# New Perspectives in Crack and Fault Dynamics

The Harvard community has made this article openly available. **Please share** how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td><a href="http://dx.doi.org/10.1007/0-306-46956-1_1">http://dx.doi.org/10.1007/0-306-46956-1_1</a></td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:2643117">http://nrs.harvard.edu/urn-3:HUL.InstRepos:2643117</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
NEW PERSPECTIVES ON CRACK AND FAULT DYNAMICS

James R. Rice
Division of Engineering and Applied Sciences and
Department of Earth and Planetary Sciences
Harvard University, Cambridge, Mass., USA
rice@esag.as.harvard.edu

Abstract Recent observations on the dynamics of crack and fault rupture are described, together with related theory and simulations in the framework of continuum elastodynamics. Topics include configurational instabilities of tensile crack fronts (crack front waves, disordering, side-branching), the connection between frictional slip laws and modes of rupture propagation in earth faulting, especially conditions for formation of self-healing slip pulses, and the rich faulting and cracking phenomena that result along dissimilar material interfaces due to coupling between slippage and normal stress alteration.

1. INTRODUCTION

The dynamics of cracking and faulting has seen much recent progress, with implications for structural mechanics, materials physics, tribology, and seismology. In this brief review, the following topics will be discussed:

- Crack front waves
- Limiting rupture speeds
- Mode of rupture on faults: enlarging shear crack or self-healing slip pulse?
- Interfacial fracture dynamics, slip, and opening

To put things in context, we will be using linear isotropic elasticity everywhere except along slip or crack zones. The governing equations of that theory are the equations of motion $\nabla \cdot \sigma = \rho \frac{\partial^2 u}{\partial t^2}$, where $n \cdot \sigma$ or $u$ or a combination are given on the boundary, together with stress-displacement gradient relations $\sigma = \lambda (\nabla \cdot u) I + \mu [(\nabla u) + (\nabla u)^T]$. 

1
leading to the Navier equations for the displacement field:

\[(\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu \nabla^2 \mathbf{u} = \rho \frac{\partial^2 \mathbf{u}}{\partial t^2}.\]  \hspace{1cm} (1)

Those equations admit a family of waves. There are primary (or dilational) and secondary (or shear) body waves, with respective speeds \(c_p = \sqrt{(\lambda + 2\mu)/\rho}\) and \(c_s = \sqrt{\mu/\rho}\), typically with \(c_p = 1.7-2.1\ c_s\), and also Rayleigh surface waves with typical speed \(c_R = 0.90-0.94\ c_s\). As we review recent progress, we shall learn of a newly discovered type of wave that lives along a moving crack front, and also see the significance of a little known generalization of Rayleigh waves for the dynamics of dissimilar materials.

Much of the numerical simulation that is reviewed here has been done by a spectral elastodynamic method \([1]-[3]\), which is of limited general utility but is very well tailored to produce good numerical solutions, without grid dispersion, for crack or fault problems that can be considered as events at the border between two homogeneous elastic half spaces. The border at \(y = 0\), which is the fault or crack plane, undergoes displacement discontinuities

\[\delta_\alpha(x, z, t) = u_\alpha(x, 0^+, z, t) - u_\alpha(x, 0^-, z, t) \quad (\alpha = x, y, z).\]  \hspace{1cm} (2)

The spectral method treats these problems by writing \(\delta_\alpha\) as a large but finite Fourier sum of terms in form \(D_\alpha(t) e^{i(kx + mz)}\); the rupture domain is replicated periodically in \(x\) and \(z\). The elastodynamic equations are solved exactly in each half space, for terms of that same \(e^{i(kx + mz)}\) space dependence, so that in the end the traction stress components \(\sigma_{y\alpha}\) are given as a corresponding Fourier sum of terms in the form \(T_\alpha(t) e^{i(kx + mz)}\). Each \(T_\alpha(t)\) can then be determined in terms of the \(D_\alpha(t)\), in an expression involving a convolution over prior \(D_\alpha(t)\). That, effectively, lets us relate the history of the \(\sigma_{y\alpha}\), at fast-Fourier-transform sample points along the interface, to the history of the \(\delta_\alpha\), and to any given loading stresses, in such a way that the elastodynamic equations are satisfied in the two half spaces. The system is then closed, so that a definite solution can be computed, by specifying another relation between the \(\sigma_{y\alpha}\) and the \(\delta_\alpha\); that is a constitutive relation for crack opening \([2, 3]\) (cohesive model) or frictional slip \([1]\). Different forms will be noted as we go along.

2. CRACK FRONT WAVES

Figure 1 shows a tensile crack growing in a 3D solid, along the plane \(y = 0\). Earlier work \([4, 5]\) addressed a simplified version of this problem for a model elastic theory with a single displacement component \(u\),
satisfying a scalar wave equation. It showed that there was a long lived response to a local perturbation of the crack front, e.g. by the crack passing through a region where the fracture energy was slightly different than elsewhere, although in that scalar model the crack ultimately recovered a perfectly straight front. That led to great interest in solving such problems in the context of actual elasticity. Willis and Movchan [6] produced the corresponding small perturbation solution, although that was difficult to interpret and the fuller implications of their solution were revealed later [7], confirming what had been suggested from spectral numerical simulations [8, 9]: for crack growth in a perfectly elastic solid with a constant fracture energy, perturbation of the crack front leads to a wave that propagates laterally, without attenuation or dispersion, along the moving crack front. The wave is called a crack front wave.

