ANNALI DI GEOFISICA, VOL. XL, N. 5, October 1997

Noncausality of numerical models of dynamic fracture growth

Leon Knopoff
Department of Physics and Astronomy and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90095, U.S.A.

Abstract
Discretization of the wave operator for purposes of solving problems in the dynamics of crack growth numerically, introduces noncausality associated with the nonlinearity of the fracture criterion at the edge of the crack, *i.e.* with an imperfect formulation of the fracture criterion in the discretized case. The noncausality can be attributed to jumps from one inertial coordinate system to another as successive particles at the edge of a digitized crack are triggered into motion. It is shown that there is an equivalent explanation in terms of the incompatibility of the short-range edge conditions and the long-range correlations of slip on the crack in the discretized case. The noncausal effects can lead to supersonic crack growth, and in some cases to infinite crack growth velocities. A proposal for amelioration of the problem is offered.

Key words *nonlinearity – fracture – noncausality*

1. Introduction

The solutions to the problems of the dynamics of fracture and crack growth are fundamental to our understanding of the seismic source and as a consequence to their application to the understanding of the problems of strong ground motions in large earthquakes, as well as to the problems of evolutionary seismicity in an active seismic region. The mathematics of dynamic growth of a crack under conditions of brittle fracture on a pre-existing fault surface for simple geometries and for specific modes of crack motions, is a well-known nonlinear problem that has been described in the published literature for over 30 years (see, *e.g.*, Kostrov, 1966).

Mailing address: Prof. Leon Knopoff, Department of Physics and Astronomy and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90095, U.S.A.; e-mail: knopoff@physics.ucla.edu

A synopsis of the physics is as follows: a prestressed elastic solid has a pre-existing fault surface that is prevented from slipping by static friction. If the prestress is equal to the friction at some point along the surface, on the brittle fracture model the stress drops instantly to a low value and crack growth is initiated. The crack continues to grow as long as the stress concentration at the edge of the crack, when added to the prestress exceeds the friction. (Since the purpose of this paper is to explore a numerical condition at the edge of the growing crack, I will not undertake to discuss the problems of healing, *i.e.* the progressive cessation of slip or healing of crack motions, which are equally important for understanding crack slip histories as dynamical models of earthquake sources). The slip on the crack surface is the analytic continuation of the displacement in the interior of the solid, where the displacement satisfies the elastic wave equation; in other words, the slip gives rise to the vibratory motions in the solid that are the earthquake waves that are transmitted to remote points (see, *e.g.*, Burridge and Halliday, 1971). The stress drop on the fracture surface is the source that drives the crack and that generates the displacements in the elastic wave radiation field. Recent attention has focused on modifications of this formulation under assumptions that the failure of the frictional contact is not instantaneous, but instead that the bond strength undergoes progressive weakening during the dynamics.

Even for the simpler cases of instantaneous failure, the solution to this nonlinear problem as a problem in dynamics is extraordinarily difficult for arbitrary fault geometry, for an arbitrary geometrical distribution of friction at the edge, and for in-plane slips in the crack surface, which generate tensor stress fields. Solutions have been given for restricted cases of problems of linear or planar faulting in one- or two-dimensions, with specialized modes of slip (usually antiplane), and for specialized distributions of stress drop and of static friction at the edge of the crack.

The problems of understanding the conditions of growth at the edge of the crack are difficult; it is at the edge of the crack where the nonlinearity arises due to the fact that the edge is a moving boundary between the fractured and unfractured regions. In this paper I will show that discretization of the pde's of dynamic elasticity introduce issues of causality that need to be addressed for problems of instantaneous failure; these problems are equally important in cases of progressive weakening of the strength of the contact at the edges of discretized cracks. I restrict this discussion to the problems of instantaneous failure. Since this problem will arise at all levels of complexity of geometry, heterogeneity and mode of rupture, it suffices for the purposes of illustration to consider the simplest possible example which is that of a problem of fracture under conditions of homogeneous one-dimensional elasticity without taking into account the energy losses due to elastic wave radiation; as remarked, there will be no loss in the generality of understanding the difficulty if we consider cracks with higher dimensionality or with slip motions other than scalar, or if we consider problems of nonplanar geometry of fault surfaces, although the solution, in most of these cases will be more complex.

