, especially for estimating maximum temperature, and the discussion in sections B6 and B7 addresses this point (see also Rempel and Rice, submitted manuscript, 2006). At the particle level, it should probably not be assumed, for a highly localized granular shear zone, that shear at any given moment in time is uniformly distributed over the thickness, such that the average slip rate between individual particles is a small fraction of the net slip rate \( V \) that is accommodated. Rather, it seems plausible that shear is highly concentrated at each moment in time, with the most active sites having slip rates more nearly comparable to \( V \), but with the locations of those active sites shifting rapidly over the nominal thickness. Numerical simulations as by da Cruz et al. [2005] of granular flows, albeit for a narrow size distribution, support this picture, particularly for the range of their dimensionless inertial shear parameter \( I < 0.003 \) which they call the “quasi-static regime” and which is characterized by intermittency of shear at the particle scale. That range is achieved in gouge, with density \( \sim 2700 \) kg/kg/m³, when fault slip at 1 m/s is accommodated over thicknesses \( > 5d \) (which should always be the case), at effective normal stress >10 MPa.

1.2. Implied Temperature Rises in Absence of a Weakening Mechanism

Unless we appeal to some mechanism to rapidly diminish fault strength as slip accumulates, that evidence for narrowness of the zone where frictional work is dissipated would imply temperature rises which far exceed those for onset of melting. For example, 7 km is a representative centroidal depth of the region slipping in crustal earthquakes, and at that depth the ambient temperature is typically \( \sim 200^\circ \text{C} \) and the effective overburden pressure is \( \sim 125 \) MPa, assuming hydrostatic pore pressure. Equating that to the effective normal stress \( \tau \) on a fault, and assuming a friction coefficient \( f = 0.6 \), the shear strength is \( \tau = 75 \) MPa for onset of slip.

Supposing, implausibly as will be seen, that \( \tau \) remains constant during slip at (for simplicity) a constant slip rate \( V \) and that the slip is accommodated by uniform shear over a fault core of thickness \( h \), the temperature rise can be calculated from equation (B7a) of Appendix B, using in its integrand the function \( B \) defined by equation (B8) in place of the function \( A \). Then, assuming that all changes in internal energy are accountable as temperature change, the temperature rise at the layer center in an event with slip \( \delta = Vt \) is accurately given by \( \Delta T = \tau \delta/(\rho C h) \) (the simple result for adiabatic heating), provided \( h > 4V/\sqrt{\alpha_{gh} V} \). The \( \Delta T \) is that amount at the layer edges, due to conductive heat loss. Here \( \rho c_p \approx 2.7 \) MPa/C and \( \alpha_{gh} = 0.70 \) mm²/s [Vøsteen and Schellschmidt, 2003; Lachenbruch, 1980] are the specific heat per unit volume and thermal diffusivity of the fault gouge, respectively.

To choose \( V \), note that Heaton [1990] reports seismic slip inversions for seven shallow earthquakes, giving for each the average slip and average slip duration at a generic point on the fault. The ratio of slip to duration defines the average slip rate \( V \) during an event; it ranges from 0.56 to 1.75 m/s, with average of 1.06 m/s for the seven events. Therefore \( V \) in this paper, when considered constant, is taken as 1 m/s. For an earthquake of slip \( \delta = 1 \) m, the condition for use of the above adiabatic heating formula at the layer center is therefore \( h > 3.3 \) mm.

Thus, for \( h = 10 \) mm, the temperature rise if \( \tau \) remains constant during slip would be 2780°C at the layer center and 1380°C at the edge. For \( h = 5 \) mm, both numbers double, to 5560°C at the layer center. For \( h = 0 \), corresponding to slip on a plane as described by equation (B11a), the temperature rise would be \( \Delta T = (\tau/\rho C)^{1/2} V/\pi \alpha_{gh} = 15,670^\circ \text{C} \). All of those numbers correspond to temperatures which are well above 1000°C, a nominal value for the equilibrium melting temperature of wet granitic compositions in the shallow crust (the most vulnerable constituents, such as biotite, would begin to melt at 750°C [e.g., Otsuki et al., 2003]). To keep the
temperature rise at the layer center at 7 km depth below 1000°C, i.e., to keep $\Delta T < 800$°C, would require $h > 35$ mm. That thickness seems inconsistent with the observation of narrow shear zones. The predicted temperature rises scale directly with $\tau$, and hence in this type of illustration, scale directly with depth. They would be twice larger at 14 km depth. For $h = 5$ mm, melt onset would occur at all depths $> 1.2$ km. Yet evidence for melting in the form of pseudotachylytes is relatively rare. Where it is found and is attributable to tectonic faulting, the inferred depths are generally toward the lower reaches of the seismogenic zone [Sibson, 1975].

Hence it seems that strength $\tau$ cannot remain constant during earthquake slip, at least at a static threshold corresponding to $f \sim 0.6$, but must weaken quite substantially to explain the apparent absence of pervasive melt. Further the weakening process must be effective enough to allow for extremely thin shear zones, as in cases that might well be described as slip on a plane ($h \approx 0$). Given the evidence for such thin shear zones, which would suffer extraordinary temperature rises if $\tau$ remained constant, it seems likely that that relevant weakening processes are of thermal origin, as now discussed.

1.3. Thermal Weakening Mechanisms

The possibility is examined in this work that the primary weakening mechanisms for shallow earthquakes on mature faults are indeed thermal in origin, but operate well before melting conditions are achieved, and sometimes preclude melting. They involve the following processes, which are assumed to act in combination: (1) Thermal pressurization of pore fluid within the fault core by frictional heating which reduces the effective normal stress and hence reduces the shear strength $\tau$ associated with any given friction coefficient $f$ [Sibson, 1973; Lachenbruch, 1980; Mase and Smith, 1985, 1987; Lee and Delaney, 1987; Andrews, 2002; Wibberley, 2002; Noda and Shimamoto, 2005; Sulem et al., 2005; Bizzarri and Cocca, 2006a, 2006b], a mechanism which has also been suggested for weakening in catastrophic landslides [Habib, 1967, 1975; Anderson, 1980; Voight and Faust, 1982; Vardoulakis, 2002], and (2) flash heating and consequent weakening at highly stressed microcontacts during rapid slip, which reduces the friction coefficient $f$, a process studied for some years in connection with high speed friction in metals [Bowden and Thomas, 1954; Archard, 1958/1959; Ettles, 1986; Lim and Ashby, 1987; Lim et al., 1989; Molinari et al., 1999] and which has recently been considered as a process active during earthquake slip [Rice, 1999; Tullis and Goldsby, 2003a, 2003b; Hirose and Shimamoto, 2005; N. M. Beeler and T. E. Tullis, Constitutive relationships for fault strength due to flash-heating, submitted to U.S. Geological Survey Open File Report, 2003, hereinafter referred to as Beeler and Tullis, submitted manuscript, 2003].

Simple, elementary models of those two processes are developed here. Neither process is new to discussions of seismicity. Nevertheless, the recent work on fault zone structure mentioned above, emphasizing the thinness of shear zones, as well as laboratory measurements of permeability and poroelastic properties of fault core materials (Lockner et al. [2000] for the Nojima fault; Wibberley [2002] and Wibberley and Shimamoto [2003] for the Median Tectonic Line fault; Sulem et al. [2004] for the Aegion fault; Noda and Shimamoto [2005] for the Hanaore fault) for mechanism 1, and of frictional properties of rocks at high slip rates [Tullis and Goldsby, 2003a, 2003b; Prakash et al., 2004; Hirose and Shimamoto, 2005; Prakash and Yuan, 2004, also private communication on work in progress at Case Western Reserve University on high speed friction studies in quartz with the torsional Kolsky bar, November 2004] for mechanism 2, makes it possible to insert with some confidence numerical values for parameters into the equations describing those processes. The focus here is on rapid and at least moderately large slips. By the latter is meant slips much larger than those of order 0.01 to 0.1 mm over which unstable sliding is thought to nucleate, according to current understanding of rate- and state-dependent friction concepts.

Predictions from the modeling of mechanisms 1 and 2 to follow are that strength drop should be nearly complete at large slip and that macroscopic melting should often be precluded over most of the seismogenic zone. Those predictions are qualitatively consistent with low heat outflow from major faults and a scarcity of glass (pseudotachylyte) that would be left from rapid recooling of silicate melts.

A more quantitatively testable prediction is that of the shear fracture energies $G$ that would be implied if actual earthquake ruptures were controlled by those two thermal mechanisms. The simple model allows prediction of stress $\tau$ versus slip $\delta$ during events, written here as $\tau = \tau(\delta)$, and thereby a prediction of the fracture energy. That fracture energy was classically defined by Palmer and Rice [1973] for slip weakening laws which had a well defined residual strength $\tau_0$, at large slip, such that $\tau(\delta) = \tau_0$, for $\delta > \delta_1$, a given constant. (In the independently formulated Ida [1972] description of slip weakening, the residual strength was taken as zero.) Then for ruptures with maximum slip greater than $\delta_1$, the fracture energy is given by

$$ G = \int_0^{\delta_1} [\tau(\delta') - \tau_0] d\delta'. $$

However, it will be seen that the mechanisms considered here lead to a $\tau = \tau(\delta)$ for which weakening continues, but at an ever decreasing weakening rate $-d\tau(\delta)/d\delta$, out to very large slip. Abercrombie and Rice [2005] have in fact argued, purely from seismic data, that a slip-weakening description with that feature is required to fit the observations. In that case, a generalization of the $G$ expression which they used, and which is consistent with the Palmer and Rice [1973] derivation, is to define $G$ for an event with slip $\delta$ as

$$ G = G(\delta) = \int_0^\delta [\tau(\delta') - \tau(\delta)] d\delta', $$

and that is used here. To write of a relation $\tau = \tau(\delta)$ is, in fact, a simplification adopted here for a first test of the proposed weakening mechanisms against seismic constraints. More generally, the physics underlying the mechanisms requires that $\tau$ be regarded as some complex functional of the slip rate history at all times up to the present. Thus, when summarizing that complex result in the form $\tau = \tau(\delta)$ here, it is to be understood that $\tau$ has been...
evaluated based on mechanisms 1 and 2 for a slip rate history \( V = \text{const} = 1 \text{ m/s} \) and that \( \delta \) is just a proxy for \( V t \).

### 1.4. Other Weakening Mechanisms

[20] Mechanisms 1 and 2 are expected to become important immediately after seismic slip initiates, but as large slip develops they may not always remain the most significant weakening mechanisms. Macroscopic melting (i.e., when a coherent melt layer has formed along the whole sliding surface) may occur too for sufficiently large combinations of slip and initial effective normal stress. It is not addressed here except to identify, within limitations of the simple modeling developed, maximum temperatures expected on the basis of processes 1 and 2, to see if they would lead to onset of melting. Melts, if sufficiently hot, have a low viscosity and may lubricate faults reducing dynamic friction [Sibson, 1975; Spray, 1993; Brodsky and Kanamori, 2001]. However, melting is not a simple weakening mechanism. During rapid shear at modest effective normal stress, in the range just prior to the transition of macroscopic melting, there is an abrupt increase in frictional strength [Tsutsumi and Shimamoto, 1997; Hirose and Shimamoto, 2005]. The strength decreases with continued shear and heating, which increases melt volume and raises its temperature, reducing its viscosity. At the start of the transition, microscopic blobs of melt form near the larger frictional contact asperities, as an extreme form of the flash heating process 2, and get smeared out along the sliding surfaces and rapidly solidify, at least when the average temperature of those surfaces is low enough, so as to locally weld the surfaces together. Fiško [2004] has therefore suggested that the onset of melting may sometimes arrest slip.

[21] Another weakening mechanism by gel formation [Goldsby and Tullis, 2002; Di Toro et al., 2004] has been identified too in silica-rich lithologies, when sliding produces fine wear debris in presence of water. It, rather than melting which it could preclude, might be the dominant weakening mechanism at large slip in some situations. Severe weakening was reported in a pure quartzite rock, Arkansas novaculite, by Goldsby and Tullis [2002] and Di Toro et al. [2004] when subjected to large slip (>0.5–1.0 m), at least at moderately rapid rates (>1 mm/s). Di Toro et al. [2004] identified the underlying process as silica gel formation, based on observations of slip surface morphologies which showed (T. E. Tullis, private communication, 2004) “now solidified... flow-like textures that make it... evident that at the time the deformation was going on, a thin layer coating the sliding surface was able to flow with a relatively low viscosity”. Further evidence of the importance of silica content is that Roig Silva et al. [2004a, 2004b] showed that the susceptibility of different rocks to this type of weakening is directly ordered as their silica content, such that the order of weakening and silica content is as follows: quartzite (Arkansas novaculite) > granite (Westerly) ≈ feldspar (Tanco albite) > gabbro. However, Tullis reports that the “solidified flow structures have so far only been seen for novaculite”. It is presumed that shear continuously disrupts silica particle bonding so that the fluidized gel mass deforms at low strength, although it would gradually consolidate into a strong, amorphous solid in the absence of continued shearing, then showing a fractionally much greater increase of strength than what a granular gouge would show on a comparable timescale. The onset of this weakening process seems to require moderately large slip according to laboratory studies thus far available, and it is not yet known whether it would normally contribute significantly to seismic weakening before the two processes analyzed here would have substantially reduced strength.

### 1.5. Poromechanical Properties of Fluid-Infiltrated Fault Core Rocks

[22] The present modeling of mechanism 1 assumes the presence of fluids, particularly water, within shallow crustal fault zones, such that the concept that the effective normal stress \( \sigma_n \) controls frictional strength will be valid (\( \sigma_n = \sigma_n - p \), where \( \sigma_n \) is the compressive normal stress on the fault and \( p \) is pore fluid pressure). That is reasonable in the sense that the fault zone is a granular material and should be fluid-infiltrated below the water table. However, it is well to note that the effective stress principle presently seems to lack laboratory verification (or disproof) at the high slip rates of interest. Mineralization and slow creep in hotter portions of the crust may isolate fluid-filled pores from one another so that the fluid phase loses the interconnectivity that it had shortly after the last earthquake, whereas the poroelastic-plastic modeling developed here for mechanism 1 assumes that gouge near the slip zone forms a porous material with an interconnected pore space. Such initial isolation of fluid pores may, however, not invalidate the modeling. Recent studies [Poliakov et al., 2002; Andrews, 2005; Rice et al., 2005] emphasize that the high local stress concentration associated with the propagating front of the slip rupture zone will cause stresses that are consistent with some nonelastic deformation in fault gouge material adjoining the slipping zone. It may assumed that such deformation acts, very near the tip, to rupture cementation and restore fluid phase connectivity as large slips develop in the wake of the passage of the rupture front. (In the 1 s that it typically takes for 1 m of slip to accumulate at a point, the rupture front will have moved 2 to 3 km. The most significant alterations of temperature and pore pressure take place in zones whose length scales occupy a few millimeters to a few tens of millimeters measured perpendicular to the fault surface. Thus the interaction with the rupture front is short lived, and the fault response can be analyzed as heat and fluid transport which has significant variation with space coordinates almost only in the direction perpendicular to the fault.) Nonelastic deformation off the fault plane will also contribute to the overall fracture energy in addition to the slip-weakening contribution identified in the expression for \( G(\delta) \) above, but that is neglected in this first study. Nevertheless, predicted \( G \) values may be smaller than actual \( G \) for that reason.

[23] Several investigations have attempted to constrain the values of hydraulic permeability and diffusivity within the fault zone as well as to estimate the value of permeability [see Lockner et al., 2000; Wibberley and Shimamoto, 2003; Sulem et al., 2004; Noda and Shimamoto, 2005, references therein]. These studies have shown that the ultracataclastic gouge zones forming the fault core have a much lower permeability than that measured in the surrounding damage zone, which can be highly variable. Permeability within the fault core (\( \sim 10^{-19} \text{ m}^2 \)) can be 3 orders of
magnitude smaller that that in the damage zone ($\approx 10^{-16} \text{ m}^3$). In both regions, permeability is reduced as the effective normal stress is increased. For example, in the case of the ultracataclastic gouge core material containing the slip zone in the MTL, after isotropic compaction to 180 MPa effective stress, at which the permeability is $3 \times 10^{-21} \text{ m}^2$, the permeability along the unloading path is approximately $4 \times 10^{-21} \text{ m}^2$ at 120 MPa, $10^{-20} \text{ m}^2$ at 70 MPa, and $10^{-19} \text{ m}^2$ at 10 MPa [Wibberley and Shimamoto, 2003, Figure 8.b.ii]. For a hydrostatic pore pressure, 126 MPa corresponds to the effective overburden stress at 7 km, a representative centroidal depth for the slip zone of crustal earthquakes, and permeability along an unloading path from that value is shown in Table 1 (along with other material parameters introduced in section 3 and Appendix A), and has been estimated as explained in Table 1.

### 2. Flash Heating and Weakening of Microasperity Contacts

[24] Flash heating at frictional asperity contacts has been suggested in engineering tribology as the key to understanding the slip rate dependence of dry friction in metals at high rates [Bowden and Thomas, 1954; Archard, 1958/1959; Ettes, 1986; Lim and Asby, 1987; Lim et al., 1989; Molinari et al., 1999]. It has also been considered recently in seismology as a mechanism that could be active in controlling fault friction during seismic slip [Rice, 1999; Tullis and Goldsby, 2003a, 2003b; Prakash, 2004; Hirose and Shimamoto, 2005; Prakash and Yuan, 2004, also private communication, 2004; Beeler and Tullis, submitted manuscript, 2003], in the range before macroscopic melting. The presentation here follows Rice [1999] but with an improved basis for choice of material parameters and also with a modification of the basic model as was proposed by Beeler and Tullis (submitted manuscript, 2003).

[25] We consider an asperity contact which has just formed and that will persist for a slip $D_s$ before it is slid out of existence. Letting $T$ be the gradually evolving average temperature along the sliding fault surface, the contact may be assumed to have that temperature $T$ when it first forms but, if sliding is fast, the temperature $T_a$ at the contacting asperity interface will undergo a local, highly transient, rise due to frictional heating (i.e., “flash” heating) during its brief lifetime $\theta$. That lifetime is $\theta = D_s/V$ where $V$ is the slip rate.