Two different fracture formulations have been used for these investigations. The first is a singular crack model, in which one sets \( \sigma_{yy} = 0 \) on the mathematical cut \( y = 0 \), which is the crack surface. That leads to a well known singular field of structure

\[
\sigma_{\alpha\beta} = \sqrt{\frac{\mu G}{r}} \Sigma_{\alpha\beta} \left( \frac{\theta, v, c_s}{c_s} \right) + \text{(terms that are bounded at } r = 0), \tag{3}
\]

where \( r, \theta \) are polar coordinates at the crack tip and the \( \Sigma_{\alpha\beta} \) are universal functions for a given cracking mode. The strength has been normalized in terms of \( G \), which is the energy release rate (energy flow to crack tip singularity, per unit of new crack area), expressed by

\[
G = \lim_{\Gamma \to 0} \int_{\Gamma} \left[ n_1 \left( W + \frac{1}{2} \rho \left| \frac{\partial u}{\partial t} \right|^2 \right) - \mathbf{n} \cdot \mathbf{\sigma} \cdot \frac{\partial u}{\partial x_1} \right] \mathrm{ds}. \tag{4}
\]
Here, \( W \) is the strain energy density, coordinate \( x_1 \) points in the direction of crack growth, \( \Gamma \) is a circuit that loops around the crack tip at the place of interest, and \( s \) is arc length along \( \Gamma \), whose outer normal is \( n \).

Owing to the great advances on unsteady crack motion in 2D elasticity by Kostrov [10] and Eshelby [11] for mode III (anti-plane shear), Freund [12] for mode I (tensile cracking), and Fossum and Freund [13] for mode II (in-plane shear), we know that \( G = g(v(t))G_{\text{rest}} \). Here \( g(v) \) is a universal function of crack speed \( v \) for each mode, normalized to \( g(0^+) = 1 \) and diminishing to \( g(c_{\text{lim}}) = 0 \) at a limiting speed (at least for fully subsonic fracturing), which is \( c_{\text{lim}} = c_R \) for modes I and II, and \( c_s \) for mode III. The variable \( G_{\text{rest}} \) is a complicated and generally untractable functional of the prior history of crack growth and of external loading, but is independent of the instantaneous crack speed \( v(t) \).

Because of that structure for \( G \), it is possible for cracks to instantaneously change \( v \) if the requisite energy \( G_{\text{crit}} \), which must be absorbed for fracture, changes discontinuously along the fracture path. Also, for a solid loaded by a remotely applied stress, it will generally be the case that \( G_{\text{rest}} \) increases as the crack lengthens, and increases quadratically with the intensity of the applied stress. Thus, if \( G_{\text{crit}} \) is bounded, then in a sufficiently large body \( g(v) \) will be driven towards 0, which means that \( v \) will accelerate towards \( c_{\text{lim}} \).

For the crack growing on a plane in 3D, as in Fig. 1, it has so far been possible to solve for the elastic field [6] only for a crack whose front position \( x = a(z,t) \) is linearly perturbed from \( x = v_0 t \), that is, from a straight front moving at uniform speed \( v_0 \).

An alternative fracture formulation, often more congenial to numerical calculation, explicitly accounts for a gradual decohesion, imposing a weakening relation between stress and displacement-discontinuity as a boundary condition on the potential crack plane. (See the inset diagram of Fig. 1.) Thus, the singularity at the crack tip is smeared into a displacement weakening zone. Its width \( R \) scales [14] as, roughly, \( \mu \delta_0 / \sigma_0 f(v) \) with \( \sigma_0 \) being the maximum cohesive strength and \( \delta_0 \) the displacement at which cohesion is lost, and with \( f(v) \) being universal for a given mode, and with \( f(0) = 1 \) and \( f(c_{\text{lim}}) = \infty \). The latter limit poses a challenge for numerical simulation of fracture at speeds very close to \( c_{\text{lim}} \). When \( R \ll \) all length scales in problem (crack length, distance of wave travel, etc.), predictions of the displacement-weakening model agree with those of the singular crack model, with \( G_{\text{crit}} \) identified (Fig. 1) as the area under the cohesive relation [14].

The spectral methodology was used with the cohesive model [8, 9] to study what happens when a crack front, moving as a straight line across \( y = 0 \) with uniform rupture speed \( v_0 \), suddenly encounters a localized
“asperity” region for which $G_{\text{crit}}$ is modestly higher than its uniform value, say, $G_{\text{crit},0}$, prevailing elsewhere. The resulting perturbation in the crack front propagation velocity, $v(z, t) - v_0$, is shown in Fig. 2. There, $\lambda$ is the periodic replication distance in the $z$ direction, as required within the spectral method. The asperity has diameter $0.02\lambda$, and has $\Delta G_{\text{crit}} = \frac{2[v(z, t) - v_0]}{v_0 \Delta G_{\text{crit}}/G_{\text{crit},0}} + \text{const} \times t$.

\begin{align*}
\text{(for discrete, equally spaced, } t \text{ values)}
\end{align*}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.png}
\caption{Numerical simulation results for perturbation of rupture propagation velocity $v(z, t)$ when the crack front, moving at uniform speed $v_0 = 0.5c_s$, encounters a small region of altered fracture energy. Adapted from Ref. [9].}
\end{figure}

$0.1G_{\text{crit},0}$; both $\sigma_0$ and $\delta_0$ were increased by 5% within it. The motivation to examine the response to such a small localized perturbation arose from understanding of the scalar model [5], for which the form of response to such isolated excitation was critical to understanding the response to (small) random excitation. Quite remarkably, when those calculations were done (Fig. 2) for the mode I crack in true elastodynamics, the perturbation of the crack front seemed to propagate as a persistent wave, for as long as it was feasible to do the numerical calculation. This is the crack front wave.

