INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/650-4

Workshop on Three-Dimensional Modelling
of Seismic Waves Generation
Propagation and their Inversion

30 November - 11 December 1992

1. *Seismic Source Models I*
2. *Seismic Source Models II*
3. *The Inverse Problem for the Seismic Source*

S. Das

University of Oxford
Department of Earth Sciences
Oxford, U.K.

Strada Costiera, 11 Tel. 22401 Telefax 224163 Telex 460392 Adriatico Guest House Via Grignano 9 Tel. 224241 Telefax 224531 Telex 460449

Lecture Notes for lectures by S. Das on

1. **Seismic Source Models I**
2. **Seismic Source Models II**
3. **The Inverse problem for the seismic source**

Three-Dimensional Modelling of Seismic Waves Generation,
Propagation and their Inversion".

International Centre for Theoretical Physics

Trieste, Italy

30 November 1992 - 11 December 1992

These lectures assume a knowledge of the Theory of Elasticity and some knowledge of seismology. A good mathematics background is essential.

The lecture notes are extracted from the following sources:

Aki, K. and P.G. Richards (1980) Quantitative Seismology: Theory and Methods. 932 pp., W.H. Freeman and Company, San Francisco, California, 1980.

Kostrov, B.V. and S. Das (1988) Principles of Earthquake Source Mechanics, Cambridge University Press, Applied Math. and Mech. Ser., Cambridge Univ. Press, N.Y., pp. 286, 1988.

Note: These notes are an attempt to provide a complete background to the lectures. Only some aspects will be discussed in the lectures. The reader is referred to the original sources for complete details and for references mentioned in the next pages.

EQUATIONS OF MOTION IN AN ELASTIC MEDIUM

The equation of motion in an elastic medium is given by

$$\tau_{ij,j} + f_i = \rho \ddot{u}_i \qquad \cdots (1)$$

where τ_{ij} is the stress tensor, $i,j,k,l = 1,2,3$ & $\tau_{ij} = \tau_{ji}$

$$= c_{ijkl}\, e_{kl}$$

where c's are the elastic constants; e_{ij} is the strain tensor; u_i is the component of displacement in the i-direction

f_i " " " " body force " " "

ρ is the density and Einstein's summation convention for tensors is used (i.e. equation (1) is a set of 3 equat

2.3 heorems of Uniqueness and Reciprocity

It is natural to introduce the discussion of uniqueness (for the displacement field **u** throughout a body with volume V and surface S) with some general remarks concerning the ways in which motion can be set up. Because the displacement is constrained to satisfy (2.17) throughout V, the application of body forces will generate a displacement field, as will the application of tractions on the surface S. We shall show that specification of the body forces throughout V, and tractions over all of S, is enough to determine uniquely the displacement field that will develop throughout V from given initial conditions. An alternative way to specify the influence of S on the displacement field is to give a boundary condition for displacement itself (on S) instead of for the traction. For example, S might be rigid. It might seem at first that the traction on S and the displacement on S are independent properties of the displacement field throughout V. This is not so, however, and it is important for an intuitive understanding of Sections 2.3–2.5 to appreciate that traction over S determines the displacement over S, and vice versa.

UNIQUENESS THEOREM

The displacement $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ throughout the volume V with surface S is uniquely determined after time t_0 by the initial values of displacement and particle velocity at t_0, throughout V; and by values at all times $t \geq t_0$ of (i) the body forces **f** and the heat $\mathcal{Q}$ supplied throughout V; (ii) the tractions **T** over any part S_1 of S; and (iii) the displacement over the remainder S_2 of S, with $S_1 + S_2 = S$. (Either of S_1 or S_2 can be the whole of S.)

PROOF

Suppose $\mathbf{u}_1$ and $\mathbf{u}_2$ are any solutions for **u** that satisfy the same initial conditions and are set up by the same values for (i) (iii). Then the difference $\mathbf{U} = \mathbf{u}_1 - \mathbf{u}_2$ is a displacement field having zero initial conditions, and is set up by zero body forces, zero heating, zero traction on S_1, and $\mathbf{U} = \mathbf{0}$ on S_2. It remains to prove that $\mathbf{U} = \mathbf{0}$ throughout V for $t > t_0$.

The rate of doing mechanical work in the displacement field **U** is clearly zero throughout V and S_1 and S_2 (see (2.22)) for $t \geq t_0$. The third equality in (2.22) can be integrated from t_0 to t, and, together with the zero initial conditions and the use of a strain-energy function (**U** involves adiabatic changes), it follows that

$$\iiint_V \frac{1}{2}\rho \dot{U}_i \dot{U}_i \, dV + \iiint_V \frac{1}{2} c_{ijkl} U_{i,j} U_{k,l} \, dV = 0$$

Both the kinetic and strain energies are positive definite, so that $\dot{U}_i = 0$ for $t \geq t_0$. But $U_i = 0$ at $t = t_0$, and hence $\mathbf{U} = \mathbf{0}$ throughout V for $t > t_0$.

BOX **2.4**

Use of the term "homogeneous," as applied to equations and boundary conditions

The equation for elastic displacement is $\mathbf{L}(\mathbf{u}) = \mathbf{f}$, where **L** is the vector differential operator defined on the components of **u** by

$$(\mathbf{L}(\mathbf{u}))_i \equiv \rho \ddot{u}_i - (c_{ijkl} u_{k,l})_{,j}$$

If body forces are absent, then the equation $\mathbf{L}(\mathbf{u}) = \mathbf{0}$ for **u** is said to be *homogeneous*. A *homogeneous boundary condition* on the surface S is one for which *either* the displacement *or* the traction vanishes at every point of the surface.

This terminology is reminiscent of linear algebra, for which a system of n equations in n unknowns, in the form $\mathbf{A}\mathbf{x} = \mathbf{0}$, is also said to be homogeneous. Here, **x** is a column vector and **A** is some $n \times n$ matrix. It is well known that nontrivial solutions ($\mathbf{x} \neq \mathbf{0}$) can exist, but only if **A** has a special property (namely, a zero determinant). The corresponding result in dynamic elasticity is that motions can occur throughout a finite elastic volume V without any body forces and with a homogeneous boundary condition over the surface of V. These are the *free oscillations* or *normal modes* of the body, which can occur only at certain frequencies. See Chapter 8.

RECIPROCITY THEOREMS

We shall state and prove several general relationships between a pair of solutions for the displacement through an elastic body V.

Suppose that $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ is one of these displacement fields, and that **u** is due to body forces **f** and boundary conditions on S and initial conditions at time $t = 0$. Let $\mathbf{v} = \mathbf{v}(\mathbf{x}, t)$ be another displacement field due to body forces **g** and to boundary conditions and initial conditions (at $t = 0$) which in general are different from the conditions for **u**. To distinguish the tractions on surfaces normal to **n** in these two cases, we shall use the notation $\mathbf{T}(\mathbf{u}, \mathbf{n})$ for the traction due to the displacement **u** and, similarly, $\mathbf{T}(\mathbf{v}, \mathbf{n})$ for the traction due to **v**.

The first reciprocal relation to note between **u** and **v** is

$$\iiint_V (\mathbf{f} - \rho\ddot{\mathbf{u}}) \cdot \mathbf{v} \, dV + \iint_S \mathbf{T}(\mathbf{u}, \mathbf{n}) \cdot \mathbf{v} \, dS = \iiint_V (\mathbf{g} - \rho\ddot{\mathbf{v}}) \cdot \mathbf{u} \, dV + \iint_S \mathbf{T}(\mathbf{v}, \mathbf{n}) \cdot \mathbf{u} \, dS \quad (2.34)$$

This result is due to Betti. It can easily be proved by substitution from (2.17) and (2.16) and then applying the divergence theorem to reduce the left side to

BOX **2.5**

Parallels

A rearrangement of Betti's relation (2.34) gives

$$\iiint_V \{v_i(c_{ijkl}u_{k,l})_{,j} - u_i(c_{ijkl}v_{k,l})_{,j}\}\,dV = \iint_S \{v_iT_i(\mathbf{u},\mathbf{n}) - u_iT_i(\mathbf{v},\mathbf{n})\}\,dS$$

This is a vector theorem for the second-order spatial derivatives occurring in the wave equation of elasticity, which is analogous to Green's theorem

$$\iiint_V (\psi\nabla^2\phi - \phi\nabla^2\psi)\,dV = \iint_S \left(\psi\frac{\partial\phi}{\partial n} - \phi\frac{\partial\psi}{\partial n}\right)dS$$

for scalars and the Laplacian operator. Green's theorem is a working tool for studying inhomogeneous equations, such as $\nabla^2\phi = -4\pi\rho$, and we shall use Betti's theorem for the elastic wave equation, in which the inhomogeneity is the body-force term.

There are many further analogies between Dirichlet problems (for potentials that are zero on S) and elasticity problems with rigid boundaries; and between Neumann problems ($\partial\phi/\partial n = 0$ on S) and traction-free boundaries.

$\iiint_V c_{ijkl}v_{i,j}u_{k,l}\,dV$. Similarly, the right-hand side reduces to $\iiint_V c_{ijkl}u_{i,j}v_{k,l}\,dV$, and (2.34) follows from the symmetry $c_{ijkl} = c_{klij}$.

Note that Betti's theorem does not involve initial conditions for $\mathbf{u}$ or $\mathbf{v}$. Furthermore, it remains true even if the quantities $\mathbf{u}$, $\ddot{\mathbf{u}}$, $\mathbf{T}(\mathbf{u},\mathbf{n})$, and $\mathbf{f}$ are evaluated at time t_1 but $\mathbf{v}$, $\ddot{\mathbf{v}}$, $T(\mathbf{v},\mathbf{n})$, $\mathbf{g}$ are evaluated at a different time t_2. If we choose $t_1 = t$ and $t_2 = \tau - t$ and integrate (2.34) over the temporal range 0 to τ, then the acceleration terms reduce to terms that depend only on the initial and final values, since

$$\int_0^\tau \rho\{\ddot{\mathbf{u}}(t)\cdot\mathbf{v}(\tau - t) - \mathbf{u}(t)\cdot\ddot{\mathbf{v}}(\tau - t)\}\,dt$$

$$= \rho\int_0^\tau \frac{\partial}{\partial t}\{\dot{\mathbf{u}}(t)\cdot\mathbf{v}(\tau - t) + \mathbf{u}(t)\cdot\dot{\mathbf{v}}(\tau - t)\}\,dt$$

$$= \rho\{\dot{\mathbf{u}}(\tau)\cdot\mathbf{v}(0) - \dot{\mathbf{u}}(0)\cdot\mathbf{v}(\tau) + \mathbf{u}(\tau)\cdot\dot{\mathbf{v}}(0) - \mathbf{u}(0)\cdot\dot{\mathbf{v}}(\tau)$$

If there is some time τ_0 before which $\mathbf{u}$ and $\mathbf{v}$ are everywhere zero throughout V (and hence $\dot{\mathbf{u}} = \dot{\mathbf{v}} = \mathbf{0}$ for $\tau \le \tau_0$), then the convolution

$$\int_{-\infty}^{\infty} \rho\{\ddot{\mathbf{u}}(t)\cdot\mathbf{v}(\tau - t) - \mathbf{u}(t)\cdot\ddot{\mathbf{v}}(\tau - t)\}\,dt$$

is zero. We deduce from Betti's theorem the important result, for displacement fields with a quiescent past, that

$$\int_{-\infty}^{\infty} dt \iiint_V \{\mathbf{u}(\mathbf{x},t)\ \mathbf{g}(\mathbf{x},\tau - t) - \mathbf{v}(\mathbf{x},\tau - t)\cdot\mathbf{f}(\mathbf{x},t)\}\,dV$$

$$= \int_{-\infty}^{\infty} dt \iint_S \{\mathbf{v}(\mathbf{x},\tau - t)\cdot\mathbf{T}(\mathbf{u}(\mathbf{x},t),\mathbf{n}) - \mathbf{u}(\mathbf{x},t)\cdot\mathbf{T}(\mathbf{v}(\mathbf{x},\tau - t),\mathbf{n})\}\,dS. \quad (2.35)$$

Green's function for elastodynamics

Green's functions are the impulse response of an elastic solid, that is, it is the displacement generated in the body by the application of an unit force $f_i = \delta(\bar{x})\delta(t)$ where $\bar{x} = (x_1, x_2, x_3)$ and t is the time.

2.4 Introducing Green's Function for Elastodynamics

A major aim of this chapter and the next is the development of a representation for the displacements that typically occur in seismology. The representation will be a formula for the displacement (at a general point in space and time) in terms of the quantities that originated the motion, and we have seen (in the uniqueness theorem) that these are body forces and applied tractions or displacements over the surface of the elastic body under discussion. For earthquake faulting, the seismic source is complicated in that it extends over a finite fault plane (or a finite volume) and over a finite amount of time, and in general involves motions (at the source) that have varying direction and magnitude. We shall find that the representation theorem is really nothing but a bookkeeping device by which the displacement from realistic source models is synthesized from the displacement produced by the simplest of sources—namely, the unidirectional unit impulse, which is localized precisely in both space and time.

The displacement field from such a simple source is the elastodynamic Green function. If the unit impulse is applied at $\mathbf{x} = \boldsymbol{\xi}$ and $t = \tau$ and in the n-direction (see (2.4), taking A = unit constant with dimensions of impulse), then we denote the ith component of displacement at general $(\mathbf{x}, t)$ by $G_{in}(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$. Clearly, this Green function is a tensor (we shall work throughout with Cartesian tensors, and therefore do not distinguish between tensors and dyadics). It depends on both receiver and source coordinates, and satisfies the equation

$$\rho \frac{\partial^2}{\partial t^2} G_{in} = \delta_{in}\delta(\mathbf{x} - \boldsymbol{\xi})\delta(t - \tau) + \frac{\partial}{\partial x_j}\left(c_{ijkl}\frac{\partial}{\partial x_l} G_{kn}\right) \tag{2.36}$$

throughout V. We shall invariably use the initial conditions that $\mathbf{G}(\mathbf{x}, t; \boldsymbol{\xi}, \tau)$ and $\partial\{\mathbf{G}(\mathbf{x}, t; \boldsymbol{\xi}, \tau)\}/\partial t$ are zero for $t \leq \tau$ and $\mathbf{x} \neq \boldsymbol{\xi}$. To specify $\mathbf{G}$ uniquely, it remains to state the boundary conditions on S, and we shall use a variety of different boundary conditions in different applications.

If the boundary conditions are independent of time (e.g., S always rigid), then the time origin can be shifted at will, and we see from (2.36) that $\mathbf{G}$ depends on t

and τ only via the combination $t - \tau$. Hence

$$\mathbf{G}(\mathbf{x}, t; \boldsymbol{\xi}, \tau) = \mathbf{G}(\mathbf{x}, t - \tau; \boldsymbol{\xi}, 0) = \mathbf{G}(\mathbf{x}, -\tau; \boldsymbol{\xi}, -t), \tag{2.37}$$

which is a reciprocal relation for source and receiver times.

If $\mathbf{G}$ satisfies homogeneous boundary conditions on S, then (2.35) can be used to obtain an important reciprocal relation for source and receiver positions. One takes $\mathbf{f}$ to be a unit impulse applied in the m-direction at $\mathbf{x} = \boldsymbol{\xi}_1$ and time $t = \tau_1$, and $\mathbf{g}$ to be a unit impulse applied in the n-direction at $\mathbf{x} = \boldsymbol{\xi}_2$ and time $t = -\tau_2$. Then $u_i = G_{im}(\mathbf{x}, t; \boldsymbol{\xi}_1, \tau_1)$ and $v_i = G_{in}(\mathbf{x}, t; \boldsymbol{\xi}_2, -\tau_2)$, so that (2.35) directly yields

$$G_{nm}(\boldsymbol{\xi}_2, \tau + \tau_2; \boldsymbol{\xi}_1, \tau_1) = G_{mn}(\boldsymbol{\xi}_1, \tau - \tau_1; \boldsymbol{\xi}_2, \tau_2) \tag{2.38}$$

Choosing $\tau_1 = \tau_2 = 0$, this becomes

$$G_{nm}(\boldsymbol{\xi}_2, \tau; \boldsymbol{\xi}_1, 0) = G_{mn}(\boldsymbol{\xi}_1, \tau; \boldsymbol{\xi}_2, 0), \tag{2.39}$$

which specifies a purely spatial reciprocity. Choosing $\tau = 0$ in (2.38) gives

$$G_{nm}(\boldsymbol{\xi}_2, \tau_2; \boldsymbol{\xi}_1, \tau_1) = G_{mn}(\boldsymbol{\xi}_1, -\tau_1; \boldsymbol{\xi}_2, -\tau_2), \tag{2.40}$$

which specifies a space-time reciprocity.

The actual computation of an elastodynamic Green function can itself be a complicated problem. We shall take up this subject in Chapter 4 for the simplest of elastic solids (homogeneous, isotropic, infinite) and also for the case of large separation between source and receiver in inhomogeneous media

2.5 Representation Theorems

If the integrated form of Betti's theorem, our equation (2.35), is used with a Green function for one of the displacement fields, then a representation for the other displacement field becomes available.

Specifically, suppose we are interested in finding an expression for the displacement $\mathbf{u}$ due both to body forces $\mathbf{f}$ throughout V and to boundary conditions on S. We substitute into (2.35) the body force $g_i(\mathbf{x}, t) = \delta_{in}\delta(\mathbf{x} - \boldsymbol{\xi})\delta(t)$, for which the corresponding solution is $v_i(\mathbf{x}, t) = G_{in}(\mathbf{x}, t; \boldsymbol{\xi}, 0)$, and find

$$\begin{aligned} u_n(\boldsymbol{\xi}, \tau) = {} & \int_{-\infty}^{\infty} dt \iiint_V f_i(\mathbf{x}, t) G_{in}(\mathbf{x}, \tau - t; \boldsymbol{\xi}, 0)\, dV \\ & + \int_{-\infty}^{\infty} dt \iint_S \{G_{in}(\mathbf{x}, \tau - t; \boldsymbol{\xi}, 0) T_i(\mathbf{u}(\mathbf{x}, t), \mathbf{n}) \\ & - u_i(\mathbf{x}, t) c_{ijkl} n_j G_{kn,l}(\mathbf{x}, \tau - t; \boldsymbol{\xi}, 0)\}\, dS. \end{aligned}$$

Before giving a physical interpretation of this equation, it is helpful to interchange the symbols $\mathbf{x}$ and $\boldsymbol{\xi}$ and the symbols t and τ. This permits $(\mathbf{x}, t)$ to be the general position and time at which a displacement is to be evaluated, regarded as an integral over volume and surface elements at varying $\boldsymbol{\xi}$ with a temporal convolution. The result is

$$u_n(\mathbf{x}, t) = \int_{-\infty}^{\infty} d\tau \iiint_V f_i(\boldsymbol{\xi}, \tau) G_{in}(\boldsymbol{\xi}, t - \tau; \mathbf{x}, 0)\, dV(\boldsymbol{\xi}) + \int_{-\infty}^{\infty} d\tau \iint_S \{ G_{in}(\boldsymbol{\xi}, t - \tau; \mathbf{x}, 0) T_i(\mathbf{u}(\boldsymbol{\xi}, \tau), \mathbf{n}) - u_i(\boldsymbol{\xi}, \tau) c_{ijkl}(\boldsymbol{\xi}) n_j G_{kn,l}(\boldsymbol{\xi}, t - \tau; \mathbf{x}, 0) \}\, dS(\boldsymbol{\xi}). \quad (2.41)$$

This is our first representation theorem. It states a way in which displacement $\mathbf{u}$ at a certain point is made up from contributions due to the force $\mathbf{f}$ throughout V, plus contributions due to the traction $\mathbf{T}(\mathbf{u}, \mathbf{n})$ and the displacement $\mathbf{u}$ itself on S. However, the way in which each of these three contributions is weighted is unsatisfactory, since each involves a Green function with source at $\mathbf{x}$ and observation point at $\boldsymbol{\xi}$. (Note that the last term in (2.41) involves differentiation with respect to ξ_l.) We want $\mathbf{x}$ to be the observation point, so that the total displacement obtained there can be regarded as the sum (integral) of contributing displacements at $\mathbf{x}$ due to each volume element and surface element. The reciprocal theorem for $\mathbf{G}$ must be invoked, but this will require extra conditions on Green's function itself, since the equation $G_{in}(\boldsymbol{\xi}, t - \tau; \mathbf{x}, 0) = G_{ni}(\mathbf{x}, t - \tau; \boldsymbol{\xi}, 0)$ (see (2.39)) was proved only if $\mathbf{G}$ satisfies homogeneous boundary conditions on S, whereas (2.41) is valid for *any* Green function set up by an impulsive force in the n-direction at $\boldsymbol{\xi} = \mathbf{x}$ and $\tau = t$.

