Chapter Number

Toward the Formulation of a Realistic Fault Governing Law in Dynamic Models of Earthquake Ruptures

Andrea Bizzarri
Istituto Nazionale di Geofisica e Vulcanologia – Sezione di Bologna
Italy

1. Introduction

Dynamic earthquake models can help us in the ambitious understanding, from a deterministic point of view, of how a rupture starts to develop and propagates on a fault, how the excited seismic waves travel in the Earth crust and how the high frequency radiation can damage a site on the ground. Since analytical solutions of the fully dynamic, spontaneous rupture problem do not exist (even in homogeneous conditions), realistic and accurate numerical experiments are the only available tool in studying earthquake sources basing on Newtonian Mechanics. Moreover, they are a credible way of generating physics-based ground motions. In turn, this requires the introduction of a fault governing law, which prevents the solutions to be singular and the crack tip and the energy flux to be unbounded near the rupture front.

Contrary to other amibts of Physics, Seismology presently lacks knowledge of the exact physical law which governs natural faults and this is one of the grand challenges for modern seismologists. While for elastic solids it exists an equation of motion which relates particle motion to stresses and forces through the material properties (the scale-free Navier–Cauchy’s equation), for a region undergoing inelastic, brittle deformations this equation is presently missed and scientists have yet to fully decipher the fundamental mechanisms of friction.

The traction evolution occurring during an earthquake rupture depends on several mechanisms, potentially concurrent and competing one with each other. Recent laboratory data and field observations revealed the presence, and sometime the coexistence, of thermally-activated processes (such as thermal pressurization of pore fluids, flash heating of asperity contacts, thermally-induced chemical reactions, melting of rocks and gouge debris), porosity and permeability evolution, elasto-dynamic lubrication, etc.

In this chapter we will analyze, in an unifying and comprehensive sketch, all possible chemico-physical mechanisms that can affect the fault weakening and we will explicitly indicate how they can be incorporated in a realistic governing model. We will also show through numerical simulations that simplified constitutive models that neglect these phenomena appear to be inadequate to describe the details of the stress release and the consequent high frequency seismic wave radiation. In fact, quantitative estimates show that in most cases the incorporation of such nonlinear phenomena has significant effects, often dramatic, on the dynamic rupture propagation, that finally lead to different damages on the free surface.
Given the uncertainties in the relative weight of the various competing processes, the range of variability of the value of some parameters, and the difference in their characteristic time and length scales, we can conclude that the formulation of a realistic governing law still requires multidisciplinary efforts from theoretical models, laboratory experiments and field observations.

2. Dynamic models of earthquake ruptures

2.1 The fault system

A fault can be regarded as the surface, or more properly the volume, where non–elastic processes take place. In Figure 1 we report a sketch illustrating the most widely accepted model of a fault, which is also considered in the present chapter. It is essentially based on the data arising from a large number of field observations and geological evidence (e.g., Chester & Chester, 1998; Sibson, 2003).

![Fig. 1. Sketch representing the fault structure suggested by geological observations.](image)

The slipping zone of thickness $2\omega$ is surrounded by highly fractured damage zone and finally by the undamaged host rocks. The inset panel illustrates the mathematical representation of the fault model adopted in the numerical simulations discussed in the present chapter.

Many recent investigations on the internal structure of fault zones reveal that coseismic slip on mature fault often occurs within an ultracataclastic, gouge–rich and possibly clayey zone (the foliated fault core), generally having a thickness of the order of few centimeters. The
fault core, which typically is parallel to the macroscopic slip vector, is surrounded by a cataclastic damage zone, which can extend up to hundreds of meters. This region is composed of highly fractured, brecciated and possibly granulated materials and it is generally assumed to be fluid–saturated. Outside the damage zone we have the host rock, basically composed of undamaged materials (e.g., Wilson et al., 2003).

Observations tend to suggest that the slip is accommodated along a single, nearly planar surface, the prominent slip surface (pss) — sometime called principal fracture surface (pfs) — which generally has a thickness of the order of millimeters (Rice & Cocco, 2007). When the breakdown process is realized (i.e., the traction is degraded down to its kinetic, or residual, level), the fault structure reaches a mature stage and the slip is concentrated in one (or sometime two) pss, which can be in the middle or near one border of the fault core (symmetric or asymmetric disposition, respectively; see Sibson, 2003). The localization to that narrow slip zone generally takes place at the early stages of the deformation. Moreover, field observations from exhumed faults indicate that fault zones grow in width by continued slip and evolve internally as a consequence of grains size reduction (e.g., Engelder, 1974). As we will see in the following of the chapter, the fault zone width, which is a key parameter for many phenomena described below, is difficult to be quantified, even for a single fault and it exhibits an extreme variation along the strike direction.

2.2 The constitutive law
The second ingredient necessary to solve the elasto–dynamic problem is represented by the introduction of a governing model which ensures a finite energy flux at the rupture tip and describes the traction temporal evolution. As an opposite of a fracture criterion — which is simply a binary condition that specifies whether there is a rupture at a given fault point and time — a governing (or constitutive) law is an analytical relation between the components of stress tensor and some physical observables. Following the Amonton’s Law and the Coulomb–Navier criterion, we can relate the magnitude \( \tau \) of the shear traction vector \( \mathbf{T}^{(n)} \) to the effective normal stress on the fault \( \sigma_{n}^{\text{eff}} \) through the well known relation:

\[
\tau = \left\| \mathbf{T}^{(n)} \right\| = \mu \sigma_{n}^{\text{eff}},
\]

\( \mu \) being the (internal) friction coefficient and

\[
\sigma_{n}^{\text{eff}} = \left\| \mathbf{\Sigma}^{(n)} \right\| = \sigma_{n} - p_{\text{fluid}}^{\text{ef}}
\]

In equation (1) an additional term for the cohesive strength \( C_{0} \) of the contact surface can also appear on the right–hand side. In equation (2) \( \sigma_{n} \) is the normal stress (having tectonic origin) and \( p_{\text{fluid}}^{\text{ef}} \) is the pore fluid pressure on the fault.

