Scale dependence in the dynamics of earthquake propagation: evidence from seismological and geological observations

Massimo Cocco and Elisa Tinti

Istituto Nazionale di Geofisica e Vulcanologia, Rome, Italy

Accepted by EPSL on the 11th June 2008
Abstract

We attempt to reconcile current understanding of the earthquake energy balance with recent estimates of fracture energy from seismological investigations and surface energy from geological observations. The complex structure of real fault zones suggests that earthquakes in such fault structures are dominated by scale dependent processes. We present a model for an inelastic fault zone of finite thickness embedded in an elastic crust represented at a macroscopic scale by a mathematical plane of zero thickness. The constitutive properties of the fault zone are governed by physical processes controlling gouge and damage evolution at meso- and micro-scale. However, in order to model and interpret seismological observations, we represent dynamic fault weakening at the macroscopic scale in terms of traction evolution as a function of slip and other internal variables defining a phenomenological friction or contact law on the virtual mathematical plane. This contact law is designed to capture the main features of dynamic fault weakening during earthquake rupture. In this study we assume that total shear traction is friction and corresponds to shear resistance of the whole fault zone. We show that seismological observations, depending on finite and limited wavelength and frequency bandwidth, can only provide an estimate of breakdown stress drop and breakdown work (a more general definition of seismological fracture energy) representing a lower bound of the total intrinsic power of dissipation on the fault zone. We emphasize that geological estimates of surface energy can be compared with seismological estimates of breakdown work only if they are representative of the same macroscopic scale. In this case, it emerges that, contrary to surface energy, seismological breakdown work represents a non-negligible contribution to the earthquake energy budget.
1. Introduction

For the first time original investigations have provided measurements of fracture energy from seismological data over a wide range of earthquake sizes [1-2] and surface energy in real fault zones from geological analyses [3, 4, 5]. Early lab observations [6] indicated that fracture and surface energies were a small part of the total energy budget during fracture, and past field observations of fracture in faults suggested that the same was true for earthquakes (e.g. [7,8]). The most recent observations challenge these perceptions. Indeed, although few authors have considered fracture energy identical to surface energy [5, 9], it is now clear that surface energy is only a small fraction of the mechanical work absorbed on the fault [2, 3, 4, 10] while fracture energy (or breakdown work as it is named by [2, 10]) does not represent a negligible contribution to the earthquake energy budget. This emerges from the comparison between these recent estimates of surface and fracture energies discussed by [2,10,11], who related them to the mechanical work absorbed on the fault plane during earthquake rupture. Moreover, several recent papers ([1, 2, 10, 12) pointed out that the seismological fracture energy (named in this study $G'$) differs from the definition in classic fracture mechanical models. This implies the necessity of providing a physical interpretation to the seismological fracture energy. We believe that a comparison between distinct energy terms should be performed in a consistent physical framework, being aware that these new observations require re-examination of the earthquake energy balance. One of the primary goals of this paper is to reconcile a theoretical understanding of the earthquake energy balance with current geologic understanding of fault zone structure and seismological measurements of fracture energy.

A common feature of any description of the mechanics of dynamic shear rupture propagation is that unstable failure is associated with dynamic fault weakening represented by the traction evolution with time or slip [11]. Traction evolution during dynamics in real fault zones can be very complex. However, it is expected that in order to have a finite stress drop and to radiate seismic waves the traction should decrease from an upper yield stress to a residual stress level usually identified with the sliding frictional level (see Fig. 1a) in a characteristic time (named the breakdown time, $T_b$) and on a spatial dimension, which is referred to as the breakdown zone, $X_b$. The condition that both $T_b$ and $X_b$ are finite (i.e., not negligible compared to any other length scale parameter involved in the process) has important implications for the physical interpretation of dynamic fault weakening (at least, it prevents use of the simplistic Orowan model for stress evolution, see Fig.2 in [12]).

Seismologists use tractions to model dynamic rupture processes. Dynamic modeling of spontaneous earthquake rupture requires the use of constitutive laws, which can prescribe the traction evolution (as in the slip weakening law, [13-15]) or can obtain it as a result of the simulations (as in the rate and state constitutive formulation, [16-18]). Fracture energy ($G$) is
commonly associated with the area below the shear traction curve and above the residual stress level (see Fig. 1a), and it is measured through the following relation:

\[ G = \int_0^{D_c} \left[ \tau(\Delta u') - \tau_{\text{res}} \right] d\Delta u' = \int_0^{\Delta u_1} \tau(\Delta u') d\Delta u' - \tau_{\text{res}} \Delta u_1, \]  

where \( D_c \) is the characteristic slip weakening distance, \( \Delta u \) is the slip, \( \tau \) is the dynamic shear traction, \( \Delta u_1 \) is the final slip and \( \tau_{\text{res}} \) is the residual shear stress. The equality on the right side holds when \( \tau_{\text{res}} \) is assumed to be independent of slip during sliding, as in the classic slip weakening model shown in Fig. 1a. In other words, the concept of \( G \) relies on the existence of a non-history dependent stress reference state. This is a useful but simplistic assumption, which might not be tenable for real fault zones.