#### 2.1. Threshold Conditions for Onset of Weakening Due to Flash Heating

[26] The local shear strength $\tau_c$ of the asperity contact interface will presumably degrade continuously with increasing $T_a$. A highly simplified yet informative first model [Rice, 1999] is to assume that $\tau_c$ is constant with respect to $T_a$ up until when $T_a$ reaches a weakening temperature $T_w$, and then decreases dramatically when $T_a > T_w$. As a simple model, Rice [1999] assumed that for $T_a > T_w$, the weakened contact shear strength $\tau_{cw}$ was negligibly small compared to its lower temperature value $\tau_c$. Beeler and Tullis (submitted manuscript, 2003) and Tullis and Goldsby [2003a, 2003b] observed that a better fit of the model to data showing strong rate weakening of friction was to assume a small but nonnegligible $\tau_{cw}$ for $T_a > T_w$. This discussion neglects the variation of $\tau_c$, with slip rate due, presumably, to thermally activated slip at the contacts [e.g., Rice et al., 2001] which is associated with the “direct effect” in rate and state friction, and also any variations in $\tau_c$ that may be associated with the maturing of contacts. While important for nucleations of instability, those effects generally correspond to modest fractional changes in the friction coefficient $f$, and tend to be dwarfed by the much larger
weakening of \( f \) that is predicted based on the thermal process now studied.)

[27] The net work per unit area dissipated at the asperity contact during its lifetime is \( \tau_{\text{c}}D_{\text{a}} \), if there is no thermal weakening. However, will there be weakening? That clearly depends on \( V \); if it is too large, heat cannot be transferred away fast enough and the asperity interface will have reached temperature \( T_{\text{a}} \) at some time \( \theta_{\text{w}} \) which is less than the contact lifetime \( \theta_{\text{c}} \). We can estimate [Rice, 1999]

\[
\theta_{\text{w}} = \frac{\pi \alpha_{\text{th}}}{V^2} \left( \frac{p \epsilon (T_{\text{a}} - T)}{\tau_{\text{c}}} \right)^2
\]

(3)

where \( \alpha_{\text{th}} \) is the thermal diffusivity and \( p \epsilon \) is the heat capacity per unit volume (\( \rho \alpha_{\text{th}} \) is the heat conductivity). That result is obtained using the equations of one-dimensional heat conduction for a planar heat source [Carslaw and Jaeger, 1959] with heat input into the solids on both sides of the contact interface at rate \( \tau_{\text{c}}V/2 \), as in the “fast moving heat source” limit of Archard [1958/1959]. Thus the total heat input \( \tau_{\text{c}}V\theta_{\text{w}} \) per unit area at the contact, up to the weakening time, is equal to the maximum thermal energy storage \( p \epsilon (T_{\text{a}} - T) \) per unit volume times the effective heated distance \( \sqrt{\pi \alpha_{\text{th}} \theta_{\text{w}}} \), which distance is defined by matching this explanation of the solution to the exact one-dimensional heat conduction analysis.

[28] The expression given for \( \theta_{\text{w}} \) is relevant if \( \theta_{\text{w}} < \theta = D_{\text{a}}/V \). That means it is relevant for \( V > V_{\text{w}} \) where the weakening velocity \( V_{\text{w}} \) so defined is obtained by setting \( \theta_{\text{w}} = D_{\text{a}}/V \), so that

\[
V_{\text{w}} = \frac{\pi \alpha_{\text{th}}}{D_{\text{a}}} \left( \frac{p \epsilon (T_{\text{a}} - T)}{\tau_{\text{c}}} \right)^2.
\]

(4)

Thus \( V_{\text{w}} \) is the critical sliding velocity, below which the contact does not weaken during its lifetime but above which it does.

[29] To estimate \( V_{\text{w}} \), recognizing that we will be concerned with temperatures ranging from ambient up to \( \sim 900^\circ \text{C} \), rough average values are \( \alpha_{\text{th}} = 0.5 \text{ mm}^2/\text{s} \) and \( p \epsilon = 2.7 \text{ MPa}\cdot\text{s} \) [Veesen and Schellenschmidt, 2003]. The slip \( D_{\text{a}} \) over which the contact exists should be identified approximately with a sliding distance to renew the asperity contact population, which means that it should be comparable to the state evolution slip distance of rate and state friction, variously denoted as \( d_{\text{c}} \) or \( L_{\text{fr}} \) at least in situations for which the frictional slip is strongly localized and not diffusely spread though a shear fault gouge [e.g., Marone, 1998]; \( D_{\text{a}} = 5 \mu \text{m} \) is taken for illustration here. The weakening temperature \( T_{\text{a}} \) cannot be in excess of the inferred temperature of pseudotachylyte melts, 1000°C to 1450°C [Spray, 1993, 1995; Ray, 1999, O’Hara and Sharp, 2001], and should be closer to the lower limit of that range since here we are considering the range prior to macroscopic melting. So \( T_{\text{a}} = 900^\circ \text{C} \) is considered here. Molinari et al. [1999] emphasize that slightly submelting \( T_{\text{a}} \) values can also greatly degrade strength, at least in metals, so that a layer of melt may not be required for weakening by flash heating. In most ceramics and rocks, it is possible that the weakened condition involves a thin layer of melt, perhaps at submicron scale, which rapidly refreezes in quenching by surrounding cooler material when the contact phase ends. Ultimately, continued heat input through the asperities causes the fault average temperature \( T \) to become too large to quench rapidly, so that such a melt layer would smear out and possibly reweld in places, beginning the complex transition to macroscopic melting discussed earlier and documented by Tsutsumi and Shimamoto [1997] and Hirose and Shimamoto [2005], but which we do not consider here. \( T = 200^\circ \text{C} \) is an ambient fault zone temperature toward the midlevel of the crustal seismogenic zone, so that \( T_{\text{w}} - T \approx 900^\circ \text{C} - 200^\circ \text{C} = 700^\circ \text{C} \). Recent measurements of contact area in transparent materials by light scattering [Dieterich and Kilgore, 1994, 1996], including quartz, confirms earlier suggestions [Boitnott et al., 1992] that in brittle rock materials \( \tau_{\text{c}} \) is of order 0.1 \( \mu \text{m} \), where \( \mu \) is the elastic shear modulus; 0.1\( \mu \) is a standard elementary estimate of the theoretical shear strength of a crystalline solid. (The result is also supported by measurements of contact indentation strength \( \sigma_{\text{t}} \) [Dieterich and Kilgore, 1994, 1996]. Because average contact stresses satisfy \( \sigma_{\text{t}} = \sigma_{\text{A}} = \sigma_{\text{A}} = \tau_{\text{A}} \), where \( \sigma_{\text{A}} \) is true contact area and \( A \) is nominal area, and \( \sigma \) and \( \tau \) are the macroscopic normal and shear stresses, a measurement of the friction coefficient \( f \) during sliding gives, through \( f = \tau/\sigma = \tau_{\text{c}}/\sigma_{\text{c}} \), a way of estimating \( \tau_{\text{c}} \). So we take \( \tau_{\text{c}} = 3.0 \text{ GPa} \). This gives \( V_{\text{w}} = 0.12 \text{ m/s} \) for onset of severe thermal weakening. The scaling of the result with the most uncertain parameters is

\[
V_{\text{w}} = 0.12 \frac{\text{m}}{\text{s}} \left( \frac{5 \mu \text{m}}{D_{\text{a}}} \right) \left( \frac{T_{\text{w}} - T}{700^\circ \text{C}} \right)^{1/2} \left( \frac{3 \text{ GPa}}{\tau_{\text{c}}} \right)^{1/2}.
\]

(5)

For example, with all else the same except for \( T = 20^\circ \text{C} \) to represent a room temperature lab experiment, the estimate is \( V_{\text{w}} = 0.20 \text{ m/s} \). During large rapid slip the average temperature \( T \) of the fault interface will slowly increase (slowly compared to the timescale over which asperity temperatures \( T_{\text{a}} \) flash up toward \( T_{\text{a}} \)) and, if all other parameters involved have only modest dependence on \( T \), the threshold velocity for onset of flash weakening should diminish.

2.2. Model for Friction Coefficient When \( V > V_{\text{w}} \)

[30] A similarly simple estimate of the effect on macroscopic friction is given by neglecting the statistics of \( D_{\text{a}} \) and other parameters, and noting that when \( V > V_{\text{w}} \) (which means when \( \theta_{\text{w}} < 0 \), the contact spends a fraction \( \theta_{\text{w}}/\theta = V_{\text{w}}/V \) of its lifetime at the initially high strength \( \tau_{\text{c}} \) and the remaining fraction at the weakened strength \( \tau_{\text{c,w}} \), so that the average contact shear strength during its lifetime is

\[
(\tau_{\text{c}})_{\text{avg}} = \tau_{\text{c}} \theta_{\text{w}} + \tau_{\text{c,w}} (1 - \theta_{\text{w}}/\theta) = (\tau_{\text{c}} - \tau_{\text{c,w}}) V_{\text{w}}/V + \tau_{\text{c,w}}
\]

(6)

whereas if the severe heating is confined to a thin layer (compared to contact size \( D_{\text{a}} \)), \( \sigma_{\text{t}} \) is unaltered from its lower temperature value. Thus, setting \( f = (\tau_{\text{c}})_{\text{avg}}/\sigma_{\text{c}} \), this simple model gives

\[
f = (f_{\text{w}} - f_{\text{c}}) V_{\text{w}} + f_{\text{c}} \text{ when } V > V_{\text{w}}.
\]

(7)
where \( f_o = \tau_c/\sigma_c \) is the low speed friction coefficient and \( f_w = \tau_{c,w}/\sigma_c = f_o \tau_{c,w}/\tau_c \) is the friction coefficient based on the weakened shear strength at high temperature. The model of Rice [1999] gave this equation for \( f \) but with \( \tau_{c,w} \) and \( f_w \) considered negligibly small; their finite values were added by Beeler and Tullis (submitted manuscript, 2003); \( \tau_{c,w} \) and hence \( f_w \), could not plausibly be independent of \( T_w \) and \( V \) (because they, together with the thickness of a thin melt layer, determine the viscous shear response at the contact), but are nevertheless treated as constant in section 2.3 to simplify the description.

[31] Equation (7) must be regarded as an expression for \( f \) in steady state sliding (i.e., sliding at effectively constant \( V \) over slip distances that are large enough to completely renew the contact population). A fuller description would have to build in well-known effects that are described within rate and state friction, e.g., with the feature that \( \partial f/ \partial V > 0 \) (positive “direct effect”) when \( V \) is changed instantaneously at a fixed state of the contact population. Such a feature is not only physically motivated, but is known to be necessary to make mathematically well-posed models of sliding, with velocity-weakening friction, on interfaces between elastically deformable continua [Rice et al., 2001].

2.3. Comparisons to Experimental Data at High Slip Rates
[32] At the time of presentation of the model discussed, the data available for \( f \) at rates \( V \) approaching the seismic range were those of Tsutsumi and Shimamoto [1997]. They were mainly concerned with melting in their annular shear experiments on gabbro, but they did identify some data points as being taken prior to melting range [see also Hirose and Shimamoto, 2005], and those showed weakening qualitatively consistent with the above concepts. Since then there has been great progress on laboratory experiments imposing rapid slip [Tullis and Goldsby, 2003a, 2003b; Prakash, 2004; Prakash and Yuan, 2004, also private communication, 2004], and the concept of substantial weakening at rapid slip rate is definitively supported. See Figure 2.

[33] The expression for \( f \) in terms of \( f_o, f_w, V_w \), and \( V \) was fit by Tullis and Goldsby [2003a, 2003b] to their experimental results for five rock types that were tested in annular shear at \( V \) up to 0.36 m/s (Figure 2a, for quartzite), of which all but one showed evidence of strong weakening at higher \( V \).

Results for \( V_w \) as read off approximately from the plots of Tullis and Goldsby [2003b] are as follow: \( V_w \approx 0.14 \) m/s for quartzite (Arkansas novaculite) and also for granite; 0.28 m/s for feldspar (Tanco albite); 0.11 m/s for gabbro; and 0.27 m/s for calcite (weakening for it was far less pronounced than for the other rocks). These \( V_w \) are all comparable to the rough estimate \( V_w \approx 0.2 \) m/s made above.

[34] Also, values of \( f_w \) and \( f_o \) associated with the Tullis and Goldsby [2003b] fits can be inferred from their plots for some of the rocks, with the results being as follows: \( f_o \approx 0.64 \) and \( f_w \approx 0.12 \) for quartzite; \( f_o \approx 0.82 \) and \( f_w \approx 0.13 \) for granite; and \( f_o \approx 0.88 \) and \( f_w \approx 0.15 \) for gabbro. Those values lead to estimates of \( f \) at 1 m/s, an average seismic slip rate, of \( f_{V=1 m/s} \approx 0.16 \) for quartzite and 0.23 for both granite and gabbro. However, those are extrapolations beyond the data; the most extreme weakening actually recorded in the experiments, i.e., at the maximum rate of 0.36 m/s, was \( f_{V=0.36 m/s} \approx 0.32 \) for quartzite, 0.40 for granite, and 0.37 for gabbro. To choose those values, and indeed to fit a curve to the experimental results, required averaging out
the rapid fluctuations as shown in Figure 2a, and those were more severe for some of the other rocks. Presumably they result because the friction is strongly velocity weakening, and instabilities result from the interaction of that type of friction with the elasticity of the loading apparatus.

While Figure 2a shows, for quartzite, $f \approx 0.30–0.35$ for $V = 0.30–0.36$ m/s, much higher speed experiments [Prakash, 2004; Prakash and Yuan, 2004, also private communication, 2004] on the same rock type (Figure 2b) show $f \approx 0.20$ or slightly less at rates $V = 2–4$ m/s. Those experiments, done by loading an annular specimen with a torsional wave from a pretwisted Kolsky bar, achieve rates higher than the average seismic slip rate, although the results are valid only for a very small slip, $\sim 0.2$ mm, prior to cracking of the wall of the annulus. The results cited from Tutsumi and Shimamoto, Tullis and Goldsby, Prakash and Yuan, and Hirose and Shimamoto suggest that despite high friction coefficients, $f \approx 0.6–0.9$ in slow slip, strong rate weakening of localized frictional sliding at seismic slip rates is a reality, probably due to flash heating, and that we might expect $f \approx 0.2–0.3$ at seismic slip rates, at least in the range of slip prior to development of a macroscopic melt layer along the friction surface (if enough slip takes place, at high enough effective stress, for such to occur). For that reason, the default value of $f$ during seismic slip in sections of the paper is taken as $f = 0.25$, and in some cases $f = 0.20$ is examined too. The weakened $f$ implies a slower heating rate, so that more slip is needed to achieve a given temperature change (the slip needed is shown in section 3 to scale as $1/f^2$, and so is 9 times larger for $f = 0.25$ than for $f = 0.75$).

3. Thermal Pressurization of Pore Fluid

This mechanism [Sibson, 1973; Lachenbruch, 1980; Mase and Smith, 1985, 1987] assumes that fluids (water, typically) are present within the fault gouge which shears and that the shear strength $\tau$ during seismic slip can still be represented by the classical effective stress law $\tau = f(\sigma_n - p)$, where $\sigma_n$ is normal stress and $p$ is pore pressure. Frictional heating then would cause the fluid, if it was unconstrained, rather than caged by the densely packed solid particles, to expand in volume much more than would the solid cage. Thus unless shear-induced dilatancy of the gouge cage overwhelms the thermal expansion effect, or unless the gouge is highly permeable, a pressure increase must be induced in the pore fluid during slip. Since $\sigma_n$ can typically be assumed to remain constant during slip, strength $\tau$ is reduced, ultimately toward zero, as shear heating continues to raise temperature so that $p$ approaches $\sigma_n$.

3.1. Conservation and Transport Equations

The shear of the material of a fault zone lying parallel to the plane $y = 0$ is considered, under constant normal stress $\sigma_n$, so that the $y$ coordinate axis is perpendicular to the fault zone. The shear rate is $\dot{\gamma} = \dot{\gamma}(y, t)$, and the spatial distribution of $\dot{\gamma}$ will shortly be considered either to be completely uniform, or else to take the form of slip that is wholly confined to the mathematical plane $y = 0$ within the fault zone, so as to most simply represent slip on the type of very thin shear zone (or principal slip surface) within the fault zone discussed in section 1. In that latter case, $\dot{\gamma}(y, t) = V(t)\delta_{y=0}(y)$, where $V(t)$ is the slip rate across the fault plane. The model of slip on a plane will be a sensible simplification for prediction of pore pressure when the diffusion distance for pore pressure change (which varies with time and slip during the event) is a few times greater than the actual shear zone thickness. Similarly, it should provide a sensible prediction of the maximum temperature rise when the diffusion distance for temperature changes is also a few times greater than shear zone thickness. Since hydraulic diffusivity is usually estimated to be greater than thermal diffusivity (see Tables 2 and 3 and sections 3.2 and 3.6 and Appendix A for the fuller basis for parameters there), the condition on pore pressure may be met even if there is no temperature failure. In such cases the model will be reasonable for prediction of pore pressure rise by thermal pressurization, and hence of fault weakening by effective stress reduction which is of primary interest here, whereas it will overestimate the maximum temperature rise in those cases. Rempel and Rice (submitted manuscript, 2006) investigate those issues more fully.

The shear stress $\tau = \tau(y, t)$, i.e., it is independent of $y$, by considerations of mechanical equilibrium. (It is true that $\partial\tau/\partial y = \rho a$, where $a$ is the fault-interparticle acceleration, but the effect of even large accelerations, say, $10–100$ g, is insignificant over the small length scales in the $y$ direction about the fault, i.e., a few millimeters to a couple tens of millimeters, over which the heat and fluid diffusion processes take place during the dynamic slip, and gigantic gradients in $p$ and $T$ occur. E.g., even with $a = 100$ g, $\partial\tau/\partial y \approx 3$ MPa/m, so $\tau$ changes by only 0.03 MPa over $10$ mm change in $y$. In comparison, the variation of shear strength $f(\sigma_n - p)$ with distance is, when there is full thermal pressurization $p \rightarrow \sigma_n$ at the fault core, of order $f(\sigma_n - p_{amb})$ over the few millimeters over which $p$ varies from $\sigma_n$ to $p_{amb}$, where $p_{amb}$ is the ambient pore pressure. Its gradient thus is typically of order 1000–5000 MPa/m at representative seismogenic depths, vastly larger than any effect of acceleration over the relevant few millimeter size scale. Thus assuming mechanical equilibrium $\partial\tau/\partial y = 0$, so that $\tau = \tau(t)$ over the relevant size scale, is certainly appropriate for present purposes.) Of course, $\tau$ and $\dot{\gamma}$ or $V$ will vary also with the $x$ and $z$ coordinates within the fault zone, but the length scales in $x$ and $z$ over which they vary will generally (except possibly at the propagating front of the rupture zone) be far larger than the scale in the $y$ direction affected by heat and fluid mass diffusion, so the problem is treated as involving the single, fault-perpendicular spatial coordinate $y$.