Returning to the singular crack model, the existence of the wave was proven, as discussed, by application [7] of the solution [6] for arbitrary, but sufficiently small (linearized) perturbation $A(z, t) \equiv a(z, t) - v_0t$. Writing the critical fracture energy as

$$G_{\text{crit}}(x, z) = G_{\text{crit},0} + \Delta G_{\text{crit}}(x, z),$$

(5)
and understanding that its perturbation at the crack front is $\Delta G(z, t) \equiv \Delta G_{\text{crit}}(a(z, t), z) = \Delta G_{\text{crit}}(v_0 t, z)$ within a strictly linearized formulation, one must have an expression of the form

$$\frac{\Delta G(k, \omega)}{G_{\text{crit}, 0}} = P(k, \omega) \tilde{A}(k, \omega),$$

where

$$[\tilde{A}(k, \omega), \Delta G(k, \omega)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{-i(kz+\omega t)} [A(z, t), \Delta G(z, t)] \, dz \, dt.$$  

The function $P(k, \omega)$ has a simple zero [7, 15] at a certain real value of $\omega/k$. That proves the existence of the wave, whose speed in the direction parallel to the moving crack front is $\omega/k$.

Figure 3 shows the resulting wave speed $c_f$, measured relative to a place on the fracture surface from which the wave was launched, or

![](image)

*Figure 3* Crack front wave speed, as a function of unperturbed crack speed $v_0$, for different values of Possion ratio $\nu$. From Ref. [15].

through which it passed; $\sqrt{c_f^2 - v_0^2}$ is the speed $\omega/k$ parallel to the front. The speed is very near to $c_R$, and consistent with the simulation result in Fig. 2. By analyzing the small perturbation problem for the case
in which $G_{\text{crit}} = G_{\text{crit}}(v)$, it was shown [7] that the wave attenuates if $dG_{\text{crit}}/dv > 0$.

The existence of the crack front wave leads to rapid disordering in crack propagation through a small random fluctuation $\Delta G_{\text{crit}}(x, z)$, about mean $G_{\text{crit},0}$, of fracture energy on the plane $y = 0$. Suppose that the variation is statistically stationary and isotropic, with correlation function

$$C\left(\sqrt{x^2 + z^2}\right) = \frac{\langle \Delta G_{\text{crit}}(x', x, z', z) \Delta G_{\text{crit}}(x^t, z') \rangle}{G_{\text{crit},0}^2},$$

and that a sample of that variation is first encountered by the crack at $t = 0$, with the crack front being straight and moving at unperturbed velocity $v_0$ for $t < 0$. Within the linearized perturbation formulation, $v_0$ remains the mean velocity after perturbation begins. We consider the random variables $A(z, t)$ and the slope $S(z, t) = \partial a(z, t)/\partial z$ of the crack front. Then, using the method of Ref. [5] for the scalar case, applied in way outlined in Ref. [9], one derives (for $v_0 t \gg$ length scales in the correlation function) that

$$\langle [A(z', t)]^2 \rangle \rightarrow F\left(\frac{v_0}{c_s}\right) c_s t \int_0^\infty C(w) \, dw,$$

and that

$$\langle A(z', t) A(z' + z, t) \rangle \rightarrow F\left(\frac{v_0}{c_s}\right) c_s t \int_0^\infty C\left(\sqrt{\frac{v_0^2 z^2}{c_f^2} + w^2}\right) \, dw,$$

where $F(v_0/c_s)$ is a certain function. Thus the variance in crack location grows in direct proportion to distance of propagation into the nonuniform region, whereas the two-point correlation between the positional fluctuations will vanish when $v_0 z/c_f$ has become large enough that there is zero correlation at such distances. This conveys a picture of rapid disordering of the crack front. That is also expressed by the associated result

$$\langle [S(z', t)]^2 \rangle \rightarrow F\left(\frac{v_0}{c_s}\right) \left(\frac{v_0}{c_f}\right)^2 c_s t \int_0^\infty \frac{1}{w} \left(\frac{-dC(w)}{dw}\right) \, dw.$$  

Note that such second order statistic will not exist if the correlation function has a non-zero slope at $0^+$ separation.

Samples of perturbed crack growth histories, within the linearized ideally elastic framework show that strong perturbations cluster in space and time along wave fronts [15]. Many features remain to be understood. These include saturation due to nonlinearities (one must, of course, have $0 \leq v \leq c_R$), damping due to $dG_{\text{crit}}/dv > 0$ and to material viscoelasticity, and, most especially, to interactions with non-planarity of growth.
3. LIMITING RUPTURE SPEED

The theoretical prediction, according to the singular crack model, for ruptures along a plane has already been mentioned. Under typical remote loading conditions we expect that \( v \to c_R \) for modes I and II, and that \( v \to c_s \) for mode III, as limiting rupture speeds.

However, to summarize observations [16]–[18] for tensile (mode I) cracks, from laboratory tests:

- The crack speed \( v \) in brittle amorphous solids (glass, PMMA) has an upper limit of order 0.50–0.60 \( c_s \) (i.e. 0.55–0.65 \( c_R \)).

- The fracture surface is mirror-smooth only for \( v < 0.3–0.4 \ c_R \). The crack surface roughens severely, and \( v \) becomes severely oscillatory, at higher (average) speeds. The crack forks at the highest speeds.