Although the model depicted in Fig.1a is widely adopted as a proxy for dynamic fault weakening at a specific position on the rupture surface, as well as the global earthquake energy budget for the whole source, several recent studies [12] have suggested that total dynamic traction decreases non-linearly with time and slip. According to these models, the slip weakening distance and the constant residual stress are not constitutive parameters, but rather might be ill-defined. Fig. 1b shows a sketch of the model proposed by [12] according to which the critical slip weakening distance and the residual stress level (\( \tau_{\text{res}} \)) correspond to the final slip (\( \Delta u_1 \)) and the final stress, respectively (\( \tau_{\text{res}}=\tau_1 \) assuming that no overshoot or undershoot occurs). For both cases shown in Fig.1 it is clear that \( G \) accounts for all the dissipation in excess of \( \tau_{\text{res}} \cdot \Delta u_1 \) (where \( \tau_{\text{res}} \) is equal to the stress at the last slip increment). This latter term is sometime named the frictional energy, but that definition is model dependent and it should be considered just an unknown fraction of the total energy dissipated in a specific fault position. The existence of different descriptions of dynamic fault weakening highlights the need to reconcile understanding of the mechanical work absorbed on the fault plane during earthquakes and to comprehend its partitioning between surface energy and other dissipative mechanisms.

From a seismological point of view it is required that a portion of the mechanical work absorbed on the fault plane must be the energy that has to sustain the dynamic earthquake rupture propagation. This has been identified in numerous theoretical models with the fracture energy \( G \) (see Fig.1), which is usually considered as the crack driving force. According to [23], the crack driving force corresponds to the energy surplus made available for unit area of crack advance and it is defined as the excess of work over a residual stress (which is assumed to be independent of slip consistently with both the models shown in Fig.1). The concept of crack driving force is commonly associated with the energy release rate (the total energy per unit area of crack advance), which is unambiguously defined for crack models having a stress singularity at the propagating tip or for non-singular cohesive rupture models characterized by an infinitesimal breakdown zone size [19, 20, 21]. In fact, Equation (1) is obtained in classic
fracture mechanical models from the application of the path-independent J-integral approach [22] to a slip weakening model under the assumption of a small breakdown zone ($X_b$) [21]. In other words, in order to maintain crack propagation it is commonly assumed that the energy release rate must exceed the energy dissipation of crack advance. However, fracture energy in seismology ($G'$) differs from that commonly used in mechanics ($G$) [10, 12], because $G'$ includes an unknown partitioning between surface energy and dissipation (heat) [2]. For these reasons the association of fracture energy with the energy release rate in realistic fault zones requires further analysis.

Another key issue in the discussion of the earthquake energy budget is that of connecting geological and seismological observations. In order to compare the energy terms contributing to the earthquake energy balance we have to properly consider the length scale at which fracture and surface energies are measured. Geological observations on surface earthquake ruptures and exhumed faults have shed light on the inner structure of fault zones [4, 24, 25]. A common assumption is that slip is localized on a principal slip zone embedded in an ultracataclastic fault core surrounded by a wider damage (highly fractured) zone. This implies that faults have a finite thickness and that coseismic, dynamic slip occurs within a highly deformed and fractured medium. The presence of a thin rupture surface within a highly deformed fault core is also confirmed by laboratory experiments on the frictional behavior at large slip of experimental faults with a gouge layer [26, 27]. The presence of inelastic deformation off the slipping plane implies that the work done in fracture is irreversible and that the mechanical energy is dissipated within the inelastic fault zone [13, 28].

Earthquakes associated with such a fault zone structure are, necessarily, scale dependent because of the inherent scale dependence of the fault zone. This requires that self-similarity, deduced from standard seismological observations, should impose constraints on the hierarchies of scale dependencies. In contrast, scale dependence should imply a departure from self-similarity of earthquake ruptures. In particular, [15] proposed that both fracture energy and the slip weakening distance ($D_c$) are scale dependent and are controlled by a characteristic length scale defined by the predominant wavelength of geometric fault irregularity (or roughness) in the slip direction. However, there are other length scales associated with dissipative processes that affect dynamic fault weakening and earthquake rupture propagation. For instance, it has been shown [29, 30] that thermal pressurization modifies the traction evolution, which in turn affects the inferred values of $G$ and $D_c$. Moreover, [31] proposed that, because earthquake processes are scale dependent, there must exist a discrete hierarchy of such characteristics length scales in order to make them consistent with the overall self-similarity of earthquakes. The debate within the scientific community concerning the scaling of stress drop and radiated energy with earthquake size focuses on similar issues. We believe that it is necessary to introduce self-consistent
definitions of energy terms accounting for the scale dependence of earthquake processes associated with seismological and geological observations of absorbed energy. Although providing a robust solution to all these issues is beyond the goals of the present study, it is important to identify the physical interpretations that have to be revised in order to merge multidisciplinary descriptions of the same processes.

2. A macroscopic description of dynamic fault weakening

Seismologists use finite wavelengths to investigate real fault zones of finite width, so we must consider the way that fracture mechanical concepts developed for perfect cracks can be adapted (or not) for study of earthquakes. Dynamic earthquake ruptures occur within a fault zone volume and on interfaces of finite thickness and involve various non linear dissipation processes coupled over a wide range of spatial and temporal scales. Because most of our understanding of dynamic earthquake ruptures relies on frequency-dependent seismological observations, this raises the question of the scale dependence and scale separation during earthquakes. No theoretical solutions are available today for a physically consistent renormalization of earthquake rupture dynamics based on an accurate representation of the physics of dissipation processes occurring at different scales. One simple reason for this lack of a complete physically-consistent description of earthquake dynamics is the rich hierarchy of scale dependencies and their intrinsic non-linearity which precludes scale separation. Moreover, the poor knowledge of the constitutive laws governing each process further impedes the achievement of such a physically-consistent scale separation. Such processes include strain localization, dynamic fault weakening and stress evolution (including fracture and friction such as asperity breaking, flash heating, gouge formation and evolution by comminution and abrasion, partial melting, frictional heating and thermal pressurization, and mechanical lubrication (see [11] for references). These different physical processes might compete with each other and contribute in a different way to the resulting dynamic fault weakening. Moreover, because a clear separation of the governing scales is not achievable, the discrete hierarchy of these characteristic length and time scales [31] cannot be analytically identified and its existence is only assumed from physical intuition.