2. One-dimensional continuum crack

In the homogeneous one-dimensional case, the problem of the slip on the fault $U(x, t)$, is given by a solution to the inhomogeneous wave equation

$$\begin{aligned} \rho U_{tt}(x, t) - \mu U_{xx}(x, t) &= f(x), \qquad x < \xi(t) \\ U &= 0, \qquad x > \xi(t), \end{aligned} \tag{2.1}$$

with the conditions

$$\begin{aligned} \mu U_x(x, t) + \dot{\xi}\rho U_t(x, t) &= -g(x), \\ U_t + \dot{\xi} U_x &= 0, \end{aligned} \tag{2.2}$$

on the edge of the crack $x = \xi(t)$, which are appropriate in the inertial frame of the coordinate system x; see Knopoff *et al.* (1973) for a derivation of these two advection equations. In the moving frame of the crack edge, the fracture criteria are simply $\mu U_x(\xi(t)) = -g(\xi(t))$ and $U(\xi(t)) = 0$. Equation (2.1) is valid in the inertial frame. The usual elastic modulus and density are (μ, ρ), and we use the usual subscript notation for partial derivatives with respect to t and x. The velocity of elastic waves is $c = (\mu/\rho)^{1/2}$. The quantity $f(x)$ is the stress drop. The edge of the crack is at coordinate $\xi(t)$. In (2.2) we only consider the condition at the rightward travelling edge of the crack; the modification for the other edge is easily made. The quantity $g(x)$ is called the «excess force» and in the dynamic case is the difference between the bond strength and the prestress at the edge of the crack, in the appropriate units; *i.e.*, the quantity $g(x)$ is the force that must be added to the prestress at the crack edge that has to be supplied by the stress concentration at the edge of the crack to cause new material to break; $g(x) = 0$ at the point of initiation of fracture.

The quantities f and g have different dimensions: $f(x)$ is a stress, which in our one-dimensional system is a force per unit length, while $g(x)$ is a force. These definitions are unique to the one-dimensional case. In two- and three-dimensions, the quantity $g(x)$ will be a stress, having the same dimensions as $\mu\partial U/\partial x$, and

f will be a stress gradient. These differences disappear when the continuum is digitized onto a discrete lattice.

It is clear that eqs. (2.1) are linear; the nonlinearity in the problem is focused on eqs. (2.2). Thus this problem is a form of Stefan problem. A closed form solution to a restricted case of the nonlinear problem (2.1)-(2.2) is given by Knopoff *et al.* (1973).

By eliminating U_x or U_t from eqs. (2.2), it is easy to see that $\dot{\xi} \leq c$, and $\dot{\xi} = c$ only if $g(x) = 0$. Thus these cracks always grow subsonically if the excess force demanded for fracture is nonzero and they grow sonically if it is zero. Thus the conditions (2.2) represent a causality condition since these cracks can never propagate supersonically. There can be no supersonic crack-edge velocities in a perfectly elastic system because of the following physics-based argument: the dynamics of the rupture is controlled by the elastic waves that are excited by the slip on the crack which is the boundary to the elastic medium outside it; the elastic waves exist only in the elastic medium outside the crack and travel outward, away from the crack, with elastic wave velocity. The dynamic stresses in these elastic waves are the triggering agents for the growth of the cracks, *i.e.* the stress waves in the elastic medium provide the increment over the prestress that takes the force up to the level needed to break additional material at the edge of the crack. The greater the amount of excess stress at the edge of the crack, the slower the rate of growth of the crack. Since the maximum speed of stress wave propagation is c, these passive cracks cannot rupture with edge velocities greater than c. In general, these conclusions are unchanged for cases of two- or three-dimensional geometries; however the form of the conditions at the edges is no longer as simple as in this example; Chatterjee and Knopoff (1983) give an example in the two-dimensional antiplane case.