The energy equation, or first law of thermodynamics, and the fluid mass conservation equation are

$$-\frac{\partial q_h}{\partial y} + \dot{\gamma} = \rho c \frac{\partial T}{\partial t}$$

and

$$\frac{\partial m}{\partial t} + \frac{\partial q_f}{\partial y} = 0.$$  \(8\)

Here $\rho c$ is the specific heat per unit volume of the fault gouge in its reference state. Also, $m$ is the current mass of pore fluid per unit of bulk volume which the porous material occupied in that reference state [Rice and Cleary, 1976], and $m = \rho_m n$, where $\rho_f$ is the density of the fluid in the pore spaces, and $n$ is the current volume of pore space per unit reference state bulk volume. (Loosely, $n$ is the
Table 2. Assumed and Resulting Parameters of the Slip-on-Plane Model, to Represent a Mature Fault Surface at 7 Km Depth, at Normal Stress of 196 MPa, Ambient Pore Pressure of 70 MPa, and Ambient Temperature of 210°C

<table>
<thead>
<tr>
<th>Models Considered</th>
<th>Intact Elastic Walls</th>
<th>Highly Damaged Walls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ambient p and T&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Average on p-T path</td>
</tr>
<tr>
<td>Common parameters assumed for all models</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific heat of fault gouge [J], [VS], p&lt;sub&gt;c&lt;/sub&gt;, MPa/°C</td>
<td>2.7</td>
<td>2.7</td>
</tr>
<tr>
<td>Starting porosity [W], n</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Friction coefficient (flash heating, see text), f</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Slip rate [see text], f, m/s</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Normal stress, σ&lt;sub&gt;n&lt;/sub&gt;, MPa</td>
<td>196</td>
<td>196</td>
</tr>
<tr>
<td>Path ranges used for property evaluations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Porous fluid pressure range, p&lt;sub&gt;a&lt;/sub&gt; = p&lt;sub&gt;amb&lt;/sub&gt; = p&lt;sub&gt;high&lt;/sub&gt;, MPa</td>
<td>70, 70</td>
<td>70, 133</td>
</tr>
<tr>
<td>Effective stress range, σ&lt;sub&gt;e&lt;/sub&gt; = σ&lt;sub&gt;a&lt;/sub&gt; + σ&lt;sub&gt;n&lt;/sub&gt; = p&lt;sub&gt;high&lt;/sub&gt;, MPa</td>
<td>126, 126</td>
<td>126, 63</td>
</tr>
<tr>
<td>Temperature range, T&lt;sub&gt;amb&lt;/sub&gt;, T&lt;sub&gt;high&lt;/sub&gt;, °C</td>
<td>210, 210</td>
<td>210, 334</td>
</tr>
<tr>
<td>Material properties (averages over path ranges)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermal diffusivity [VS], α&lt;sub&gt;th&lt;/sub&gt;, mm/s&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.70</td>
<td>0.65</td>
</tr>
<tr>
<td>Fluid thermal expansivity [B], β&lt;sub&gt;f&lt;/sub&gt;, 10&lt;sup&gt;-3&lt;/sup&gt;/°C</td>
<td>1.08</td>
<td>1.21</td>
</tr>
<tr>
<td>Pore space thermal expansivity, β&lt;sub&gt;n&lt;/sub&gt;, 10&lt;sup&gt;-3&lt;/sup&gt;/°C</td>
<td>−0.19</td>
<td>−0.18</td>
</tr>
<tr>
<td>Pore fluid compressibility [B], β&lt;sub&gt;p&lt;/sub&gt;, 10&lt;sup&gt;-10&lt;/sup&gt;/Pa</td>
<td>0.64</td>
<td>0.74</td>
</tr>
<tr>
<td>Pore space pressure expansivity [W], β&lt;sub&gt;n&lt;/sub&gt;, 10&lt;sup&gt;-10&lt;/sup&gt;/Pa</td>
<td>0.65</td>
<td>0.77</td>
</tr>
<tr>
<td>Fluid viscosity [K], μ&lt;sub&gt;f&lt;/sub&gt;, 10&lt;sup&gt;−4&lt;/sup&gt; Pa s</td>
<td>1.48</td>
<td>1.26</td>
</tr>
<tr>
<td>Permeability [WS], k, 10&lt;sup&gt;−20&lt;/sup&gt; m&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.65</td>
<td>1.38</td>
</tr>
<tr>
<td>Resulting material properties</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undrained pressurization factor, Λ, (MPa/°C)</td>
<td>0.98</td>
<td>0.92</td>
</tr>
<tr>
<td>Hydraulic diffusivity, α&lt;sub&gt;hy&lt;/sub&gt;, (mm/s)&lt;sup&gt;2&lt;/sup&gt;</td>
<td>0.86</td>
<td>1.81</td>
</tr>
<tr>
<td>Resulting parameters of slip-on-plane model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weakening length parameter, L&lt;sub&gt;*&lt;/sub&gt;, mm</td>
<td>1.51</td>
<td>2.55</td>
</tr>
<tr>
<td>Maximum possible temperature rise, ΔT&lt;sub&gt;max&lt;/sub&gt;, °C</td>
<td>271</td>
<td>366</td>
</tr>
<tr>
<td>Maximum possible T, T&lt;sub&gt;max&lt;/sub&gt; (for p&lt;sub&gt;amb&lt;/sub&gt; = 210°C), °C</td>
<td>481</td>
<td>576</td>
</tr>
</tbody>
</table>

<sup>a</sup>“Intact elastic walls” models use laboratory-constrained data for permeability k and pore space pressure expansivity β<sub>n</sub> = β<sub>p</sub> of undisturbed gouge, and “highly damaged walls” models account approximately but arbitrarily for gouge damage at the rupture front, and during slip and thermal pressurization, by using a differently defined pore space pressure expansivity β<sub>n</sub> = β<sub>n</sub><sup>amb</sup> and increasing k by k<sub>n</sub> = 2β<sub>n</sub><sup>amb</sup> and the two β<sub>n</sub><sup>amb</sup> measures are assumed to vary only with σ<sub>a</sub> − p<sub>a</sub>. Fluid and other T or p<sub>a</sub> dependent properties are evaluated as averages along straight line paths from p<sub>a</sub> = p<sub>amb</sub> = p<sub>high</sub>, T<sub>high</sub>. Codes are B, Burnham et al. [1969]; K, Keenan et al. [1978]; L, Lachenbruch [1980]; T, Todheide [1972]; VS, Voseen and Schellenschmidt [2003]; W, Wibberley [2002] and C. A. Wibberley (private communication, 2003); WS, Wibberley and Shimamoto [2003]. Fluid properties estimated from B, K, and T with the collaboration of A. Rempel.

<sup>b</sup>The “high” values are set to ambient values for the columns. Results are used to approximately estimate the respective p<sub>high</sub>, T<sub>high</sub> for the “average on p-T path” columns, as rough spatial averages, over the part of the wall that is actively participating in the heat and mass transfer, at the stage when p at the fault has been elevated to σ<sub>n</sub>.

“Porosity” of the fault gouge, although that term should strictly be reserved when pore volume is divided by current, rather than reference state, bulk volume.) Further, γ is here treated as a Lagrangian, or material, coordinate, which is measured off in that reference state, and γ = dγ/dτ where γ is the local fault-parallel velocity. Work by the normal stress σ<sub>n</sub> is negligible in comparison to that by τ in the large shear considered.

The energy flux q<sub>f</sub> and fluid mass flux q<sub>f</sub> in the above balance equations are given by

\[ q_f = -K \frac{dT}{dy} \]  
\[ q_f = -\rho_f \alpha_f \frac{dp}{dy} \]  (9)

The latter is the Darcy law where k is permeability and \( \alpha_f \) is fluid viscosity (\( \alpha_f \sim 10^{-4} \) Pa s for water in the elevated temperature range of interest). In the former, all nonnegligible energy flux q<sub>f</sub> is assumed to be in the form of heat conduction, where \( K \) is the thermal conductivity. That is, advective energy transport by the moving hot fluid has been neglected, justifiably, because the pore volume fraction n is low (e.g., of order 5% for Median Tectonic Line, Japan, ultracataclasite at effective confining stress of order 100 MPa (C. Wibberley, personal communication, 2003), values that are consistent with results of Sulem et al. [2004], for samples from borehole penetration the Aegion fault in the Gulf of Corinth), and the ultracataclasite permeability is exceptionally low, of order \( 10^{-20} \)–\( 10^{-19} \) m<sup>2</sup> [Lockner et al., 2000; Wibberley and Shimamoto, 2003; Sulem et al., 2004] (Tables 1 and 2). See Mase and Smith [1987] and Lee and Delaney [1987] for a fuller discussion of the advective terms; the latter suggest that they should be negligible if \( k < 10^{-16} \) m<sup>2</sup>.

3.2. Porothermoelastic Parameters, Governing Equations for p and T, and Types of Solutions Allowed

[41] With standard procedures (see Appendix A), including the Segall and Rice [1995, 2004] treatment of inelastic dilatancy, increments dm in fluid mass content can be written in terms of increments dp in pore pressure and dT in temperature, and of the inelastic (or “plastic”) porosity increments d\( \rho <sup>el</sup> \), as

\[ dm/\rho_f = \beta(dp - \Delta T) + d\rho <sup>el</sup> \]  (10)

where it is understood that \( \rho_f \) and the fault-parallel extensional strain components are held constant during the process considered. The new parameters \( \beta \) and \( \Lambda \), defined by the way in which they appear in that equation, are expressed in terms of compressibility and thermal expansion properties of the fluid and of the pore space later in this section, following the derivation in Appendix A, and values are given in Tables 2 and 3.
Table 3. Study of Effects of Damage of Fault Gouge Due to Concentrated Stressing Near the Rupture Front and Due to Inelastic Response to Thermal Stressing Along the Fault Walls

<table>
<thead>
<tr>
<th>βp, 10⁻⁵ Pa⁻¹</th>
<th>1.0βp</th>
<th>1.5βp</th>
<th>2.0βp</th>
</tr>
</thead>
<tbody>
<tr>
<td>k[^dmgw^]</td>
<td>1.04</td>
<td>1.77</td>
<td>2.49</td>
</tr>
<tr>
<td>ΔL, MPa°C</td>
<td>0.64</td>
<td>0.44</td>
<td>0.34</td>
</tr>
</tbody>
</table>

[^dmgw^] = k = 0.65 × 10⁻²⁰ m²

| αp, mm²/s    | 0.65  | 0.46  | 0.35  |
| L[^p^], mm   | 8.17  | 13.2  | 19.0  |
| ΔT[^max^], °C| 517   | 635   |       |
| Path L[^p^], mm | ~4.6 | ~6.6 | ~8.4 |

[^dmgw^] = 5k = 3.25 × 10⁻²⁰ m²

| αp, mm²/s    | 2.28  | 1.76  |       |
| L[^p^], mm   | 13.5  | 21.2  | 29.8  |
| ΔT[^max^], °C| 809   | 1200  |       |
| Path L[^p^], mm | ~25  | ~36   | ~49 |

[^dmgw^] = 10k = 6.50 × 10⁻²⁰ m²

| αp, mm²/s    | 7.56  | 4.57  | 3.52  |
| L[^p^], mm   | 13.5  | 21.2  | 29.8  |
| ΔT[^max^], °C| 809   | 1200  |       |
| Path L[^p^], mm | ~25  | ~36   | ~49 |

[^dmgw^] = 10k = 6.50 × 10⁻²⁰ m²

As explained in Appendix A, we must consider the possibility of other than elastic response of the material bordering the slip zone. Inelastic response is there relevant for the following reasons: The material will have been damaged and deformed inelastically by large concentrated stresses, e.g., locally in excess of a Mohr-Coulomb yield criterion, as the rupture front passed by it [Andrews, 1976, 2005; Poliakov et al., 2002; Rice, 1980; Rice et al., 2005]. Also, the process of shearing in granular materials (section 1.1), and sliding on fault surfaces that are not perfectly flat [Power and Tullis, 1991], necessarily forces some accommodation deformations in the fault walls that would not be likely to elicit only elastic response. Finally, the assumption of elastic response of the fault walls, even next to a perfectly smooth and planar fault, leads ultimately to large fault-parallel compressional effective stresses near the sliding zone, and hence to large effective stress differences which cannot be sustained elastically as p → σ[^g^] (Appendix A). Thus Tables 1 and 2 gives values of parameters corresponding not only to elastic response of intact wall material, but also to nonelastic response due to highly damaged walls (but with damage simply represented in terms of increased drained compressibility and permeability of the wall material, and relaxation of fault-parallel stresses to the fault-normal value). Table 3 further explores a range of assumptions for representing effects of damage. The present inability to precisely characterize fault wall response leads to a broad range of possible model parameters here. (The induction of large fault-parallel compressive effective stresses very near the sliding zone suggests that at least locally, all principal stresses will be compressive and at least as large as p, except possibly in the zone affected by stress concentration at the rupture front [Rice et al., 2005], so that hydraulic cracking by the pressurized fluid is not generally expected. In any event, the fluid realistically available to feed such a fracture is limited because gouge porosity and permeability are quite small.)

[43] There are few guidelines on how to represent the magnitude of inelastic dilatancy. A possibly important concept is that the gouge which deforms during an earthquake on a mature fault has been subjected to previous similar deformations, and thus it might be regarded as if it is in a “normally consolidated” state just before the presently considered episode. That means that it is then subjected to the largest effective stress experienced since its structure was last disturbed and unloaded (by thermal pressurization) to a low effective stress during the prior event. After that prior event it was then reloaded monotonically to its effective stress at the start of the presently considered event, as the excess pore pressure diffused away. Because cases of large dilatancy are generally associated with a densely packed state achieved through overconsolidation, that would suggest that only modest inelastic dilatation might take place during each shear episode on a mature fault. However, the real situation is complicated by the unknown effects of healing and cementation processes over the time interval between earthquakes.

[44] Note that dm[^p^]/dT in equation (10) corresponds to the volume of fluid, per unit reference state bulk volume, that would have to be absorbed (or expelled if dm is negative) in order for there to be changes dp, ΔT, and Δ[^f^] (again, at fixed σ[^n^] and fault-parallel extensional strains). Thus β is the volumetric pore fluid storage coefficient (i.e., dm[^p^]/dT for elastic storage at fixed T, with “elastic” implying dm[^f^]=0). Alternatively, if an increment of deformation is undrained (i.e., dm = 0), then the expression tells us how p varies with the other parameters, namely, dp[^undrained^] = Δ[^f^] - Δ[^f^]/β[^e^]. That shows that ΔT (called ΔT by Mase and Smith [1987]) is the thermoelastic value of dp[^f^]/dT under undrained conditions. Appendix A shows that β[^e^] = n(β[^e^] - β[^e^]), where β[^e^] is the isothermal compressibility of the pore fluid (dp[^f^]/dT = β[^e^]dp) and β[^n^] (equivalent to β[^e^] by Segall and Rice [1995]) is the isothermal pressure expansivity of the pore space (dn[^n^]/n = β[^n^]dp), assuming no inelastic dilatant contribution. Also, Δ is given as Δ = (Δ[^e^] - Δ[^e^])/(β[^e^] + β[^e^]), where the λ are isobaric, volumetric thermal expansion coefficients, defined such that dp[^f^]/dT[^f^] = -λ[^f^]dT[^f^] and dn[^n^]/n = λ[^n^]dT[^f^], again assuming no dilatant contribution.

[45] Using the above results, the energy and fluid mass conservation equations are

\[
\frac{\partial T}{\partial t} - \frac{\tau_T}{\rho c_v} = \frac{1}{\rho c_v \alpha_p} \frac{\partial T}{\partial y} \\
\frac{\partial p}{\partial t} - \Lambda \frac{\partial T}{\partial y} + \frac{1}{\gamma} \frac{\partial \theta}{\partial y} \left( \frac{\rho_c}{\beta_p^2} \frac{\partial \phi}{\partial y} \right)
\]

(11)

where α[^g^] = K/(ρc) is the thermal diffusivity and α[^f^] = k(ρc) is the hydraulic diffusivity. See Tables 2 and 3 for values under different assumptions.
Finally, it would generally be assumed that the shear stress transmitted across the fault zone is consistent with the effective stress law \( \tau = f(\sigma_n - p) \) with, in the simplest modeling, a constant friction coefficient \( f \). In fact, that implies remarkable limits to the types of solutions allowed by the above differential equations. Then the value of \( p(y, t) \) in any one place \( y \) which undergoes shear must be the same as in any other place \( y' \) that also shears, because \( f, \sigma_n \) and \( \tau \) have the same value at both places. That is discussed further in Appendix B. When \( \partial n^p / \partial t = 0 \) and \( f \) is constant, it implies that the types of solutions that can exist are such that either one or the other of the following is the case:

1. The \( p(y, t) = p(t) \) (i.e., \( p \) is spatially uniform) throughout the zone that shears, which implies, through equations (11), that \( T(y, t) \) has the form of a function of \( y \) plus a function of \( t \). If shear commences at all values of \( y \) at the same time, and at the same initial temperature, that reduces to \( T(y, t) = T(t) \), and then \( \gamma(y, t) = \gamma(t) \) too. Such describes a solution with no heat or mass transport and with spatially homogeneous straining. It is a state of homogeneous, adiabatic, undrained deformation, an obviously very idealized case discussed in section 3.3.

2. Shear occurs at a single value of \( y \) at which \( p(y, t) \) has its global maximum. That is, the thermal pressurization process naturally leads to shear localization. The situation corresponds to the model of slip on a plane, as considered by Mase and Smith [1987] and Lee and Delaney [1987]. It is unrealistic in that, as discussed in section 1.1, slip on a mathematical plane is not strictly possible in a disordered granular system, but it is to be interpreted physically as slip over such a small thickness that the disordered particulate state controls the size of that thickness (at scales of order of a few tenths of a millimeter, it is suggested in section 1). To resolve that thickness further, as is not attempted here, it would be necessary to formulate a refined model of granular friction which had built into it some mathematical structure to limit localization [e.g., Muhlhaus and Vardoulakis, 1987; Tordesillas et al., 2004]. Exceptionally, a finite set of 2 or more \( y \) values might share the hosting of the global maximum of \( p(y, t) \), in which case, there would be slip on a finite set of planes, but each would have to undergo exactly the same pore pressure history and that seems too special to address further here.