In the absence of a detailed physical description of scale dependent processes, we are forced to use classical continuum mechanics and a phenomenological approach to describe dynamic fault weakening and earthquake rupture propagation. The main implication is that the shear traction used in seismological constitutive laws cannot be considered as the shear stress acting on individual gouge fragments or micro-cracks within the slipping zone [15]. Instead, the shear stress, slip, and slip velocity used in the “seismological” constitutive formulation should be considered to be macroscopic or phenomenological quantities. In practice, we must consider the physical quantities characterizing dynamic fault weakening
(shear stress, slip, slip rate) as “equivalent” physical quantities acting on the walls of a fault zone of finite thickness, whereas calculations are carried out on a virtual mathematical fault plane at the macroscopic scale. At this macro-scale we cannot distinguish the diverse physical processes acting at the micro- and meso-scale within the fault zone (that is, at length scales ranging between the gouge grain size and the 30-500 m fault zone thickness). Therefore, we describe dynamic fault weakening in terms of macroscopic traction evolution as a function of macroscopic slip or slip rate.

This phenomenological description is a generalization of non-ideally brittle models (see [20], p.55) to fault zones of finite thickness. Such a generalization requires precise definition of the meaning of cohesive forces [32]. That is, at this macroscopic level the term “cohesive stress” should be considered equivalent to the shear stress resisting friction. Thus, frictional forces originate from processes controlling gouge evolution at the mesoscale, but we represent them as tractions on the “equivalent” (mathematical) fault plane at a macroscopic scale. This implicitly means that we have selected a scale of macroscopic description for which fracture occurs in a continuous and non-ideally brittle medium. A corollary is that friction should be considered as a macroscopic quantity and described using a phenomenological approach, at least for seismological applications. Furthermore, in this formulation friction is defined to be total shear traction (in agreement with [10]).

Constitutive laws govern shear stress evolution and may describe multiple physical processes and parameters, including pore pressure, fault roughness, and state variables. However, at the macroscopic scale dynamic fault weakening is represented by traction evolution. The latter represents a phenomenological constitutive law, which is defined in continuum mechanics at the macroscopic scale. For this reason, we point out that such a phenomenological law is different from a real friction law, because the latter should be defined both for loading and unloading of the dynamic system. In other words, a real constitutive law should control both the fast dynamic failure process (coseismic slip) and the restrengthening process during the coseismic and interseismic phases. On the contrary, seismological observations of a single earthquake provide constraints on the macroscopic traction evolution during the breakdown time, controlled by distinct scale-dependent processes, and they only represent a phenomenological contact law for dynamic weakening. Therefore, two fundamental parameters characterizing dynamic fault weakening, fracture energy (G) and slip weakening distance (Dc), are intrinsically scale dependent. This implicitly entails that frictional strength in this phenomenological description, associated with the drop from the initial to the final stress, is also a scale dependent quantity.

In this context, the rupture process is not scale invariant; this has important implications for the scaling of stress drop, radiated energy and fracture energy with earthquake size (see the discussion in [12]).
3. The earthquake energy balance

The common formulation of the earthquake energy balance states that the total (elastic and gravitational) strain energy variation ($\Delta W$) is partitioned between that radiated from the source and that dissipated on the fault plane:

$$\Delta W = E_S + E_\Sigma$$

(2)

where $E_S$ is energy flow through a surface $S$ containing the fault plane, $\Sigma$, (it includes the seismic energy $E_R$ and the elastostatic work of initial stresses on $S$, [33]) and $E_\Sigma$ is the energy absorbed on the rupture plane by fracture and frictional dissipation. The energy terms appearing in (2) are all global estimates for the whole fault plane, hence, equation (2) defines the earthquake energy balance from a seismological point of view. It is commonly believed that $E_\Sigma$ contains the energy consumed in overcoming fault friction and the energy consumed for expanding the rupture surface area (that is, to maintain the rupture front propagation). Indeed, different authors propose that these two energy terms represent two distinct contributions to $E_\Sigma (= E_F + E_G)$: that is, $E_\Sigma$ is commonly assumed to be partitioned into fracture energy $E_G$ and frictional dissipation $E_F$ ([8,12]), where the total fracture energy for the whole fault is written as:

$$E_G = \iint Gd\Sigma,$$

(3)

and $G$ is the fracture energy at a specific target point on the fault as defined by (1). This implicitly requires that $E_F$ is identified (see Fig.1 and equation 3) and separated from fracture energy in order to associate the latter with the crack driving force. We believe that this separation is model dependent and we disagree with its broad application to real fault zones.

