UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY
Office of Earthquake Studies

PROCEEDINGS OF
CONFERENCE III
FAULT MECHANICS AND ITS
RELATION TO EARTHQUAKE PREDICTION

Convened Under Auspices of

NATIONAL EARTHQUAKE HAZARDS REDUCTION PROGRAM

1-3 December, 1977

OPEN-FILE REPORT 78-380

Menlo Park, California

1978

Conferences to Date

Conference I	Abnormal Animal Behavior Prior to Earthquakes, I
Conference II	Experimental Studies of Rock Friction with Application to Earthquake Prediction
Conference III	Fault Mechanics and Its Relation to Earthquake Prediction

Conferences in Process

Conference IV	Use of Volunteers in the Earthquake Hazards Reduction Program
Conference V	Methodology for Identifying Seismic Gaps and Soon-To-Break Seismic Gaps
Conference VI	Measurements of Stress and Strain Pertinent to Earthquake Prediction

UNITED STATES
DEPARTMENT OF THE INTERIOR
GEOLOGICAL SURVEY

Office of Earthquake Studies

PROCEEDINGS OF
CONFERENCE III
FAULT MECHANICS AND ITS
RELATION TO EARTHQUAKE PREDICTION

Convened Under Auspices of

NATIONAL EARTHQUAKE HAZARDS REDUCTION PROGRAM

1-3 December, 1977

Co-Organizers

William D. Stuart
United States Geological Survey
Office of Earthquake Studies
Menlo Park, California 94025

Amos Nur
Department of Geophysics
Stanford University
Stanford, California 94305

Convener

Jack F. Evernden
United States Geological Survey
Office of Earthquake Studies
Menlo Park, California 94025

OPEN-FILE REPORT 78-380

Menlo Park, California

1978

CONTENTS

Introduction

On 1 - 3 December 1977, a Conference on "Fault Mechanics and Its Relation to Earthquake Prediction" was held at Stanford University, Stanford, California. This Conference was the third in a continuing series of Conferences to be held under the auspices of the Earthquake Hazards Reduction Program.

These Conferences and publication of their proceedings are a vital part of the EHRP, as they are intended to be state-of-knowledge Conferences and documents. They are intended to provide intellectual stimulus both to participants and to recipients of the proceedings.

Participants in the EHRP are encouraged to suggest topics for future Conferences. The subject should be of comparatively narrow focus, should be based on active participation of all attendees, all of whom must be conducting research appropriate to the Conference theme. Total attendance is restricted to 25 or 30.

SUMMARY

by

Amos Nur and William Stuart

Discussion of theoretical mechanical models for deformation associated with faulting fell into five overlapping categories: (1) large scale quasi-static fault models, (2) fault instability (earthquakes), (3) dynamic rupture propagation, (4) use of seismic radiation to infer source parameters, and (5) heterogeneity of fault constitutive properties. Each category was distinguished not only by the topic discussed, but also by the mathematical and conceptual approaches commonly used. In each category recent work was described and discussed and main outstanding problems identified; in most the current state of knowledge was reviewed. We summarize here the scientific results receiving general acceptance, and identify promising areas of future work emerging from the conference.

Current Results

Consititutive (stress-deformation) properties, especially those of fault zones seemed to be the dominant recurrent subject. In particular, heterogeneity or spacial variation of constutive behavior was considered to be a major factor to be considered in realistic models of both quasi-static and dynamic faulting.

In large scale quasi-static plate models the lithosphere is adequately described by elasticity, but the deeper asthenosphere is better approximated

by an inelastic law such as viscous, visco-elastic, or plastic. Geodetic observations are too scarce and imprecise to distinguish the finer points of models at present, however.

Earthquake precursor and instabilty models require either a post-peak strain (slip) or strain rate (velocity) weakening fault zone to produce instability. Models also postulate that the fault zone properties vary with position. In models growing pre-instability deformtion rates arise from two sources: (1) the above mentioned fault weakening on a time scale determined by remote forcing, (2) temporary pore fluid stabilization due to Biot type or dilatant hardening on a diffusion controlled time scale.

Strong evidence from diverse observations indicates that faults are often very heterogeneous. This is reflected in the frequency - magnitude relation, random ground motion, and strongly non uniform fault offsets often observed following sizeable earthquakes. These heterogeneities are responsible for uncertainties in using seismic moment for stress estimates. Several analyses of the seismic radiation field indicate that the stress state along faults may be highly irregular, both in space and time. Furthermore, earthquake rupture often appears to be a complex succession of smaller ruptures leading to multiple events. Inspection of seismically produced fractures in mines and at the surface also suggests faulting complexity.

Consequently, one problem in fault modelling is to understand how faults remain heterogeneous with time. It appears that spatially nonuniform fracture energy may account for the maintenance of heterogeneity, as well as irregular rupture velocity and irregular slip.

Laboratory measurements of frictional slippage show increasing rates prior to stick-slip events. Transiently higher rates also occur without stick-slip. Several theoretical models show similar phenomena. Such

accelerated fault slip before earthquakes may cause observable precursory deformation fields. A key problem which remains unresolved, however, is the conditions for rupture initiation. Although the point of initiation appears central for earthquake prediction, it is unclear what the initiation process actually is. Much work is needed, particularly in observing seismicity patterns before the main event in and around the epicentral region.

The state of crustal shear stress is unclear: Some seismic and crustal deformation data as well as laboratory results suggest high stress, whereas the lack of heat flow anomalies imply low stress.

Future Work

It now seems possible to formulate and solve analytically or numerically a wide variety of boundary value problem representations of quasi-static and dynamic faulting. The principal uncertainties are the form and spacial variation of fault constitutive laws and physical properties. Non-linear rate dependent and independent weakening as well as healing laws need elaboration. Relevant laboratory data, either from homogeneously deformed samples or from scaled simulations is scarce, and the connection between laboratory and natural condition remains unclear. To a lesser extent, geometric and boundary conditions are poorly known. Field observations such as surface deformation, pore pressure, and deviation stress state are too sparce to test theoretical models in detail.

One prodedure is to construct a quasi-static numerical (e.g. finite element) model of a portion of an active fault like the San Andreas. Geodetic, seismic, and other data might then be inverted to estimate the stress and strain near the fault. Though resolution is likely to be poor, one might be able to infer the fault consitutive law form and coefficient values, as

well as eleatic and friction heterogeneities nearby.

Perhpas the clearest conclusion is that heterogeneities are important, and perhaps essential feature of the earthquake mechanism, and must be included in future modelling work.

The determination of the state of crustal shear stress - at present unclear - must be accomplished in order to understand the process of faulting.

In conclusion, it is clear that earthquake studies are shifting from geometrical or kinematic source modesl to mechanical and dynamic models. The nature of the faulting instability, the fundamental link between rupture velocity and rise time, slip and stress will receive much more attention than in the past, leading eventually to a source model of self generating and propagating rupture which can also stop on its own.

A QUANTITATIVE MODEL OF STRESS IN A SEISMIC REGION AS A BASIS FOR EARTHQUAKE PREDICTION

by

Keiiti Aki

Department of Earth and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

This paper outlines a framework for combining available data from plate tectonics, geodetic, seismic, and laboratory measurements for the purpose of estimating stress in a seismic region. Applying a new earthquake source model containing a fault plane with unbroken barriers to major earthquakes, we find the importance of distinguishing between tectonic stress and self stress, as pointed out by Andrews [1978]. The self stress is due to irregular slip function along fault plane and may show great variation in the fault zone after a major earthquake. The stress may drop to a low value on the slipped segment of fault, but may be elevated considerably near the unbroken barriers. Because of the resistance by unbroken barriers, the stress drop for the whole fault zone as seen from outside by geodetic methods is only a small fraction of the initial stress. This implies an ambient tectonic stress much higher than the range from 10 to 100 bars and adds support to the argument of Hanks [1977] of the importance of traction at the base of a plate as a cause of tectonic stress. Methods for determining the basal traction by downward continuation of stress from observations on the surface as well as a preliminary quantification of self stress in the fault zone are discussed

Introduction

Since the stress in the earth is believed to be the cause of an earthquake, it is desirable to know its space-time distribution accurately and in detail for the purpose of earthquake prediction.

There are three distinct approaches to estimating stress in a seismic region. One is to use the seismic waves radiated from an earthquake. This gives the stress drop which occurred along the fault plane during the earthquake. Another approach is to measure the deformation of the earth's surface. This gives the secular change in strain, which may be related to stress change. The third method uses artificial perturbations of the stress field in the earth by hydraulic fracturing or some other mechanical operation and, in principle, gives the absolute value of existing stress.

Various observations have been used in estimating the stress change associated with earthquakes. Large-scale geodetic measurements, such as the leveling and triangulation surveys in the epicentral areas of major shallow earthquakes, usually indicate the strain change in the immediate vicinity of the epicenters in the range 2×10^{-5} to 2×10^{-4}, as summarized by Rikitake [1976]. Even if these strains are entirely elastic, the above range of strain corresponds to the range of stress drop from 6 to 60 bars, assuming a crustal rigidity of 3×10^{11} dyne cm^{-2}.

Similar low values of stress drop have been obtained from seismic observations for major earthquakes. If the

seismic moment (measured from long period waves) and the fault area (estimated from aftershock area, Tsunami source area, geodetic data, seismic wave-form, and spectrum, etc.) are combined and interpreted in terms of a simple crack model, the result usually gives the stress drop in the range from 10 to 100 bars as summarized by Aki [1972] and Kanamori and Anderson [1975].

Revaluation of stress drop on the basis of the barrier model

Recently, Das and Aki [1977] made numerical experiments on the rupture propagation over a fault plane with distributed barriers, to explain a variety of seismic observations on the complexity of earthquake source such as the so-called multiple events.

Aki et al. [1978] further proposed that major earthquakes may be modeled by a rupture over a fault plane with distributed barriers, some of which will remain unbroken after the earthquake. One important consequence of this new model is that if this model applies to major earthquakes, the stress drop estimated by assuming a simple crack (without barriers) may be an order of magnitude too small.

The new model is supported by some geologists' observations on the slip distribution along the fault. Wallace [1968] discovered in the Carrizo Plain in central California streams with no apparent offset lying along fault segments on which 10 meters of offset had occurred during the great earthquake of 1857. For the maximum slip of 10 meters over

a fault segment of estimated length 8 km, the corresponding stress drop is nearly 1 kilobar.

Such a segmentation of a fault appears to be rule rather than an exception. The maps published by the U.S. Geological Survey of recently active breaks along the San Andreas and related faults show that the pattern of surface faulting is not a single continuous fracture but consists of multiple breaks. The longest individual segment is about 10 to 18 km. Earthquakes with greater fault lengths must then be multiple fractures.

The segmentation of faults is not restricted to the San Andreas fault. Tchalenko and Berberian [1975] studied the fault slip of the Dasht-e Bayaz earthquake of 1968 and found clear evidence for a segmented fault. Similar evidence is found by Matsuda [1972, 1974, 1976] and Matsuda and Yamashina [1974] from detailed studies of the Nobi (1891), Tango (1927), Izu-Hanto-Oki (1974), and Kita-Izu (1930) earthquakes. It appears that the fault slip distribution is always irregular whenever a detailed observation is made. This irregularity may be due to complication by soft surface material, but some features can only be explained by the real segmentation of the fault at depth. One such example is the Parkfield earthquake of 1966. Not only the fault trace jumps from the northeast to the southwest rim of the fault-zone valley, but also the initial linear trend of the aftershock zone is terminated at the point of fault trace jump, and a new segment is started along a line about 1 km apart from the initial

trend [Eaton et al., 1970]. Since these aftershocks are located down to 15 km deep and believed to define the fault plane of the San Andreas fault, the segmentation of fault cannot be due to secondary effects of the surficial layer.

High stress drop has been reported for some major earthquakes. For example, the stress drop in the area of nucleation of rupture in the San Fernando earthquake is estimated as high as 500 bars [Hanks, 1974; Bouchon, 1978].

Bouchon [1978] found a kinematic similarity between the uniform dislocation model widely used in earthquake seismology and the crack model over a fault plane with distributed unbroken barriers. The two models are roughly equivalent in seismic radiation, if the rise time of the uniform dislocation model is equated to the barrier interval divided by the rupture velocity. Using this approximate equivalence relation, one can obtain the barrier interval for earthquakes for which the rise time and rupture velocity is known.

Figure 1 shows the summary of relations between the barrier interval and maximum slip determined by the geological and seismic methods (reproduced from Aki et al., 1978). The stress drop associated with an individual fault segment can be estimated from the barrier interval and the maximum slip. They lie in the range between 100 and 1 kilo bars, an order of magnitude higher than the range obtained by the use of a simple crack model.

Large-scale tectonic stress versus stress in the fault zone

From the above discussion, we must conclude that the local stress drop over a fault segment can be an order of magnitude higher than the large-scale stress release estimated from geodetic data. Accompanying this large stress drop in the slipped segment of the fault, the stress near the unbroken barrier will be increased after the earthquake. Thus, the stress will vary strongly along the fault plane after the earthquake.

At this point, it is convenient to separate the stress in the fault zone into two terms, tectonic stress and self stress, following Andrews [1978]. The tectonic stress is due to distant causes and should be smoothly varying along the fault plane. The self stress is due to irregular slip function and varies strongly along the fault plane.

Andrews [1978] pointed out, from a consideration of energetics, that the stationary occurrence of a large number of small earthquakes cannot be explained by the load of smoothly varying tectonic stress alone, but requires a generation of short wavelength self stress by a large earthquake, unless fault creep, varying in amplitude at all length scales prepares the fault for small earthquakes. Our barrier model offers a physical mechanism for such a roughening of self stress in the fault zone after a major earthquake.

The strong spatial variation of self stress after a major earthquake becomes smoother with time through the occurrence of aftershocks, creep, and other stress relaxation mechanisms. On the other hand, the tectonic stress in the

fault zone drops only by a small amount. This drop is recovered slowly by the time of the next major earthquake. By that time, the self stress in the fault zone is probably diminished by healing and stress relaxation. Within the framework of our barrier model, the initial stress before the major earthquake must be at least as high as the stress drop in the fault zone. Then, it follows that the absolute value of tectonic stress must be an order of magnitude greater than its drop during an earthquake (observed as a large scale stress change by geodetic means). The stress may drop to a low value on the slipped segment of the fault, but it will be elevated considerably near the unbroken barriers. Because of the resistance by unbroken barriers, the stress drop for the whole fault zone as seen from outside by geodetic methods is only a small fraction of the initial stress.

Absolute value of tectonic stress

Brune et al. [1969] argued for a low ambient tectonic stress along the San Andreas fault on the basis of lack of local heat flow anomaly near the fault. They obtained the upper limit of 200∿250 bars if the fault slip occurs primarily by earthquakes. If our barrier model is applicable to the San Andreas fault, the stress drop in the fault zone may amount to 1 kilobar as discussed above. However, since the lack of heat flow anomaly only contrains the dynamic friction on the fault to about 100 bars over a fault depth of 20 km, the initial stress can be as high as 1.1 kilobar, for a

kilobar stress drop, without contradicting the heat flow observations.

Evidence of a very high ($\sim$1 kilobar) tectonic stress has been presented by Hanks [1971] for the plate near the Kuril trench from the high apparent stress drop in shallow earthquakes and the bulge of the ocean floor seaward of the trench axis which was interpreted as due to flexure of elastic plates [Walcott, 1970]. Watts and Talwani [1974] obtained a similar result for other trenches in the Pacific using gravity data in addition to topography.

Another evidence for high stress comes from rock mechanics laboratories. It is now well established that a single simple law exists relating frictional strength to normal stress irrespective of mineralogy, pressure, temperature to about 600°C, and a wide range of displacement rates. Unless pore pressure is unusually high in the fault zone, the frictional strength must be higher than 1 kilobar at the depth of foci even for the shallow earthquakes of a transform fault.

Hanks [1977] argues that if the frictional strength of transform faults is higher than a kilobar, the boundary forces such as gravitational pull by descending slabs or the push from mid-ocean ridges cannot drive the plate motions. Traction at the base of a plate exerted by convection current in the asthenosphere is needed to overcome the resistance at the transform fault. Our results from seismic studies of earthquake faulting processes give additional support to

the high ambient stress. It supports Hanks' argument which opposes the idea advocated by Forsyth and Uyeda [1975] and Solomon et al. [1975] that the boundary forces are the main driving forces and the asthenosphere is only passively exerting weak drag on the plates. It also suggests the importance of traction at the base of the plates as a source of tectonic stress.

Downward continuation of tectonic stress

If the traction at the base of a plate is an important cause of tectonic stress, it is worthwhile to attempt a downward continuation of stress from observations at the surface in a manner similar to the density determination from gravity observed on the surface.

Our problem is not well posed. The uniqueness theorem for elastic fields states that if the body force inside an elastic body V and the displacement or traction is known on the surface enclosing V, then the displacement inside the body is uniquely determined. Our problem is ill-posed, because we know the displacement and traction on the free surface but neither the displacement nor the traction is known at the base or side of the volume of plate under consideration. Furthermore, non-elastic strain or stress-free strain [Eshelby, 1957; Backus and Mulcahy, 1976] such as dilatancy, seismic and aseismic slip, etc. will contribute in general as unknown body forces distributed in the body. These difficulties, however, may not prevent a geophysicist from working on the

problem, because though many geophysical problems are of this nature, a solution may be found under some reasonable assumptions, giving better insight to the phenomena than the mere accumulation of data.

Our problem can be approached by several different methods. If we can assume a vertically heterogeneous earth model, then the Fourier transformation with respect to the two horizontal coordinates (x,y) and the propagator method, such as used by Brown [1975], can be applied to find a solution in z with the boundary condition that the displacement is known and traction vanishes at z = 0. This approach was attempted by Aki (1953, unpublished) for the three components of displacement data for the Tokyo earthquake of 1923 using a half-space model. The result was unpublished because of the undue amplification of short-wave length with increasing depth. Recent advances in inversion techniques, such as the stochastic inversion [Franklin, 1970], can be used to obtain an optimal solution by taking into account the signal-to-noise ratio at each wave length.

The finite element method would be more versatile. Dividing the plate into appropriate elements, we shall distinguish elements in which non-elastic processes such as seismic slip, aseismic slip, or dilatancy is occurring from purely elastic elements. Non-elastic processes will be modeled by assigning the stress-free strain to each non-elastic element.

Putting the nodal displacement vector U, the nodal load

vector R, and the stiffness matrix K, the usual finite-element formulation can be written as

$$KU = R \tag{1}$$

where R contains the contributions from the stress-free strain (or initial strain in the terminology of finite-element method) and elastic strain.

In our problem, U is measurable for the nodes on the free surface by geodetic survey. Since traction is free on the free surface, R is partially known there. The part of R due to stress-free strain may be known from independent measurements. If not, they must be considered as unknown. Plate tectonics may constrain the value of U for the nodes on the outside boundary. On the other hand, the upper bound of absolute value of R at the nodes on the bottom surface may be constrained from laboratory measurements on stress relaxation mechanisms for rocks under appropriate pressure and temperature.

Then, equation (1) will present a linear inverse problem for unknown parts of U and R. The equation is most likely singular, and the solution will be non-unique. We can, however, determine the range of possible solutions. The method of linear programming would be useful for finding the range of possible solutions. If the result is meaninglessly arbitrary, then it would at least suggest what new observations can be made in order to reduce the non-uniqueness.

Fault zone stress

The self stress in the fault zone can be calculated if the slip function is known [Andrews, 1974]. In principle, the slip function for seismic slips may be determined by seismic methods. For example, the distribution of self stress after the mainshock may be inferred from the waveform of the mainshock. Then, the occurrence of aftershocks, which are relaxation of self stress, may be related with the waveform of the mainshock. In fact, Tsujiura (1977, personal communication) shows some correlation between the envelope shape of P waves from large earthquakes and the frequency of their aftershocks. Earthquakes with P waves of gradually rising envelope, suggesting more complex rupture propagation, are accompanied by a large number of aftershocks. On the other hand, very few aftershocks follow a mainshock that shows a P wave form with shorter duration and presumably has a smoother fault plane.

As described earlier, our earthquake model suggests the rougher stress distribution after the mainshock. Then, the distribution of fault zone stress may be smoother during the period of foreshock occurrence. Then, we anticipate more abundant small earthquakes relative to larger ones in the former than in the latter, in accordance with observations that the b-value for foreshocks is sometimes lower than that for aftershocks. Our model may give a physical basis for the difference in magnitude distribution between foreshocks and aftershocks found by Kagan and Knopoff [1976], who analyzed

several earthquake catalogues as a branching point process.

An adequate starting point for developing the fault zone stress model is the study of aftershock phenomena because a large amount of data is available from seismic, geodetic, and geologic observations.

Here, I shall make a preliminary attempt to construct a quantitative stress model for a fault zone during an aftershock sequence.

Immediately after the mainshock the self stress generated by irregular slip function shows the roughest distribution. The stress will be elevated at unbroken barriers and drop at slipped segments of the fault. We shall introduce a single parameter S(t) as a function of time to describe the level of elevated stress. Then, we shall make the following assumptions. (1) When an aftershock occurs, the fault-zone stress S(t) is decreased by a certain amount ΔS. Following Kostrov [1974], we shall assume that an earthquake with seismic moment M_o will introduce a strain change spread over a "volume of seismic region." The corresponding stress change may be estimated simply as M_o/V, where V is the volume of seismic region, or

$$\Delta S = M_o/V \tag{2}$$

We may consider equation (2) as defining the fault-zone stress S in terms of volume of seismic region V, or vice versa. (2) According to Mogi [1962], the probability of occurrence

of fracture in a rock sample increases exponentially with the applied stress S. When a constant stress is applied at t = 0, he found that the probability of occurrence of fracture between t and t + dt is independent of t and given by

$$\mu(t)dt = \mu_o e^{\beta S} dt \tag{3}$$

where β is determined as 0.37 bar^{-1}.

This formula has been used by Mogi [1962], Scholtz [1968], and Utsu [1962, 1970] in the discussion of aftershock phenomena. Hagiwara [1974] applied this formula to the statistical distribution of "ultimate strain" obtained by geodetic measurements and obtained the value of β to be 0.3 bar^{-1}.

Since $e^{\beta\Delta S} \underset{\sim}{} 1 + \beta\Delta S$ for a small increase of stress ΔS, the probability of earthquake occurrence is increased by a fraction $\beta\Delta S$. For example, for a stress increase by 1 bar, the probability is increased by 30-37%.

Denoting a cumulative moment of aftershocks as

$$Q(t) = \Sigma M_o, \tag{4}$$

the fault zone stress S(t) at time t can be written as (assuming no supply from the causes of tectonic stress during the aftershock period),

$$S(t) = S(t_o) - Q(t)/V \tag{5}$$

where t_o is the time immediately after the occurrence of the main shock. Then, from equation (3), we have

$$\mu(t) = \mu_o \exp \{\beta S(t_o) - \beta Q(t)/V\} \qquad (6)$$

Assuming that the fracture rate $\mu(t)$ is independent of magnitude of the aftershock, we can estimate the corresponding rate of total moment increase. The result can be written in the form

$$\frac{dQ(t)}{dt} = R(t_o) \exp(-\beta Q(t)/V) \qquad (7)$$

where $R(t_o)$ is independent of t. The solution of the above differential equation is given by

$$Q(t) = \frac{V}{\beta} \left(\ln\left(\frac{R(t_o)}{V}\right) + \ln t\right) \qquad (8)$$

This logarithmic time dependence of cumulative moment was actually observed by Eaton et al. [1970] for the aftershocks of the Parkfield earthquake of 1966. Their results show that the value of $(\ln 10)V/\beta$, slope of the Q vs. log t relation, is $1 \sim 3 \times 10^{23}$ dyne-cm. Using the value of β obtained by Mogi [1962] and Hagiwara [1974], we find that the volume of aftershock zone of the Parkfield earthquake is $V = 1.5 \sim 4 \times 10^{16}$ cm^3. Using the fault area estimated by Eaton et al. [1970], the fault zone thickness d is given by

$$d = \frac{1.5\sim4 \times 10^{16} \text{ cm}^3}{2 \times 10^{12} \text{ cm}^2} = 0.7\sim2 \times 10^4 \text{ cm}$$

or 70 to 200 meters. The narrow width of aftershock zone is consistent with the planar distribution of aftershock hypocenters determined by Eaton et al.

Using the above value of V, we can estimate the drop of fault-zone stress S by a given earthquake in the case of the Parkfield aftershocks. For example, an earthquake with magnitude 4 and moment 10^{23} dyne-cm, S drops by $M_o/V = 2.5\sim7$ bars.

The period of observation covered by the work of Eaton et al. is from 3 to 80 days from the time of the main shock. During this period the total seismic moment is increased by $1.5\sim3.5 \times 10^{23}$ dyne-cm. This corresponds to the drop of fault-zone stress by about 10 bars during this period.

According to McEvilly et al. [1967], there were three aftershocks with $M > 4$ in the first three days and the total moment due to these events in roughly 4×10^{24}. This corresponds to the drop of fault-zone stress by $100\sim300$ bars during the first three days.

Conclusion

The barrier model of Das and Aki [1977] developed from numerical experiments on rupture propagation over a fault plane with distributed barriers has been shown to explain various observations on major earthquakes. In the model with unbroken barriers, the stress may drop to a low value

on the slipped segment of the fault, may be elevated considerably near the unbroken barriers. Because of the resistance by unbroken barriers, the stress drop for the whole fault zone as seen from outside by geodetic methods is only a small fraction of the initial stress. On the other hand, the stress in the fault zone varies strongly in time and space. The fault zone stress will become smoother with time through aftershocks, creep, and other stress relaxation mechanisms. Thus, the large scale tectonic stress and the stress in the fault zone show quite different behavior during a cycle of earthquake recurrance. We need different strategies for developing quantitative models for them. The strategy for smoothly varying tectonic stress was outlined by dual approaches. One is based on the downward continuation of stress by the propagator method, and the other is a generalized inverse or linear programming solution of observational equations for surface displacements formulated by the finite element method. For the fault zone stress, a preliminary analysis was made on the stress in the fault zone of the Parkfield earthquake of 1966 during the aftershock period.

Finally, we emphasize that the determination of stress is quite a non-unique problem and the testing of stress models by deep-drilling into the fault zone is essential. A deep hole or holes can be used to monitor the stress by the hydraulic fracturing method or the seismic velocity method.

Acknowledgement

This work was supported by the National Science Foundation under grant NEV 76-17160

References

Aki, K., Scaling law of earthquake source time function, Geophys. J. Roy. Astron. Soc., 31, 3-25, 1972.

Aki, K., M. Bouchon, B. Chouet, and S. Das., Quantative prediction of strong motion for a potential earthquake fault, Annali di Geofisica, in press, 1978.

Andrews, D.J., Evaluation of static stress on a fault plane from a Green's function, Bull. Seism. Soc. Am., 64, 1629-1633.

Andrews, D.J., Coupling of energy between tectonic processes and earthquakes, Geophys. J. Roy. Astron. Soc., in press, 1978.

Backus, G.E. and M. Mulcahy, Moment tensors and other phenomenological descriptions of seismic sources, I. Continuous displacements, Geophys. J. Roy. Astron. Soc., 46, 341-361; II. Discontinuous displacements, Geophys. J. Roy. Astron. Soc., 47, 301-329, 1976.

Bouchon, M., A dynamic source model for the San Fernando earthquake, Bull. Seism. Soc. Am., submitted 1978.

Brown, R.L., A dislocation approach to plate interaction, Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, 449 pp., 1975.

Brune, J.N., T.L. Henyey, and R.F. Roy, Heat flow, stress and the rate of slip along the San Andreas fault, California, J. Geophys. Res., 74, 3821-3827, 1969.

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

Eaton, J.P., M.E. O'Neill, and J.N. Murdock, Aftershocks of the 1966 Parkfield-Cholame, California, earthquake: a detailed study, Bull. Seism. Soc. Am., 60, 1151-1197, 1970.

Eshelby, J.D., The determination of the elastic field of an ellipsoidal inclusion and related problems, Proc. Roy. Soc. London, A 241, 376-396, 1957.

Forsyth, D. and S. Uyeda, On the relative importance of the driving forces of plate motion, Geophys. J. Roy. Astron. Soc., 43, 163-200, 1975.

Franklin, J., Well-posed stochastic extensions of ill-posed linear problems, J. Math. Anal. Appl., 21, 682-716, 1970.

Hagiwara, Y., Probability of earthquake occurrence estimated from results of rock fracture experiment, Tectonophysics, 23, 99-103, 1974.

Hanks, T.C., The Kuril trench-Hokkaido rise system: large shallow earthquakes and simple models of deformation, Geophys. Jour., 23, 173-189, 1971.

Hanks, T.C., The faulting mechanism of the San Fernando earthquake, J. Geophys. Res., 79, 1215-1229, 1974.

Hanks, T.C., Earthquake stress drops, ambient tectonic stress, and stresses that drive plate motions, Pageoph., 115, 441-458.

Kagan, Y. and L. Knopoff, Statistical search for non-random features of the seismicity of strong earthquakes, Phys. Earth Planet. Int., 12, 291-318, 1976.

Kagan, Y. and L. Knopoff, Statistical study of the occurrence of shallow earthquakes, in preparation, 1978.

Kostrov, B.V., Seismic moment and energy of earthquakes and seismic flow of rock, Izv. Earth Phys., 23-44, 1974.

Matsuda, T., Surface faults associated with Kita-Izu earthquake of 1930 in Izu Peninsula, Japan, in The Izu Peninsula, edited by M. Hoshino and H. Aoki, Tokai Univ. Press, Tokyo, 1972.

Matsuda, T., Surface faults associated with Nobi (Mino-Owari) earthquake of 1891, Japan, Sokuho (in Japanese), Earthq. Res. Inst., Univ. Tokyo, 13, 85-126, 1974.

Matsuda, T., and K. Yamashino, Surface faults associated with the Izu-Hanto-Oki earthquake of 1974, Japan, Sokuho (in Japanese), Earthq. Res. Inst., Univ. Tokyo, 14, 135-158, 1974.

Matsuda, T., Active faults and earthquakes: the geological aspects, Chishitsugaku-ronshi (in Japanese), 12, 15-32, 1976.

McEvilly, T.V., W.M. Bakun, and K.B. Casaday, The Parkfield, California, earthquakes of 1966, Bull. Seism. Soc. Am., 57, 1121-1244, 1967.

Mogi, K., Study of elastic shocks caused by the fracture of heterogeneous material and its relation to earthquake phenomena, Bull. Earthq. Res. Inst., Univ. Tokyo, 40, 125-173, 1962.

Rikitake, T., Earthquake Prediction, Elsevier Scientific Publishing Co., 357 pp., 1976.

Scholtz, C.H., Microfractures, aftershocks, and seismicity, Bull. Seism. Soc. Am., 58, 1117-1130, 1968.

Solomon, S.C., N.H. Sleep, and R.M. Richardson, On the forces driving plate tectonics: inferences from absolute plate velocities and intraplate stress, Geophys. J. Roy. Astron. Soc., 42, 769-801, 1975.

Tchalenko, J.S. and M. Berberian, Dasht-e Bayaz fault, Iran: earthquake and earlier related structure in bedrock, Geol. Soc. Am. Bull., 86, 703-709, 1975.

Utsu, T., On the nature of three Alaskan aftershock sequences of 1957 and 1958, Bull. Seism. Soc. Am., 52, 279-297, 1962.

Utsu, T., Aftershocks and earthquake statistics II., J. Faculty Science, Hokkaido Univ., Series VII, Geophysics, 3, 197-266, 1970.

Walcott, R.I., Flexural rigidity, thickness, and viscosity of the lithosphere, J. Geophys. Res., 75, 3941-3954, 1970.

Wallace, R.E., Notes on stream channels offset by the San Andreas fault, southern Coast Ranges, California, in Proc. Conf. Geologic Problems of San Andreas Fault System, XI, 6-21, Stanford Univ. Press, 1968.

Watts, A.B. and M. Talwani, Gravity anomalies seaward of deep-sea trenches and their tectonic implications, Geophys. J. Roy. Astron. Soc., 36, 57-90, 1974.

Figure Caption

Figure 1. Relation between the barrier interval and the maximum slip obtained by various methods. The lines correspond to constant stress drop assuming a circular crack with the diameter equal to the barrier interval. Reproduced from Aki *et al*. [1978].

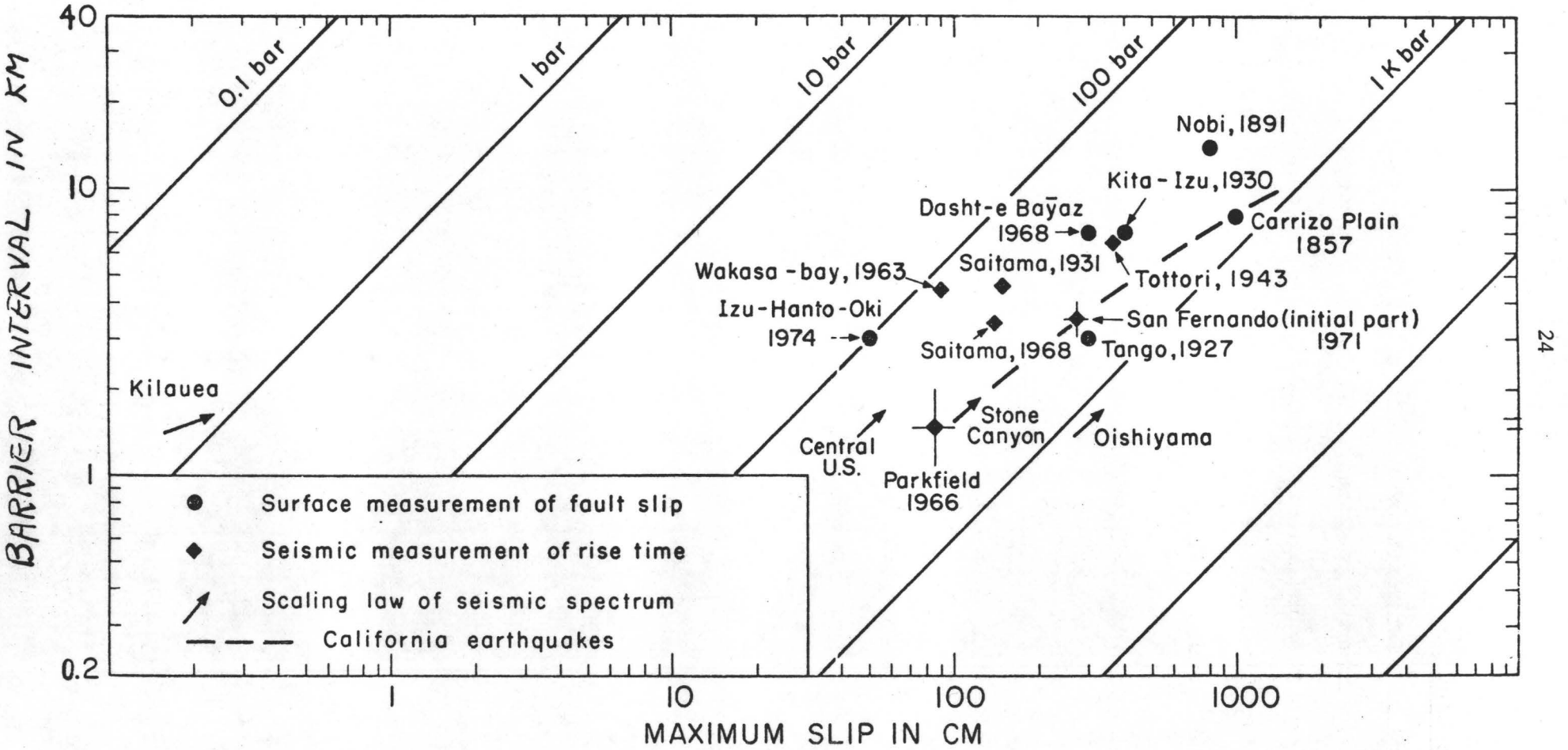

40
10
1
0.2
BARRIER INTERVAL IN KM
0.1 bar
1 bar
10 bar
100 bar
1 K bar
Nobi, 1891
Kita - Izu, 1930
Dasht-e Bāyaz
1968
Carrizo Plain
1857
Wakasa - bay, 1963
Saitama, 1931
Tottori, 1943
Izu-Hanto-Oki
1974
San Fernando (initial part)
1971
Saitama, 1968
Tango, 1927
Kilauea
Stone
Canyon
Central
U.S.
Oishiyama
Parkfield
1966
Surface measurement of fault slip
Seismic measurement of rise time
Scaling law of seismic spectrum
California earthquakes
1
10
100
1000
MAXIMUM SLIP IN CM

Figure 1

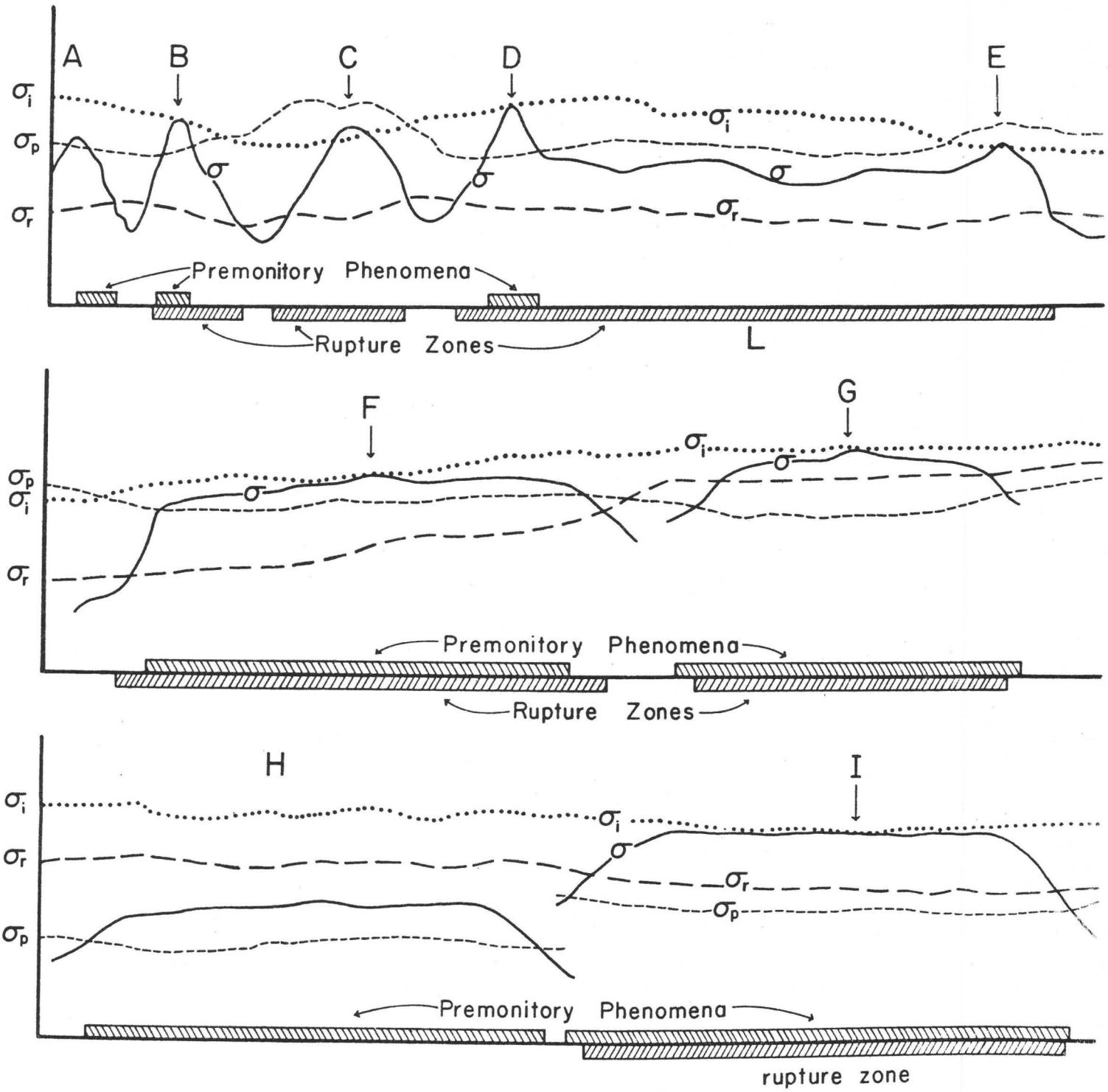

A
B
C
D
E
σ_i
σ_p
σ_r
σ
Premonitory Phenomena
Rupture Zones
L
F
G
H
I
rupture zone

Figure 2

APPLICATION OF RUPTURE PROPAGATION THEORIES TO EARTHQUAKES

by

D. J. Andrews

U.S. Geological Survey

Menlo Park, Ca. 94025

ABSTRACT

Determination of a rupture or seismic source from initial conditions is, in general, a nonlinear problem. However, a special class of rupture models, in which sliding takes place at constant stress after slip is initiated, may be approximated as a mixed boundary value problem. For prescribed motion of the rupture front such a problem is linear. The velocity of the rupture front is determined by the energy absorbed within the nonlinear rupture front region. This nonlinear region must be considered to establish the magnitude of the energy absorbed and its scaling law in going from laboratory experiments to earthquakes. Outstanding problems dependent on the nonlinear rupture front include mixed-mode rupture propagation and the determination of the thickness of the zone that is heated by nonelastic energy loss.