We shall examine two different cases. Suppose, first, that Green's function is determined with S as a rigid boundary. We write $\mathbf{G}^{\text{rigid}}$ for this function and $G_{in}^{\text{rigid}}(\boldsymbol{\xi}, t - \tau; \mathbf{x}, 0) = 0$ for $\boldsymbol{\xi}$ in S. Then (2.41) becomes

$$u_n(\mathbf{x}, t) = \int_{-\infty}^{\infty} d\tau \iiint_V f_i(\boldsymbol{\xi}, \tau) G_{ni}^{\text{rigid}}(\mathbf{x}, t - \tau; \boldsymbol{\xi}, 0)\, dV - \int_{-\infty}^{\infty} d\tau \iint_S u_i(\boldsymbol{\xi}, \tau) c_{ijkl} n_j \frac{\partial}{\partial \xi_l} G_{nk}^{\text{rigid}}(\mathbf{x}, t - \tau; \boldsymbol{\xi}, 0)\, dS. \quad (2.42)$$

Alternatively, we can use $\mathbf{G}^{\text{free}}$ as Green's function, so that the traction $c_{ijkl} n_j (\partial/\partial \xi_l) G_{kn}^{\text{free}}(\boldsymbol{\xi}, t - \tau; \mathbf{x}, 0)$ is zero for $\boldsymbol{\xi}$ in S, finding

$$u_n(\mathbf{x}, t) = \int_{-\infty}^{\infty} d\tau \iiint_V f_i(\boldsymbol{\xi}, \tau) G_{ni}^{\text{free}}(\mathbf{x}, t - \tau; \boldsymbol{\xi}, 0)\, dV + \int_{-\infty}^{\infty} d\tau \iint_S G_{ni}^{\text{free}}(\mathbf{x}, t - \tau; \boldsymbol{\xi}, 0) T_i(\mathbf{u}(\boldsymbol{\xi}, \tau), \mathbf{n})\, dS. \quad (2.43)$$

Equations (2.41)–(2.43) are all different forms of the representation theorem, and each has its special uses. Taken together, they seem to imply a contradiction to the question of whether $\mathbf{u}(\mathbf{x}, t)$ depends upon displacement o S (see (2.42)) or traction (see (2.43)) or both (see (2.41)). But since traction and displacement cannot be specified independently on the surface of an elastic medium, there is no contradiction.

The surface on which values of traction (or displacement) are explicitly required has been taken, in this chapter, as external to the volume V. It is often useful to take this surface to include two adjacent internal surfaces, being the opposite faces of a buried fault. Specialized forms of the representation theorem can then be developed, which enable one to analyze the earthquakes set up by activity on a buried fault. This subject is central to earthquake source theory, and is taken up in the following chapter.

So far, we have considered only Cartesian coordinate systems. In practice, the seismologist is often required to use different coordinates that allow the physical relationship between components of displacement, stress, and strain to be simplified for the geometry of a particular problem. In parti ular, it is often found that a boundary condition must be applied on a surface that can be chosen as the surface on which a general curvilinear coordinate is constant. Vector operations grad, div, curl, and ∇^2 are derived for general orthogonal coordinates in many texts, but rather more is needed to analyze the vector operations required in elasticity.

Equation (2.42) and (2.43) are the most important for our purposes. If body forces are absent, then (2.42) + (2.43) become

$$u_n(\bar{x},t) = -\int_{-\infty}^{\infty} d\tau \iint_S u_i(\bar{\xi},\tau)\, c_{ijkl}\, n_j \frac{\partial}{\partial \xi_l} G_{nk}^{\text{rigid}}(\bar{x},t-\tau;\bar{\xi},0)\, dS$$

and

$$u_n(\bar{x},t) = \int_{-\infty}^{\infty} d\tau \iint_S G_{ni}^{\text{free}}(\bar{x},t-\tau;\bar{\xi},0)\, T_i\left(\bar{u}(\bar{\xi},\tau),\bar{n}\right) dS$$

When the first of these equations is used, i.e. if the LHS is determined when u_i and G_{nk}^{rigid} are known, the problem is called the DISLOCATION model for the seismic source. i.e. given the displacements u_i on S we find the displacement at some point using this equation.

When the second of these equations is used, ie. the LHS is determined when G_{ni}^{free} and the stress T_i are given, the problem is called the CRACK model for the seismic source.

The "dislocation" model is a kinematic model whereas the "crack" model is a dynamic model.

BOX **5.3**

The distinction between kinematics and dynamics

Kinematics is the branch of mechanics that deals purely with motion, without analyzing the underlying forces that cause or participate in the motion. Dynamics is the branch of mechanics that deals directly with force systems, and with the energy balance that governs motion. From these fundamental definitions, two useful conventions have developed for applying the words "kinematic" and "dynamic."

First, in the analysis of displacements alone, kinematic properties are those that may be derived from the eikonal equation (4.41), whereas dynamic properties are those related to displacement amplitudes. Thus the existence of particular wavefronts and ray paths is part of the kinematics of the problem in hand. As an example of a dynamic problem, we might ask if a certain approximation is adequate for the displacements observed at a given receiver at some given distance from a localized source.

Second, in those problems in which we have a direct interest both in the displacement and the associated system of stresses, then kinematic properties are properties of the displacement field and dynamic properties are related to the stresses. For example, if the relative displacement between opposite faces of a fault surface is known as a function of space and time, we say that we have a kinematic description of the fault motion. If the stresses (i.e., traction components) are known on the fault surface, we have a dynamic description. As another example, one refers to boundary conditions as being kinematic or dynamic, in the sense developed in the present section

FORWARD and INVERSE PROBLEMS.

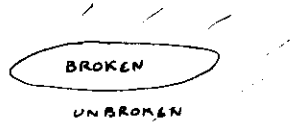
Forward problems in mechanics are when the physical laws governing motion are given together with a system of forces acting on some body and we are required to find the resulting motion. In the earthquake problem, we may be given the material properties at the place in the earth's crust where rupture is occurring together with the equations of motion in an elastic body and we are required to find the motion on the fault.
ANOTHER forward problem is when the motion on the fault is determined, to calculate the ground motion at seismometers located all over the world.

Inverse problems are when the resulting motion due to a system of forces are known to determine the forces which caused the motion. In seismology, we measure the ground motion at seismic stations and reconstruct the motion at the hypocentre of the earthquake

∴ Forward problem → KNOWN cause; FIND effect
Inverse problem → KNOWN effect; FIND cause.

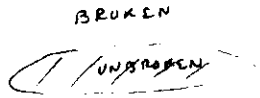
We shall consider inverse problems in the 3rd lecture.

INTERIOR CRACK PROBLEM



"CRACK" IN SEISMOLOGY

EXTERIOR CRACK PROBLEM



"ASPERITY" IN SEISMOLOGY

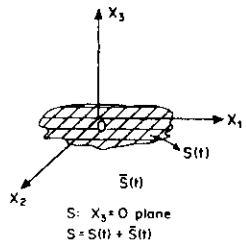
THE BOUNDARY-INTEGRAL EQUATION METHOD

5.1 Representation relations

In order to reduce the earthquake source problem to the solution of a boundary-integral equation, it is necessary first to represent the stress and displacement fields throughout the medium in terms of the displacement discontinuity or the traction perturbations on the fault plane. In the case of the kinematic description of the earthquake source ("dislocation" model), such representation relations provide an explicit solution, as we saw in Section 3.2. However, as we shall show later, the boundary conditions for the dynamically described source ("crack" model) are of mixed type, and such representation relations do not give the required solution explicitly; instead, they relate the boundary values of the traction perturbations and slip, neither of which are known on the entire boundary (fault plane). Together with boundary conditions, these relations comprise the integral equations. When these integral equations are solved, the displacement and stress fields inside the body can be obtained using the same representation relations.

The geometry of the problem is shown in Figure 5.1. The plane $X_3 = 0$ is taken as the crack plane and denoted by S. Let $S(t)$ be that portion of the $X_3 = 0$ plane where slip is nonvanishing at time t and let $\bar{S}(t)$ be its complement: $S = S(t) + \bar{S}(t)$. Generally $S(t)$ is unknown and is to be determined as part of the solution. The relation between the displacement $u_k(\mathbf{X}, t)$ at any point $(\mathbf{X}, t)$ in the medium and the displacement discontinuity $a_i(\mathbf{X}', t')$ across S was given by equation (3.2.9). For a homogeneous, isotropic medium and the case of a planar shear crack, this relation becomes (with slightly different notation for convenience)

$$u_k(\mathbf{X}, t) = \int_0^t dt' \int_{S(t')} K_{\alpha 3k}(\mathbf{X} - \mathbf{X}', t - t') a_\alpha(\mathbf{X}', t')\, dS \qquad (5.1.1)$$



5.1 The geometry of the problem. $S(t)$ is the cracked portion of the $X_3 = 0$ plane.

where Latin subscripts take values 1, 2, 3 and Greek ones the values 1, 2. $K_{\alpha 3k}$ is given by

$$K_{\alpha 3k}(\mathbf{X}, t) = -\mu \left(\frac{\partial U_{k3}}{\partial X_\alpha} + \frac{\partial U_{\alpha k}}{\partial X_3} \right) \qquad (5.1.2)$$

where the Einstein summation convention is assumed and U_{ik} is given by (3.3.1). Equation (5.1.1) gives the representation relation throughout the medium. To express the traction perturbation on the fault plane in terms of the slip on it, we evaluate the traction perturbation components $\tau_{\alpha 3}$ by differentiating (5.1.1) and taking the limit as $X_3 \to 0$ to obtain

$$\tau_{\alpha 3}(\mathbf{X}, t) = \int_0^t dt' \int_{S(t')} T_{\alpha\beta}(\mathbf{X} - \mathbf{X}', t - t') a_\beta(\mathbf{X}', t')\, dS \qquad (5.1.3)$$

where $\mathbf{X}$ and $\mathbf{X}'$ are now the two-dimensional vectors on the fault plane S. The kernel $T_{\alpha\beta}$, obtained by differentiating $K_{\alpha 3k}$, is given by

$$T_{\alpha\beta} = -\mu \left(\frac{\partial^2 U_{33}}{\partial X_\beta \partial X_\alpha} + \frac{\partial^2 U_{\beta 3}}{\partial X_3 \partial X_\alpha} + \frac{\partial^2 U_{\alpha 3}}{\partial X_\beta \partial X_3} + \frac{\partial^2 U_{\beta\alpha}}{\partial^2 X_3} \right) \qquad (5.1.4)$$

Kernel $T_{\alpha\beta}$ has strong singularities, so the integral in (5.1.3) must be considered to be a principal value. Hence, the solution of (5.1.3) is unique only under additional conditions. Usually, it is sufficient to assume that the slip a_α is smooth everywhere except at the crack edge, where it must be finite. This implies the square-root behavior of the slip a_α near the crack edge (see Section 2.3).

An alternative representation relation is obtained by exploiting the symmetries in the problem. For planar shear cracks, the solution can be shown to be antisymmetric in X_3. That is, the displacement components

u_α and traction perturbation τ_{33} are odd in X_3, whereas u_3 and $\tau_{\alpha 3}$ are even in X_3, and it is sufficient to solve the problem for the upper half-space $X_3 \geq 0$. Furthermore, from the continuity of tractions across $X_3 = 0$ [equation (1.3.25)], it follows that $\tau_{33} = 0$ everywhere on $X_3 = 0$. To obtain the required representation relation, let us first reproduce the Green–Volterra formula (3.1.5) after replacing t_1 by t and t by t' for notational convenience. Then we have

$$\int_0^t dt' \int_S (\sigma_{ij} u_i' - \sigma_{ij}' u_i) n_j \, dS + \int_0^t dt' \int_V \rho (f_i u_i' - f_i' u_i) \, dV + \int_V \rho (u_i \dot{u}_i' - \dot{u}_i u_i') \, dV \Big|_{t'=0}^{t'=t} = 0 \tag{5.1.5}$$

Let us choose as u_i' the three solutions corresponding to the three concentrated unit forces f_i' directed along the X_1, X_2, and X_3 axes and given by

$$f_i' = \delta_{ik} \delta(\mathbf{X} - \mathbf{X}') \delta(t - t') \qquad \text{for} \quad X_3' > 0, \quad X_3 \geq 0 \tag{5.1.6}$$

where δ_{ik} is the Kronecker delta and $\delta(\mathbf{X})$ and $\delta(t)$ represent the Dirac delta function, and with the initial and boundary conditions

$$u_i' = \dot{u}_i' = 0 \qquad \text{for} \quad t \leq t'; \qquad \sigma_{i3}' = 0 \qquad \text{at} \quad X_3 = +0 \tag{5.1.7}$$

Let us denote the solution u_i' by G_{ik}. Then in (5.1.5) the last term vanishes due to initial conditions, and the terms containing f_i and σ_{i3}' vanish due to the absence of body forces f_i and by (5.1.7), respectively. Taking σ_{ij} as the stress perturbation tensor τ_{ij} and evaluating the term with the δ function, we get

$$u_k(\mathbf{X}, t) = \int_0^t dt' \int_S G_{k\alpha}(\mathbf{X} - \mathbf{X}', t - t') \tau_{\alpha 3}(\mathbf{X}', t') \, dS \tag{5.1.8}$$

Letting $X_3 \to 0$ and accounting for the symmetry of the displacement components, we obtain the required alternative representation relation as

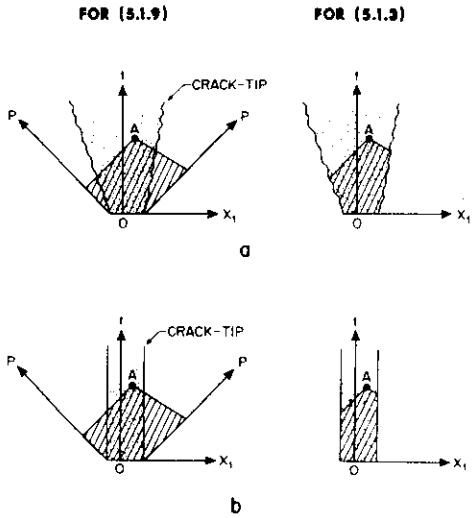
$$a_\alpha(\mathbf{X}, t) = 2 \int_0^t dt' \int_S G_{\alpha\beta}(\mathbf{X} - \mathbf{X}', t - t') \tau_{\beta 3}(\mathbf{X}', t') \, dS \tag{5.1.9}$$

where $\mathbf{X}$ and $\mathbf{X}'$ are now two-dimensional vectors on S. The required components of G are the solution to Lamb's problem and can be expressed in terms of elementary functions. The expressions for $G_{\alpha\beta}$ for the two- and three-dimensional cases are given in Appendix 1. The kernel $G_{\alpha\beta}$ possesses only weak singularities and can be directly discretized for numerical computation, as we shall discuss in Section 5.2.

Thus, we have obtained two representation relations, (5.1.3) and (5.1.9), both of which relate the slip on the crack plane to the traction perturbation on this plane. These relations comprise a set of mutually inverse integral transforms. If either the slip a_α or the traction perturbation $\tau_{\alpha 3}$ were known everywhere on S, the other of these two quantities could be obtained using (5.1.3) or (5.1.9), respectively. But in a dynamically described source (crack) problem, neither one of them is known everywhere on S. The systems (5.1.3) and (5.1.9) thus simply provide relationships between a_α and $\tau_{\alpha 3}$. Some additional relations between a_α and $\tau_{\alpha\beta}$ on the crack plane are needed to provide the required boundary-integral equations. These relations can be obtained from the constitutive relations on S, as we shall discuss in Section 5.2.

Relations (5.1.3) and (5.1.9) are equivalent in that either of them can be used together with the constitutive relations to solve the crack problem. The variables $\mathbf{X}$ and $\mathbf{X}'$ in (5.1.3) are confined to $S(t)$ since a_α vanishes outside the crack. On the other hand, the integration domain in (5.1.9) covers all points influenced by disturbances that propagate with the fastest wave velocity of the problem (e.g., the compressional wave speed of the medium for general three-dimensional problems). Therefore, one of these two relations may be more efficient than the other for a given problem. For problems in which $S(t)$ is much smaller than S, for example, "interior" crack problems, it is more advantageous to use (5.1.3). For some "exterior" crack problems, the region of traction perturbations is limited, and then (5.1.9) is the more efficient relation. Of course, the domains of integration may coincide for some particular situations – for example, for the case of a self-similar crack propagating with the fastest wave speed of the medium. The domains of integration for the two representation relations are shown in Figure 5.2 for two cases of the interior crack problem. The domain of integration for the exterior problem will be discussed in detail in Section 5.6 under "A circular asperity on an infinite fault plane."

The relation (5.1.3) was essentially formulated by Budiansky and Rice (1979). Burridge (1969) used its two-dimensional form, and Burridge and Moon (1981) its three-dimensional scalar form. The two-dimensional form of (5.1.9) was first used by Kostrov (1966) for dynamic elastic problems and by Das (1976, 1980) for general three-dimensional problems.



5.2 Cross sections of the domains of integration in the $(X_1 - t)$ plane for the relations (5.1.3) and (5.1.9) for (a) a propagating "interior" crack and (b) a stationary "interior" crack. The crack regions are stippled. The striped areas are the domains of integration for determining the slip at some representative point A, say, within the crack.

In the next section, we discuss the discretization of the representation relations developed here.

5.2 Discrete representations

Let u first discuss the method of discretization of relation (5.1.9). We introduce a regular network of grids centered at the points

$$\begin{aligned} X_1 &= i\,\Delta X, & i &= -\infty, \ldots, -1, 0, 1, \ldots, \infty \\ X_2 &= j\,\Delta X, & j &= -\infty, \ldots, -1, 0, 1, \ldots, \infty \\ t &= \left(k + \tfrac{1}{2}\right)\Delta t, & k &= 0, 1, \ldots, \infty \end{aligned} \tag{5.2.1}$$

In each element of the network, we replace the tractions by their average values over the grid given by (the subscript 3 for the τ_α's being implicit from now on)

$$\tau_{\alpha ijk} = \frac{1}{\Delta t\,\Delta X^2} \times \int_0^{\Delta t} d\tau \iint_{-\Delta X/2}^{\Delta X/2} \tau_\alpha(i\,\Delta X + \xi_1, j\Delta X + \xi_2, k\,\Delta t + \tau)\, d\xi_1\, d\xi_2 \tag{5.2.2}$$

We replace the slip a_α by its value at a point within the grid,

$$a_{\alpha ijk} = a_\alpha(i\,\Delta X, j\Delta X, k\,\Delta t + \delta t), \qquad 0 \le \delta t \le \Delta t \tag{5.2.3}$$

and we replace the discrete Green function $F_{\alpha\beta}$ by

$$F_{\alpha\beta}(i, j, k) = -2\int_0^{\Delta t} d\tau \int_{-\Delta X/2}^{\Delta X/2} \times G_{\alpha\beta}(i\,\Delta X + \xi_1, j\Delta x + \xi_2, k\,\Delta t + \delta t - \tau)\, d\xi_1\, d\xi_2 \tag{5.2.4}$$

Since $G_{\alpha\beta}$ possesses integrable singularities, $F_{\alpha\beta}$ is easily evaluated using (5.2.4). The properties of $F_{\alpha\beta}$ are discussed in detail in Appendix 1. The Green function $F_{\alpha\beta}$ vanishes outside the region $v_P^2 t^2 = X_1^2 + X_2^2$. In particular, if $v_P(\Delta t - \delta t) \le \Delta X/2$, $F_{\alpha\beta}(i, j, 0)$ is nonvanishing only for $i = 0 = j$. This value is

$$F_{\alpha\beta}(0,0,0) = -2\int_0^{\delta t} d\tau \iint_{-\infty}^{\infty} G_{\alpha\beta}(\xi_1, \xi_2, \tau)\, d\xi_1\, d\xi_2 \tag{5.2.5}$$

or, taking into account the homogeneity and symmetry properties of $G_{\alpha\beta}$,

$$F_{\alpha\beta}(0,0,0) = F_0\,\delta_{\alpha\beta}\,\delta t \tag{5.2.6}$$

where

$$F_0 = -2\int_0^1 d\tau \iint_{-\infty}^{\infty} G_{11}(\xi_1, \xi_2, \tau)\, d\xi_1\, d\xi_2 \tag{5.2.7}$$

Here F_0 is a positive constant [the minus sign in equations (5.2.4) and (5.2.7) was included to make F_0 positive] independent of the grid size and $\delta_{\alpha\beta}$ is the Kronecker delta. The constant F_0 is the largest element in absolute value of the matrix $F_{\alpha\beta}$. Substituting the discrete forms of τ_α,

a_α, and $G_{\alpha\beta}$ into (5.1.9), we obtain

$$a_{\alpha ijk} = -\sum_{k'=0}^{k}\sum_{i'=-\infty}^{\infty}\sum_{j'=-\infty}^{\infty} F_{\alpha\beta}(i-i', j-j', k-k')\tau_{\beta i'j'k'} + \text{approximation error} \tag{5.2.8}$$

In this expression, with the same order of approximation, the constant F_0 can be replaced by some other positive constant G_0, say, independent of the grid size, which is to be chosen on the basis of stability considerations.

Accordingly, (5.2.8) can be rewritten as

$$a_{\alpha ijk} + \Delta t G_0 \tau_{\alpha ijk} = -\sum_{k'=0}^{k-1}\sum_{i',j'\in S^P_{(k)}} F_{\alpha\beta}(i-i', j-j', k-k')\tau_{\beta i'j'k'} \quad \text{for} \quad i, j \in S^P_{(k)} \tag{5.2.9}$$

where $S^P_{(k)}$ is the union of all grids influenced by disturbances at time $k\,\Delta t$. The stability and approximation error of (5.2.9) are discussed in Appendix 2, where it is shown that

$$0 < \frac{F_0}{G_0} < 2 \tag{5.2.10}$$

is a necessary condition for stability. Relation (5.2.9) is the required discrete form of the representation relation (5.1.9). For a given k, the right side of (5.2.9) depends only on the solution at previous times ($k < k'$) and (5.2.9) is an explicit scheme.