While it would be the case that strong rate strengthening of \( f \) with increasing shear rate \( \dot{\gamma} \) would allow delocalization of shear from a plane, any feature leading to rate weakening of \( f \), including the flash heating mechanism discussed earlier, would tend to promote localization to a plane. Rate weakening and severe localization are well correlated experimentally [Beeler et al., 1996; Scruggs and Tullis, 1998], at least in low strain rate experiments. Thus the two weakening mechanisms given primary focus here may be intimately connected with why slip can be so highly localized on mature faults.

### 3.3. Adiabatic, Undrained Deformation

This case 1. It corresponds to homogeneous simple shear strain \( \gamma \) at constant \( \sigma_n \), on a spatial scale of the sheared layer that is broad enough to effectively preclude heat or fluid transfer. It seems unlikely to be achieved even approximately in practice, except for gouge showing an extremely strong positive increase of \( f \) with shear rate, because, as shown in a companion paper (J. R. Rice and J. W. Rudnicki, Stability of spatially uniform, adiabatic, undrained shear of a fault zone, manuscript in preparation, 2006), such deformation is unstable to severe flow localization, after which the heat and mass transfer effects must become important. Neglect of transport terms in equations (11) gives

\[
dp - \Lambda dT + d\rho \dot{\gamma} / \beta = 0 \quad \text{and} \quad \tau d\gamma = \rho c dT,
\]

and describes a special case considered by Lachenbruch [1980]. The equations can be rearranged, and integrated, starting at ambient conditions \( p_{\text{amb}} \) and \( T_{\text{amb}} \), and assuming for simplicity that \( \rho, \Lambda, \beta \) can be considered as constant (or using averaged values for the integration interval), to obtain

\[
\int_0^t \tau d\gamma = \rho c (T - T_{\text{amb}}) = (\rho c / \Lambda) (p - p_{\text{amb}} + \Delta \rho \dot{\gamma} / \beta),
\]

Remarkably, this result is independent of any assumption about how \( \tau \) varies with \( \sigma_n - p \), \( \gamma \), \( d\gamma / dt \), or their history. Making only the weak constitutive assumption that the shear strength reduces to 0 when \( p \) is raised to the level \( \sigma_n \), there results expressions for the maximum possible work which can be absorbed per unit volume, and the maximum possible temperature rise, in sufficiently large adiabatic, undrained shear. These are

\[
\int_0^{T_{\text{max}} - T_{\text{amb}}} \tau d\gamma = \frac{\rho c}{\Lambda} (\sigma_n - p), \quad \text{and} \quad T_{\text{max}} - T_{\text{amb}} = \frac{\sigma_n - p}{\Lambda},
\]

where \( p_o = p_{\text{amb}} - \Delta \rho \dot{\gamma} / \beta \).

The term \( p_o \) here has a simple interpretation; if the net dilatancy during the shear is \( \Delta \rho \dot{\gamma} \) (physically, we expect nearly all of that to accumulate in the earliest stages of deformation), then \( p_o \) just involves an effective resetting of the initial pore pressure from ambient, by reducing it by the suction \( \Delta \rho \dot{\gamma} / \beta \) [Segall and Rice, 1995].

A specific form for the weakening law when \( \tau = f(\sigma_n - p) \) with constant \( f \) results by eliminating \( T \) and \( p \) as explicit variables in equation (12) to get the Lachenbruch [1980] equation

\[
\frac{d\tau}{d\gamma} + \frac{\Lambda}{\rho c} \tau = \frac{1}{\beta} \frac{d\rho \dot{\gamma}}{d\gamma}.
\]

Assuming as an approximation that all inelastic dilatancy occurs in the earliest phases of shear, the right side is effectively a Dirac delta function at strain \( \gamma = 0 \) and the solution is

\[
\tau = f(\sigma_n - p) = f(\sigma_n - p_o) \exp \left( -\frac{f \Lambda}{\rho c} \gamma \right)
\]

\[
= f(\sigma_n - p_o) \exp \left( -\frac{f \Lambda \delta}{\rho c} \right),
\]

where the latter form applies to slip \( \delta \) over a layer thickness \( h \) that is large enough to justify the neglect of heat and fluid transport (with the caveats that homogeneous deformation of that type appears to be unstable against localization and that there are inconsistencies, discussed in Appendix B, related to the feature that \( p \) will not be spatially uniform.
within any layer of finite thickness with realistic boundary conditions. The associated temperature rise is

$$T - T_{amb} = \frac{p - p_o}{\Lambda} = \left( 1 - \frac{\tau}{f(s_n - p_o)} \right) \frac{s_n - p_o}{\Lambda}.$$  \hfill (17)

Also, the fracture energy associated with such shear in a layer of thickness $h$ is

$$G = G(\delta) = \int_0^\delta \left[ (\delta(\delta') - \delta(\delta)) \right] \, d\delta'.$$

$$= \frac{pc(s_n - p_o)h}{\Lambda} \left[ 1 - \left( 1 + \frac{f \Delta \delta}{p \Delta h} \right) \exp \left( -\frac{f \Delta \delta}{p \Delta h} \right) \right].$$  \hfill (18)

which approaches $G = (pc/\Lambda)(s_n - p_o)h$ at large slip; that value is independent of $f$, as anticipated from equation (14).

[52] The characteristic shear strain for the adiabatic, undrained weakening is $pc/\Lambda$. It may be estimated from values for $pc$ and $\Lambda$ in Table 2. Because the entire gouge domain is assumed to be deforming in shear, the “highly damaged” model seems appropriate for it, at least for the ambient conditions at 7 km depth assumed for Table 2. Then $pc/\Lambda \approx 7.3/f$. For slip on a plane at seismic rates, a weakened friction $f = 0.25$ seems justified to represent flash heating and is used subsequently here. However, for broadly distributed shear as in this section, the contacts would not likely shear fast enough to be susceptible to flash heating and so $f \approx 0.6$ may be more appropriate, leading to the characteristic weakening strain $pc/\Lambda \approx 12$. The characteristic slip distance in the exponential slip-weakening law is therefore estimated as $12h$. The associated value of $D$, as defined in the linear slip-weakening relation assumed by many seismologists, with property that $G = f(s_n - p_o)Dc/2$ at large slip, is therefore estimated to be $D_c = \approx 2(pc/\Lambda)h \approx 24h$, i.e., $D_c \approx 0.24$ m for $h = 10$ m.

[53] Rice [2003] and Rice et al. [2005] discussed interpretations of seismically inferred earthquake fracture energies in terms of such estimates of $G$ and $D_c$. However, the new perspectives reached in the present work, based on the extreme localization of shear, leads to the conclusion that the model of slip on a plane discussed next (including modifications to it to deal with effects of the actual small but finite thickness of the shear zone) provides a preferred physical basis for predicting slip weakening and fracture energy associated with thermal pressurization.

3.4. Model of Slip on a Plane, General Results

[54] In the case of interest here, all slipping is on the plane $y = 0$ and $\tau = f(s_n - p_v)(y - a)$ where $p_v = 0$ is the pore pressure on that sliding plane. Thus the equations which must be satisfied are as follows:

$$\text{In } |y| > 0, \quad \frac{\partial T}{\partial t} = \frac{1}{\rho c_p} \frac{\partial}{\partial y} \left( \rho c_p \frac{\partial T}{\partial y} \right)$$

and

$$\frac{\partial p}{\partial t} + \frac{\partial p}{\partial t} = \frac{1}{\beta} \frac{\partial}{\partial y} \left( \beta \frac{\partial p}{\partial y} \right).$$  \hfill (19a)

On $y = 0^+$,

$$- \rho c_p \frac{\partial T}{\partial y} = \pm \frac{1}{2} \tau V = \pm \frac{1}{2} \left( s_n - p_v \right) V$$

and

$$\frac{\partial p}{\partial y} = 0.$$  \hfill (19b)

[55] For simplicity, and in keeping with the spirit of making elementary estimates, in the sequel the transport terms on the right sides of the energy and fluid mass equations are linearized to $\alpha_{th} \partial^2 T/\partial y^2$ and $\partial \rho \partial^2 p/\partial y^2$, respectively, choosing $\alpha_{th}$ and $\alpha_v$ as representative values for the range of $p$ and $T$ experienced during a typical seismic event that is modeled (Tables 2 and 3).

[56] It is assumed here for simplicity that essentially all inelastic dilatancy $\Delta s^{\text{pl}}$ in the fault zone takes place in the concentrated stress field associated with passage of the rupture tip, and in the earliest increments of localized slip, with no further dilatancy as significant slip develops. Thus the inelastic dilatancy rate can be written approximately for our purposes as $\partial s^{\text{pl}}/\partial t = \Delta s^{\text{pl}} / \text{Dirac}(t)$. The effect of this approximation is to make the above set of partial differential equations become

$$\text{In } |y| > 0, \quad \frac{\partial T}{\partial t} + \frac{\partial T}{\partial t} = \frac{1}{\rho c_p} \frac{\partial}{\partial y} \left( \rho c_p \frac{\partial T}{\partial y} \right)$$

and

$$\frac{\partial p}{\partial t} = \frac{\partial p}{\partial t} + \frac{\partial p}{\partial t} = \frac{1}{\beta} \frac{\partial}{\partial y} \left( \beta \frac{\partial p}{\partial y} \right).$$  \hfill (19b)

where $p_{amb}$ is the ambient pore pressure, and $T_{amb}$ is the ambient temperature, just before rupture. The same conditions apply on $y = 0$ as in equations (19b).

[57] The PDE set can be solved in general to show (Appendix B) that at the fault plane, $y = 0$, there is a universal proportionality between the temperature rise from its initial value and the pore pressure rise from its suction-reduced value just after onset of failure. That relation is

$$\left( T - T_{amb} \right)_{y=0} = \left( 1 + \sqrt{\frac{\alpha_{th}}{\alpha_v}} \right) \frac{p - p_v}{\Lambda}.$$  \hfill (21)

and it is universal in that it applies no matter how slip rate $V$ varies with $t$, or how $\tau$ or $f$ varies with slip rate and slip history. In fact, the derivation does not even require the assumption that the shear strength law be given in the form $\tau = f(s_n - p_v)$. However, if the strength law at least has the property that $\tau = 0$ (or is negligible) when $p_v = a \to s_n$, then the maximum possible temperature which can be achieved on the fault plane is

$$\Delta T_{\text{max}} = T_{\text{max}} - T_{amb} = \left( 1 + \sqrt{\frac{\alpha_{th}}{\alpha_v}} \right) \frac{s_n - p_v}{\Lambda}$$  \hfill (22)

and results for it are given in Tables 2 and 3. It is a factor $1 + \sqrt{\alpha_{th}/\alpha_v}$ times the maximum possible rise for the adiabatic undrained model of section 3.3. That $T_{\text{max}}$ can be approached only at sufficiently large slip such that $p$ there is driven to $s_n$ (at which point the heating rate vanishes and there can be no further temperature rise). In cases for which it exceeds the melting temperature, say, $T_{\text{max}} > 1000^\circ$, the model must be abandoned at some smaller amount of slip, corresponding to a pore pressure rise to a value $p < s_n$, which just brings $T$ on the slip plane to the melting temperature, according to the above expression for $T - T_{amb}$.

[58] Mase and Smith [1987] derived equation (22) through an ingenious analysis of an asymptotic limit, for
which a fault had been sliding for a long time at $p \approx \sigma_n$, and for which the temperature rise was long saturated at a value $\Delta T_{\text{max}}$ to be determined. They then made what they seem to have thought to be an approximation, that equation (21) might roughly apply during the entire process, to derive equation (23) to follow as a corresponding approximation, albeit one that they had already shown to be an exact result when there was no fluid transport so that $\alpha_{h'} \alpha_{oh} = 0$ (see Appendix B). The present work shows that in fact, equations (21) and (23) are exact results.

The utility of these expressions for temperature rise is marred only by the fact that the model of slip on a plane tends to overestimate the temperature of even a very thin but finite thickness shear zone. See Appendix B for discussion and some approximate corrections for that. Mase and Smith [1987] and Rempel and Rice (submitted manuscript, 2006) provide a fuller analysis.

### 3.5. Model of Slip on a Plane, Solution for Constant Friction and Slip Rate

A linear integral equation for $\tau(t)$, given general histories of slip rate $V(t)$ and friction coefficient $f(t)$, is derived in Appendix B equation (B15) based on equations (B11a) and (B11b). A generalization of the integral equation is also given there for a Gaussian shear rate distribution [Andrews, 2002] over a narrow zone.) The equation thus defines the constitutive laws for the fault surface. It can be solved numerically in general, but in a simple closed form for the case of constant $V$ and $f$ to obtain, when the solution is written in terms of slip $\delta(=Vt)$,

$$\tau = f(\sigma_n - p) = f(\sigma_n - p_o) \exp \left( \frac{\delta}{L^*} \right) \text{erfc} \left( \frac{\delta}{L^*} \right),$$  \hspace{1cm} (23a)

where

$$L^* = \frac{4}{\sqrt{\pi}} \frac{\rho c}{\Lambda} \left( \sqrt{\frac{\alpha_{h'}}{\alpha_{oh}}} + \sqrt{\frac{\alpha_{h}}{\alpha_{oh}}} \right)^2 \frac{1}{V}.$$  \hspace{1cm} (23b)

This parameter $L^*$ (related to $\psi_o$, of Mase and Smith [1987] by $L^* = V \psi_o$) has the unit of length, descended from diffusivity divided by slip rate, and nicely wraps into a single parameter the dependence of scaled friction strength $\tau[f(\sigma_n - p_o)]$ on all of the porothermoelastic properties of the solid and fluid, and on the slip rate and friction.

Although the equation contains a single length parameter, a plot as in Figure 3 shows a remarkable multiscale feature, in that continued weakening takes place at an ever decreasing rate over a very broad range of slip $\delta$, which is scaled by $L^*$ in Figure 3. A modest rearrangement of equation (23a) lets us solve for pore pressure rise at the fault plane and, using equation (21), the fault plane temperature as a function of $\delta$ is then

$$T - T_{\text{amb}} = \left( 1 + \sqrt{\frac{\alpha_{h'}}{\alpha_{oh}}} \frac{\sigma_n - p_o}{\Lambda} \right) \left( 1 - \exp \left( \frac{\delta}{L^*} \right) \text{erfc} \left( \frac{\delta}{L^*} \right) \right),$$  \hspace{1cm} (24)

where the bracketed term starts at 0 at $\delta = 0$, and slowing approaches 1 as $\delta$ increases to many times $L^*$, so that its prefactor is $\Delta T_{\text{max}}$.

The prefactor can also be expressed in terms of $L^*$, writing

$$\Delta T_{\text{max}} = \left( 1 + \sqrt{\frac{\alpha_{h'}}{\alpha_{oh}}} \frac{\sigma_n - p_o}{\Lambda} - \frac{f}{\rho c} \frac{V L^* \sigma_n - p_o}{\alpha_{oh}} \right),$$  \hspace{1cm} (25)

where the dependence on $f$ and $V$ cancels because $L^* \propto 1/V^2$. This further emphasizes that $L^*$ encapsulates all of the thermoelastic and poroelastic parameters that enter the model.

### 3.6. Model of Slip on a Plane: How Large is $L^*$?

The estimates, given in Tables 2 and 3, are based on material properties (Table 1 for the gouge, and references cited in Table 2 for water) evaluated so as to represent 7 km depth, which is chosen because it is a representative centroidal depth of the region which slips in large crustal earthquakes. The ambient conditions there are taken as $T_{\text{amb}} = 210^\circ\text{C}$ and $p_{\text{amb}} = 70 \text{ MPa}$, and the fault-normal stress $\sigma_n$ is identified as the overburden, 196 MPa (so that $\sigma_n - p_{\text{amb}} = 126$ MPa). Dilatancy is neglected ($p_o = p_{\text{amb}}$) in the standard model and for the estimates in Tables 2 and 3, although later an illustration will be given for a large, 50 MPa dilatant suction, so that $p_o = 20 \text{ MPa}$. $L^*$ depends on $f$ and $V$, and in the standard model, $f = 0.25$ based on the earlier discussion of flash heating and $V = 1 \text{ m/s}$ as explained. Because $L^* \propto 1/V^2$, the weakening of $f$ by flash heating is very important to the numerical size of $L^*$; use of a slow slip rate $f$ of order 0.9–0.6 would decrease $L^*$ to only...
8–17% of the values shown. Use of $f = 0.20$, also plausible, increases $L^*$ by 56%, and use of $f = 0.15$ by 180%.

[64] In Table 2, approximate end-members are considered, as follows: (1) intact elastic fault walls and (2) highly damaged walls (see Appendix A). “Ambient $p$ and $T$ columns” evaluate all relevant porothermomechanical properties based on the assumed $P_{amb}$, $T_{amb}$ and $\sigma_n - P_{amb}$ and then use those to obtain the resulting $\Delta \alpha$ and $\Delta \alpha_{th}$, and thus $L^*$, $\Delta T_{max}$ and $T_{max}$. However, the range of $p$ and $T$ thus predicted is large, and the porothermomechanical properties on which the results depend vary with $p$ and $T$. To retain the simplicity of the linear model, with single relevant parameter $L^*$, the properties are recalculated as averages along the $p$-$T$ paths implied by the ambient $p$ and $T$ results. This is done by identifying a representative $p$ and $T$ at each stage of the process, which $p$ and $T$ are chosen recognizing that each vary across the few millimeters distance where the process is active, and range from $p(0, t)$, $T(0, t)$ at the slip plane to $P_{amb}$, $T_{amb}$ at the outer limit of the affected zone. Thus, at time $t$ in the process, properties are evaluated using representative $p = p(0, t) + P_{amb}/2$ and $T = T(0, t) + T_{amb}/2$. Using those for $p$ and $T$, we then march along the $p$-$T$ paths corresponding to the ambient $p$ and $T$ results, respectively, which paths then run from $P_{amb}$, $T_{amb}$ to $p_{high}$, $T_{high}$ at full pressurization, where $p_{high} = (\sigma_n + P_{amb})/2$ and $T_{high} = (T_{max} + T_{amb})/2$, and calculate averages of all material properties along those paths (that done with the collaboration of A. Rempel). Those path-averaged properties are used as the basis of the “average on $p$-$T$ path” refined estimates, respectively. It is seen that $L^*$ increases (by factors of 78% and 66%, respectively, principally because the predicted $\alpha_{th}$ roughly doubles, but also because $\Delta \alpha$ diminishes about 10%), and corresponding to the increases in $L^*$ and modest decreases in $\alpha_{th}$, the predicted temperature rise $T_{max}$ in very large slip increases by 10–25%. Again, in average on $p$-$T$ path highly damaged walls cases, for which the temperature rise, even after correcting for the small but finite thickness of the shear zone, is in excess of melting, the model predicts weakening continues over a broad range of slip at an ever decreasing rate. Abercrombie and Rice [2005] showed that seismic evidence for how fracture energy scales with slip, if attributed to a single slip weakening law, the same for all earthquakes in their data set, implied qualitatively similar features to what emerges here, in the sense that weakening continues over a broad range of slip at an ever decreasing rate. That is an alternative to suggesting [Ohnaka, 2003] that $D_c$ scales with slip in the event (or, restated, that faults which host large slip events have large $D_c$).