The conditions (2.2) show that immediately behind the crack edge the particle velocity U_t jumps instantly to a finite value from the zero value it has in front of the crack. The jump in slip velocity does not violate conditions for continuity of momentum since an infinitesimal element at the crack edge has infinitesimal mass.

3. Noncausality of the discretized crack

In order to solve the general problem $g(x) \neq 0$, $f(x) \neq$ constant, we can have recourse to the Green's function methods (see Chatterjee and Knopoff, 1983), or we can try to solve the system (2.1)-(2.2) numerically. In the latter case, the discretization of the spatial second-derivative operator $\partial^2/\partial x^2$ in eqs. (2.1), and more generally the operator ∇^2 arising from elasticity leads to the difference equations

$$m\ddot{U}_n - k\,[U_{n+1} - 2U_n + U_{n-1}] = f_n, \qquad (3.1a)$$

where we have set $k = \mu/a$, $m = \rho a$, $x = na$, and a is the lattice spacing or digitization interval. For convenience we rewrite (3.1a) in the obvious way

$$m\ddot{U}_n - k\,[U_n - U_{n-1}] + k\,[U_{n+1} - U_n] = f_n. \qquad (3.1b)$$

Equation (3.1b) exhibits the force balance of a system undergoing dynamic slip among a chain of particles with mass m interconnected in a linear array by a set of linear springs with spring constants k; we can place subscripts n on the masses and spring constants if they are spatially variable quantities. Stresses in the ruptured section are described by the system of difference eqs. (3.1a,b) and are transmitted along the ruptured segment with elastic wave velocities, albeit with dispersion due to the discretization; the dispersion is well-known (*e.g.*, Brillouin, 1946). (In the discretized system, the quantities f and g are both forces, the difference of the units of these quantities in the continuum having been accommodated by the lattice spacing a. Thus the critical fracture threshold of a lattice site is $(f+g)$, and the fracture needs to develop a force g or more to allow the fracture to grow.)

Our attention is focused on the last bracketed term on the left hand side of eqs. (3.1b). Assume that the discretized fracture has reached lattice site (or particle number) n and that particle $(n+1)$ has not yet begun to move. Consider the transmission of stresses from the last moving element, particle n, to the adjacent unbroken element, particle $(n+1)$. Then the force on particle $(n+1)$ at time t is $kU_n(t)$. The usual fracture criterion is then taken to be that rupture occurs at time t (measured from the time that the n-th particle began to move) such that

$$kU_n(t) = g_{n+1}. \tag{3.2}$$

The stress at particle $(n+1)$, which is $kU_n(t)$, no longer satisfies the wave equation, since the $(n+1)$st particle does not move and hence there are no inertial effects. We can restore the wave property if we suppose that there is a perfect image wave travelling leftward that arrives from sites to the right of site $(n+1)$, so that site $(n+1)$ is a nodal point in a wave system. But this would violate causality, since it would imply that sites to the right of $(n+1)$ have a foreknowledge of the stress system at times in advance of the arrival of signals from the left that can only travel with wave velocity c. The difficulty is that the particle at site n has stresses on it that do not satisfy the wave equation *within the framework of eqs.* (3.1a,b).

As each particle begins to move from rest under the sudden application of the force f, its velocity increases linearly with time because of its inertia, as all first-year physics students know. Thus the discrete system differs from the continuum markedly: the continuum begins to move initially with finite velocity while in the discrete system the particles begin to move quadratically. A given particle only approaches the constant velocity continuum limit after the edge has moved some lattice sites away.

Although the edge moves in jerks, we define the crack edge velocity to be the quotient of the lattice spacing and the time interval between initiation of motion at successive lattice sites.