OVERVIEW

I would like to start this discussion of rupture propagation with some philosophical perspectives. Most of the work done in geophysics quite properly consists of natural science, with the researchers being observers of nature. The observers may at times use rather complex mathematics to interpret their data but are still basically interested in perceiving the earth in all its complexity. The rest of geophysics consists of applied physical science, in which the goal of the researchers is to construct models, supposedly representing some idealized version of the earth, that works according to the principles of physics. Both the observers and the modelers may make heavy use of applied mathematics; that is not the basis of the distinction I am making. The observers want to perceive the world, and the modelers want to understand some of its essential mechanisms in terms of physics.

I am speaking today as a modeler, to seek to understand rupture propagation as a physical process determined by initial conditions and material properties in the medium. Still it is necessary to be aware, as a natural scientist would quite naturally be aware, that the initial conditions and material properties in the earth may be quite heterogeneous. I will return to this point at the end of the talk.

A rupture, such as an earhquake, is an inelastic

process that takes place in a solid continuum that to a good approximation is linearly elastic outside the rupture. Indeed, a rupture is a source of elastic stress and displacement fields in the medium.

Elasticity has been compared to other linear field theories and worked out as an example of a tensor field by Morse and Feshbach (1953). It is curious that in Morse and Feshbach and in many other standard texts elasticity is treated as a field theory without a source. What is the source of elastic fields? What is the analogue of electric charge?

The answer is that the source of an elastic field is the deviation of the medium from linear elastic behavior (Eshelby, 1956, 1957; Backus and Mulcahy, 1976a, 1976b). If an element of the medium follows a stress-strain curve that deviates from Hooke's Law, then that deviation ("stress-free strain", a second order tensor), through the effect of strain incompatibility with neighboring elements, is the source density of the resulting elastic field. When multiplied by the elastic constant tensor, this source density is the "moment tensor density", and its volume integral is the seismic moment tensor. The tensor divergence of this source density is the more familiar, though less physically meaningful, "equivalent body force". A volume distribution of stress-free strain that is concentrated to become a delta function on a surface represents a

displacement discontinuity, as on a crack or fault.

Rupture propagation, then, is concerned with the question of the development in time of a space-time distribution of stress-free strain, a breakdown of elasticity. An analogue is dielectric breakdown and the propagation of a lightning stroke.

In this talk I want to classify modeling efforts into four conceptual levels. Level I consists of the determination of the displacement and stress fields arising from a prescribed source distribution, for instance slip on a fault prescribed as a function of space and time. The electromagnetic analogy would be the determination of the electric and magnetic fields from a prescribed charge and current distribution. In the elastic case Level I modeling can lead to quite complicated mathematics, for compressional and shear waves are coupled together at any interface or heterogeneity.

If the source distribution is described in terms of an assumed functional form characterized by a small number of parameters, then Level I modeling can indicate to the modeler how observed field variables are related to those parameters. More generally, linear inverse theory can be used to infer from observed field variables the seismic moment tensor, and, with limited resolution, the space- time distribution of the moment tensor density. A unique inverse does not exist from far-field data alone, regardless of the

completeness of azimuthal coverage or accuracy of the data (Aki, personal communication; Kostrov, 1975).

From the point of view of a pure natural scientist, an observer, inverse Level I modeling is entirely sufficient to study seismic sources. However, a modeler will not be satisfied and will want a model based on physical principles of continuum mechanics that is consistent with laboratory observations of rock properties. Laboratory rock mechanics experiments have not yet been done with fast enough time resolution to record motion from a dynamically propagating rupture. Therefore the modeling effort must be largely theoretical.

The term "rupture" is taken to mean a more or less abrupt drop in stress accompanied by an increase in stress-free strain that might be concentrated on a surface as a displacement discontinuity. It can be either fracture of virgin rock or unstable frictional sliding on a pre-existing fault.

I define as Level II those models in which stress is assumed to drop abruptly at the rupture front and then remains at a prescribed level as stress-free strain or slip develops. An example is Archambeau's (1968) model in which stress is prescribed on an expanding surface and the nonelastic source develops within the enclosed volume. If the source volume is collapsed onto a surface then we have a crack model, which is a mixed boundary value problem. Beyond

the edge of the crack displacement discontinuity is zero and on the crack surface inside the edge stress is prescribed. For prescribed motion of the crack edge this is a linear boundary value problem, and it is possible to attack such problems with analytic methods.

I want to jump ahead now to define Level III modeling. This is true physical modeling in which both the source (stress-free strain or slip) and the dynamic elastic field are determined by initial conditions. The stress-free strain at each point in space-time is consistent with the stress history of the particulat material element through a nonelastic constitutive relation. At the same time the stress and displacement fields are consistent with elastic field equations with stress-free strain included as a source density. This is analogous to an antenna problem in electromagnetism, in which source and radiated field must be self-consistent. The dynamic development of rupture is determined by assumed initial stress and material constitutive relations that may depend on position. In order to have an instability, such as spontaneous rupture nucleation and propagation, it is necessary that the constitutive relation allow for decreasing elastic strain energy as stress-free strain increases. A Level III model is, by definition, nonlinear, and cannot in general be attacked by analytic methods.

The importance of Level II modeling, the mixed boundary

value problem, is that if a physical criterion is provided to determine the propagation of the crack tip then Level II models comprise a special subgroup of Level III models. For a sharp-tipped crack propagating at less than a limiting velocity, inverse square root singularities of stress and particle velocity exist at the crack tip, at the juncture of the different types of boundary condition, up to a critical crack velocity. The tip of a prpopagating crack is an energy sink. The partition of released strain energy between absorption at the crack tip and radiation is a function of crack velocity. If the energy absorbed at the crack tip were known, either as a material constant or as determined by dynamic processes within the crack tip, then the propagation of the crack would be determined. In this way analytic solution techniques have come close to the solution of an essentially nonlinear problem. All the nonlinearity is concentrated at a singularity, the crack tip.

More generally, we may consider some constitutive model in which some arbitrary nonelastic breakdown process occurs followed by sliding at constant stress. Then the region in which the stress drops is no longer a singularity, the crack tip, but is spread out into a rupture front, analogous to a boundary layer in fluid mechanics. The inner solution in the rupture front must be matched to the large scale outer solution, which to a good approximation may be taken to be a sharp tipped crack solution. The important parameter in the

matching is the energy lost in the rupture front. In this way analytic Level II solutions can be generalized to a special class of Level III problems.

An outstanding difficulty in Level III modeling is that the constitutive law governing the failure process is not known. Conventional rock mechanics experiments provide only gross information about the initial and final states. Dynamic measurements with very fast time resolution that can reveal the structure of the rupture front are needed.

So far the initial conditions determining an event have been dicussed as if they were arbitrary choices. Here we reach the conceptual limitation of Level III modeling. Initial stress and material properties cannot be uniform if a rupture is to stop. Furthermore, the detailed variation of stress and material properties from point to point within the earth is essentially unknowable. The initial conditions are determined, of course, by the past history of faulting and tectonic loading. Statistical mechanical modeling of these initial conditions I define as Level IV modeling. Virtually nothing has been done at this level of modeling. This is what is required, however, if physical modelers are to contribute anything toward understanding the questions of seismicity statistics, foreshocks, aftershocks, and the high frequency spectrum of ground motion.

SELF-SIMILAR SHARP-TIPPED CRACK

Let us return now to Level II, the mixed boundary value

problem, and consider rupture propagation in the case of a uniform initial stress field and a uniform elastic medium with no reflecting boundaries. In this case if a rupture initiates it will not stop, but this case might provide an approximation to the initial growth of an actual rupture. The prescription of the problem is that on an elliptical area expanding at a constant velocity shear stress drops to a prescribed constant level. This problem has a self-similar solution, and if the propagation velocity of all points of the crack edge is less than the Rayleigh wave velocity then the slip function plotted as a function of position on the crack surface is an ellipsoid (Kostrov 1964, Burridge and Willis 1969, Dahlen 1974, Richards, 1976). A cross section through the solution is shown schematically in Figure 1. Ahead of the crack edge stress rises from the initial value and has an inverse-square-root singularity as the crack edge approaches. Inside the crack edge stress is at a constant level lower than the initial value. What is plotted is the stress change; the zero level of absolute stress would presumably lie somewhere below the sliding stress level. The slip function has a square root dependence near the crack edge, which means that the particle velocity has an inverse-square-root singularity. On the crack plane itself the particle velocity singularity does not overlap the stress singularity.

The dynamic solution for a crack growing self-similarly

is qualitatively the same as a static solution. If the dynamic solution is compared at some instant of time with a static crack of the same size and equal stress drop, it is seen that the amplitude of both the slip function and the stress singularity ahead of the crack are smaller in the dynamic case. The dependence of these amplitudes on rupture velocity is shown in figure 2. The quantity Q is the amplitude of the slip function of a dynamic circular crack relative to that of a static crack of the same size and stress drop. It is a smooth curve that falls off about 20 percent as rupture velocity increases from zero to the Rayleigh wave velocity. The intensity of the singularity in particle velocity is this function times the rupture velocity itself.

The stress intensity factor may vary around the edge of the elliptical crack. For the sake of definiteness let the y axis of a Cartesian coordinate system be normal to the crack plane and let the slip be in the x direction. It is the shear stress component sxy that is prescribed to drop inside the crack edge, and it is the amplitude of the singularity of this component that is called the stress intensity factor. Locate the origin of the coordinate system at the center of the ellipse. The solution on the x axis as the crack edge is approached is a two-dimensional plane strain configuration (mode II crack). The stress intensity factor at this point relative to that of a static crack of

the same size and stress drop is labeled k_{II} in the figure. It decreases smoothly as a function of the propagation velocity of this point of the crack edge and crosses zero at the Rayleigh wave velocity. On the z axis near the crack edge the solution approaches a two dimensional antiplane strain configuration (mode III crack). The stress intensity factor at this point (labeled k_{III} in the figure) falls smoothly as a function of rupture velocity down to zero at the shear wave velocity.

The singular crack edge is a sink for energy. To see this, consider the energy flux through the surface of a tube of radius r surrounding the crack edge. Energy flux is stress times particle velocity. Although the peaks of particle velocity and stress do not overlap on the crack plane, they do overlap on the tubular surface and each is proportional to $r^{-1/2}$. Their product, proportional to $1/r$, times the surface area of the tube is independent of r. Therefore the energy flowing through the tubular surface is independent of the radius of the tube, and the crack edge is an energy sink. Since the tube is moving through the material, one must also consider the flux of kinetic and internal energy by material transport, but there is no net contribution from this source (Freund 1972). The energy flux to the crack edge is proportional to the product of the Q and k factors shown in Figure 2, which is a function of rupture velocity alone, and is also proportional

to the distance the crack has propagated in the self-similar solution.

If the elastic material inside a tube of some particular radius were replaced by a region in which nonelastic deformation takes place, then the elastic solution outside the tube could still hold to a good approximation if the energy flowing through the tube matches the energy absorbed by nonelastic deformation inside the tube. In this way the mixed boundary value problem can be generalized to a case where the crack tip is smeared out over a rupture front region in which nonelastic deformation occurs. The relevant parameter in matching an inner solution to the outer solution is the energy absorption in the rupture front. The outer solution is a mixed boundary value problem and can be approached analytically. Assuming a value for fracture surface energy obviates the need for an inner solution. However, this assumption needs to be examined, and the energy absorption is determined by the inner solution in the rupture front region, which is a nonlinear problem.

The nonlinear inner solution can be found from numerical calculations, if a constitutive relation is specified to determine the development of stress-free strain in the early stage of rupture before sliding takes place at constant stress. One may expect in general that nonelastic deformation will take place in a volume with finite

thickness. However, I will first show some results for an artificial model in which all nonelastic deformation is confined to the crack plane.

SLIP-WEAKENING MODEL

In the slip-weakening model shown in Figure 3 the stress level at which sliding takes place is a function of the amount of slip at each point on the crack plane. Slip does not occur until shear stress reaches the limiting stress, analogous to a plastic yield stress. As slip increases the limiting stress decreases to the sliding friction level, and then further slip takes place at constant stress. The shaded area in the Figure is the energy absorbed on the crack plane in excess of the energy that would be absorbed if stress were constant at the sliding friction level (assuming that there is no singular energy sink). Palmer and Rice (1973) have shown that it is the energy represented by the shaded area that is relevant to rupture propagation; the absolute level of stress is relevant only to the heat produced on the fault. In this artificial constitutive model the fracture surface energy is constant and is independent of the dynamic solution.

The ratio of fracture surface energy to energy released per unit volume by the stress drop determines a critical crack length. A critical crack solution for a slip weakening model is shown in Figure 4. The crack tips are not singular, but are spread over a significant fraction of the crack

length. The solution shown is a two dimensional plane strain case. This numerical solution, which was found by trial, is slightly beyond the verge of instability.

Rupture propagation in the dynamic solution developing from this nearly critical solution is shown in Figure 5, where the location of the rupture front in space-time is shown by the shaded region. The rupture front is initially fairly wide, and the rate of growth is quite slow. By the time the crack has doubled its length the rupture velocity is a significant fraction of the Rayleigh wave velocity, and it approaches the Rayleigh velocity asymptotically.

Ida (1972) has examined rupture propagation with slip-weakening models in antiplane strain semi-analytically. He finds that the length of the rupture front varies inversely with the length the crack has propagated. The same tightening up of the rupture front is seen in this numerical calculation in plane strain.

A snapshot of the solution is shown in Figure 6 at the instant of time when the rupture has propagated to 5.5 times the critical half length. The length interval in which stress drops is much shorter than for the critical crack. The stress peak at the rupture front is propagating at slightly less than the Rayleigh velocity. There is another stress peak ahead of the rupture front propagating at the shear wave velocity, which leads to another story (Andrews 1976b) that I will not go into now.

As the crack grows, the ratio of fracture surface energy to energy released per unit advance of the rupture front decreases. The slip function becomes more singular as the size of the rupture front decreases, and the rupture velocity approaches the Rayleigh velocity, which is the velocity at which a plane strain crack with zero fracture energy will propagate. As the crack length becomes much longer than the critical length the solution approaches the solution for a cohesionless crack.

The motivation for using the slip-weakening model was to examine a simple rupture model in which stress is finite. The shear stress component sxy is indeed finite and is given by the slip weakening law. However, as the crack gets longer, the rupture front gets shorter, the derivative of the slip function gets larger, and strain components other than exy get larger. Any finite elastic stress limit will be exceeded for a sufficiently large crack length.

Typical values of fracture surface energy measured in laboratory experiments are 0.5e3 erg/cm^2 for single crystals and 0.5e5 erg/cm^2 for polycrystralline rocks. These values, together with assumed stress drops ranging from 100 bars to 1 kilobar, give critical crack lengths ranging from microns to centimeters. Clearly we are dealing with much larger length scales in earthquakes. If fracture surface energy is constant at a laboratory value, then it is negligible at the length scale of an earthquake, and the dynamic solution

approaches that for a cohesionless crack.

For either a cohesionless crack or a crack with constant fracture surface energy that has propagated a distance much larger than the critical length shear stress components other than sxy will exceed any finite elastic limit in some volume around the rupture front. In the cohesionless case a singularity in sxy is avoided by the crack propagating at the limiting velocity, the Rayleigh velocity in plane strain or the shear velocity in antiplane strain. But other components of shear stress are singular at any rupture velocity due to the slip function having a singular derivative.

These singular stress components will produce microcracks at orientations different from the main crack, as shown in Figure 7. In antiplane strain the microcracks will be perpendicular to the main crack, and in plane strain they will be at 45 degrees to it. Slip on the microcracks is in opposite directions on opposite sides of the main crack, so that they do not contribute to the distant solution. The significance of the microcracks is that additional energy is absorbed in a larger rupture front volume. Even in the case of stick-slip friction, there will be an effective fracture surface energy due to microcracking at the rupture front. The basic assumption underlying this discussion is that if any solid material is strained at a sufficiently high stress it will absorb energy.

STRAIN-WEAKENING MODEL

To test the idea that inelastic energy loss will increase as the crack lengthens a numerical calculation was done with a strain-weakening constitutive law, rather than the slip-weakening law used before (Andrews 1976a). Nonelastic deformation takes place in a volume of finite thickness. In this calculation rupture takes place in a continuous uniform medium, and the prospective fault plane has no special properties. Fracture energy absorbed per unit area of rupture is not a material constant, but is proportional to the thickness of the rupture front, which is determined as part of the dynamic solution. In Figure 8 the shaded region shows where plastic strain exists at an instant of time in the dynamic solution. Near the origin the plastic zone is only one finite element thick. Farther out the plastic zone thickness increases linearly with rupture length. The solution grows self-similarly in time. The energy absorbed at the rupture front is a constant fraction of the energy released by the stress drop, so the fracture energy per unit area on the fault plane increases with crack length.

SCALING

The result that fracture surface energy is proportional to the distance rupture has propagated will hold in general if the constitutive law does not determine a length scale. This will be true if plastic strain is determined by the

past stress-strain path but is independent of strain rate, and if layer dimensions, gravity, and heat conduction are not important. In that case if length, time, and displacement are multiplied by the same scale factor another valid dynamic solution is obtained in which stress and strain at scaled space-time points are unchanged. If in addition the initial state is uniform then the dynamic solution will be self similar.

Clearly scaling is important in applying laboratory results to the field. The scaling laws that apply depend on the rate dependence of the constitutive law and upon whether the failure process is confined to a surface or takes place through a volume. For instance, the slip-weakening and the strain-weakening models discussed here scale differently. Also Burridge has pointed out that if stress on a crack surface is a function of slip velocity a self similar solution is to be expected, in contrast to the rate-independent slip-weakening model discussed here.

MIXED-MODE RUPTURE

The limiting rupture propagation velocity of a pure antiplane shear crack is the shear wave velocity and of a pure plane strain shear crack is the Rayleigh velocity. What is to be expected for a mixed mode shear crack? This is still an outstanding problem. What is the solution if a crack propagates between the Rayleigh velocity and the shear wave velocity?

In his excellent review of self similar elliptical shear cracks Richards (1976) was concerned that the energy absorbed at the crack edge could be embarassingly larger than laboratory values of fracture energy. He proposed that the solution for sub-Rayleigh rupture velocity be applied above the Rayleigh velocity where the plane strain stress intensity factor is negative. Then the plane strain component could generate energy at the crack edge while the antiplane component absorbs energy. Then the rupture velocity could be adjusted until the generation and absorption of energy nearly cancel.

Freund and Clifton (1974) have shown that a solution in which energy is generated at the crack edge is not unique, and Burridge (1973) has found a different plane strain solution for rupture velocity between the Rayleigh and shear wave velocities. Still, Richards' proposal can not be disproved from within the context of mixed boundary value problems.

To examine the reasonableness of Richards' proposal we must ask whether his singular solution is the limit of a class of nonsingular solutions satisfying physically realistic constitutive relations. If Richards' solution is convolved with a function that has the effect of smearing it out slightly on the crack plane then the solution to some slip-weakening law is obtained. What are the properties of such a slip-weakening model? Associated with the antiplane

component of slip is a component of traction in the opposite direction, but the plane strain component of slip is associated with a traction in the same direction in the rupture front region. At a rupture velocity with no net energy absorption the traction vector will be perpendicular to the slip velocity vector on the average within the rupture front. This is not a physically realistic model.

If you want to talk about nonelastic yielding in a volume instead of on a surface, you must consider whether a plastic flow rule could allow energy to be transferred from plane strain to antiplane strain deformation. In principal stress space a plastic strain increment vector need not be exactly parallel to the stress vector, but it seems unlikely that they could be at right angles. Therefore Richards' proposal is not physically reasonable. A solution for mixed-mode rupture propagation between the Rayleigh and shear velocities that is consistent with a physically realistic constitutive law inside the rupture front has yet to be found. The nonlinear inner solution is essential to the question.

STOPPING

So far my discussion has been limited to a uniform medium with uniform stress. In such a case a rupture will not stop. The important question of the size of an earthquake is determined by nonuniformities of stress and material properties.

Burridge and Halliday (1971) have obtained an analytic solution that stops for a cohesionless frictional sliding model in which the difference between the friction level and the initial stress level varies as a function of position. Rupture begins in a region where initial stress is above the sliding friction level. As the rupture propagates out into a region where there is no longer a net stress drop, slip slows down and stops. In this model the rupture must penetrate into a region where the final stress is larger than the initial stress.

Husseini et al. (1975) have proposed that rupture is stopped by a barrier, a large increase in fracture surface energy. The result of their analysis is that the total fracture energy required to stop a rupture for a given stress drop amounts to a certain fraction of the total elastic strain energy released. Dividing the total fracture energy by the fault area ruptured gives an average fracture surface energy proportional to crack length. Husseini et al. interpreted their results in terms of fracture surface energy being a material property. Then for a barrier of given fracture energy to stop earthquakes of different sizes average stress drop of an earthquake would have to be inversely proportional to the square root of length of rupture.

From the point of view that fracture surface energy is not a material property but is determined by the dynamic

solution itself, one cannot escape the conclusion that the difference between stress and sliding friction must vary on the fault surface. This difference must vary at all length scales to explain earthquakes of different sizes. Furthermore, the heterogeneity must be reproduced after a cycle of earthquakes of all sizes.

ENERGY BUDGET

The energy budget of an earthquake is indicated schematically in Figure 9, which is a plot of stress vs. slip for a typical point on a fault plane. Area on this plot represents energy per unit fault surface, so the plot must be imagined to be integrated over the fault surface to get the total energy budget. Of course, initial stress, sliding friction, final stress, and final slip vary over the fault surface. The dashed line connecting the initial and final states is a hypothetical quasistatic process, and the trapezoidal area under this line is the difference between the elastic strain energy of the initial and final static states. The rectangular area under the sliding friction level is energy lost to frictional heating on the fault. The difference between the trapezoidal and rectangular areas, indicated by the diagonal hatching, is available energy, which may be partitioned between fracture energy and radiation of elastic waves. The ratio of the hatched area to the trapezoidal area is conventionally defined to be the seismic efficiency.

If the rupture is a smooth event, with most of the slipping area in communication with itself while slip proceeds, such as the case calculated by Burridge and Halliday, the final static stress is below the sliding friction stress. If, on the other hand, a patch slips and then locks while slip still continues farther down the fault, then the final static stress is larger than the sliding friction stress. In this case more energy can be radiated for a given static stress drop.

The total energy absorbed in nonelastic deformation is represented by the area under the stress-strain curve that rises to the upper yield stress and then falls to and follows the sliding friction stress. In addition to the frictional heating, fracture energy, represented by the black area, is included. This energy loss, due to focusing of energy at the rupture front, is quite schematic in the diagram; it includes work done against other components of stress-free strain on microcracks.

The energy that is radiated is the energy available, the hatched area, minus the fracture energy, the black area. Husseini (1977) defines radiation efficiency to be the ratio of radiated energy to available energy. Radiation efficiency determines rupture velocity, so it may be inferred from seismic observations. On the other hand, the dynamic motion is independent of the seismic efficiency, so the absolute stress level cannot be inferred from ground motion.

A possible contribution that modelers might make to the question of the absolute stress level concerns thermal effects on the fault. Heating from nonelastic deformation could change the mechanical properties of the medium and reduce the friction level as slip increases. If melting occurs or if fluid pressure rises sufficiently, stress drop may be nearly complete. The essential variable that is not known in such considerations is the thickness of the zone in which heat from nonelastic deformation is deposited. So far theoretical modeling has contributed virtually nothing to this question.

In conclusion it should be noted that this discussion has been restricted to generalizations of the mixed boundary value problem, in which sliding is assumed to take place at constant stress after an initial breakdown process is completed. This is an assumption that needs to be questioned. There has been much discussion at this meeting of the irregularity of faulting. Theoretical treatments of heterogeneous conditions may have to be statistical in nature in general. Some interesting conclusions about radiated waveforms have been reached by Das and Aki (1977) from a simple deterministic model in which a rupture jumps barriers that do not break.

REFERENCES

Andrews, D. J., Rupture propagation with finite stress in antiplane strain, J. Geophys. Res., 81, 3575-3582, 1976a.

Andrews, D. J., Rupture velocity of plane strain shear cracks, J. Geophys. Res., 81, 5679-5687, 1976b.

Archambeau, C. B., General theory of elastodynamic source fields, Rev. Geophys., 6, 241-288, 1968.

Backus, G. E. and M. Mulcahy, Moment tensors and other phenomenological descriptions of seismic sources - I. Continuous displacements, Geophys. J., 46, 341-361, 1976a.

Backus, G. E. and M. Mulcahy, Moment tensors and other phenomenological descriptions of seismic sources - II. Discontinuous displacements, Geophys. J., 47, 301,329, 1976b.

Burridge, R. and J. Willis, The self-similar problem of the expanding elliptical crack in an anisotropic solid, Proc. Camb. Phil. Soc., 66, 443-468, 1969.

Burridge, R. and G. S. Halliday, Dynamic shear cracks with friction as models for shallow focus earthquakes, Geophys. J., 25, 261-283, 1971.

Burridge, R., Admissible speeds for plane-strain self-similar shear cracks with friction but lacking cohesion, Geophys. J., 35, 439-455, 1973.

Dahlen, F. A., On the ratio of P-wave to S-wave corner

frequencies for shallow earthquake sources, Bull. Seismol. Soc. Amer., 64, 1159-1180, 1974.

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

Eshelby, J. D., The continuum theory of lattice defects, Solid State phys., 3, 79-144, 1956.

Eshelby, J. D., The determination of the elastic field of an ellipsoidal inclusion, and related problems, Proc. Roy. Soc. Lond. A, 241, 376-396, 1957.

Freund, L. B., Energy flux into the tip of an extending crack in an elastic solid, J. Elasticity, 2, 341-349, 1972.

Freund, L. B. and R. J. Clifton, On the uniqueness of plane elastodynamic solutions for running cracks, J. Elasticity, 4, 293-299, 1974.

Husseini, M. I., D. B. Jovanovich, M. J. Randall, and L. B. Freund, The fracture energy of earthquakes, Geophys. J., 43, 367-385, 1975.

Husseini, M. I., Energy balance for motion along a fault, Geophys. J., 49, 699-714, 1977.

Ida, Y., Cohesive force across the tip of a longitudinal shear crack and Griffith's specific surface energy, J. Geophys. Res., 77, 3796-3805, 1972.

Kostrov, B. V., Self-similar problems of propagation of shear cracks, J. Appl. Math. Mech. (PMM), 28,

1077-1087, 1964.

Kostrov, B. V., The Mechanism of the Focus of Tectonic Earthquakes, Ch. 3 Sec. 6., Nauka, Moscow, 1975.

Morse, P. M. and H. Feshbach, Methods of Theoretical Physics, McGraw-Hill, New York, p. 142, 1953.

Palmer, A. C. and J. R. Rice, The growth of slip surfaces in the progressive failure of over-consolidated clay, Proc. Roy. Soc. Lond. A, 332, 527-548, 1973.

Richards, P. G., Dynamic motions near an earthquake fault: a three- dimensional solution, Bull. Seismol. Soc. Amer., 66, 1-32, 1976.

FIGURE CAPTIONS

1. Slip and stress as a function of position on crack plane for a static or sub-Rayleigh dynamic crack.
2. Dynamic intensity factors relative to static values for slip, Q, antiplane stress, k_{III}, and plane strain stress, k_{II}, shown schematically as a function of rupture velocity.
3. Slip-weakening model.
4. Critical crack. Solid curve is slip function, dashed curve is stress change.
5. Space-time plot of rupture front.
6. Dynamic crack.
7. Orientation of microcracks. Top, antiplane strain; bottom, plane strain.
8. Snapshot of rupture calculated with strain-weakening model. Stress-free strain exists in shaded region. Contours show particle velocity, which is nonsingular.
9. Energy budget illustrated on plot of stress vs. slip for a typical point on fault.