The formulation of the earthquake energy budget stated in equation (2) is consistent with the phenomenological description proposed in the previous section, but the separation between fracture energy and frictional dissipation (Fig.1) requires further discussion. For instance, several recent studies [2, 10] have proposed that seismological estimates of fracture energy include frictional dissipation. In our phenomenological description the total macroscopic shear traction is friction and similarly to [20], we can derive an expression for the mechanical work absorbed on the mathematical fault plane (eq. 2.2.16 in [20]):

$$\tau_i \Delta \dot{u}_i = 2 \dot{\gamma} + \Delta q$$

(4)

where $\tau_i$ is the macroscopic traction (which in our nomenclature is friction) and $\Delta \dot{u}_i$ is the slip velocity on the fault plane $\Sigma$; $\dot{\gamma}$ is the surface energy rate, $\Delta q$ is the heat flux per unit area generated on the fault ($\Delta q$ contains not only the actual heat, but also the “radiation loss”, the energy radiated from the crack as high frequency stress waves that do not get to the far field,
as well as other dissipative processes). We emphasize that equation (4) differs from that proposed by [20] because it contains the total shear traction \((\tau_i)\) and not an unknown frictional stress. Equation (4) allows the definition of the macroscopic frictional work rate for a point on the fault plane and it states that it is partitioned between surface energy and heat generation. We discuss this issue further below. To obtain the macroscopic frictional work \((\mathcal{I}_f)\) at a target point, we integrate (4) through time

\[
\mathcal{I}_f = \int_0^{t_m} \tau_i \Delta \dot{u}_i \, dt = \int_0^{t_m} d_{\text{int}} \tau_i (\Delta u_i),
\]

where \(D_{\text{tot}}(=\Delta u_i)\) is the final slip (see Fig. 1), \(t_m\) is the duration of slip, \(\tau_i = \sigma_i n_j\) is the shear traction (i.e., friction). \(\mathcal{I}_f\) is the frictional work density (work per unit area). It is a function of the position on the fault plane, because the final slip is a function of fault position. By integrating equation (5) on the fault plane we obtain a global (i.e., for an extended source) estimate of the macroscopic frictional work

\[
\mathcal{E}_\Sigma = \iint_{\Sigma} \mathcal{I}_f \, dS = \iint_{\Sigma} \tau_i \Delta \dot{u}_i \, dt = \iint_{\Sigma} d_{\text{int}} \tau_i (\Delta u_i)
\]

where \(\Sigma\) is the final ruptured area. Equation (5) shows that the area below the slip weakening curve represents a measure of the frictional work density at a single point on the fault plane and its surface integral (6) yields the estimate of the total frictional work. Equation (6) expresses the total mechanical work absorbed on the fault plane because this model does not include any stress singularity at the crack-tip (i.e., there is a finite stress over a finite area).

The mechanical work absorbed on the fault plane is the irreversible work of macroscopic shear tractions acting on the fault surface representing the walls of the fault zone and it is equal to the work done by external forces less the total kinetic energy and the strain energy (see [34]). In other words, in a realistic fault zone model described through a phenomenological description, the macroscopic frictional work contains all the mechanical energy absorbed within the fault zone, including breakdown work (i.e., seismological fracture energy). We will provide a definition of the breakdown work (see Fig.2) in the next section and we will try to explain why it is convenient to avoid calling it fracture energy [2].

4. The breakdown work

Recently, [2] have defined an alternative measure of work that characterizes traction evolution in kinematic earthquake models. In real earthquakes the traction-change and slip-velocity vectors are usually not collinear, making the use of the scalar equations for fracture energy problematic. These authors defined breakdown work \(W_b\) to be the excess of work relative to that defined from a minimum traction level achieved during slip \(\tau_{\text{min}}\) (Fig. 2):
where $\Delta \tilde{u}(t)$ is slip velocity and $\bar{\tau}(t)$ is shear traction; $t_b$ is the time at which the minimum traction $\bar{\tau}_{\text{min}}$ is reached at the target point (which we consider an estimate of the breakdown time $t_b \approx T_b$). As part of the definition of $W_b$, [2] specified a way to select the initial traction vector required to calculate $\bar{\tau}(t)$ from the traction-change vectors derived from the kinematic slip models. $W_b$ is an energy density ($J/m^2$), but [2] called it breakdown work for simplicity. Breakdown work $W_b$ is the energy density (or work) associated with the breakdown phase (i.e., traction changes from the initial level to the minimum value) and it has been interpreted as the energy density spent to allow the rupture to advance at a determined rupture velocity similarly to "seismological" fracture energy ($G'$) [12]. Breakdown work includes the energy lost during any initial slip-hardening phase, consistent with the definition of [35]. [2] have also defined the excess work $W_e$ as the sum of breakdown work and restrengthening work ($W_b$ and $W_r$, respectively), where restrengthening work is defined as (see Fig. 2):

$$W_r = \int_{t_b}^{t_m} (\bar{\tau}(t) - \bar{\tau}_{\text{min}}) \cdot \Delta \tilde{u}(t)dt,$$

where $t_m$ is the total duration of slip at the point; $W_r$ is also an energy density.

For traction evolution curves characterized by a restrengthening phase after the end of the breakdown process and before the healing of slip ($t_b \leq t \leq t_m$) (Fig.2) the excess of work over a minimum stress level is given by $W_e (= W_b + W_r)$. However, because in most cases $W_b$ is much larger than $W_r$ [2] and because the initial traction evolution during the weakening phase controls slip acceleration and rupture propagation (the rupture front is around the target point during the breakdown time) it is reasonable to concentrate attention on the breakdown work $W_b$. The definition of breakdown work [2] is very similar to the definition of fracture energy given by [23] in their equation 7, similar to our (1). These latter authors defined the fracture energy (although they did not use that term) as the energy surplus made available for unit area of rupture advance, the surplus being the excess of work of applied forces over the energy stored in deformation and the dissipation against the residual part of shear resistance on the slipping plane. [23] used an asymptotic residual stress level ($\tau_{\text{res}}$), while [2, 36, 37] used a minimum stress value, which can be more easily identified on complex traction evolution curves with possible restrengthening expected for real earthquakes in complex fault zones. [36] and [37] defined quantities similar to $W_b$ using a minimum stress level, but their definitions are suitable only for the situation in which slip velocity and traction are collinear.