- There are exceptions for which \( v \to c_R \) (or a large fraction of \( c_R \)); brittle highly anisotropic single crystals (W, mica); and incompletely sintered solids [19].

Up to very recently, the only observations for shear ruptures (modes II and III) have come from inversions of seismograms for slip histories on faults:

- The range \( v \approx 0.7–0.9 \ c_s \) is a commonly inferred range of rupture speeds, although it is not well constrained.

- Rarely, bursts of intersonic rupture with \( c_s < v < c_p \) have been inferred [20, 21].

In the case of tensile cracks in brittle materials, interferometric studies [16] of propagating fractures in PMMA sheets showed that the propagation process became quite intermittent as the limit speed was approached. That was evidenced by strong stress waves being radiated from the tip in distinctly separated pulses. (It is actually the mean value of \( v \), averaged over a few such pulses, and not necessarily the local \( v \) itself, which should be thought to have a well defined limit, much less than \( c_R \).) Later studies [17] clearly tied the strongly intermittent \( v \), and intensity of fluctuation of \( v \) about the mean, to a side branching process. In that, small fractures seem to have emerged out of the walls of the main fracture near to its tip, such that the main crack and one or a pair of such side-branching cracks coexisted for a time, although the branches were ultimately outrun by the main fracture, at least at low mean propagation speeds. It may instead [22] be that the side branched cracks begun their life as part, not well aligned with the main crack plane, of a damage cluster of microcracks developing ahead of the main crack tip.
(Fig. 4(c)), and grew into the main crack walls. In any event, the density of cracks left as side-branching damage features increases substantially

with increase of the loading that drives the fracture. That process also results in the well-known increased roughening of the fracture surface with increasing crack speed. Ultimately, the response of the fracture to further increase in loading is no longer to increase the average $v$, which is the result expected from the theoretical analysis of a crack moving on a plane, but rather to absorb more energy by increasing the density of cracks in the damage cluster formed at the tip of the macroscopic fracture. Thus the energy adsorption rises steeply with crack speed [17].

The deviation of the fracture process from a plane is tied yet more definitively to the intermittency and low limiting crack speeds by studies [19] of weakly sintered plates of PMMA. The resistance to fracturing along the joint was much smaller than for the adjoining material, and that led to smooth propagation of the fracture (no evidence of jaggedness of interferograms by stress wave emission), with $v$ reaching $0.92 c_R$.

The attempts to explain low limiting speeds and fracture surface roughening begin with the first paper, by Yoffe [23], in which the elastodynamic equations were solved for a moving crack. That was done for a crack that grew at one end and healed at the other, so that the field depended on $x - vt$ and $y$ only (2D case), but sufficed to reveal the structure of the near tip singular field, including the $v$ dependence of the $\Sigma_{\alpha\beta}$ functions above. Yoffe found that $\sigma_{\theta\theta}$ (Fig. 4(a)), at fixed small $r$ within the singularity-dominated zone, reached a maximum with
respect to \( \theta \) for \( \theta \neq 0 \) when \( v > 0.65 c_R \). That gives a plausible explanation of limiting speeds and macroscopic branching, but does not explain the side-branching and roughening that sets in at much smaller speeds.

Eshelby [24] posed the question of what is the \( v \) at which a fracture must be moving so that, by slowing down, it can provide enough energy to drive two fractures (Fig. 4(b)). The extreme is to slow down to \( v' = 0^+ \). That question still has no precise answer when the angle \( \phi \) (Fig. 4(b)) is non-zero, but Eshelby answered it for mode III—Freund [12] had not yet solved the mode I case—in the limit \( \phi \to 0 \), which reduces to sudden change of speed of a fracture moving along the plane. The condition is \( g(v) < 1/2 \) (assuming no \( v \) dependence of \( G_{\text{crit}} \)), which gave \( v > 0.60 c_s \approx 0.65 c_R \) for mode III. Freund addressed the same issue based on his mode I solution [12], again in the limit \( \phi \rightarrow 0 \); that, with recalculation [25], leads to \( v > 0.53 c_R \). By working out the static version of the \( \phi \neq 0 \) problem, which has at least approximate relevance since \( v' \) is near 0, and using approximate considerations of wave travel times to limit the interaction of one branch with the other, Adda-Bedia and Sharon [25] optimize with respect to \( \phi \) to estimate that the lowest \( v \) for branching is \( \approx 0.50 c_R \). Material response with \( dG_{\text{crit}} / dv > 0 \) will reduce that threshold but, at present, it seems unlikely that the inferred onset of roughening at speeds in the range 0.3–0.4 \( c_R \) can be explained by a branching instability at a crack tip in a solid that is modeled, otherwise, as a linear elastic continuum.

Nevertheless, it is interesting that some discrete numerical models of cracking do give branching at roughly realistic speeds. A transition to zig-zag growth was shown to set in around 0.33 \( c_R \) in molecular dynamics simulations [26]. It is not yet clear to what extent the \( v \) at its onset may depend on details of the force law. Cohesive finite elements [27] or cohesive element interfaces [28] that fully model the separation process allow, within constraints of the mesh, for self-chosen fracture paths, and have shown off-plane fracturing. Xu and Needleman [28] nicely reproduce the observations that when the crack is confined to a weak plane, it accelerates towards \( c_R \), whereas for the uniform material, side branches form at 0.45 \( c_R \) or at lower speed [29] if statistical variation in cohesive properties is introduced. Recent work [30] suggests that such propensity for low-speed side branching may not be universal for all cohesive finite element models. In particular, models that involve linear behavior until a strength threshold is reached [31], after which there is displacement-weakening (much like for the inset in Fig. 1) seem less prone to development of side-branching. The procedure of Ref. [28] allows substantial nonlinear deformation at the element boundaries before achieving their peak cohesive strength. Thus local nonlinear features of the pre-peak
deformation response may be critical to understand the onset of roughening [32].