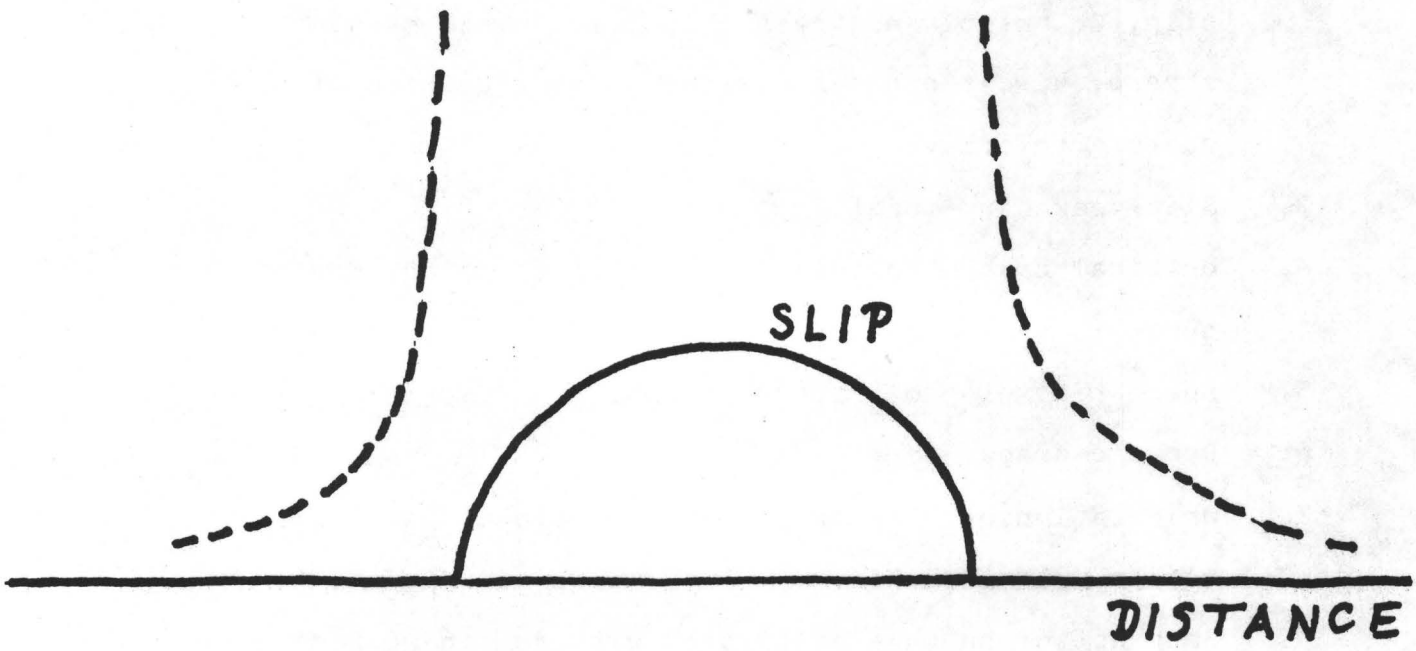


Figure 1

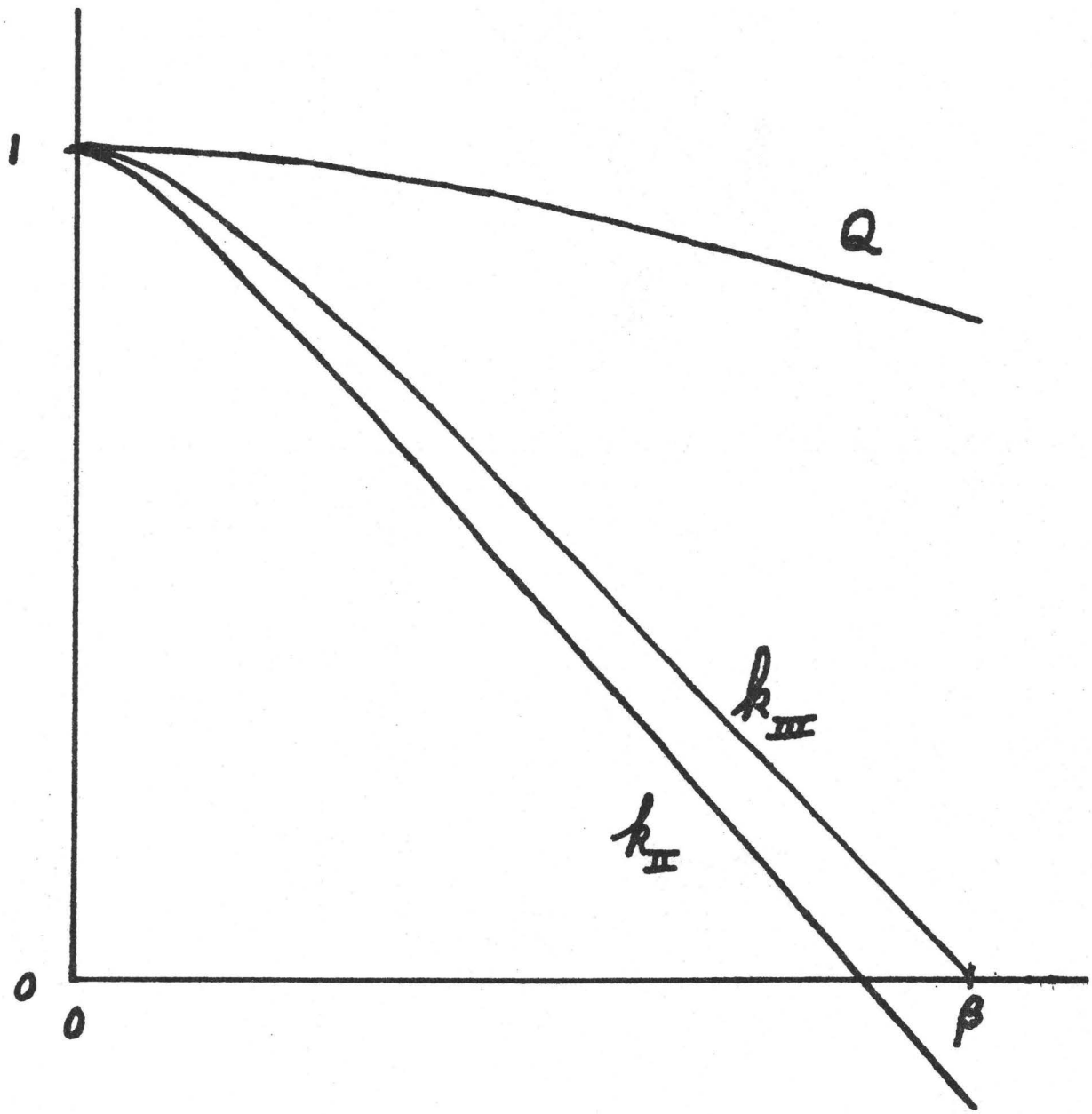


Figure 2

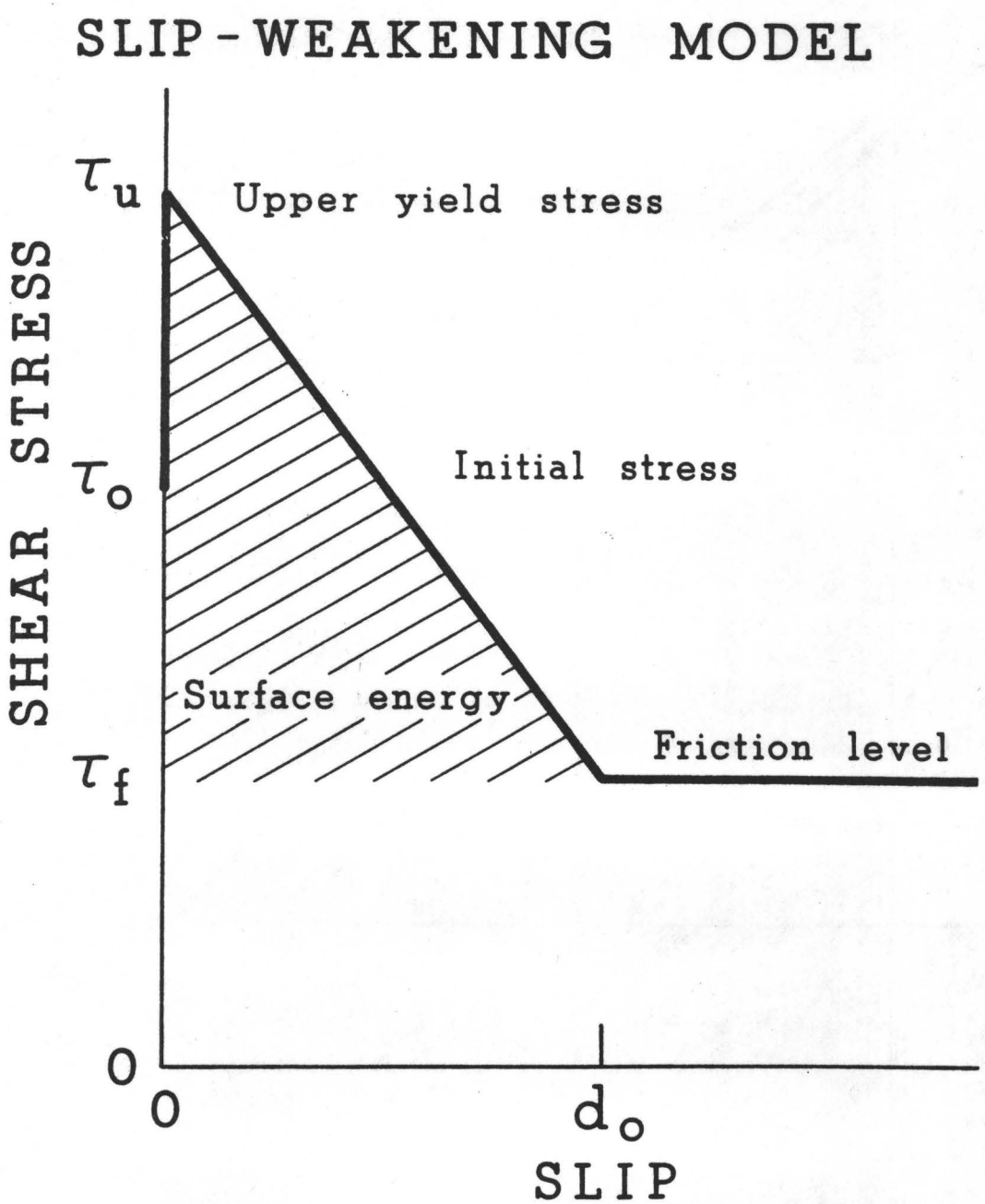

SLIP-WEAKENING MODEL
SHEAR STRESS
τ_u
τ_o
τ_f
Upper yield stress
Initial stress
Surface energy
Friction level
0
0
d_o
SLIP

Figure 3

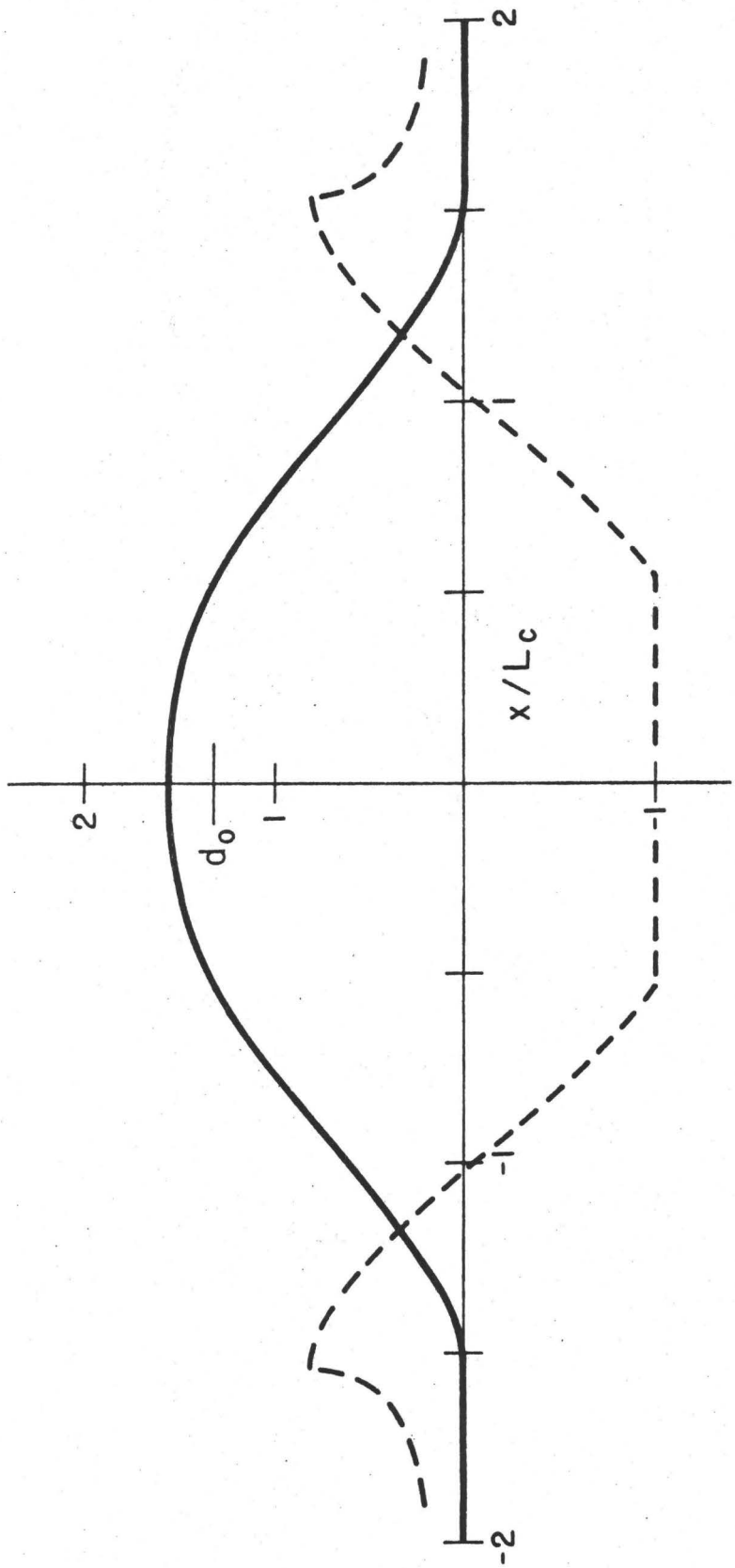


Figure 4

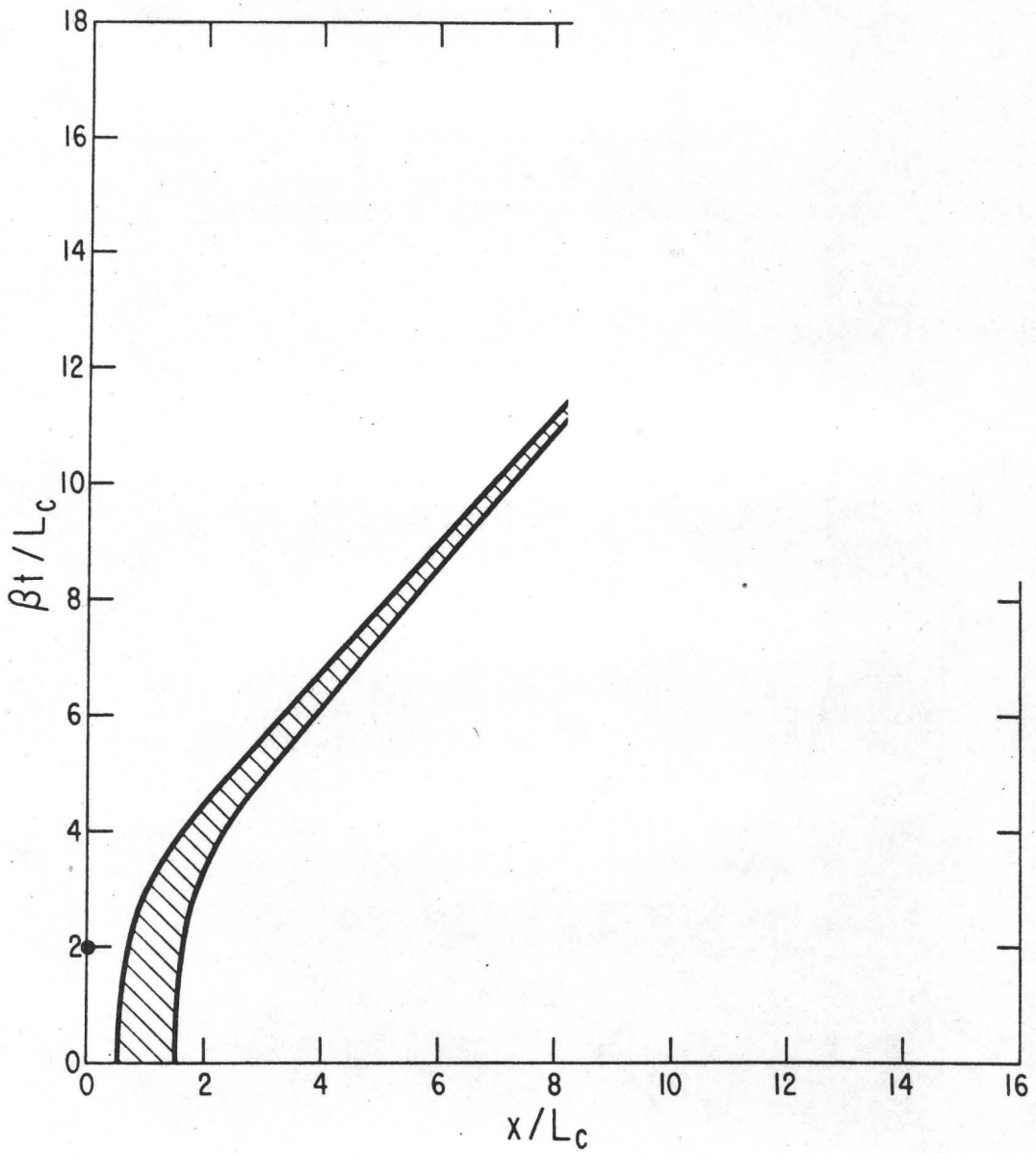


Figure 5

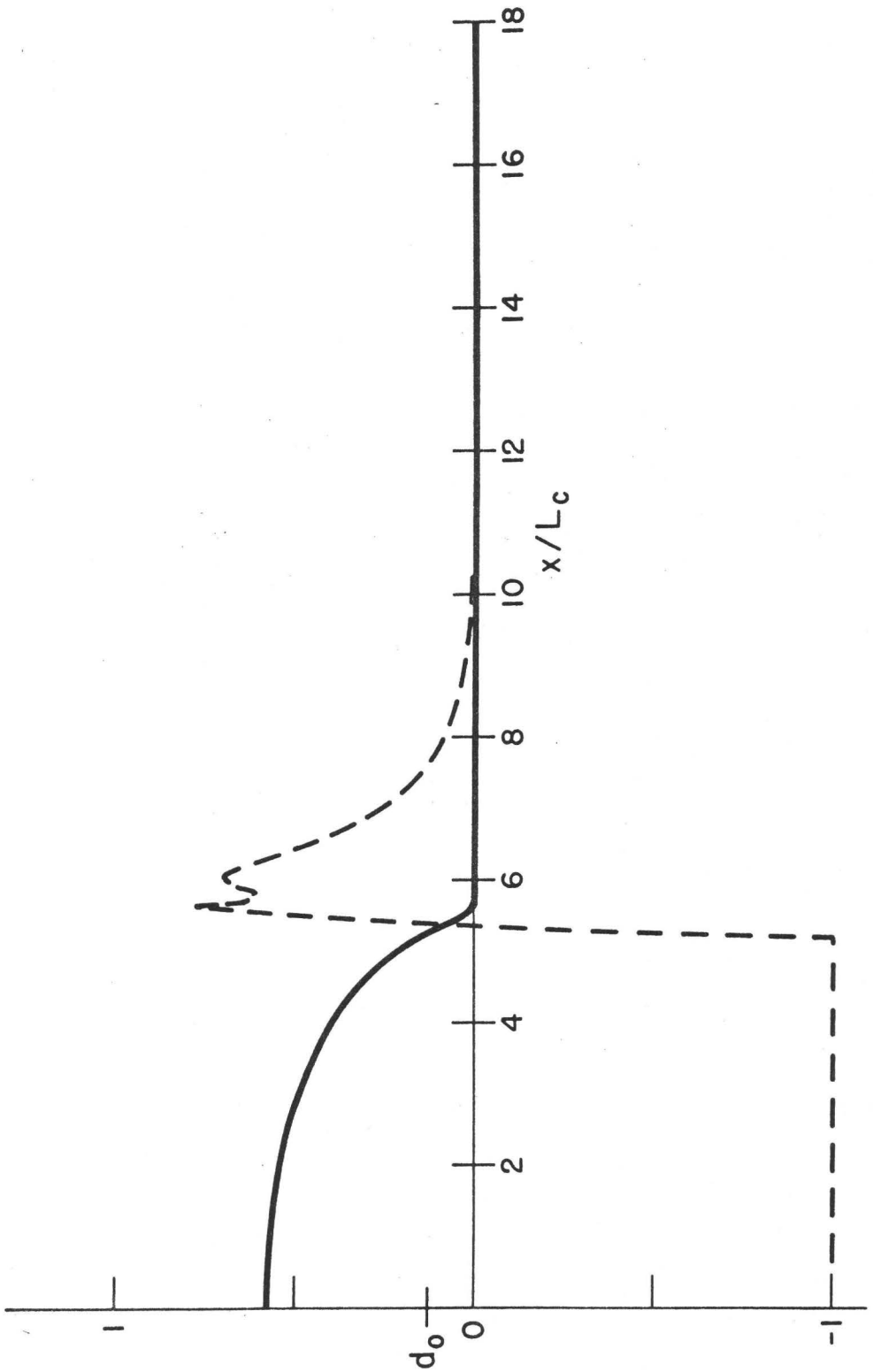


Figure 6

Figure 7

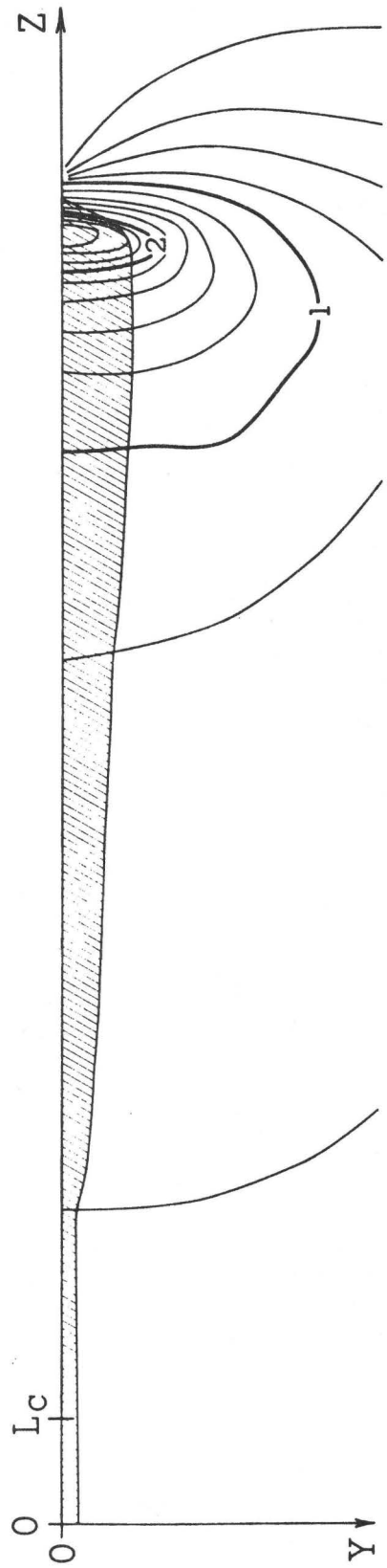


Figure 8

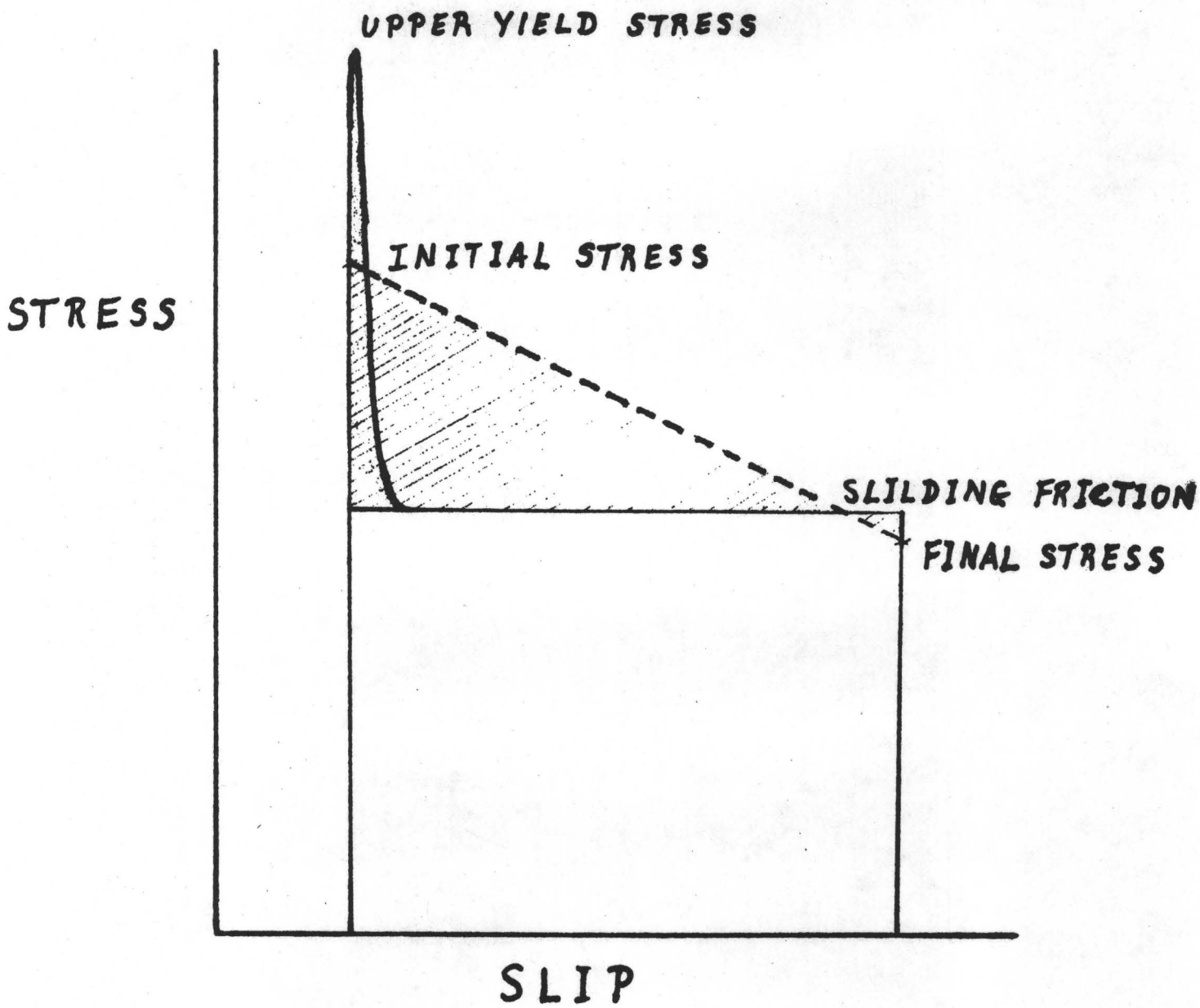

UPPER YIELD STRESS
INITIAL STRESS
STRESS
SLILDING FRICTION
FINAL STRESS
SLIP

Figure 9

IMPLICATIONS OF EARTHQUAKE TRIGGERING AND RUPTURE PROPAGATION FOR EARTHQUAKE PREDICTION BASED ON PREMONITORY PHENOMENA

James N. Brune
Scripps Institute of Oceanography
University of California
La Jolla, California 92037

One approach to earthquake prediction is to think up earthquake models for which prediction will work, and test to determine if these models correspond to nature. An associated task might be to think up earthquake models for which prediction won't work and test them. One risk of the first approach is that one might go on indefinitely discarding models and not hit on one that corresponds to nature. On the other hand, in the second approach you might find that your model corresponds to nature and thus save some money and effort to apply to other means of earthquake hazard reduction. Since there have been a number of optimistic models proposed which have not been verified, this paper attempts to define a reasonable model for which earthquake prediction based on premonitory phenomena won't work and see if it can be refuted. The model discussed here is called the triggering rupture propagation model (Brune, 1974). The model is suggested by the following observations:

1. Many large earthquakes appear to be made of a series of multiple events successively triggered, the initial event, or hypocenter often being at the edge of the region of energy release.

2. Some earthquakes appear to have triggered small anounts of motion considerably outside the region of main energy release (Allen *et al.*, 1968; Allen, *et al.*, 1972).

3. Earthquake mechanism studies, particularly studies of earthquake spectra, suggest that earthquakes can be modeled as ruptures that start at a point and grow

initially outward from this point and then continue either bilaterally or unilaterally. There is often continued slower slip after the main slip.

4. Spectral studies of small earthquakes suggest a complex small-scale stress field.

These facts suggest that a fault may rest in a state of stress considerably below that necessary to initiate slip, and yet be triggered and caused to slip by nearby earthquakes or by a propagating rupture. Thus, in order to predict slip on a particular fault section, one must predict when nearby parts of that fault or adjacent faults will slip and whether or not they will trigger motion or cause a rupture to propagate into the section. However, the same reasoning can be applied to these adjacent parts or faults, and so on back to the "initial event". The "initial event" or initial rupture presumably occurs where the state of stress is high enough to spontaneously initiate rupture, or high enough to allow triggering by much weaker strain sources such as tides, atmospheric loading or gradual or episodic tectonic strain build up. This is the region where one might expect premonitory phenomena such as dilatancy, microfracturing, etc., and associated anomalous behavior.

In this model, to predict large earthquakes, we must predict rupture initiations. However, presumably every small earthquake begins as a rupture initiation. Thus, one must either predict every small earthquake and determine which of these will in fact grow into a large earthquake, or know in some way that only a small subset of rupture initiations (hopefully one) can grow into a large earthquake, and be able to predict this one.

To be more precise in defining this model, we define the following stress parameters: σ_r, the stress necessary for rupture propagation (σ_f, the sliding frictional stress, is presumably close to σ_r [Andrews, 1976]); σ_p, the stress necessary for premonitory phenomena to be produced; σ_i, the stress necessary for rupture initiation, and σ, the actual ambient stress (a function of both time and position). Note that the definition of σ_r is not precise since it depends on the nature of the rupture, - a rupture can propagate to some extent into a zone where the stress resists rupture propagation ($\sigma < \sigma_r$) but the actual extent depends on which direction the rupture comes from and how long of a rupture has preceded.

A Model for Non-Predictable Earthquakes

Earthquake prediction using premonitory phenomena will not be possible in the state defined by the following conditions:

1. At any point along the fault the material stress parameters are ordered as follows: (a) $\sigma_f \tilde{=} \sigma_r$; (b) $\sigma_r < \sigma_i$; (c) $\sigma_r < \sigma_p$; (d) σ_i can either be greater than or less than σ_p.
2. All but a small (perhaps zero) volume of the medium is under too low a stress to be associated with premonitory phenomena ($\sigma < \sigma_p$).
3. A certain section of the fault in this volume is under high enough strain so that a rupture can propagate through it releasing enough energy to cause an earthquake, i.e., the stress is sufficiently above the dynamic sliding friction, σ_f, to allow rupture and energy release if triggered from outside ($\sigma > \sigma_r$).

4. At some point a small volume, V_i, somewhere in the region is stressed high enough to initiate spontaneous slip ($\sigma = \sigma_i$) and is located where it can grow, by rupture propagation or triggering, and release the energy stored up on the larger fault area.

General Models

We can imagine that after a large earthquake has occurred and its aftershocks have subsided, the actual stress level is low and no section of the fault shows premonitory phenomena, i.e., $\sigma < \sigma_p$. As the general tectonic stress level increases, it may remain erratic along the fault, but at some points the rupture initiation stress is exceeded, causing small earthquakes, or creep, redistributing the stress, so that stress concentrations or stress spikes occur within a generally erratic stress pattern. As the general stress level increases, several interesting conditions can occur, illustrated by diagrams A through H, Figure 1.

A. Premonitory effects occur without an earthquake (false alarm). An earthquake might occur later as stress builds up or creep may relieve the stress so that no earthquake occurs.

B. A small earthquake occurs, preceded by premonitory effects.

C. A small earthquake occurs with no premonitory effects.

D. A large earthquake occurs along L triggered off by initial slip at D which was preceded by localized premonitory effects.

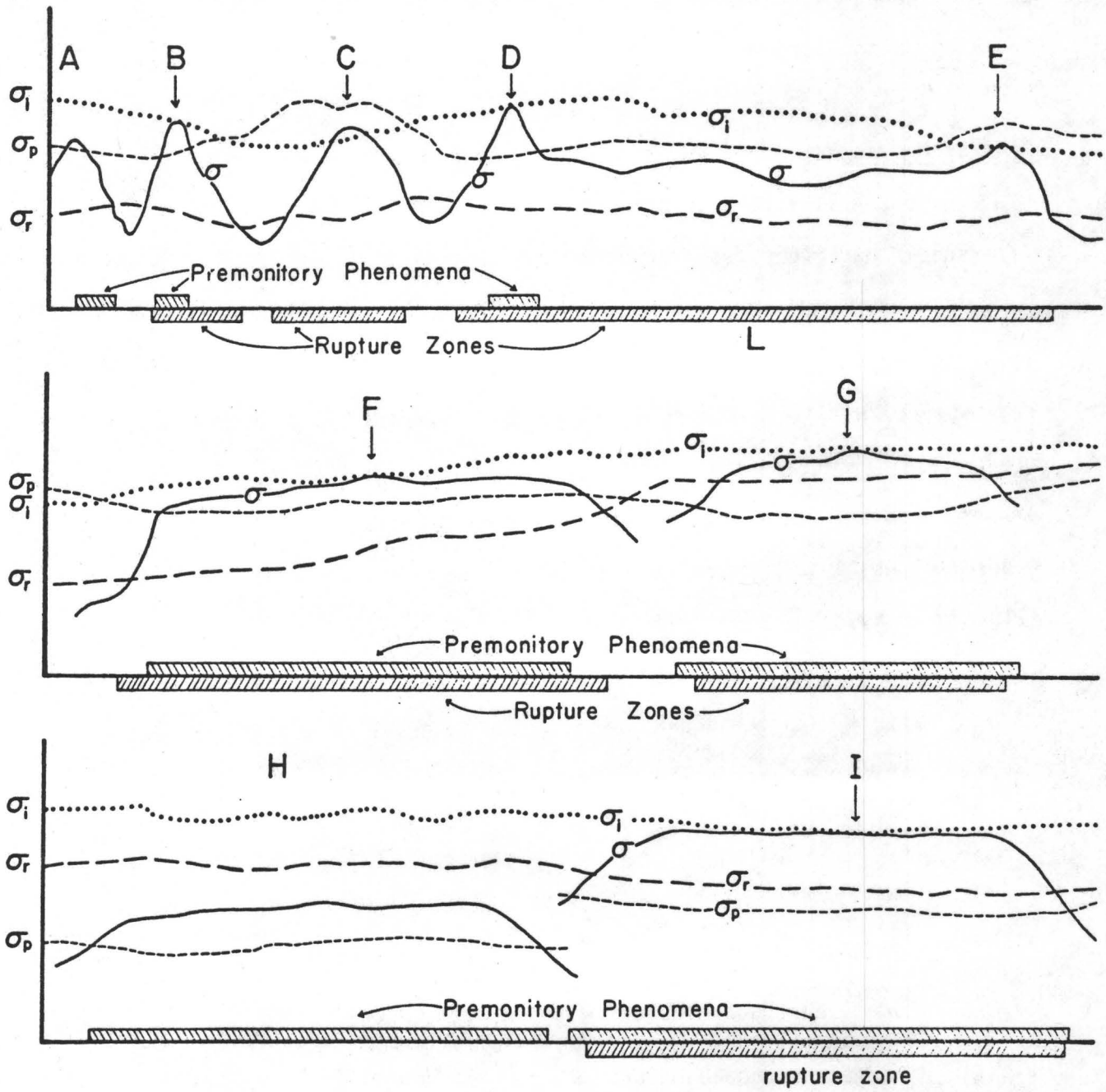

A
B
C
D
E
σ_i
σ_p
σ_r
σ
Premonitory Phenomena
Rupture Zones
L
F
G
H
I
rupture zone

Figure 1.

E. A large earthquake occurs along L triggered off by initial slip at E with no premonitory effects (the non-predictable earthquake described above).

F. - I. Potentially predictable earthquakes, preceded by widespread premonitory phenomena.

In case E prediction using premonitory effects is impossible because there are none. In case D prediction using premonitory effects is difficult because the zone of premonitory effects is very small and difficult to detect. It is also difficult to know if the event will grow or not, i.e., it is difficult to distinquish D from B, without knowing the details of the stress distribution around the zone of premonitory effects.

In case F, a predictable earthquake, rupture over an extensive region is initiated by rupture at F, and the prior stress level over a large region around F (approximately the same area as the rupture zone) is high enough to cause premonitory effects.

Although it seems probable that $\sigma_p > \sigma_r$, this is not known for sure and we should consider the opposite case. Earthquake prediction will be potentially effective if σ is relatively smoothly varying and $\sigma_r > \sigma_p$ at all points (G in Figure 1), i.e., if in order to reach the stress required for rupture propagation every section of the fault must first pass through σ_p and thus demonstrate observable premonitory phenomena. This situation does not eliminate the possibility of false alarms because a large region could reach σ_p without reaching σ_i or σ_r. A case where earthquake prediction without false alarms is possible is when σ is relatively smoothly varying and σ_i, σ_r and σ_p are all nearly equal, with σ_i infinitesimally higher than σ_r. However, in this case the lead time for

prediction will be very small. A more realistic case where prediction without false alarms would be possible is when $\sigma_p < \sigma_r$ and σ is relatively smoothly varying along the fault, and increasing at a known constant rate so that σ can be predicted to reach σ_i at a certain time after σ_p is observed to have been reached; thus the approximate length of rupture can also be estimated. Another model for which earthquake prediction might be successful is the dilatancy-fluid flow model, which has been widely discussed (Nur, 1972; Scholz *et al.*, 1973; Anderson and Whitcomb, 1973).

Reid's Model

The simplest model where prediction will work is when σ_i and σ_r are known at every point of the fault and σ is known and predictable, i.e., when tectonic strain build-up as well as rock properties are known. This corresponds to the method for earthquake prediction suggested by Reid (1910, 1969) as a result of his studies of the San Francisco earthquake of 1906. In this method, prediction is not dependent on observation of premonitory phenomena. The method will probably remain quite unreliable for the near future because the required information cannot be obtained with sufficient accuracy. However, it may ultimately be the only workable model of earthquake prediction if no reliable premonitory phenomena are established.

Can We Refute the Model for Non-Predictable Earthquakes?

The critical condition in the model for non-predictable earthquakes is that a large section of the fault can exist in a state of stress such that $\sigma_r < \sigma < \sigma_p$, i.e., the stress is greater than the stress necessary to

allow triggering or rupture propagation through the area, but less than the stress necessary to cause premonitory effects. In this state a large section of the fault can be triggered off by a small event occurring within the section or near its edge.

Unfortunately, neither σ_r nor σ_p is accurately known. Laboratory results suggest that dilatancy, one of the mechanisms associated with premonitory effects requires stress of the order of a kilobar or greater at depths corresponding to shallow earthquake strain energy release (5 - 15 km). On the other hand, σ_r might be identified with stress drop which is the order of a hundred bars, perhaps less than the stress required to cause dilatancy at depth. Because of these possibilities, it does not appear that we can reject the triggering-rupture propagation model without further study, and thus we have to accept the possibility that in many cases prediction of earthquakes may be difficult or impossible.

Two- and Three-Dimensional Rupture Zones

For simplicity, the foregoing description has been done in one dimension, along the strike of the fault zone. Of course, actual fault planes are approximately two-dimensional or in complex cases even three-dimensional. However, the basic reasoning remains the same. In the case of a simple planar fault we would describe the various stress levels by contour lines on the fault plane. In three dimensions, the stress distributions could be described by surfaces of constant stress.

Statistical Considerations

It will be noted that one of the critical differences between the case of a predictable earthquake and the case of a non-predictable one is the smoothness of the stress distributions. In case F the stress was relatively smoothly distributed so that when point F reached the stress required for rupture initiation it was surrounded by a large zone of relatively constant high stress, sufficient to cause observable premonitory phenomena, whereas in case E there was a sharp spike of stress sufficient to cause rupture initiation, but no contiguious large zone demonstrating premonitory phenomena. This suggests that the reliability of earthquake prediction might be estimated if the scale and amplitude (statistical properties) of stress variations were known.

Experimental Tests

Stress Values

The most crucial evidence necessary to test the implications of the above model is of course experimental values for σ_r, σ_p, σ_i and σ. The values for σ_r are not well known for real rocks. Although values of σ necessary to cause dilatancy, perhaps related to σ_p, are known from laboratory experiments to be some considerable fraction of the breaking strength (30-50%) it is not known for certain what values are appropriate to the low strain rates and other *in situ* conditions in the earth. Similarly, values of σ_i can be estimated from rupture strength of rocks in the laboratory, but it is not known whether these values are appropriate for large faults with highly developed gouge zones. Levels of the actual *in situ* stress σ are not known. Debate still continues between those who feel absolute shear stresses are comparable to stress drops (~ 100 bars) or more like fracture strengths in the laboratories (kilobars).

It will be of prime importance to estimate the above parameters in the next few years if we are to advance our understanding of earthquake mechanism. At present we can only speculate on their values.

Evidence from Earthquake Spectra and Time Functions

As mentioned above, studies of earthquake spectra and time functions have supported the idea that earthquakes are propagating ruptures and can be complex multiple ruptures in many cases. This suggests large variability for the various stress parameters discussed above. Similarly, evidence from the large variation in stress drops observed for small earthquakes suggests that stress parameters may be extremely complex in a given region. However, the evidence for variations in stress drop for small earthquakes is not as reliable as that from larger earthquakes in which actual fault offsets can be measured, and needs to be further studied. Kanamori and Anderson (1975) have suggested that the variation in stress drop for large earthquakes is considerably less (range 10-30 bars). This has led Kanamori (1977 - Oral communication at meeting of the John Muir Geophysical Society, Yosemite Valley, Ca.) to suggest that the large range in stress drops for small earthquakes as contrasted to the smaller range for large earthquakes, might be evidence for statistical variations in stress and strength, e.g., asperities, along the fault.