Successful discretization of equation (5.1.3), which would lead to a convenient numerical scheme, has not been achieved. Some possible approaches are discussed by Burridge (1969) and by Burridge and Moon (1981). Instead of discretizing (5.1.3), we shall use the fact that (5.1.3) is an integral transform, inverse to (5.1.9), and construct an inverse of the discrete transform (5.2.8). This inverse transform, like the direct transform (5.2.8), must be a discrete convolution transform; that is, its kernel will depend only on differences:

$$\tau_{\alpha ijk} = -\sum_{k'=0}^{k}\sum_{i'=-\infty}^{\infty}\sum_{j'=-\infty}^{\infty} S_{\alpha\beta}(i-i', j-j', k-k')a_{\beta i'j'k'} \tag{5.2.11}$$

Then, application of transform (5.2.8) to this kernel must give the unit

kernel:

$$\sum_{k'=0}^{k}\sum_{i'=-\infty}^{\infty}\sum_{j'=-\infty}^{\infty} F_{\alpha\beta}(i-i', j-j', k-k') \times S_{\beta\gamma}(i'-i'', j'-j'', k'-k'') = \delta_{\alpha\gamma}\delta_{ii''}\delta_{jj''}\delta_{kk''} \tag{5.2.12}$$

To determine $S_{\alpha\beta}$, we obtain the explicit numerical scheme from (5.2.12):

$$\Delta t G_0 S_{\alpha\gamma}(i, j, k) = -\sum_{k'=0}^{k-1}\sum_{i'=-\infty}^{\infty}\sum_{j'=-\infty}^{\infty} F_{\alpha\beta}(i-i', j-j', k-k') \times S_{\beta\gamma}(i', j', k') + \delta_{\alpha\gamma}\delta_{i0}\delta_{j0}$$

Equation (5.2.11) can be written in a form similar to (5.2.9) as

$$\tau_{\alpha ijk} + S_0 a_{\alpha ijk} = -\sum_{k'=0}^{k-1}\sum_{i',j'\in S_{(k')}} S_{\alpha\beta}(i-i', j-j', k-k')a_{\beta i'j'k'} \quad \text{for} \quad i, j \in S_{(k)} \tag{5.2.13}$$

where $S_{(k)}$ is the union of all grids with nonvanishing slip and the tip element $S_0 = S_{11}(0,0,0) = 1/(\Delta t G_0)$ and is a positive constant. (The sign of F_0, G_0, and S_0 becomes important in problems where friction acts on the fault faces.) Note that using (5.2.9) is exactly equivalent to using (5.2.13) for a particular problem; that is, the solutions using the two algorithms must coincide, apart from different rounding error accumulation. The properties of $S_{\alpha\beta}$ are discussed in Appendix 2.

The approach of discretizing $T_{\alpha\beta}$ by inverting $F_{\alpha\beta}$ has two advantages. First, one can use the simple discrete representations of τ_α and a_α, that is, (5.2.2) and (5.2.3), which can be shown to be good approximations by comparing the results with analytical solutions of simple problems (an example of which is given in the next section). Direct discretization of $T_{\alpha\beta}$ may require some different representation for τ_α and a_α. Second, and more important from the practical point of view, $F_{\alpha\beta}$ is a well-behaved matrix and can be inverted without reservation. Furthermore, since $F_{\alpha\beta}$ is quite sparse, one does not actually have to invert a very large matrix [$F_{\alpha\beta}$ will consist of $(N^2T)^2$ elements if N^2 is the number of perturbed spatial grids on the fault plane and T the maximum time level for which $S_{\alpha\beta}$ is desired] but may determine $S_{\alpha\beta}$ by an explicit time-stepping procedure.

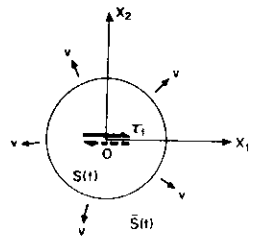
It was mentioned in Section 5.1 that representation (5.2.13) is more economical than representation (5.2.9) for "interior" crack problems,

whereas the situation is reversed for "exterior" crack problems. Let us quantify this efficiency for the numerical solution of a stationary crack – that is, a finite crack that slips without growing. In general, in seismology one is interested only in the slip distribution on the crack, because it completely determines the radiation from the crack through the asymptotic form of (5.1.3), that is, equation (3.2.11). The traction distribution outside the crack, in this context, is of interest at most in the vicinity of the crack edge. The number of grid values that are relevant at a fixed time step is proportional to the crack area in this case, whereas the number of traction values involved in algorithm (5.2.9) is proportional to the square of time. The number of arithmetic operations necessary to obtain the solution at the kth time step is proportional to k^6, and the necessary storage is proportional to k^3, when using (5.2.9). So every doubling of the number of time steps requires eight times more storage and sixty-four times longer computation time. For the stationary crack problem under discussion here, the number of slip values is proportional to k and the number of arithmetic operations when using (5.2.13) is proportional to k^2, and the storage required is proportional to k. Thus, for this problem, using (5.2.13) would be k^4 times faster and require k^2 less storage. Since the two algorithms are equally efficient for a self-similar crack propagating at the fastest wave speed of the medium, the economy in using (5.2.13) over (5.2.9) for interior problems increases with slower and slower crack speeds. Similar comparisons of efficiency can be made for exterior problems, with the above considerations applying in reverse. Algorithm (5.2.9) was first used by Hamano (1974) for two-dimensional crack problems and by Das (1980) for three-dimensional problems.

We shall now proceed to apply the algorithms developed here to some specific dynamic crack problems. In general, we shall consider only three-dimensional cases except when appropriate solutions do not exist, in which case a two-dimensional illustration will be used.

5.3 The circular self-similar shear crack

We consider the three-dimensional problem of a self-similar circular shear crack, which is initiated at a point and propagates at a known constant velocity, v say. The assumption of a constant fracture speed is rather unphysical, because it violates principles of fracture mechanics. A stress singularity that grows in time, as the stress singularity at the growing crack edge does, is unlikely to result in a constant fracture speed under any of the fracture criteria discussed in Chapter 2,



5.3 Geometry of the self-similar circular shear crack.

unless the fracture toughness distribution on the fault plane is rather pathological. Spontaneous crack problems are more physical, and we shall devote a large part of this chapter to them. It is instructive, however, to discuss the self-similar problem since it is the simplest possible case and is useful for demonstrating the numerical method, including its accuracy and stability. Historically, more than two decades ago, this problem was the first dynamic three-dimensional shear problem to be solved.

The crack region $S(t)$ is known and given by

$$S(t): X_1^2 + X_2^2 \le v_p^2 t^2$$

The geometry of the problem is shown in Figure 5.3. We shall solve the problem when the stress drop on $S(t)$ is prescribed to be a constant, $\Delta\sigma$, say.[1] Without loss of generality for the circular crack problem, we may assume that the stress drop is directed in the X_1 direction. Then we have the mixed boundary value problem

$$\tau_1 = \Delta\sigma, \tau_2 = 0 \quad \text{on} \quad S(t); \qquad a_\alpha = 0 \quad \text{on} \quad \bar{S}(t)$$

After discretization, we obtain

$$\tau_{1ijk} = \Delta\sigma, \tau_{2ijk} = 0$$

$$\text{for} \quad S_{(k)}: (i\Delta X)^2 + (j\Delta X)^2 \le v^2\left(\frac{k+1}{2}\right)^2(\Delta t)^2 \quad (5.3.1)$$

$$a_{\alpha ijk} = 0 \quad \text{for} \quad \bar{S}_{(k)}: (i\Delta X)^2 + (j\Delta X)^2 > v^2\left(\frac{k+1}{2}\right)^2(\Delta t)^2$$

[1] The term "stress drop" has traditionally been used in seismology to mean traction drop, and we shall continue to use the term in this sense

Taken together with (5.2.9) or (5.2.13) this gives the complete formulation of the problem.

We first use (5.2.9) to solve the problem. We denote its right side by, say, $L_{\alpha i j(k-1)}$, that is,

$$L_{\alpha i j(k-1)} = -\sum_{k'=0}^{k-1} \sum_{i', j' \in S^{P}_{(k-1)}} F_{\alpha\beta}(i-i', j-j', k-k')\tau_{\beta i' j' k'}$$

for $i, j \in S^{P}_{(k)}$

where $S^{P}_{(k)}$ was defined in the last section. The term $L_{\alpha i j(k-1)}$ depends on the values of τ_α at all previous time steps and is thus known at any time step if the traction history up to the previous time step is known. The summation in $L_{\alpha i j(k-1)}$ extends over the entire cone of dependence of the grid point $(i\Delta X, j\Delta X, (k+\frac{1}{2})\Delta t)$ except the grid point itself. Then we can rewrite (5.2.9) as

$$a_{\alpha i j k} + \Delta t G_0 \tau_{\alpha i j k} = L_{\alpha i j(k-1)}$$

It follows very simply from this that, under the mixed boundary conditions (5.3.1), the solution is given by

$$\begin{aligned} a_{\alpha i j k} &= -\delta_{\alpha 1}\,\Delta t G_0\,\Delta\sigma + L_{\alpha i j(k-1)} \\ \tau_{\alpha i j k} &= \delta_{\alpha 1}\,\Delta\sigma \qquad \text{on} \quad S_{(k)} \end{aligned} \tag{5.3.2}$$

and by

$$\tau_{\alpha i j k} = \frac{L_{\alpha i j(k-1)}}{\Delta t G_0}; \qquad a_{\alpha i j k} = 0 \qquad \text{on} \quad \bar{S}_{(k)} \tag{5.3.3}$$

where $\delta_{\alpha 1}$ is the Kronecker delta. Since the initial $L_{\alpha i j(k-1)}$ is zero by definition, this is an explicit scheme to determine slip and traction perturbation everywhere on the crack plane.

Let us now solve the same problem using the discrete representation (5.2.13). We denote its right side by, say, $M_{\alpha i j(k-1)}$, that is,

$$M_{\alpha i j(k-1)} = -\sum_{k'=0}^{k-1} \sum_{i', j' \in S_{(k')}} S_{\alpha\beta}(i-i', j-j', k-k')a_{\beta i' j' k'}$$

for $i, j \in S_{(k)}$

where $S_{(k)}$ was defined earlier. The term $M_{\alpha i j(k-1)}$ depends on the values of a_α at all previous time steps and is thus known at any time step if the slip history up to the previous time step is known. The summation in $M_{\alpha i j(k-1)}$ extends over the intersection of the cone of dependence of the point $(i\Delta X, k\Delta X, (k+\frac{1}{2})\Delta t)$ with the crack area $S_{(k)}$ but excluding

this point itself, that is, over a smaller region than when (5.2.9) was used. We can now rewrite (5.2.13) as

$$\tau_{\alpha i j k} + S_0 a_{\alpha i j k} = M_{\alpha i j(k-1)}$$

The solution to this under the mixed boundary conditions (5.3.1) is very simply

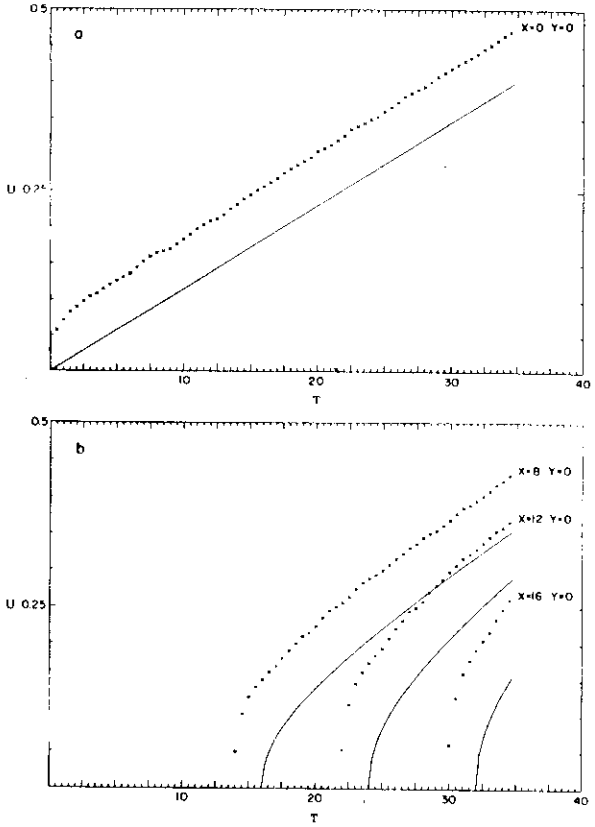
$$\tau_{\alpha i j k} = \delta_{\alpha 1}\,\Delta\sigma; \qquad a_{\alpha i j k} = -\frac{\delta_{\alpha 1}\,\Delta\sigma + M_{\alpha i j(k-1)}}{S_0} \qquad \text{on} \quad S_{(k)} \tag{5.3.4}$$

This is an explicit scheme for finding $a_{\alpha i j k}$ on $S_{(k)}$ since $M_{\alpha i j(k-1)}$ is known to be zero initially and gives the required slip on the crack. If the solution for the stress $\tau_{\alpha i j k}$ on $\bar{S}(k)$ is desired, it can be easily obtained as

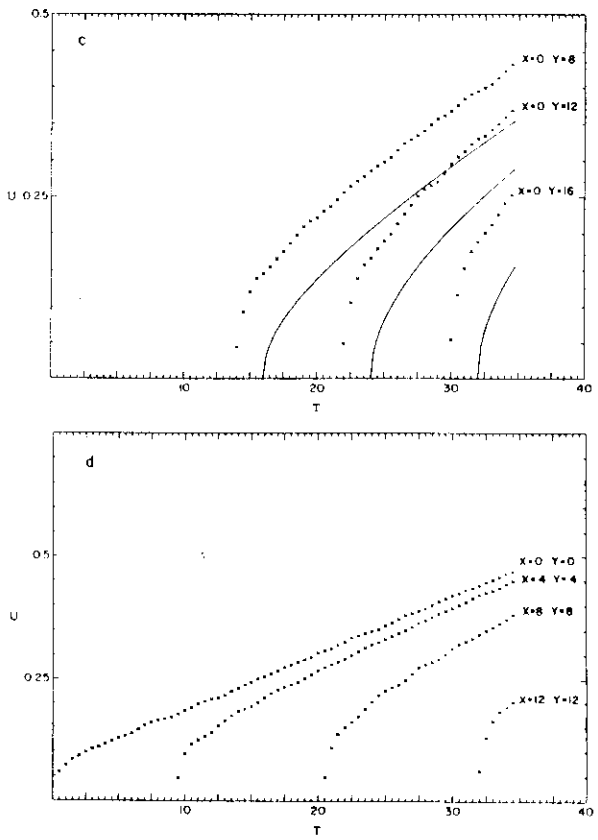
$$a_{\alpha i j k} = 0; \qquad \tau_{\alpha i j k} = M_{\alpha i j(k-1)} \tag{5.3.5}$$

It is important to note that the solutions for the traction perturbation and slip are homogeneous functions of zeroth order of the coordinates and time for self-similar problems (definition of self-similarity).

We shall now compare the numerical solution using the two forms of the representation relations with the corresponding analytic self-similar solution. Let us consider the case in which the fracture velocity $v = v_p/2$. The results are shown in Figure 5.4 for the case of a Poisson solid. (In the remainder of the book all numerical results will be illustrated for the Poisson solid only.) The required analytic solution for the slip a_α is obtained by integrating equation (3.4.18) with respect to time and substituting $v = v_p/2$. The numerical solutions are determined from (5.3.2), (5.3.3) and from (5.3.4), (5.3.5) for the two forms of the representation relations. The two numerical methods yield solutions that are identical except at the last decimal place, as expected. Since the initial crack in a numerical method cannot be infinitesimal, one cannot in fact numerically study a self-similar crack. In this example, an initial crack of radius ΔX is assumed to appear instantaneously and start extending at a speed of $v_p/2$. In other words, the analytic and numerical solutions are really solutions of different problems. However, at larger and larger times, the effect of this initial difference is expected to become less and less significant. The normalized half-slip U is plotted in units of $(20\,\Delta\sigma\,\Delta X)/3\mu$ and the normalized time T is plotted in units of $v_p/\Delta X$. The analytic solution for the half-slip is shown by the continuous line in Figure 5.4; its numerically calculated values are shown by crosses at



5.4 Normalized half-slip U versus normalized time T for a circular crack propagating at a speed $v_p/2$. The analytic solution is indicated by solid lines, and the numerical solution for the instantaneously appearing circular crack of radius ΔX is indicated by crosses. $X = X_1/\Delta X$; $Y = X_2/\Delta X$. The solution is plotted at points (a) and (b) along the X_1 axis, (c) along the X_2 axis, (d) along a line at 45° to the axes [the analytic solution is omitted to increase clarity in (d)]. The spatial to temporal grid size ratio is given by $v_p \Delta t/\Delta X = .5$.



5.4 (*cont.*)

points along the X_1 and X_2 axes and along a line at 45° to the two axes. The numerical solution always lies above, that is, is larger than the analytic solution owing to the above-mentioned initial difference. What is most important, however, is that the *rate* of increase of slip with time is the same for the analytic and numerical solutions except for the first few time steps of the solution. To ensure stability, G_0 was taken as $2F_0$, which satisfies the necessary criterion for stability given by (5.2.10). We note that the square-root form of the slip a_1 is well approximated numerically. The slip component a_2 on the crack was always found to be less than 3 percent of a_1; in other words, it is practically negligible (it is identically zero in the analytic case). In fact, the maximum values of a_2 are concentrated in grids near the crack edge, which in a numerical scheme is necessarily smeared out, and in the interior of the crack its values are even smaller. The values of a_2 can be considered a measure of the numerical noise in the solution. Thus, the slip is in the direction of the stress drop on the fault plane for the circular self-similar crack. Figure 5.4 also shows that the numerically determined slip is azimuthally symmetric (as he analytic one is) without the *a priori* imposition of such a condition. The three-dimensional self-similar circular shear crack problem was numerically studied by Madariaga (1976), Archuleta (1976), and Das (1980), among others.

So we have used the numerical boundary-integral method in its two forms to study a simple self-similar problem and showed that the results compare well with its analytic solution for one particular fracture speed. In a similar way, any crack problem with given stress drop on the fault plane and crack speed (neither of these need be constant and the speed need not be the same in all directions) can be solved following the above development. It includes as a special case the stationary crack problem studied by Madariaga (1976) and by Das (1980).

5.4 The finite circular shear crack

Next we consider the problem of a circular shear crack that initiates at a point, propagates at a preassigned constant velocity v, say, and stops when it reaches some finite radius r, say. Let the stress drop on the crack be assigned a constant, $\Delta\sigma$, say, and directed in the X_1 direction. The crack region $S(t)$ is defined by

$$S(t): \begin{cases} X_1^2 + X_2^2 \le v^2 t^2 & \text{for } vt \le r; \\ X_1^2 + X_2^2 = r & \text{for } vt > r \end{cases} \tag{5.4.1}$$

Then we have the mixed boundary value problem

$$\tau_1 = \Delta\sigma, \tau_2 = 0 \quad \text{on} \quad S(t); \qquad a_\sigma = 0 \quad \text{on} \quad \bar{S}(t) \tag{5.4.2}$$

This problem can be solved numerically following the procedure outlined in detail in the last section, the discrete crack area $S_{(k)}$ now being defined as

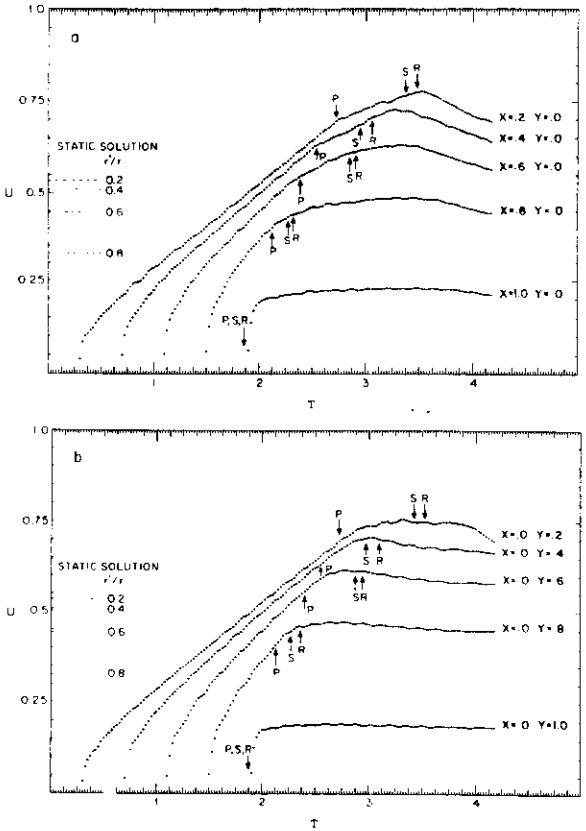
$$S_{(k)}: \begin{cases} (i\Delta X)^2 + (j\Delta X)^2 \le v^2\left(k + \tfrac{1}{2}\right)^2(\Delta t)^2 & \\ \qquad\qquad\qquad\qquad \text{for} \quad v\left(k + \tfrac{1}{2}\right)\Delta t \le r & \\ (i\Delta X)^2 + (j\Delta X)^2 = r^2 \qquad \text{for} \quad v\left(k + \tfrac{1}{2}\right)\Delta t > r & \end{cases}$$

Even this relatively simple problem of a finite dynamic crack cannot be solved analytically, though a kinematic description using the results of the self-similar problem was considered by Sato and Hirasawa (1973).