Once the boundary conditions (initial conditions, geometrical settings and material properties) are specified, the value of the fault friction \( \tau \) fully controls the metastable rupture nucleation, the further (spontaneous) propagation (accompanied by stress release, seismic wave excitation and stress redistribution in the surrounding medium), the healing of slip and finally the arrest of the rupture (i.e., the termination of the seismogenic phase of the rupture), which precedes the re–strengthening interseismic stage. With the only exception of post–seismic and interseismic phases of the seismic cycle, all the above–mentioned stages of
the rupture process are accounted for in fully dynamic models of an earthquake rupture, provided that the exact analytical form of the fault strength is given. The possibility to explicitly include all the previously-mentioned physical processes that can potentially occur during faulting is a clear requisite of a realistic fault governing law. In the light of this, equation (1) can be rewritten in a more comprehensive form as follows (generalizing equation (3.2) in Bizzarri & Cocco, 2005):

$$\tau = \tau(w_1O_1, w_2O_2, \ldots, w_NO_N)$$

where \(O_i\) are the physical observables, such as cumulative fault slip \((u)\), slip velocity modulus \((v)\), internal variables (such as state variables, \(\Psi\); Ruina, 1983), etc. (see Bizzarri & Cocco, 2005 for further details). Each observable can be associated with its evolution equation, which is coupled to equation (3).

Fig. 2. Scheme of the mechanisms potentially occurring within the coseismic time scale. Each color path represent a distinct phenomenon. Processes occurring in the slipping zone are written in black; processes potentially involving the damage zone are written in purple.

In Figure 2 we present in a unifying sketch all phenomena that can potentially occur during a faulting episode and that can lead to changes to the fault traction. In the following sections we will follow each single color path, which identifies a specific mechanism. It is unequivocal that the relative importance of each single process (represented by the weights \(w_i\) in equation (3)) can change depending on the specific event we consider. Therefore it would be very easily expected that not all independent variables \(O_i\) will appear in the expression of fault friction for all natural faults. Moreover, each phenomenon is associated with its own characteristic length and duration (spatial and temporal characteristic scale) and it is controlled by some parameters, some of whom are sometime
poorly constrained by observational evidences. As we will discuss in the following of the chapter, the difference in the length (and time) scale parameters of each chemico–physical process potentially represents a theoretical and computational complication in the effort to include different mechanisms in the governing law.

2.3 The numerical approach

Unless some explicit, restrictive hypotheses are introduced (e.g., assuming a constant rupture speed, neglecting inertial effects, considering homogeneous condition in the seismogenic region of interest) it is not possible to obtain closed-form analytical solutions of the elasto–dynamic problem. As a consequence, fully dynamic, spontaneous (i.e., with not prior-assigned rupture speed), realistic (i.e., structurally complex) models of earthquakes require the exploit of numerical codes. In some situations free surface topography, anisotropy, non-planar interfaces, spatially variable gradients of velocity, density and quality factors are necessary ingredients for a faithful description of the real–world events.

We can regard computer simulations as a type of experimental approach in the case of conditions that can be not reproduced in laboratory experiments of intact rock fracturing and/or sliding friction on pre-existing surfaces.

The overall requirement for a numerical code is to satisfy the three basic properties: i) the consistency of the discretized (algebraic) equations with respect to the original differential equations, ii) the stability and iii) the convergence of the numerical solution. The goodness of the obtained synthetic solution has to be validated through a systematic comparison against other numerical solutions, obtained independently and with different numerical algorithms (e.g., Bizzarri et al., 2001; Harris et al., 2009). Another essential feature of a numerical code is represented by the computation requests (or the computational efficiency), expressed in terms of memory requirements and CPU time. The latter can be successfully reduced by the utilization of optimized mathematical libraries and parallelization paradigms, such as MPI and OpenMP.

In the literature (see for instance Moczo et al., 2007 for a review) several numerical codes have been used to simulate dynamic earthquake ruptures, some of them belonging to the class of boundary elements approaches (boundary elements (BE), boundary integral equation methods (BIEM)), as well as to the class of domain methods (finite differences (FD), finite elements (FE), spectral elements (SE) and pseudospectral elements (pSE), combined (hybrid) FD and FE).

The results of the numerical experiments presented and discussed in the following of the present chapter have been obtained by using the FD, conventional grid code described in detail in Bizzarri & Cocco (2005). They refer to a strike slip fault, as illustrated in the inset panel of Figure 1. The adopted numerical code — which is under continuous development — is 2nd–order accurate in space and in time, OpenMP-parallelized, it contains various absorbing boundary conditions (in order to minimize spurious numerical pollutions arising from the reflection at the borders of the computational domain) and includes the free surface condition. It also fully manages the time–weakening friction (fault traction is released over a finite time interval), the slip–dependent laws (either linear and non linear), various formulations of the rate– and state–dependent friction laws (including regularizations at low slip velocities). Moreover, it also incorporates the thermal pressurization of pore fluids (see section 3.1), the flash heating of asperity contacts (section 3.2), porosity and permeability variations (sections 4.1 and 4.2). The fault boundary conditions is implemented
Dynamic Modelling

by using the Traction–at-Split–Node (TSN) technique, which has been proved to be one of the most accurate numerical schemes to incorporate the non-elastic response of the fault. Finally, the code can handle multiple faults, in order to simulate stress interaction and fault triggering phenomena (e.g., Bizzarri & Belardinelli, 2008).

3. Thermally activated processes

3.1 Thermal pressurization of pore fluids

The role of fluids and pore pressure relaxation on the mechanics of earthquakes and faulting is the subject of an increasing number of studies, based on a new generation of laboratory experiments, field observations and theoretical models. The interest is motivated by the fact that fluids play an important role in fault mechanics; they can affect the earthquake nucleation and the earthquake occurrence (e.g., Sibson, 1986; Antonioli et al., 2006), they can trigger aftershocks (Nur & Booker, 1972 among many others) and they can control the breakdown process through the so-called thermal pressurization phenomenon (Bizzarri & Cocco, 2006a, 2006b and references therein). Here we will focus on the coseismic time scale, but we want to remark that pore pressure can also change during the interseismic period, due to compaction and sealing of fault zones.