If small earthquakes can give reliable information about variations in stress along faults the information would be very useful in testing some of the ideas presented in this paper. Therefore, it seems important to carry out more careful experiments using large numbers of broad band digital recorders.

REFERENCES

Allen, C. R., A. Grantz, M. M. Clark, R. V. Sharp, T. G. Theodore, E. W. Wolf, J. N. Brune and M. Wyss, The Borrego Mountain, California, earthquake of April 9, 1968: A preliminary report, Bull. Seismo. Soc. Amer., 58, 1183-1186, 1968.

Allen, C. R., M. Wyss, J. N. Brune, A. Grantz and R. E. Wallace, Displacements on the Imperial, Superstition Hills, and San Andreas faults triggered by the Borrego Mountain earthquake, in U.S.G.S. Prof. Paper #787, 87-104, 1972.

Anderson, D. L. and J. H. Whitcomb, The dilatancy-diffusion model of earthquake prediction, in Proceedings of Conference on Tectonic Problems of the San Andreas Fault System, Geol. Sci. Ser., 13, Ed. by R. L. Kovach and A, Nur, 417-426, Stanford University, Stanford, California, 1973.

Andrews, D. J., Rupture propagation with finite stress in antiplane strain, J. Geophys. Res., 81, 3575, 1976.

Brune, J. N., Current status of understanding quasi-permanent fields associated with earthquakes, EØS, 55, No. 9, 1974.

Kanamori, H. and D. L. Anderson, Theoretical basis of some empirical relations in seismology, Bull. Seismo. Soc. Amer., 65, 1073-1095, 1975.

Nur, A., Dilatancy, pore fluids, and premonitory variations of t_s/t_p travel times, Bull. Seismo. Soc. Amer., 62, 1217-1222, 1972.

Reid, H. F., Mechanics of the earthquake, in The California Earthquake of April 18, 1906, Carnegie Inst. of Washington, Washington, D. C., 1969.

Scholz, C. H., L. R. Sykes and Y. P. Aggarwal, Earthquake prediction: A physical basis, Science, 181, 803-810, 1973.

Wyss, M. and J. N. Brune, The Alaska earthquake of 28 March 1964: A complex multiple rupture, Bull. Seismo. Soc. Amer., 57, 1017-1023, 1967.

FAR-FIELD S-WAVE SPECTRA, CORNER FREQUENCIES AND PULSE SHAPES

James N. Brune

Ralph J. Archuleta

Stephen Hartzell

ABSTRACT

Recent results concerning the relationship of S-wave far-field corner frequencies to stress drop are discussed. It is suggested that the corner frequency as picked by Madariaga (1977) from his theoretical spectra is not consistent with the way in which corner frequencies were picked by Tucker and Brune (1973, 1977). When the difference is taken into account, the stress drops inferred from the Madariaga relationship and his method of picking corner frequencies are roughly the same as inferred using the Brune (1971) relationship and the Tucker and Brune method of picking corner frequencies (for those relatively few experimental spectra which are similar in shape to the Madariaga theoretical spectra).

Far-field spectra for a number of new finite element models of fault ruptures in a half-space are presented, and for these data, new values for the relationship between source dimension and corner frequency are obtained. Fault models used include semi-circular faults with rupture initiation at the surface and at depth, and rectangular faults with unilateral and bilateral rupture propagation. The results indicate that there is a considerable variation in corner frequency with respect to type of rupture and position around the rupture. Because of the variation it is not possible to conclude, without more calculations, what the "best" average relationship between corner frequency and source dimension is; however, a value for K about 1/3 is reasonable where $K = f_c r/\beta$ (f_c = corner frequency, r = radius and β = shear wave velocity).

Dahlen (1974) speculated that the corner frequencies picked experimentally could be significantly altered by scattering. For the San Fernando aftershocks, it is possible to make a case that this is not so. Many of the San Fernando aftershocks show very simple pulse shapes,

with a pulse duration consistent with the spectral corner frequency and little later arriving energy - a direct indication that scattering is not radically affecting the results.

INTRODUCTION

Studies of far-field spectra and pulse shapes of earthquakes offer the possibility of estimating the level and variability of tectonic stress. This possibility is especially attractive if small earthquakes can be used, since they occur so frequently. However, to test the reliability and usefulness of such studies will require more experimental and theoretical work. We need reliable near-field recordings at a large number of observation points on the focal sphere of earthquakes to obtain accurate information about source parameters such as dimension, rupture velocity and stress drop. For comparison with theory, we need to examine more fault models to determine the effects of stress drop, friction, and rupture velocity on spectra and pulse shapes. In this paper we present some further discussion of the experimental digital data obtained from aftershocks of the San Fernando earthquake, and theoretical data obtained from finite-element numerical models of earthquake rupture.

SAN FERNANDO AFTERSHOCK DATA

The main experimental data discussed in this paper is the data obtained by Tucker and Brune (1973, 1977) from aftershocks of the San Fernando earthquake. In those studies high-dynamic range, broad-band digital data were used to determine near source earthquake spectra. These were then interpreted to infer seismic moment, source dimension and stress drop using formulas given by Brune (1970, 1971). The results suggested a wide range in stress drops, with an upper limit of about 200 bars, and many cases of stress drop less than 1 bar. A number of the larger events (M = 3.5 - 4) apparently had two corner frequencies, one between .1 and

1 Hz, below which the spectra were approximately constant and another between 3 and 10 Hz above which spectra were proportional to about ω^{-2} to ω^{-3}. It was suggested that for these larger events with two corner frequencies, a two stage rupture process may have occurred, an initial sharp high stress drop followed by a more slowly growing rupture. Similar results were found by Hartzell and Brune (1977) for earthquakes occurring in the Brawley swarm of January, 1975. We will not discuss these results further in this paper, but will use the experimental data to investigate the method of determining corner frequencies and the relationship of corner frequency to source dimension and hence stress drop. All the experimental seismograms and spectra are taken from Tucker and Brune (1973) using their numbering. The horizontal arrow under each seismogram indicates the time window for the Fourier transform used to obtain the spectra (Figures 2-7).

MADARIAGA'S THEORETICAL MODEL RESULTS

Madariaga (1976, 1977) used a numerical finite-difference method to compute the dislocation rate on a growing circular fault which stopped at a certain radius r (in a full space). With the dislocation rate and the far-field representation integral, Madariaga computed the far-field spectra at different azimuths. From these results he obtained an average spectrum at an angle of 60° from the normal to the fault (Figure 1). For this spectrum, he drew a low frequency asymptote and a higher frequency asymptote to determine the corner frequency f_c (indicated by M in Figure 1). From f_c he determined $K = f_c r/\beta = .21$, where r is the fault radius and β the shear wave velocity. This value of K is considerably lower than the value of K for the Brune model, K = .37. Since the formula for determining stress

drop from corner frequency depends on the 3rd power of corner frequency, Madariaga's relationship implies a stress drop 5.47 times higher than Brune's. This would mean that for the San Fernando aftershocks the estimated upper bound on stress drops would be closer to 1 kb than to 200 bars as inferred by Tucker and Brune. Since the values of stress drop determined from earthquakes have played a considerable role in recent discussions about earthquake source mechanisms, this uncertainty is important, and hence we wish to investigate further the interpretation of earthquake corner frequencies.

In a second paper, Madariaga (1977) also calculated a high frequency asymptote which if used to infer corner frequency, gave a value for K of .40 (indicated by B in Figure 1). If this value of K were used, nearly the same values of stress drop as found by Tucker and Brune would be obtained.

It appears that the method of picking corner frequencies and, in particular, the weight put on high frequencies, is important. This is particularly true in the case of numerical modelling since it is difficult and costly to adequately represent frequencies much above a factor of 2 or 3 times the corner frequency. This is why Madariaga had to use other techniques to estimate the higher frequency asymptote. At this point it is necessary to decide whether the method of picking corner frequencies used by Madariaga is consistent with that used by Tucker and Brune.

COMPARISON OF THE TUCKER AND BRUNE EXPERIMENTAL SPECTRA WITH THE MADARIAGA THEORETICAL SPECTRA

In order to compare the theoretical and experimental spectra, we have plotted the Madariaga spectrum and a number of Tucker and Brune spectra on the same scale (Figure 2). Comparison with the more than 100 spectra presented by Tucker and Brune shows that one of the characteristics of the Madariaga spectrum is a more flattened corner than observed for most of the Tucker and Brune spectra. Figure 2 was purposely selected to show examples of some of the Tucker and Brune spectra which have this character, i.e., a flattened corner or missing energy near the corner frequency (top three spectra in Figure 2). Most of the Tucker and Brune spectra have sharper corners, like the bottom two examples in Figure 2, and hence there is little uncertainty in determining a corner frequency. Because of the difference in shape between the Madariaga spectrum and most of the Tucker and Brune spectra, there is some doubt about the growing and stopping circular crack as a model for the San Fernando aftershocks.

The dashed lines in Figure 2 show the asymptotes used to infer corner frequencies by Tucker and Brune, and these can be compared with the asymptotes used by Madariaga to interpret the theoretical spectra as shown in the upper part of Figure 2. It is immediately obvious that the corner frequency picked by Madariaga, indicated by an M in the upper theoretical spectrum, is considerably lower than would have been picked by Tucker and Brune. Whereas, the corner frequency determined from the "high frequency asymptote" of Madariaga, indicated by B, corresponds nearly exactly with the corner frequency as determined by Tucker and Brune.

Since the Madariaga high frequency asymptote gives a corner frequency to fault radius relationship almost identical to that used by Tucker

and Brune (K = .40 vs. K = .37), it is evident that the interpretation of stress drops by Tucker and Brune is in accordance with the theoretical spectra of Madariaga. The further question of why the shape of most of the Tucker and Brune spectra are different from the shape of the Madariaga spectra is left unanswered. Some recent results pertinent to this are discussed in the next section.

NEW RESULTS FOR FAR-FIELD PULSE SHAPES AND SPECTRA BASED ON FINITE ELEMENT MODELS OF FAULT RUPTURE

Archuleta and Frazier (1977) have presented results for finite element models of faulting based on stress relaxation. The resulting time functions for fault slip at various points on the fault can be used to compute the far-field radiation. Half-space Green's functions, computed using a program of Johnson (1974), for each point on the fault-surface, are convolved with the slip functions on the fault and summed to give the complete far-field pulse (Hartzell *et al.*, 1978).

Preliminary results using this method were presented orally by Hartzell and Archuleta (1976). We have now completed calculations for far-field SH pulses observed on the surface of an elastic half-space for four different fault models:

1. A semi-circular fault in which rupture initiation begins at the surface (r = 0) and propagates radially to a specified fault radius, with 10 elements on a radius.
2. A semi-circular fault in which rupture initiation begins at the deepest point on the fault and symmetrically propagates over the semi-circle, breaking out at the surface, also with 10 elements on a radius.

3. A rectangular fault 3 elements deep and 15 elements long in which rupture propagation proceeds unilaterally from one end to the other.
4. A rectangular fault 5 elements deep and 20 elements long in which rupture proceeds bilaterally from the center to both ends.

A rupture velocity of 0.9 β is used for each of the four fault models above. Results for the far-field time functions (azimuthal component) are shown in Figures 8a and 8b and the corresponding spectra (S phase only) at the bottom of Figures 3 through 7 and in Figures 9 through 12. All of the numerical results shown here are for surface observations, with the azimuthal angle relative to the normal to the fault as a variable. Interpreted corner frequencies are indicated by the intersecting asymptotes drawn on the spectra.

For each fault model, the points of observation are at a constant radial distance, R, from the center of the fault. R is equal to 10 r_o for both semicircular faults, 5 ℓ_o for the bilateral rupture, and 6.6 ℓ_o for the unilateral rupture. Here r_o is the radius of the semicircular faults and ℓ_o is the length of the corresponding rectangular fault. The azimuth of the point of observation is measured from the normal to the fault. In the case of the rectangular unilateral rupture, the time function at 90° is in the direction of rupture propagation. The far-field time functions are plotted in dimensionless format (Figures 8a and 8b), where the dimensionless displacement is given by $u\ \mu\ 10^3/\sigma\ r_o$ for the semicircular faults and $u\ \mu\ 10^3/\sigma\ \ell_o$ for the rectangular faults. Here u is the displacement, μ the rigidity, and σ is the effective stress.

Because of the finite grid size, frequencies greater than 2Hz are

not accurately synthesized (Archuleta and Frazier, 1977), and this must be taken into account in interpreting corner frequencies and pulse shapes. Archuleta and Frazier (1977) have shown that their results have somewhat better frequency range than the results of Madariaga. In most cases the frequency range was sufficient to establish the true corner frequency with confidence. To assure objectivity in picking corner frequencies, both Brune and Archuleta independently picked the corner frequencies without knowledge of the other's picks. In nearly all cases the picked corner frequencies agreed within 20%. Since Brune was one of the investigators picking corner frequencies in the Tucker and Brune experimental study, there is some assurance that corner frequencies determined there were determined in the same manner as in this study. As a comparison with observed spectra, some of the results for the rectangular faults are shown along with representative spectra of San Fernando aftershocks, chosen to show similar features, in Figures 3 through 7.

Study of the time functions and spectra shows a range of pulse durations, and a corresponding range of corner frequencies (inverse proportionality) depending primarily on the azimuth between the direction of rupture propagation and the direction of observation. For bilateral ruptures, the pulse widths are narrower normal to the fault than in the direction of rupture; however, the rise times are shorter in the direction of rupture propagation than normal to the fault. The pulse width is controlled primarily by the difference in travel times from different points on the fault surface. The rise time is related to focussing caused by a propagating rupture. In the propagating stress relaxation models of Archuleta and Frazier (1977) the slip velocity increases in amplitude in the direction of propagation. Thus, in addition to the common Doppler effect due to a moving source, we have the effect of increasing slip

velocity in the direction of rupture. The combined effect is strong focussing of energy in the direction of rupture propagation. It is this combined effect which leads to the shorter rise times in the direction of rupture propagation. This focussing has also been observed in a laboratory model of propagating stress relaxation (Archuleta and Brune, 1975; Archuleta, 1976). For unilateral ruptures, both the pulse width and rise time decrease as one moves from normal to the fault to in line with the direction of rupture propagation. Short rise times and narrow pulse widths increase the high frequency content and yield higher corner frequencies and larger values of K. In the direction away from the rupture propagation, for a unilateral rupture, both the pulse width and rise time are greater, leading to lower values of K. In Figure 13 we show the value of $K = f_c D/\beta$ as a function of azimuth for our fault models and those of Savage and Madariaga. We have also re-interpreted the corner frequencies for the Madariaga spectra using the considerations outlined in the first part of this paper and these are shown as M in Figure 13. This figure can be used as an aid in interpreting experimental data.

Since, for a given fault, the corner frequency, and consequently the relationship between corner frequency and source dimension, are strong functions of azimuth, or position on the focal sphere, it is desirable to know the fault orientation and direction of rupture propagation before interpreting the spectra and time functions in terms of fault parameters such as fault dimension, moment and stress drop. However, in many cases, especially for small earthquakes, this is not possible, and hence it is helpful to have some approximate relationship between corner frequency and fault dimension. Besides Brune and Madariaga, Savage (1972, 1974), Randall (1973), and Dahlen (1974) have also obtained such relationships.

To summarize the results of various models, we have shown in Table I the values of K for $\theta = 60°$, or in some cases average values. The value of $\theta = 60°$ was selected by Madariaga because it represents the angle which divides the focal sphere roughly in half (equal areas). Thus, for numerous random observations around the fault, approximately half of the observations should have higher corner frequencies. The value of K for $\theta = 60^{\circ}$ should not be equated with the corner frequency averaged over the focal sphere nor with the value of K obtained by Brune (1970, 1971), which corresponds to the corner frequency for a spectrum which, if constant over the focal sphere, would, at high frequencies, give the same total energy radiation as the actual variable spectrum. However, the values of K determined in the various ways should be roughly comparable. The values for K in Table I range from .13 to .49. The low value of .13 corresponds to a bilateral rupture, and the corresponding spectrum (Figure 6, bottom) has a broad intermediate slope proportional approximately to ω^{-1}. At other azimuths the bilateral rupture produces values of K near .3, Figure 13. This spectrum is quite different from most of the other spectra and from most of the spectra observed by Tucker and Brune. However, some of the Tucker and Brune spectra did show such a broad ω^{-1} region and as would be expected gave low stress drops when interpreted using a K value of .37. Considering the rest of the results, this low value can be considered to a certain extent anomalous.

Although the results of Table I are derived from a variety of different earthquake models, an average value of K for ruptures with $v = .9\beta$ is .32. If the models of Brune (1970,1971) and Randall (1973) ($v = \infty$) are included, then K becomes .33. In veiw of the many uncertainties in relating theoretical models to actual earthquakes, a reasonable average value of K is 1/3. It

TABLE I

$K = f_c/(\beta/D)$

FAR-FIELD S-WAVE $\Theta = 60°$ $v = .9\beta$

Semicircular Faults in a Halfspace		$D = r_o$
This Study		
origin (0, 0, 0)	$K = .26$	
origin $(0, r_o, 0)$	$K = .29$	
Rectangular Faults		$D = W$
This Study: Halfspace		
L = 4W bilateral	$K = .13$	
L = 5W unilateral	$K = .32$	
SAVAGE (HASKELL) Full Space		
$L \simeq W$ bilateral	$K = .32$	
$L >> W$ bilateral	$K = .29$	
Circular Faults in a Full Space		$D = r_o$
MADARIAGA	$K = .21$	
high frequency asymtote	$K = .40$	
DAHLEN	$K = .45$	
BRUNE ($v = \infty$)	$K = .37$	Average Values
RANDALL; ARCHAMBEAU ($v = \infty$)	$K = .42$	Average Values
Dislocation Models: MOLNAR, TUCKER, BRUNE ($\Theta = 55°$)	$K = .32$	
Dislocation Models: SAVAGE	$K = .49$ (.1)	

must be noted that for a given station recording an earthquake, K could possibly range between .15 and .5, i.e., K = 1/3 $\pm$ (.5) (1/3). A 50% error in K corresponds to an error in inferred stress drops of a factor of 3.3. This factor is consistent with the variation in stress drops computed by Tucker and Brune between two different stations for the San Fernando aftershocks.

SCATTERING AND CORNER FREQUENCY

Dahlen (1974), observing that the ratio of corner frequencies of P-waves and S-waves observed by Molnar *et al.* (1973) was inconsistent with a source mechanism theory he had developed, suggested that the observed corner frequencies may have been seriously perturbed by scattering. A small amount of scattered energy arriving slightly after the S-wave would not cause a frequency shift in the spectrum, since the reflected pulses would have nearly the same spectrum as the direct energy. The existence of a certain amount of scattering is one of the main reasons for using the spectrum rather than the direct time function - the time function can be seriously distorted by scattering without seriously affecting the amplitude spectrum (which is based on the modulus of the Fourier transform and not the phase). In order for scattering to seriously affect the corner frequencies, a significant fraction of the energy would have to be scattered to travel paths that cause the scattered energy to arrive considerably later than the main energy. This would imply that the S-wave pulse would be considerably distorted or complicated. Thus, in a qualitative way we can estimate the importance of scattering by looking at pulse shapes. A large number of seismograms of aftershocks of the San Fernando

earthquake published by Tucker and Brune (1973) can be used for this purpose.

Perusal of the seismograms of the San Fernando aftershocks immediately reveals a wide range of apparent pulse shapes, ranging from simple pulses of about 0.1 sec duration, with little scattered energy, to very complex looking signals of a second or more duration. Figure 14 presents a selection of seismograms with simple pulse shapes. A curious thing is observed by comparing these seismograms with the more complex seismograms (e.g., 87 and 57 in Figure 2; 138 in Figure 3; 111 and 112 in Figure 5; 129 and 130 in Figure 6; and 53 and 57 in Figure 7). Seismograms with simple S-waves generally have low P-wave amplitudes relative to the S-wave, while the seismograms with complex-looking S-waves generally have high P-wave amplitudes relative to the S-waves (all the seismograms presented by Tucker and Brune were normalized to the same peak amplitude for plotting purposes; they actually represent a wide range in magnitude).

The obvious explanation for this phenomenon is that it is an effect of radiation pattern (Tucker and Brune, 1977). Near a node for S-waves there will tend to be an antinode for P-waves (relatively high P-wave amplitudes) whereas S-waves will appear complex because scattered energy will be large relative to the main S-wave energy. On the other hand, near an antinode for S-waves, the P-wave will tend to be small and the S-wave will appear simple because it will stand out from the smaller scattered energy. If we accept this explanation, we can estimate of the amount of scattering by observing the time functions of the simpler events. Figure 15 shows two simple time functions constructed as examples of predicted S-wave pulses when no scattering is present. These are compared with two of the simple observed pulse shapes. The theoretical

pulses were diagramatically constructed so as to correspond to a simple smooth spectrum with the same corner frequency as determined experimentally, with a time function similar to the theoretical pulse used by Brune (1970), as seen through a simple velocity transducer of the type used by Tucker and Brune (1977). For the purposes of determining the approximate amount of scattered energy, it is not important whether or not this pulse is exactly correct; we are interested in the amount of late arriving scattered energy relative to the main pulse. As can be seen from the seismograms in this figure, when the S-pulse is simple, the scattered late arriving energy is quite small relative to the main energy, probably not enough to seriously affect the inferred corner frequencies. The effect of scattering could be more serious for the more complex looking seismograms. However, study of the results of Tucker and Brune does not indicate any obvious correlation of stress drop with complexity, suggesting that, as expected, the scattered energy has about the same spectrum as the direct energy. Further evidence of this is the fact that in a number of trials to determine the effect of record length on the shape of the spectrum, Tucker and Brune found very little effect, i.e., adding more or less of the S-wave coda, presumably scattered energy, had little effect on the specturm shape. We conclude that at least for the San Fernando aftershocks studied by Tucker and Brune, scattering did not seriously affect the observed spectra. Other effects such as variation in source parameters, source complexity, and direction of rupture propagation were probably more important in causing the variations in spectral shape (and inferred stress drop) observed by Tucker and Brune.

REFERENCES

Archuleta, R. J., and J. N. Brune, Surface strong motion associated with a stick-slip event in a foam rubber model of earthquakes, Bull. Seismo. Soc. Amer., 65, 1059-1071, 1975.

Archuleta, R. J. , Experimental and Numerical Three-Dimensional Simulations of Strike-Slip Earthquakes, Ph.D. Dissertation, University of California, San Diego, 1976.

Archuleta, R. J., and G. A. Frazier, Three-dimensional numerical simulations of dynamic faulting in a half-space, submitted to Bull. Seismo. Soc. Amer., 1977.

Brune, J. N., Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res., 75, 4997-5009, 1970.

Brune, J. N., Tectonic stress and the spectra of seismic shear waves from earthquakes (Correction), J. Geophys. Res., 76, 5002, 1971.

Brune, J. N. and R. J. Archuleta, Near-field and far-field spectra for dynamic 3-D finite element fault models, Abstract, Spring, 1977 Meeting of American Geophysical Union, EØS, 1977.

Dahlen, F. A., On the ratio of P-wave to S-wave corner frequencies for shallow earthquake sources, Bull. Seismo. Soc. Amer., 64, 1159-1180, 1974.

Hartzell, S. H., and R. J. Archuleta, Study of relationships between effective stress, fault size and corner frequency derived from numerical simulation of an earthquake in a half space, Abstract, Seismo. Soc. Amer., 1976.

Hartzell, S. H., and J. N. Brune, Source parameters for the January, 1975 Brawley - Imperial Valley earthquake swarm, PAGEOPH, 115, 1/2, 1977.

Hartzell, S. H., G. A. Frazier and J. N. Brune, Earthquake modelling in a homogeneous half-space, submitted to Bull. Seismo. Soc. Amer., 1978.

Johnson, L. R., Green's function for Lamb's problem, Geophys. J. R. astr. Soc., 37, 99-131, 1974.

Kostrov, B. V., Self-similar problems of propagation of shear cracks, J. Appl. Math. Mech., 28, 1077-1087, 1964.

Kostrov, B. V., Crack propagation of variable velocity, J. Appl. Math. Mech., 38, 511-519, 1974.

Madariaga, R., Dynamics of an expanding circular fault, Bull. Seismo. Soc. Amer., 66, 639-666, 1976.

Madariaga, R., Implications of stress-drop models of earthquakes for the inversion of stress drop from seismic observations, Pure and Appl. Geophys., 115, No. 112, 301-316, 1977.

Molnar, P., B. E. Tucker and J. N. Brune, Corner frequencies of P and S waves and models of earthquake sources, Bull. Seismo. Soc. Amer., 63, 2091-2105, 1973.

Randall, M. J., The spectral theory of seismic sources, Bull. Seismo. Soc. Amer., 63, 1133-1144, 1973.

Savage, J. C., Relation of corner frequency to fault dimensions, J. Geophys. Res., 77, 3788-3795, 1972.

Savage, J. C., Relation between P- and S-wave corner frequencies in the seismic spectrum, Bull. Seismo. Soc. Amer., 64, 1621-1627, 1974.

Tucker, B. E. and J. N. Brune, Seismograms, S-wave spectra and source parameters for aftershocks of the San Fernando earthquake of February 9, 1971, NOAA Special Report, 1973.

Tucker, B. E. and J. N. Brune, Source mechanism and m_b - M_s analysis of aftershocks of the San Fernando earthquake, Geophys. J. R. astr. Soc., 49, 371-426, 1977.

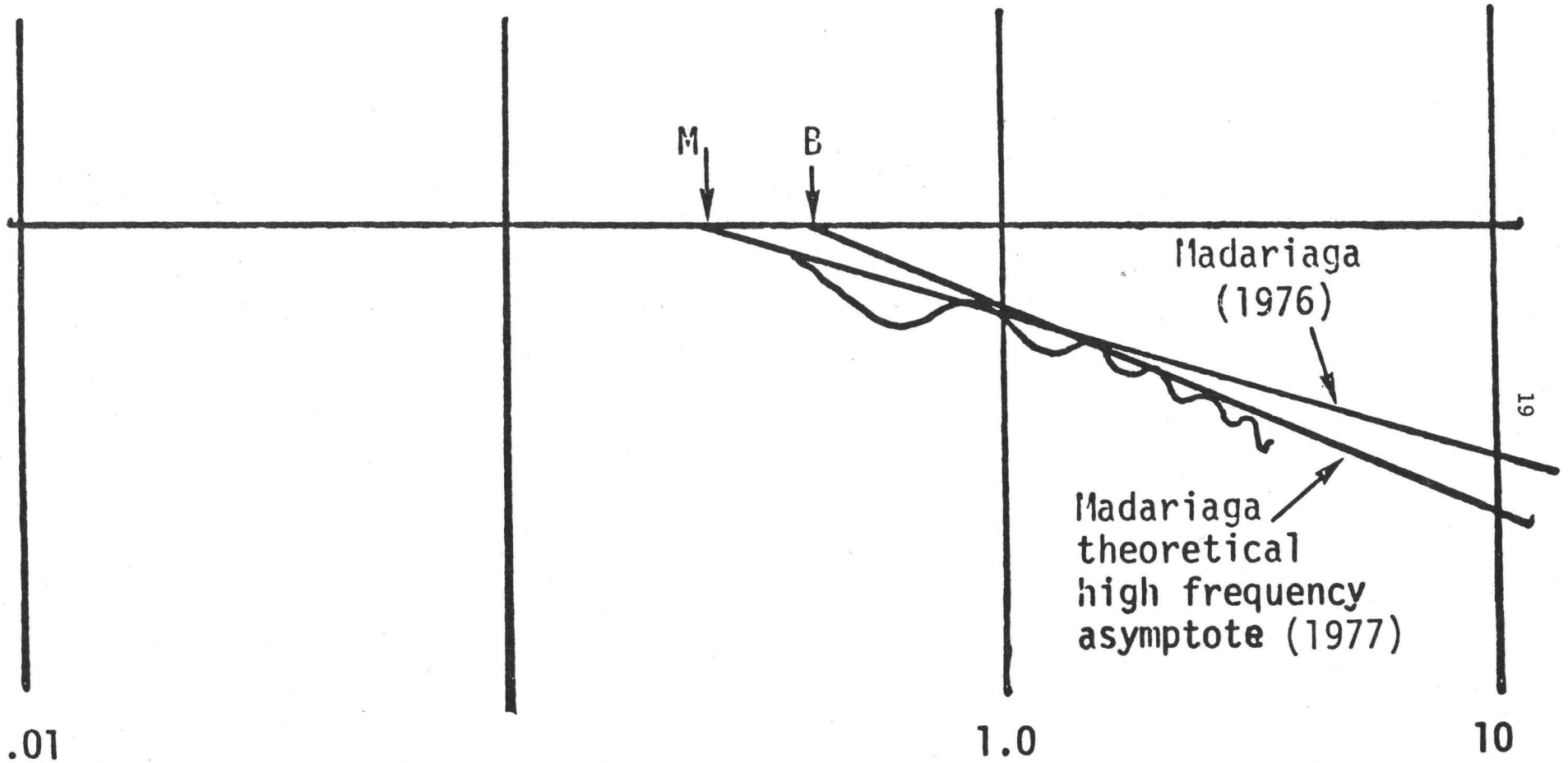

M
B
Madariaga
(1976)
Madariaga
theoretical
high frequency
asymptote (1977)
.01
1.0
10
K = $f_c/(B/D)$

Figure 1

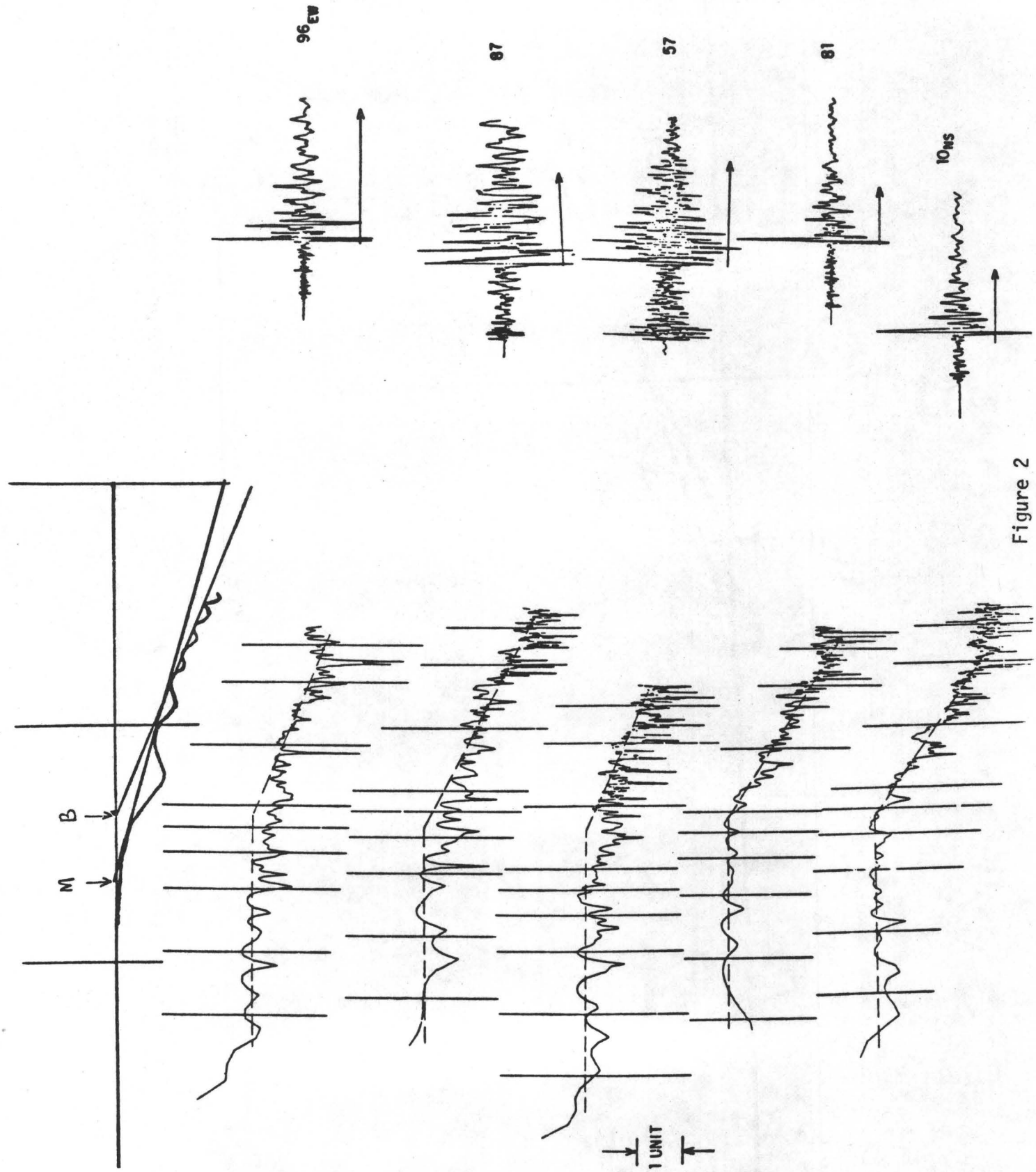
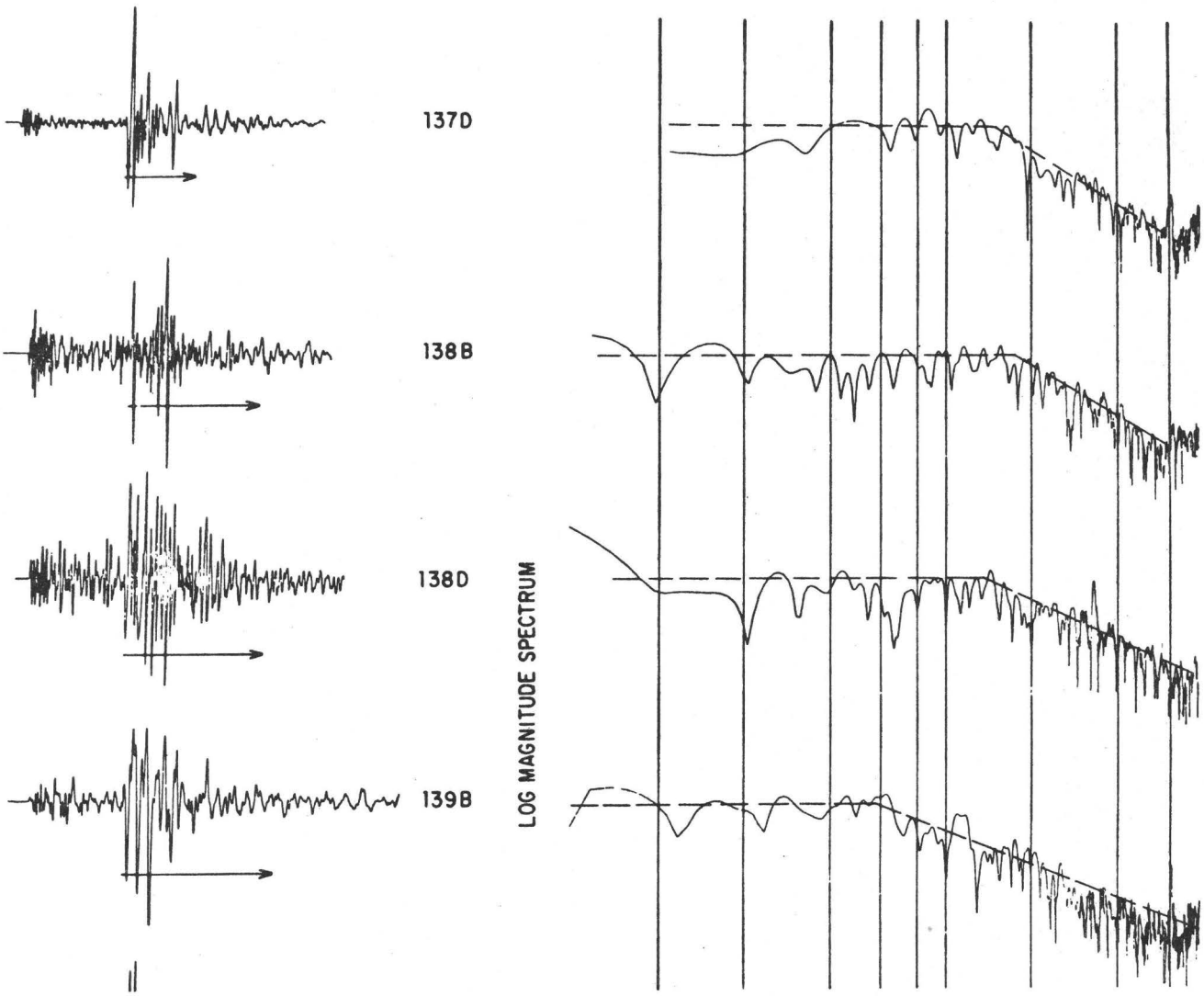