Let us apply the numerical algorithm (5.2.13) to this problem. We consider an instantaneously appearing crack of diameter $3\Delta X$, which grows to a final diameter of $41\,\Delta X$ at a speed $v = v_R/2$. We shall allow backslip to occur on the crack in this example. The half-slip $a_1/2$ is plotted against time in Figure 5.5 and is normalized by $(r\Delta\sigma)/3\mu$ for this problem. The slip on the crack is found to be essentially in the direction of the stress drop even after the crack has stopped, and it coincides with the solution of the corresponding self-similar problem until the first diffracted waves from the crack edge arrive and decrease the slip rate. The rise time and slip at the center are larger than at the crack edge. The dynamic slip is found to overshoot the static value, but when backslip is allowed on the crack, as it is here, it decreases from its maximum value and approaches the static value.

From some simple geometric considerations, it is possible to obtain a rough estimate of the dynamic overshoot expected at a point within the crack. A point on the crack continues slipping until some (diffracted) wave from the crack edge returns to it. Let this wave have velocity v_H. Then the displacement expected at a point on the crack is given by the solution of the dynamic self-similar circular crack propagating at the speed v. The overshoot OV on the crack varies with position on the crack, being largest at the center and smallest at the edge. At the center, the overshoot is given by

$$OV \approx \frac{A(v)}{A(0)}\left(1 + \frac{v}{v_H}\right) - 1$$



5.5 Normalized half-slip U versus normalized time $T = v_P' t/r$ for a finite circular crack for (a) points along the X_1 axis and (b) points along the X_2 axis $X = X_1/r$, $Y = X_2/r$, r being the final crack radius, and $v_P \Delta t/\Delta X = 5$. The theoretical arrival times of the diffracted P, S, and Rayleigh waves from the crack edge are indicated on each curve by arrows. The static solutions for the points $r'/r = .2, .4, .6,$ and $.8$, where r' is the distance of the point from the center of the crack, are given by $U = .55\sqrt{1 - r'^2/r^2}$ and marked along the abscissas.

where $A(v)/A(0)$ was plotted in Figure 3.6 as a function of v. For $v = v_P/2$, this overshoot is $0.85(1 + v_P/2v_H) - 1$. A lower bound of this overshoot would be obtained for $v_H = v_P$, that is, if the wave that stops the slip is the P wave. This lower bound at the center of the crack is ~ 28 percent above the corresponding static slip from the above formula. Figure 5.5 shows that the maximum slip is reached at or soon after the time when the Rayleigh wave arrives from the crack edge. Then, taking $v_H = v_R$, the Rayleigh wave velocity, one obtains a very rough estimate of the upper bound of the overshoot at the crack center as ~ 65 percent. This is a very rough estimate since it assumes that the P and/or S waves from the crack edge did not modify the slip determined by the dynamic self-similar solution (an assumption that is seen from Figure 5.5 to be not quite valid!). At the crack edge the overshoot is

$$OV \approx \frac{A(v)}{A(0)}\sqrt{1 + v/v_H} - 1$$

so that for $v_H = v_P$, the edge overshoot is ~ 4 percent, and for $v_H = v_R$, it is ~ 18 percent of its static value. The static values of U are shown in Figure 5.5, and the central overshoot in the numerical case was found to be ~ 36 percent. Estimates of this overshoot, obtained by Madariaga (1976), Archuleta (1976), Das (1980), and others using different numerical methods and/or different grid sizes and without allowing backslip were found to lie between 20 and 27 percent. The dynamic overshoot of slip in the interior of the crack may, of course, be interpreted as the overshoot of the static stress drop there. The above formulas also show that the dynamic overshoot increases with increasing crack speed v. Obviously for a given problem (solved by the same method and using the same discretization of the problem), the overshoot must be greater when backslip is disallowed on the fault. Hence, in obtaining an estimate of the static solution, allowing backslip would give a better estimate than disallowing it.

ANALYITICAL SOLUTION FOR SLIP RATE ON A SELF-SIMILAR PROPAGATING CRACK, RUPTURE VELOCITY BEING LESS THAN THE RAYLEIGH WAVE SPEED OF THE ELASTIC MEDIUM (OR EQUAL).

CIRCULAR CASE

For a self-similar, circular shear crack with constant stress drop $\Delta\sigma$, the slip rate distribution is (Kostrov, 1964)

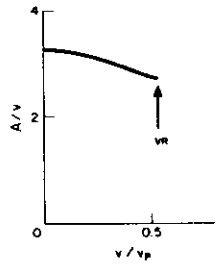
$$\dot{d} = \frac{tA}{\left(t^2 - x_i x_i / v^2\right)^{1/2}} \qquad \text{for} \quad x_i x_i < t^2 v^2 \tag{3.4.18}$$

Here, v is the constant fracture velocity, which is $\leq v_R$, the Rayleigh wave velocity, and A is given by

$$A = \frac{4\Delta\sigma}{\mu I(v)} \tag{3.4.19}$$

where $I(v)$ is a smooth function of v given by

$$I(v) = \left(\frac{v^2}{v_S^3}\right)\int_0^\infty \left\{(2\nu + 1)^2 - 4\nu\left(\gamma^2 + \nu\right)^{1/2}(1 + \nu)^{1/2} + 1 + \nu\right\} \times \frac{d\nu}{\left[1 + \nu\left(v^2/v_S^2\right)\right]^2(1 + \nu)^{1/2}} \tag{3.4.20}$$



3.6 A/v, normalized by its static value, versus fracture speed v.

and $\gamma = v_S/v_P$.[1] Here A/v gives the ratio of slip versus crack radius and is plotted against v in Figure 3.6, its value being normalized by the static value $24\,\Delta\sigma/7\pi\mu$.

[1] Performing the integration in (3.4.20), we can write $I(v)$ in closed form as

$$I(v) = \alpha^4\left\{12(\gamma - 1) - \frac{12\alpha^4 - 21\alpha^2 + 8}{(\alpha^2 - 1)^{3/2}}\tan^{-1}\sqrt{\alpha^2 - 1} + \frac{12\alpha^2 - 8\gamma^2}{\sqrt{\alpha^2 - \gamma^2}}\tan^{-1}\frac{\sqrt{\alpha^2 - \gamma^2}}{\gamma} + \frac{\alpha^2 - 2}{\alpha^2(\alpha^2 - 1)}\right\} \tag{3.4.20a}$$

where $\alpha^2 = v_S^2/v^2$.

ELLIPTICAL CRACK

Burridge and Willis (1969) solved a more general problem, in which the fracture velocity depended on the direction, so that the crack at all times had an elliptical shape, given by

$$\frac{x_1^2}{v_1^2} + \frac{x_2^2}{v_2^2} \leq t^2 \tag{3.4.24}$$

where v_1 and v_2 are the fracture speeds in the x_1 and x_2 directions, respectively. It was found that the slip direction coincides with the direction of stress drop, when it is parallel to one of the axes of the ellipse. The distribution of slip rate magnitude is

$$\dot{d}(\mathbf{x}, t) = \frac{At}{\left(t^2 - x_1^2/v_1^2 - x_2^2/v_2^2\right)^{1/2}} \tag{3.4.25}$$

where A is proportional to the stress drop $\Delta\sigma$. The expression for A has the form of a complicated surface integral, which depends on the velocities v_1 and v_2.

Far-field radiation

For "crack" problems.

It can be proved that the far-field P and S wave pulse shapes are given by the sum of the slip rate on the fault at a given time t i.e.

$$S(t) = \frac{d}{dt} \sum a(t)$$

where $S(t)$ is the pulse shape, and the summation (integral for the continuous problem) extends over the broken area of the fault

For "asperity" problems, it is the sum of the stress distribution on the broken part of the fault

$$S(t) = \sum \tau(t)$$

6.3 The heterogeneous faulting process

We saw in the last chapter that the fracture process depends on the initial stress distribution on the fault and on the physical properties of the fault such as its strength and its static and dynamic friction levels. Variations in any one of these parameters over the crack plane would produce variations in the fracture velocity, slip rate, and stress drop distribution over the fault. This heterogeneity would be manifested in the complexity of the radiated pulse shapes. Such observations of "multiple shocks" led to the introduction and subsequent acceptance of models with heterogeneous stress drop and strength over the fault plane. Two idealizations of this situation have been considered in the past decade. In one, known as the "barrier" model, the stress drop on the fractured part of the fault plane is essentially uniform and the critical stress level has large variations. In the other, known as the "asperity" model, the stress drop is highly variable over the fault. Obviously, every conceivable variation and combination of these two extreme cases is plausible in reality. Also, instead of one unique crack edge, there may be multiple crack edges due to the locking of regions behind the main crack edge. The stress drop in this case becomes inhomogeneous not only in space but in time as well. A problem of random variation of stress drop and strength over the fault was studied numerically by Mikumo and Miyatake (1979), though with a somewhat simplified model. The fracture process was found to be quite chaotic, with no clearly distinguishable fracture front. In such cases, a stochastic or a fractal approach may be instrumental.

It is now well known that some aftershocks occur off the main fault plane. Obviously, a complex seismic event may be accompanied by such shocks, occurring during the main earthquake rather than after it. This implies that at least part of the complexity of seismic radiation cannot be assigned to the main fault plane as is assumed in the models mentioned here. This is especially true when one is considering the high-frequency radiation from an earthquake.

In this section, we shall consider only some simple examples of these variations confined to the fault plane and determine the far-field seismic radiation due to fracture propagation on such a plane.

The barrier model

A barrier may be characterized by some measure of its areal extent and some measure of its strength. We may use the parameter S, defined in Section 5.6, to denote the relative strength of the barrier. If the areal extent of the barrier is large, the crack edge propagation will be arrested. But if its areal extent is small compared with the instantaneous crack dimension at the time it is encountered by the crack edge, the crack edge and the barrier will interact in the three different ways, depending on the value of S:

1. If S is small, the barrier will be broken as the crack edge encounters it.

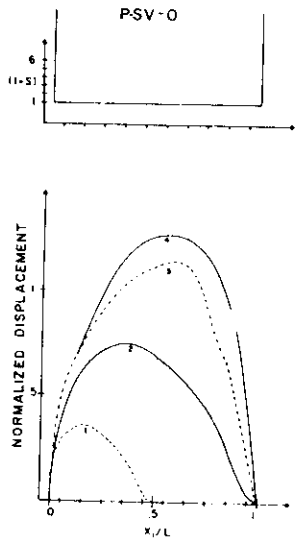
2. If S is very large, the crack edge will propagate around it, leaving behind an unbroken region.
3. If S has some intermediate value, the barrier will not be broken the first time it is encountered by the crack front but will eventually break due to the subsequent concentration of stress on it during the dynamic growth and slip of the surrounding areas.

The presence of such barriers on the fault will introduce diverse slip functions over the fault, which in turn will be seen as complexity in the radiated seismic wave forms and will modify the seismic moment of the earthquake. This problem was first studied two dimensionally by Das and Aki (1977b) and by Das (1985) for the three-dimensional case. Since the two-dimensional results are very complete and are now well known, we shall include only these results here. The three-dimensional calculations show that the two-dimensional results correctly predict the complexity of the far-field waveforms for a fault with barriers.

Let us confine our discussion to the unilateral propagation of an inplane crack. The four cases studied are listed below. The total crack length is taken as $10\,\Delta X$ and $S = 0$ in the areas without barriers for all the cases. The latter parameter value means that a critical crack length of zero is needed for dynamic propagation, which makes the calculations very economical. Backslip will not be allowed in all cases so that $a_{mjk_2'} > a_{mjk_1'}$, where $k_2' > k_1'$. Hence $A^{P}_{\text{initial}}(k\,\Delta t, \mathbf{m})$ is positive for all time and the maximum value of the amplitude spectrum is at zero frequency.

Case P-SV-0: There are no barriers on the fault, this case being included purely for the purpose of comparison. The crack extends at a speed close to v_p due to S being chosen as zero. The distribution of S and the resulting slip on the fault are shown in Figure 6.3. The far-field P pulse shape determined from equation (6.1.4) and the corresponding amplitude spectra are given in Figure 6.4.

Case P-SV-1: One strong barrier exists on the fault, and it remains unbroken when the dynamic fracture process on the fault is completed. The distribution of S and the slip on the fault are shown in Figure 6.5, and the far-field radiated field is plotted in Figure 6.6. The spectra for the case P-SV-0 is indicated in the latter figure by dashed lines.



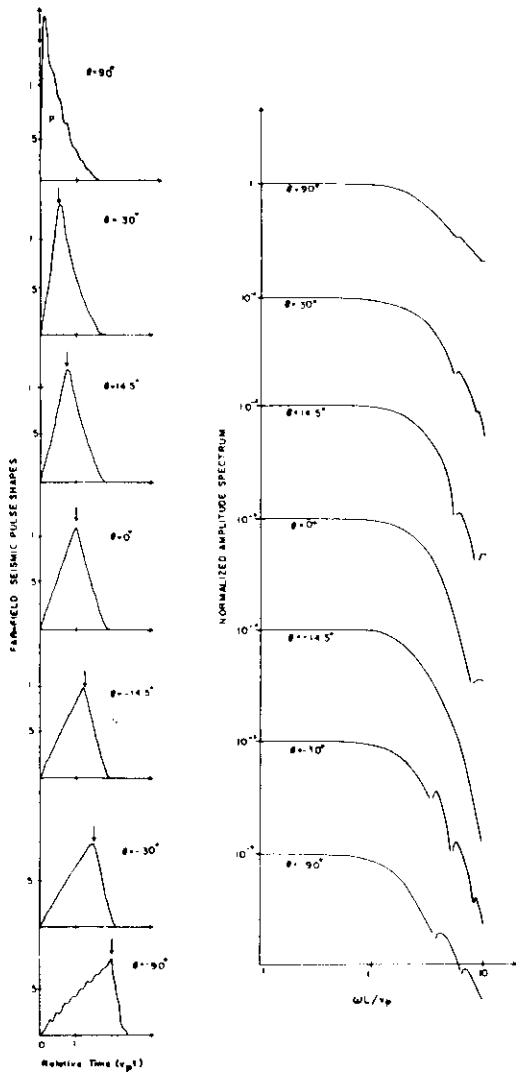
6.3 Distribution of the parameter S and "snapshots" of the distribution of the normalized half-slip $a_1/2$ over the fault length for case P-SV-0. The half-slip is normalized by $L\,\Delta\sigma/3\mu$, where L is the fault length and the integer next to each curve indicates the time measured in units of $.5L/v_p$. (*From Das and Aki, 1977b. © Am. Geophys. Union.*)

Case P-SV-2: Two unbreakable barriers exist on the fault. Figure 6.7 and Figure 6.8 show the corresponding S, the slip on the fault, and the far-field radiation. The dashed lines again give the P-SV-0 spectra.

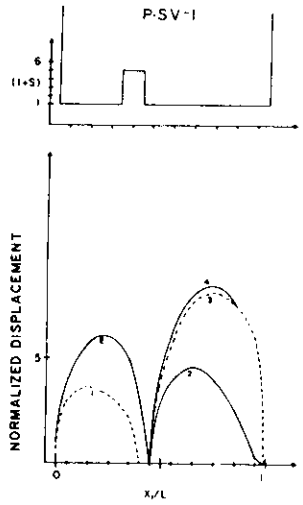
Case P-SV-3: The two barriers on the fault, having intermediate S values, do not break at the initial passage of the fracture front but break before the completion of the dynamic fracture and slipping process is completed. The related parameters and results are shown in Figures 6.9 and 6.10.

The major conclusions drawn from this example can be summarized as follows:

1. The smooth fault P-SV-0 and the P-SV-3 fault result in single earthquakes, whereas the heterogeneous faults P-SV-1 and P-SV-2 result in multiple shocks.
2. The time history of slip on the fault and the resulting far-field radiation are most complicated in the case when the initially unbreakable barrier eventually breaks (P-SV-3). In this case the duration of the fracture and slipping process are longer than in the other cases for the same fault length.
3. The final slip on the fault and hence the seismic moment are largest for the smooth crack (P-SV-0) and smallest for the case of the fault with two unbroken barriers (P-SV-2). In the case of the barrier that eventually breaks, the final slip and moment are almost as large as those for the smooth fault. The slip for the fault with two unbreakable barriers has the most uniform value over the fault, whereas the fault with no barriers at the end of the fracture process (P-SV-0 and P-SV-3) shows the largest amount of variation in slip distribution over the fault! This may explain why the uniform dislocation model (Haskell, 1964) has often been able to explain observed overall features of seismograms satisfactorily.
4. Clear directivity effects in the seismic radiation are seen in all cases, these effects being stronger for the fault with unbreakable barriers than for the smooth fault. However, when the barriers eventually break the directivity effect is even weaker than that for the smooth fault.
5. The time domain pulses are more sensitive to the complexity of the fracture process than the spectral shapes. In particular, when the barriers eventually break the pulses show complexity in all directions from the source, but the spectra are not particularly revealing.
6. When the barriers remain unbroken, the spectra at the highest frequencies for which the numerical results are meaningful (this



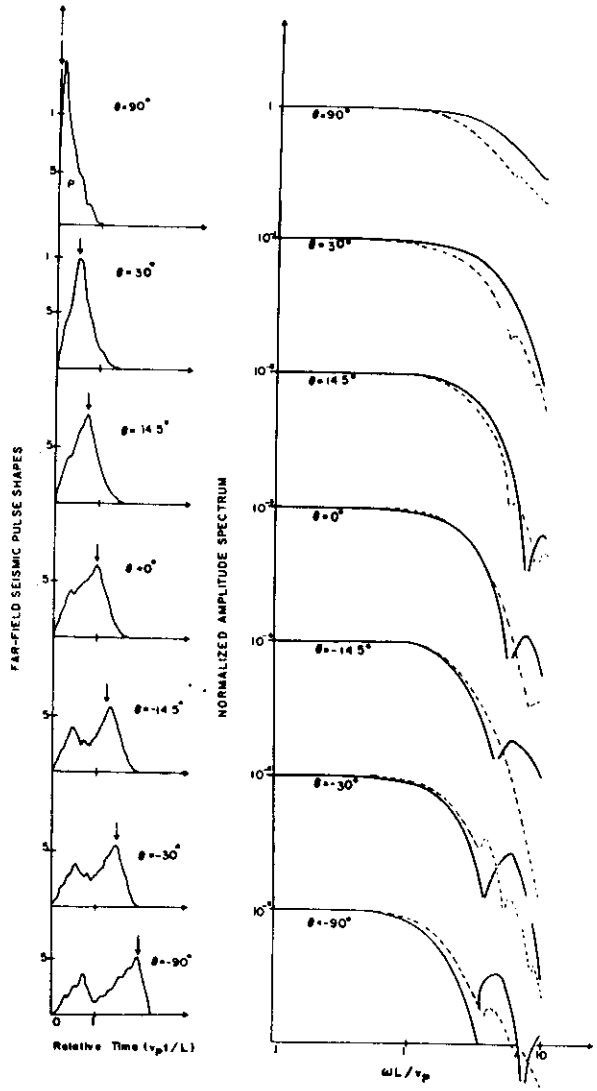
6.4 Far-field P-wave displacement pulse shape and amplitude spectra for various directions from the fault for case P-SV-0. The angle θ is measured from the normal to the fault. The arrows indicate the arrival of the first diffracted wave when the crack tip stops. For $\theta = 0°$ the P- and S-wave pulse shapes coincide. The amplitude spectra are normalized by their value at zero frequency (*From Das and Aki, 1977b. © Am. Geophys. Union.*)



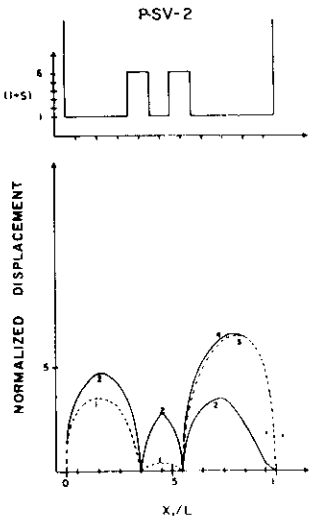
6.5 Same as Figure 6.3 but for case P-SV-1. There is now one barrier on the fault that remains unbroken at the completion of the dynamic fracture of the fault. (*From Das and Aki*, 1977*b*. © *Am. Geophys. Union*.)

limit can be obtained by comparing the numerical solution for some simple case with an analytic solution, the spectra in all the cases plotted in this example being shown only up to the frequency where the numerical results are valid) have more energy than that for the smooth fault.