The temperature variations caused by frictional heating,

\[ T^w(\xi, \zeta, \eta, t) = T_0 + \frac{1}{4c} \int_0^{t-c} \left\{ \text{erf}\left(\frac{\zeta - w(z, \eta)}{2\sqrt{\chi(t-t')}}\right) - \text{erf}\left(\frac{\zeta - w(z, \eta)}{2\sqrt{\chi(t-t')}}\right) \right\} d(t-t') \]

(Bizzarri & Cocco, 2006a; \( \chi \) is the thermal diffusivity, \( c \) is the heat capacity of the bulk composite and erf(.) is the error function), heats both the rock matrix and the pore fluids; thermal expansion of fluids is paramount, since thermal expansion coefficient of water is greater than that of rocks. The stiffness of the rock matrix works against fluid expansion, causing its pressurization. Several in situ and laboratory observations show that there is a large contrast in permeability \((k)\) between the slipping zone and the damage zone: in the damage zone \(k\) might be three orders of magnitude greater than that in the fault core (see also Rice, 2006). As a consequence, fluids tend to flow in the direction perpendicular to the fault. Pore pressure changes are associated to temperature variations caused by frictional heating, temporal changes in porosity and fluid transport through the equation:

\[ \frac{\partial}{\partial t} P_{\text{fluid}} = \alpha_{\text{fluid}} \frac{\partial}{\partial t} T - \frac{1}{\beta_{\text{fluid}}} \frac{\partial}{\partial t} \Phi + \frac{\partial^2}{\partial \omega^2} P_{\text{fluid}} \]

where \( \alpha_{\text{fluid}} \) is the volumetric thermal expansion coefficient of the fluid, \( \beta_{\text{fluid}} \) is the coefficient of the compressibility of the fluid and \( \omega \) is the hydraulic diffusivity, expressed as:

\[ \omega \equiv \frac{k}{n_{\text{fluid}} \beta_{\text{fluid}}} \]

\( n_{\text{fluid}} \) being the dynamic fluid viscosity and \( \Phi \) the porosity, which potentially can evolve through the time. The solution of equation (5), coupled with the Fourier’s heat conduction equation, can be analytically determined and assumes the following form:
Toward the Formulation of a Realistic Fault Governing Law in Dynamic Models of Earthquake Ruptures

In previous equations $p_{\text{fluid}0}$ is the initial pore fluid pressure (i.e., $p_{\text{fluid}0} \equiv p_{\text{fluid}}(\xi_1, \zeta, \xi_3, 0)$) and $\gamma \equiv \alpha_{\text{fluid}}/\beta_{\text{fluid}}c$. In (7) the term involving $\Phi$ accounts for compaction or dilatation and it acts in competition with respect to the thermal contribution to the pore fluid pressure changes. Additionally, variations in porosity will modify, at every time instant (see equation (6)), the arguments of error functions which involve the hydraulic diffusivity.

As a consequence of equations (1) and (2), it follows from equation (7) that variations in pore fluid pressure lead to changes in fault friction:

$$
\tau = \left[ \mu_* + a \ln \left( \frac{v}{v_*} \right) + b \ln \left( \frac{\Psi v_*}{L} \right) \right] \sigma^{\text{eff}}_w
$$

$$
\frac{d}{dt} \Psi = - \frac{\Psi v}{L} \ln \left( \frac{\Psi v}{L} \right) - \left( \frac{a_{LD} \Psi}{b \sigma^{\text{eff}}_w} \right) \frac{d}{dt} \sigma^{\text{eff}}_w
$$

(Linker & Dieterich, 1992; $a$, $b$, $L$ and $a_{LD}$ are constitutive parameters and $\mu_*$ and $v_*$ are reference values for friction coefficient and sliding velocity, respectively).

In their fully dynamic, spontaneous, 3-D earthquake model Bizzarri & Spudich (2008) showed that the inclusion of fluid flow in the coseismic process strongly alters the dry behavior of the fault, enhancing instability, even causing rupture acceleration up to super-shear rupture velocities for rheologies which do not allow this transition in dry conditions. For extremely localized slip (i.e., for small values of slipping zone thickness) or for low value of hydraulic diffusivity, the thermal pressurization of pore fluids increases the stress drop, causing a nearly complete stress release (Andrews, 2002; Bizzarri & Cocco, 2006b). It also changes the shape of the slip–weakening curve and therefore the value of the so–called fracture energy. This is important, since fracture energy, defined physically as the amount of energy (for unit fault surface) necessary to maintain an ongoing rupture which propagates on a fault, is recognized to be one of the most important parameter in the context of the physics of the earthquake source. It directly influences the earthquake dynamics, since its value controls the rupture propagation and its arrest and it affects the radiation efficiency.

In Figure 3 we report slip–weakening curves obtained in the case of Dieterich–Ruina law (Linker & Dieterich, 1992) for different values of $2w$ and $\omega$. In some cases (Bizzarri & Cocco, 2006b) it is impossible to determine the equivalent slip–weakening distance (in the sense Cocco & Bizzarri, 2002) and the friction exponentially decreases as recently suggested by several papers (Abercrombie & Rice, 2005; Mizoguchi et al., 2007).
Fig. 3. Traction versus slip curves for wet faults obeying to equation (8). (a) Effect of different slipping zone thickness, $2w$. (b) Effect of different hydraulic diffusivities, $\omega$. In all panels blue line refers to a fault where fluid migration is not allowed (i.e., dry faults where $\sigma_{neff}$ is constant through the time).