Figure 2

Comparison with San Fernando aftershocks



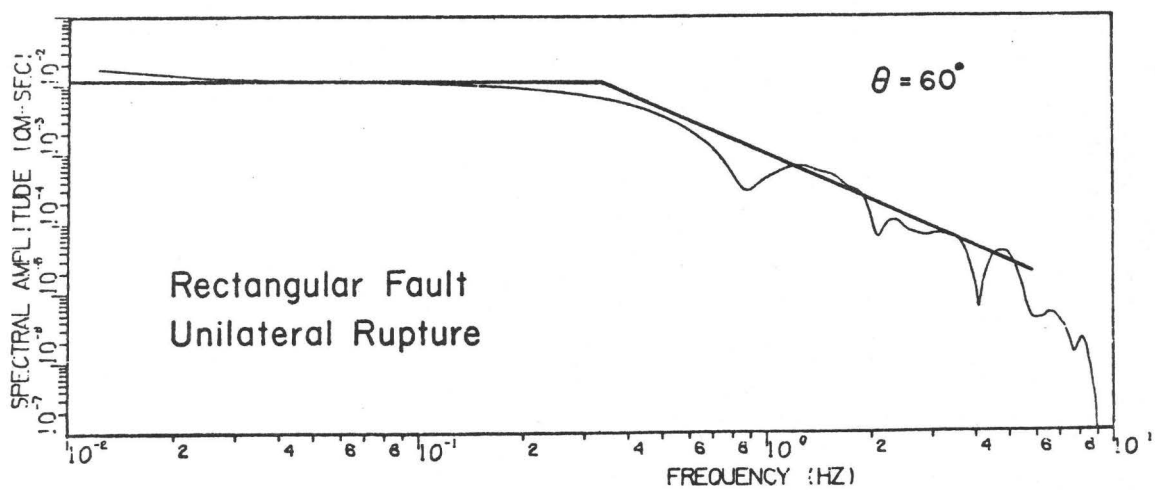
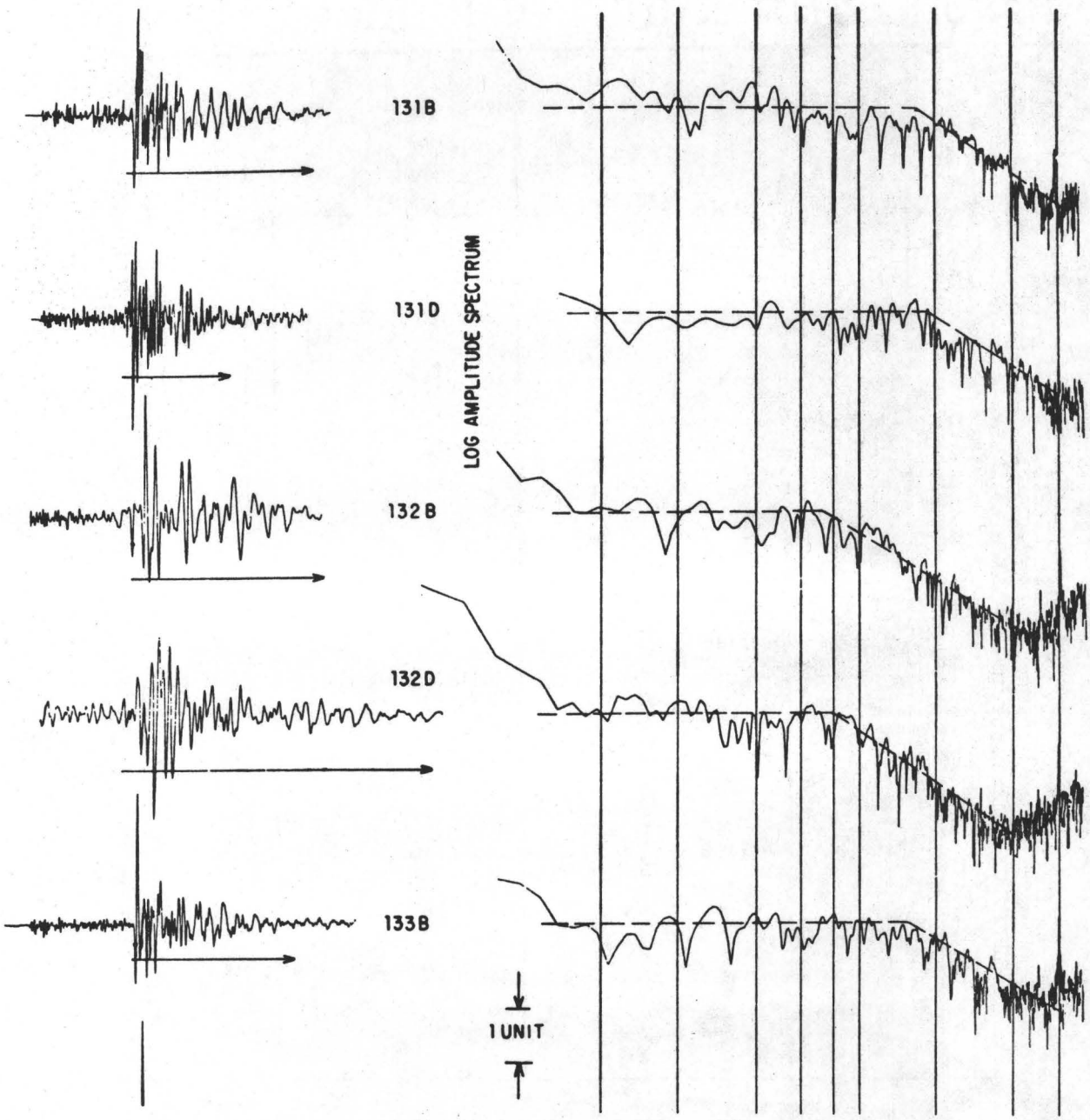


Figure 3

Comparison with San Fernando aftershocks



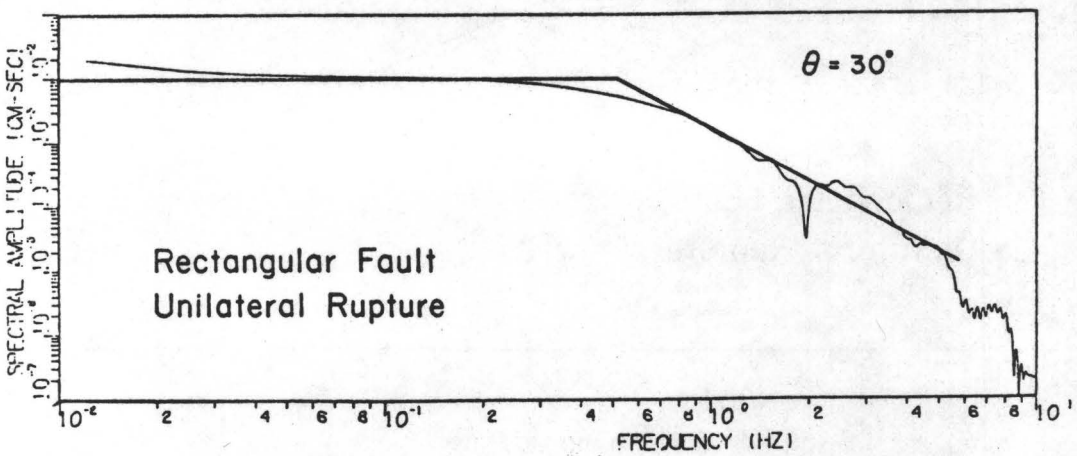
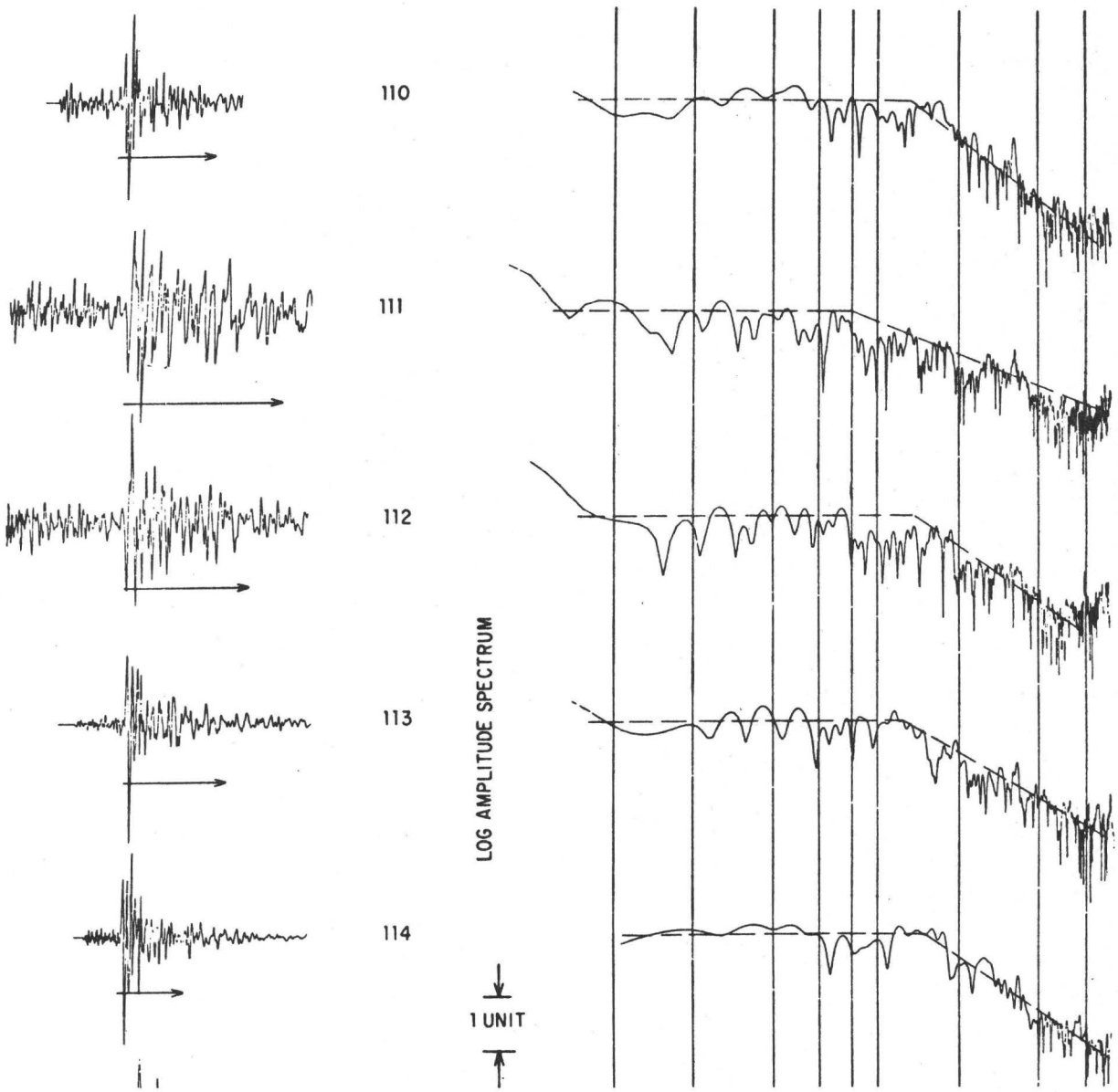


Figure 4

Comparison with San Fernando aftershocks



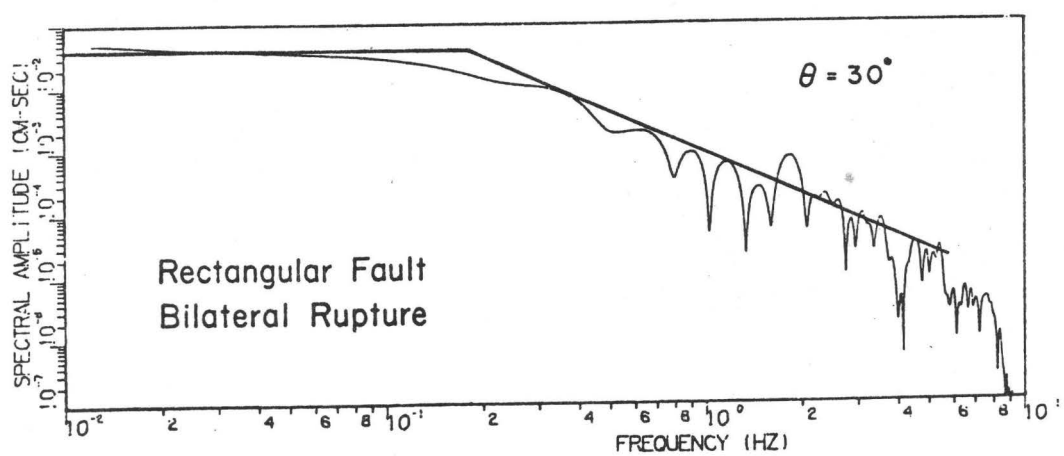
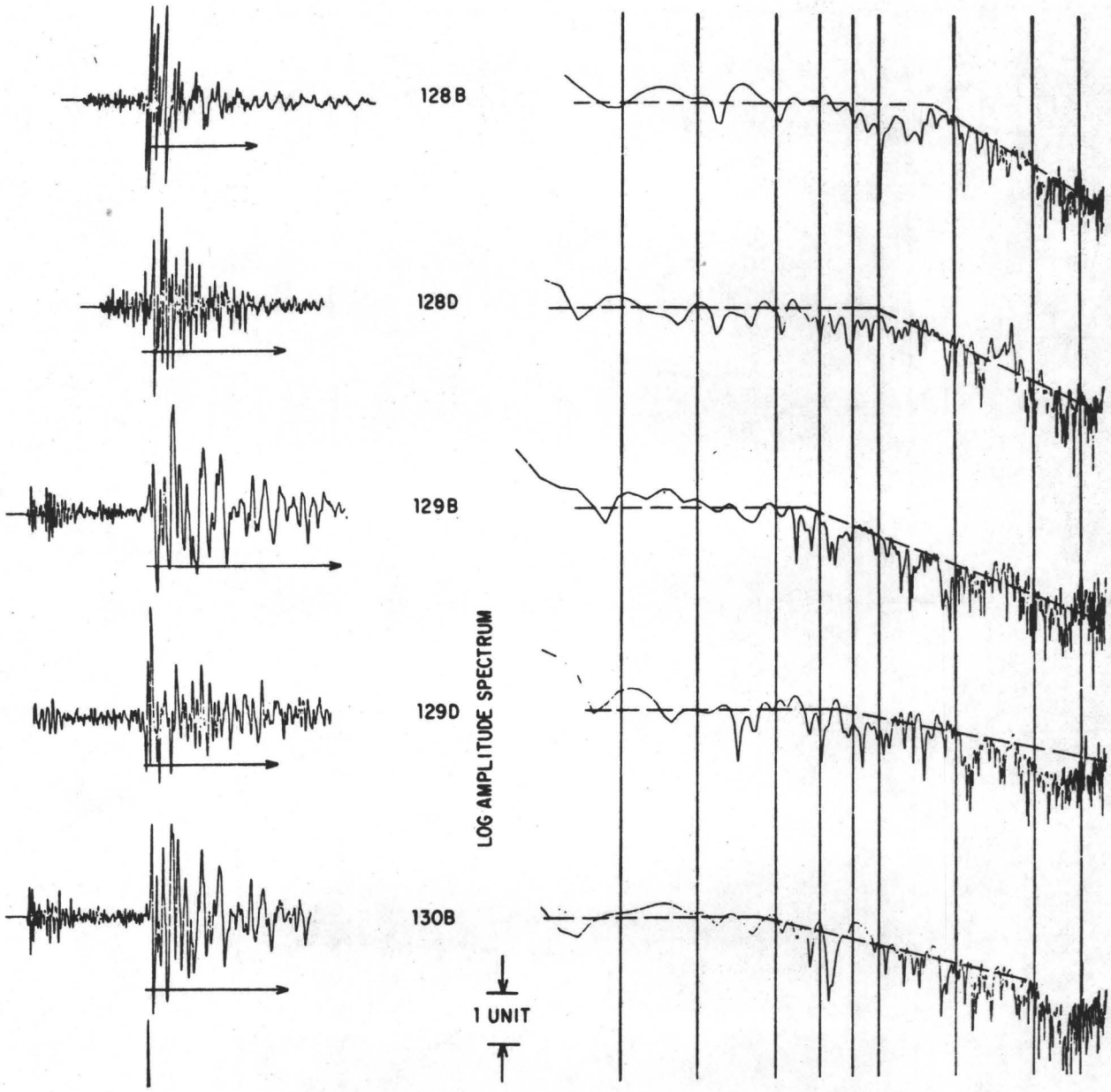


Figure 5

Comparison with San Fernando aftershocks



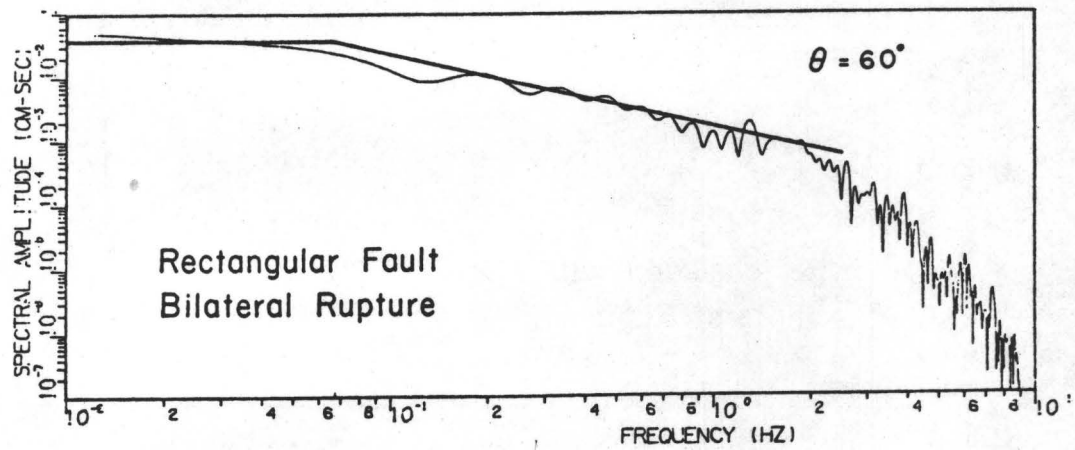
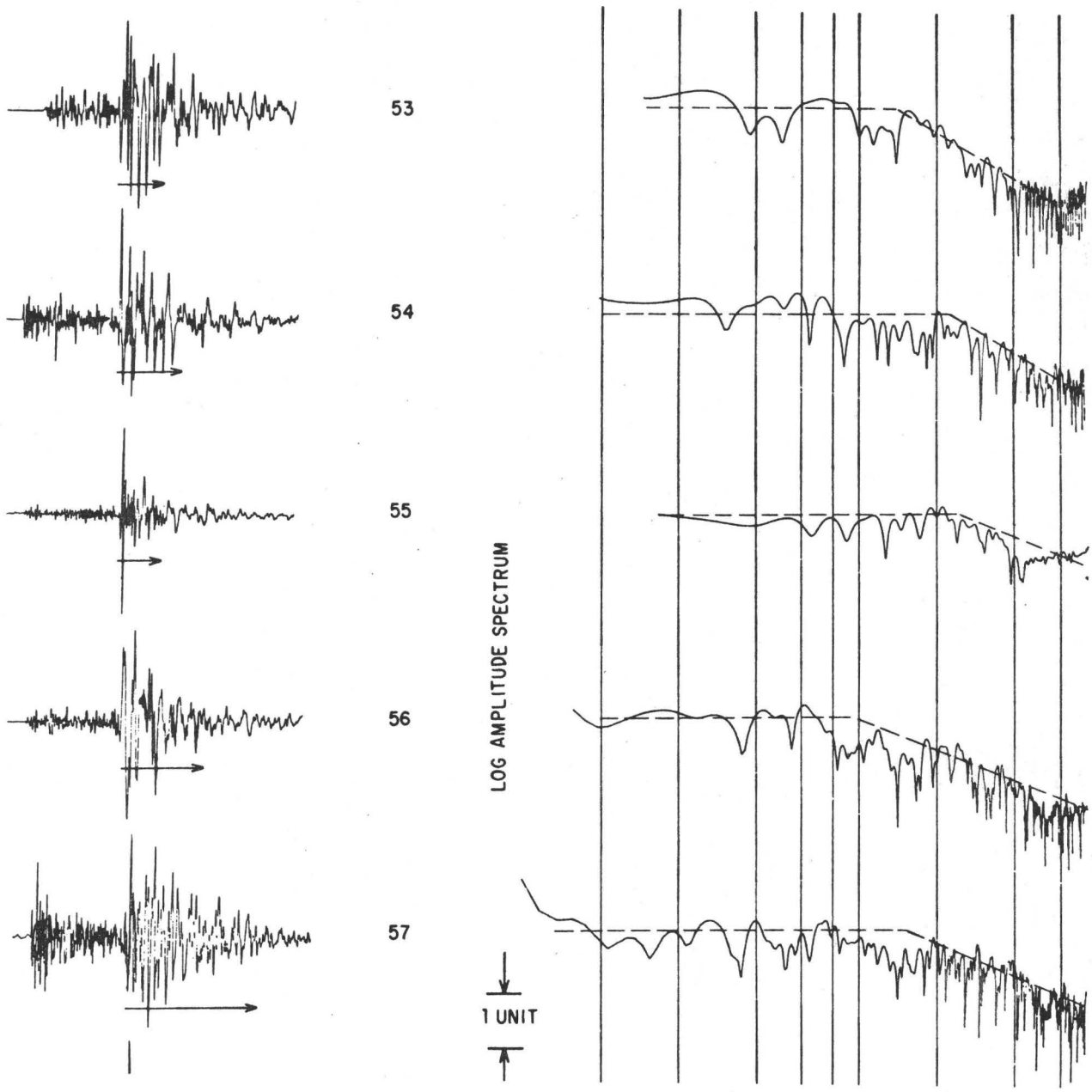


Figure 6

Comparison with San Fernando aftershocks



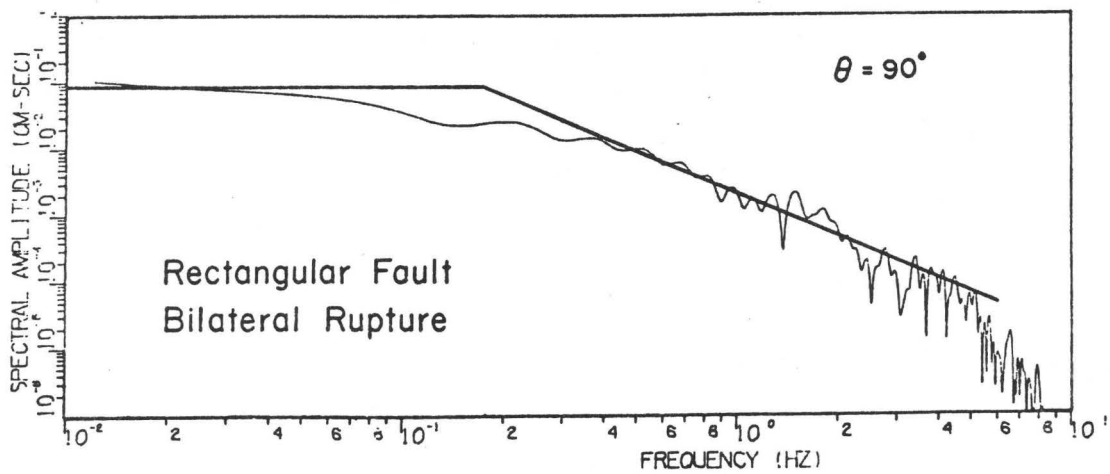


Figure 7

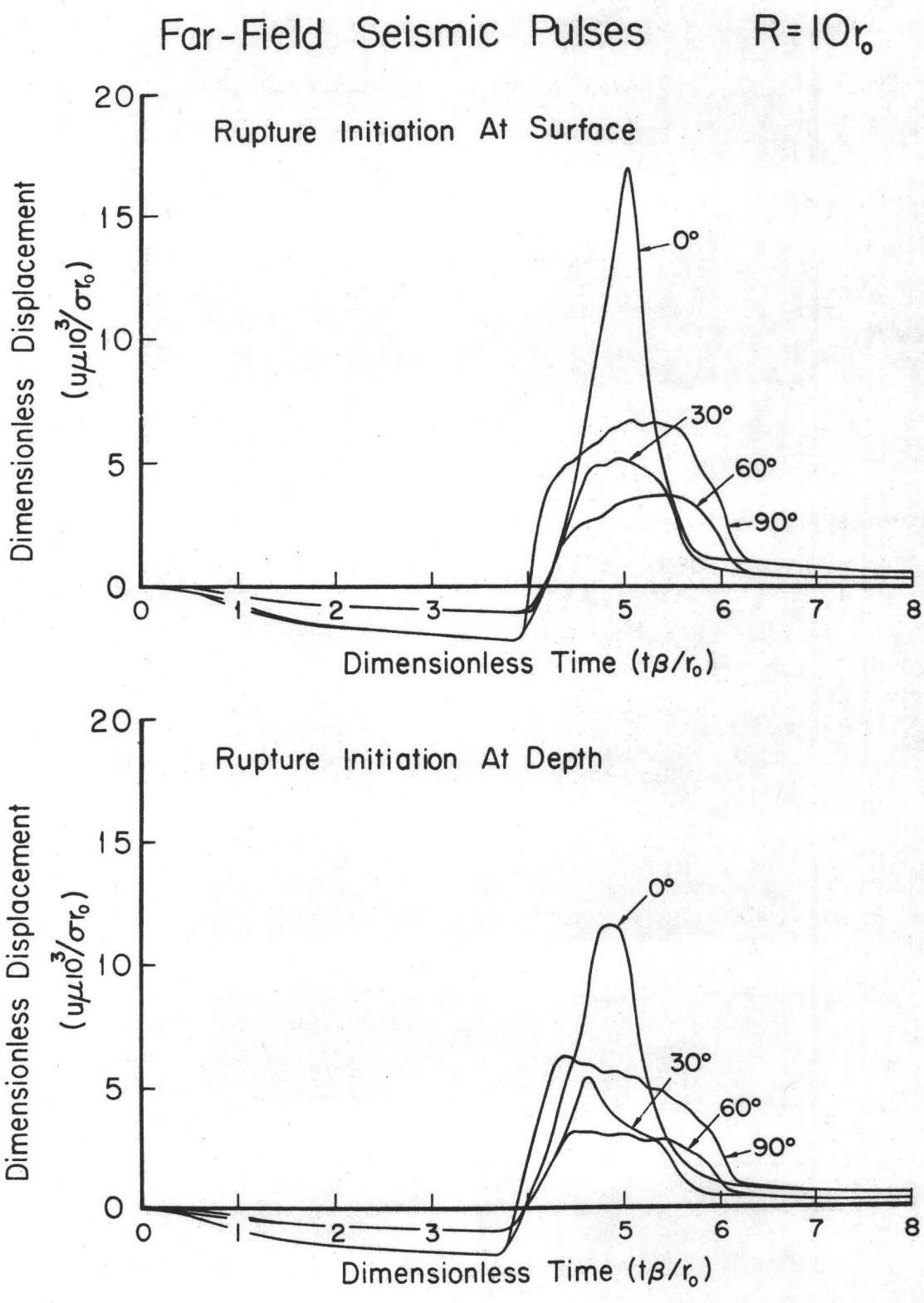

Far-Field Seismic Pulses R=10r0
Rupture Initiation At Surface
Dimensionless Displacement (uμ10^3/σr0)
Dimensionless Time (tβ/r0)
0°
30°
60°
90°
Rupture Initiation At Depth
Dimensionless Displacement (uμ10^3/σr0)
Dimensionless Time (tβ/r0)
0°
30°
60°
90°

Figure 8a

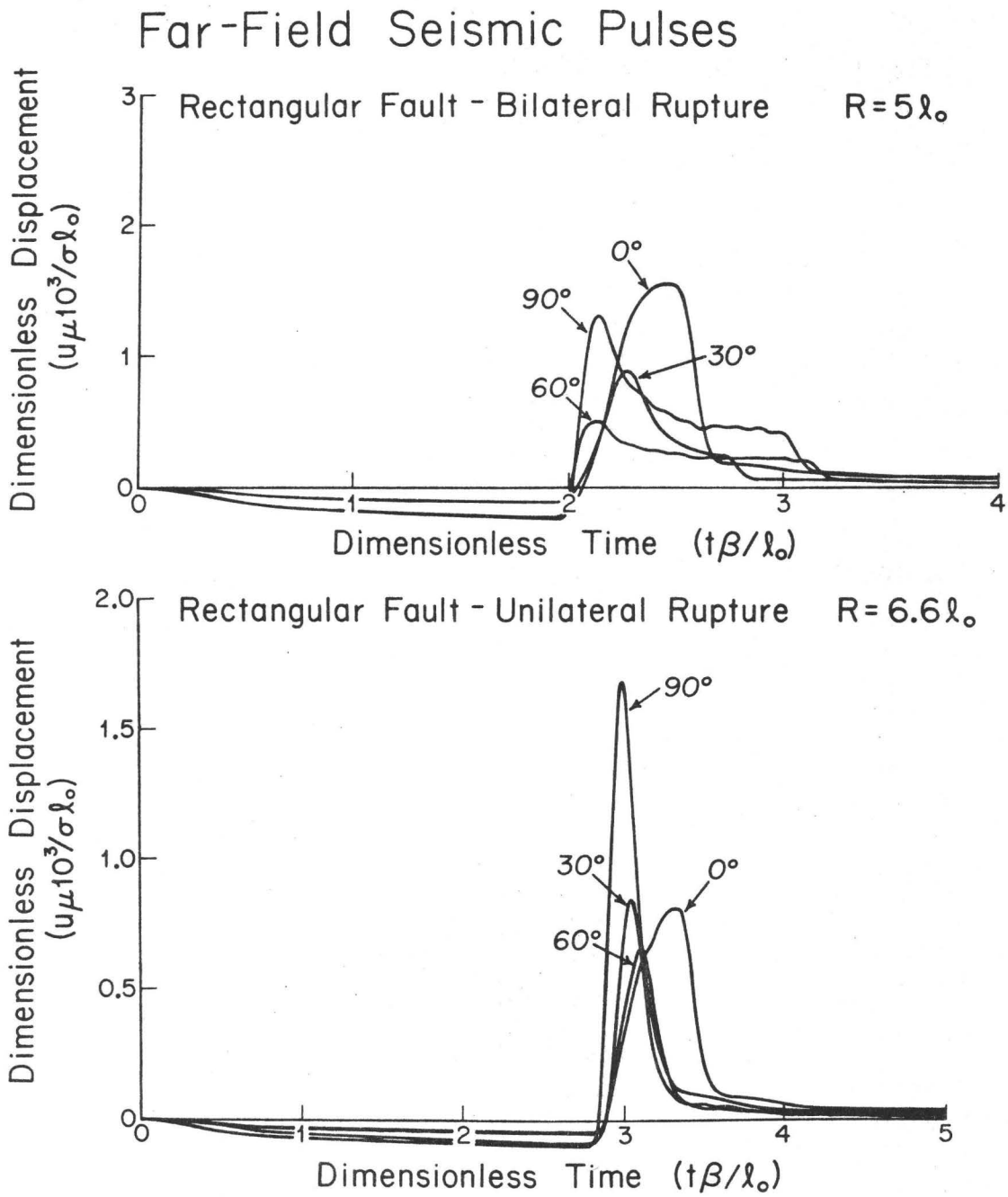

Far-Field Seismic Pulses
Rectangular Fault - Bilateral Rupture R=5ℓo
Dimensionless Displacement (uμ10³/σℓo)
0°
90°
30°
60°
Dimensionless Time (tβ/ℓo)
Rectangular Fault - Unilateral Rupture R=6.6ℓo
Dimensionless Displacement (uμ10³/σℓo)
90°
30°
0°
60°
Dimensionless Time (tβ/ℓo)