7. The corner frequency averaged over all directions from the source is unaffected by the presence of unbreakable barriers.
8. The stress drop averaged over the total fault length (including the barriers) is lower for the case with unbroken barriers than the other cases. In fact, there is a stress increase on these unbroken regions due to the earthquake. Thus, a complex earthquake with lower average stress drop can generate waves of

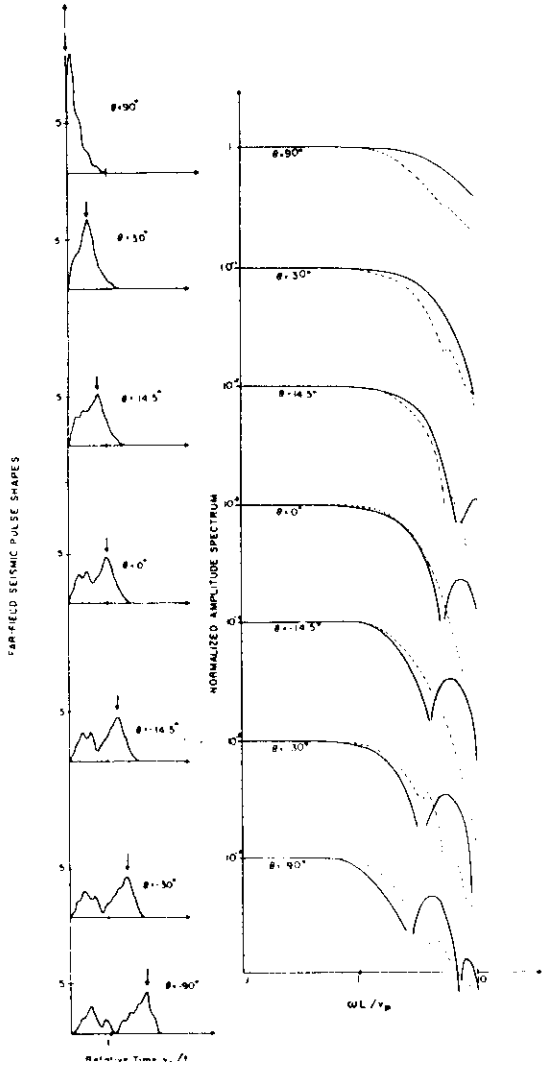


6.6 Same as Figure 6.4 but for the case P-SV-1. The dashed lines on the spectra are the curves for the case P-SV-0 and are included for the purpose of comparison. (*From Das and Aki*, 1977*b*. © *Am. Geophys. Union*.)

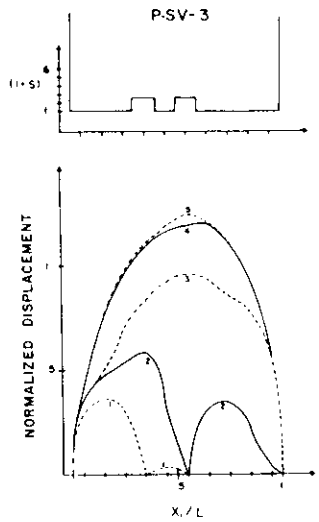


6.7 Same as Figure 6.3 but for case P-SV-2. The two barriers on the fault remain unbroken. (*From Das and Aki*, 1977*b*. © *Am. Geophys. Union*.)

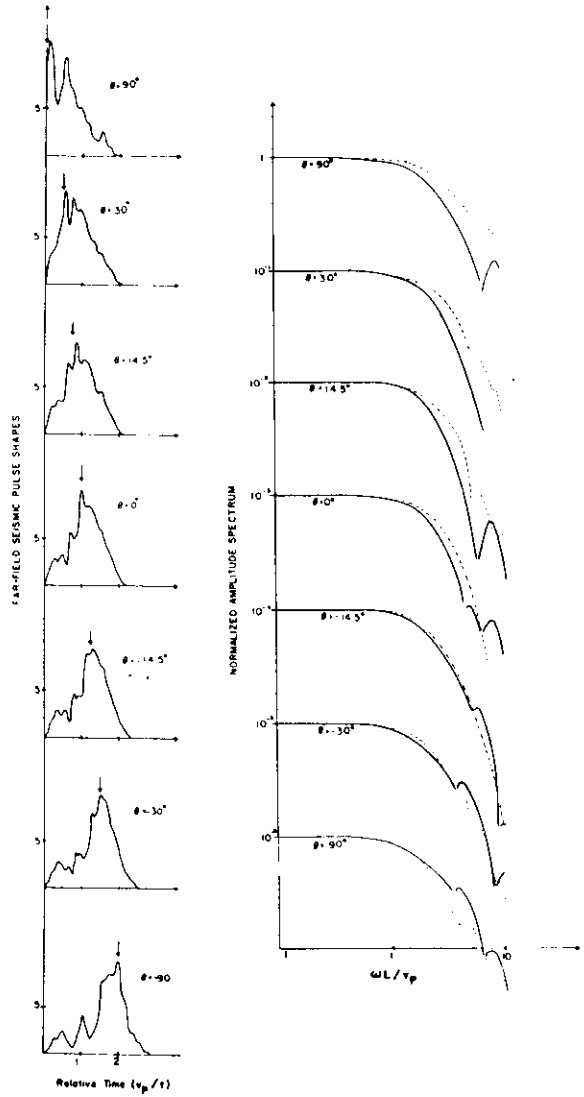
relatively higher frequency than a simple earthquake with relatively higher stress drop.



6.8 Same as Figure 6.4 but for the case P-SV-2. The dashed lines give the curve for the case P-SV-0. (*From Das and Aki*, 1977*b*. © *Am. Geophys. Union*.)



6.9 Same as Figure 6.3 but for case P-SV-3. The two barriers on the fault are of intermediate strength and eventually break while dynamic fracturing of other parts of the fault is continuing. (*From Das and Aki*, 1977*b*. © *Am. Geophys Union*.)



6.10 Same as Figure 6.4 but for case P-SV-3. (*From Das and Aki* 1977*b*. © *Am. Geophys. Union*.)

The observational support for complex faulting models came from both seismology and geology. Observations of multiple shocks on seismograms were mentioned at the beginning of this section. The measured

surface slip after large earthquakes often shows a form similar to the fault slip found for P-SV-1 and P-SV-2. Direct evidence from fractures on mine faces showed that faults are usually very complex, with side steps and highly deformed but unbroken ligaments in the stepover regions (Spottiswoode and McGarr, 1975; McGarr et al., 1979). The impact of this model, in spite of its idealizations, on the understanding of the earthquake faulting process was significant. It led to the characterization of barriers as being material (large S) or geometric (when the fault plane deviated from planarity) by Aki (1979). It also led to the identification of barriers in the field by structural geologists and by seismologists in various locations around the world (Lindh and Boore, 1981; King and Yielding, 1984; Nabelek and King, 1985; Sibson, 1986; Barka and Kadinsky-Cade, in press; Bruhn, Gibler, and Parry, 1987, to name only a few). Major projects are under way in many countries to identify barriers along faults and to try to understand the origin and geochemical characteristics of barriers. The primary reason for this general interest is that earthquakes often nucleate and terminate at barriers.

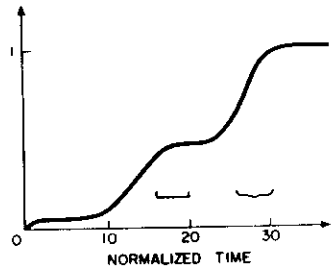
Since the unbroken barrier with its high residual stress concentration can become the "asperity" of a future earthquake on the same fault, it is important to consider the radiation due to the fracturing of such an unbroken barrier. In Section 5.6, we studied the dynamic fracture of isolated asperities of different shapes on infinite faults. In the next subsection, we will look at the far-field radiation generated by such a model.

Radiation due to the failure of an isolated asperity

The far-field displacement pulse shapes can be conveniently calculated for this case using (6.1.5). The corresponding radiation patterns were given in Section 4.6. Let us consider the far-field pulse shape for the circular asperity along the direction of the normal to the fault. In this direction, the P- and the S-wave pulses coincide. The pulses are given by the term

$$\sum_{i,\, j \in \Sigma_k} \tau_{\alpha ijk},$$

of (6.1.5) and plotted in Figure 6.12. The normalized time in the figure is $v_p t/\Delta X$, and it is measured from the time of arrival of the first wave at the receiver. The most striking feature of this pulse is that there is a permanent offset, in contrast to what we saw in the previous examples in



6.12 Far-field displacement pulse shape for the P and S waves along the normal to the asperity, due to the fracturing of a circular asperity (Figure 5.18). The bracket and brace are explained in the text. (*From Das and Kostrov, 1983. © Am Geophys. Union.*)

this chapter for the conventional crack model. This is not surprising when we recall that the problem was formulated such that the two half-spaces on either side of the fault plane remain permanently shifted after the asperity has fractured and disappeared. The rise time of the displacement from zero to this final value is the time required for the asperity to fracture. The brace in the figure indicates the time when the number of grids broken per unit time is the highest. The fracture process for this case (Figure 5.18) shows that this indeed took place toward the end of the breaking process. The square bracket indicates the period when the breaking rate is high but the displacement does not increase. This is because, although the number of grids breaking per unit time is large, these points are situated far from one another on the asperity; also, they do not have large stress drops associated with them and hence do not contribute significantly to the increase in the far-field displacement. If we looked at the acceleration pulse shape (obtained simply by twice differentiating the displacement pulse in Figure 6.12), we would find that the high accelerations correspond in time to the (relative) times when the breaking rate of grids on the asperity is the highest. Thus, the far-field displacements are very sensitive to the location of fracturing points on the asperity, whereas the accelerations are sensitive to the rate of increase of the broken area but not to its distribution over the fault. The pulse shapes in other directions from the source have essentially similar characteristics, the rise times being shortest in the (general) direction of fracture propagation and longest in the opposite direction.

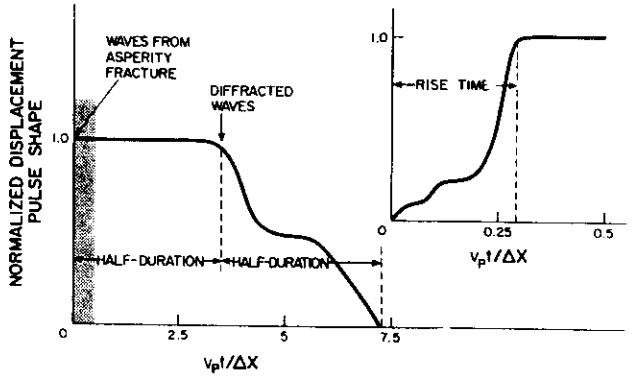
6.13 Geometry of the problem of failure of an isolated asperity on a finite fault.

The pulses for elliptical asperity fracture (Das and Kostrov, 1985) are similar and are not included here.

The asperity model

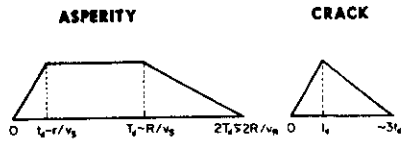
The basic idea of this model was suggested by Madariaga (1979) and by Rudnicki and Kanamori (1981). According to the model, an earthquake is caused by the failure of isolated, highly stressed regions of the fault, the rest of the fault having little or no resistance to slip (being partially broken and preslipped, say) and contributing little or no stress drop to the earthquake process. This results in a nonuniform stress drop over the fault. Since the regions without slip are able to withstand the high stresses concentrated on it until the moment the earthquake begins, it must be assumed that the parameter σ'' for these regions is higher than that for the rest of the fault. The spontaneous, dynamic fracturing of one or more such isolated asperities of general shape and size on a finite fault has not yet been studied. The simpler problem of radiation from the fracturing of a circular asperity at the center of a circular fault was studied by Das and Kostrov (1986), and we shall discuss the result here.

In this model, a circular crack of radius R, say, has a circular asperity of radius r, say, at its center (Figure 6.13). The annular region between the crack and the asperity is broken and assumed to be at or very close to the kinetic frictional level. When the central asperity breaks, this annular region exhibits no (or little) dynamic stress drop. It also has little or no resistance to slip. For the numerical calculations, r/R is taken as



6.14 Far-field S-wave displacement pulse shape due to the failure of the isolated asperity shown in Figure 6.12. Inset shows details of pulse shape in stippled region. (*From Das and Kostrov, 1986. © Am. Geophys. Union.*)

.1 and the asperity is taken as a single spatial grid. The asperity is released, and the ensuing dynamic slip is allowed to spread out over the entire circular fault. The slip is calculated using algorithm (5.2.13) and the P- and S-wave pulse shapes in different directions from the source are found using (6.1.4). The normalized S-wave displacement pulse shape looking down at the fault along the normal as a function of normalized time $v_s t/\Delta X$ is shown in Figure 6.14 as a representative example. The displacement pulse immediately reaches its maximum value and remains flat until the first diffracted waves from the crack edge arrive at the observer at time $T_d \simeq R/v_S$, measured from the time of arrival of the first S wave. The displacement then starts decreasing and finally reaches zero at time $\simeq 2T_d$ ($\simeq 2R/v_R$). The minor oscillations following this that occur due to backslip being permitted on the fault are ignored in this figure. The duration of the flat part of the pulse is thus controlled by the size of the large crack of radius R. Since the asperity was released instantaneously in this problem, this picture does not represent the rising part of the pulse correctly. But this was calculated in the last subsection, and using those results and adjusting the time scale, we obtain the pulse shape in the stippled region of Figure 6.14, as shown in the inset of this figure. The rise time for the failure of a single asperity was shown in the



6.15 Schematic representation of the far-field displacement pulse due to the failure of an asperity on a finite fault and due to a propagating crack. (*From Das and Kostrov, 1986. © Am. Geophys. Union.*)

last subsection to be controlled by the asperity size and is given by $t_d = r/\beta$.

Since th Details of the pulses of Figure 6.14 depend on the parameters of the particular problem (crack and asperity size and shape, fracture velocity, etc), one may neglect the details and construct a schematic representation of the pulse shape due to the fracture of an isolated asperity on a finite fault, as shown in Figure 6.15. The main features of this pulse are a steeply rising part followed by a flat portion of long duration and then a gradual return of the displacement pulse to zero. The triangular pulse from a circular crack (Figure 6.1) is also shown in the figur for comparison. If this circular crack is taken to be the same size as the asperity of the model under discussion here, then the pulse shape would have the same rise time as the asperity model. However, once the maximum amplitude is reached, the two pulses become very different in character, the crack pulse immediately starting to decrease toward zero and reaching zero at time about twice the rise time, as we saw earlier in this chapter.

Thus, the asperity pulse has an anomalously large seismic moment (area under the pulse) and anomalously large duration compared with a crack pulse for a crack of radius r. Such earthquakes have been called "slow" or "weak" earthquakes (Kanamori and Cipar, 1974; Kuznetsova et al., 1976). The spectrum of the pulse shown in Figure 6.14 was found to have the same general form as that for the conventional crack model (Figure 6.1).

It must be pointed out here that this model is not the only possible model for slow earthquakes. Clearly, such earthquakes could also be modeled as a very slowly propagating crack, due, for example, to very low stress drop.

In geophysics, as a whole, forward problems are of lesser interest serving only to elucidate underlying physical phenomena More important are inverse problems — problems that require the distribution of material parameters and motions in the earth's interior to be determined from surface observations.

The inverse problem for the earthquake source has been formulated as one of reconstructing the displacement discontinuity and its history over the fault surface at the source. A common feature of inverse problems in geophysics is that many different distributions at the source can produce the same surface observation. Due to this instability, it is not possible to find the motions at the source without additional a priori information. In the paper by Das + Kostrov (1990), attached next, the problem is set up and additional meaningful information to include in the solution are searched for. Finally, the method developed is applied to the 1986 Andreanof Islands, Alaska earthquake.

You will recognize equation (1) from our previous lectures.

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 95, NO. B5, PAGES 6899-6913, MAY 10, 1990

Inversion for Seismic Slip Rate History and Distribution With Stabilizing Constraints: Application to the 1986 Andreanof Islands Earthquake

SHAMITA DAS[1]

Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York

B. V. KOSTROV

Institute of Physics of the Earth, Academy of Sciences, Moscow, USSR

We develop and present a method of inverting seismograms, including their absolute amplitudes, to obtain the complete seismic slip time history and distribution, using the method of linear programming. A rectangular fault is discretized into cells and the source duration into time steps. The inversion for the slip rate distribution on the fault as a function of time reduces to the solution of a system of linear equations. The influence on the solution of constraining different source parameters such as the sign of the slip rate and the seismic moment is investigated. It is found that the constraint of positiveness of slip rate and the constraint that the seismic moment determined from the solution should be close to that obtained from the centroid moment tensor (CMT) solution are very strong constraints and are essential to obtaining a physically reasonable solution. Since the inverse problem is unstable, instead of presenting only one particular solution we examine several solutions equally well fitting the data for all practical purposes. We find that when both constraints are applied the solution becomes quite stable. The method is applied to a data set consisting of the vertical component of available digitally recorded teleseismic P waves at stations within a chosen distance window for the 1986 Andreanof Islands earthquake ($M_w = 8.0$). The preferred solution is the one in which the slip rate is constrained to be positive and the moment is set to the CMT value. The average rupture speed in the 90-km segment to the east of the hypocenter was ≈ 3.2 km/s and in the 190-km section to the west was ≈ 2 km/s. The maximum slip in the area of large moment release to the west was ≈ 5 m and the maximum slip rate, which occurs in the same place as the maximum slip, was ≈ 6.5 m/s. The slip distribution on the fault shows a corrugated pattern with the long axis in the direction of the plate convergence, rather than normal to the arc, which was the slip vector direction for the earthquake, with most of the slip occurring on a region of the fault to the west of the rupture nucleation zone. This region continued to slip for about 55 s, the total source duration being about 80 s. The results of this analysis suggest that the long-term motion on this portion of the Aleutian arc is indeed in the direction of plate convergence.

INTRODUCTION

During the last decade several attempts were undertaken to solve the inverse problem for the source of particular earthquakes, that is, to determine the spatial and temporal distribution of slip (or slip rate) over the fault area, using teleseismic as well as near-field waves. From a very extensive literature on this subject, we refer here only to papers which describe a method for a true inversion for areal as well as temporal distribution of slip or moment on the fault area, namely, *Olson and Apsel* [1982], *Hartzell and Heaton* [1983], *Kikuchi and Fukao* [1985], and *Beroza and Spudich* [1988]. In addition, much work has been done on determining the source time function for specific earthquakes and the result interpreted in terms of spatial moment release using a constant rupture velocity as well as on inversion considering the source as a line source. In these categories, one may consider the papers of *Kikuchi and Kanamori* [1982], *Ruff* [1983], *Ekstrom* [1987, 1989], and *Boyd and Nabelek* [1988]. Finally, the inverse problem for the static case, namely, the determination of the final slip distribution on the fault using geodetic data, has been studied by *Ward and Barrientos* [1986].

The solution of all these problems are far from trivial. It is well known that this inverse problem is unstable, even in the imaginary case of continuous distribution of seismic stations over the surface of Earth and its stability and uniqueness properties have been discussed extensively by *Kostrov and Das* [1988]. From the computational point of view, this instability is equivalent to nonuniqueness of the solution. The real situation is even worse because the number of stations with appropriate records is very limited (about 10-20 global stations, at present). Consequently, to obtain a definite solution of such a problem, one needs some physical constraints on the source process, in addition to the requirement of fitting the observed seismograms. In principle, these constraints should be inferred from the physics of faulting during the earthquakes, that is, from fracture and frictional mechanics. Unfortunately, our knowledge of the physics of the earthquake process is still rather limited and the only comparatively well-established constraint is the limitation on the fracture propagation speed. This is a weak constraint because the duration of the earthquake process is greater than the seismic wave propagation time across the fault whereas the limiting fracture speed is comparable to the seismic wave velocities. To implement this constraint numerically, one requires very fine spatial gridding on the fault. A less physically founded constraint would be the requirement that the slip rate vector be directed in accordance with the average stress drop direction (inferred from the fault plane solution), that is, the projection of the slip rate vector in the stress drop direction must be positive. While, in principle, slip in the opposite direction is possible due to interference of waves on the fault, it does not seem likely and experience from three-dimensional forward modelling shows that the slip direction almost coincides with that of the stress drop [*Das*, 1981]. This suggests that with sufficient practical accuracy, one may assume the slip direction to be constant over the fault during the process and coincide with the stress drop direction and the only component of the slip rate vector to be nonnegative. This constraint has been used by *Hartzell and Heaton* [1983], *Kikuchi and Fukao* [1985], *Ward and Barrientos* [1986], and *Ekstrom* [1987, 1989], among others. In what follows, we shall refer to this constraint as the "no backslip" constraint. As we shall see later in the paper, other possible constraints may be considered. For example, one may require the solution to be in agreement with the seismic moment obtained from the centroid moment tensor solution or from geodetic measurements. Or one may require the maximum slip rate on the fault to be limited by some considerations from fracture mechanics. *Hartzell and Heaton* [1983] minimized the seismic moment, which is another possible constraint on the solution. We shall investigate the effect of different constraints on the inversion process, in this paper.

To invert for the three-dimensional slip rate distribution (two spatial dimensions on the fault and time) is a computationally difficult task. With sufficiently fine gridding of the fault, it is almost impracticable in a regular way, even on existing supercomputers. Not surprisingly, simplified methods of solving the problem have been used. Basically, such simplifications are always a version of trial and error fitting. In any case, the method of the solution imposes some implicit constraints. Then, when one obtains a unique solution, it is not clear if it is unique due to the explicitly formulated constraints or as a result of the method used in the inversion. As an example, the generalized inverse (or pseudoinverse) leads to a unique solution which minimizes the root mean square residual even without additional constraints. In that case the implicit constraint consists of the requirement that the solution has the minimum norm. Or, in the method of *Kikuchi and Fukao* [1985] the no backslip constraint was explicitly imposed, but the method consisted of successive reduction of the root mean square residual by fitting a series of essentially instantaneous point sources. This again imposes an implicit requirement that the solution must be concentrated in as small a number of discrete patches on the fault, as possible. The interpretation of the solution obtained in this manner in terms of the asperity model is only natural, but this interpretation as being the only and unique interpretation is not supported by the data because it was implicitly imposed by the method, and the question whether there exists another, more smoothly distributed solution, equally well fitting the data for all practical purposes and satisfying the no backslip constraint, remains open.