3.2 Flash heating of asperity contacts

Another thermically–activated phenomenon is the flash heating (FH thereinafter; Tullis & Goldsby, 2003; Rice, 2006; Bizzarri, 2009) which might be invoked to explain the reduction of the friction coefficient $\mu$ from typical values at low slip rate ($\mu = 0.6$–0.9 for almost all rock types; e.g., Byerlee, 1978) down to $\mu = 0.2$–0.3 at seismic slip rate. It is assumed that the macroscopic fault temperature ($T_{wf}$) changes much more slowly than the temperature on an asperity contact, causing the rate of heat production at the local contact to be higher than average the heating rate of the nominal area. In the model, flash heating is activated if sliding velocity is greater than the critical velocity

$$v_{fh} = \frac{\pi \chi}{D_{ac}} \left( \frac{T_{weak} - T_{w'}}{\tau_{ac}} \right)^2$$

where $\tau_{ac}$ is the local shear strength of an asperity contact (which is far larger than the macroscopic applied stress), $D_{ac}$ is its diameter and $T_{weak}$ (near the melting point) is a weakening temperature at which the contact strength of an asperity begin to decrease. We want to remark that $v_{fh}$ changes in time as macroscopic fault temperature $T_{wf}$ does. When fault slip exceeds $v_{fh}$ the governing equations are (Bizzarri, 2009):

$$\tau = \left[ \mu_s + a \ln \left( \frac{v}{v_s} \right) + \Theta \right] \sigma_{neff}$$

$$\frac{d}{dt} \Theta = \left[ \frac{v}{L} \left[ \Theta + b \frac{v_{fh}}{v} \ln \left( \frac{v}{v'} \right) + \left( 1 - \frac{v_{fh}}{v} \right) \left( a \ln \left( \frac{v}{v'} \right) + \mu_s - \mu_{fh} \right) \right] \right]$$

being $\mu_{fh}$ a reference value for friction coefficient at high slip velocities. For $v < v_{fh}$ the governing equations retain the classical form (Ruina, 1983):
We note that thermal pressurization of pore fluids and flash heating are inherently different mechanisms because they have a different length scale: the former is characterized by a length scale of the order of few micron ($D_m$), while the length scale of the latter phenomenon is the thermal boundary layer ($\delta = (2\chi t_{\text{pulse}})^{1/2}$, where $t_{\text{pulse}}$ is the duration of slip, of the order of seconds), which is ~ mm up to few cm. Moreover, while thermal pressurization affects the effective normal stress, flash heating causes changes only in the analytical expression of the friction coefficient at high slip rates. In both cases the evolution equation for the state variable is modified: by the coupling of variations is $\sigma_{\text{eff}}$ for the first phenomenon, by the presence of additional terms involving $v_{\text{fh}} / v$ in the latter.

Fig. 4. Temperature change (computed from equation (4)) as a function of cumulative fault slip. The inset shows the time evolution of temperature change. Dashed lines refer to models without FH.

Numerical results of Bizzarri (2009) demonstrate that, compared to classical models where FH is neglected, the inclusion of FH considerably increases the degree of instability of the fault; the supershear rupture regime is highly favored, peaks in slip velocity are greater (nearly 50 times), as well as the stress drop (more than 3 times). Moreover, the fault traction exhibits larger weakening distances, leading to a greater (nearly 4 times) fracture energies. It is also found that for highly localized shear ($2w \leq 10$ mm for $a$ in between 0.016 and 0.018) the modification of the governing law due to FH causes a fast re-strengthening, leading to a self-healing of slip. In self-healing cases, the strength recovery for increasing slip and slip velocity is quite similar to that previously obtained by Tinti et al. (2005) and it is such that the final traction is in a steady state and therefore it is not sufficient to originate a further failure event in absence of external loading. Finally, Bizzarri (2009) indicates that the FH increases the propensity of the fault to melt earlier and can not prevent it from occurring (see Figure 4): the decrease of the sliding resistance is counterbalanced by enhanced slip velocities (recall equation (4)). This results leaves us with the mystery of why actual evidence of melting is so rare.
3.3 Melting of gouge and rocks
As first pointed out by Jeffreys (1942), melting should probably occur during coseismic slip, typically after rocks comminution. Rare field evidence for melting on exhumed faults (i.e., the apparent scarcity of glass or pseudotachylytes, natural solidified friction melts produced during coseismic slip) generates scepticism for the relevance of melt in earthquake faulting. However, several laboratory experiments have produced melt, when typical conditions of seismic deformation are attained (Spray, 1995; Tsutsumi & Shimamoto, 1997). Moreover, as previously mentioned in section 3.1, it has been demonstrated by theoretical models that for thin slipping zones (i.e., \(2\omega/\delta < 1\)) melting temperature \(T_m\) can be easily exceeded in dynamic motion (Bizzarri & Cocco, 2006a, 2006b). Even if performed at low (2–3 MPa) normal stresses, the experiments of Tsutsumi & Shimamoto (1997) demonstrated significant deviations from the predictions obtained with the usual rate- and state-friction laws (e.g., Ruina, 1983). Fialko & Khazan (2005) suggested that fault friction simply follows the Coulomb–Navier equation (1) before melting and the Navier–Stokes constitutive relation, \(\tau = \eta_{\text{melt}} v / (2\omega_{\text{melt}})\), after melting (\(2\omega_{\text{melt}}\) being the thickness of the melt layer). Nielsen et al. (2008) theoretically interpreted the results from high velocity (\(v > 0.1 \text{ m/s}\)) rotary friction experiments and derived the following relation expressing the fault traction in steady state conditions when melting occurs:

\[
\tau = \sigma_n^{\text{eff}} \frac{K_{\text{NEA}}}{R_{\text{NEA}}} \left[ \ln \left( \frac{2v}{v_m} \right) \right] \left( \frac{2v}{v_m} \right)
\]

where \(K_{\text{NEA}}\) is a dimensional normalizing factor, \(R_{\text{NEA}}\) is the radius of the sample and \(v_m\) is a characteristic slip rate.