We show in Figure 3 the scaling of $W_b$ with seismic moment. Table 1 lists the average breakdown work estimates (average work density on the fault plane of each earthquake)
revised from those previously published by [2] and new estimates from [38], who applied the same method as [2, 51] to kinematic slip models of the 1999 Hector Mines [39], 2002 Denali [40], 2004 Parkfield [41], and 2005 Fukuoka [42] earthquakes. This figure confirms the finding of [1] and [2] suggesting that breakdown work density (or fracture energy) scales with seismic moment following a power law whose slope is nearly 0.6. Several recent studies attempted to find a relation between breakdown work and rupture velocity. However, such a scaling is poorly understood because rupture velocity in kinematic slip models is not well constrained.

In the framework of the phenomenological description proposed in this study we point out that these seismological estimates of breakdown work represent the only measurable portion of the mechanical work dissipated within the fault zone, since the absolute stress level on the fault is unknown. Because the selected scale of the macroscopic description coincides with the thickness of the damage zone, the seismological estimates of breakdown work should also contain the energy lost outside the principal slipping zone for off-fault cracking and plastic deformation [1, 28]. As described above, in the phenomenological description the fault zone volume is replaced by a fictitious contact surface of zero thickness. This mathematical surface is characterized by a phenomenological friction law or contact law that is supposed to capture the main features of dynamic fault weakening during the earthquake rupture. In this framework, this contact law can be interpreted in terms of a surface dissipation potential which is controlled by the macroscopic quantities defined on that surface such as slip, slip rate, internal variables, etc… Therefore, the macroscopic frictional work is the total intrinsic power of dissipation of the whole fault zone.

The most important implication of this reasoning is that we are not allowed to link the main seismological variables and the dissipation mechanism directly to the physics of processes at the micro- and meso-scales. The seismological breakdown work estimates can be expressed in terms of a virtual dissipation potential on the mathematical surface (because they represent estimate of dissipation at the macroscopic scale on a virtual mathematical plane). This allows the application of classic fracture mechanics concepts to this phenomenological description of a macroscopic process, as we discuss below, but it does not allow the interpretation of this virtual dissipation potential in terms of the real physical processes controlling gouge formation and damage evolution at the meso- and micro-scales. This raises the question of the partitioning of the total intrinsic power of dissipation between heat and real surface energy absorbed within the fault zone. In other words, the problem is to discuss the validity of the partitioning stated in equation (4) in the framework of the phenomenological description proposed in this study.

5. Breakdown work and seismological fracture energy
In the present study we have used the term $G'$ to identify seismological fracture energy and to emphasize that this measure is different from classic fracture energy $G$ (as defined in 1). Few papers provided estimates of seismological fracture energy for different real earthquakes. Among them, [12] defined and measured an energy density, originally using the term $G'$, which can be evaluated from seismic moment, corner frequency and radiated energy estimates. Their $G'$ coincides with $G$ when overshoot or undershoot are negligible, that is final stress coincides with dynamic frictional stress ($\tau_{res}=\tau_1$ as in models shown in Fig.1). However, despite [12] measure an energy density (measured in J/m$^2$), their $G'$ is by definition a global estimate (for the whole fault). Moreover, $G'$ measured by [12] is model dependent, since it relies on the scaling of stress drop with seismic moment and source radius as well as of corner frequency with source radius. [12] pointed out that $G'$ scales with slip according to a power law with slope nearly equal to 1.3.

[43] examined the scaling properties of distributions of fracture energy and stress drop for a suite of dynamic rupture models of well-recorded earthquakes for which reliable kinematic source inversions exist. The spontaneous dynamic models fit the inverted slip- and rupture-time distributions for individual events. These authors found that fracture energy scales with seismic moment according to a power law whose slope is 0.55 (thus in agreement with [2, 51] and Figure 3) and with slip following a power law with slope roughly equal to 1. The estimates provided by [43] cannot be considered seismological estimates of fracture energy, because they are model dependent since they are obtained through a forward modeling approach in which a slip weakening law is imposed in spontaneous dynamic simulations.

We propose that breakdown work is fittingly a reliable estimate of seismological fracture energy. Contrary to $G'$ defined by [12], $W_b$ is an energy density measured on the fault plane, whose distribution can be imaged if kinematic models have the appropriate resolution. Moreover, our $W_b$ measures are identical to $G'$ defined by [12] if their power law model is adopted (see Fig.1-b). $W_b$ scales with seismic moment and final slip (the latter follows a power law with slope equal to 2, see [2, 10]), because it is consistent with self-healing pulse models [1]. Finally, breakdown work is not model dependent, since it does not rely on specific assumptions on the traction evolution, even if the inferred dynamic traction depends on the parameterization of kinematic source models which might be not dynamically consistent.