The problem for inversion for spatial and temporal distribution of seismic slip was first formally solved by *Olson and Apsel* [1982], and their paper deserves detailed discussion here. In that paper, the need for additional constraints apart from observations is most explicitly discussed and the mathematical aspects of the inverse problem are extensively presented. To quote from them: [*Olsen and Apsel*, 1982 p.1969]

> In order for a particular slip distribution to be an acceptable solution to the inverse problem it must satisfy the following three conditions.
>
> 1. The solution must explain the data.
> 2. The solution must be physically reasonable (consistent with independent constraints).
> 3. If more than one solution fits the data equally well, additional information must be supplied to uniquely define which solution is being obtained.

The third condition means that the physical constraints may be insufficient to specify an unique solution of the problem and they proposed to include some additional nonphysical assumptions to make the solution unique. In their paper they used the requirement that the solution has minimum norm. We suggest in this paper that a more reasonable approach to this difficulty would be to describe the whole set of equally acceptable solutions by obtaining some extreme representatives of this set, say.

For physical constraints, *Olson and Apsel* [1982] used form of the no-backslip constraints as well as limiting of the rupture propagation speed. Additionally, they required slipping to be confined to grids intersecting the rupture front and limited the number of times each grid could slip. Effectively, their model included two preassigned fronts, the rupture front and the healing one between which the whole slip process was confined. This last assumption permitted drastic reduction of the number of unknowns in the model. Numerical solutions of the forward problem show, however, that with an inhomogeneous strength and friction on the fault, the rupture front can become very distinct from a single line and subsequent slip behind the main rupture front frequently obtained [*Das and Aki*, 1977; *Mikumo and Miyatake*, 1979]. As we shall see later, such reslipping of previously slipped regions on the fault is found in this study. Consequently, this assumption can be considered only another example of the above mentioned simplification the problem. *Hartzell and Heaton* [1983] and *Beroza and Spudich* [1988] also included this simplification in their inversions thus decreasing the size of the problem. In this paper, we shall not confine the slippage to a region near the rupture front, nor shall we a priori assign the number times each grid can slip but shall determine the entire slip time history at every spatial grid on the fault.

The discretization of the problem can itself imply unwanted constraints on the solution. Namely, with a coarse grid the number of unknowns can be reduced so much to make the corresponding numerical problem stable, that is, well conditioned, without additional physical constrain

[1] Now at Department of Earth Sciences, University of Oxford, Oxford, UK.

Paper number 90JB00701.
0148-227/90/90JB-00701$05.00

the extreme cases being point source fitting, and fitting the source to a rectangular dislocation with constant slip and constant slip rate.

The main purpose of this paper is to develop and test a method for determining the complete slip time history on a fault during an earthquake and to examine the stability of such a solution by examining not only a "best fitting" solution but also solutions close to it in data space.

Formulation of the Discrete Problem

We briefly summarize the formulation of the discrete problem under study for completeness and clarity, particularly of the notations used in this paper. Using the representation theorem (e.g., equation (3.2) of *Aki and Richards* [1980]; equation (3.2.18) of *Kostrov and Das* [1988]) the displacement record at a station located at a point $\mathbf{x}$, on the earth's surface can be expressed in terms of the slip distribution over a fault Σ as

$$u_k(\mathbf{x}_1, t_1) = \int_0^{t_1} dt \iint_\Sigma K_{ik}(\mathbf{x}_1, \mathbf{x}, t_1, t) a_i(\mathbf{x}, t) dS \qquad (1)$$

where $i, k = 1, 2, 3$, $u_k(\mathbf{x}_1, t_1)$ are the components of the displacement vector, $a_i(\mathbf{x}, t)$ are the components of the slip and $K_{ik}(\mathbf{x}_1, \mathbf{x}, t_1, t)$ are the components of the impulse response of the medium at $(\mathbf{x}_1, t_1)$, due to a dislocation point source at $(\mathbf{x}, t)$. The observed seismograms do not represent the displacement vector u itself but are filtered by the instrument. Convolving both sides of (1) with the instrument response for a given station and assuming the slip direction to be constant, we obtain, after some transformations:

$$S_j(t_1) = \int_0^{t_1} dt \iint_\Sigma W_j(\xi, t_1 - t) \dot{a}(\boldsymbol{\xi}, t) dS \qquad (2)$$

where j identifies the station and components of the seismogram $S(t_1)$, $W_j(\boldsymbol{\xi}, t)$ is the impulse response at $(\boldsymbol{\xi}, t)$ corresponding to a fixed slip direction and convolved with the instrument response, $\dot{a}(\boldsymbol{\xi}, t)$ is the slip rate, the two-dimensional vector $\boldsymbol{\xi}$ gives the position on the fault relative to some reference point (for example, the earthquake hypocenter) and t is the time measured from the origin time of the earthquake. With a continuous distribution of stations, equation (2) would represent an integral equation of the first kind. Such equations are known to be unstable.

In this paper we shall consider the effect of the following additional constraints:

$$\dot{a}(\boldsymbol{\xi}, t) \geq 0 \text{ for all } (\boldsymbol{\xi}, t) \qquad (3a)$$

$$\dot{a}(\boldsymbol{\xi}, t) = 0 \text{ for } t < T(\boldsymbol{\xi}) \qquad (3b)$$

where, $t = T(\boldsymbol{\xi})$ gives the boundary of the area where slip is permitted (due to a causality condition) at time t,

$$\int_0^\infty dt \iint_\Sigma \mu(\boldsymbol{\xi}) \dot{a}(\boldsymbol{\xi}, t) dS = M_0 \qquad (3c)$$

where M_0 is the seismic moment and μ is the modulus of rigidity of the medium, together with other constraints to be discussed later in the paper. The constraint (3a) is the "no-backslip constraint," the constraint (3b) is the "causality constraint" and the constraint (3c) is the "seismic moment constraint." Note that $T(\boldsymbol{\xi})$ as defined above is not the usual rupture front, but in what follows we shall refer to this boundary as the "rupture front" to distinguish it from the usual definition.

We restrict ourselves to long-period body waves at teleseismic distances in this paper because they are relatively insensitive to the details of earth structure [*Hwang*, 1985]. The functions W_j can then be easily calculated using the ray approximation and the Thomson-Haskell technique to account for the layered structure at the source and the station. We follow the procedure described by *Nabelek* [1984] to determine the source and station crustal responses and follow *Ekstrom* [1987] to construct the synthetic seismograms by accounting for the geometrical spreading and attenuation factors.

For the numerical solution, the integrals in (2) must be discretized. We divide the fault area Σ into a number of rectangular cells and approximate $\dot{a}$ within each cell by linear functions in time and along strike and by a constant along dip. W_j is then integrated over each cell analytically, and the integrals over the fault are replaced by sums. The time at the source is discretized by taking a fixed time step, Δt say, and assuming that the slip rate $\dot{a}$ during the time step varies linearly with time. Since we use only long-period body wave records, we may use comparatively large Δt, this being desirable to reduce the number of unknowns. Strictly speaking, both the fault area and the total source duration should be determined as part of the inversion process. Due to limited computer power, however, we shall assume a finite fault size and a finite source duration which cannot be longer than the longest record used and cannot be much shorter than that obtained from the centroid moment tensor solution. The seismograms S_j are sampled with a time step Δt_1, say, usually 1 s for long periods. This value might seem redundant, but it is worth keeping to obtain better control of noise, if of course the computer power allows it. Let us renumber the observations in a one-dimensional way, ordering them by component, station number and time. Denote the value of $S_j(t_1)$ by b_k, say, k being the index in the ordering adopted. Similarly, let us renumber the values of $\dot{a}$ by cell number and time, denoting the vector by x_i, say. In addition, to allow for the possibility of weighting the different stations and components differently in the inversion, let us include this weighting into equation (2). Then equation (2) takes the form

$$A\mathbf{x} \approx \mathbf{b} \qquad (4)$$

where A is the matrix obtained by integration of W_j, each column of A being a set of synthetic seismograms for all stations corresponding to different cells and time instants of the source duration, ordered in the same way as the observed seismograms and A and b are appropriately weighted. This arrangement of the system of linear equations is very similar to that used by *Hartzell and Heaton* [1985, Figure 11]. The condition (3a) then becomes:

$$\mathbf{x} \geq 0 \qquad (5a)$$

the inequality meaning that every component of x is nonnegative. The condition (3b) can be replaced by

$$x_i = 0 \qquad (5b)$$

for those i corresponding to cells and time samples outside the "rupture front" and the condition (3c) becomes:

$$\Sigma c_i x_i \approx M_0 \qquad (5c)$$

where c_i is time-independent and for each cell is equal to the product of the average rigidity times the area of the cell. So, the inverse problem has been reduced to the solution of the linear system (4) under one or more of the constraints (5). In the system (4), the number of equations m is equal to the total number of samples taken from all the records involved and the number of unknowns n is equal to the number of cells times the number of time steps at the source. We shall take m greater than n to reduce the influence of the noise contained in the observations b on the solution. Then, the system (4) is overdetermined and we can only obtain a solution x which provides a best fit to the observations, under constraints (5).

It is well known that the matrix A is ill conditioned which implies that the system (4) admits more than one solution, equally well fitting the observations. The constraints (5) are introduced just for the purpose of reducing the set of permissible (feasible) solutions. It is to be investigated if these conditions are sufficient to make the solution unique. If not, then we have to obtain some description of the set of permissible solutions of the problem. In any case, even if an unique solution does exist, there may be many other solutions that almost satisfy the equations. Since the data used in geophysical applications often contains experimental noise and the models used are themselves approximations to reality, solutions almost satisfying the data are also of great interest.

The system of equations (4) together with the constraints (5) do not yet comprise a complete mathematical problem. It remains to formulate in exact form what the "best fit" to observations means. In the next section, we only include constraint (5a) in the mathematical formulation for the sake of simplicity, the inclusion (5b) being trivial and the inclusion of (5c) will be discussed later. We have to minimize the vector of residuals:

$$\mathbf{r} = \mathbf{b} - A\mathbf{x} \qquad (6)$$

For this purpose, some norm of the vector r must be adopted Usually in such problems one may choose to minimize the l_1, the l_2 or the l_∞ norm [*Noble and Daniel*, 1977; *Tarantola*, 1987], all three being equivalent in the sense that they tend to zero simultaneously. In this paper we shall use the linear programming method to solve the system (4) and minimize the l_1 norm subject to the condition (5a), but we shall also evaluate the other two norms of the solution to investigate how they behave.

Linear Minimization Problem

To express the l_1 minimization problem in the standard form of linear programming, we represent the residual vector r as the difference of two vectors with nonnegative components:

$$\mathbf{r} = \mathbf{y}^+ - \mathbf{y}^-$$

Then

$$\Sigma|\mathbf{r}| \leq \Sigma(\mathbf{y}^+ + \mathbf{y}^-)$$

Obviously

$$\min\Sigma|\mathbf{r}| = \min\Sigma(\mathbf{y}^+ + \mathbf{y}^-)$$

Introducing additional unknowns as

$$\left.\begin{aligned} x_{n+i} &= y_i^+ \\ x_{n+m+i} &= y_i^- \end{aligned}\right\} \text{ for } i = 1, 2, \ldots, m$$

the problem can be rewritten as follows:

$$\text{minimize } f = (1/m) \sum_{i=n+1}^{n+2m} x_i \qquad (7)$$

subject to the constraints:

$$\sum_{j=1}^{n} a_{ij} x_j - x_{n+i} + x_{n+m+i} = b_i, \quad i = 1, 2, \ldots, m \qquad (8)$$

$$x_i \geq 0, \quad i = 1, 2, \ldots, n + 2m \qquad (9)$$

In equation (7), f is the mean absolute residual.

The constraints (8) and (9) define a convex polytope in $(n + 2m)$ dimensions, each point of which represents a feasible solution. In general, the feasible set is a continuum containing an infinite number of feasible solutions. It is well known [*Franklin*, 1980] that the feasible solution which gives the minimum (7) corresponds to a vertex of the polytope. We use the simplex method of solving the linear programming problem [*Press et al.*, 1986].

Observational Data. The 1986 Andreanof Islands Earthquake

To illustrate the ideas and the method of inversion developed in this paper, we select a recent earthquake, namely the May 7, 1986, Andreanof Islands earthquake ($M_w = 8.0$; origin time = 2247), for which digital data were available for a several stations in the required epicentral distance range. This was the largest earthquake to have occurred in the central Aleutians since the March 9, 1957, Aleutian Islands earthquake. The along-arc rupture length of 280 km for this earthquake was contained entirely within the 1150 km long rupture of the 1957 event. Thus a second great event occurred only 29 years later on a plate boundary which is believed to have a much larger repeat time [*Sykes et al.*, 1981]. This earthquake has been studied by several authors using teleseismic data [*Hwang and Kanamori*, 1986; *Ekstrom*, 1987; *Boyd and Nabelek*, 1988; *Houston and Engdahl*, 1988, 1989] and the centroid moment tensor (CMT) solution has been published [*Dziewonski et al.*, 1987]. The tectonic setting of the earthquake and its two-week aftershock distribution along with the focal mechanism obtained by *Boyd and Nabelek* [1988] is shown in Figure 1. The earlier studies of this earthquake suggest that nonuniform along arc moment release occurred during this event, which partially motivated our choice of this event for detailed analysis.

In this paper, we shall consider only the vertical component of P waves recorded by long-period stations with epicentral distances between $\Delta = 33°$ and $\Delta = 75°$. The upper bound of the range was chosen because beyond this distance other unmodelled phases may influence the seismograms. The lower bound was chosen because at epicentral distances shorter than this the uncertainties in crustal structure may have significant effect. Several Global Digital Seismic Network (GDSN) stations in this range were found but on close examination of the seismograms only 10 stations had usable P wave records. This is because records were either clipped or the instrument response was found to be obviously nonlinear. We also examined digital data from Network of Autonomously Recording Stations (NARS) operated by the Dutch and from GEOSCOPE operated by the French. Data from neither of these networks were used either because the few stations which did record this event were located close to GI SN stations used (for example,

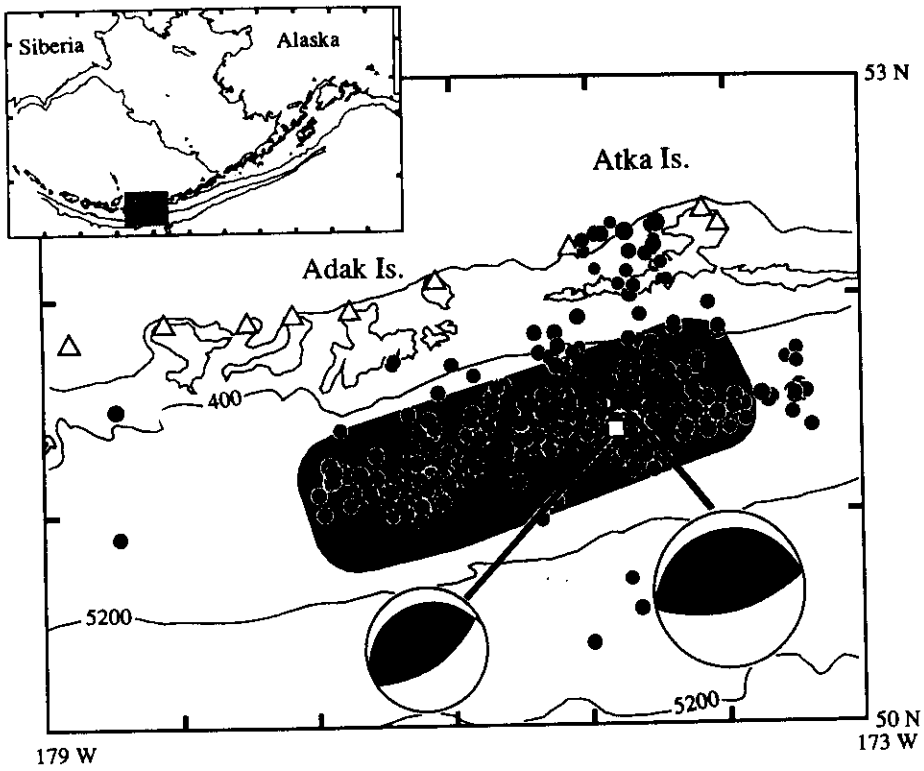


Fig. 1. Tectonic setting of the 1986 Andreanof Islands earthquake ($M_w = 8.0$) and its 2-week aftershock distribution. The focal mechanisms for the main shock (indicated by a star) and the large foreshock (indicated by a square) are also shown. The shaded region indicates the 1-day aftershock area. (Modified from *Boyd and Nabelek* [1988].)

Scandinavia and northeastern United States, where several GDSN stations were available in each region) or because the stations were well beyond our chosen distance window. In particular, no GEOSCOPE stations were then operational in the southern Pacific, which is also the quadrant in which no other stations were available (see Figure 2). We consider 100 s of the seismogram from the arrival of the P wave. Though the seismogram at GDH was clipped 25 s after the P wave arrival, it was retained to improve azimuthal coverage and the length of the record used was 25 s. It must be noted that though many stations had to be rejected from the P wave analysis in this paper, these stations along with stations at distances larger than $\Delta = 75°$ were used in the CMT solution, a detailed log of which was examined by us. The CMT solution used 150-s waves at 22 stations, with distances between $\Delta = 7°$ and $\Delta = 123°$ and azimuths between 7° and 357° for this particular event. Instruments which became nonlinear during the body wave arrivals often recovered before the mantle waves arrived so stations which had to be excluded from the body wave analysis were included into the CMT analysis. Thus this solution had very good azimuthal coverage and the fit between synthetics and records for the mantle waves was also good. Finally, the fact that the strike, dip and rake of the CMT solution generally agrees with that obtained by *Boyd and Nabelek* [1988] from the analysis of both P and SH waves suggests that the well-known ambiguities of the CMT moment due to instability of the solution for $M_{r\theta}$ and $M_{r\phi}$ for shallow depth earthquakes [*Dziewonski and Woodhouse*, 1983] are not a problem and the CMT solution together with its moment is reliable for this earthquake. Figure 2 shows all the seismograms used plotted around the CMT solution, and it is seen that the azimuthal coverage is not uniform around the epicenter. The work of *Hartzell and Heaton* [1983], *Hwang* [1985], and *Mendoza and Hartzell* [1988, 1989] shows that reasonable estimates of the slip history and distribution on a fault can be obtained with nonuniform azimuthal coverage. In particular, *Hartzell and Heaton* [1983] and *Mendoza and*

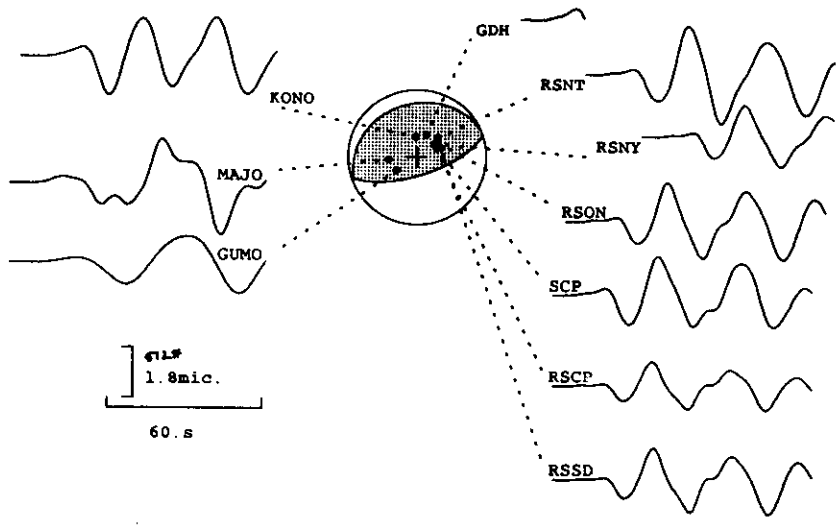


Fig. 2. The focal mechanism, obtained by *Boyd and Nabelek* [1988], together with the long-period seismograms used in the inversion for the 10 GDSN stations. Only the portion of the seismogram used in the inversion is plotted. The traces are adjusted for instrument magnification and normalized by the maximum value among all stations, correctly representing the relative ground motion at the stations.