Figure 8b

Circular Fault
Rupture Initiation at the Surface

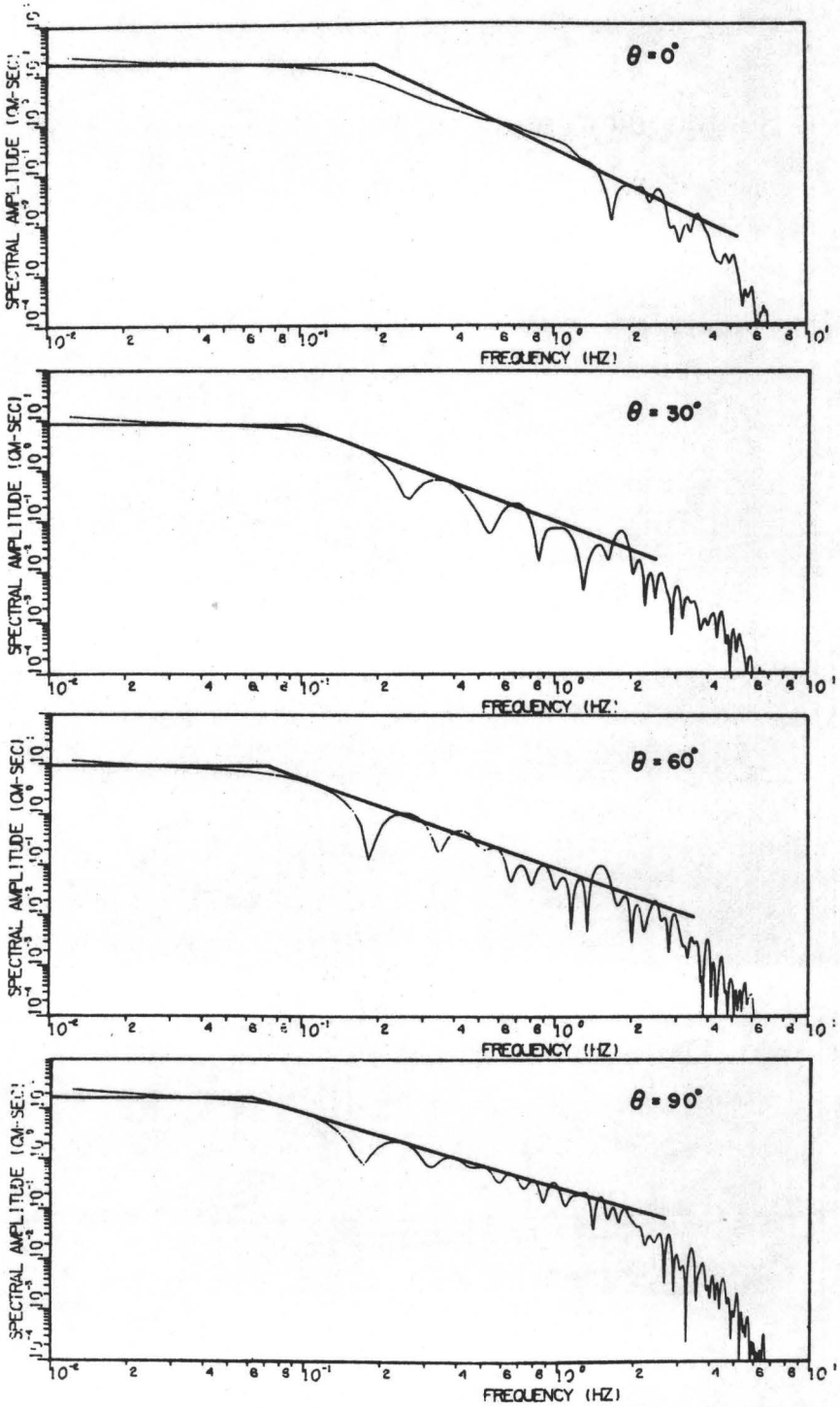


Figure 9

Circular Fault
Rupture Initiation at Depth

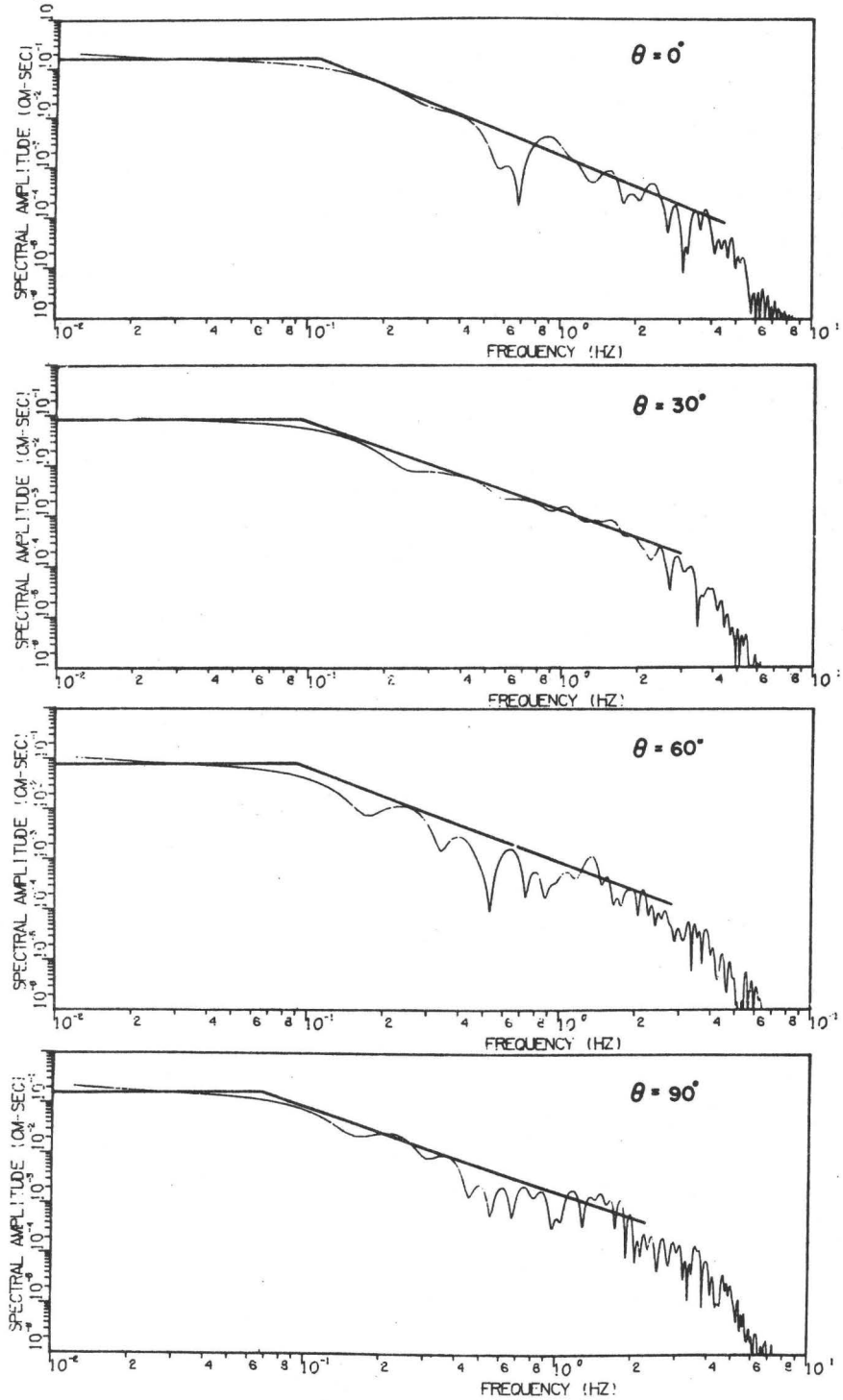


Figure 10

Rectangular Fault
Bilateral Rupture

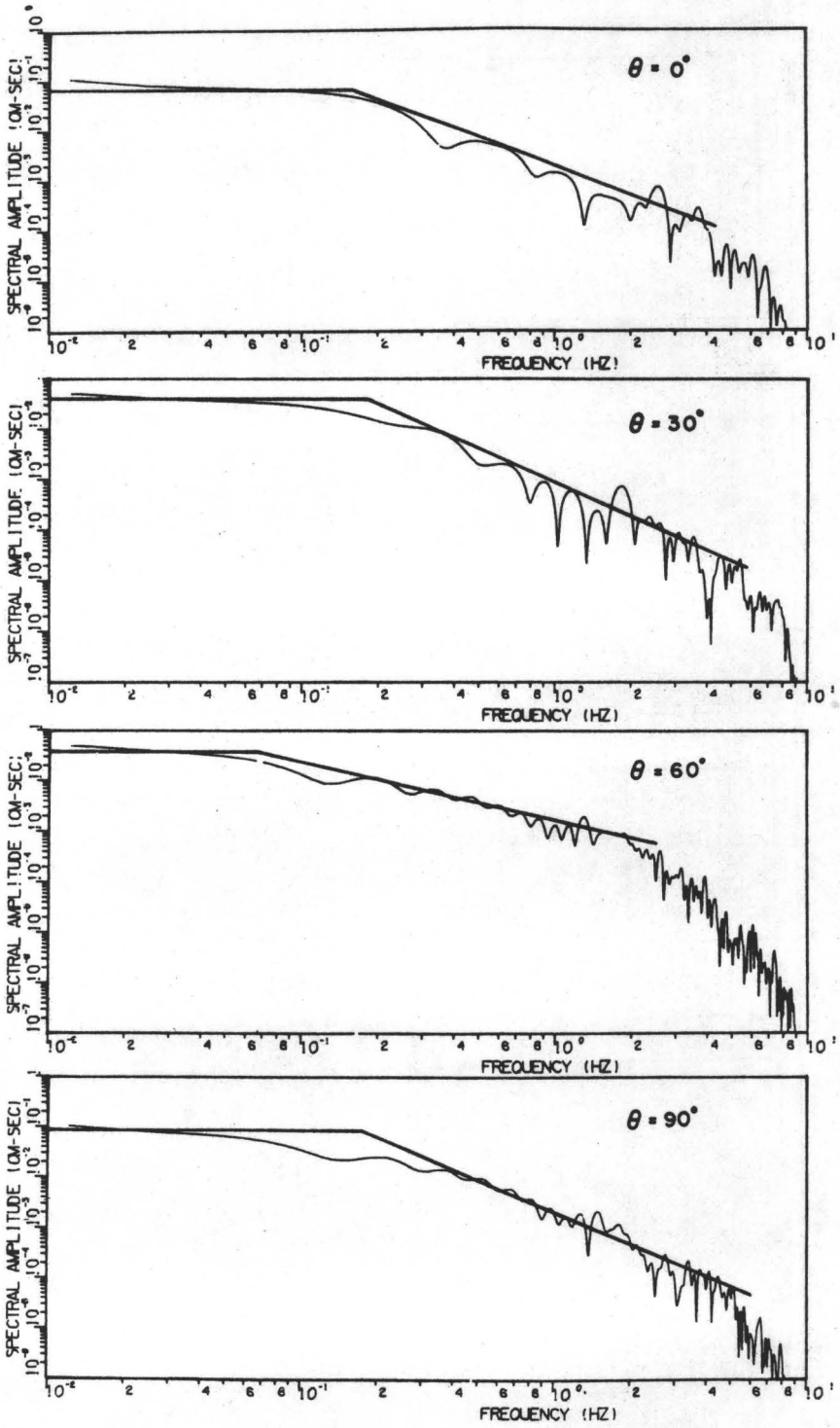


Figure 11

Rectangular Fault
Unilateral Rupture

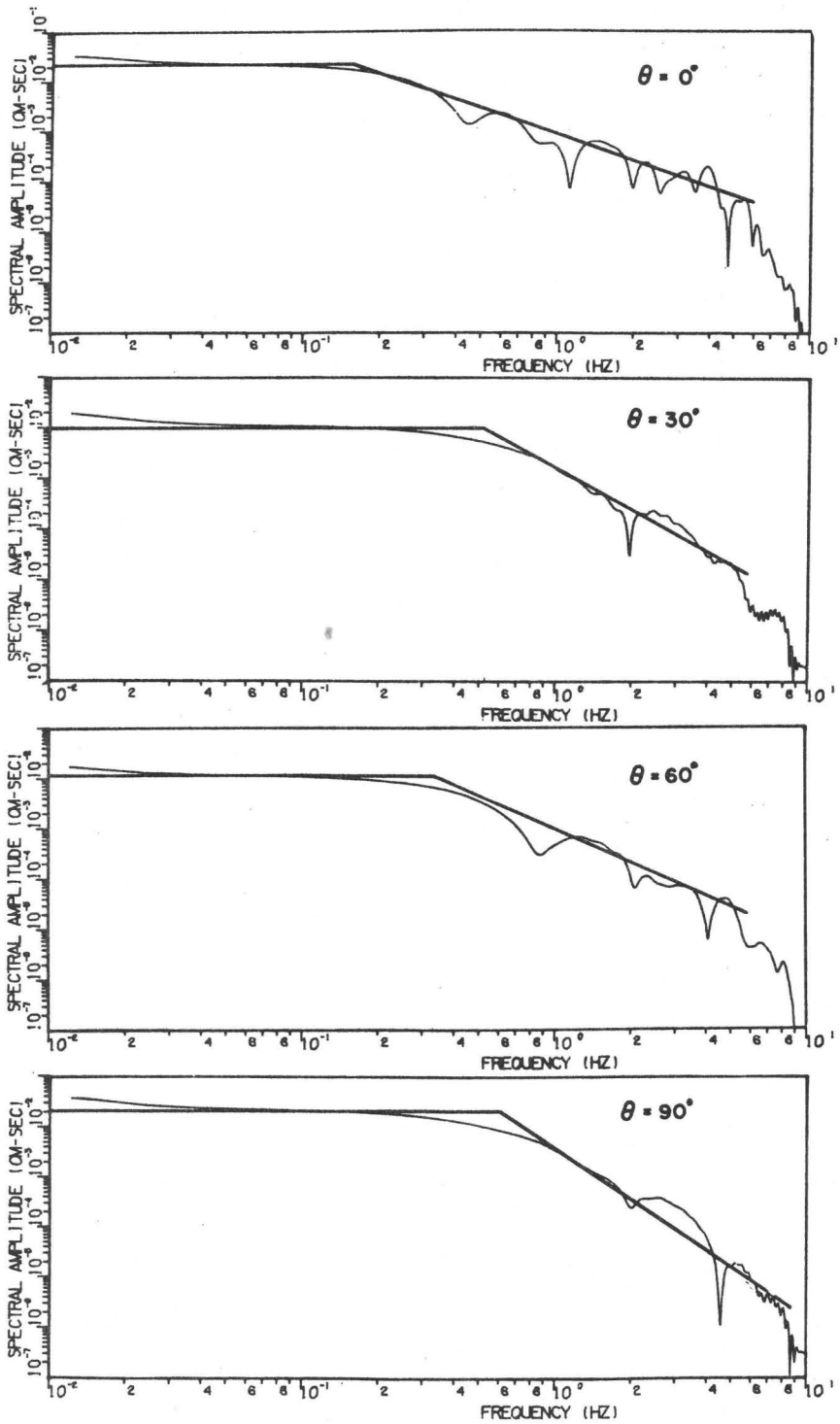


Figure 12

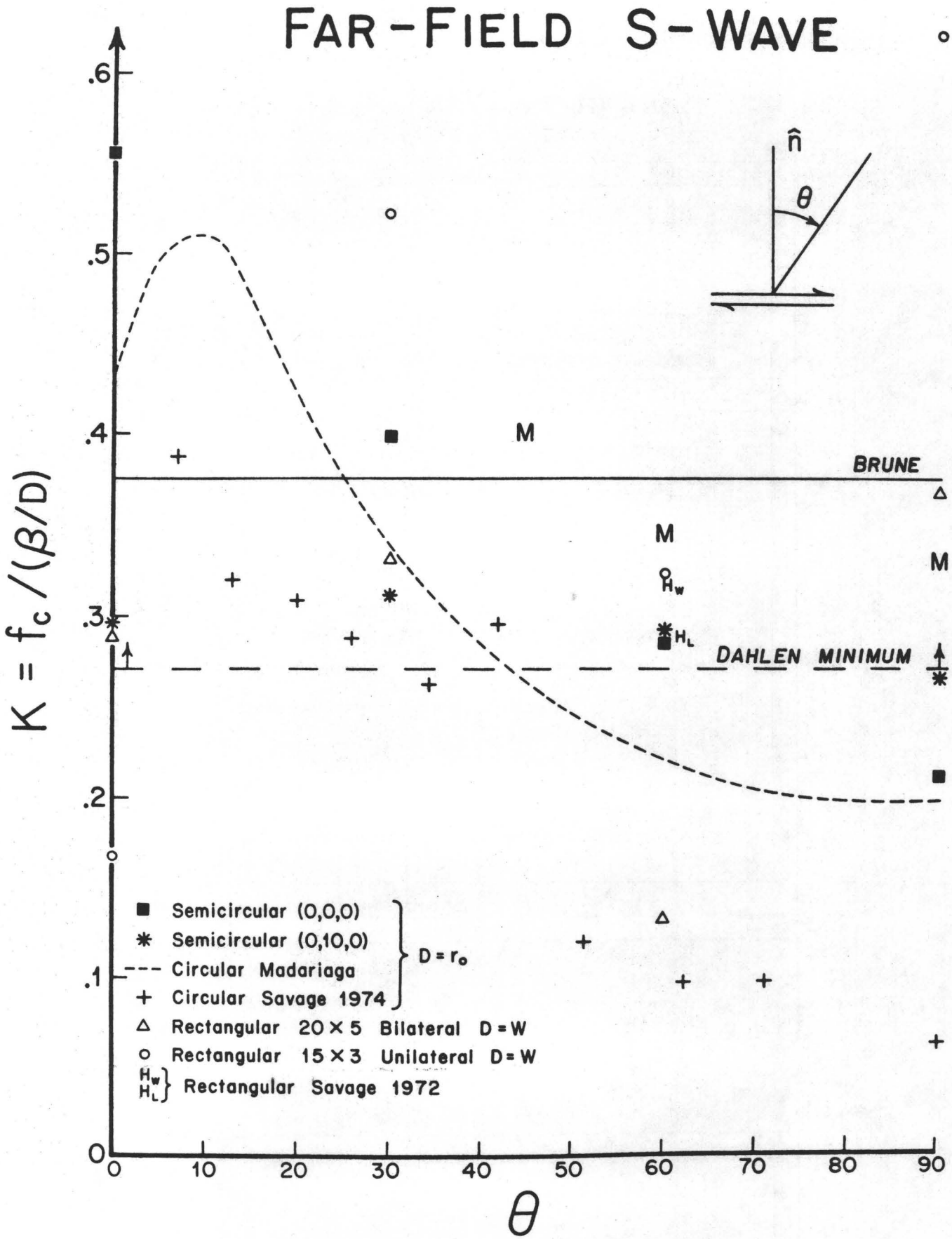

FAR-FIELD S-WAVE
K = f_c /(β/D)
θ
n̂
θ
BRUNE
DAHLEN MINIMUM
M
M
M
H_W
H_L
Semicircular (0,0,0)
Semicircular (0,10,0)
Circular Madariaga
Circular Savage 1974
D = r_o
Rectangular 20 × 5 Bilateral D = W
Rectangular 15 × 3 Unilateral D = W
H_W / H_L Rectangular Savage 1972
0
.1
.2
.3
.4
.5
.6
0
10
20
30
40
50
60
70
80
90

Figure 13.

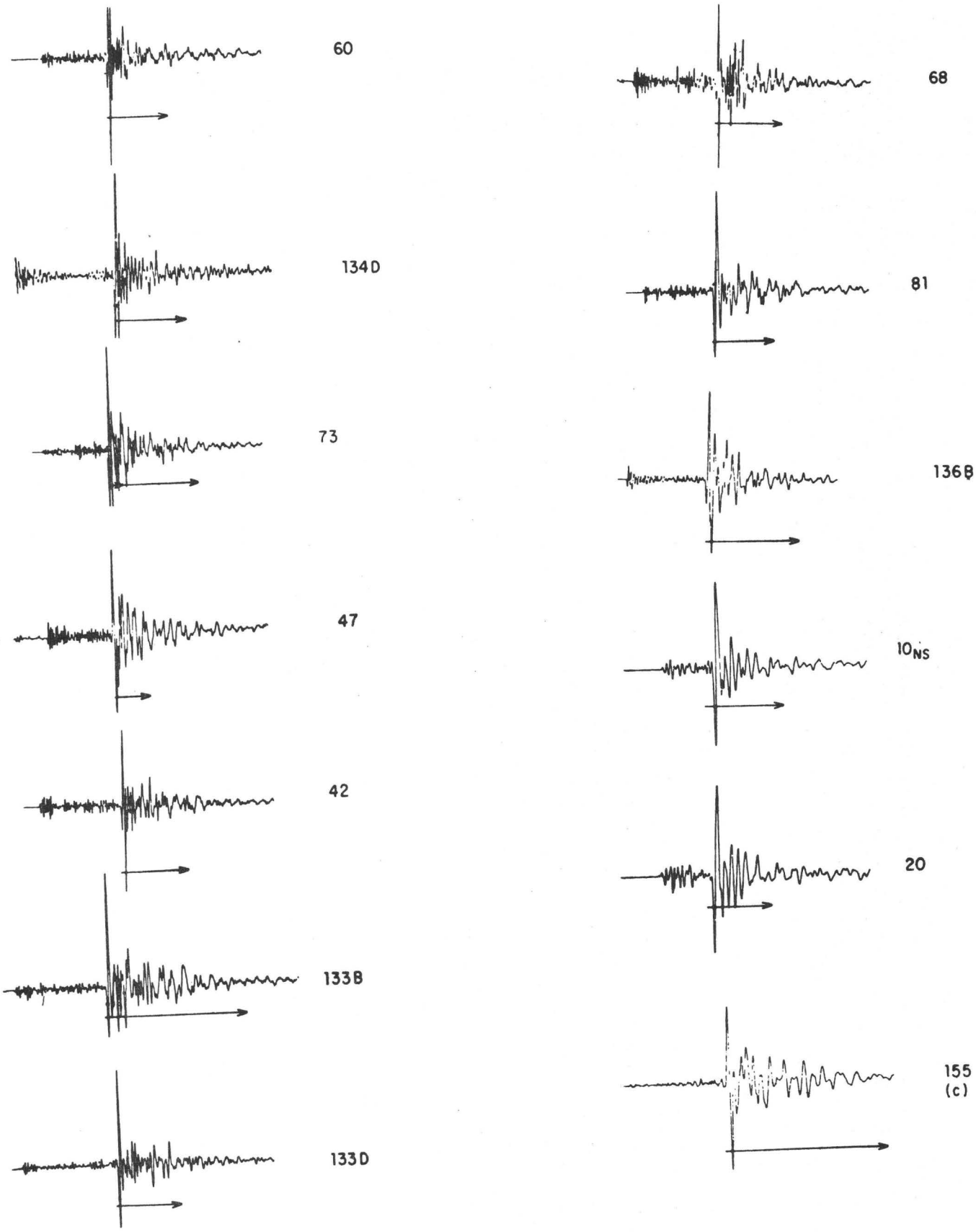


Figure 14.

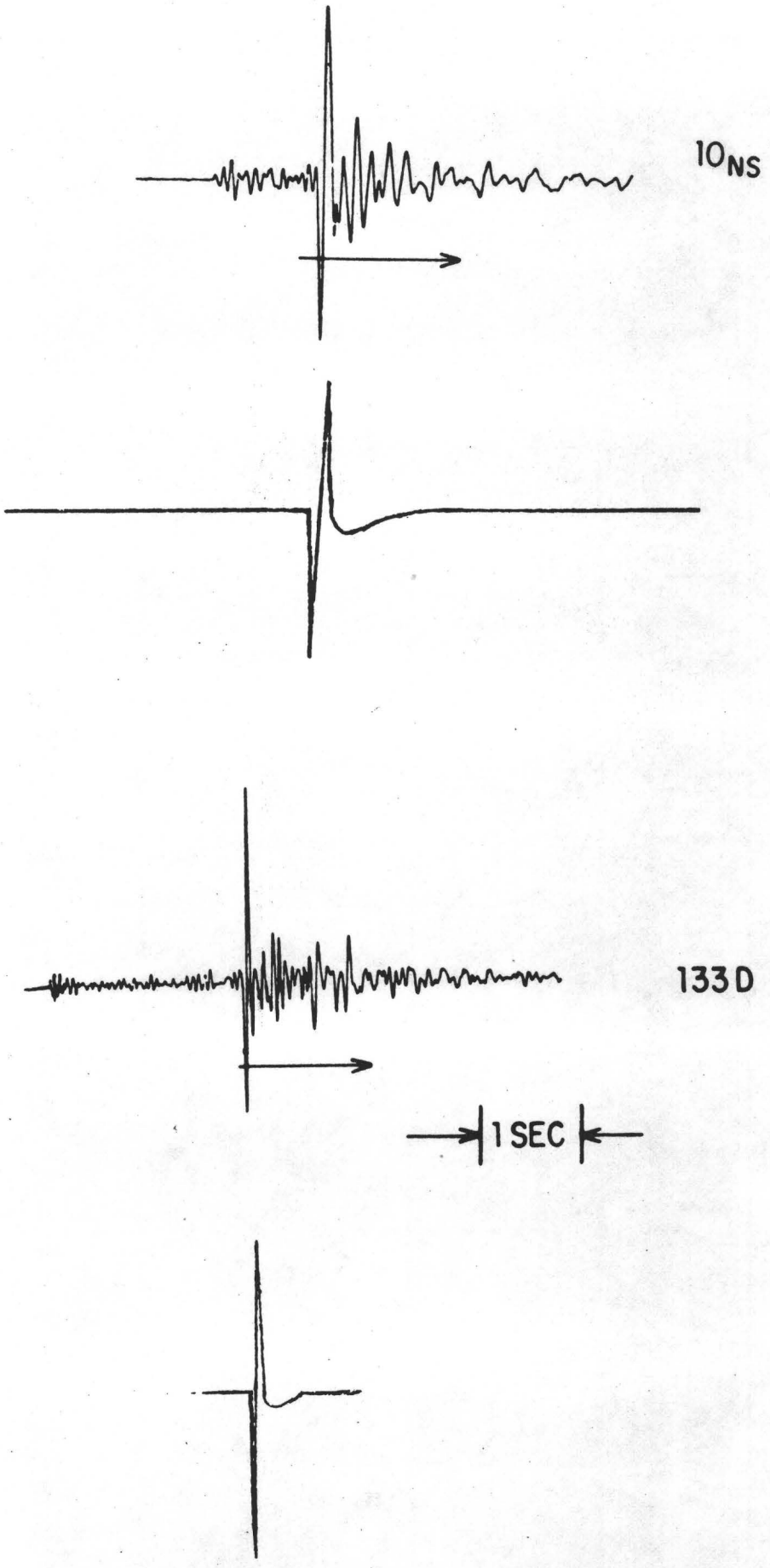


Figure 15.

MODELING OF ROCK FRICTION,
PART I: EXPERIMENTAL RESULTS AND CONSTITUTIVE EQUATIONS

James H. Dieterich
U.S. Geological Survey
345 Middlefield Road
Menlo Park, CA 94025

INTRODUCTION

The principal source of uncertainty and variability found in theories for the earthquake source is the designation of mechanical properties to represent interactions along a fault prior to and during an earthquake. It is asserted here that physical parameters and phenomena observed in laboratory fault friction experiments can provide a rational means for identifying the critical mechanical properties controlling fault slip.

Of course faults as studied in the laboratory do not have many of the complexities of natural faults. It could be argued therefore that important controlling processes that operate in natural fault systems have been simplified out of the experiments. However, there exist many analogies between the response of simple laboratory faults and real fault phenomena - a situation that invites detailed analysis of laboratory fault friction processes for possible application to faulting. The most widely noted analogy is the possible relevance of unstable frictional slip sliding (stick-slip) to the mechanism of crustal earthquakes caused by unstable fault slip (Brace and Byerlee, 1966). In addition to the qualitative similarities between stick-slip and earthquake fault slip, Dieterich (1974) proposes that stick-slip can account for the relatively low stress drops of earthquakes when differences in the geometry of slip are taken into account. Similarly, experimental observations of stable frictional slip have been equated with aseismic fault creep (Scholz et al., 1969). Elevated temperatures are found to enhance the tendency for stable sliding and may account for the absence of earthquakes below 15 km in California (Brace and Byerlee, 1970). A critical test of the use of

friction data to earthquake faulting was provided by the earthquakes at Rangely, Colorado. It was found that the Rangely earthquakes could be accounted for and modeled (Dieterich et al., 1972) at the stresses and fluid pressures, measured in the focal zone (Raleigh et al., 1972) using friction data for the onset of slip as a function of confining pressure and fluid pressure (Byerlee, 1975). At a more speculative level, observations of time-dependent friction have been used to explain the mechanism of aftershocks (Dieterich, 1972a) and experimental observations of preseismic slip may explain certain earthquake precursors (Dieterich, 1978a).

The general insensitivity of friction measurements to rock type, test conditions and characteristics of the sliding surface further suggest that laboratory friction may be relevant to natural faults under more complex conditions. Similar values for the coefficient of friction and qualitatively similar slip phenomena are obtained for slip on clean, machine finished surfaces; surfaces with simulated gouge; and for slip on fracture surfaces. This indicates that the processes controlling the coefficient of friction and slip instability are intrinsic to slip on discontinuities in rocks and are not greatly dependent on complexities of geometry, structure or composition.

Instability theories for the earthquake source have an essential feature in common. Some form of displacement (or strain) weakening for the fault or focal region must be postulated to give rise to the instability and stress drops of earthquakes. At present there is no consensus as to the mechanism of displacement weakening and the number and variety of instability models that have been proposed is noteworthy. For repeated slip along a fault, the requirement of displacement weakening has a corollary - some type of healing

mechanism must operate to restore the fault strength following earthquake slip. Otherwise the fault strength would eventually fall to zero with repeated earthquakes.

The purpose of this paper is to: 1) present experimental data that is relevant to understanding displacement weakening and healing processes, and 2) develop constitutive equations that account for the experimental details. The companion paper employs the constitutive equations to model experimental observations of preseismic fault slip.

PREVIOUS WORK

The friction model and constitutive relationships developed in a later section of this paper extend the results of Dieterich (1978b) on time- and velocity-dependence of friction and displacement weakening effects. Earlier observations showing time-dependence of friction (Dieterich, 1972b; Scholz et al., 1972; Tuefel and Logan, 1977) give convincing evidence on the characteristics of fault healing in friction experiments. It is found that the coefficient of static friction increases with the time of stationery contact. The following empirical law has been proposed (Dieterich, 1978b) for the time dependency:

$$\mu = \mu_o + A \log (Bt + 1) \tag{1}$$

where μ is the coefficient of friction is given by the ratio of shear to normal stress, τ/σ; t is the time of contact; and μ_o, A and B are constants with values of approximately 0.6 - 0.8, 0.01 - 0.02 and 1.0 - 2.0 respectively. Measurement of sliding friction at different velocities of slip (Dieterich, 1978b) has shown an analogous velocity dependency:

$$\mu = \mu_o + A \log \left(B \frac{d_c}{\dot{\delta}} + 1\right) \tag{2}$$

where d_c is an experimental displacement parameter and $\dot{\delta}$ is slip velocity. Hence, friction decreases with increasing slip velocity. Note that (2) is equal to (1) if the constants μ_o, A and B are the same and time t is replaced by

$$t = \frac{d_c}{\dot{\delta}} \tag{3}$$

The identification of the parameter d_c was suggested by records giving friction as a function of displacement. Those records show that if the slip velocity is increased the coefficient of friction does not drop immediately to a value characteristic of the new slip velocity, but that the friction changes with displacement and stabilizes at the lower value only after a critical displacement, d_c, has taken place. The magnitude of d_c appears to be independent of normal stress, σ, and the magnitude of the change of velocity, but does correlate with surface roughness. Values of 5×10^{-4} cm and 1×10^{-4} cm were obtained for surfaced lapped with #240 and #600 abrasives respectively.

These observations were interpreted (Dieterich, 1978b) to be of importance in understanding the surface interactions that control friction and cause slip instability. The dependence of d_c on surface roughness and the apparent success of using (3) to replace time in equation (1) suggest that the t in (3) is properly the average lifetime of a population of contacts and that d_c is the displacement required to completely change the population of contact points. Hence, the friction observations noted above result from the competition of two distinct processes. First, as a population of contacts ages, friction increases according to (1). Second, displacement acts to destroy an existing population of contacts which is then replaced with new and consequently weaker contacts. Displacement weakening occurs whenever the average lifetime of the population of contacts decreases because of displacement.

Based on this interpretation, a simple spring and slider model was proposed to explain experimental observations of the transition from stable sliding to stick-slip (Dieterich, 1978b). With this model the spring, with stiffness K, represents the combined stiffnesses of the sample and test

apparatus. The friction of the slider satisfies equations (1), (2) and the drop of μ from the static to the sliding value is assumed to be linear over the displacement d_c. Unstable slip occurs when the decrease of friction with displacement has a slope that exceeds the slope, -K, of the unloading curve of the system. In that case the stress acting on the slider will be greater than the frictional resistance to slip causing an acceleration of the slider and instability. Stable sliding takes place if the decrease of friction with displacement has a lesser slope than the characteristic unloading curve of the system. The transition from stable sliding to stick-slip occurs when the slope of the unloading curve first equals the slope for displacement weakening of friction. The model quantitatively accounts for experimental observations that show the transition from stable to unstable slip depends on normal stress, stiffness and surface roughness. The principal shortcoming of the model is that it does not predict nor account for laboratory observatons of preseismic slip. Modeling of preseismic slip which is discussed in the companion paper appears to require a more complete constitutive relationship for friction as a function of displacement, velocity and time. The procedure followed below is to build up constitutive equations that permit detailed simulation of experimental records for μ for different time, displacement and velocity conditions.

EXPERIMENTAL RESULTS AND CONSTITUTIVE EQUATIONS

The experimental results presented here were obtained using the "sandwich" type direct shear configuration described previously (Dieterich, 1972). The sample material is gray "granite" from the Raymond, California quarry. Sliding surfaces were lapped using #60, #240 or #600 abrasive.

Figure 1 gives results for the coefficient of friction, μ, as a function of displacement. Normal stress was held constant at 60.7 bars. For each of the two experimental runs shown in Figure 2, slip velocity varied in a stepwise manner from $\sim 10^{-6}$ cm/sec to $\sim 10^{-3}$ cm/sec. Each curve represents a continuous record in which the velocity of slip was held constant for a displacement of $\sim 5 \times 10^{-3}$ cm then abruptly increased by a factor of ten and held constant for another displacement of $\sim 5 \times 10^{-3}$ cm and so on. The small irregularities in the experimental curves are caused by electronic noise. The apparently greater noise at the slower sliding velocities arises because of lower recording pen velocity which compresses the irregularities on the record. In Figure 1 note that μ stabilizes at progressively lower values for each increase in sliding velocity as indicated by equation (2). The transient increase in μ, observed when the velocity is increased, is discussed below. Critical displacement, d_c, for slip on the #90 surfaces shown in Figure 1 is taken to be $\sim 2 \times 10^{-3}$ cm. Table 1 lists the center-line-average surface roughnesses and critical displacements for the #60, #240 and #600 surfaces.

Figure 2 gives μ as a function of displacement for a different type of experiment with the #60 surfaces. In these experiments the driving ram was held stationary for approximately 400 seconds then advanced at different

velocities as shown at the top of the figure. Normal stress was held constant at 57.2 bars. The peak of the curve gives the "static" friction. These experiments show that static friction is sensitive to loading rate - a feature that had not been noted previously. Higher rates of loading give higher static friction values. This effect appears to be analogous to the well-known effect of strain rate on critical stress for yield or fracture of silicates. Below, it will be shown that this effect is probably related to the transient increase in μ when velocity is increased as shown by Figure 1. Note again in Figure 2 that sliding friction tends to stabilize at lower values for the higher slip velocities. The points on the curves labeled o.v. and c.v. correspond to the opening and closing respectively, of the hydraulic valve that controls motion of the driving ram. The jump in the curve at c.v. probably arises because of a slight pressure surge in the hydraulic system. Small amounts of fault creep that partially relax stress in the sample cause the decrease in amplitude of the curves between c.v. and o.v. During creep the rate of slip rapidly decays with time. It is of possible significance that the creep takes place at stresses below the stresses for steady-state slip (equation 2). Previously, Johnson (1975) reported a similar surface creep phenomena at stresses below the nominal stress for static friction and having a slip velocity that decays with time.

For the purposes of developing a more complete constitutive relation for friction it is assumed that friction of rocks like friction of metals and most other materials is largely controlled by adhesion at actual points of contact between sliding surfaces. Several lines of evidence indicate that adhesion is the principal source of the frictional force in rocks (Dieterich, 1978b).

Reliance on adhesion is not essential for the following discussion, however it does provide a plausible framework for discussion of parameters and for interpretation of mechanisms.

Bowden and Tabor (1964) propose that when surfaces are brought into contact minute irregularities will prevent uniform contact over the entire surface, even for flat, well polished surfaces. Actual contact is limited to scattered points (asperities) where the contact stresses are very high. An increase of the normal stress pushing the surfaces together causes the points of contact to yield and results in an increase of the real area of contact. For a unit area of surface, the real area of contact A, may by approximated by

$$A = C\sigma \tag{4}$$

where σ is the average normal stress applied over the entire surface (assuming 100% contact) and C is a material constant inversely proportional to indentation hardness or yield stress. Bowden and Tabor assert that the resistence of the surface to slip is controlled by the adhesive strength of the junctions. Hence, the average shear stress, τ, (again assuming uniform contact) for slip is proportional to the real area of contact:

$$\tau = FA \tag{5}$$

where F is the strength per unit area of a contact. The coefficient of friction, μ, is given by τ/σ:

$$\mu = \tau/\sigma = CF \tag{6}$$

Note that C x F is dimensionless because C^{-1} and F are proportional to strength.

The time-dependence of equation (1) is clearly related to an increase in area of contact, or perhaps in some situations depth of penetration of asperities (Scholz and Engelder, 1976; Tuefel and Logan, 1977; Dieterich, 1978b). Creep deformation of the asperities apparently causes the increase in area. This suggests that C in equation (4) is time dependent:

$$C = c_1 + c_2 \log (c_3 t + 1) \tag{7}$$

giving for the coefficient of friction from (6):

$$\mu = [c_1 + c_2 \log (c_3 t + 1)][F] \tag{8}$$

which is the same as (1) with:

$$c_1 F = \mu_o$$

$$c_2 F = A$$

$$c_3 = B$$

The reader will note that (8) applies to static friction if t is the time of stationary contact or to sliding friction if t is the average lifetime of the population of contacts. The use of (8) with relationship (2) for the average time of contact as a function of velocity gives a steplike change in friction whenever the sliding velocity abruptly changes. Figure 3 curve A is a simulation of the experimental run (top curve) using equations (2) and (8). Comparison of curve A with the experimental results demonstrates the unsuitability of this approach for representation of friction as a function of displacement. The steps in curve A arise because t is assumed to jump to a

new value whenever the velocity changes. The observations suggest however that t relaxes to the new value over the characteristic displacement d_c. A possible relationship for the relaxation of t as a function of displacement δ, is given by:

$$t = \left(\frac{d_c}{\dot{\delta}}\right)\left(\frac{\dot{\delta}}{d_c}\, t_o\right) e^{\frac{\delta_o - \delta}{d_c}} \tag{9}$$

where t_o and δ_o are the average contact time and displacement respectively when the slip velocity is changed to the new value. Use of (9) with (8) for step changes in velocity gives curve B of Figure 3.

The transient increase in friction observed when the velocity is increased seems not to be associated with variation of surface area with displacement, but apprears to result from another process. The interpretation is offered here that this effect results from a loading rate dependence of the strength term, F. The following relationship appears to give satisfactory results:

$$F = f_1 + \frac{1}{f_2 \log\left(\frac{f_3}{\dot{\delta}} + 10\right)} \tag{10}$$

where f_1, f_2 and f_3 are constants. For the simulations discussed here and in Part II of this study f_3 has been equated with d_c. Several other relationships might serve as well. The essential characteristic of (10) is that if the area of contact is held constant, the strength of a contact increases as the velocity of loading increases. As noted above, this effect is analogous to the effect of strain rate on yield strength of silicates. In addition, it is noted that the creep that apparently causes the time-dependent increase in area (equation 7) implies a similar strain rate dependence in strength.

Use of (10) and (7) gives the following expression for the coefficient of friction as a function of velocity and time of contact.

$$\mu = \left[c_1 + c_2(\log \; c_3 t + 1)\right]\left[f_1 + \frac{1}{f_2 \log\left(\frac{f_3}{\dot{\delta}} + 10\right)}\right] \qquad (11)$$

Equation (11) used with (9) for the variation of contact time with displacement and velocity appears to adequately represent the static, transient and steady state sliding friction observations described above.

Curve C in Figure 3 gives the variation of μ as a function of displacement and velocity using equations (11) and (9). Overall, curve C appears to be in good agreement with the data. An increase in sliding velocity first causes an increase in friction because of the velocity dependence of F. As sliding proceeds however, the average time of contact relaxes to the new value, causing a decrease in the real area of contact and an overall lowering of the total resistance to slip. In detail the simulation shown by curve C (Fig. 3) differs from the experiments, mainly in the sharp peak in friction obtained when the velocity is increased.

Figure 4 shows further refinement of the simulation that tends to smooth out the peaks. In this case the velocity of slip accelerates and temperarily overshoots the driving velocity. This type of velocity overshoot is evident in the displacement vs. time records for the experiments and apparently arises because of elasticity of the sample and load-bearing anvils in the apparatus.

Figure 5 gives a simulation of the experiment shown in Figure 3 for static friction as a function of loading velocity. Again, equation (11) was used to determine μ. The initial time of static contact was taken to be 400 seconds and equation (9) was used to give the variation of contact time with

displacement and velocity. Comparison of Figure 5 with Figure 3 shows that the simulation provides a reasonably good quantitative representation of principal features of the experiment. No attempt was made here to represent the variations in slip velocity due to elastic effects in the apparatus or to model the fault creep that occured in the experiments when the hydraulic system feeding the driving ram was valved off. Incorporation of a more realistic velocity history that accounts for accelerations of slip rates when the surface begins to slide would result in flattening of the peaks in the friction curves and yield a better agreement with the experimental curves.