3.4 Chemical environment changes
It is known that fault friction can be influenced also by chemical environment changes. Chemical analyses of gouge particles formed in high velocity laboratory experiments by Hirose & Bystricky (2007) showed that dehydration reactions (i.e., the release of structural water in serpentine) can take place. Moreover, recent experiments on Carrara marble performed by Han et al. (2007) showed that thermally activated decomposition of calcite (into lime and CO\(_2\) gas) occurs from a very early stage of slip, in the same temporal scale as the ongoing and enhanced fault weakening. Thermal decomposition weakening may be a widespread chemico-physical process, since natural gouges commonly are known to contain sheet silicate minerals. The latter can decompose, even at lower temperatures than that for calcite decomposition, and can leave geological signatures of seismic slip (Han et al., 2007), different from pseudotachylytes. Presently, there are no earthquake models where chemical effects are incorporated within a governing equation. We believe that some efforts will be spent to this goal in the next future.

4. The importance of porosity and permeability
4.1 Temporal evolution of porosity
The values of permeability \((k)\), porosity \((\Phi)\) and hydraulic diffusivity \((\omega)\) play a fundamental role in controlling the fluid migration and the breakdown processes on a seismogenic wet
fault. During an earthquake event the frictional sliding tends to open (or dilate) cracks and pore spaces (leading to a decrease in pore fluid pressure), while normal traction tends to close (or compact) cracks (therefore leading to a pore fluid pressure increase). Stress readjustment on the fault can also switch from ineffective porosity (i.e., closed, or non-connected, pores) to effective porosity (i.e., catenary pores), or vice versa. Both ductile compaction and frictional dilatancy cause changes to $k$, $\Phi$ and therefore to $\omega$. It is clear from equation (11) that this leads to variations to $p_{\text{fluid}}$.

Starting from the theory of ductile compaction of McKenzie (1984) and assuming that the production rate of the failure cracks is proportional to the frictional strain rate and combining the effects of the ductile compaction, Sleep (1997) introduced the following evolution equation for the porosity:

$$
\frac{d}{dt}\phi = \frac{v \beta_p \mu_*}{2w} - \frac{\sigma_{\text{eff}}^*}{C_\eta (\phi_{\text{sat}} - \phi)^m}
$$

(13)

where $\beta_p$ is a dimensionless factor, $C_\eta$ is a viscosity parameter with proper dimensions, $n$ is the creep power law exponent and $m$ is an exponent that includes effects of nonlinear rheology and percolation theory. Equation (13) implies that porosity can’t exceeds a saturation value $\phi_{\text{sat}}$.

As noticed by Sleep & Blanpied (1992), frictional dilatancy is associated also with the formation of new voids, as well as with the intact rock fracturing (i.e., with the formation of new tensile micro-cracks). In fact, it is widely accepted that earthquakes result in a complex mixture of frictional slip processes on pre-existing fault surfaces and shear fracture of initially intact rocks. This fracturing will cause a change in porosity; fluid within the fault zone drains into these created new open voids and consequently decreases the fluid pressure. The evolution law for the porosity associated with the new voids is (Sleep, 1995):

$$
\frac{d}{dt}\phi = \frac{v \beta_{ov} \mu_*}{2w}
$$

(14)

where the factor $\beta_{ov}$ is the fraction of energy that creates the new open voids.

Sleep (1997) also proposed the following relation that links the increase of porosity to the displacement, which leads to an evolution law for porosity:

$$
\frac{d}{dt}\Phi = \frac{v \Phi_{\text{fluid}} \tau}{2w c}
$$

(15)

Finally, Segall & Rice (1995) proposed two alternative relations for the evolution of $\Phi$. The first mimics the evolution law for state variable in the Dieterich–Ruina model (Beeler et al., 1994 and references therein):

$$
\frac{d}{dt}\Phi(\xi_1, \xi_2, \xi_3, t) = -\frac{v}{L_{SR}}\Phi(\xi_1, \xi_2, \xi_3, t) - \varepsilon_{SR} \ln\left(\frac{c_1 v + c_2}{c_3 v + 1}\right)
$$

(16)

where $\varepsilon_{SR}$ and $L_{SR}$ are two parameters representing the sensitivity to the state variable evolution (in the framework of rate- and state-dependent friction laws) and a characteristic
length-scale, respectively, and \( \{c_i\}_{i=1,2,3} \) are constants ensuring that \( \Phi \) is in the range \([0,1]\). In principle, \( \varepsilon_{SR} \) can decrease with increasing effective normal stress, but at the present state we do not have detailed information about this second-order effect.

The second model, following Sleep (1995), postulates that \( \Phi \) is an explicit function on the state variable \( \Psi \):

\[
\Phi(\xi, \zeta, \zeta, t) = \Phi_0 - \varepsilon_{SR}\ln\left(\frac{\Psi}{L_{SR}}\right)
\]

\( \Phi_0 \) being a reference value, assumed to be homogeneous over the whole slipping zone thickness.

\[
\text{Exponential decay}
\]

Charateristic slip–weakening distance

Fig. 5. Comparison between solutions of the thermal pressurization problem in case of constant (black curve) and variable porosity (as in equation (17); gray curve). (a) Traction vs. slip curve. (b) Traction evolution of the effective normal stress.

Considering the latter equation, coupled with \( (7) \), and assuming as Segall & Rice (1995) that the scale lengths for the evolution of porosity and state variable are the same, we have that, even if the rupture shape, the dynamic stress drop and the final value of \( \sigma_{neff} \) remain unchanged with respect to a corresponding simulated event in which a constant porosity was assumed, the weakening rate is not constant for increasing cumulative slip. Moreover, the equivalent slip–weakening distance becomes meaningless. This is clearly visible in Figure 5, where we compare the solutions of the thermal pressurization problem in cases of constant (black curve) and variable porosity (grey curve).