The simplest possible example is an extreme case. Consider the homogeneous case in which f_n is the constant f_0, and let $g_n = \varepsilon \ll f_0$, except for $g_0 = 0$, which allows for initiation of rupture at $n = 0$. As we have seen, the edge of the fracture must travel with elastic wave speed in the continuum limit of $g(x) = 0$, *i.e.* $\varepsilon \to 0$, which is causally consistent.

In the discrete case, the situation is different: at the time that the n-th particle begins to move with quadratic increase of slip as a function of time, it transmits a force kU_n to the $(n+1)$st particle, thereby triggering it into motion at the same instant as well, since the excess-force threshold for rupture is negligibly small. Thus the crack edge in this discrete case travels with infinite velocity as $\varepsilon \to 0$, a result in violation of causality. The result is independent of the wave properties of the system, because any individual classical connecting spring transmits a force $k(U_n - U_{n+1})$ instantly without the phase delays implicit in the wave equation. The phase delays arise in the wave equation because of the long-range correlations implicit in eqs. (3.1a,b), even though only the equations describe the coupling between nearest neighbour particles. Thus the wave/phase delay is a property of the (slip) motions at wavelengths that are long compared with the lattice spacing, while crack growth is a short-wavelength property of the system. Thus the failure to be causal in this discretized case can be described as a consequence of the incompatibility of the long-range forces implied in the slip on a crack and the short-range fracture criterion.

If $g_n \neq 0$, the discretized crack grows at finite velocity, although this velocity may still be supersonic; in any case, the situation improves with increasing g_n, but there is always a noncausal term at the edge of the crack that influences the motion that is due to the discretization. In most cases $g_n \neq 0$, the slip of the particles leave their parabolic regime quickly and may reach their terminal velocity quickly; if the first moving particle behind the crack edge reaches its terminal velocity state after or about the time that the next particle begins to move, then these cases closely resemble the continuum.

4. Amelioration of the noncausality

The amelioration of the incompatibility between the long-range correlations of the slip and the short-range properties of the fracture condition at the edge in the numerical case is difficult. An approximate solution to our problem can be found in eqs. (2.2). On a continuum theory, the static criterion for fracture is

$$\mu U_x = -g(x), \tag{4.1}$$

which is the same as its discretized version, eqs. (4.1). (Recall that U_x is negative at the right hand edge of the crack.) The first of eqs. (2.2) shows that on the basis of a dynamic theory, the criterion is

$$\mu U_x(x, t) = -g(x) - \dot{\xi}\rho U_t(x, t). \tag{4.2}$$

In other words, the dynamic continuum theory states that the effective fracture strength is increased due to the advection of momentum across the crack boundary. To put it another way, the static theory (4.1) would have been perfectly applicable, had we evaluated the fracture criterion on the moving boundary, rather than on the fixed (inertial) coordinate system in which $U_n = 0$.

Let us eliminate the velocity $\dot{\xi}$ between the two equations of (2.2); we get an energy version of the fracture criterion

$$\mu U_x^2 = -g(x)\,U_x + \rho U_t^2. \tag{4.3}$$

This expression is a statement of energy conservation. Thus the strain energy in the elastic deformation field not only must be used to provide the energy needed to break the frictional bonds at the crack tip, but also it must supply the kinetic energy of motion of the sliding material just behind the crack tip. (This expression also has the form of a Lagrangian if we transpose the second term on the right hand side of (4.3) to the left; this not unexpected if we appreciate again that (4.3) is an expression of the energy balance in the stationary and not in the moving coordinate system). Thus we have two forms of the fracture criterion, one being a critical stress condition (4.2) and the other being a critical energy criterion (4.3). These are identical fracture criteria in the dynamic regime, but have often been identified as being at odds with one another for static cracks.