In contrast to tensile cracks, earthquake ruptures (large scale shear cracks) do seem to approach much more closely to $c_{\text{in}}$ and, as mentioned, bursts of intersonic rupture have been inferred [20, 21]. In fact, theoretical prediction [33, 34] of intersonic rupture from rupture simulations with slip-weakening models (mode II or III versions of the tensile displacement-weakening model discussed above) preceded the observations. A concentration of stress in a shear wave peak develops ahead of a mode II shear rupture. In a model with a finite strength to get slip started, that allows for the possibility that slip-weakening will initiate ahead of the main crack tip. As part of the linking up with the main slip-weakening zone, that allows a fracture to emerge at a high speed in the range of $\sqrt{2} c_s$ (which is the unique intersonic speed at which a mode II rupture can propagate in the singular crack model [18]). Such had never been seen in the laboratory until Rosakis et al. [35] demonstrated that for weakly sintered Homalite-100 plates, impact loaded in shear, a mode II fracture propagated intersonically with a speed that fluctuated to high values, near to $c_p$, but approached at greater propagation distances a speed near $\sqrt{2} c_s$. The importance of the weak channel for the rupture is that typical attempts to form a mode II (or III) rupture in the laboratory, except under quite high pressure [36], lead to mode I cracking from the rupture tip.

4. **MODE OF RUPTURE ON FAULTS: ENLARGING SHEAR CRACK OR SELF-HEALING SLIP PULSE WHEN THERE IS VELOCITY-WEAKENING FRICTION?**

In an influential paper, Heaton [37] argued that the mode of rupture in large earthquakes, as inferred from seismic inversion studies for well recorded events, was such that the duration of slip at a point on the fault was much shorter than the overall duration of the rupture. He argued that a point on the fault begins to slip as the rupture front arrives but that this slipping phase is of short duration and that the fault “heals” (by which is meant that it stops slipping) after a time that is much less than the overall duration of the rupture. That is called the “self-healing” rupture mode.

That is to be contrasted with what may be called the “crack-like” mode. For it, a point on the fault plane is assumed to slip for a significant fraction of the overall rupture duration, i.e. beginning when the rupture
front arrives and continuing until waves generated in the arrest of the front, at a barrier of some sort, carry back signals to stop slipping. The crack mode has been widely observed in numerical simulations of spontaneous rupture. These include cases of prescribed uniform strength drop on the fracture surface in singular crack models [38][41], non-singular slip-weakening models [33, 42, 43, 44], and simple, if grid-size dependent, models with a critical stress failure condition at the rupture front [45][47].

Heaton’s paper thus went against a widely accepted view of how rupture occurs. It launched an active body of theoretical research, to understand what does indeed determine the mode of rupture, and what could lead to the self-healing mode, which seems to be supported by observations. There are now a few candidates [48], one to be discussed in this section and one in the next. Here we focus on recent theoretical understanding of a possibility already suggested by Heaton, namely that strong velocity-weakening of friction could allow self-healing. In fact, simulations of rupture with strong velocity weakening sometimes showed crack-like rupture and sometimes self-healing [1, 49, 50], and it has been an important goal to understand what controls. One requirement is that the friction strength increase with time on slipped portions of the fault that are momentarily in stationary contact [1]. Another is that the overall driving stress be low [48] in a way that we analyze here.

Velocity-weakening friction on faults is interpreted in the rate and state framework, which includes laboratory-based state evolution features. Those also regularize ill-posed or paradoxical features of models of sliding between two identical elastic solids [51]. Thus, as in Fig. 5, we shall think of the heavy solid line as giving the basic response \( \tau = \tau_{ss}(V) = \sigma f_{ss}(V) \) between shear stress \( \tau \) and slip rate \( V \); here we regard normal compressive stress \( \sigma \) as constant. A full constitutive description, which must be used in numerical simulations, involves strength expressed as \( \tau = \sigma f(V, \theta) \) where the state variable is \( \theta \)—e.g. representing lifetime of current population of asperity contacts, or of current gouge packing—and where \( \partial f(V, \theta) / \partial V \geq 0 \) (note path of instantaneous change, i.e. change at constant \( \theta \), in Fig. 5), and \( \partial f(V, \theta) / \partial \theta \geq 0 \). The state variable follows an evolution law, e.g. of form \( \theta / dt = 1 - V \theta / L \), although other forms are sometimes used instead [1, 48]. The characteristic slip distance \( L \) for state evolution is typically found to be in the 1-50 \( \mu \)m range. The evolution law has the property that \( \theta \rightarrow \theta_{ss}(V) (= L/V \) for the law above) in sustained slip at fixed \( V \), so that \( \tau \rightarrow \tau_{ss}(V) = \sigma f_{ss}(V) \), where \( f_{ss}(V) = f(V, \theta_{ss}(V)) \) and, for the cases of interest here, \( df_{ss}(V)/dV < 0 \).