6. Mechanical work partitioning

Recent geological investigations have suggested [3, 4] that the surface energy produced during the propagation of a dynamic earthquake rupture at a given point on the fault is given by:
\[ \tilde{U}_S = (A_{SZ} + A_{DZ}) \gamma \quad [\text{J m}^{-2}], \]  

where \( A_{SZ} \) and \( A_{DZ} \) are the new surface per unit fault area (\( A_{SZ} \) and \( A_{DZ} \) are dimensionless) produced in the slipping zone and in the damage zone, respectively, and \( \gamma \) is the specific surface energy. It follows that, in order to estimate \( \tilde{U}_S \) from microstructural analyses, the new surface area produced during seismic slip both in the slipping and damage zones has to be determined. According to (9), the geological observations should yield estimates of surface energy for the whole fault zone thickness. Therefore, they might be considered as measures obtained at the same scale of the seismological macroscopic description. In this case, the comparison between geological estimates of surface energy and seismological measures of breakdown work is physically consistent. We point out that a comparison between energy estimates relies on the assumption that they are both representative at the macroscopic scale.

In this framework, the energy absorbed in the fault zone can be partitioned differently from that discussed above. Indeed, the energy absorbed on the fault plane can be written as:

\[ E_S = U_S + Q \]

where \( U_S \) is the surface energy created by the coseismic rupture (a global estimate for the whole fault plane) and \( Q \) is heat for the whole fault. Relation (10) is appropriate over the time scale of hours, not decades, because over interseismic time scales some (although not all, see [44]) of the coseismically created surfaces will heal, releasing surface energy as heat. The energy partitioning stated in (10) is equivalent to that stated in (4) but defined for the whole fault, while (4) is defined at a specific position on the fault (and for a unit time).

The comparison between observed surface energies and breakdown work estimates for recent earthquakes shown in Fig. 3 supports the suggestion of [2] that 85 – 98% of the average breakdown work is heat and that only 2 – 15% is surface energy. Indeed, this figure shows surface energy estimates from the Punchbowl fault (California) [4] and from [3], who determined heat and surface energy from field and microstructural analyses of an exhumed segment of Gole Larghe fault zone located in Italian Alps characterized by the presence of pseudotachylyte. They found surface energy, estimated from microcrack density within a clast of the pseudotachylyte and in the fault wall rock to be smaller than 0.4 MJ·m\(^{-2}\) for a paleoearthquake having an estimated moment between 6-7. We have refrained from adding the surface energy measurements of [5] from the San Andreas fault at Tejon Pass to the plot because these measurements are likely a significant overestimate owing to problems in the particle size analyzer [45].

While the results for surface energy in Fig. 3 are straightforward to interpret, estimates of frictional heat in Fig. 3 are less easy to interpret. That figure shows frictional heat, estimated by [3] from the amount of pseudotachylytes, to be \( \sim 27 \text{ MJ·m}^{-2} \), which exceeds by almost a factor of 10 the breakdown work of comparable size earthquakes like the Tottori
event. A similar but smaller discrepancy exists for the 1995 Kobe event (Hyogo-ken Nanbu), for which [51] inferred an average breakdown work of 3.32 MJ·m$^{-2}$. This value can be compared with the estimate of heat production provided by [46] based on electron spin resonance measurement of partial defect annealing in quartz obtained from a borehole at 389 m depth on a branch of the Nojima fault near Ogura (Awaji Island, Japan) that probably slipped during the Kobe earthquake. They found that the heat generated in that location was 8 MJ·m$^{-2}$, exceeding the average breakdown work by a factor of ~2. Thus, the comparison with estimates of frictional heat generation is not easy to interpret.

7. Discussion and concluding remarks

We have presented a model for an elastic material outside the fault zone and an inelastic fault zone of finite thickness represented at the macroscopic scale by a mathematical plane (i.e., a fault of zero thickness) suitable for interpreting seismological observations. Although the constitutive properties of the fault zone are governed by the physical processes controlling gouge and damage evolution at the mesoscopic and microscopic scales, we represent dynamic fault weakening at this scale in terms of traction evolution as a function of slip and other macroscopic internal variables representing a phenomenological friction or contact law. We emphasize that seismological observations, which depend on selected frequencies and wavelengths, can provide a lower bound on the virtual surface dissipation potential. The proposed model allows a physically consistent interpretation of breakdown work (seismological fracture energy, $G'$) as the only measurable portion of the total macroscopic intrinsic power of dissipation.

The physical interpretation of the total macroscopic intrinsic power of dissipation requires some clarification. Indeed, in order to interpret part of the virtual surface dissipation potential as real surface energy measured from particle size distribution of fault gouge (therefore accounting for absorbed energy measures derived from geological investigations), we have to face the problem of scale dependent observations. If we rely on the assumption that geological estimates of surface energy are also representative of the whole fault zone thickness, we can compare them with seismological breakdown work estimates (since they are considered representative of the same macroscopic scale). This comparison confirms that surface energy is a negligible contribution to the mechanical work dissipated on the fault and corroborates that breakdown work (or seismological fracture energy) is not a negligible contribution to the earthquake energy budget. Indeed, the total breakdown energy (the surface integral of breakdown work on the whole fault plane, [2]) for a target earthquake of $M_o = 10^{19}$ N·m ranges between $10^{14}$ and $10^{15}$ J [2, 10, 51], which is similar to the energy radiated by an earthquake of similar size [47]. This underscores the need for improved understanding of dissipative processes on the fault plane, because the magnitude of dissipative processes
influences the radiated field by changing the energy available for radiation. However, it is important to point out that the mechanical work partitioning between surface energy and other dissipative mechanisms does not affect earthquake dynamics.