Hartzell [1988, 1989] studied the 1979 Imperial Valley earthquake and the 1985 Michoacan earthquake, respectively, using both teleseismic P waves with poor azimuthal coverage and strong ground motion data and in both cases found that the P wave solution was in remarkable agreement with the strong motion solution, apart from some details. This suggests that the nonuniformity of azimuthal coverage is not a major obstacle for our purpose. The seismograms were lined up by the arrival time of the short-period P wave on the vertical component, when available or by using the expected P wave arrival times according to the PREM when the short period data for the particular station was not available. Table 1 lists the station codes for these ten stations, the type of instrument, the epicentral distances Δ, the forward and back azimuths Az and Az' and the arrival times picked. The crustal structure at the source and receiver are the same as that used by *Boyd and Nabelek* [1988]. At the source, the crust is approximated by a half-space overlain by a water layer of thickness 1.5 km. The P and S wave velocities of the half-space, v_P and v_S, are 6.6 and 3.8 km/s, respectively and the density is 2.9 g/cm^3. The water layer is taken to have a compressional wave speed of 1.45 km/s and a density of 1.02 g/cm^3.. The receiver structure is taken as a half-space with v_P and v_S of 6.0 and 3.5 km/s and a density of 2.8 g/cm^3.

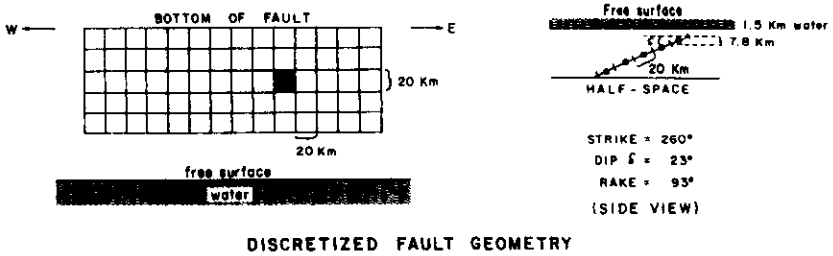
TABLE 1. List of Stations

Station	Instrument	$\Delta°$	$Az°$	$Az'°$	Arrival Time
GDH	DWWSSN	51.9	22.4	317.8	2256:18
RSNT	RSTN	33.2	46.9	279.4	2253:49
RSNY	RSTN	61.9	52.2	316.0	2257:30
RSON	RSTN	48.1	56.5	304.3	2255:51
SCP	DWWSSN	63.0	57.2	316.0	2257:37
RSCP	RSTN	62.4	66.0	315.2	2257:34
RSSD	RSTN	46.2	69.4	305.3	2255:36
GUMO	SRO	49.8	235.0	31.9	2255:59
MAJO	ASRO	36.2	264.1	50.8	2254:12
KONO	ASRO	68.8	357.3	3.4	2258:15

Note that the PcP phase arrives at some stations within the 100-s data window used here but at all the stations PP arrives outside this window. A simple calculation using PREM shows that the expected PcP amplitudes are only a few percent of the P wave amplitude for the source-station distances used in this analysis and we may therefore neglect it. Body wave inversions often use time windows which do not exclude large unmodelled phases and use source-station distance up to the P wave shadow zone but still model the body wave synthetics used in the inversion only as the P group ($P + pP + sP$). This maps unmodelled effects into the solution and contaminates it. To avoid this problem, we restricted our choice of source-station distance range and time window. In any case, no usable stations exist between $\Delta = 75°$ and 90° and using stations beyond 90° introduces problems of unreliability in Earth models [*Ekstrom*, 1987].

Discretization of the Fault

A 280 km × 100 km section of the fault is divided into 20 km × 20 km cells with 14 cells along the fault strike and 5 cells along the fault width. The top of the fault is located at a dept 7.8 km below the bottom of the water layer. The distribution of the seventy cells on the fault is shown in Figure 3. The strike, dip and rake of the fault, also



DISCRETIZED FAULT GEOMETRY

Fig. 3. (Left) Distribution of the seventy 20 km × 20 km cells on the fault, oriented as in Fig. 1. E and W denote the easterly and westerly directions, along the arc. The cell within which rupture nucleated is shaded. (Right) Side view of discretized fault. The top of the fault is located at a depth of 7.8 km below the bottom of the water layer.

shown in Figure 3, are taken from the analysis of *Boyd and Nabelek* [1988], who used both P and SH waves to constrain the source mechanism. Based on their solution, we take the source duration as 80 s and sample it at 5-s intervals using a linear approximation to obtain 16 time steps in the discrete source duration. Ideally, one should consider a fault size much larger than the fault size expected, for example, from aftershock studies and source duration as long as the longest seismogram duration and determine the actual source duration as part of the solution but to reduce computation time, we do not use this approach. The grid size of 20 km is $\approx 1/6$ of the P wavelength and $\approx 1/4$ of the S wavelength.

We now include a further, rather weak, constraint into the problem. Namely, we assume that the slip rate at any cell and time step which would produce a signal before the first arrival at any station is zero. Then the number of unknowns $n = 909$ and the number of equations $m = 925$, for this case.

Results

We solve the linear minimization problem, that is, we solve equation (4) subject to conditions (5a) and (5b). In the simplex method of linear programming, the size of the matrix involved is $n \times (n+m)$ and the problem must generally be solved using a supercomputer. Both sides of (4) are divided by the maximum value of the instrument response at the corresponding station to adjust for different instrument magnifications and all stations are uniformly weighted. We determine the complete slip rate time history and distribution which occurs during the faulting process. The moment and moment rate distributions as a function of time are derived from the calculated slip rates. Figure 4 shows the moment rate versus time (source time function) and Figure 5 shows the comparison of the data with the solution. Let us define the misfit in the l_1 sense as the ratio of the mean absolute error of fit to the mean absolute amplitude of the data. Similarly, the misfit in the l_2 sense would be the ratio of the rms error of fit to the rms amplitude of the data, and so on. (Note that our definition of the l_2 misfit is the square root of the misfit usually used by other authors, for example, *Boyd and Nabelek* [1988] or *Kikuchi and Fukao* [1985].) For this case, the l_1 misfit is found to be ≈ 0.04, the l_2 misfit is ≈ 0.08 and the l_∞ misfit is $\approx .25$. Thus our fit in both the l_1 and l_2 sense are better than that obtained by

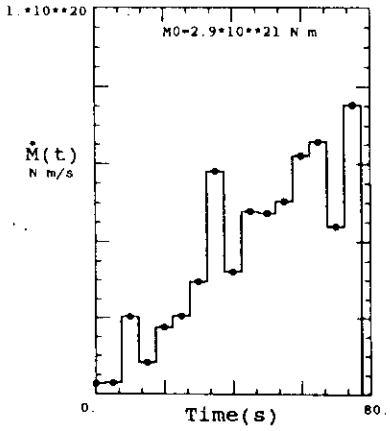


Fig. 4. Discrete source time function for the 1986 Andreanof Islands earthquake with positiveness constraint on slip rate. M_0 denotes the seismic moment obtained for this solution.

Boyd and Nabelek [1988], who are the only ones who have published their results for this earthquake in sufficient detail as yet so as to make such a comparison. Since we use a larger number of parameters in our fit than these authors, it is not at all unexpected that the fit is better. The seismic moment obtained from this solution is 2.9×10^{21} N m, which is larger than the CMT moment of $1. \times 10^{21}$ N m and the centroid depth determined as part of the solution is 22 km, that obtained for the CMT solution being 31 km. This indicates that inspite of the good fit the solution may not be physically meaningful and further examination of this point is carried out next.

Investigation of Stability of Seismic Moment

The minimum value ϵ, say, of (7) estimates the level of noise in the observations and/or errors in the model. Obviously, all the solutions which give misfit close to ϵ are as acceptable as the solution of (1)-(9). Let us take a value

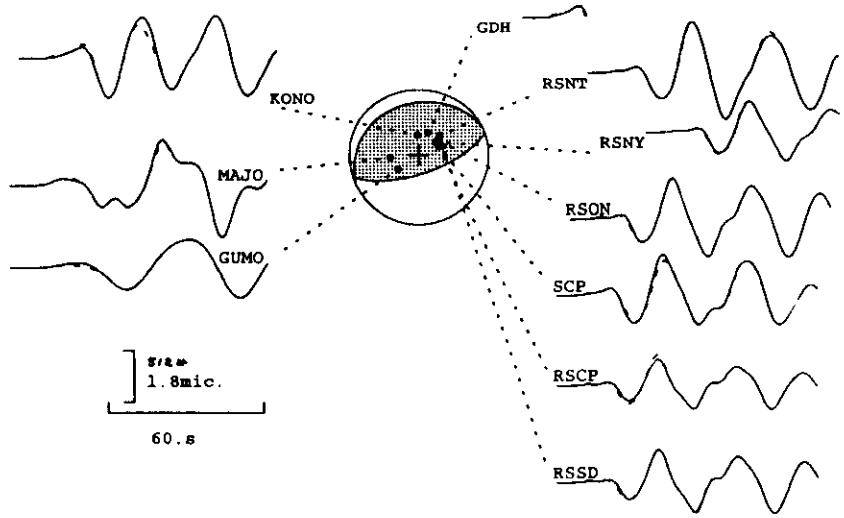


Fig. 5. Comparison of the data (solid lines) with the synthetics (dashed lines) for the solution of Figure 4. All traces are normalized by the maximum value of the data among all stations, plotted on the same scales as in Figure 2.

$\epsilon_1 > \epsilon$ ($\epsilon_1 = 1.5\epsilon$ or 2ϵ, for example). Assuming that the solutions with mean absolute misfit not greater than ϵ_1 are as reliable as the minimum residual solution, it is possible to investigate how different they can be. The inequality $f \le \epsilon_1$ reduces the set of feasible solutions to a slice around the previous optimal vertex. This new feasible set is also a continuous set. Let us select a norm of a feasible solution and find a solution giving, for example, the minimum or maximum of this norm. The seismic moment M_0 of the event is a linear combination of slip rate variables x_i, $i = 1, \ldots, n$ with positive coefficients and can be chosen as such a norm. Denoting these coefficients as c_i, $i = 1, \ldots n$, we have

$$M_0 = \sum_{i=1}^{n} c_i x_i$$

and we arrive at two new linear programming problems, one of which is to

$$\text{minimize } M_0 = \sum_{i=1}^{n} c_i x_i$$

subject to the constraints

$$\sum_{j=1}^{n} a_{ij} x_j - x_{n+i} + x_{n+m+i} = b_i, \quad i = 1, 2, \ldots, m \qquad (10)$$

$$\sum_{i=n+1}^{n+2m} x_i \le \epsilon_1 m \qquad (11)$$

and the positivity constraints (9), and the other problem consists of maximizing M_0 under the same set of constraints. Solving either of these two problems, we could obtain the maximum or minimum values of seismic moment consistent with the data within the given tolerance ($\epsilon_1 - \epsilon$). Since the seismograms contain no information on the very low frequency radiation from the fault, we should expect the difference between the previous solution and either of these two solutions to be large, even for small tolerance of misfit.

Let us consider a tolerance of 100%, that is, $\epsilon_1 = 2\epsilon$ and investigate only the minimum M_0 solution. Figure 6

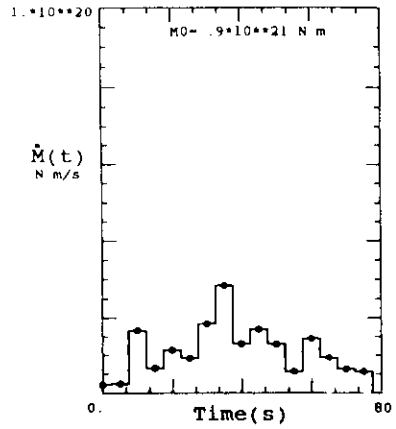


Fig. 6. Same as Figure 4 but for the solution with positiveness constraint on slip rate and moment minimized within 100% tolerance of misfit (that is, l_1 misfit ≈ 0.08).

shows the corresponding source time function and Figure 7 the comparison of the data with the solution. The l_1 misfit is twice as large as before as required by the problem, that is ≈ 0.09 and the l_2 and l_∞ misfits are $\approx .12$ and $\approx .19$, respectively. Imposing additional constraints has deteriorated the l_1 and l_2 fits but the moment and centroid depth are now 0.9×10^{21} N m and 29 km, which are closer to the CMT solution.

As mentioned earlier, the long-period data used in this paper do not constrain the seismic moment well. The CMT estimate for the seismic moment is more reliable for it is obtained using longer period waves and the procedure includes extrapolation to zero frequency. (The ideal value of M_0 would be obtained from geodetic data, provided measurements are made immediately following the event so as not to be contaminated by post-seismic processes.) Consequently, it appears to be important to use some reliable value of seismic moment as an additional constraint on the inversion, and in this paper, we shall take it to be the CMT solution. This would reduce the feasible set, and we may select another norm to investigate the stability of the corresponding optimal solution, as we shall discuss in a later section.

We consider the solution with positiveness constraint and with moment constrained to be equal to the CMT solution as our preferred solution and we shall discuss it in detail. Figure 8 shows the evolution of moment and moment rate distribution on the fault at several time steps, the fault being oriented in the same way as in Figure 1, and Figure 9 shows the source time function. The comparison of the data with the solution is similar to that in Figure 7 and is not shown. Plots of the along arc and along dip moment release are shown in Figure 10. It was found that the optimal solution with proper M_0 required a misfit tolerance of at least 75%, that is, $\epsilon_1 = 1.75\epsilon$. The l_1 misfit is therefore ≈ 0.08, the l_2 and l_∞ misfits are found to be .10 and .21, respectively and the centroid depth is 27 km.

Our solution can be compared with that of *Boyd and Nabelek* [1988], who modelled the source as a bilaterally propagating line source with constant speed. Keeping in mind that our cells are 20 km $\times$20 km in size, we may make a rough estimate of the rupture speed. It is found that the average speed in the 90-km section of the fault to the east of the hypocenter is $\approx$ 3.2 km/s ($0.85v_S$) and is $\approx$ 2 km/s ($0.5v_S$) in the 190-km section lying to the west. The propagation speed in the eastern direction is about twice the average speed of 1.5 km/s obtained by these authors but is closer to their speed in the western direction. Looking at our results in more detail, we find that the rupture speed through the central region of low moment release was greater than v_S and through the region of higher moment release to the west was less than $0.5v_S$, suggesting that these represent mechanically stronger and weaker regions of the fault relative to the state of stress on the fault at the earthquake origin time. Comparison of the along arc moment release shows that both solutions obtain a region of large moment release at the western side of the fault but the region of somewhat lesser moment found near the hypocenter by *Boyd and Nabelek* [1988] is less peaked in our case. Both solutions however seem to have very low moment release near the middle of the fault. Within the region of highest moment release, the maximum slip was found to be 5 m and the maximum slip rate on the fault which also occurred at the same place was 6.7 m/s. The maximum slip in *Boyd and Nabelek's* [1988] solution also occurred in the western region of highest moment release and was 4.5 m. Our solution shows that most of the activity on the fault was confined to the region of high moment release in the west and this region continues to slip from about 25 s after rupture nucleation till the almost complete cessation of all motion on the fault.

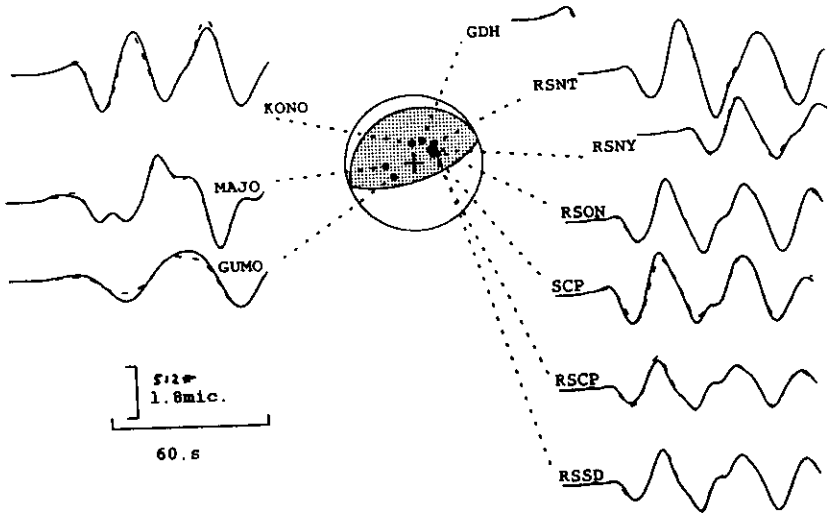


Fig. 7. Same as Figure 5 but for the solution with positiveness constraint on slip rate and moment minimized.

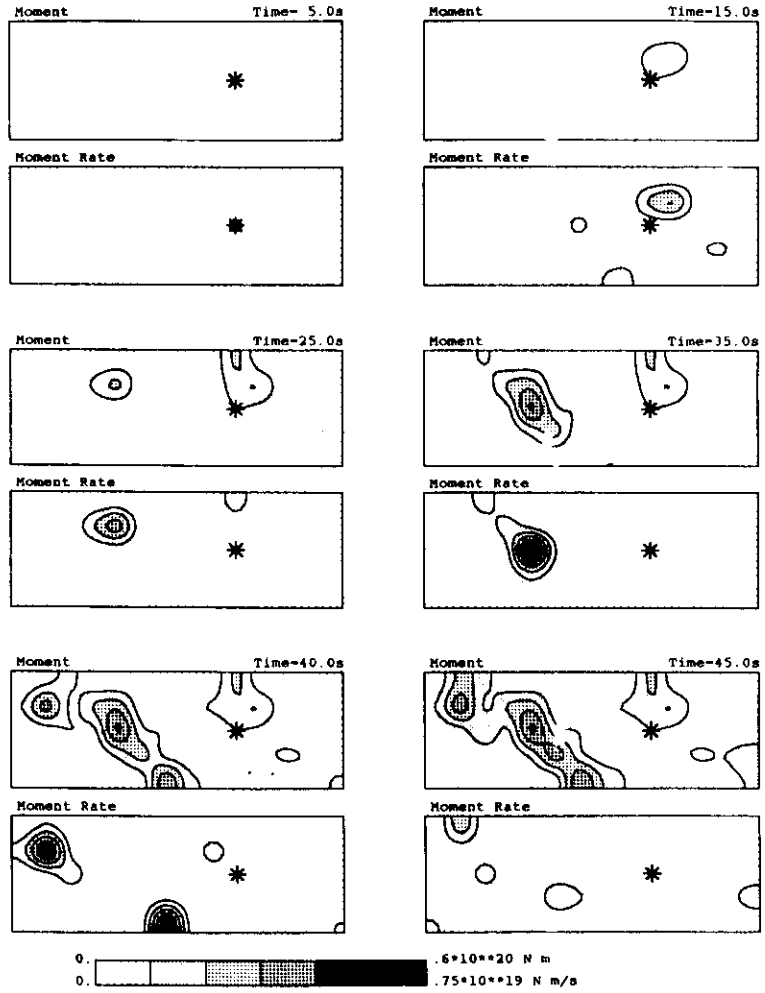


Fig. 8. Evolution in time of moment and moment rate distribution on the fault for the optimal solution with positiveness constraint and with the final moment constrained to the CMT value. The fault view is the same as in Figure 1. The numerical results are smoothed by convolution with a cosine bell with a half width equal to the cell size of 20 km. The center of the cell within rupture nucleated, obtained by *Boyd and Nabelek* [1988], is indicated by a star. The solid arrow indicates the direction of the plate convergence and the broken arrow the direction of the slip vector for this earthquake.

In fact, the plots of moment rate in Figure 8 show that the part of the fault to the east has little activity from 35 s on after rupture initiation. The shape of the slipping region to the west is seen to be clearly developed 40 s after initiation time. Figure 8 also shows some regions which have slipped and stopped slipping being reactivated but in all cases the amount of reslip is only a few percent of the final slip at that place. We do not compare our results in detail with those of *Houston and Engdahl* [1988, 1989] since the moment determined by them is a factor of 2 lower than the CMT moment, and the reliability of the details of the moment release process obtained by them is therefore in question, but we note

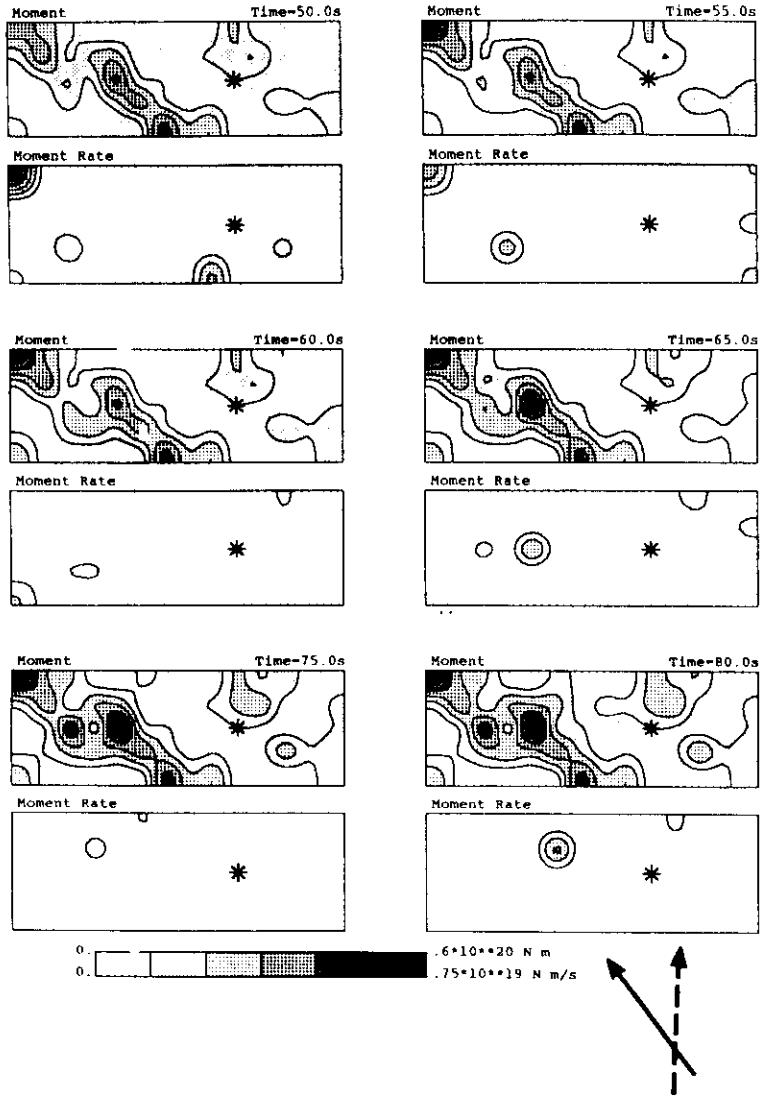


Fig. 8. (continued)

that the regions of highest moment release in our solution is not inconsistent with their rolution. The spatial moment distribution found by *Hwang and Kanamori* [1986] is generally similar to that obtained in this paper. The details of the spatial moment release are discussed next.