That is a somewhat complicated formulation, and the state evolution is over in micron range changes of slip, which would be very small
for applications to tectonic faulting. So the question arises, why use rate- and state-dependent friction instead of a pure rate-dependent friction law, say, $\tau = \tau_{\text{ss}}(V)$? The answer is that in addition to attaining consistency with laboratory evidence and with the microphysical understanding of friction, we eliminate the following [51]:

- Ill-posedness of the pure rate-dependent formulation when $-\infty < \frac{d\tau_{\text{ss}}(V)}{dV} < -\frac{\mu}{2c_s}$. An $e^{ikx}$ perturbation of a steady sliding state, with $V = V_0$, then elicits response $V(x,t) - V_0 \sim e^{ikx}e^{\alpha|yc_s|}$, $\alpha > 0$; see the discussion of a similar issue in the next section.

- Supersonic propagation of rupture fronts [52] when $-\frac{\mu}{2c_s} < \frac{d\tau_{\text{ss}}(V)}{dV} < 0$. An $e^{ikx}$ perturbation elicits response $V(x,t) - V_0 \sim e^{ik(x+\tau)}$ where $r > c_p$ for mode II and $r > c_s$ for mode III slip. (While supersonic propagation of rupture fronts seems to be precluded in the rate and state formulation, phase velocities at sufficiently low $|k|$, within the range for which there is unstable response to $e^{ikx}$ perturbations, do become supersonic [53]).

These shortcomings of the pure rate-dependent model are not widely known, probably because friction studies directed to machine technology have often focused on sliding rigid blocks rather than deformable elastic continua. The rate and state formulation provides a regularization.

Now consider a fault surface, which we treat as the boundary $y = 0$ between two identical half spaces (Fig. 6). An initial shear stress $\tau_0(x,z) = \tau_0^b$, a constant level too small to cause failure, acts everywhere on $S_{\infty}$ (the $xz$ plane) except in small nucleation region $S_{\text{nuc}}$, which will
be overstressed to start the rupture. An important stress level is given by $\tau_{\text{pulse}}$ as marked in Fig. 5. It is understood as the highest value that $\tau^b_0$ could have if we were to require $\tau^b_0 - \mu V/2c_s \leq \tau_{\text{ss}}(V)$ for all $V \geq 0$.

Suppose that $\tau^b_0 < \tau_{\text{pulse}}$. As will be seen, that effectively precludes the possibility that rupture could occur on $S_{\infty}$ in the form of an indefinitely expanding shear crack [48]. Note that

$$\tau^b_0 < \tau_{\text{pulse}} \Rightarrow \tau_{\text{ss}}(V) - (\tau^b_0 - \mu V/2c_s) > 0 \quad \text{for all } V \geq 0. \quad (12)$$

Important along the way is an elastodynamic conservation theorem [48]

$$\iint_{S_{\infty}} \left[ \tau(x, z, t) - \tau_0(x, z) + \frac{\mu V(x, z, t)}{2c_s} \right] \, dx \, dz = 0, \quad (13)$$

which holds throughout the rupture. This can be derived from relations involving spatial Fourier transforms of elastodynamic fields [2] in the zero wavelength limit. As an aside, it provides an interpretation for the seismic moment release rate,

$$\frac{dM_0(t)}{dt} \equiv \mu \iint_{S_{\infty}} V(x, z, t) \, dx \, dz = 2c_s \iint_{S_{\infty}} [\tau_0(x, z) - \tau(x, z, t)] \, dx \, dz, \quad (14)$$

which has apparently not appeared before in seismology.

Let us now assume that with the loading $\tau^b_0 < \tau_{\text{pulse}}$, rupture has been locally nucleated and grows on $S_{\infty}$ in the form of an indefinitely expanding shear crack. We shall try to develop a contradiction. The integrand everywhere along the rupturing surface $S_{\text{rup}}(t)$, except for $S_{\text{nuc}}$ and for small regions at the rupture front affected by the rate/state regularization, is equal to

$$\tau_{\text{ss}}(V) - \tau^b_0 + \frac{\mu V}{2c_s} = \tau_{\text{ss}}(V) - \left( \tau^b_0 - \frac{\mu V}{2c_s} \right) > 0, \quad (15)$$

where the inequality follows from the above consequence of $\tau^b_0 < \tau_{\text{pulse}}$. Thus, letting $S_{\text{out}}(t) = S_{\infty} - S_{\text{rup}}(t) = \text{region of } S_{\infty} \text{ lying outside the}
rupture at \( t \), and noting that \( V = 0 \) there, we must then have

\[
\iint_{S_{\text{net}}(t)} (\tau - \tau_0^b) \, dx \, dz < 0 \quad \text{(and \( \rightarrow -\infty \) as \( S_{\text{rupt}} \rightarrow \infty \))}.
\]

That seems to be a contradiction: We expect cracks to increase the net force carried outside themselves, or at least to not decrease it. (However, such non-decrease of force is a proven result, thus far, only for 2D anti-plane shear crack solutions [48].) The result does nevertheless suggest that the crack-like rupture mode should not occur if \( \tau_0^b < \tau_{\text{pulse}} \). That is consistent with a range of calculations [48, 54] for different stress levels and forms of \( \tau_{SS}(V) \). Crack-like ruptures are not found—only self-healing slip pulses—when \( \tau_0^b < \tau_{\text{pulse}} \). In the loading regime \( \tau_0^b > \tau_{\text{pulse}} \), one may focus on a dimensionless measure \( T \) of the strength of the velocity weakening at a representative slip rate [48]. When \( T \) is very small, the rupture surface could reasonably be expected to respond somewhat like that for a model with constant stress drop, which is in a crack-like mode; that is what is found in simulations.