The most important parameter for interpreting earthquake dynamics is the energy necessary to maintain rupture propagation on the fault plane. If we want to recognize this energy in the framework of the phenomenological description proposed in this study, we need to identify and interpret the crack driving force for the earthquake rupture: that is, the portion of the mechanical energy controlling rupture speed. At this level of macroscopicity energy has to be absorbed near the virtual rupture front and shear stress is finite at the virtual crack tip and over the macroscopic slipping region. Therefore, to face this problem we can apply a slip weakening model, as those depicted in Figure 1, to the virtual mathematical plane; this allows us to associate the seismological fracture energy \( G_r' \) or the breakdown work with the crack driving force:

\[
W_b = G_r' = g(v_r)G_{stat} = g(v_r)\frac{K_{r}^2}{2\mu}
\]

where \( K_r \) is the macroscopic stress intensity factor (subscript refers to distinct crack modes) and \( \mu \) is the rigidity. Here, \( g(v_r) \) is a function of crack speed \( v_r \) depending on the crack mode and \( G_{stat} \) is the static energy release rate (see \([1,8]\) and references therein). Equation (11) is widely adopted in the literature and it provides an analytical relation between fracture energy and rupture velocity. However, it is important to emphasize that the application of this relation to real earthquakes relies on classic fracture mechanics concepts \([33, 48]\) applied to a virtual mathematical fault of zero thickness in the framework of a phenomenological model based on continuum mechanics. The use of this relation to interpret seismological data does not allow recovery of fracture energy \( G_r' \) in terms of the effective fracture energy \( (\gamma_{eff}) \). In the same way, the stress intensity factor appearing in (11) should not have the same physical meaning of the analogous parameter defined in classic fracture mechanics \([49]\).

There are several important implications rising from this reasoning. (i) The partitioning between measurable seismological fracture energy \( (W_b \text{ or } G_r') \) and the remaining intrinsic power of dissipation (named frictional heating in Figure 1 and expressed as to \( \tau_{res} \Delta u_i \) in equation 1) is only necessary for identifying the crack driving force in the framework of this phenomenological description. Therefore, the separation between fracture energy and frictional dissipation, commonly adopted in the literature to express the mechanical work partitioning, is misleading since friction (or the total intrinsic power of dissipation) contains by definition fracture energy. (ii) The rupture velocity imaged from modeling seismological data is a macroscopic parameter, which allows the representation of the propagation of a virtual rupture front on a mathematical plane. In reality and at smaller scales, the earthquake rupture propagation is a much more complex process that probably
involves heterogeneous geometrical fracture paths (kinks, bending, branching, etc.) as well as the interaction and coalescence of micro-cracks within the fault zone volume. (iii) Seismological fracture energy $G'$ (and breakdown work $W_b$) is a scale dependent parameter and it cannot be associated with any physical process occurring at smaller scales without properly solving a scale separation problem. Therefore, its association with the on-fault effective fracture energy ($\gamma_{\text{eff}}$) is not physically correct. This also because in a non-singular stress model $\gamma_{\text{eff}}$ vanishes [34]. (iv) The analytical form of the function $g(\nu_r)$ has been derived from classic fracture mechanics identifying the limiting speed of a propagating crack-tip singularity and therefore its application to real earthquakes is limited to the validity of this macroscopic description. At this macro-scale we can assume that strain rate is localized on the virtual mathematical plane (which might not be tenable for natural faults) and because rupture velocity from seismological observations is a macroscopic parameter.

The physical interpretation of the phenomenological friction or contact law also needs to be discussed in detail. In particular, the interpretation of seismological observations using application of laboratory rock friction results requires careful discussion. We must face the problem of extrapolating results of rock physics experiments, at the laboratory scale, to real earthquakes in complex fault zones. It is evident that the understanding of source properties requires contributions from different research fields, such as geology and experimental rock physics [48], not limited to seismology. Therefore, laboratory investigations are necessary to access the particular processes and to perform tests aimed at characterizing the associated intrinsic time and length scales so that the proper constitutive relations can be inferred and used in numerical simulations. However, the intrinsic scale dependence of the earthquake process in natural faults makes this task a difficult challenge. At present, we can use the available constitutive laws, such as rate and state friction (R&S) or slip weakening (SW), and apply them as possible representations of the phenomenological contact law. However, the application of rate- and state-dependent constitutive laws to model seismological observations of earthquake dynamics [11, 18] at this macroscopic scale requires that we think carefully about the common interpretations of the state variable and friction coefficient as evidence for evolution of contact properties of the sliding surfaces. Similarly, the characteristic slip weakening distance in the SW model cannot be associated with any physical process occurring at the meso- or micro-scale. In the R&S case, the state variable should be considered just as a variable representing gouge and damage evolution and accounting for all our lack of knowledge in modeling physical processes at their proper scales. The main point is that once the scale of the macroscopic representation is selected, the contact law on the virtual mathematical plane cannot be associated to any physical process occurring at other scales.
The phenomenological model presented in this study to interpret seismological observations is consistent with consideration of the slip weakening curve as a global representation of dynamic fault weakening that accounts for processes controlling fault zone evolution at the meso- and microscopic scales. This is different from considering the slip weakening behavior as a unifying constitutive law (as proposed by [15]). Moreover, because the physical quantities characterizing dynamic fault weakening (G or $D_c$) are scale dependent, we cannot associate them with any fixed length scale parameter without properly solving a scale separation problem. For instance, $D_c$ inferred from seismological observations may be associated with the predominant wavelength of fault roughness in the slip direction ([15]) only if the selected scale of the macroscopic representation coincides with the principal slipping zone (i.e., the sliding surface), and not with the whole fault zone thickness.

We emphasize that identification of the main physical processes controlling fault behavior requires connecting the constitutive relationships to the appropriate time and length scales. The achievement of this task requires the identification of dominant physical processes through a new generation of laboratory experiments and the definition of analytical laws describing the constitutive behavior. This is in our opinion an extraordinary challenge for future research.