3.7. Model of Slip on a Plane, Fracture Energy and Net Heating

[68] With the aim of testing the theoretical predictions, based on geological characterizations of mature fault zones and laboratory studies of poromechanical and frictional properties, against seismological constraints, the predicted law $\tau = \tau(\delta)$ is used to calculate the fracture energy $G$:

$$G = G(\delta) = \int_{0}^{\delta} [\tau(\delta') - \tau(\delta)]d\delta' = f(\sigma_n - p_o)L^* \cdot \left[\exp\left(\frac{\delta}{L^*}\right) \text{erfc}\left(\sqrt{\frac{\delta}{L^*}}\right) - 1 + 2\sqrt{\frac{\delta}{\pi L^*}}\right].$$  \hspace{1cm} (26)

A convenient three-term asymptotic expansion of that, accurate to three significant figures when $\delta > 20L^*$, is

$$G(\delta) \approx f(\sigma_n - p_o)L^* \left[\frac{\delta}{\pi L^*} - 1 + 3 \sqrt{N_{th}} + \ldots\right];$$  \hspace{1cm} (27)

it is useful also because some software packages cannot accurately evaluate the $\exp(D)\text{erfc}(D)$ product when $D \gg L^*$. Figure 4 shows the predicted $G$ versus slip $\delta$ for a wide range of $L^*$ values, ranging from 1 to 120 mm. The two solid curves of the set correspond to the low-end (blue curve, $L^* = 4$ mm) and high-end (red curve, $L^* = 30$ mm) models.

[69] The leading term in the asymptotic expression for $G$, when $\delta \ll V_T \gg L^*$ is

$$G \approx f(\sigma_n - p_o) \sqrt{\frac{L^*\delta}{\pi}} = 2(\sigma_n - p_o)\left(\frac{\Delta \alpha}{\pi}\right)\left(\sqrt{\frac{\Delta T_{th}}{\pi}} + \frac{\Delta \alpha_{th}}{\pi}\right).$$  \hspace{1cm} (28)
where here \( t \) is to be interpreted as the duration of slip at a point on the fault. As the last form makes clear, this result for \( G \) at large slip is, for a given duration \( t \), actually independent of the friction coefficient \( f \) and slip rate \( V \) that are assumed (although the derivation has, nevertheless, assumed that both are constant in time during slip). That independence of \( f \) at large \( \delta L^* \) can be seen in Figure 5. If this \( G \) expression, which applies only for very large slips (when the curves cluster in Figure 5), is written as \( G = f(\sigma_n - p_0)D_c/2 \), as for the frequently assumed linear slip-weakening law, then for such large slip events the \( D_c \) would correspond to

\[
D_c = \frac{4\rho c}{\sqrt{\pi A}} \left( \sqrt{\alpha_{hy}} t + \sqrt{\alpha_{th}} t \right) . \tag{29}
\]

Thus, using values somewhat comparable to highly damaged walls, ambient \( p \) and \( T \), in Table 2, e.g., \( \rho c = 2.7 \text{ MPa}/\text{s}^3, \Lambda = 0.37 \text{ MPa}/\text{K}, \alpha_{hy} = 4.0 \text{ mm}/\text{s}, \) and \( \alpha_{th} = 0.64 \text{ mm}/\text{s} \), this would give \( D_c = 0.18 \text{ m} \) for an event with 1 s duration, i.e., 1 m slip at 1 m/s. The result scales as \( \sqrt{\delta}/1 \text{ m} \) in this large slip range. However, for events that involve slips that are comparable to or just a few times \( L^* \), the scaling factor is not a power law; it can be seen in Figures 4 to 6 to be not so far from linear in slip on the logarithmic plots over a substantial slip range, although the curves have a significant curvature.

Figure 6 illustrates the effect of a significant initial suction, 50 MPa (i.e., reducing \( p_0 \) to 20 MPa), developed in the pore fluid as a representation of strong inelastic dilatancy, a possibility ignored in the earlier plots. It simply resets the initial effective stress, at the 7 km depth considered, from \( \sigma_n - p_0 = 126 \text{ MPa} \) to 176 MPa, so that the stress at any slip increases by a factor 176/126 \( \approx 1.4 \). On the basis of the ambient \( \beta_f \) and range of possible \( \beta_n \) values in Tables 2 and 3, to induce that suction would require that \( \Delta u^{pl}/n \approx 5-15\% \).
Table 4. Fracture Energies G of Large Earthquakes, Arranged in Terms of Average Slip δ

<table>
<thead>
<tr>
<th>Event</th>
<th>M&lt;sub&gt;m&lt;/sub&gt;</th>
<th>10&lt;sup&gt;18&lt;/sup&gt; N m</th>
<th>l, km</th>
<th>w, km</th>
<th>δ, m</th>
<th>G, MJ/m&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Ref&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Michoacan 1985</td>
<td>(M = 8.1)</td>
<td>1.500</td>
<td>150</td>
<td>120</td>
<td>2.8</td>
<td>6.63</td>
<td>1</td>
</tr>
<tr>
<td>Landers 1992</td>
<td>(M = 7.1)</td>
<td>56</td>
<td>60</td>
<td>14</td>
<td>2.2</td>
<td>5.0</td>
<td>3</td>
</tr>
<tr>
<td>Landers 1992</td>
<td>(M = 7.3)</td>
<td>97</td>
<td>79</td>
<td>15</td>
<td>2.2</td>
<td>17.4</td>
<td>2a</td>
</tr>
<tr>
<td>San Fernando 1971</td>
<td>(M = 6.5)</td>
<td>7</td>
<td>12</td>
<td>14</td>
<td>1.4</td>
<td>6.9</td>
<td>1</td>
</tr>
<tr>
<td>Northridge 1994</td>
<td>(M = 6.7)</td>
<td>12</td>
<td>18</td>
<td>24</td>
<td>0.99</td>
<td>5.2</td>
<td>2b</td>
</tr>
<tr>
<td>Borah Peak 1983</td>
<td>(M = 7.3)</td>
<td>23</td>
<td>40</td>
<td>20</td>
<td>0.96</td>
<td>2.9</td>
<td>1</td>
</tr>
<tr>
<td>Tottori 2000</td>
<td>(M = 6.8)</td>
<td>13</td>
<td>29</td>
<td>18</td>
<td>0.80</td>
<td>3.7</td>
<td>2e</td>
</tr>
<tr>
<td>Kobe 1995</td>
<td>(M = 6.9)</td>
<td>22</td>
<td>48</td>
<td>20</td>
<td>0.78</td>
<td>1.5</td>
<td>4</td>
</tr>
<tr>
<td>Kobe 1995</td>
<td>(M = 6.9)</td>
<td>24</td>
<td>60</td>
<td>20</td>
<td>0.62</td>
<td>1.7</td>
<td>2d</td>
</tr>
<tr>
<td>Imperial Valley</td>
<td>(M = 6.5)</td>
<td>5</td>
<td>30</td>
<td>10</td>
<td>0.56</td>
<td>1.3</td>
<td>1</td>
</tr>
<tr>
<td>Imperial Valley</td>
<td>(M = 6.6)</td>
<td>7.7</td>
<td>35</td>
<td>12</td>
<td>0.6</td>
<td>0.81</td>
<td>5</td>
</tr>
<tr>
<td>Imperial Valley</td>
<td>(M = 6.6)</td>
<td>8.6</td>
<td>42</td>
<td>11</td>
<td>0.6</td>
<td>1.8</td>
<td>2e</td>
</tr>
<tr>
<td>Morgan Hill 1984</td>
<td>(M = 6.2)</td>
<td>2.1</td>
<td>20</td>
<td>8</td>
<td>0.44</td>
<td>2.0</td>
<td>1</td>
</tr>
<tr>
<td>Morgan Hill 1984</td>
<td>(M = 6.3)</td>
<td>2.7</td>
<td>30</td>
<td>10</td>
<td>0.26</td>
<td>2.0</td>
<td>6</td>
</tr>
<tr>
<td>Morgan Hill 1984</td>
<td>(M = 6.3)</td>
<td>2.6</td>
<td>30</td>
<td>10</td>
<td>0.25</td>
<td>1.4</td>
<td>2f</td>
</tr>
<tr>
<td>Colfiorito 1997</td>
<td>(M = 5.9)</td>
<td>0.71</td>
<td>10</td>
<td>7</td>
<td>0.38</td>
<td>0.83</td>
<td>2g</td>
</tr>
<tr>
<td>North Palm Springs 1986</td>
<td>(M = 6.0)</td>
<td>1.8</td>
<td>18</td>
<td>10</td>
<td>0.33</td>
<td>0.15</td>
<td>1</td>
</tr>
<tr>
<td>Coyote Lake 1979</td>
<td>(M = 5.9)</td>
<td>0.35</td>
<td>6</td>
<td>6</td>
<td>0.32</td>
<td>0.57</td>
<td>1</td>
</tr>
</tbody>
</table>

<sup>a</sup>References are 1, Rice et al. [2005] based on slip inversions by Heaton [1990]; the G values are averages of G<sub>min</sub> and G<sub>max</sub>(=2 G<sub>min</sub>) of Rice et al. [2005], i.e., G=1.5 G<sub>min</sub> or 0.75 G<sub>max</sub>; 2. Tinti et al. [2005]; 2a. average of two models, G values +116% of mean; 2b. average of two models, G values +111% of mean; 2c. average of four models, of which one is an average of three models, G values +92% to −54% of mean; 2d. single model; 2e. average of two models, G values ±3% of mean; 2f. single model; 2g. average of three models, G values +34% to −52% of mean; 3. Peyrat et al. [2001]; 4. Guatteri et al. [2001]; 5. Favreau and Archuleta [2003]; 6. Beroza and Spudich [1988].

An additional quantity of interest is the net frictional work per unit area, equal to the total heat supplied per unit area of fault, \( \int_0^L \tau(\delta') d\delta' \). It is convenient to normalize that by the total heat that would have been supplied if there was no thermal pressurization process, i.e., if stress stayed constant at \( f(\sigma_p - p_0) \). The result is

\[
\frac{\int_0^L \tau(\delta') d\delta'}{f(\sigma_p - p_0)} = \frac{L*L}{\delta} \exp \left( \frac{6}{L*} \right) \operatorname{erfc} \left( \sqrt{\frac{\delta}{L*}} + 1 + 2 \sqrt{\frac{\delta}{\pi L*}} \right)
\]

and it shows that in large slip events, the heating is notably reduced from what would be the case if stress stayed constant at \( f(\sigma_p - p_0) \), and that level is already much smaller than is often assumed, because the flash heating effect has reduced \( f \). This means that conventional estimates of heating in large events as in section 1.2, even if they involved slip confined to a plane as in the present modeling, would in practical terms result in smaller heat outflow.

4. Predictions of Fracture Energy Compared to Seismological Estimates

The predictions of the slip on a plane model are that \( G \) increases substantially with slip during an event (Figures 4–6). That is a feature shared by recent assemblies on data for large numbers of earthquakes [Abercrombie and Rice, 2005; Rice et al., 2005; Tinti et al., 2005], all of which show a clear tendency for the average \( G \) in an event, as inferred from seismic slip inversions, to increase with the average slip in the event. The same feature was recognized previously by Ohnaka [2003] and interpreted by him as an increase of \( D_S \) with slip in the event. Further, Mikumo and Yagi [2003], Zhang et al. [2003], and Tinti et al. [2005] inferred the variation of \( G \) (or \( D_S \)) with position on the rupture plane within an individual events, and found that it did increase with slip. An alternative explanation for such an increase has been proposed on terms of inelastic energy dissipation off the main fault plane [Andrews, 2005], and such dissipation is surely a part of the overall fracture dissipation, although at present it cannot be ascertained how much is attributable to dissipation in slip-weakening processes within the narrow shear zone (the part modeled here), and how much to inelastic straining of nearby material.

Table 4 assembles data from two recent attempts to use a systematic single method to infer fracture energies for a large set of earthquakes. In one study, Rice et al. [2005] used their model of a propagating slip pulse to convert the slip inversion parameters for seven earthquakes presented by Heaton [1990] into estimates of fracture energies for those events. Their method establishes a factor of two range for the result, i.e., \( G_{\text{min}} < G \leq G_{\text{max}} \), where \( G_{\text{max}} = 2 G_{\text{min}} \), due to a parameter of the model ("R/L" of Rice et al. [2005]) being left undetermined by the slip inversion information. That factor of two difference is not large on the logarithmic scale on which the results are plotted (Figure 7) and thus Table 4 lists the average of the bounds, \( G = 1.5 G_{\text{min}} = 0.75 G_{\text{max}} \) for that event set. Also, Tinti et al. [2005] started with kinematic slip inversions for seven earthquakes (two of their events being members also of the previous set of seven) and applied a systematic smoothing procedure to estimate stress histories associated with the slip, and hence to estimate fracture energies. Their results for each event are represented by a single entry in Table 4, although, as the footnotes there explain, they often produced a set of possible models which, usually, were not very different from one another, at least considering the logarithmic scale. For the two events common to both the Tinti et al. and the Heaton and Rice et al. sets (Imperial Valley 1979, and Morgan Hill 1984), the results are reasonably...
Tinti et al. 2005. Their approach gives results for a vastly broader slip range, from submillimeter slips up to those with slips on the 1 m scale, for the large events as in Table 4 and Figure 7. The approach is based on the processing of seismic recordings to obtain radiated seismic energy $E_s$, moment $M_o$, and a measure of rupture source area $A$, typically by corner frequency, to get estimates also of slip $\delta$ and static stress drop (from initial stress $\tau_d$ to final static stress $\tau_1$) for a large collection of earthquakes. The smallest members were recorded at depth in the Cajon Pass borehole. The data set also includes Northridge aftershocks, and various larger events for which radiated energy, moment, and stress drop parameters have been estimated. For each event a parameter $G'$ was inferred which is generally expected to be comparable to $G$ and to coincide exactly with $G$ when the final dynamic friction stress $\tau_d$ during the last increments of seismic slip is negligibly different from the final static stress $\tau_1$ (i.e., when there is negligible dynamic overshoot or undershoot); $\tau_d$ would correspond to what has been written as $\tau(\delta)$ here, evaluated at the average final slip $\delta$ of the event.

$$G' = G + (\tau_d - \tau_1)\delta = (\tau_0 - \tau_1 - 2\mu E_s/M_o)\delta/2;$$

see Abercrombie and Rice (2005) for attempts to estimate how different $G'$ and $G$ might be due to differences between $\tau_d$ and $\tau_1$; e.g., dynamic overshoot models have $\tau_d > \tau_1$, whereas partial stress drop, or undershoot, models have $\tau_d < \tau_1$. The results for $G'$ for the larger slip events are generally consistent with what has been inferred for $G$ based on seismic slip inversions and summarized in Table 4. See the upper right portion of Figure 8, where the ovals are the $G$ values for the 12 events of Table 4 (with each event now being represented by a single $\delta$ and $G$ value, by averaging results where there is more than one $\delta$ and $G$ set in Table 4). The results also show a clear tendency for $G'$ to increase with slip in the event, at least for slips up to the order of ~1 m.

The curves drawn in Figure 8 are for the same three models as in Figure 7. It is seen that the two curves (red and black) based on the high-end choice for the thermoporoelastic parameters provide a plausible reproduction of main trends in the data, not just at large slip as in Figure 7 but also at slips of order of a few millimeters (where the two curves based on high-end properties essentially overlap one another). In contrast, the predictions based on the low-end parameters (blue curve) give a less suitable description of the data. There are many uncertainties in the seismic data that cannot be addressed here and of course in knowing what are defensible model parameters.

5. Conclusion

Earthquakes occur because fault strength weakens with increasing slip or slip rate. Here the question has been addressed of what physical processes determine how that weakening occurs during rapid, large slip, with the focus being on mature crustal faults, capable of producing large earthquakes. Field observations suggest that slip in individual events then occurs primarily within a thin shear zone, <1–5 mm, within a finely granulated (ultracataclastic) fault...
core. In absence of a strong weakening mechanism, temperature rise would lead to widespread melting, yet evidence of glass (pseudotachylyte) that would be left from rapid recooling is not pervasive on most exhumed faults. Relevant weakening processes in large crustal events are therefore likely to be thermal, and it is proposed here that the two primary processes are (1) thermal pressurization of pore fluid and (2) flash heating at highly stressed frictional microcontacts. For sufficiently large slip, a macroscopic melt layer may form too at high enough normal stress or, in silica-rich lithologies, weakening by gel formation may occur instead.

Theoretical modeling of mechanisms 1 and 2 has been constrained with lab-determined poroelastic and transport properties of intact fault core material, and frictional properties at high slip rates. Predictions are that strength drop should often be nearly complete at large slip and that the onset of melting should be precluded over much of the seismogenic zone. These are qualitatively consistent with low heat outflow from major faults.

A more quantitatively testable prediction is that of the shear fracture energies \( G \) that would be implied if actual earthquake ruptures were controlled by those thermal mechanisms. Seismic observations have recently been processed to allow inference, albeit with much uncertainty, of the \( G \) of large crustal events, with focus on its variation with slip in an event. The results are plausibly described by theoretical predictions based on the above mechanisms, within a considerable range of uncertainty of gouge material property choices, thus allowing the possibility that such thermal weakening prevails in the Earth. The property ranges which provide the best descriptions are those chosen to represent the effects of significant damage and inelastic straining in the concentrated stress field associated with passage of the rupture front, represented simply by increasing, relative to intact gouge, the permeability and, less so, the compressibility of the fault wall material during the slip and thermal heating process which then begins.

Among the uncertainties in interpreting the seismic data, to which the theoretical results have been compared, is that it has been assumed that all of what is inferred as \( G \) can be accounted by stressing and slip on the main fault plane. However, it is presently unknown what partitioning there is between contributions to the total \( G \) from such slip weak-
ening on the main fault plane and from distributed inelastic deformation in regions bordering that plane. Also, once the latter are recognized, there is uncertainty concerning how the values of $G$ inferred from seismic slip inversions, which purport to constrain the slip weakening part of $G$ but which assume elastic response off the main fault, actually relate to that slip weakening part.