5. MATERIAL PROPERTY DISSIMILARITY ACROSS A FAULT PLANE

Here we consider two dissimilar solids in sliding frictional contact of mode II type (Fig. 7). We consider first the simplest case of a constant friction coefficient \( f \), independent of slip rate \( V \) or its history. Thus

\[
\text{Slower wave speed, } c_{s1} < c_{s2}, \text{ etc.}
\]

\[
\text{Slip } = \delta
\]

\[
\sigma(x,t) \quad \tau(x,t)
\]

\[
\delta(x,t)
\]

\[
\text{Figure 7 Frictional sliding of dissimilar elastic solids on one another.}
\]

whenever \( V(x,t) > 0 \) and normal stress \( \sigma(x,t) \geq 0 \), the shear stress \( \tau(x,t) = f\sigma \). The general effect that can drive highly unstable response in this case is that spatially inhomogeneous slip \( \delta(x,t) \) induces a local change in \( \sigma(x,t) \), and such reduction of \( \sigma \) allows easier slip.
Interesting results have been found in the following two types of analyses:

- Dynamic stability of steady sliding [55]-[59] on the surface between dissimilar elastically deformable materials, loaded with \( \tau_\infty = f \sigma_\infty \)

- Slip rupture dynamics on faults between dissimilar materials that are loaded below threshold [60]-[65], \( \tau_\infty < f \sigma_\infty \)

For the first type of problem, we assume that the two solids are sliding at a uniform rate \( \frac{\partial \delta}{\partial t} = V_0 > 0 \) and are loaded by remote stresses that just meet the sliding conditions, \( \tau_\infty = f \sigma_\infty \), and that the stress state is uniform within each solid. We then consider the response of this system to perturbation, such that the friction law \( \tau = f \sigma \) remains valid everywhere along the interface and there is no opening. Solutions to the Navier equations are then sought in the form

\[
\begin{align*}
  u_x &= \frac{1}{2} \text{sgn}(y) V_0 t + \Re \left[ U_x(y) e^{i k x + p t} \right], \\
  u_y &= \Re \left[ U_y(y) e^{i k x + p t} \right].
\end{align*}
\]

(17)

For real \( k \), one finds that \( p \) is of the form \( p = a|k| - i c k \), where \( a \) and \( c \) are real. Most importantly, over a wide range of parameters [56] (values of \( f \), ratios of elastic constants, and densities), it turns out that \( a > 0 \) for at least one such solution.

That remarkable fact not only implies that the dissimilar material system is modally unstable, but also implies ill-posedness of response to general perturbation. That is, the response to a perturbation that has Fourier strength \( \overline{W}(k) \) in that unstable mode is, formally,

\[
\begin{align*}
  V(x, t) - V_0 &\sim \int_{-\infty}^{+\infty} \overline{W}(k) e^{a|k| t} e^{i k (x - \xi) t} \, dk,
\end{align*}
\]

(18)

which diverges when \( t > 0 \) for generic \( \overline{W}(k) \) if \( a > 0 \). Response in a given \( e^{i k x} \) mode does not diverge in finite time, but the solution ultimately loses validity because either \( V < 0 \) or \( \sigma < 0 \) is predicted.

The following has been shown [59] for dissimilar materials:

- Response is unstable \((a > 0)\) for all \( f > 0 \) when a generalized Rayleigh (GR) wave\(^1\) exists, and for \( f > f_{\text{crit}} > 0 \) when a GR wave does not exist.

- For sufficiently large \( f \), two unstable solutions (i.e. both with \( a > 0 \)) may exist, with different growth rates \( a \) and propagation

\(^1\)A GR wave corresponds to a motion with free slip, \( \tau = 0 \), but no opening gap at the interface. It exists when there is a real solution \( c \) (the wave speed \( c_{GR} \)) to \( \mu_1 c_{GR}^3 + R_1 |c| \gamma_2 |c| + \)
speeds \( c \) of different magnitudes and signs (directions), and neutrally stable solutions \((\alpha = 0)\) with supersonic propagation speeds may exist.

The second type of problem envisions that the two solids in contact (Fig. 7) are loaded below the friction threshold, \( \tau_\infty < f\sigma_\infty \), and that an event is somehow nucleated (e.g. by local reduction of \( \sigma \)) and the focus is on how it propagates into previously locked locations along the interface. Weertman [60] suggested that steady rupture propagation, \( \delta = \delta(x - vt) \), in a form of what is now called a self-healing pulse, might exist in that circumstance, and such has been confirmed [61]. That solution exists only when the GR wave exists. Rupture propagates at \( v = c_{GR} \) (its possible significance for rupture propagation had been noted earlier [67]), in a unique direction along the interface, which is that of slip in the slower material—i.e. in the +\( x \) direction in Fig. 7, for slip as shown. The solution has the remarkable feature that \( \tau \) remains unaltered from \( \tau_\infty \), and that slip occurs because \( \sigma \) reduces in the slipping region so as to just meet \( \tau_\infty = f\sigma \). The speed \( V \) is constant within the pulse, and proportional to \( f\sigma_\infty - \tau_\infty \). Its stability and possibility of emergence from initial conditions remain unclear.