Fault creep occurs at stresses as much as 3 percent below the coefficient of friction for steady state slip (i.e. when $t = d_c/\dot{\delta}$) and can be accounted for by equation (11). The effect arises becaue of the velocity dependence of F. The characteristics of fault creep observed in the experiments of Figure 3 are illustrated with the aid of Figure 6. Figure 6 plots μ from equation (11) against the logarithm of the contact-time for different velocities of slip. Constants for C, F and d_c are those used for the above simulations. The heavy dashed curve in Figure 6 gives μ as a function of contact time assuming $t = d_c/\dot{\delta}$. Hence, the dashed curve gives the value for μ and t for steady-state slip at constant velocity. During slip at constant velocity if the parameters for μ, and t fall on the dashed curve, those parameters will be stable and will not change unless $\dot{\delta}$ is changed. If the conditions during slip plot above the dashed curve, then slip will be at a velocity in excess of the steady-state velocity and therefore cause a decrease in t with displacement. If the velocity of slip is held constant then t and consequently μ must

decrease with displacement until both parameters coincide with the steady state values. If the stress is held constant, then acceleration of slip must result. Similarly, slip may take place for conditions plotting below the dashed curve. In this case the velocity of slip is less than the steady state value and t must increase. If velocity is held constant, μ will increase with t until the steady-state value is reached. If stress (μ), is held constant or decreases then $\dot{\delta}$ must decay with total elapsed time.

Curves labeled by points a-i in Figure 6 show the approximate paths followed for the experiment of Figure 3. Point (a) in Figure 6 corresponds to the first peak in friction for slip at 1.8×10^{-3} cm/sec. Contact time is approximately 400 seconds. As sliding progresses, μ decreases along the path a-b and stabilizes at b. Closure of the hydraulic valve prevents further slip at b. Any additional slip results in a decrease in stress in the sample which drops μ into the field below the dashed curve. Because $\dot{\delta}$ is less than the steady-state value, contact time increases along the path b-c. Assuming that slip is small during creep, displacement will not change the population of contacts significantly and contact time will be approximately equal to the duration of creep. At c the valve is opened and stress rapidly rises to the second peak, d, where the slip velocity is 1.2×10^{-4} cm/sec. In response to slip, t and μ follow the path d-e and stabilize at e. Again at e, the valve is closed and creep occurs following the path e-f. Path f-g is the loading of the sample at a velocity of 1.4×10^{-3} cm/sec and g-h is slip at that velocity. Creep following closure of the value follows h-i.

SUMMARY AND DISCUSSION

Equation (11) used with equation (8) appear to give an adequate representation of the experimental observations reported here for the variation of friction with time, displacement and slip velocity. It is quite possible that other mathematical relationships could give as good or perhaps better representation of the data. However, the results presented here clearly establish the qualitative effect of several experimental parameters affecting rock friction. Specifically, the coefficient of friction can be expressed as the product of the parameters C and F. Parameter C depends on the time of contact as given approximately by equation (7) and F depends on velocity of slip as given approximately by equation (10). During slip, time of contact depends on $\dot{\delta}^{-1}$ and for changes in velocity t relaxes to a new value over the characteristic displacement, d_c, equation (8). The decrease in contact time with increasing velocity causes a decrease in friction for steady-state slip. Hence, there are two competing velocity effects. The first is a transient effect giving an immediate increase in friction for an increase in slip velocity. The second arises indirectly from decrease in contact time with increased velocity and becomes evident only after finite displacement at the new velocity.

The effect of contact time on friction apparently arises because creep at contacting asperities increase the actual area of contact between the sliding surfaces. While the transient velocity effect given by parameter F is identified here with the breakdown of adhesive junctions at the points of contact, it is not clear what specific processes control asperity failure. Simple pulling apart of the adhesive junctions without surface damage, brittle

failure of the contacting asperities and ploughing may all be important processes. The presence of wear-generated gouge along the surfaces gives evidence for at least some brittle processes during slip. The observations on time of contact suggest that creep determines the size, i.e. cross sectional area, of the junctions that subsequently fail, at least in part by velocity dependent brittle processes.

Additional experiments are needed to explore the details of time, displacement and velocity dependence as a function of rock type, pressure and especially temperature. Generally, friction data in the literature are inadequate to test for the time, velocity and displacement dependence reported here for normal stress above approximately 1 kilobar. The similarity of observations for frictional instabiltiy for different types of rocks and for very different normal stresses suggests that the effects noted above are general characteristics of rock friction at room temperatures. Time-dependent effects that agree with the general form of (7) have been reported for normal stresses from 19 bars to ∿700 bars for a variety of rocks (Dieterich, 1972; Tuefel and Logan, 1977). Evidence for the parameter d_c is limited and has been noted only in the study by Dieterich (1978b) and the present study. However, the model for the transition from stick-slip to stable sliding by Dieterich (1978b) specifically requires a characteristic displacement and successfully accounts for data to normal stresses of 1.2 kilobars. Data for the velocity dependence of the type given in equation (10) is reported here for the first time and needs additional experimental corroboration. The similarity of the velocity dependence of F to the widely observed dependency of the brittle strength of silicates suggests that this effect may also be general.

REFERENCES

Bowden, F. P., and Tabor, D., The friction and lubrication of solids, vol. 2, Clarendon, London, 1964.

Brace, W. F., and J. D. Byerlee, California earthquakes: why only shallow focus?, Science, 168, p. 1573-1575, 1970.

Brace, W. F., and J. D. Byerlee, Stick-slip as a mechanism for earthquakes, Science, 153, p. 990, 1966.

Byerlee, J. D., The fracture strength and frictional strength of the Weber sandstone, Int. J. Rock Mech. Min. Sci., 12, p. 1-4, 1975.

Dieterich, J. H., Time-dependent friction as a possible mechanism for aftershocks, J. Geophys. Res., 77, p. 3771-3781, 1972a.

Dieterich, J. H., Time-dependent friction in rocks, J. Geophys. Res., 77, p. 3690-3697, 1972b.

Dieterich, J. H., Earthquake mechanisms and modeling, Ann. Rev. of Earth and Planetary Sciences, 2, p. 275-301, 1974.

Dieterich, J. H., Preseismic fault slip and earthquake prediction, J. Geophys. Res., in press, 1978a.

Dieterich, J. H., Time-dependent friction and the mechanics of stick-slip, PAGEOPH, in press, 1978b.

Dieterich, J. H., C. B. Raleigh, and J. D. Bredehoeft, Earthquake triggering by fluid injection at Rangely, Colorado, Proc. Int. Soc. Rock Mech. and Int. Assoc. Eng. Geol. Symp.: Percolation through fissured rock. Pap. T2-B, 12 pp., 1972.

Johnson, T. L., A comparison of frictional sliding on granite and dunite surfaces, J. Geophys. Res., 80, p. 2600-2605, 1975.

Raleigh, C. B., J. H. Healy, and J. D. Bredehoeft, Faulting and crustal stress at Rangely, Colorado. In: Flow and Fracture of Rocks, Geophys. Monogr., 16, p. 275-284, 1972.

Scholz, C. H., and Engelder, J. T., The role of asperity indentation and ploughing in rock friction: I. Asperity creep and stick-slip, Int. J. Rock Mech. Sci. and Geomech. Abstr., 13, p. 149-154, 1976.

Scholz, C. H., P. Molnar, and T. Johnson, Detailed studies of frictional sliding of granite and implications for the earthquake mechanism, J. Geophys. Res., 77, p. 6392-6406, 1972.

Scholz, C. H., M. Wyss, and S. W. Smith, Seismic and aseismic slip on the San Andreas fault, J. Geophys. Res., 74, p. 2049-2069, 1964.

Teufel, L. W., and J. M. Logan, Effect of shortening rate on the real area of contact and temperatures generated during frictional sliding, PAGEOPH, in press, 1978.

FIGURE CAPTIONS

Figure 1. Coefficient of friction, μ, vs. displacement. Slip velocities are shown by the arrows above the experimental curves.

Figure 2. Coefficient of friction vs. displacement. The driving ram was held fixed for $\sim$400 seconds, then advanced at the velocities indicated above the experimental curve.

Figure 3. Comparison of experimental results from Fig. 1 with empirical friction laws. Curve (A) employs equations (2) and (8) with $c_1F = 0.72$, $c_2F = 0.005$ and $c_3 = 0.1$. Curve (B) employs equations (8) and (9). Curve (C) employs (9) and (11) with $d_c = 2 \times 10^{-3}$ cm, $c_1 = 0.69$, $c_2 = 0.010$, $c_3 = 0.5$, $f_1 = 1.0$, $f_2 = 25.$ and $f_3 = 2 \times 10^{-3}$.

Figure 4. Comparison of experimental results with empirical friction laws (9) and (11). The upper curves give μ as a function of displacement from the experiments of Fig. 1 and for the model. Lower curve gives velocity as a function of displacement used for the model.

Figure 5. Simulation of the experimental data of Fig. 2. Friction parameters for equations (9) and (11) are: $d_c = 2 \times 10^{-3}$ cm, $c_1 = 0.69$, $c_2 = 0.013$, $c_3 = 0.5$, $f_1 = 1.0$, $f_2 = 20.$, $f_3 = 2 \times 10^{-3}$.

Figure 6. Coefficient of friction vs. time of contact from equation (11) using the parameters given for Fig. 5. The solid curves give slip velocity. The dashed curve gives μ vs. t for steady state slip, $t = d_c/v$. The dotted curves a-i show the path followed by the experiment in Fig. 2.

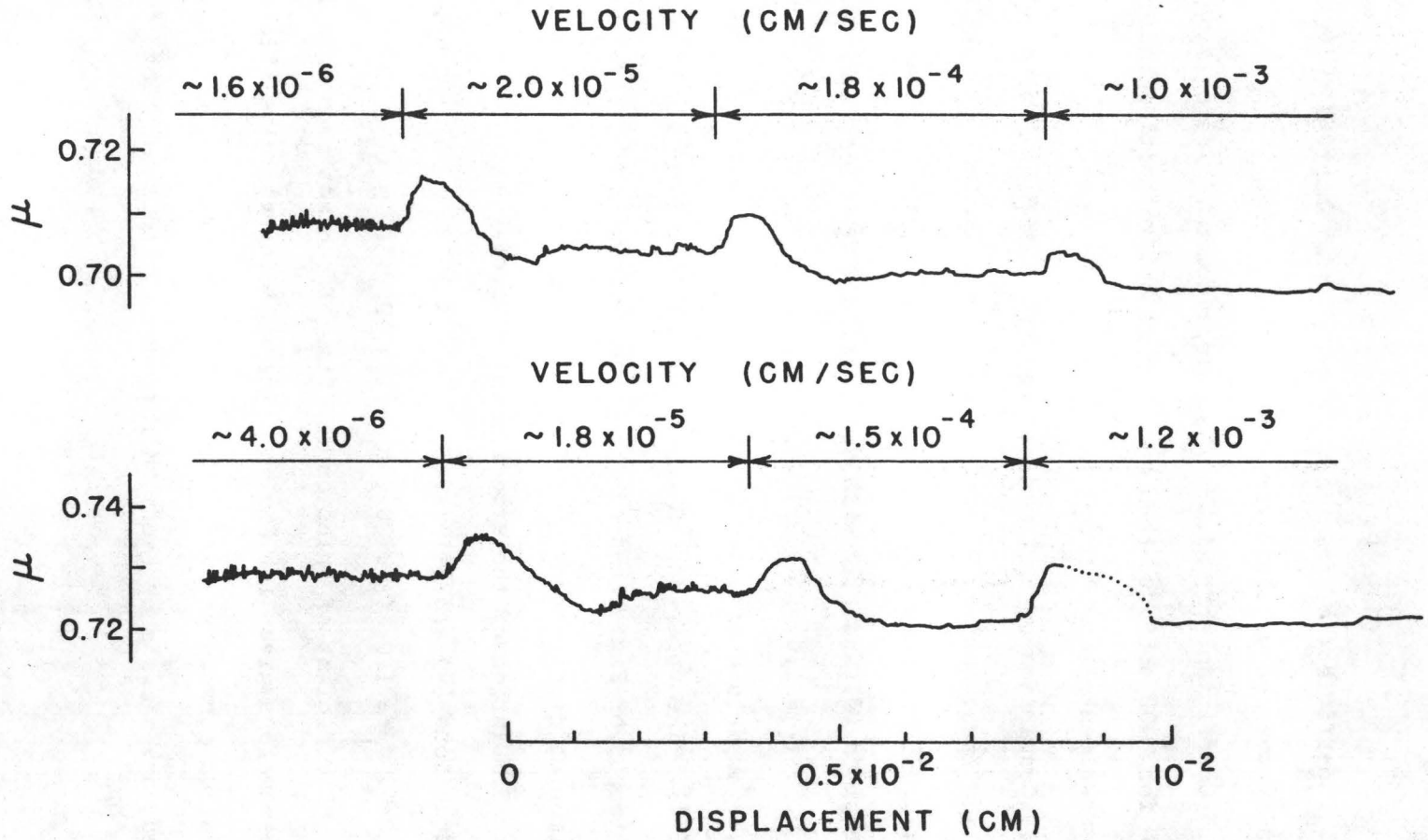

VELOCITY (CM/SEC)
~1.6x10^-6
~2.0x10^-5
~1.8x10^-4
~1.0x10^-3
0.72
μ
0.70
VELOCITY (CM/SEC)
~4.0x10^-6
~1.8x10^-5
~1.5x10^-4
~1.2x10^-3
0.74
μ
0.72
0
0.5x10^-2
10^-2
DISPLACEMENT (CM)

Figure 1

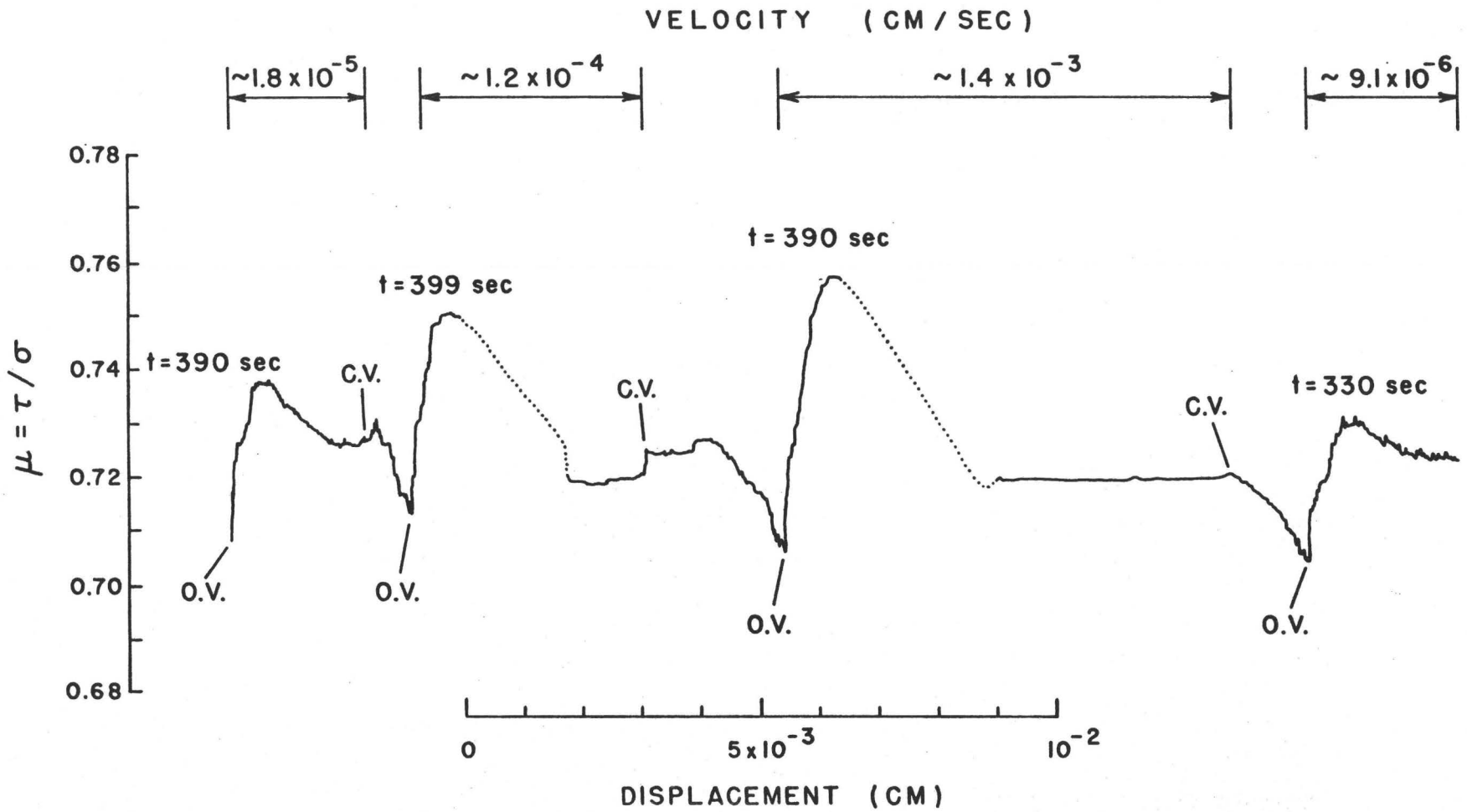

VELOCITY (CM / SEC)
~1.8 x 10^-5
~1.2 x 10^-4
~1.4 x 10^-3
~ 9.1 x 10^-6
0.78
0.76
0.74
0.72
0.70
0.68
μ = τ / σ
t = 390 sec
t = 399 sec
t = 390 sec
t = 330 sec
C.V.
C.V.
C.V.
C.V.
O.V.
O.V.
O.V.
O.V.
0
5x10^-3
10^-2
DISPLACEMENT (CM)

Figure 2

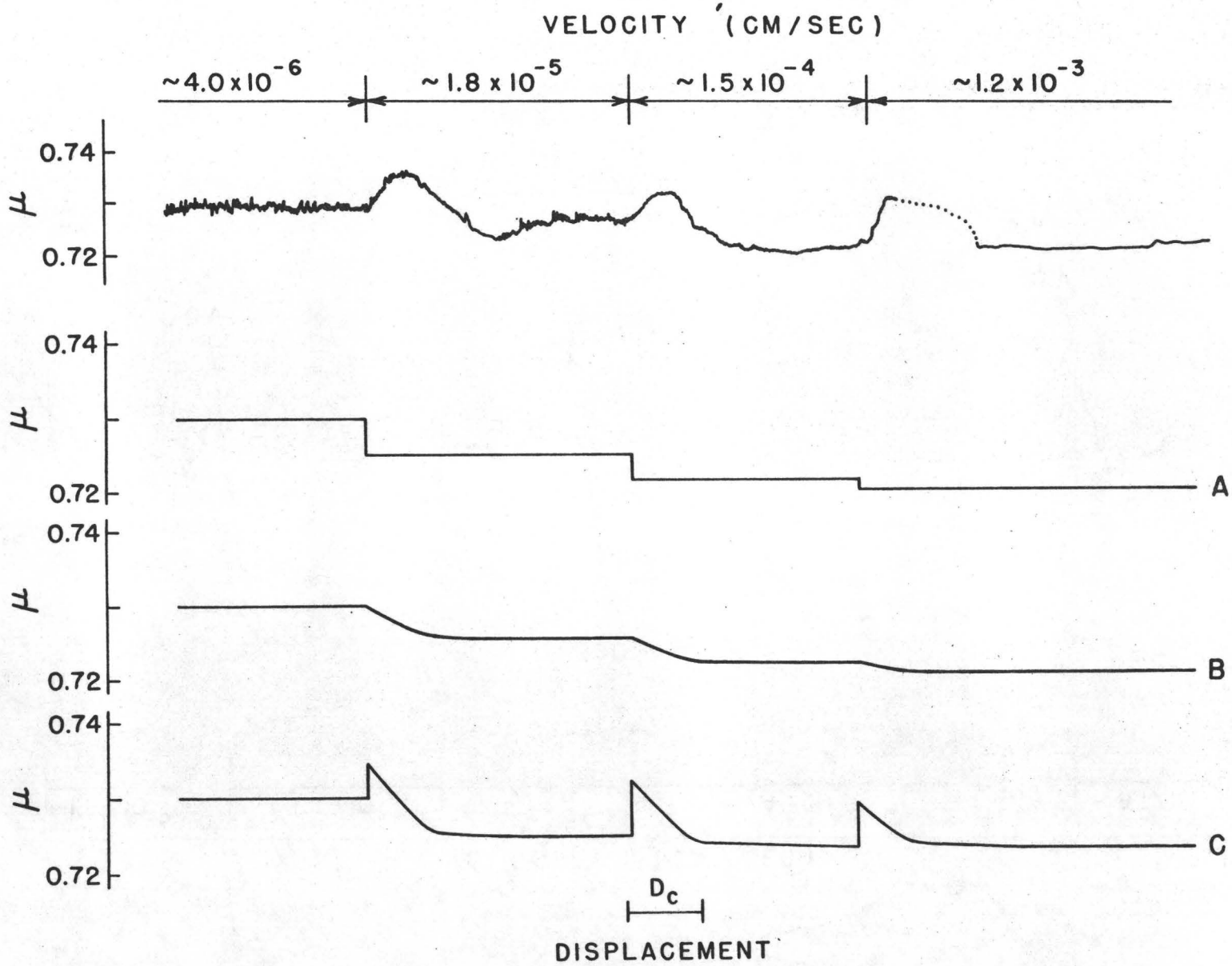

VELOCITY (CM/SEC)
~4.0x10^-6
~1.8x10^-5
~1.5x10^-4
~1.2x10^-3
0.74
μ
0.72
A
B
C
Dc
DISPLACEMENT

Figure 3

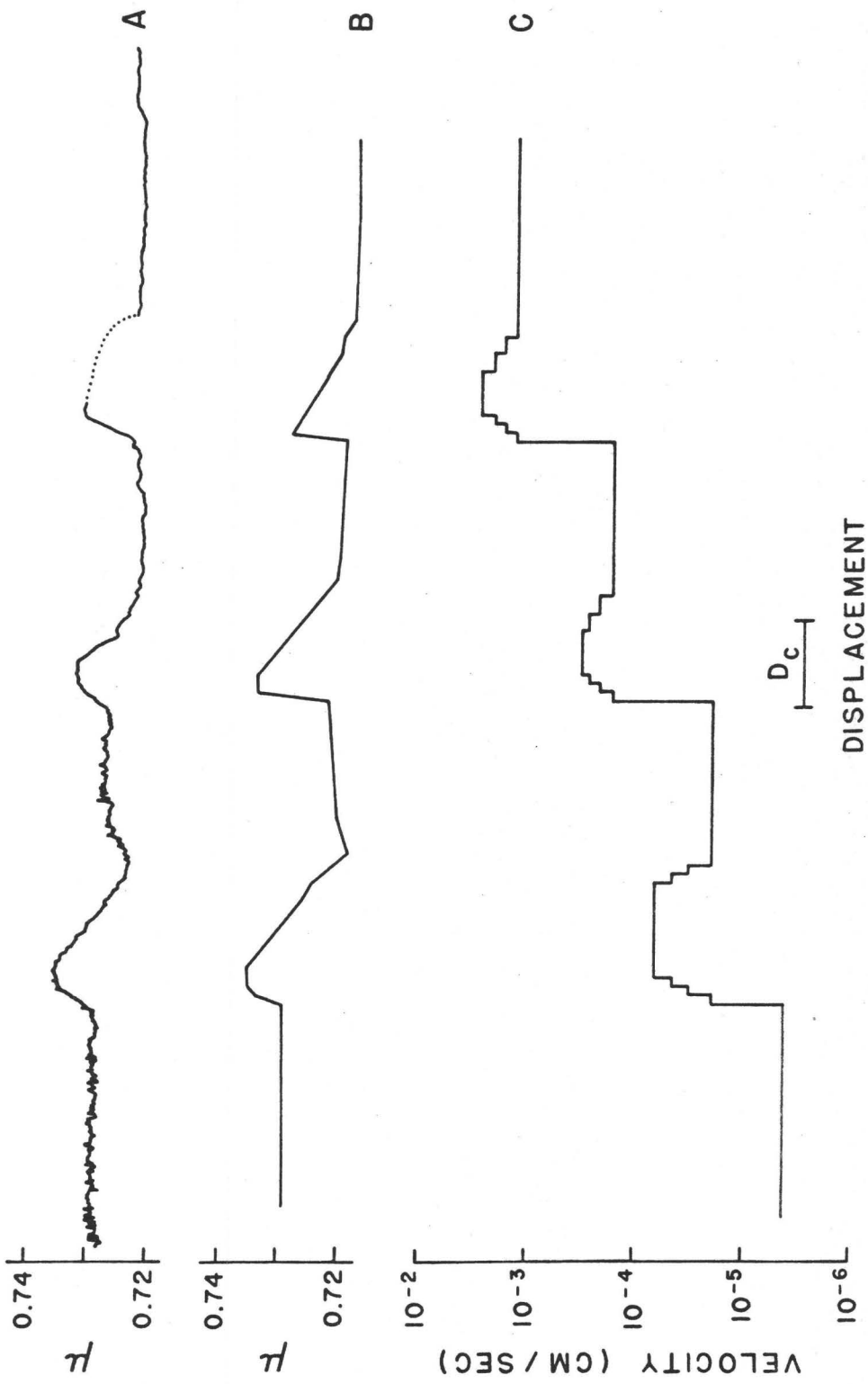

A
B
C
0.74
0.72
μ
VELOCITY (CM/SEC)
10-2
10-3
10-4
10-5
10-6
Dc
DISPLACEMENT

Figure 4

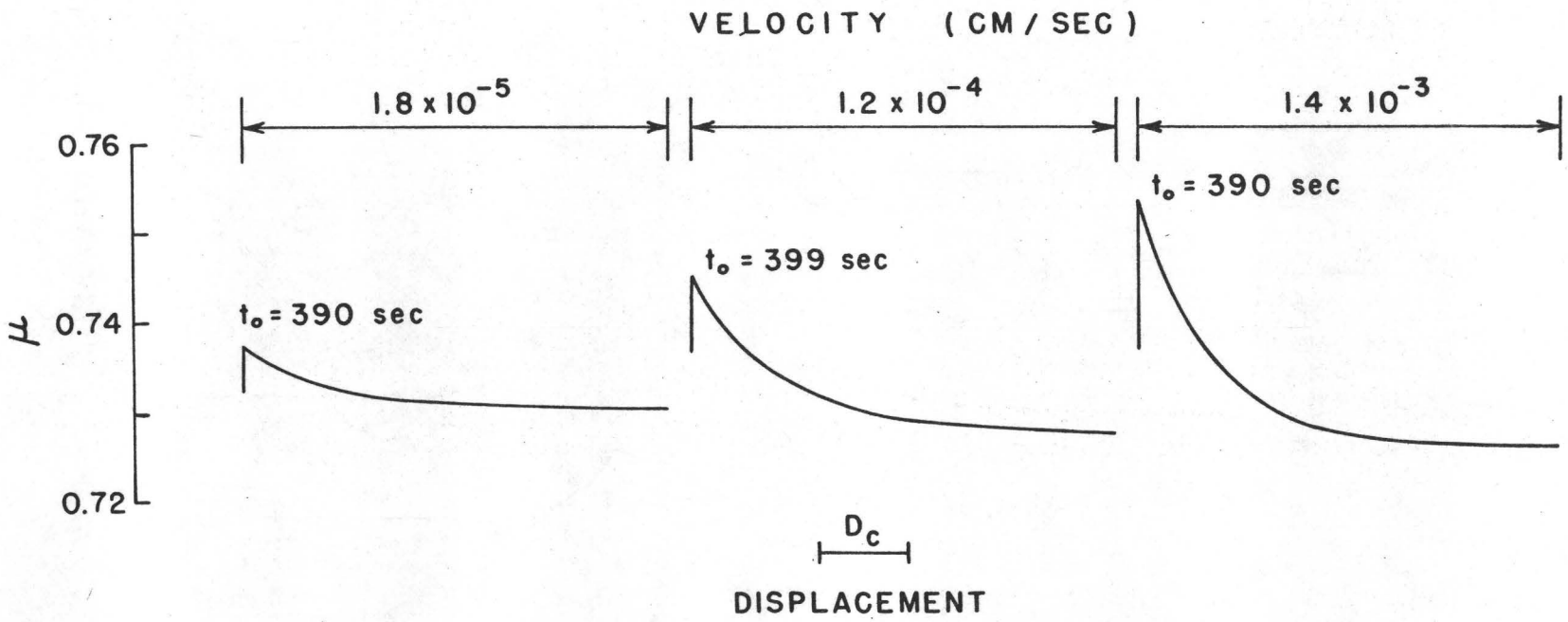

VELOCITY (CM/SEC)
1.8 x 10^-5
1.2 x 10^-4
1.4 x 10^-3
0.76
0.74
0.72
μ
t_o = 390 sec
t_o = 399 sec
t_o = 390 sec
D_c
DISPLACEMENT

Figure 5

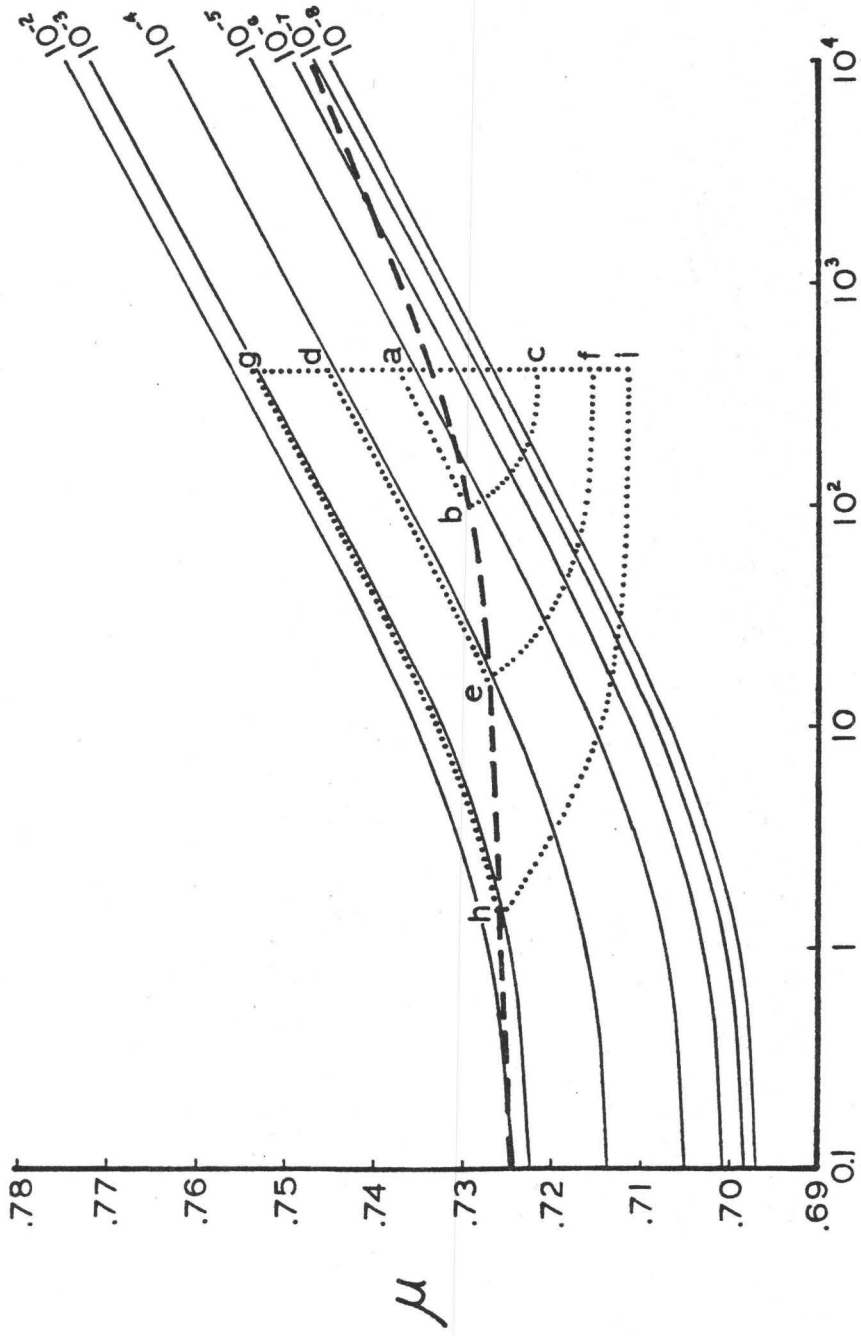

10^{-2}
10^{-3}
10
10^{-4}
10
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10
a
b
c
d
e
f
g
h
i
.78
.77
.76
.75
.74
.73
.72
.71
.70
.69
μ
0.1
1
10
10^2
10^3
10^4
CONTACT TIME

Figure 6

MODELING OF ROCK FRICTION, PART 2:

SIMULATION OF PRESEISMIC SLIP

James H. Dieterich
U.S. Geological Survey
345 Middlefield Road
Menlo Park, CA 94025

Part I of this study first presents experimental results for friction as a function of time, displacement and velocity and then develops constitutive relationships that permit a fairly accurate simulation of the experimental phenomena. This portion of the study applies those results to the modeling of preseismic slip and the initiation of unstable slip. The motivation for this analysis is twofold. First, some recent experiments (Dieterich, 1978a) provide detailed data for preseismic slip that afford an opportunity to further test the constitutive equations in a system with fairly complicated mechanical interactions. The model presented below has the mechanical elements needed to represent the principal interactions that have been identified in the laboratory experiments. Simulation of earthquake faulting probably requires analogous interactions. The model is two-dimensional physical properties and stresses along the sliding surface are permitted to vary as a function of position and the elastic properties of the sample and test apparatus are represented. Second, the process of preseismic slip in laboratory experiments holds obvious interest as a possible earthquake precursor. At present, the applicability of experimental preseismic slip data for earthquake precursor models is quite conjectural because the mechanisms and parameters controlling preseismic slip are poorly understood. The simulations presented here may provide some insight into the mechanics of preseismic slip in laboratory experiments.

PRESEISMIC SLIP

Stable slip as a premonitor of unstable (seismic) slip occurs for a variety of experimental conditions and for different types of sliding surfaces (for example: Scholz et al., 1972; Logan et al., 1972; Byerlee and Summers, 1975). While the magnitude of the preseismic slip displacements observed for the studies cited above is extremely variable, it appears that at least small amounts of stable slip always precede slip instability.

Scholz et al. (1972) have examined preseismic slip on ground surfaces of Westerly granite under biaxial loading conditions. They observe preseismic displacements on the order of 10^{-3} cm, independent of strain rate. The duration of preseismic slip is inversely proportional to strain rate. Plots of fault displacement against time show accelerating displacement rates up to the time of unstable slip.

The study by Dieterich (1978a) employed a biaxial configuration with Westerly granite (Figure 1) similar to that of Scholz et al. (1972). The principal refinement was the addition of several strain gages adjacent to the surface that permit the propagation of preseismic slip along the surface to be monitored. The gages at positions 1, 2 and 3 in Figure 1 record strains normal to the slip surface and the remaining gages 4, 5, ... 16 record shear strains parallel to the surface.

Two distinct stages of preseismic slip were consistently observed in those experiments. The first stage consists of a slip event that begins at some point on the surface and slowly propagates over most or all of the sliding surface. Figure 2 is a representative example given by Dieterich

(1978a). Shear strains are plotted as a function of time for several locations adjacent to the surface. The curves are arranged by position on the sliding surface and the numbers labeling the curves correspond to the strain gages shown in Figure 1. Prior to slip, the rams that load the sample increase the shear stress at an approximately constant rate. A leveling or a decrease in strain amplitude indicates slip on the fault adjacent to the strain gage. For the example in Figure 2 slip begins at gage 15 and propagates across the sample. Once slip starts at a location on the fault, stable slip continues at that point until the time of the instability.

The breakout of the first stage of slip at the end of the sliding surface initiates the second stage of slip which is of very short duration, 0.01 to 0.001 seconds. Note in Figure 2 that unstable slip occurs a short time after the slip event reaches the end of the sample at gage 5. While the first stage of slip appears to be intrinsically stable, the second stage rapidly becomes unstable and is characterized by acceleration of slip rates and falling stresses. Other preseismic slip events vary in detail and sometimes show greater complexity than the example in Figure 2. However, all have the common characteristic that unstable slip follows a short interval of falling stresses that is triggered after stage I reaches the end of the sample. Hence, most or all of the surface is slowly sliding at constant or slightly decreasing stress at the time the second stage of slip begins.

Figure 2 gives evidence for velocity dependent friction because the stresses are higher during the low velocity stage I slip than during the high velocity slip of the instability. Part I of this study explains that type of velocity dependence as originating from a decrease in the time of asperity

contact with increasing velocity. Independent observations show that friction decreases with time of contact. Slip instability appears to be triggered as a velocity perturbation when the slip event breaks out at the end of the surface. This causes a jump in slip velocity that in turn causes the friction to suddenly drop.

The stability of the first stage of slip appears to be associated with the discovery that inhomogeneity of stress relative to the critical stress for slip initiation controls the amount of stage I slip. The greater the inhomogeneity the greater the amount and duration of preseismic slip. For two experiments arranged to give homogeneous stress with respect to the frictional strength, no stage I preseismic slip was observed. It was concluded by Dieterich (1978a) that during the first stage of slip, increasing external load is required to drive the boundary between the slipping and unslipped portions of the surface into the regions where the applied stress is less than the frictional strength.