At a more fundamental level, the physics underlying the thermal weakening processes studied here implies a strong dependence on slip rate and slip rate history. For simplicity, that dependence has been mapped into slip weakening here, with stress dependent on slip, by evaluating the models at what is thought to be a representative seismic slip rate. At a more fundamental level, fracture energy is a precisely defined concept in the frictional rupture case [Palmer and Rice, 1973; Rice, 1980] only for slip weakening with a well defined residual strength that is approached at large slip. The slip weakening laws dealt with here have the feature of an ever decreasing slope with increasing slip (Figure 3), so that it is probably suitable to regard that as a proxy for a well defined residual strength. However, the physical phenomena are fundamentally rate and history dependent and then, even assuming perfectly elastic response off the fault plane, there is no rigorous definition of $G$ unless the rate and history dependence of stress is compatible with the shear strength on the fault settling to a constant residual level (which could be zero) well behind the rupture front. Instead, strong strength variations with position there could be expected with certain velocity weakening laws (because $V$ is changing from large to zero values at say, the arresting tail of a slip pulse). A constant residual level at effectively zero strength could be expected for large slip events which achieve close to full thermal pressurization and reduce strength to a negligible level, but not more generally, especially due to the strong sensitivity of $f$ to $V$ implied by flash heating.

There is a clear need for laboratory tests of the thermal mechanisms discussed. Because they provide estimates of fracture energy that agree plausibly with seismic data only when speculative corrections for damage and inelastic deformation of the fault wall at the rupture front, and during subsequent slip, are made, as in Appendix A here, it is essential also to obtain better laboratory constraints on the instantaneous permeability and poroelastic moduli, and on dilatancy, for gouge materials undergoing shear. At present such properties are known only for intact gouge that is not actively shearing, and then only from a limited number of studies.

Also, it is important to understand how the thermal pressurization mechanism operates in conditions that might exist at mid crustal depths, possibly involving a mineralized pore space with pressure solution at gouge particle contacts, and unconnected pockets of liquid water. Shear of such a zone, initiated by the propagating rupture tip, would then seem to be necessary for fluid connectivity to be reestablished so that the effective stress characterisation of strength, $\tau = f(\sigma_n - p)$, holds.

The thermal mechanisms discussed here are those which are expected to be dominant from the start of seismic rupture, but in some cases there will be transitions to melting, or perhaps to gel formation, at large enough slip. Clearly, the theory of the transitions to the larger slip mechanisms needs to be developed and tested. Further, there is need to better understand how to represent the shear strength of a dense, rapidly shearing gouge with $p \approx \sigma_n$. Is it sufficiently small to be negligible compared to $\tau = f(\sigma_n - p)$? Or does it represent a substantial fraction of that value? Can the condition of $p$ approaching $\sigma_n$ actually be achieved on a real fault? It is sometimes speculated that highly pressurized pore fluids would hydraulically crack their way into the already damaged walls of the fault. The induction of large fault-parallel compressive effective stress very near the slip surface (Appendix A) should generally mitigate against that. However, local stress alterations near the rupture front [Poliakov et al., 2002; Rice et al., 2005], associated with stress concentration, could allow some local hydraulic cracking at a fault wall in that region, particularly on the extensional side of a mode II rupture.

The type of modeling presented describes a fault which is statically strong, in that a high static friction coefficient controls nucleation, but once dynamic slip is nucleated, there is a rapid and extreme weakening of the fault, down to essentially negligible strength in sufficiently large slip. So it is dynamically weak and, in fracture mechanics terminology, it is “brittle”. Such scenarios require enormous strength drops, say, at the 7 km depth focused on here, from a static strength of perhaps 75 MPa to get slip started (with $f \sim 0.6$) at a generic point of the rupture surface, to effectively zero if large enough slip occurs.

It is often assumed that such scenarios would also imply a seismic stress drop comparable to that strength drop, which would then be much larger than stress drops inferred seismically or geodetically, and so could not be a realistic process for the Earth. However, a strength drop is not the same as a stress drop. The later is, rather, the change in the shear stress from a static prestress acting shortly before rupture nucleates somewhere on the fault to a time shortly after the rupture when waves have propagated away from the ruptured region, and a static state has been restored. There is no need for the average static shear prestress over all the fault to be at all comparable to the average static strength. Rather, if “defect regions” [Rice, 1996; Lapusta and Rice, 2003] exist where rupture is readily nucleated, then the propagating rupture fronts so nucleated will concentrate stress on parts of the fault which were prestressed well below the strength, and thus sometimes enable the rupture zone to grow well beyond its region of nucleation. That is a well known process at the heart of brittle crack theory. In the fault mechanics context, those defect regions could be sites of concentrated shear stress as at boundaries between locked and creeping zones, or regions of low normal stress due, e.g., to nonplanarity of faults with releasing bends or with rupture branches that have locally perturbed stress, or local regions of significantly elevated pore pressure, perhaps due to a low permeability cap to an upflowing pore fluid.

The only requirement on the prestress, then, is that if a significant earthquake is to occur, the prestress must be sufficiently large to drive rupture over a substantially greater region than that where it nucleated. That does not require prestress comparable to the static strength. Both the Rice [1996] and Lapusta and Rice [2003] studies verified such concepts in numerical modeling of sequences of earth-
quakes in a simple crustal plane model of a major strike-slip fault. That assumed a Lachenbruch [1980] type of exponential slip weakening, from high static to negligible residual strength in one study, and strong rate weakening at high slip rates, roughly based on the flash heating concept, in another study. The results showed that a fault zone, with defect regions allowing easy nucleation, could operate at average prestress levels which could be far below the average static strength, and produce great ruptures of seismically reasonable stress drops that corresponded to negligible long term heat outflow.

Appendix A: Porothermoelastic Background

A1. Volumetric Porothermoelastic Response Under Isotopic Confining Stress

[90] Wibberley [2002] studied the Median Tectonic Line gouge and showed in his Figure 3b the change in pore volume $V_{pore}$ normalized by the original total sample volume $V_{tot,o}$ as a function of the confining pressure $p_c$ (positive in compression), where $p_c$ was applied isotropically (i.e., applied stress state $\sigma_{ij} = -\sigma_i \delta_{ij}$) on a sample which was maintained at constant pore pressure $p$. The ratio $V_{pore}/V_{tot,o} = n$, where $n$ is the apparent fluid volume fraction of Rice and Cleary [1976]; $n$ is essentially the porosity $V_{pore}/V_{tot}$ if $V_{tot}$ is nearly equal to its reference state value $V_{tot,o}$. C. A. J. Wibberley (private communication, 2003) provided the slope $-\partial n(\sigma_c, p, T)/\partial \sigma_c$ (difficult to read from the plot) at different $\sigma_c = p$, leading to the $-\partial n(\partial \sigma_c$ values shown in Table 1 for the ultracataclastic clayey gouge containing the principal slip surface, and also provided the values of $n (\beta_{c,ij})$ in Table 1), determined by measuring the absolute $V_{pore}$. The gouge was compacted to $\sigma_c = p = 180$ MPa and then studied at various state of unloading (unloading and reloading below 180 MPa are then approximately reversible) to provide the results. Similar data are shown by Noda and Shimamoto [2005] for ultracataclastic, clayey gouge of the Hanaore fault in southwest Japan, which has higher porosity and permeability.

[91] To interpret the measurements, note that the trace of the strain tensor is $\varepsilon_{kk} = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz} = (V_{tot} - V_{pore})/V_{tot,o}$ and the porothermoelastic constitutive relations, linearized about a reference state at $\sigma_{c,ref}, p_{ref}, T_{ref}$, where $T$ is temperature, are

$$
\varepsilon_{kk} = -\beta_d (\sigma_c - \sigma_{c,ref}) + (\beta_d - \beta_s)(p - p_{ref}) + \lambda_s (T - T_{ref}),
$$

$$
n - n_{ref} = -(\beta_d - \beta_s)(\sigma_c - \sigma_{c,ref}) + [\beta_d - (1 + n_{ref})\beta_s](p - p_{ref}) + \lambda_n (T - T_{ref}).
$$

(A1)

Here the coefficients are constrained by the requirement that $-\sigma_{c,ref} + p_{ref} + \rho_n$ is a perfect differential at fixed $T$ [Rice and Cleary, 1976], and $\beta_d$ is the drained compressibility of the porous solid. Also, to simplify, it has been assumed that all pore space is interconnected and that the material of each solid element of the gouge would exhibit an isotropic volumetric response to isotropic local stressing, with the same compressibility $\beta_s$ in each element (implying $(n - n_{ref}) = -\beta_s (p - p_{ref})\rho_n$ when the confining stress equals the pore pressure), and an isotropic response to unconstrained temperature change, with the same volumetric thermal expansion $\lambda_s$. That allows the various coefficients to be expressed in a standard manner in terms of $\beta_d, \beta_s$, and $\lambda_s$. The dependence on $\sigma_c$ and $p$ can be grouped as $\sigma_c - (1 - \beta_s/\beta_d)p$, showing the expected Biot factor $(1 - \beta_s/\beta_d)$ for elastic response. Thus $-\partial n(\sigma_c, p, T)/\partial \sigma_c = \beta_d - \beta_s$ and since $\beta_s$ is generally small compared to $\beta_d$, we need estimate it only approximately and use $\beta_s = 1.6 \times 10^{-11}/Pa$, a compromise between $1.2 \times 10^{-11}/Pa$ for mica [Wibberley, 2002] and $\sim 2 \times 10^{-11}/Pa$ for granite. Thus elastic changes in $n$ at fixed confining stress $\sigma_c$ can be written

$$
\frac{dn}{d\sigma_c} = n((\beta_d dp + \lambda_s dT) \left[ \frac{\beta_d - \beta_s}{n} - \beta_s \right] and \lambda_n = \lambda_s, \right)
$$

(A2)

and $\beta_s$ is shown in Table 1. Also, $\lambda_s = 2.4 \times 10^{-5}/^\circC$ (three times the linear thermal expansion coefficient for granite) is adopted.

A2. Porothermoelastic Expressions for Conditions of Constant Fault-Normal Stress and Vanishing Fault-Parallel Strains

[92] A form for $dn$ like that just given will be required in formulating the fluid mass conservation equation. However, in our application it is not $\sigma_c$ which should be regarded as constant but, instead, only the fault-normal compressive stress component $\sigma_n (\sigma_n = -\sigma_{yy})$. The fault-parallel components $\sigma_{xx}$ and $\sigma_{zz}$ near the sliding surface will vary during the slip history; however, they must in order to keep the fault-parallel strain components $\varepsilon_{xx}$ and $\varepsilon_{zz}$ at essentially zero values (at least once a material point is out of influence of the strong stress perturbations at the rupture tip). The goal in this section is therefore to estimate the factors $\beta_n$ and $\lambda_n$ which enter into

$$
\frac{dn}{d\sigma_n} = n((\beta_n dp + \lambda_n dT) \left[ \text{at fixed } \sigma_{yy}, \varepsilon_{xx} \text{ and } \varepsilon_{zz} \right]. \right)
$$

(A3)

[93] We can distinguish two end-member cases for that as follows:

[94] 1. In the first case, the wall material adjoining the sliding fault is modeled as if it remains intact and elastic during the heating and pressurization. The notations $\beta_n, \lambda_n$ are then used to denote $\beta_n, \lambda_n$. Expressions for them are derived just after explanation of case 2, and they are tabulated in Table 1 for the MTL gouge. The difficulty with this first case is that because of the very large temperature excursions that will be calculated based on it, the fault-parallel stresses $\sigma_{xx}$ and $\sigma_{zz}$ are predicted to become implausibly large in compression. One must then expect the wall material to fail to remain elastic, especially since the pore pressure is being elevated toward $\sigma_c$. That means that fault-parallel plastic strains would develop (with the sum of fault-parallel thermal, elastic and plastic strains necessarily being zero), which would have the effect of letting $\sigma_{xx}$ and $\sigma_{zz}$ relax toward equality with $\sigma_{yy}$ (i.e., with $-\sigma_n$), a plastic relaxation which is neglected in this first case. This is all the more a concern because theoretical stress modeling for propagating ruptures and other considerations (section 3.2) suggest that the fault wall material is likely to have been freshly damaged and inelastically deformed as it begins the large shear slippage process being modeled here.
[92] The second case accounts very approximately for that damage and inelastic response, which must be expected to open flow passages within the pore space and increase permeability. It does so by addressing an opposite end-member for which the fault-parallel stress relaxation is modeled as if it was complete, so that all normal stresses remain equal to one another (i.e., as in the isotropically confined situation, but with \( \sigma_r = \sigma_n \)). The simple representation of damage effects adopted in conjunction with that is to assume increased values of drained compressibility and permeability, denoted \( \beta_{dmg} \) and \( K_{dmg} \), respectively, relative to those (denoted \( \beta_d \) and \( k \)) for the undamaged gouge. In Tables 1 and 2, results are shown for \( \beta_{dmg} = 2\beta_d \) and \( K_{dmg} = 10 \) \( k \), whereas in Table 3 results are shown for various combinations of \((\beta_{dmg}, K_{dmg})\), ranging from \((\beta_d, k)\) to \((2\beta_d, 10 \) \( k \)). For this case with damage considered, the notations \( \nu_{dmg} \), \( \lambda_{dmg} \) are used to denote \( \beta_d \), \( \lambda_n \), and since this end-member assumes fault-parallel stress relaxation, we take \( \beta_{dmg} = \beta_d \) as would be calculated based on \( \beta_{dmg}, \) i.e.,

\[
\beta_{dmg} = \beta_d
\]  

and take \( \lambda_{dmg} = \lambda_d = \lambda_n \). That makes \( \beta_{dmg} \) approximately twice as large as \( \beta_d \) of case 1 even when \( \beta_{dmg} = \beta_d \).

[96] To analyze case 1 requires a further condition on details of the porothermoelastic stress-strain description of the fault wall material, and for simplicity, that is taken as elastically isotropic. In that case, the constitutive relations for \( \varepsilon_{kk} \) and \( n - n_o \) will have exactly the same form as above if we reinterpret \( \sigma_r \) as \(-\sigma_{dd}/3\), and we need only supplement them with the expressions relating deviatoric stress and strain, namely,

\[
\sigma_{ij} - \sigma_{kk} \frac{\varepsilon_{ij} \sigma_{kk} - \sigma_{kk} \varepsilon_{ij}}{3} = 2\mu (\varepsilon_{ij} - \varepsilon_{kk} \delta_{ij})/3,
\]

where \( \mu \) is the shear modulus and \( \sigma_{dd} \) denotes stresses in the reference state before the slip process begins. Given the constraints

\[
\varepsilon_{kk} = 0 + \varepsilon_{yy} + 0 = \varepsilon_{yy}, \quad -\sigma_{kk} - \sigma_{kk} \frac{\varepsilon_{kk} - \sigma_{kk}}{3} = 2\mu (\varepsilon_{kk} - \varepsilon_{kk} \delta_{kk})/3,
\]

and substituting into the above volumetric constitutive relations, and rearranging to eliminate \( \sigma_{kk} \)/3, then gives

\[
\varepsilon_{yy} = -\frac{\beta_d}{1 + r} (\sigma_r - \sigma_n) + \frac{\beta_d - \beta_d}{1 + r} (p - p_o) + \frac{\lambda_n}{1 + r} (T - T_o),
\]

\[
(n - n_o) = -\frac{\beta_d}{1 + r} (\sigma_r - \sigma_n) + \frac{(\beta_d - \beta_d)(\beta_d + r\beta_d)}{(1 + r)\beta_d} \frac{n_o \beta_d}{(1 + r)\beta_d} \cdot (p - p_o) + \lambda_n \left( n_o \frac{r(\beta_d - \beta_d)}{(1 + r)\beta_d} - n_o \beta_d \right) (T - T_o),
\]

where \( r = 4\mu\beta_d/3 = 2(1 - 2\nu_d)/(1 + \nu_d) \) and \( \nu_d \) is the Poisson ratio under drained conditions. This model neglects any fault-parallel plastic straining, even though the thermally induced fault-parallel compressive stresses will be seen to ultimately take on implausibly large values.

[97] Thus, when we write \( dn = n(\sigma'_{dd}dp + \nu'_{dd}dT) \) at fixed compressive normal stress \( \sigma_n \) the parameters are given, under the case 1 assumption of elastic fault wall behavior, as

\[
\beta_{d}^{el} = \left( \beta_d - \beta_d \right) \frac{\beta_d + r\beta_d}{n(1 + r)\beta_d} - \beta_d, \quad \lambda_{n}^{el} = \lambda_n \left( 1 - \frac{r \beta_d - \beta_d}{n\beta_d} \right).
\]

\[
(\text{A8})
\]

\( \beta_d^{el} \) and \( \lambda_{n}^{el} \) are given in Table 1. Typically, \( \nu_{d} \approx 0.20 \) for gouge, corresponding to \( r = 1 \), which is used to estimate \( \beta_d^{el} \) and \( \lambda_{n}^{el} \) in Table 1. It can be seen that \( \lambda_{n}^{el} \) is of opposite sign to \( \lambda_n \), and is 5 to 8 times larger in magnitude. That is because the normal effect of temperature rise to expand pore space is strongly counteracted by a second effect, particular to the constraint of zero fault-parallel strain change, which is that the thermal expansion induces large fault-parallel compressive stresses which act to reduce porosity.