The problem of the noncausality in the discretized case can be restated: since the edge moves ahead by jerks in the discretized case, the edge of the crack is at rest most of the time in the inertial coordinate system, but the edge advances abruptly to a new inertial coordinate system every time a new particle is promoted to the edge of the crack. Since the second of eqs. (2.2) is always satisfied at the edge of the crack, we only need consider the first of these equations. We rewrite the edge condition (4.2) in a form appropriate to the local or discrete system,

$$kU_n - \frac{\dot{\xi}}{c^2}\langle \dot{U}_n \rangle = g_{n+1}, \tag{4.4}$$

where the second term on the left is the dynamic correction to the usual static stress rule; the quantity in angular brackets is the value of the particle velocity averaged over the time since initiation of motion of the n-th particle. Since $\dot{\xi} = a/t$, where t is the time of fracture of the $(n+1)$st particle, $\langle \dot{U}_n \rangle = U_n/t$, and $c^2 = \frac{k}{m}a^2$, the result is

$$k\left(1 - \frac{a^2}{c^2 t^2}\right)U_n = g_{n+1}. \tag{4.5}$$

This fracture criterion model (4.5) is a satisfactory solution to the problem of noncausality, since it guarantees that the rupture velocity cannot be greater than c, even in the extreme case $g_n = 0$. Unfortunately (4.5) does not give the correct rupture velocity, even though it is always subsonic, for cases $g \neq 0$. This is due to dynamic corrections that must be introduced into (4.5) that arise from the failure of the motion of the n-th particle to satisfy the wave equation for fixed $(n+1)$st particle. These dynamic effects are manifested as lattice oscillations; although these oscillations are strongly damped in the interior of the growing crack, nevertheless they have a profound influence on the rate of growth of the crack. This important issue will be discussed in a second paper on this subject.

Unfortunately, we do not have similar analytic expressions as (4.3) for the edge condition for more complex geometries such as those of in-plane fractures in three dimensions. The one-dimensional case is the only one having a local fracture criterion; all others involve long-range interactions because of the nature of the Green's function in several dimensions. For example, the fracture condition in two-dimensions for antiplane motions is known for general $(f(x), g(x))$ (Chatterjee and Knopoff, 1983), and includes the effects of the long-range terms; an application of the local model (4.5) to discretized 2D antiplane cases gives ambiguous results, even though of necessity, they do not have supersonic edge velocities.

An application of model (4.5) to the problems of two-dimensional in-plane fractures, with the imposition of a local fracture criterion, gives the reasonable result that cracks that grow in the direction of slip have edge velocities that are bounded by the P-wave velocity, while growth in the crack plane perpendicular to the slip is bounded by the S-wave velocity. Details of this application are not given in this paper.

Acknowledgements

I have had the opportunity to try out some of these ideas with H.-J. Xu, M.S. Abinante, and X.X. Ni and I am indebted to them for their willingness to discuss them with me. This paper is Publication No. 4912 of the Institute of Geophysics and Planetary Physics, University of California, Los Angeles.

REFERENCES

BRILLOUIN, L. (1946): *Wave Propagation in Periodic Structures* (McGraw-Hill Book Co., New York), 31-33.

BURRIDGE, R. and G.S. HALLIDAY (1971): Dynamic shear cracks with friction as models for shallow focus earthquakes, *Geophys. J. R. Astron. Soc.*, **25**, 261-283.

CHATTERJEE, A.K. and L. KNOPOFF (1983): Bilateral propagation of a spontaneous two-dimensional anti-plane shear crack under the influence of cohesion, *Geophys. J. R. Astron. Soc.*, **73**, 449-473.

KNOPOFF, L., J.O. MOUTON and R. BURRIDGE (1973): The dynamics of a one-dimensional fault in the presence of friction, *Geophys. J. R. Astron. Soc.*, **35**, 169-184.

KOSTROV, B.V. (1966): Unsteady propagation of longitudinal shear cracks, *J. Appl. Math. Mech.*, **30**, 1241-1248.