Transient dynamic analysis of locally nucleated rupture has also been addressed, first in finite difference simulations [62]-[64]. They did indeed find rupture in the form of a self-healing pulse but results indicated problems of convergence with grid refinement, or with time of rupture propagation, which is now understood to be related to the ill-posedness mentioned above. That ill-posedness was confirmed [65] by using the spectral numerical formulation, as generalized to bimaterials [3]. For a fixed replication distance along strike, the numerical results became progressively more jagged with increase from 256 to 2048 terms in the underlying Fourier series representation of \( \delta(x, t) \), in a way that was consistent with theoretical understanding of the instability.

Some possible regularizations of the problem have been discussed [58], e.g. basing the friction law on the average of \( \sigma \) over a finite patch size around the position of interest. Another approach [59, 65] was based

\[
\mu_2 c_{\alpha_2}^2 R_2(c) \gamma_1(c) = 0 \text{, where}
\]

\[
R_k(c) = 4 \sqrt{\left(1 - \frac{c^2}{c_{pk}^2}\right) \left(1 - \frac{c^2}{c_{\alpha_k}^2}\right) - \left(2 - \frac{c^2}{c_{\alpha_k}^2}\right)^2}, \quad \gamma_k(c) = \sqrt{1 - \frac{c^2}{c_{pk}^2}},
\]

with \( k = 1, 2 \) to denote the different materials. GR waves were discovered independently by several investigators [66]-[68]. They satisfy \( \min(c_{R_1}, c_{R_2}) < c_{GR} < \max(c_{R_1}, c_{R_2}) \) and \( c_{GR} < c_{\alpha_1} < c_{\alpha_2} \), and exist only for modestly different materials, typically for \( c_{\alpha_2} < 1.30-1.35 c_{\alpha_1} \).
on results of oblique shock wave experiments [69, 70]. In those, a shock reflected from the back surface of the target specimen causes an abrupt decrease of normal stress on the sliding interface. As shown schematically in Fig. 8, there is then only a gradual evolution of shear strength (over a few micrometers of slip, or few tenths of microseconds time)

![Graph showing stress vs. slip](image)

*Figure 8  Schematic of gradual evolution of shear strength (over a few micrometers of slip, or few tenths of microseconds time) in oblique shock wave experiments of Refs. [69, 70]. A shock reflected from the back surface of the target specimen causes an abrupt decrease of normal stress.*

...towards the new level consistent with the altered normal stress. A simple representation of this result is to replace \( \tau = f \sigma \) with [59]

\[
\frac{d\tau}{dt} = -\frac{V}{L}(\tau - f \sigma),
\]

assuming \( V > 0, \sigma > 0 \), where \( L = \text{const} > 0 \). This regularizes the first group of problems above, i.e. \( e^{ikx} \) perturbation of steady frictional sliding, as follows [59] (now \( a \) and \( c \) in \( p = a|k| - ick \) do depend on \( k \), at least when \( k c_{\lambda L} L/V \) is of order 1 or larger):

- When a GR wave exists, the unstable response approaches neutral stability, \( a \to 0 \), as \( |k| \to \infty \) (i.e. as wavelength \( \lambda \to 0 \)). In practical terms, that means stability at length scales \( \lambda \ll 2\pi c_{\lambda L}/V \).

- When a GR wave does not exist, there is a critical wavenumber \( k_{cR} \) such that response is stable \( (a < 0) \) when \( |k| > k_{cR} \) (or when \( \lambda < \lambda_{cR} \), where \( \lambda_{cR} \) scales in proportion to \( 2\pi c_{\lambda L}/V \).) Nothing in those results would change if \( V/L \) was replaced by \( a + \beta V \), where \( \alpha, \beta \geq 0 \) with at least one \( > 0 \), or indeed by any positive factor.

The same regularization was then used [65] to address the second class of problems above, systems that are frictionally locked with \( \tau_\infty < f \sigma_\infty \), but for which rupture is nucleated somewhere by local unclamping. With
the regularization, the problem becomes well posed, in that convergent numerical solutions occur upon refinement of the Fourier basis set. Further, the rupture is in the form of a self-healing pulse propagating at a speed close to $c_{GR}$. In cases for which $c_{GR}$ does not exist, the propagation speed seems to be close to $c_4$.

These are all problems of slip rupture along bimaterial interfaces, but there are also challenging problems, with remarkable features seen experimentally and rationalized theoretically, involving combined tensile fracture and slip rupture along bimaterial interfaces. For example, for polymer–metal systems (PMMA–steel [71] and Homalite 100–aluminum [72], respectively), impact loading created fractures that began their life as shear slippage over a millimeter scale at the rupture front, and subsequently opened. Their propagation speed was faster than $c_3$ for the polymer, leading to shock wave structures in high speed photographs of a photoelastic fringe pattern [72]. The essential features of those complex near tip slip and opening features, and the propagation speed, were rationalized in spectral numerical calculations [3], based on a non-singular numerical model with combined displacement weakening in tension and shear.

6. CONCLUSION

It is hoped this set of short examples conveys some idea of the new discoveries and scientific excitement in understanding the dynamics of rupture. We have seen the discovery of previously unsuspected waves along crack fronts, of reasons why a previously inferred speed limit for fracturing (the Rayleigh speed) is sometimes too high and sometimes too low, of unsuspectedly ill-posed problems in frictional slip dynamics, and of rupture modes, such as the self-healing mode, that were largely unsuspected a relatively short time ago. These are a small sample from a large and vigorously growing body of science.

Acknowledgment

The studies discussed were supported by the Office of Naval Research, the U.S. Geological Survey, and NSF through the Southern California Earthquake Center.
References


New perspectives on crack and fault dynamics