[98] To solve for the predicted fault-parallel stresses \( \sigma_{xx} \) and \( \sigma_{zz} \), note that the deviatoric stress-strain relation above gives

\[
\sigma_{xx} = \sigma_{xx.o} = \sigma_{zz} = \sigma_{zz.o} = -\left( \sigma_n - \sigma_{n.o} \right) - 2\mu\varepsilon_{yy}
\]

and from that

\[
\sigma_{xx} - \sigma_{xx.o} + (1 - \beta_d/\beta_d) (p - p_o) - \frac{\nu_d}{1 - \nu_d} \left( \sigma_n - \sigma_{n.o} \right) - (1 - \beta_d/\beta_d)(p - p_o)
\]

\[
- \frac{2(1 + \nu_d)}{3(1 - \nu_d)} \mu \lambda_n (T - T_o).
\]

\[
(\text{A10})
\]

In the case of interest here \( \sigma_n = \sigma_{n.o} \) and we assume \( \sigma_{xx.o} \approx \sigma_{zz.o} \approx -\sigma_{n.o} \). Arranging the result in terms of effective compressive stresses like \( \sigma_n - p \), this implies

\[
(-\sigma_{xx}) = p = (\sigma_n - p) + \frac{1 - 2\nu_d}{1 - \nu_d} (1 - \beta_d/\beta_d)(p - p_o)
\]

\[
+ \frac{2(1 + \nu_d)}{3(1 - \nu_d)} \mu \lambda_n (T - T_o)
\]

\[
\approx (\sigma_n - p) + 0.55(p - p_o) + 0.31(T - T_o) \text{MPa}/\text{C}.
\]

\[
(\text{A11})
\]

In the last form, \( \beta_d, \lambda_n, \nu_d \) and \( \nu_d \) are chosen as above, and \( \beta_d = 5.8 \times 10^{-11}/\text{Pa} \), approximately correct for the MTL gouge at 126 MPa effective stress (Table 1), leading to \( \beta_d/\beta_d \approx 0.27 \) and \( \mu \approx 13 \text{ GPa} \). For the model of slip on a plane, it is shown in section A3 that the temperature and pore pressure changes immediately adjacent to the slipping plane are related by (section 3.4)

\[
T - T_o = \left( 1 + \frac{\alpha_{hy}/\alpha_h}{\alpha_h} \right) (p - p_o) / \Lambda \approx 2.90(p - p_o) \text{C/MPa},
\]

\[
(\text{A12})
\]

where the numerical term corresponds to intact elastic walls average on p-T path parameters as in Table 2 (i.e., estimating parameters appropriate to elastic response of the fault walls over a range of \( p \) and \( T \), starting at ambient conditions of 70 MPa and 210°C, and effective normal stress of 126 MPa, which are which are approximately appropriate conditions for a fault surface at 7 km depth).
Inserting that expression for temperature rise into the expression for effective stress,

\[ (-\sigma_{xx}) - p \approx (\sigma_n - p) + 1.45(p - p_o). \]  

(A13)

Clearly, as \( p \to \sigma_n \), which is what the thermal pressurization analysis suggests very near the fault, the predicted effective stress state approaches a purely biaxial state at \( \sim 1.4 \) times the initial \( \sigma_n - p_o \), and that could not be sustained elastically. Recognizing that during the weakening process, shear stress \( \tau = f(\sigma_n - p) \) where the \( f \) is reduced by flash heating, and taken to be 0.25 here, one can calculate all in-plane stress components, made nondimensional with \( \sigma_n - p \), in terms of the ratio \((p - p_o)/(\sigma_n - p)\). It is thus possible to determine when the Mohr-Coulomb failure condition is reached in the gouge of the fault wall. Assuming \( f = 0.60 \) governs Mohr-Coulomb failure (gouge particle contacts in the wall will not be shearing rapidly enough to be susceptible to weakening by flash heating), failure is expected to commence when \( p - p_o \) reaches 58% of \( \sigma_n - p_o \). Thus, if the fault wall gouge has been damaged at the rupture tip, so that it behaves as a cohesion-less granular material, then subsequent response cannot remain elastic during substantial thermally induced pressure increase.

[99] Also, as commented above in connection with case 2, such damage will increase the permeability. That will lessen the slip surface pore pressure rise associated with a given temperature rise, or rested, will increase the fault plane temperature rise associated with a given pore pressure rise above the 2.67°C/MPa quoted above. The result is to assure that Mohr-Coulomb failure of the wall will commence at yet smaller pore pressure rise. For those reasons, several evaluations of how wall damage would affect the principal results are given (Table 3), with that damage being parameterized simply in terms of values of \( \partial\theta/dT \) and \( k^{\text{mig}}/k \) for case 2.

A3. Fluid Mass Content in Terms of \( p \) and \( T \)

[100] The above developments for cases 1 and 2 are in aid of evaluating the term \( dm/d\tau \) in the fluid mass conservation equation. There, \( m \) is defined as the mass of fluid divided by \( V_{tot,0} \), and therefore \( m = \rho_f n \) where \( \rho_f \) is the local density of fluid in the pore space. Clearly, \( dm = \rho_f n (dp_f/\rho_f + dn/\rho_f) \); we have seen how to write \( dn \) (except for an inclusion of inelastic dilatancy, done in the manner of Segall and Rice [1995]), so that

\[ \dot{n} = n(\beta d\rho_f + \lambda dT) + d\rho_f', \]  

(A14)

where the additional “plastic” term here represents inelastic dilatancy. Also, because \( \rho_f = \rho_f(p, T) \),

\[ dp_f = \rho_f (\beta dp_f - \lambda dT), \]  

(A15)

where \( \beta \) is the fluid compressibility and \( \lambda \) is its volumetric thermal expansivity; see the Table 2 caption for the data sources for those used here. Thus

\[ dm/d\tau = n((\beta + \beta_d) dp - (\lambda_f - \lambda_m) dT) + d\rho_f' \]  

\[ = \beta \left( dp - \Lambda dT + \frac{d\rho_f'}{\beta} \right), \]  

(A16)

where \( \beta \) and \( \Lambda \) are defined by this expression and written explicitly in section 3.2.

Appendix B: Notes on Formulation and Some Solutions for Thermal Pressurization Model


[101] Here we are concerned with solving the pair of equations

\[ \frac{\partial T}{\partial t} - \frac{\tau(y, t)}{\rho c} = \frac{\partial^2 T}{\partial y^2} \]  

and

\[ \frac{\partial p}{\partial t} - \frac{\Lambda}{\rho c} = \frac{\partial^2 p}{\partial y^2} \]  

(B1)

for arbitrary histories of shear stress \( \tau(t) \), independent of \( y \) as required by the shear equilibrium equation, or equation of motion at the small spatial scale considered, and shear strain rate histories \( \dot{\gamma}(y, t) \). In this phase of the presentation, we do not consider any specific friction law, like \( \tau = f(\sigma_n - p) \), or more general constitutive law relating the stress to the strain rate or, if confined to a plane, to the slip rate. Further, we adopt the simplified representation of dilatancy discussed earlier (sections 3.2–3.4), so that it just resets the initial value assumed for pore pressure from \( p_{amb} \) to \( p_o = p_{amb} - \Delta n^0/\beta \).

[102] The solution to the energy equation is well known [Carslaw and Jaeger, 1959] when a layer of unit thermal energy per unit area is inserted at \( t = 0 \) along the plane \( y = 0 \); that defines the Green’s function \( G(y, t; \alpha, \alpha_h) \) such that the solution is

\[ T(t) - T_{amb} = \left( \frac{1}{\rho c} \right) G(y, t; \alpha) \]  

where

\[ G(y, t; \alpha) = \frac{1}{2\sqrt{\pi \alpha t}} \exp \left( -\frac{y^2}{4 \alpha t} \right) \]  

when \( t > 0 \),

and

\[ G(y, t; \alpha) = 0 \]  

when \( t < 0 \). (B2)

Note that \( \int_0^\infty G(y, t; \alpha)dy = 1 \) for \( t > 0 \). The corresponding pore pressure Green’s function is [Lee and Delaney, 1987]

\[ p(y, t) - p_o = \frac{\Lambda}{\rho c} \left( \frac{\alpha_{hy}}{\alpha_h} - \frac{\alpha_{hy}}{\alpha_h} \right) \]  

(B3)

if \( \alpha_{hy} \neq \alpha_{ih} \). If they coincide, then \( \partial[\alpha G(y, t; \alpha)]/\partial\alpha \) replaces the last fraction. Thus for arbitrary histories of dissipation rate \( \tau(t)\dot{\gamma}(y, t) \),

\[ T(t) - T_{amb} = \frac{1}{\rho c} \int_0^t \int_{-\infty}^\infty \tau(t')\dot{\gamma}(y', t') G(y - y', t - t'; \alpha) dy'dt', \]  

(B4a)

\[ p(y, t) - p_o = \frac{\Lambda}{\rho c} \int_0^t \int_{-\infty}^\infty \tau(t')\dot{\gamma}(y', t') \left( \frac{\alpha_{hy}}{\alpha_h} - \frac{\alpha_{hy}}{\alpha_h} \right) \]  

\[ \cdot dy'dt', \]  

(B4b)

It can be noticed that the solution for \( p \) has a symmetric dependence on the thermal and hydraulic diffusivities.
These solutions are for the model in which all inelastic dilatancy is represented by the simple resetting of the initial pore pressure. When that approximation is not appropriate, the solution for \( p(y, t) - p_{amb} \) is the same as in the expression given above for \( p(y, t) - p_{o} \), except that to the right side of the expression must be added the term

\[
-\frac{1}{5} \int_{0}^{t} \int_{-\infty}^{t} \partial \gamma^2(y', t') G(y - y', t - t'; \alpha_{h}) dy' dt'.
\]

### B2. Examples of Distributed Shear

Consider the Andrews [2002] assumption of a Gaussian shape, with mean-square half width \( w \), for the shear rate distribution, namely,

\[
\dot{\gamma}(y, t) = \frac{V(t)}{\sqrt{2\pi w^2}} \exp \left( -\frac{y^2}{2w^2} \right),
\]

where \( V(t) \) is the net slip rate accommodated. Then, doing the convolution over \( y' \) of \( \dot{\gamma}(y', t)V(t) \) with \( G(y - y', t; \alpha) \) to define a function arising in the Andrews [2002] solution,

\[
A(y, t; \alpha) = \frac{1}{\sqrt{2\pi \sqrt{w^2 + 2\alpha t}}} \exp \left( -\frac{y^2}{2(w^2 + 2\alpha t)} \right),
\]

the results are

\[
T(y, t) - T_{amb} = \frac{1}{\rho c} \int_{0}^{t} \tau'(t') V(t') A(y, t - t'; \alpha_{h}) dt',
\]

\[
p(y, t) - p_{o} = \frac{\Lambda}{\rho c} \int_{0}^{t} \tau'(t') V(t') \left[ \frac{\alpha_{h} A(y, t - t'; \alpha_{h}) - \alpha_{h} A(y, t - t'; \alpha_{h})}{\alpha_{h} - \alpha_{h}} \right] dt'.
\]

The second case is a spatially uniform shear rate within a layer of thickness \( h \) centered on \( y = 0 \), with no shear outside the layer, i.e., \( \dot{\gamma}(y', t) = V(t)h \) for \( y' \) within the layer. Convoluting that \( \dot{\gamma}(y', t)V(t) \) with \( G(y - y', t; \alpha) \) defines a new function

\[
B(y, t; \alpha) = \frac{1}{2h} \left[ \text{erf} \left( \frac{h + 2y}{4\sqrt{\alpha t}} \right) + \text{erf} \left( \frac{h - 2y}{4\sqrt{\alpha t}} \right) \right],
\]

and then the solutions for \( p \) and \( T \) are given by the same expressions as in the two previous equations, but with \( B(y, t - t'; \alpha) \) replacing \( A(y, t - t'; \alpha) \).

### B3. Possible Inconsistency of Assuming Distributed Shear

It may be proven that both of those solutions (and presumably any solution for a shear distribution that is symmetric about \( y = 0 \) and for which \( \partial \gamma \partial y \partial t \leq 0 \)), have a maximum in the pore pressure distribution at \( y = 0 \) at least for some finite time interval after the shear history begins; an analysis of that is given below by solving for \( \partial^2 p/\partial y^2 \) at \( y = 0 \) and ascertaining when it is provably negative. As noted in section 3, the assumption of such type of strain rate distribution leads, strictly, to an inconsistency if we insist that in all deforming regions the Coulomb law \( \tau = f(\sigma_n - p) \) be met, with a friction coefficient that is constant, or that weakens with increasing strain rate. (Here sustained shear is being considered, so that the rate weakening mentioned refers to the steady state \( f \); there is always rate strengthening when we look at the effects of sudden change of shear rate, and in some cases rate strengthening is observed in sustained, or steady state, shear too.

Consider the model of uniform shear within a layer of thickness \( h \). By direct calculations one finds that

\[
\frac{\partial p(y, t)}{\partial y^2} = -\frac{\Lambda}{\rho c} \int_{0}^{t} \tau'(t') \left[ G(h/2, t - t'; \alpha_{h}) - G(h/2, t - t'; \alpha_{h}) \right] dt'.
\]

Observing that \( \tau'(t') V(t') \) and \( t - t' \) are nonnegative, it is possible, after a little analysis, to show that the integrand is positive when

\[
t - t' < t_{pos} \equiv \frac{\alpha_{h} - \alpha_{h}}{8\theta_{h}} \frac{h^2}{\ln(\alpha_{h}/\alpha_{h})}
\]

and negative when \( t - t' > t_{pos} \). Thus \( p \) will have a maximum at \( y = 0 \) for all \( t < t_{pos} \), since no values of \( t' \) allowing the integrand to be negative will then have entered the integral. Also, \( p \) will continue to have such a maximum for at least some (and possibly all) \( t > t_{pos} \). No more precise statement can be made without specifying a particular \( V(t') \), \( \tau'(t') \) history because then the integrand will take both positive and negative values, and numerical evaluation would be needed in any event. As representative values, take \( \alpha_{h} = 0.6 \text{ mm/s} \) and \( \alpha_{h} = 3.0 \text{ mm/s} \). There results \( t_{pos} = 0.10(\text{h/mm})\text{s} \), which is 10 s when \( h = 10 \text{ mm} \) and 2.5 s when \( h = 5 \text{ mm} \). Recalling that 1 m/s is an average slip rate during a seismic event, those times correspond to slips of 10 m and 2.5 m, respectively, so for all practical purposes such thick shearing zones would have \( p \) maximum at \( y = 0 \) throughout the event, raising questions, as above, about the plausibility of the assumption of such thick shear zones.

### B4. Solution for Slip on a Plane

The special case of most interest here, motivated by the preceding discussion, is that of slip on a plane, \( \dot{\gamma}(y', t) = \delta_{\text{Dirac}}(y') V(t) \). The general solutions then reduce to

\[
T(y, t) - T_{amb} = \frac{1}{\rho c} \int_{0}^{t} \tau'(t') V(t') G(y, t - t'; \alpha_{h}) dt',
\]

\[
p(y, t) - p_{o} = \frac{\Lambda}{\rho c} \int_{0}^{t} \tau'(t') V(t') \left[ \frac{\alpha_{h} G(y, t - t'; \alpha_{h}) - \alpha_{h} G(y, t - t'; \alpha_{h})}{\alpha_{h} - \alpha_{h}} \right] dt'.
\]

\[
= \frac{\Lambda}{\alpha_{h}} \left[ \sqrt{\alpha_{h}} \left( T(y, \sqrt{\alpha_{h}}/\alpha_{h}, t) - T_{amb} \right) - \sqrt{\alpha_{h}} \left( T(y, t) - T_{amb} \right) \right].
\]
Remarkably, if we examine the changes of \( p \) and \( T \) on the fault plane \( y = 0 \) itself, the latter version of equation (B11b) shows that there is a universal relation

\[
T(0, t) - T_{\text{amb}} = \left( 1 + \frac{\sqrt{\alpha_{th}}}{\alpha_{th}} \right) \frac{p(0, t) - p_0}{\Lambda} \quad (B12)
\]

between them, valid no matter what the history of \( \tau(t) \) and \( V(t) \), where, to give one of the two,

\[
p(0, t) - p_0 = \frac{\Lambda}{2\sqrt{\pi} \left( \sqrt{\alpha_{th}} + \sqrt{\alpha_{d}} \right) pc_0} \int_0^t \frac{\tau(t')V(t')}{\sqrt{t - t'}} dt'. \quad (B13)
\]

For consistency it should be checked that the fault plane, where slip is assumed to take place, really is the location of highest pore pressure. It may be verified from the above representation that \( p, \partial p/\partial t, \) and \( \partial^2 p/\partial y^2 \) are continuous at \( y = 0 \) (and that \( \partial p/\partial y = 0 \) there). That is generally not so for higher derivatives but suffices for use of \( (\partial^2 p/\partial y^2)_{y=0} < 0 \) as the condition for \( p \) to have at least local maximum there. Thus, observing from the above expression that \( \Lambda(\partial T/\partial t)_{y=0} = (1 + \sqrt{\alpha_{th}}/\alpha_{d})\partial p/\partial t \), the PDE expressing fluid mass conservation shows that

\[
\alpha_{th} \left( \frac{\partial p}{\partial y} \right)_{y=0} = \left( \frac{\partial p}{\partial t} \right)_{y=0} - \Lambda \left( \frac{\partial T}{\partial t} \right)_{y=0} = -\sqrt{\alpha_{th}}/\alpha_{d} \frac{\partial p}{\partial t} \quad (B14)
\]

Thus, so long as \( \partial p/\partial t > 0 \) on the fault plane, which means that \( \partial T/\partial t > 0 \) there too, we have \( (\partial^2 p/\partial y^2)_{y=0} < 0 \), assuring that \( p \) has at least a local maximum. Hence no immediately neighboring plane is stressed beyond the friction threshold.

### B5. Constitutive Relation for Shear Strength at the Fault Plane

The above integral representation shows how to calculate \( p(y = 0) = p(0, t) \) along the fault, given \( \tau(t) \) and \( V(t') \) for \( 0 < t' < t \), but \( \tau \) is, of course, not known a priori and must be chosen to meet the friction law \( \tau(t) = \int_0^t \sigma_n(t') - p(t') dt' \). Here the notation \( f = f(t) \) is meant to emphasize that the friction coefficient may itself evolve with time, e.g., through a dependence of \( f \) on the slip rate \( V \) or, more precisely, through a dependence on \( V \) and its prior history, of the type described in rate and state friction (including the extension, yet to be made specific, of rate and state concepts to the rapid slip range of interest here, in which flash heating occurs). In this presentation it will be assumed that, given the history of \( V(t) \), the history of \( f(t) \) is then determined, as in the most common discussions of rate and state friction or, more simply, of velocity-dependent friction (which is a valid concept only when considering slips that are large enough, and have \( V \) changing slowly enough, that \( f \) can be equated to its steady state value at rate \( V \) [Rice et al., 2001]). However, that is not a complete description because \( f \) will also be dependent on the history of \( T \) and of effective stress \( \sigma_n - p \); those effects are ignored here.