All the equations presented in this section clearly state that porosity evolution is concurrent with the breakdown processes, since it follows the evolution of principal variables involved in the problem \( (v, \tau, \sigma_{neff}, \Psi) \). However, in spite of the above-mentioned profusion of analytical relations, porosity is one of the biggest unknowns in the fault structure and presently available evidence from laboratory, and from geological observations as well, do not allow us to discriminate between different possibilities. Only numerical experiments performed by coupling one of the equations \((13)\) to \((17)\) with \((7)\) can show the effects of different assumption and suggest what is the most appropriate. Quantitative results will plausibly give some useful indications for the design of new laboratory experiments.
4.2 Permeability changes
As mentioned above, changes in hydraulic diffusivity can be due not only to the time evolution of porosity, but also to variations of permeability. $k$ is known to suffer large variations with type of rocks and their thermo–dynamical state (see for instance Turcotte & Schubert, 1982) and moreover local variations of $k$ have been inferred near the fault. Several laboratory results (e.g., Brace et al., 1968) supported the idea that $k$ is an explicit function of $\sigma_{\text{eff}}$. A reasonable relation (Rice, 1992) is:

$$k = k_0 e^{\frac{\sigma_{\text{eff}}}{\sigma_c}}$$  \hspace{1cm} (18)

where $k_0$ is the permeability at zero effective normal stress and $\sigma_c$ is a constant. For typical changes in $\sigma_{\text{eff}}$ expected during coseismic ruptures we can guess an increase in $k$ at least of a factor 2 within the temporal scale of the dynamic rupture. In principle, this can counterbalance the enhancement of instability due to the fluid migration out of the fault. This is particularly encouraging because seismological estimates of the stress release (almost ranging from about 1 to 10 MPa; e.g., Aki, 1972) do not support the evidence of a nearly complete stress drop, as predicted by numerical experiments of thermal pressurization.

Another complication may arise from the explicit dependence of permeability on porosity and on grain size $d$. Following one of the most widely accepted relation, the Kozeny–Carman equation (Kozeny, 1927), we have:

$$k = K_{KC} \frac{\phi^3}{(1 - \phi)^2} d^2$$  \hspace{1cm} (19)

Previous equation therefore states that gouge particle refinement and temporal changes in $\phi$, such as that described in equations (13) to (17), affect the value of $k$.

As in the case of porosity evolution, permeability changes also occur during coseismic fault traction evolution and consequently equations (18) or (19) can be easily incorporated in the thermal pressurization model (i.e., coupled with equation (7)).

5. Elasto–dynamic lubrication
Another important effect of the presence of pore fluids within the fault structure is represented by the mechanical lubrication (Sommerfeld, 1950; Ma et al., 2003). In the model of Brodsky & Kanamori (2001) an incompressible fluid obeying the Navier–Stokes equations flows around the asperity contacts of the fault, without leakage, in the direction perpendicular to the fault surface. In absence of elastic deformations of the rough surfaces, the fluid pressure in the lubrication model is:

$$p_{\text{fluid}}^{\text{(lub)}}(\xi_1) = p_{\text{res}} + \frac{3}{2} \eta_{\text{fluid}} V \int_0^{\xi_1} \frac{w^* - w(\xi_1^*)}{(w(\xi_1^*))^3} d\xi_1^*$$  \hspace{1cm} (20)

where $p_{\text{res}}$ is the initial reservoir pressure (which can be identified with quantity $p_{\text{fluid}^{\text{ref}}}$ of equation (7)), $V$ is the relative velocity between the fault walls (2$v$ in our notation), $w^* \equiv w(\xi_1^*)$, where $\xi_1^*$ is such that $(dp_{\text{fluid}}^{\text{(lub)}}/d\xi_1)|_{\xi_1=\xi_1^*} = 0$, and $\xi_1$ maps the length of the lubricated zone $L^{\text{(lub)}}$. Qualitatively, $L^{\text{(lub)}}$ is equal to the total cumulative fault slip $u_{\text{tot}}$. 
Interestingly, simple algebra shows that if the slipping zone thickness is constant along the strike direction also the lubrication pore fluid pressure is always equal to $p_{\text{res}}$. The net result of the lubrication process is that the pore fluid pressure is reduced by an amount equal to the last member of equation (20). This in turn can be estimated as

$$P^{(\text{lab})} \approx 12\eta_{\text{fluid}}^2 \frac{ru_{\text{tot}}}{(<2w>)^3}$$

where $r$ is the aspect ratio constant for roughness and $<2w>$ is the average slipping zone thickness. Therefore equation (2) is then rewritten as:

$$\sigma_{\text{eff}} = \sigma_n - p_{\text{fluid}} - P^{(\text{lab})}.$$ (22)

The fluid pressure can also adjust the fault surface geometry, since

$$2w(\xi_1) = 2w_0(\xi_1) + u^{(\text{lab})}(\xi_1),$$ (23)

where $2w_0$ is the initial slipping zone profile and $u^{(\text{lab})}$ is elasto-static displacement caused by lubrication. Equation (23) can be approximated as

$$<2w> = <2w_0> + \frac{P^{(\text{lab})}L}{E}.$$ (24)

$E$ being the Young’s modulus. $u^{(\text{lab})}$ is significant if $L^{(\text{lab})}$ (or $u_{\text{tot}}$) is greater than a critical length, defined as (see also Ma et al., 2003):

$$L^{(\text{lab})} = 2 <2w_0> \left(\frac{<2w_0>E}{12\eta_{\text{fluid}} vr}\right)^\frac{1}{3};$$ (25)

otherwise the slipping zone thickness does not widen. If $u_{\text{tot}} > L^{(\text{lab})}$ then $P^{(\text{lab})}$ is the positive real root of the following equation

$$P^{(\text{lab})}\left(<2w_0> + \frac{P^{(\text{lab})}u_{\text{tot}}}{E}\right)^3 - 12\eta_{\text{fluid}} vr u_{\text{tot}}^2 = 0.$$ (26)

It is clear from equation (22) that lubrication contributes to reduce the fault traction (and therefore tends to increase the fault slip velocity, which in turn further increases $P^{(\text{lab})}$, as stated in equation (21)). Moreover, if the lubrication increases the slipping zone thickness, then it will reduce asperity collisions and the contact area between the asperities (which in turn will tend to decrease $P^{(\text{lab})}$, as still expressed by equation (21)).