The final moment distribution shows a corrugated pattern with two regions of higher moment release on the eastern and western parts of the fault and a region of low moment release in between. The convergence vector obtained by *Minster and Jordan* [1978] for this plate boundary at the epicenter is indicated by the solid arrow and the slip vector for this earthquake is indicated by the broken arrow. The trends of the regions of higher moment along the downdip direction both in the eastern and western parts of the fault and the region of low moment release in between are aligned with the plate convergence direction but not with the direction of the slip vector for this earthquake. The hypocenter of the great earthquake that occurred in this region on March 9, 1957 (M_w = 8.6), falls in the central region of low moment release. It is interesting to note that the region of seismic quiescence observed by *Kisslinger* [1988] before this earthquake overlaps in large part the region of high moment release to the west, but the earthquake did not nucleate in this region. Instead it nucleated 90 km to the east at a region which was not the region of highest moment release on the fault but was adjacent to a local maximum in moment release in the eastern part of the fault. Thus quiescence correctly identified the region of high moment release of the future earthquake but not the zone of rupture initiation.

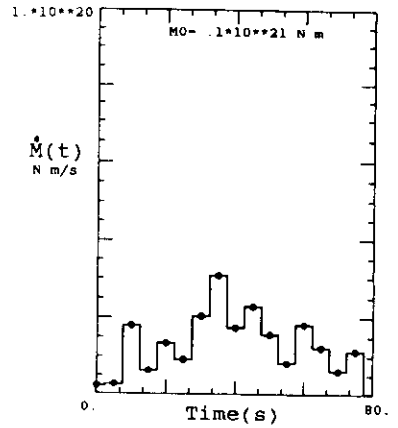


Fig. 9. Discrete source time function corresponding to the optimal solution shown in Figure 8.

The general results of this study can also be compared with those of *Mendoza and Hartzell* [1988, 1989], who studied the 1985 Michoacan earthquake, which is another large subduction zone earthquake which occurred since the installation of worldwide digital stations. For that event, these authors identified three regions of large moment release and the maximum slip of 5.5 m and the rupture speed of about $0.7v_S$ are not dissimilar to those obtained for the Andreanof Islands event.

We do not plot the dynamics of the source process for the case with moment minimized but the solution shows that the evolution of moment and moment rate for this case is very similar to that shown in Figure 8 for moment set to the CMT value. The moments in these two solutions are very close, and the source time functions are very similar. This shows that once the moment is constrained to some chosen value, the solution is quite stable. Further, it implies that our choice of the CMT moment as a constraint instead of some other value of moment does not significantly affect the solution as long as the CMT moment is not incorrect by a factor of 2 or more, say. In fact, even the first solution presented in this paper, namely, the solution with positiveness constraint but without any moment constraint has the feature that there are two main regions of moment in the eastern and western sections of the fault with a region of low moment release in the middle.

The above analysis shows that it is important to include moment as an a priori constraint into the problem to obtain a physically reasonable solution, just as the positiveness constraint (3*a*) or (5*a*) must be imposed on the solution for the same reason. In fact, experiments without including the positiveness constraint which are not reported in detail here, showed that the solution can be very unstable with neighboring cells having large positive and negative values and yet fitting the seismograms very well.

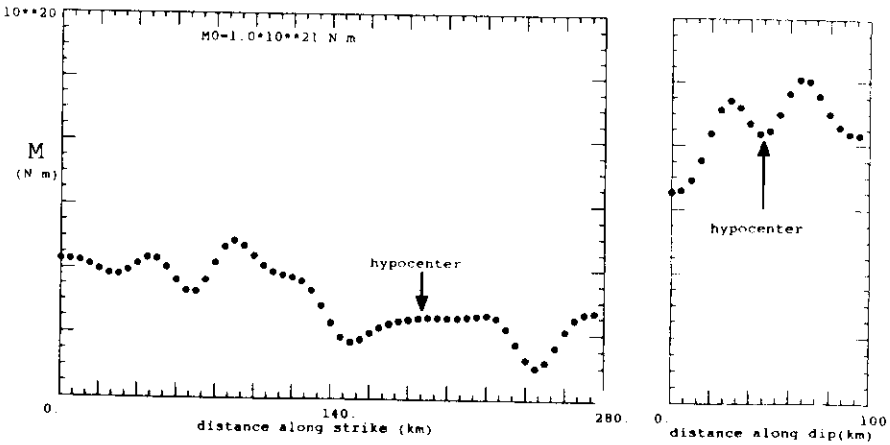


Fig. 10. Plots of along strike and along dip moment release for our preferred solution, with same smoothing as in Fig. 8.

Tectonic Implications of the Solution

Transverse physiographic and seismogenic features of subduction interfaces are usually thought to be due to subduction of seamounts, aseismic ridges and other features of the seafloor [*Kelleher and McCann*, 1976; *Vogt et al.*, 1976; *McCann and Sykes*, 1984]. For the corrugation of slip in the Andreanof Islands earthquake to have resulted from such a mechanism requires that the net interface slip be in the direction of relative plate motion, rather than in the arc normal direction, which is the direction of slip for this earthquake. This property that the slip vector is normal to the arc rather than in the direction of relative plate motion is a common feature of zones of oblique subduction [*Fitch*, 1972; *Jarrard*, 1986]. The generally accepted solution to this conundrum is to hypothesize the existence of a strike-slip zone behind the trench that supplies the transverse component of motion. The puzzle then remains on the origin of the transverse component of slip on the fault. The existence of sufficiently large transcurrent movement on the continental side of a subduction zone to provide the transverse component of motion on the plate interface has been well established for some subduction zones, for example, in southwestern Japan and in western Sunda [*Fitch*, 1972]. A similar hypothesis has been suggested for the Adak Islands region of Alaska [*Ekstrom and Engdahl*, 1989]. However, though there are strike-slip earthquakes in the back arc region in this area, these may not be large enough (magnitudes are less than about 6.5) to provide the necessary transverse component of slip. An alternative hypothesis may be that the transverse component of slip is supplied by aseismic slip on the interface.

DISCUSSION

One may next investigate how different the distribution of slip rate in space and time can be. For example, one may minimize the maximum value of the slip rate producing a slip rate distribution which is most uniformly spread over the fault area and over the duration of the event. Conversely, one could maximize the maximum value obtaining the solution that is most concentrated in space and time. These two solutions would give an indication of the possibilities of imaging the fault in terms of the asperity model. The minimization problem is easily formulated as a linear programming problem as follows. Introduce a new variable x_0, say, denoting an upper bound of slip rate value. Then the problem is to minimize x_0 subject to (10) and (11), the additional constraint for seismic moment:

$$\sum_{i=1}^{n} c_i x_i = M_0 \tag{12}$$

the constraints $x_i \le x_0, i = 1,...n$, and the positivity constraints (9). Maximizing the maximum value of the slip rate is not easily formulated as a linear programming problem. Actually one would have to maximize each one of the x_i's, $i = 1,...n$ separately (which is a linear programming problem) and then take the solution corresponding to the greatest of these maxima. Alternately, one may limit oneself only to maximizing that variable which has the maximum value in the previous optimal solution of the problem described by (10) and (11), together with the positiveness constraint (9) and the fixed moment constraint (12).

Continuing in this way, one may obtain independent constraints on the maximum slip rate value and use it as a constraint and re-solve the problem. For a sufficiently large spatial fault grid size, a reasonable estimate of the upper bound of the slip rate may be obtained as the slip rate value at the centre of a (finite) crack with instantaneous stress drop, $\Delta\sigma$, as

$$\dot{u} \le \frac{\Delta\sigma}{\rho v_s} \tag{13}$$

where $\dot{u}$ is the slip rate, ρ is the density and v_s is the shear wave speed. For $\Delta\sigma \sim 100$ bars, $\rho \sim 2.5$ g/cm^3 and $v_s \sim 3$ km/s, this bound is ~100 cm/s. An alternate estimate may be obtained by averaging the stress and slip rate near the edge of a shear crack in an ideally brittle medium, propagating at a velocity v (using equation (15.6) and (15.7) of *Aki and Richards* [1980]) as

$$\text{for Mode III cracks} \quad \dot{u} \le \frac{\Delta\sigma}{\rho v_s} \frac{v}{\sqrt{v_s^2 - v^2}} \tag{14}$$

$$\text{for Mode II cracks} \quad \dot{u} \le \frac{\Delta\sigma}{\rho v_s} \frac{v}{R(v)} \tag{15}$$

where $R(v)$ is the Rayleigh function. Interpreting $\Delta\sigma$ in the above formulae as the strength of rock or fault gouge, one can obtain the required constraint on the slip rate. The presence of the velocity factor in (14) and (15) makes it less useful than (13). Since in our discretization of the fault, it was assumed that slip can occurr simultaneously over the entire cell, the constraint that would be applicable in our case is (13). If, however, the slip is allowed to propagate through the cell at speed v, then (14) or (15) should be the constraint relevant to that particular case. The slip rate constraint is important from the point of view of the strong ground motion occurring due to an earthquake, since the faster the fault slips the greater is the ground acceleration and the intensity of the earthquake. In fact, the relation between strong motion acceleration and fault slip rate can be used to constrain the slip rate on a fault.

Next, if some independent estimate of the final slip on the fault is available, then the solution could be made consistent with such a constraint, and so on. These types of constraints can be easily incorporated into the algorithm but are not carried out in this paper. The above discussions show how physically meaningful constraints can be found and incorporated into the solution of the inverse problem of the earthquake source process one by one, as desired.

CONCLUSIONS

In summary, we suggest that the well known instability in the inverse problem of earthquake source can be most reasonably dealt with by describing the whole set of equally acceptable solutions and by producing some extreme representatives of this set, for example. The linear programming technique is used in this paper to solve this problem, and includes additional constraints on the solution. These constraints simply fill the void created by the lack of sufficient data. The advantage of such constraints is that they are stated explicitly. Obviously, instability is an inherent property of the problem and cannot be removed, but whatever is supplied in place of the data can be made explicit. Stated in other words, the known bias of the seismologist is preferable to the unknown bias of some implicit constraint!

The feature of our solution that the moment release is concentrated in two regions on the fault separated by a region of little or no moment release seems to be quite stable with respect to the choice of constraints. In fact, experiments with twice as large grids, which are not reported here, also showed this feature, but the details of the downdip linear features with trend similar to the plate convergence direction were, however, not as clear in the coarser grid solution.

An advantage of the linear programming approach is that additional constraints can be very easily incorporated into the solution at any stage of the problem. Further, since the inclusion of additional constraints reduces the number of vertices in the hyperspace where the solution lies, the solution at each stage is generally arrived at more rapidly by the simplex method. This is an advantage of the direct method over alternative iterative methods.

The method developed here can be refined in many obvious ways. The synthetics used in this paper were very simple and better crustal responses, particularly at the source region, which is characterized by a downgoing slab, should be obtained. Only P waves were used here but the inclusion of later phases into the synthetics would allow longer lengths of records to be analyzed. Finer gridding of the fault and finer sampling of the source duration together with using broadband data are also obvious improvements.

The results of the study of the 1986 Andreanof Islands earthquake undertaken in this paper imply that the long term plate motion on this portion of the Aleutian arc is indeed in the direction of the plate convergence.

Acknowledgments. This work was performed under area IX of the U.S.-USSR agreement in the Field of Environmental Protection and was supported by the National Science Foundation under grants RII-85-03099 and EAR-87-20630 and by USGS grant 14-08-0001-G1667. We thank Klaus Jacob, Art Lerner-Lam, and Adam Dziewonski for reviewing the manuscript and Goran Ekstrom and Tom Boyd for helpful discussions. In addition, Tom Boyd provided us a copy of his figure which was modified for Figure 1 of our paper. We are also very grateful to D.G.M. Anderson of the Division of Applied Sciences, Harvard University, for helpful discussions and suggestions and for a careful review of the paper. Finally, John Nabelek reviewed the paper for JGR and provided an in-depth review. We thank him for several discussions and an anonymous reviewer, both of whom helped to improve our manuscript. The numerical calculations were performed on the Cray-YMP at NSF's Pittsburgh Supercomputer Center and the figures were produced on the seismology computer at the Earth and Planetary Sciences Department at Harvard University, and we thank them for the use of their facilities. Lamont-Doherty Geological Observatory contribution 4598.

REFERENCES

Aki, K., and P. G. Richards, *Quantitative Seismology: Theory and Methods*, 932 pp., W.H. Freeman, New York, 1980.

Berosa, G. C., and P. Spudich, Linearized inversion of fault rupture behavior: Application to the 1984 Morgan Hill, California earthquake, *J. Geophys. Res., 93*, 6275-6296, 1988.

Boyd, T. M., and J. L. Nabelek, Rupture process of the Andreanof Islands earthquake of May 7, 1986, *Bull. Seismol. Soc. Am., 78*, 1653-1673, 1988.

Das, S., Three-dimensional rupture propagation and implications for the earthquake source mechanism, *Geophys. J. R. Astron. Soc., 67*, 375-393, 1981.

Das, S., and K. Aki, Fault plane with barriers: A versatile earthquake model, *J. Geophys. Res., 82*, 5658-5670, 1977.

Dziewonski, A. M., and D. L. Anderson, Preliminary reference earth model, *Phys. Earth Planet. Inter., 25*, 297-356, 1981.

Dziewonski, A. M., and J. H. Woodhouse, An experiment in systematic study of global seismicity: Centroid-moment tensor solutions for 201 moderate and large earthquakes of 1981, *J. Geophys. Res., 88*, 3247-3271, 1983.

Dziewonski, A. M., G. Ekstrom, J. E. Franzen, and J. H. Woodhouse, Centroid-moment tensor solutions for April-June 1986, *Phys. Earth Planet. Inter., 45*, 229-239, 1987.

Ekstrom, G. A., A broad band method of earthquake analysis, Ph.D. thesis, 216 pp., Harvard Univ., Cambridge, Mass., 1987.

Ekstrom, G. A., A very broad band inversion method for the recovery of earthquake source parameters, *Tectonophysics, 166*, 73-100, 1989.

Ekstrom, G. A., and E. R. Engdahl, Earthquake source parameters and stress distribution in the Adak Island region of the central Aleutian Islands, Alaska, *J. Geophys. Res., 94*, 15,499-15,519, 1989.

Engdahl, E. R., and S. Billington, Focal depth determination of central Aleutian earthquakes, *Bull. Seismol. Soc. Am., 76*, 77-93, 1986.

Engdahl, E. R., S. Billington, and C. Kisslinger, Teleseismically recorded seismicity before and after the May 7, 1986 Andreanof Islands, Alaska, earthquake, *J. Geophys. Res., 94*, 15,481-15,493, 1989.

Fitch, T. J., Plate convergence, transcurrent faults, and internal deformation adjacent to Southeast Asia and the western Pacific, *J. Geophys. Res., 77*, 4432-4460, 1972.

Franklin, J., *Methods of Mathematical Economics. Linear and Nonlinear Programming, Fixed Point Theorems*, 297 pp., Springer-Verlag, New York, 1980.

Hartzell, S. H., and T. H. Heaton, Inversion of strong-ground motion and teleseismic waveform data for the fault rupture history of the 1979 Imperial Valley, California earthquake, *Bull. Seismol. Soc. Am., 73*, 1553-1583, 1983.

Houston, H., and E. R. Engdahl, A comparison of the spatial distribution of moment release with relocated seismicity for the 1986 Andreanof Islands earthquake, (abstract), *EOS Trans. AGU, 69*, 1303, 1988.

Houston, H., and E. R. Engdahl, A comparison of the spatio-temporal distribution of moment release for the 1986 Andreanof Islands earthquake with relocated seismicity, *Geophys. Res. Lett., 16*, 1421-1424, 1989.

Hwang, L. J., and H. Kanamori, Source parameters of the May 7, 1986 Andreanof Islands earthquake, *Geophys. Res. Lett., 13*, 1426-1429, 1986.

Hwang, P. Y., Focal depths and mechanisms of mid-ocean ridge earthquakes from body waveform inversion, Ph.D. thesis, 301 pp., Mass. Inst. of Technol., Cambridge, Mass., 1985.

Jarrard, R. D., Relations among subduction parameters, *Rev. Geophys., 24*, 217-284, 1986.

Kelleher, J. and W. McCann, Bouyant zones, great earthquakes, and unstable boundaries of subduction, *J. Geophys. Res., 81*, 4885-4896, 1976.

Kikuchi, M., and Y. Fukao, Iterative deconvolution of complex body waves from great earthquakes - The Tokachi-Oki earthquake of 1968, *Phys. Earth Planet. Inter., 37*, 235-248, 1985.

Kikuchi, M., and H. Kanamori, Inversion of the complex body waves, *Bull. Seismol. Soc. Am., 72*, 491-506, 1982.

Kisslinger, C., An experiment in earthquake prediction and the May 7, 1986 Andreanof Islands earthquake, *Bull. Seismol. Soc. Am., 78*, 218-229, 1988.

Kostrov, B. V., and S. Das, *Principles of Earthquake Source Mechanics, Appl. Math. Mech. Ser.*, 286 pp., Cambridge University Press, New York, 1988.

McCann, W.R. and L.R. Sykes, Subduction of aseismic ridges beneath the Caribbean plate: Implications for the tectonics and seismic potential of the northeastern Caribbean, *J. Geophys. Res., 89*, 4493-4519, 1984.

Mendoza, C., and S. H. Hartzell, Inversion for slip distribution

using teleseismic P waveforms: North Palm Springs, Borah Peak, and Michoacan earthquakes, *Bull. Seismol. Soc. Am.*, *78*, 1092-1111, 1988.

Mendoza, C., and S. H. Hartzell, Slip distribution of the 19 September 1985, Michoacan, Mexico, earthquake: Near-source and teleseismic constraints, *Bull. Seismol. Soc. Am.*, *79*, 655-669, 1989.

Mikumo, T., and T. Miyatake, Earthquake sequences on a frictional fault model with non-uniform strengths and relaxation times, *Geophys. J. R. Astron. Soc.*, *59*, 497-522, 1979.

Minster, J. B., and T. H. Jordan, Present-day plate motions, *J. Geophys. Res.*, *83*, 5331-5354, 1978.

Nabelek, J., Determination of earthquake source parameters from inversion of body waves, Ph.D. thesis, 360 pp., Mass. Inst. of Technol., Cambridge, 1984.

Noble, B., and J. W. Daniel, *Applied Linear Algebra*, 477 pp., Prentice Hall, Englewood Cliffs, N. J. 1977.

Olson, A. H., and R. J. Apsel, Finite faults and inverse theory with applications to the 1979 Imperial Valley earthquake, *Bull. Seismol. Soc. Am.*, *72*, 1969-2001, 1982.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical recipes: The Art of Scientific Computing*, 818 pp., Cambridge University Press, New York, 1986.

Ruff, L. J., Fault asperities inferred from seismic body waves, in *Earthquakes: Observation, Theory and Interpretation*, edited by H. Kanamori and E. Boschi, pp. 251-276, North-Holland, Amsterdam, 1983.

Sykes, L. R., J. B. Kisslinger, L. House and K. H. Jacob, Rupture zones and repeat times of great earthquakes along the Alaska-Aleutian arc 1784-1980, *Earthquake Prediction*, AGU Monograph, Maurice Ewing Series, *4*, 73-80, 1981.

Tarantola, A., *Inverse Problem Theory. Methods for Data Fitting and Model Parameter Estimation*, 613 pp., Elsevier, New York, 1987.

Vogt, P.R., A. Lowrie, D.R. Bracey, and R.N. Hey, Subduction of aseismic oceanic ridges: Effect on shape, seismicity, and other characteristics of consuming plate boundaries, *Geol. Soc. Am., Spec. Pap. 172*, 59 pp., 1976.

Ward, S. N., and S. E. Barrientos, An inversion for slip distribution and fault shape from geodetic observations of the 1983 Borah Peak, Idaho earthquake, *J. Geophys. Res.*, *91*, 4909-4919, 1986.

S. Das, Department of Earth Sciences, University of Oxford, Oxford OX1 3PR, UK

B. V. Kostrov, Institute of Physics of the Earth, Bolshaya Gruzinskaya 10, Moscow D232, USSR.

(Received February 23, 1989;
revised January 8, 1990;
accepted March 1, 1990.)