Observing that \( f(t)[\sigma_n(t) - p_0] = f(t)(\sigma_n - p_0) - \tau(t) \), the foregoing integral representation becomes the integral equation

\[
f(t)(\sigma_n - p_0) - \tau(t) = \frac{f(t)\Lambda}{2\left( \sqrt{\alpha_{th}} + \sqrt{\alpha_d} \right)pc_0} \int_0^t \frac{\tau(t')V(t')}{\sqrt{t - t'}} dt', \quad (B15)
\]

which defines the constitutive relation for the fault plane, in the sense that for any given slip rate history \( V(t) \), which implies also the function \( f(t) \), we can solve the equation to determine the corresponding shear stress history \( \tau(t) \). For any given \( V(t) \), it is a linear equation in \( \tau(t) \), although it is easy to solve other than numerically only in special cases, one to be discussed.

The corresponding integral equation can be stated but not solved here for a distributed shear rate in the form of the Andrews [2002] Gaussian distribution. This makes the approximation that although the entire layer shears, the constitutive relation is based on the highest pore pressure, presumed to be at the layer center \( y = 0 \), which, as remarked, is not consistent with distributed shear. Thus its thickness scaled \( w \) should be regarded as some microstructurally demanded thickness, in recognition that slip on a mathematical plane is not a physical possibility in a granular aggregate. Taking strength \( \tau(t) = f(t)(\sigma_n - p(0, t)) \), the preceding integral representation of the Andrews solution leads to

\[
f(t)(\sigma_n - p_0) - \tau(t) = \frac{f(t)\Lambda}{2\sqrt{\pi} \left( \sqrt{\alpha_{th}} + \sqrt{\alpha_d} \right)pc_0} \int_0^t \left[ \frac{\alpha_{th}\tau(t')V(t')}{\sqrt{w^2 + 2\alpha_{th}(t - t')}} - \frac{\alpha_d\tau(t')V(t')}{\sqrt{w^2 + 2\alpha_d(t - t')}} \right] dt', \quad (B16)
\]

and, once solved, the temperature rise at the center of the shearing layer (no longer directly proportional to the pore pressure rise) is

\[
T(0, t) - T_{\text{amb}} = \frac{1}{2\sqrt{\pi}pc_0} \int_0^t \frac{\tau(t')V(t')}{\sqrt{w^2 + 2\alpha_d(t - t')}} dt'. \quad (B17)
\]

The solution to the integral equation for slip on a plane, in the simplest case, for which \( V \) and \( f \) are treated as being constant throughout the event, is used in the paper to predict the stress versus slip relation \( \tau = \tau(\delta) \). Slip \( \delta = Vt \) in this case and, rewritten in terms of slip, the integral equation reduces to

\[
f(\sigma_n - p_0) - \tau(\delta) = \frac{1}{\sqrt{L^*}} \int_0^\delta \frac{\tau(\delta')}{\sqrt{\pi(\delta' - \delta')}} d\delta', \quad (B18)
\]

where

\[
L^* = \frac{4}{f^2 V} \left( \frac{pc_0}{\Lambda} \right)^2 \left( \sqrt{\alpha_{th}} + \sqrt{\alpha_d} \right)^2.
\]

Note that the length parameter \( L^* \) inherits its length dimension from the ratio of a composite diffusivity to the
slip rate \( V \). The solution, found by steps outlined in section B6, is

\[
\tau = \tau(\theta) = f(\sigma_y - p_o) \exp \left( \frac{\theta}{L^*} \right) \text{erfc} \left( \frac{\theta}{\sqrt{L^*}} \right),
\]

(B19)

where \( \text{erfc}(z) = (2/\sqrt{\pi}) \int_z^\infty \exp(-u^2) \, du \) and \( \text{erfc}(0) = 1 \).

**B6. Solution Details**

The Laplace transform \( \hat{\tau}(s) \) is defined by

\[
\hat{\tau}(s) = \int_0^\infty \tau(\theta) \exp(-s \theta) \, d\theta,
\]

so that the integral equation, when transformed (noting that the transform of a convolution is a product of transforms), becomes

\[
f(\sigma_y - p_o)/s - \hat{\tau}(s) = \frac{\hat{\tau}(s)}{\sqrt{L^*} s},
\]

(B20)

giving

\[
\hat{\tau}(s) = \frac{f(\sigma_y - p_o)}{1 + \sqrt{L^*} s} \frac{\sqrt{L^*}}{s}.
\]

That \( \hat{\tau}(s) \) is analytic in the complex \( s \) plane except for the singular point at \( s = 0 \) and the branch cut, which can be taken to coincide with the negative \( \text{Re}(s) \) axis. Then the contour \( \Gamma \) in the inversion formula, which properly runs along the \( \text{Re}(s) > 0 \) side of the \( \text{Im}(s) \) axis, can be distorted to run along the branch cut, i.e., from \( -\infty \) to \( 0 \) along the \( \text{Im}(s) < 0 \) side of the \( \text{Re}(s) \) axis, and from \( 0 \) to \( -\infty \) along the \( \text{Im}(s) > 0 \) side of that axis. Thus with \( s \) written as \( \exp(\theta i) \),

\[
\tau(\theta) = \frac{1}{2\pi i} \int_\Gamma \frac{f(\sigma_y - p_o)}{1 + \sqrt{L^*} s} \exp(\theta i) \, ds
= \frac{f(\sigma_y - p_o)}{2\pi} \int_0^\infty \frac{\exp(-\theta i) L^* \, ds}{1 + \sqrt{L^*} r}
\]

(B21)

which can be rewritten, using \( u = \sqrt{L^*} r \), as

\[
\tau(\theta) = \frac{2}{\pi} \int_0^\infty \exp \left( \frac{-u^2}{1 + u^2} \right) \frac{du}{1 + \sqrt{L^*} u^2}
\]

(B22)

Now, call the last integral \( I(\theta) \), noting that \( I(\infty) = 0 \), and calculate \( dI(\theta)/d\theta \). The result is some factors times a new integral, recognizable as \( \sqrt{\pi}/2 \), so that

\[
\frac{dI(\theta)}{d\theta} = -\frac{\sqrt{\pi}}{2\sqrt{L^*}} \exp(-\theta/\sqrt{L^*})
\]

(B23)

Thus we obtain \( \tau(\theta)/[f(\sigma_y - p_o)] = \exp(\theta L^*/\sqrt{L^*}) \text{erfc}(\sqrt{L^*/\sqrt{L^*}}) \).

The \( \exp(D \text{erfc}(\sqrt{D})) \) solution (with \( D = \sqrt{L^*} \)) just derived for arbitrary values of \( \alpha_{lb} \) and \( \alpha_{by} \) is of the same form as one presented by Mase and Smith [1987] for slip on a plane in the special case \( \alpha_{by} = 0 \). That is a case in which there is no fluid motion, so that response is locally undrained. The present solution coincides precisely with theirs when we evaluate \( L^* \) by setting \( \alpha_{by} = 0 \). To see why that should be so, note that when \( \alpha_{by} = 0 \), the differential equations governing for slip on a plane reduce to

\[
\frac{\partial T}{\partial y} = \alpha_{lb} \frac{\partial^2 T}{\partial x^2};
\]

(B24a)

On \( y = 0 \),

\[
\frac{\partial T}{\partial y} (y = \pm 0, t) =
\]

\[
\frac{1}{2} \tau(t) V = \frac{1}{2} \int \left[ (\sigma_y - p_o) - (p(0, t) - p_o) \right] V;
\]

(B24b)

For all \( y \),

\[
p(y, t) - p_o = \Lambda(T(y, t) - T_{amb}) \Rightarrow p(0, t) - p_o
\]

\[
= \Lambda[(0, t) - T_{amb}];
\]

(B24c)

Expression (B24c) converts (B24b) into a boundary condition on \( T \) alone, and as Mase and Smith noted, that converts the problem into one solved in another context by Carslaw and Jaeger [1959], with solution in that \( \exp(D \text{erfc}(\sqrt{D})) \) form. Once we realize that in our case with \( \alpha_{by} = 0 \), there does nevertheless remain the strict proportionality derived above between \( T(0, t) - T_{amb} \) and \( p(0, t) - p_o \) (which had in fact been assumed as a plausible approximation by Mase and Smith [1987] based on their derivation of the correct relation between \( T_{max} - T_{amb} \) and \( \sigma_y - p_o \); see discussion in section 3.4), the same adaptation of the Carslaw and Jaeger [1959] solution is good for the more general \( \alpha_{by} \neq 0 \) case. That has a further benefit because Carslaw and Jaeger [1959] also derive the complete expression for the temperature field \( T(y, t) \), reported also by Mase and Smith [1987]. The result is

\[
T(y, t) - T_{amb} = \left( 1 + \sqrt{\frac{\alpha_{lb}}{\alpha_{by}}} \frac{\sigma_y - p_o}{\Lambda} \right)
\]

\[
\cdot \left[ \text{erfc} \left( \frac{Y}{2\sqrt{D}} \right) - \exp(Y + D)\text{erfc} \left( \frac{Y + \sqrt{D}}{2\sqrt{D}} \right) \right],
\]

(B25)

where \( Y = |y|/\sqrt{\alpha_{lb}L^*/V} \) and \( D = \sqrt{L^*} \).

**B7. General Series Solution for Temperature Near the Slip Plane and Finite Shear Layer Effects on Predicted Temperatures**

[117] The above solution is exact, at least within linearization of the PDE pair for \( p \) and \( T \), but it applies only for the simplified case in which \( f \) and \( V \) are constant during the slip. We can learn something about the more general case as follows: Consider the domain \( y > 0 \), and the Taylor series in \( y \),

\[
T(y, t) - T_{amb} = T(0, t) - T_{amb} + \frac{\partial T}{\partial y} (0^+, t)
\]

\[
+ \frac{1}{2!} \frac{\partial^2 T}{\partial y^2} (0^+, t) + \ldots .
\]

(B26)

From the PDE for \( T \) and its boundary condition, as well as from the direct proportionality of \( p(0, t) - p_o \) to \( T(0, t) - T_{amb} \),
\[ T_{amb}, \text{ which was shown to hold no matter how } f \text{ and } V \text{ vary with time,} \]
\[ -\rho c_{oh} \frac{\partial T}{\partial t}(0^+, t) = \frac{1}{2} f(t)(\sigma_n - p(0, t))V(t), \tag{B27} \]
\[ \frac{\partial^2 T}{\partial y^2}(0^+, t) = \frac{1}{\alpha_{oh}} \frac{\partial T}{\partial t}(0, t) = -1 + \frac{\sqrt{\gamma_{oh}/\alpha_{oh}}}{\sigma_n} \frac{d}{dt}(\sigma_n - p(0, t)). \]

These are substituted into the series and some terms are rearranged to give the general result
\[ T(y, t) - T_{amb} = \Delta T_{\text{max}} \left[ 1 - \frac{\sigma_n - p(t)}{\sigma_n - \rho_0} - y \frac{\sigma_n - p(t)}{\sigma_n - \rho_0} \right. \]
\[ \left. - \frac{Y^2}{2(\sigma_n - \rho_0)} \frac{L^*(t)}{V(t)} \frac{d}{dt}(\sigma_n - p(t)) + \ldots \right] \tag{B28} \]

where \( \Delta T_{\text{max}} = (1 + \sqrt{\gamma_{oh}/\alpha_{oh}})(\sigma_n - \rho_0)/\Lambda \) as before, where \( p(t) \) is a shorthand for \( p(0, t) \), i.e., the pressure on the slip plane, where \( L^*(t) \) is the same \( L^* \) as before but now defined in terms of the time dependent \( f(t) \) and \( V(t) \) (which means that \( L^*(t) \propto 1/\sqrt{f(t)V(t)} \)), and \( Y = |y|/\sqrt{\gamma_{oh}L^*(t)/V(t)} \).

[118] An application of this result is to find the average temperature over a small distance \( h \) about the fault plane. That is of interest because if the slip zone is not a mathematical plane but rather is better idealized as a narrow shearing layer of thickness \( h \), then the maximum temperature rise at any given time will be less than that predicted for slip on a plane. A full exploration of that is in preparation (Rempel and Rice, submitted manuscript, 2006), but a simple estimate is obtained by identifying the maximum temperature \( T(t) \) at any given time with the average of \( T(y, t) \) over the strip \(-h/2 < y < +h/2\), so that
\[ T(t) - T_{amb} = \Delta T_{\text{max}} \left[ 1 - \frac{\sigma_n - p(t)}{\sigma_n - \rho_0} - H \frac{\sigma_n - p(t)}{\sigma_n - \rho_0} \right. \]
\[ \left. - \frac{2H^2}{3(\sigma_n - \rho_0)} \frac{L^*(t)}{V(t)} \frac{d}{dt}(\sigma_n - p(t)) + \ldots \right] \tag{B29} \]

where \( H = h/\sqrt{16\alpha_{oh}L^*(t)/V(t)} \). The series can be used with reliability if the ratio \( r^* \) of the magnitude of the quadratic in \( H \) term to that of the linear in \( H \) term is small compared to 1, and truncated with the linear in \( H \) term if \( r^* \gg 1 \), where
\[ r^* = -\frac{2H}{3(\sigma_n - p(t))} \frac{L^*(t)}{V(t)} \frac{d}{dt}(\sigma_n - p(t)). \tag{B30} \]

To estimate \( r^* \) in a simple case, let us return to the model with constant \( f \) and \( V \). Then \( \sigma_n - p(t) \) follows the \( \exp(D)\text{erfc}(\sqrt{D}) \) form, which falls off as \( \propto 1/\sqrt{D} \) at large \( D \), so that when it is noted that \( Vdt \text{d}b \), an increment of slip, there results \( r = (2H/3)(L^*/V)(dV/db) \approx HL^*/3b \) when \( b \) is a few or more times \( L^* \).

[119] Consider, the Punchbowl fault as studied by Chester and Chester [1998] and Chester et al. [2003]; see Figure 1. The candidate values of \( h \) are \( \sim 0.2 \text{ mm} \) if to represent the heavily sheared core, and \( h \sim 1.0 \text{ mm} \) if to represent the zone that has been visibly affected (in polarized light) by the shear. For Table 2, these values give the following: \( h = 0.2 \text{ mm} ; H = 1.69, 1.31, 0.35, \text{ and } 0.32, \text{ respectively;} \) and \( h = 1.0 \text{ mm} ; H = 8.45, 6.57, 1.73, \text{ and } 1.59, \text{ respectively.} \) Thus, assuming we are looking at large enough slips so that \( r^* \ll 1 \), we can estimate the maximum temperature in the fault zone of finite but small thickness \( h \) by
\[ T(t) = T_{amb} + \Delta T_{\text{max}} \left( 1 - (1 + H) \frac{\sigma_n - p(t)}{\sigma_n - \rho_0} \right) \]
versus \( T(0, t) = T_{amb} + \Delta T_{\text{max}} \left( 1 - \frac{\sigma_n - p(t)}{\sigma_n - \rho_0} \right) \tag{B31} \)
for the model of slip on a plane. That is, the model of slip on a plane would then overestimate the maximum temperature at time \( t \) by approximately
\[ T(0, t) - T(t) = \Delta T_{\text{max}}(H) \frac{\sigma_n - p(t)}{\sigma_n - \rho_0}. \tag{B32} \]

The overestimate becomes negligible when \( p \to \sigma_n \), but can be numerically significant before that. Consider the intact elastic walls, average on \( p-T \) path, case of Table 2, and suppose that melting should be assumed to begin when \( T = 900^\circ \text{C} \). At the 7 km depth considered in Table 2, for the model of slip on a plane, melting would then begin when \( (\sigma_n - p)/(\sigma_n - \rho_0) = 0.425 \), corresponding to slip \( \delta = 1.01 \text{ L}^* \approx 30 \text{ mm}. \) However, for thickness \( h = 0.2 \text{ mm} (H = 0.35), \) melting would instead begin when \( (\sigma_n - p)/(\sigma_n - \rho_0) \approx 0.31, \) which is achieved at slip \( \delta \approx 2.4 \text{ L}^* \approx 71 \text{ mm}, \) over twice as large. For \( h = 1.0 \text{ mm} (H = 1.73), \) it would begin when \( (\sigma_n - p)/(\sigma_n - \rho_0) \approx 0.16, \) achieved at \( \delta \approx 12 \text{ L}^* \approx 360 \text{ mm}. \) These are very rough estimates, but the differences in the estimated slips to the onset of melting are quite significant. The forthcoming work of A Rempel and Rice (submitted manuscript, 2006) should be consulted for more precise corrections of temperature rises for finite, if small, shear zone thicknesses.

[120] Acknowledgments. Support is gratefully acknowledged to NSF grants EAR-0125709 and 0510193 and to the NSF/USGS Southern California Earthquake Center, funded by NSF Cooperative Agreement EAR-0106924 and USGS Cooperative Agreement 02HQAG0008 (this is SCEC contribution 918). Also, short-term support was provided by funding to participate in research programs on granular materials at the Newton Institute in Cambridge, UK, 2003, and the Institute Henri Poincare in Paris, 2005. Thanks for discussions on the background for these studies, and in some cases, collaboration on related studies, are due to Rachel Abercrombie, Michael Ashby, Nick Beeler, Yehuda Ben-Zion, Judi Chester, Massimo Cocco, David Goldsby, Laurent Jacques, Nadia Lapusta, Anaël Lemaître, Alan Rempel, John Rudnicki, Charles Sammis, Paul Segall, Toshi Shimamoto, Jean Sulen, Elisa Tinti, Victor Tsai, Terry Tullis, Ioannis Vardoulakis, Chris Wibberley, and Teng-Fong Wong. Also, Nick Beeler, Judi Chester, David Goldsby, Vikas Prakash, and Terry Tullis are thanked for allowing me to quote their studies in progress; Alan Rempel for major help in choosing parameters and performing the \( p-T \) path averages for Tables 2 and 3; and Alan Rempel, John Rudnicki, Charles Sammis, and Paul Segall for helpful comments on earlier drafts of the text.

References
Anderson, D. L. (1980), An earthquake induced heat mechanism to explain the loss of strength of large rock and earth slides, in Engineering for Protection from Natural Disasters, Proceedings of the International Con-
B05311

RICE: HEATING AND WEAKENING OF FAULTS

B05311

Bizzarri, A., and M. Cocco (2006a), A thermal pressurization model for the...

Biegel, R. L., and C. G. Sammis (2004), Relating fault mechanics to fault zone...

Chester, F. M., and C. G. Sammis (1989), The frictional properties of a simulated...

Di Toro, G., D. L. Goldsby, and T. E. Tullis (2004), Friction falls toward...

Favreau, P., and R. J. Archuleta (2003), Direct seismic energy modelizing...

Fialko, Y. A. (2004), Temperature fields generated by the...
RICE: HEATING AND WEAKENING OF FAULTS


Voigt, B., and C. Faust (1982), Frictional heat and strength loss in some rapid landslides, Geotechnique, 32, 43–54.