In many papers it has been generally assumed that when effective normal stress vanishes then material interpenetration and/or tensile (i.e., mode I) cracks (Yamashita, 2000) develop, leading to the superposition during an earthquake event of all three basic modes of fracture mechanics (Atkinson, 1987; Petit & Barquins, 1988). An alternative mechanism that can occur when $\sigma_{\text{eff}}$ falls to zero, if fluids are present in the fault zone, is that the frictional stress of contacting asperities described by the Amonton’s Law (1) becomes negligible with respect to the viscous resistance of the fluid and the friction can be therefore expressed as
$$\tau = \frac{<2w>}{u_{tot}} p^{(lub)} \quad (27)$$

which describes the fault friction in the hydrodynamic regime. Depending on the values of total cumulative fault slip and fault slip velocity, in equation (27) $p^{(lub)}$ is alternatively expressed by (21) or by the solution of (26). For typical conditions ($<2w> = 1 \text{ mm}$, $E = 5 \times 10^4 \text{ Pa}$, $v = 1 \text{ m/s}$, $u_{tot} = 2 \text{ m}$, $r = 10 \times 10^{-3} \text{ m}$), if the lubricant fluid is water ($\eta_{\text{fluid}} = 1 \times 10^{-3} \text{ Pa s}$), then $u_{tot} < L^{(lub)}$ and (from equation (21)) $p^{(lub)} \approx 4.8 \times 10^4 \text{ Pa}$. Therefore the lubrication process is negligible in this case and the net effects of the fluid presence within the fault structure will result in thermal pressurization only. On the contrary, if the lubricant fluid is a slurry formed form the mixture of water and refined gouge ($\eta_{\text{fluid}} = 10 \text{ Pa s}$), then $u_{tot} > L^{(lub)}$ and (from equation (26)) $p^{(lub)} \approx 34.9 \text{ MPa}$, which can be a significant fraction of tectonic loading $\sigma_n$. In this case hydro–dynamical lubrication can coexist with thermal pressurization; in a first stage of the rupture, characterized by the presence of ample aqueous fluids, fluids can be squeezed out of the slipping zone due to thermal effects. In a next stage of the rupture, the gouge, rich of particles, can form the slurry with the remaining water; at this moment thermal pressurization is not possible but lubrication effects will become paramount. This is an example of how two different physical mechanisms can be incorporated in a single frictional model.

6. Bi–material Interfaces

Traditional and pioneering earthquake models (see for instance Brace & Byerlee, 1966) simply account for the reduction of the frictional coefficient from its static value to the kinetic frictional level, taking the effective normal stress constant over the duration of the process. Subsequently, Weertman (1980) suggested that a reduction in $\sigma_n$ during slip between dissimilar materials can influence the dynamic fault weakening. Considering an asperity failure occurring on a bi–material, planar interface separating two uniform, isotropic, elastic half–spaces, Harris & Day (1997) analytically demonstrated that $\sigma_n$ can change in time. On the other hand, a material property contrast is not a rare phenomenon in natural faults: Li et al. (1990) and Li & Vidale (1996) identified some strike–slip faults where one side is embedded in a narrow, fault parallel, low–velocity zone (having width of a few hundred of meters). At the same time several authors (Lees, 1990; Michael & Eberhart–Phillips, 1991) inferred the occurrence of significant velocity contrasts across faults, generally less than 30%.

Even if Renardy (1992) theoretically demonstrated that Coulomb frictional sliding is unstable if occurs between materials with different properties, there is not a general consensus about the importance of the presence of bi–material interface on natural earthquakes (Ben–Zion, 2006 versus Andrews & Harris, 2005). More recently, Dunham & Rice (2008), showed that spatially inhomogeneous slip between dissimilar materials alters $\sigma_{\text{eff}}$ (with the relevant scale over which poro–elast–ic properties are to be measured being of order the hydraulic diffusion length, which for large earthquakes is mm to cm). Moreover, it is known that the contrast in poro–elastic properties (e.g., permeability) across faults can alter both $\sigma_n$ and $p_{\text{fluid}}$ (while the elastic mismatch influences only $\sigma_n$).
7. Characteristic lengths and scale separation

It has been previously mentioned that each nonlinear dissipation process that can potentially act during an earthquake rupture has its own distance and time scales, that can be very different from one phenomenon to another. The difference in scale lengths, as well as the problem of the scale separation, can represent a limitation in the attempt to simultaneously incorporate all the mechanisms described in this chapter in a single constitutive model.

We have previously seen that thermal pressurization (section 3.1) can coexist with mechanical lubrication (section 5) as well as with porosity (section 4.1) and permeability evolutions (section 4.2). The same holds for flash heating and thermal pressurization. This simultaneous incorporation ultimately leads to numerical problems, often severe, caused by the need to properly resolve the characteristic distances and times of each single process. The concurrent increase in computational power and the development of new numerical algorithms can definitively assist us in this effort.

In Table 1 we report a synoptic view of the characteristic length scales for the processes described in the present chapter. Two important lengths (see Bizzarri et al., 2001) involved in the breakdown process, are the breakdown zone length (or size, $X_b$) and the breakdown zone time (or duration, $T_b$). They quantify the spatial extension, and the time duration, of the cohesive zone; in other words they express the amount of cumulative fault slip, and the elapsed time, required to the friction to drop, in some (complicated) way, from the yield stress down to the residual level.

<table>
<thead>
<tr>
<th>Process</th>
<th>Characteristic distance</th>
<th>Typical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macrosopic decrease of fault traction from yield</td>
<td>$d_0$</td>
<td>~ few mm in the lab</td>
</tr>
<tr>
<td>traction to residual level</td>
<td>$X_b$</td>
<td>~ 100 of m</td>
</tr>
<tr>
<td>Temporal evolution of the state variable in the</td>
<td>$L$</td>
<td>~ few μm in the lab</td>
</tr>
<tr>
<td>framework of the rate- and state-dependent friction laws</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thermal pressurization (section 3.1)</td>
<td>$2\omega$</td>
<td>≤ 1 cm</td>
</tr>
<tr>
<td></td>
<td>$\delta = (2\chi t_{pulse})^{1/2}$</td>
<td>~ few cm</td>
</tr>
<tr>
<td>Flash heating (section 3.2)</td>
<td>$D_{ac}$</td>
<td>~ few μm</td>
</tr>
<tr>
<td>Gouge and rocks melting (section 3.3)</td>
<td>$2\omega_{melt}$</td>
<td>~ 100 of μm in the lab</td>
</tr>
<tr>
<td>Porosity evolution (section 4.1):</td>
<td>$2\omega$</td>
<td>≤ 1 cm</td>
</tr>
<tr>
<td>- equations (13) to (15)</td>
<td>$L_{SR}$</td>
<td>assumed to be equal to $L$</td>
</tr>
<tr>
<td>- equations (16) and (17)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Synoptic view of the characteristic lengths of the processes described in the chapter.
Another open problem is related to the difficulty to move from the scale of the laboratory (where samples are of the order of several meters) up to the scale of real faults (typically several kilometer long). A large number of the phenomena described above have been measured in the lab: this raises the problem of how to scale the values of the parameters of the inferred equations to natural faults.