**Acknowledgments**

We are indebted to Paul Spudich for the scientific discussions and the valuable contributions that allowed us to submit this paper. We thanks J.P. Vilotte who reviewed this manuscript; his contributions, suggestions, criticisms were extremely important to better focus the main outcomes of this study. We have appreciated the comments of the Editor Rob van der Hilst during the editorial process. We thank Chris Marone, and Giulio Di Toro for helpful discussions.
References


[38] Spagnuolo, E., 2006, Evoluzione della trazione dinamica sulla faglia durante I forti terremoti, Thesi di Laurea in Geofisica, Univ. degli Studi di Roma “La Sapienza.”


[48] Beeler, N. M., 2006, Inferring earthquake source properties from laboratory observations and the scope of lab contributions to source physics, in Earthquakes: Radiated Energy and the Physics of Faulting, Geophysical Monograph Series 170, American Geophysical Union, 10.1029/170GM12


[51] Tinti, E., P. Spudich, and Cocco M., 2008. Correction to “Earthquake fracture...
Figure 1. (a) Heavy line: Classical slip weakening curve proposed in the Ida’s [32] and Andrews’ [13-14] models. In this plot $\tau_0$ is the initial stress, $\tau_p$ and $\tau_{res}$ are the yield and the residual stress values, respectively. $D_{tot}$ is the final slip and $D_c$ is the slip weakening distance. (b) Heavy line: power-law traction evolution model proposed by [12] characterized by a non-linear decay of traction with slip. Here $\tau_1$ is the dynamic traction at the last slip increment and $\tau_{min}$ is the minimum stress. Both models are characterized by $\tau_{res} = \tau_{min} = \tau_1$. In many fracture mechanical models the shaded area below the residual stress level is identified with the frictional dissipation (named heat by [20]) and $G$ is the fracture energy defined in (1) (see text for explanations).

Figure 2. Slip weakening curve obtained by [50]. $\tau_{min}$ and $\tau_1$ are the minimal and the final stress values. The latter is measured at the time of the last slip increment. The breakdown stress drop is defined as $\Delta \tau_b = \tau_p - \tau_{min}$. 
Figure 3. Scaling of breakdown work with seismic moment and comparison with heat and surface energy estimates. Values of average breakdown work for extended source are taken from [2, 51], except the estimates for the 1999 Hector Mine, the 2000 Denali and the 2004 Parkfield earthquakes (USA) and the 2004 Fukuoka (Japan) event, which are taken from [38]. See table for references.
Table 1. Breakdown work, seismic moment, surface energy and heat measured for different earthquakes. Data are plotted in Figure 3.

<table>
<thead>
<tr>
<th>Earthquake</th>
<th>Wb (MJ/m²)</th>
<th>M₀ (Nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984 Morgan Hill ¹</td>
<td>2.72</td>
<td>2.62e18</td>
</tr>
<tr>
<td>1997 Colfiorito 0033 ¹</td>
<td>0.80</td>
<td>4.38e17</td>
</tr>
<tr>
<td>1997 Colfiorito 0940 ¹</td>
<td>1.94</td>
<td>1.04e18</td>
</tr>
<tr>
<td>1997 Colfiorito-Oct ¹</td>
<td>2.22</td>
<td>6.478e17</td>
</tr>
<tr>
<td>1979 Imperial Valley ¹</td>
<td>2.12</td>
<td>6.4e18</td>
</tr>
<tr>
<td>1979 Imperial Valley ¹</td>
<td>3.64</td>
<td>8.64e18</td>
</tr>
<tr>
<td>2000 Western Tottori ¹</td>
<td>3.38</td>
<td>1.1e19</td>
</tr>
<tr>
<td>2000 Western Tottori ¹</td>
<td>14.26</td>
<td>1.9e19</td>
</tr>
<tr>
<td>2000 Western Tottori ¹</td>
<td>6.04</td>
<td>1.158e19</td>
</tr>
<tr>
<td>2000 Western Tottori ¹</td>
<td>5.46</td>
<td>1.195e19</td>
</tr>
<tr>
<td>1995 Kobe ¹</td>
<td>3.32</td>
<td>2.44e19</td>
</tr>
<tr>
<td>1992 Landers ¹</td>
<td>40.52</td>
<td>9.26e19</td>
</tr>
<tr>
<td>1992 Landers ¹</td>
<td>29.14</td>
<td>1.02e20</td>
</tr>
<tr>
<td>1994 Northridge ¹</td>
<td>11.5</td>
<td>1.22e19</td>
</tr>
<tr>
<td>1999 Hector Mine by [39]²</td>
<td>81.2</td>
<td>6.7e19</td>
</tr>
<tr>
<td>2002 Denali by [40]²</td>
<td>41.4</td>
<td>7.57e20</td>
</tr>
<tr>
<td>2004 Parkfield by [41]²</td>
<td>0.42</td>
<td>1.08e18</td>
</tr>
<tr>
<td>2005 Fukuoka by [42]²</td>
<td>10.68</td>
<td>1.15e19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Surface energy (MJ/m²)</th>
<th>Heat (MJ/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudotachylyte [3]</td>
<td>27</td>
</tr>
<tr>
<td>Punchbowl fault [4]</td>
<td>0.5</td>
</tr>
<tr>
<td>1995 Kobe [46]</td>
<td></td>
</tr>
</tbody>
</table>

¹ Measures taken from [2, 51];
² From [38].