It is apparent that both geological observations and improvements in laboratory machines are necessary elements in the understanding of earthquake source physics and in the capability to reproduce it numerically.

8. Summary and conclusions

The dynamic modelling of an earthquake rupture on a fault surface is extremely challenging not only from a merely numerical point of view, but also because of the lack of knowledge of the state of the Earth crust and of the law which describes the earthquake physics.

In this chapter we have described a large number of physical mechanisms that can potentially take place during a faulting episode. These phenomena are macroscopic, in that the fundamental variables (i.e., the physical observables) describing them have to be regarded as macroscopic averages (see also Cocco et al., 2006) of the solid–solid contacts properties. As a result, the fault friction, expressed analytically in terms of a governing law, does not formally describe the stress acting on each single asperity, but the macroscopic average of the stress acting within the slipping zone (see Figure 1). Unlikely, a link between the microphysics of materials, described in terms of lattice or atomic properties, and the macrophysical description of friction, obtained from stick–slip laboratory experiments, is actually missed. On the other hand, we can not expect to be able to mathematically describe (either deterministically or statistically) the evolution of all the surface asperities and of all micro-cracks in the damage zone.

Recent laboratory experiments and geological investigations have clearly shown that different dissipative processes can lead to the same steady state value of friction. In the simple approximation which considers only one single event on an isolated fault, some authors claim that the slip dependence is paramount (Ohnaka, 2003). On the other hand, the explicit dependence of friction on sliding velocity (Dieterich, 1986) is unquestionable, even at high slip rates (Tsutsumi & Shimamoto, 1997). In fact, in the literature there is a large debate (see for instance Bizzarri & Cocco, 2006) concerning the most important dependence of fault friction. Actually, the problem of what is (are) the dominant physical mechanism(s) controlling the friction evolution (i.e., the quantitative estimate of the weights $w_i$ in equation (3)) is still unsolved. Given this fact, we have to regard Figure 2 as a schematic representation of the logical links existing between the different phenomena. It is clear that in a realistic situation only a few colour paths will survive; the scope of that diagram is to emphasize the degree of complexity of the rupture process, which contains more ingredients than the so-called first-order observables (such as slip, slip velocity and state variable(s)).

We believe that seismic data presently available are not sufficient to clarify what specific mechanism is operating (or dominant) during a specific earthquake event. The inferred traction evolution on the fault, as retrieved from seismological records (e.g., Ide & Takeo, 1997), gives us only some general information about the average weakening process on an idealized mathematical fault plane. Moreover, it is affected by the unequivocal choice of the source time function adopted in kinematic inversions and by the frequency band limitation in data recording and sometime could be inconsistent with dynamic ruptures. On the other
hand, we have seen in previous sections that we do not have any physical basis to neglect \textit{a priori} the insertion of additional physical and chemical mechanisms in the analytical expression of a fault governing equation. The first reason is that, compared to results obtained by adopting a simplified (or in some sense idealized) constitutive relation, numerical experiments from models where additional physical mechanisms are accounted for show a significant, often dramatic, change in the dynamic stress drop (and therefore in the resulting ground motions), in the distance over which it is realized, in the so-called fracture energy and in the total scalar seismic moment. The second reason is that, as we have shown (recall the effects of gouge and rocks melting and those of hydro–dynamic lubrication), the inclusion of different mechanisms in some case requires a modification of its classical analytical expression.

As a future perspective, it would be intriguing try to compare synthetics obtained by assuming that one particular physical mechanism is paramount with respect to the others, in order to look for some possible characteristic signatures and specific features in the solutions. The next step would eventually be try to envisage such features in real seismological data.

The above–mentioned approaches are not mutually exclusive and the contributes from each field can lead to the answer of the following key questions: 1) what are the predictions arising from different mathematical and physical descriptions of rupture dynamics that can be observed in the real world?, and 2) what can data illuminate us about earthquake faulting?

In the present chapter we have underlined that some different, nonlinear, chemico–physical processes can potentially cooperate, interact, or even compete one with each other. We have also seen that in most cases we are able to write equations describing them and we have explicitly indicated how they can be incorporated into a fault constitutive model. It is clear that in order to reproduce quantitatively the complexity of the inelastic and dissipative mechanisms occurring on a fault during a failure event a “classical” constitutive relation appears to be nowadays inadequate. To conclude, we are inclined to think that only a multidisciplinary approach to source mechanics, which systematically combines results from accurate theoretical models, advanced laboratory experiments, field observations and data analyses, can hopefully lead in the future to the formulation of a realistic and consistent governing model for real earthquakes. This is an ambitious task of great urgency, and it has to be pursued in the next future.

9. Acknowledgements

The author would like to thank Emily Brodsky, Renata Dmowska, Eric Dunham, Yann Klinger, Chris Marone, Jim Rice and Paul Spudich for fruitful discussions.

10. References

Aki, K. (1972), Earthquake mechanism, \textit{Tectonophys.}, 13, pp. 423–446


Bizzarri, A., & M. Cocco (2006c), Comment on “Earthquake cycles and physical modeling of the process leading up to a large earthquake”, *Earth Planets Space*, Vol. 58, pp. 1525–1528


Toward the Formulation of a Realistic Fault Governing Law in Dynamic Models of Earthquake Ruptures