MODEL

Figure 3 illustrates the finite element model used to simulate the experimental configuration of Figure 1. Slip between the two triangular blocks is represented in this model by the motion of a single triangular block on a planar surface. During slip, friction at points along the surface satisfy the relationship developed in the first part of this study (Dieterich, this volume):

$$\mu = \left[c_1 + c_2 \log (c_3 t + 1)\right]\left[f_1 + \frac{1}{f_2 \log\left(\frac{f_3}{\dot{\delta}} + 10\right)}\right] \qquad (1)$$

where μ is the coefficient of friction which is given by the ratio of shear stress, τ, to normal stress, σ, during slip. The parameters c_1, c_2, c_3, f_1, f_2, f_3 are constants; t is the time of contact; and $\dot{\delta}$ is the slip velocity. During slip at constant velocity t was found to be a function of displacement, δ, velocity, $\dot{\delta}$, and initial time of contact, t :

$$t = \frac{d_c}{\dot{\delta}}\left(\frac{\dot{\delta}}{d_c} t_o\right)^{e\,\frac{\delta_o - \delta}{d_c}} \qquad (2)$$

where δ_o is the initial displacement and d_c is an experimental parameter with values characteristic of the surface roughness.

For the experiments, hydraulic rams apply shear and normal stress to the sliding surface. Those rams have finite stiffness and therefore interact with the blocks during loading and slip. The series of springs shown in Figure 3 represent the rams in the model. The total stiffness of the springs was specified to give a shear stiffness for slip of 400 bars/cm which falls within the range of stiffnesses measured for the hydraulic system under various loads.

In the model motion of the rigid boundaries distorts the springs and applies uniform loads to the sides of the block. To simplify the analysis the left boundary compresses the springs while the right boundary extends the springs by an equal amount. Hence, there is no change in the total normal stress acting across the surface. For these simulations, normal stress is independently prescribed and may vary systematically by position on the surface. Frictional strength is the product of prescribed normal stress and the coefficient of friction. It will be recalled that inhomogeneity of shear stress relative to strength controls preseismic slip in the experiments. Hence, specified variations of normal stress simulate the inhomogeneity that controls preseismic slip.

Quasi-static slip and deformation of the block and loading system are computed with the finite element method. The triangular finite element mesh shown in Figure 3 represents the block. The mesh consists of an array of elastic triangular elements connected at nodal points. The Lamé elastic constants are equal to 2.5×10^{11} dynes/cm^2. Strains within each element are constant giving a displacement field that varies linearly by position within the element. Each spring is represented by simple one-dimensional elastic elements with two nodal points that connect the element to the block and to the rigid driving boundary.

Nodal displacements are determined from a system of equations having the form:

$$\{F\} = [K]\{\delta\} \qquad (3)$$

where $\{F\}$ is a vector listing the x- and y-components of the nodal forces, $\{\delta\}$

is the corresponding vector of nodal displacements and [K] is the stiffness matrix. Procedures for the determination of [K] are found in the standard references. Stress and displacement boundary conditions are incorporated into (3) by specifying equivalent nodal forces and nodal displacements. This yields a system of simultaneous equations for the unknown nodal displacements.

Frictional stresses on the sliding surface are represented by lumped nodal forces. For the computations reported here, the sliding surface is assumed to be rigid. Hence the y-component of nodal displacement at each node on the surface is fixed at zero. It is noted that non-uniform displacements on an initially planar surface will cause warping and the appropriate boundary condition is for constant normal stress instead of fixed displacements perpendicular to the surface. Comparison of results using both boundary conditions yielded indistinguishable slip histories. The displacement condition was chosen because iterative solution of (3) was significantly more rapid than for models with the stress boundary condition and therefore permitted more economical simulations.

A time marching procedure is used to find the displacement fields at successive time steps. The duration of the time steps, ΔT, is variable and depends on the rate of loading or the rate of slip. Because μ changes over the characteristic displacement, d_c, ΔT is selected to give slip at any point that is much less than d_c. For the initial stages of slip when displacement rates are low:

$$\Delta T \approx 10^7/(K \times R) \qquad (4)$$

where K is the total stiffness of the springs and R is the rate of displacement of the rigid boundaries that distort the springs.

During the later stages of preseismic slip, the slip rates accelerate and the time steps from (4) give displacements that greatly exceed d_c. Therefore, if the maximum slip displacement, $\Delta\delta$, during a time step is greater than $d_c/5$. then the time step, $\Delta T'$, for the next step is set at

$$\Delta T' = d_c \, \Delta T/(\Delta\delta \times 5) \qquad (5)$$

Motion of the rigid boundary connected to the springs is at constant velocity. For each time step those nodal displacements are appropriately incremented and entered as displacement boundary conditions in (3). Along the sliding surface the x-component of either the nodal force or nodal displacement are specified. Initially, the simulation begins with all displacements on the surface held fixed. The boundary condition changes, permitting the nodes to slide when the applied force equals the frictional force. Conversely, if the velocity of a sliding node is less than or equal to zero, then the displacement of the node is again held fixed.

An iterative procedure coupled to the solution of (3) gives the frictional force from equations (1) and (2) that satisfy the displacements and velocities. An initial estimate of displacement and sliding velocity of each surface node for the current time step is first obtained by extrapolation of the velocity from the previous time step. From equation (2) velocity and displacement give an initial estimate of contact time that is used in (1) to obtain μ at the end of the time step. This approximation of μ is used to find the frictional force for the boundary conditions for equation (3). Solution of (3) yields an improved estimate of displacement, velocity and t which are used again to obtain a better estimate of μ. The procedure continues until the solution no

longer changes. The simulation is then continued for another time increment. Because the computations are quasi-static, the analysis terminates when unstable slip begins. The criterion for instability in these simulations was for slip velocity in excess of 10 cm/sec.

RESULTS

For these simulations the shear stress on the surface prior to slip varies somewhat with position, especially near the ends. Apparently the truncated corners of the triangular blocks prevent homogeneous loading by the springs. Inhomogeneity of shear stress on the sliding surface relative to the frictional strength was found to be a principal determinant of the amount of preseismic slip in the experiments of Dieterich (1978a). Figure 4 gives the results of a simulation in which the prescribed normal stress distribution was set to give a friction distribution that is everywhere identical to the shear stress. The average shear stress at the beginning of slip is 60 bars. Parameters for μ, equation (1), were obtained using the results of the first part of this study (Dieterich, this volume), and have the following values: $c_1 = 0.69$, $c_2 = 0.015$, $c_3 = 0.5$, $f_1 = 1.0$, $f_2 = 20.$ and $f_3 = 5 \times 10^{-4}$. Initial time of contact, $t_o = 10^4$ sec., and $d_c = 5 \times 10^{-4}$ cm, which is appropriate for the measured surface roughness of 2×10^{-5} cm.

No significant preseismic slip occurs in this simulation. Figure 4 plots the component of shear stress parallel to the surface as a function of time for the row of elements in Figure 3 that have a single node on the sliding surface. The position of those elements corresponds approximately to the position of the strain gages used in the experiment. As with the experimental results, each curve in Figure 4 is arranged by position on the sliding surface. A leveling or decrease in amplitude of the curves indicates slip adjacent to the element. As shown by Figure 4, slip rapidly accelerates giving an instability without a propagating preseismic slip event.

Figure 5 shows the results of a simulation with a friction distribution similar to that of the experiment shown in Figure 2. For this model, the difference between the frictional strength and shear stress prior to slip was specified to increase linearly along the sliding surface. The frictional strength of the extreme ends of the surface differ by a factor of 1.5. Otherwise, all conditions and parameters for μ are identical to the previous simulation. In this case a distinct period of stable sliding precedes the slip instability. The details of this simulation closely resemble the details seen in the experimental result of Figure 2. Slip begins at the end of the surface (bottom curve) where the strength is least and propagates uniformly across the surface. The upturn of the curves immediately prior to slip is caused by slip on the nearby segment of the surface which increases the rate of loading on the adjacent unslipped portion of the surface.

Once slip begins at a node, slip continues with slightly decreasing stress. Instability and rapid stress drop occur only after the propagating front of the boundary between the unslipped and the stably slipping segments of the surface breaks out at the end of the sample.

Figure 6 gives the results for the model with frictional strength arranged to give a bilateral propagation of preseismic slip. Again the parameters for μ, d_c and t_o are identical to the previous models. In this case the weakest point on the fault is near the middle of the surface and strength (measured relative to the shear stress) increases linearly toward each end. A sequence of events similar to that of Figure 5 is also seen here. The length of the zone of preseismic slip grows approximately linearly with time and instability occurs after stable slip has propagated across the entire surface.

The experimental results of Scholz et al. (1972) show an acceleration of preseismic slip rates prior to instabilty. Slip displacements were not recorded in the experiments of Dieterich (1978a). However, an analysis based on the characteristics of strain records suggests that the smoothly accelerating rates of displacement observed by Scholz et al. (1972) result from growth of the area of preseismic slip (Dieterich, 1978a). Figure 7 plots preseismic displacements for a point near the center of the surface against time for the simulations of Figures 5 and 6. The displacement curves for the simulated events show an acceleration of displacement rates similar to that seen by Scholz et al. (1972). The agreement of the results of Figures 5, 6 and 7 with the strain and displacement records of Dieterich (1978a) and Scholz et al. (1972) respectively, indicate that the slip phenomena of the simulations give a good approximation to the phenomena of those experiments. Although slip rates in the simulations accelerate with time, it is interesting to note that the rates of slip at the time of breakout of the slip events are over an order of magnitude less than the rate of loading. Because the entire surface is sliding at that time, the block must then greatly accelerate to the loading velocity, just to keep the stress from increasing. This extreme jump in slip rates appears to cause a sudden drop in friction because of the decrease in contact time with velocity given by equation (2).

The correlation between triggering of unstable slip and breakout of preseismic slip at the end of the surface was tested with the simulation shown in Figure 8. The conditions for this model are indentical to the model of Figure 6 except that the ends of the sliding surface are pinned. Therefore, the zone of preseismic slip was not allowed to break out at the ends of the sliding

surface, and slip could not homogenize the stresses to the frictional strength. No instability occured in this simulation. Unlike the simulations of Figures 5 and 6, slip rates for this model ceased to accelerate when the slip boundary encountered the pinned ends.

Surface roughness is a principal experimental determinate of slip instability with smoother surfaces having a greater tendency for unstable slip than rough surfaces (Dieterich, 1978b). Dieterich (1978b) gives a quantitative explanation for this effect based on the correlation between surface roughness and critical displacement, d_c. The simulation of Figure 9 has conditions identical to that of Figure 6 except that $d_c = 5.0 \times 10^{-3}$ cm compared to 5×10^{-4} cm for the previous simulations. While rapid slip occurs following breakout of the zone of stable slip, no instability occurs in this simulation. The maximum slip velocity is 6.4×10^{-2} cm/sec. Another simulation with $d_c = 5.0 \times 10^{-2}$ cm greatly smoothed out the accelerated slip event seen in Figure 9 and yielded a maximum slip rate of 2.5×10^{-3} cm/sec. The rate of loading in these simulations would give a stable slip rate with no change in stress of 7.1×10^{-4} cm/sec.

DISCUSSION

Some of the general characteristics of preseismic slip have been explored with the above simulations. Overall, the results agree with the phenomena reported by Dieterich (1978a) and also appear to be compatable with the data of Scholz et al. (1972). It is concluded that constitutive relationships (1) and (2), which are developed in the companion paper (Dieterich, this volume), give an adequate representation of the coefficient of friction as a function of time, displacement, velocity and surface roughness. It appears that precise simulations of experimental observations can be obtained by careful selection of the model and friction parameters.

Inhomogeneity of stress relative to friction strength causes preseismic slip. In experiments (Dieterich, 1978a) and in these simulations, little or no preseismic slip takes place if the stress is homogeneous with respect to the strength. For inhomogeneous distributions, stable slip begins at a point and propagates along the surface. The observations by Scholz et al. (1972) that the duration of preseismic slip is inversely proportional to the loading rate while the amount of slip and the form of the displacement vs. time curves are independent of loading rate indicate that preseismic slip is controlled by external loading. An analysis by Dieterich (1978a) that is confirmed by these simulations shows that the preseismic displacements observed by Scholz et al. (1972) are associated with the growth of the area of preseismic slip. If the surface stress is inhomogeneous with respect to the strength it is clear that an incremental increase of external loading is required to incrementally expand the area of slip into regions on the surface where the stress is less than the

friction. Hence, preseismic slip is intrinsically stable and is driven by external loading.

Slip instability in the simulations and experiments occurs after preseismic slip breaks out at the ends of the sample. Prior to the breakout, slip rates are less than rates of boundary loading. Hence, at the time slip reaches the end of the surface slip rates must jump to the loading rate in order to keep the stress from increasing. However, equations (1) and (2) predict that friction must decrease with displacement at the higher slip velocities because the time of contact decreases with increasing velocity. The displacement weakening leads to unstable slip.

Conditions for the occurence of unstable slip have been outlined by Dieterich (1978b) using a simple block and spring model and somewhat more primitive constitutive relationships than those employed here. The simple model would appear to approximate the conditions in the experiments only after slip breakout when the stress and friction are homogeneous and the entire block is uniformly sliding. According to the theory, instability occurs if

$$K \quad \frac{\Delta\mu\sigma}{d_c} \tag{6}$$

where K is the shear stiffness of the sample and loading system, σ is the normal stress and $\Delta\mu$ is the change in the coefficient of friction that takes place over the characteristic displacement, d_c. As originally employed, $\Delta\mu$ is the difference between friction using the time of static contact and the steady-state friction at the loading velocity. However, the results discussed in this study suggest that time of static contact might be altered by preseismic slip. Therefore, it would be more appropriate to use contact time when

slip breaks out at the end of the sample rather than time of static contact prior to slip.

This raises an interesting point. Considering the uncertainty in the specification of contact time the analysis of Dieterich (1978b) using static time of contact is in surprisingly good agreement with experimental data for the transition from stable to unstable slip as a function of normal stress, stiffness and surface roughness. For those experiments it is considered likely that preseismic slip took place prior to instability. In the preseismic slip experiments considerable trial and error were required to give a few events with homogeneous stresses and no preseismic slip. No such care was exercised for the experiments used to analyze the transition from stable to unstable slip. It appears that the analysis of Dieterich (1978) was successful because preseismic displacements were small and at a low velocity. If this were the case, the static time of contact would not be much different than the time of contact when slip reached the end of the sample. This interpretation is supported by the model simulations of this study. For example, in the simulations of Figures 5 and 6, the time of static contact was 10^4 sec while the average times of contact at slip breakout were 3×10^3 and 5×10^3 seconds respectively. By comparison the time of contact for steady slip at the loading velocity was 1.4 seconds. Using static time and steady state time with loading velocity in equation (1) gives $\Delta\mu = 0.055$ while use of contact times at slip breakout and steady state contact time gives $\Delta\mu = 0.046 - 0.050$. The difference between these values is well within experimental error. Because preseismic displacements increase with the length of the slipping zone (e.g. Fig. 7) larger sliding surfaces would have proportionally larger preseismic

displacements and the use of static contact time would probably lead to significant error for the prediction of the transition from stable sliding to stick-slip.

A principal reason for this detailed study of preseismic slip is the potential relevance to earthquake prediction (Dieterich, 1978a). The consistency with which preseismic slip is observed in laboratory experiments suggests by analogy that preseismic fault displacements may be a regular premonitor to earthquake instability. A source of uncertainty in applying experimental results to earthquake faulting is the question of scaling. Of particular importance is the possibility that for earthquake faults, preseismic slip may take place over relatively small fault dimensions with earthquake slip propagating well beyond the zone of initial slip (Dieterich, 1978a). If this is the general case then preseismic slip might be of little practical interest for earthquake prediction. Unfortunately the scaling question may not be accesible to direct experimentation. The good agreement between the results of this study and experimental data suggests that simulations might yield reliable results for earth faulting.

REFERENCES

Byerlee, J. D. and R. Summers, Stable sliding preceding stick-slip on fault surfaces in granite at high pressure, PAGEOPH, 113. p. 63-68, 1975.

Dieterich, J. H., Preseismic fault slip and earthquake prediction, J. Geophys. Res., in press, 1978a.

Dieterich, J. H., Time-dependent friction and the mechanics of stick-slip, PAGEOPH, in press, 1978b.

Logan J. M., T. Iwasaki, M. Friedman, and S. Kling, Experimental investigation of sliding friction in multilithologic specimens, in: Engineering Geology Case History No. 9, p. 55-67, 1972.

Scholz, C. H., P. Molnar, and T. Johnson, Detailed studies of frictional sliding of granite and implications for the earthquake mechanism, J. Geophys. Res., 77, p. 6392-6406, 1972.

FIGURE CAPTIONS

Figure 1. Schematic diagram of the experiment by Dieterich (1978a). Arrows indicate loading directions.

Figure 2. Shear strain vs. time record from Dieterich (1978a). Numbers refer to the strain gages shown in Fig. 1. Arrows mark the beginning of slip at each gage.

Figure 3. Finite element model of the experiment shown in Fig. 1. Arrows indicate the direction of motion of the rigid boundaries that distort the springs and load the block.

Figure 4. Computed stress vs. time for elements adjacent to the sliding surface. Shear stress prior to slip is homogeneous with respect to the frictional strength.

Figure 5. Computed stress vs. time for elements adjacent to the sliding surface showing unilateral propagation of stable slip prior to instability. The difference between frictional strength and shear stress prior to slip increases linearly along the surface.

Figure 6. Computed shear stress vs. time for elements adjacent to the siding surface showing bilateral propagation of stable slip prior to instability.

Figure 7. Displacement vs. time for the center nodes in the simulations of Fig. 5 and Fig. 6 (A and B, respectively). The arrows mark the breakout of slip at the end of the sliding surface.

Figure 8. Computed shear stress vs. time for the elements adjacent to the sliding surface. This simulation is identical to the simulation of Fig. 6 except the ends of the sliding surface are pinned. Only the first third of the simulation is shown.

Figure 9. Computed shear stress vs. time for elements adjacent to the sliding surface. The conditions for this simulation are identical to Fig. 6 except $d_c = 5 \times 10^{-3}$ cm instead of 5×10^{-4} cm used for Fig. 6.

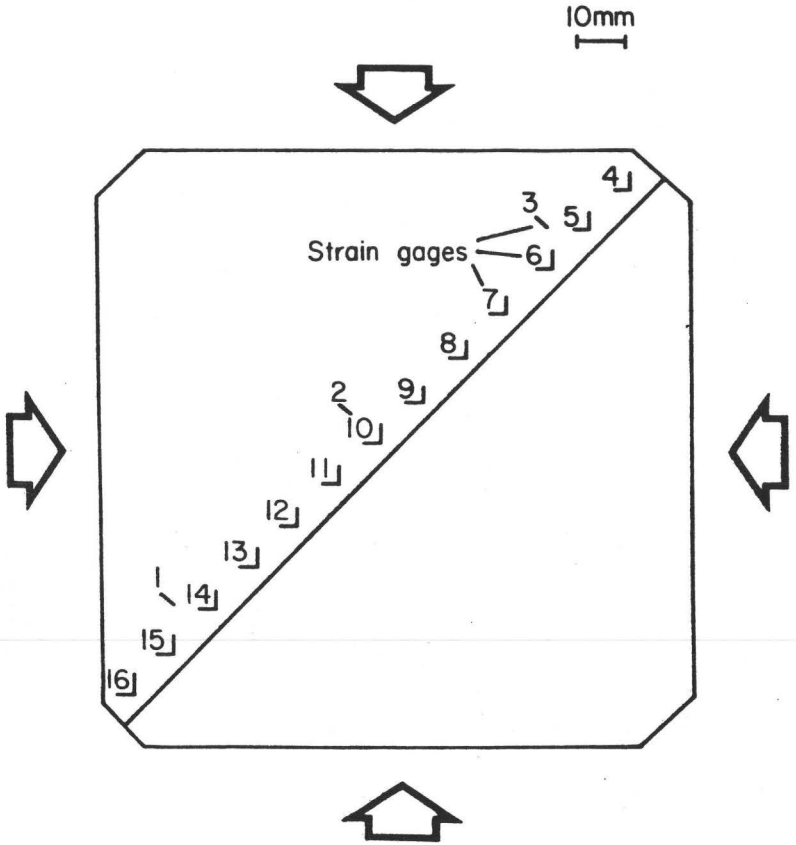


Figure 1

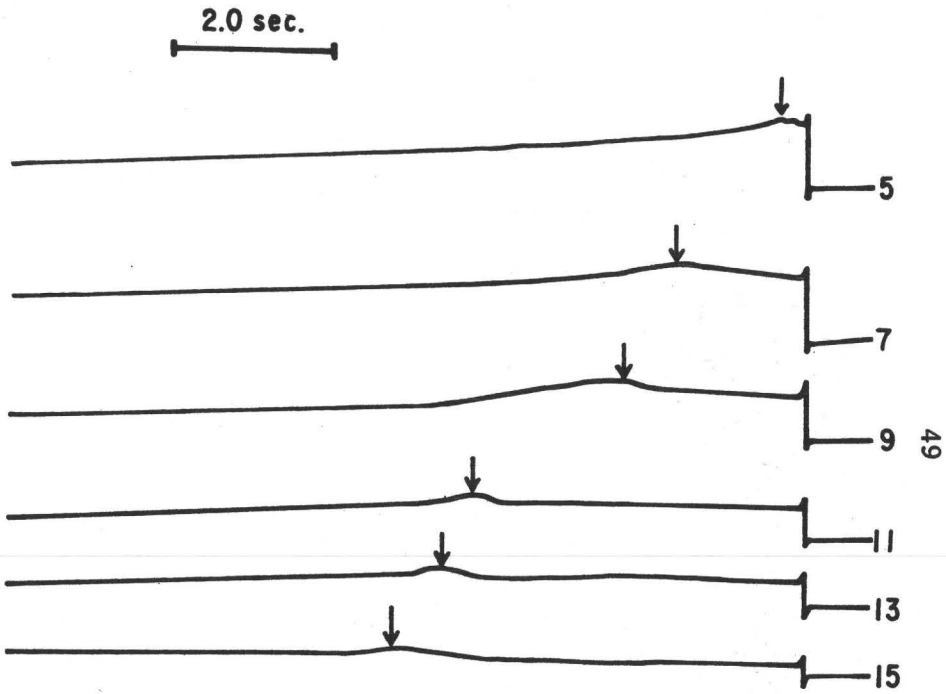


Figure 2

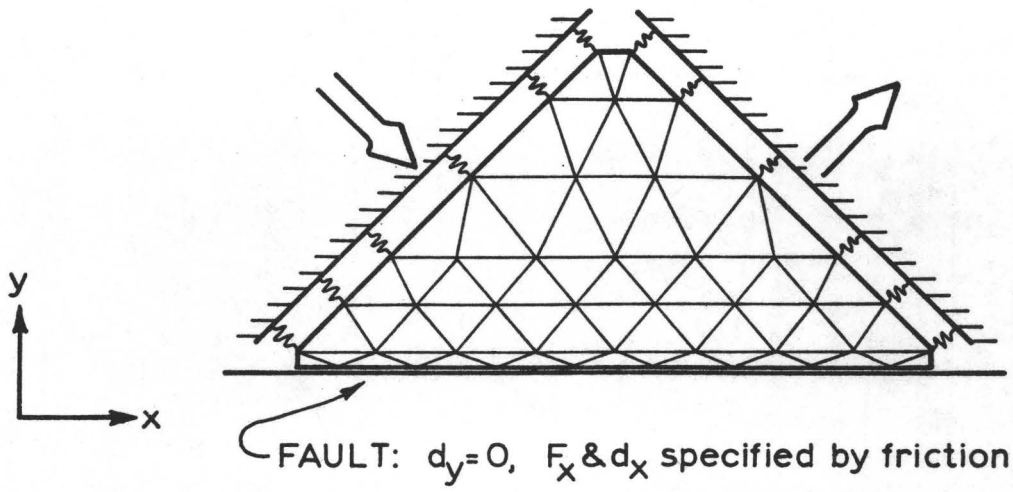


Figure 3

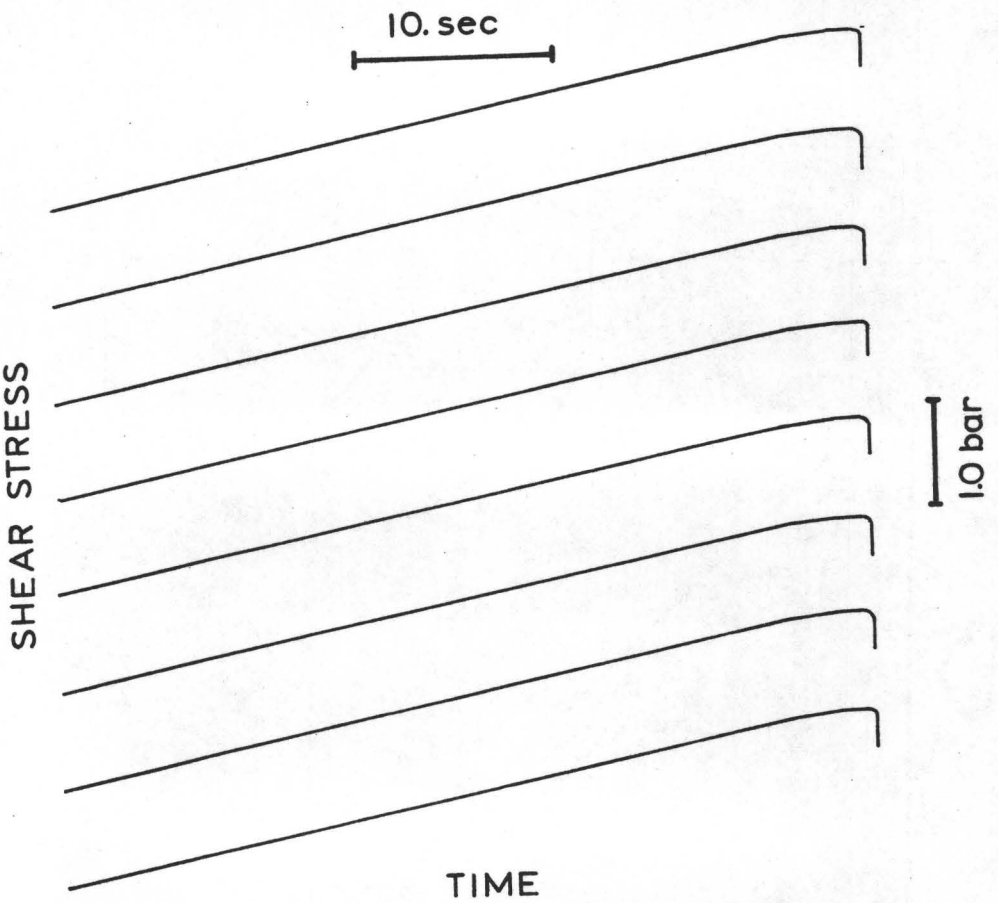


Figure 4

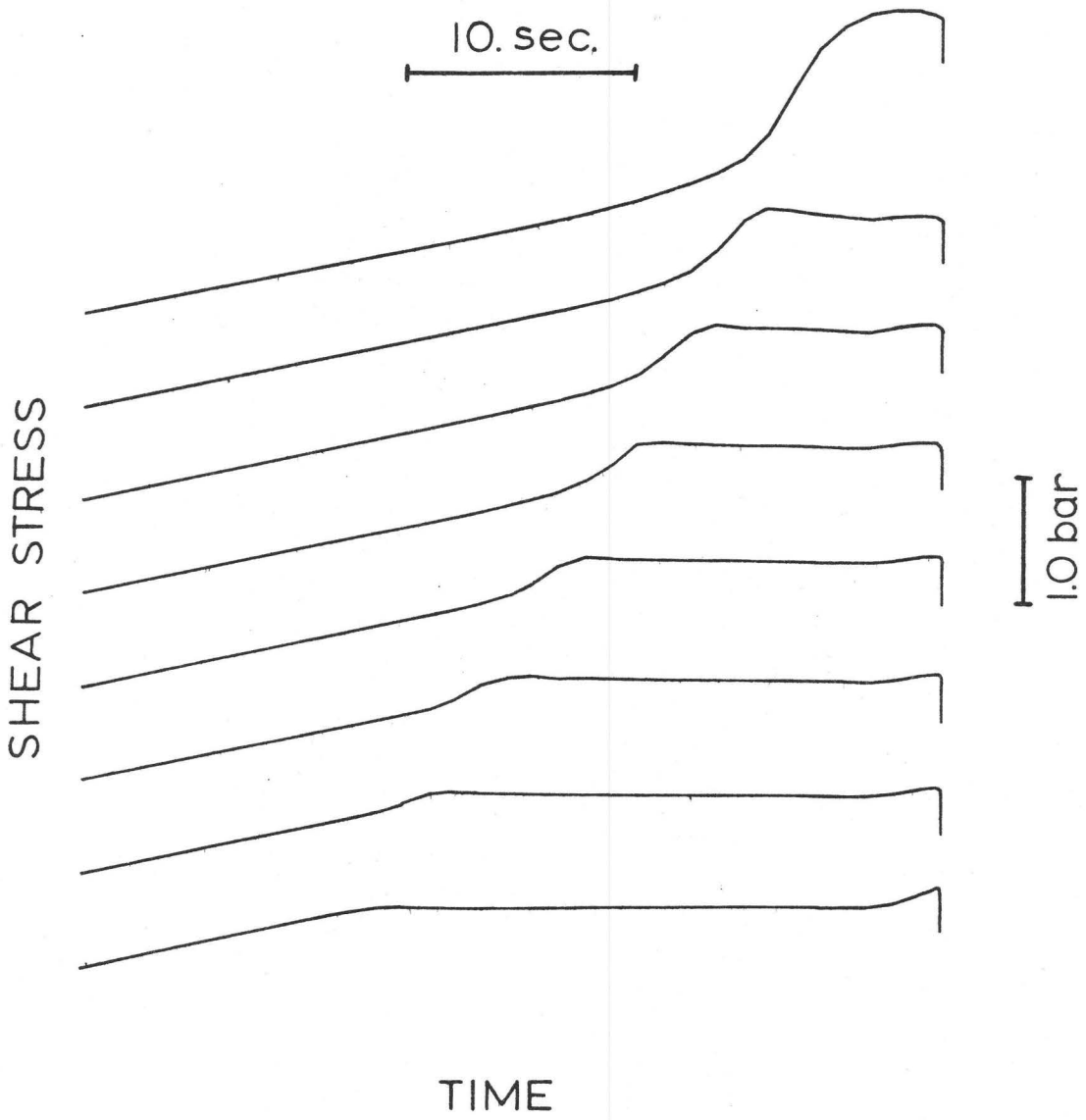


Figure 5

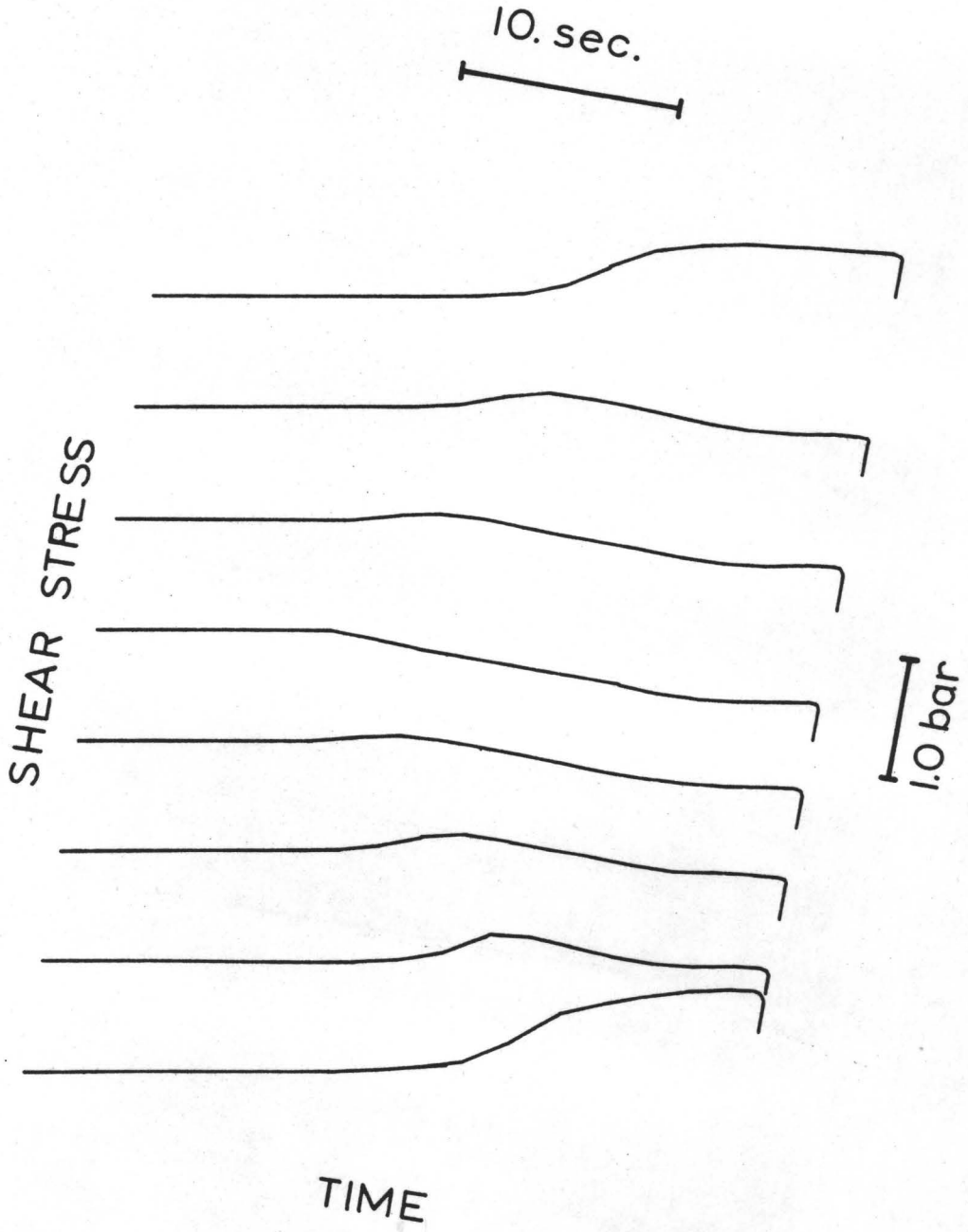


Figure 6

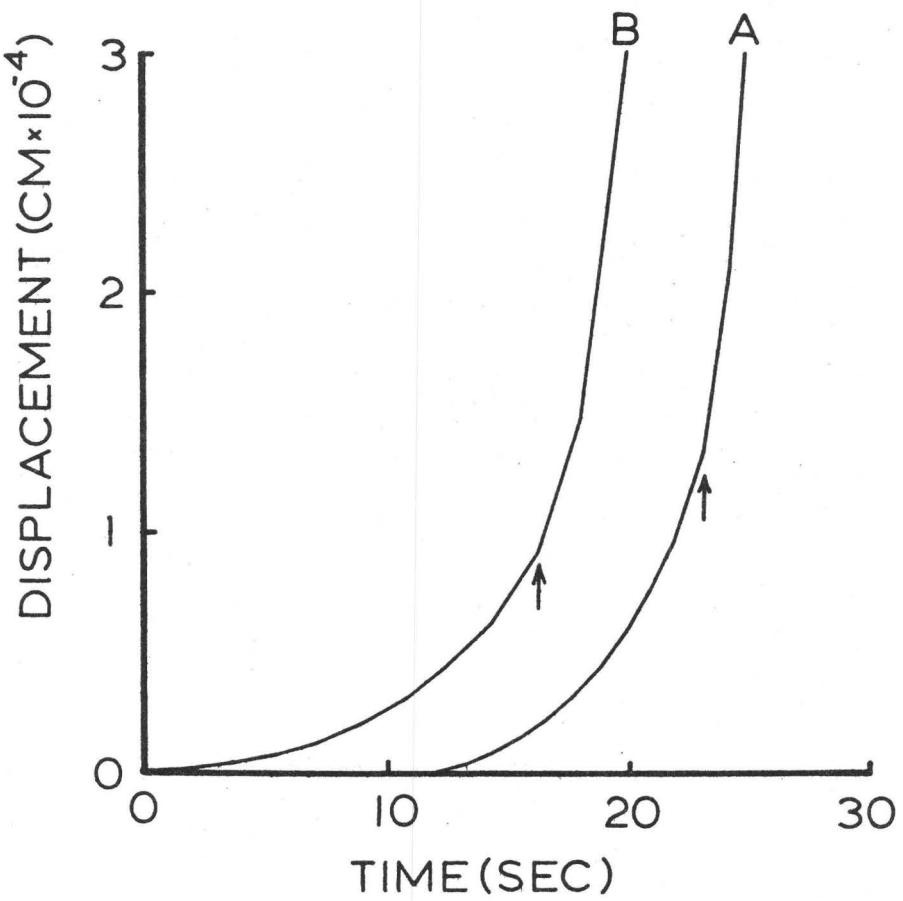


Figure 7

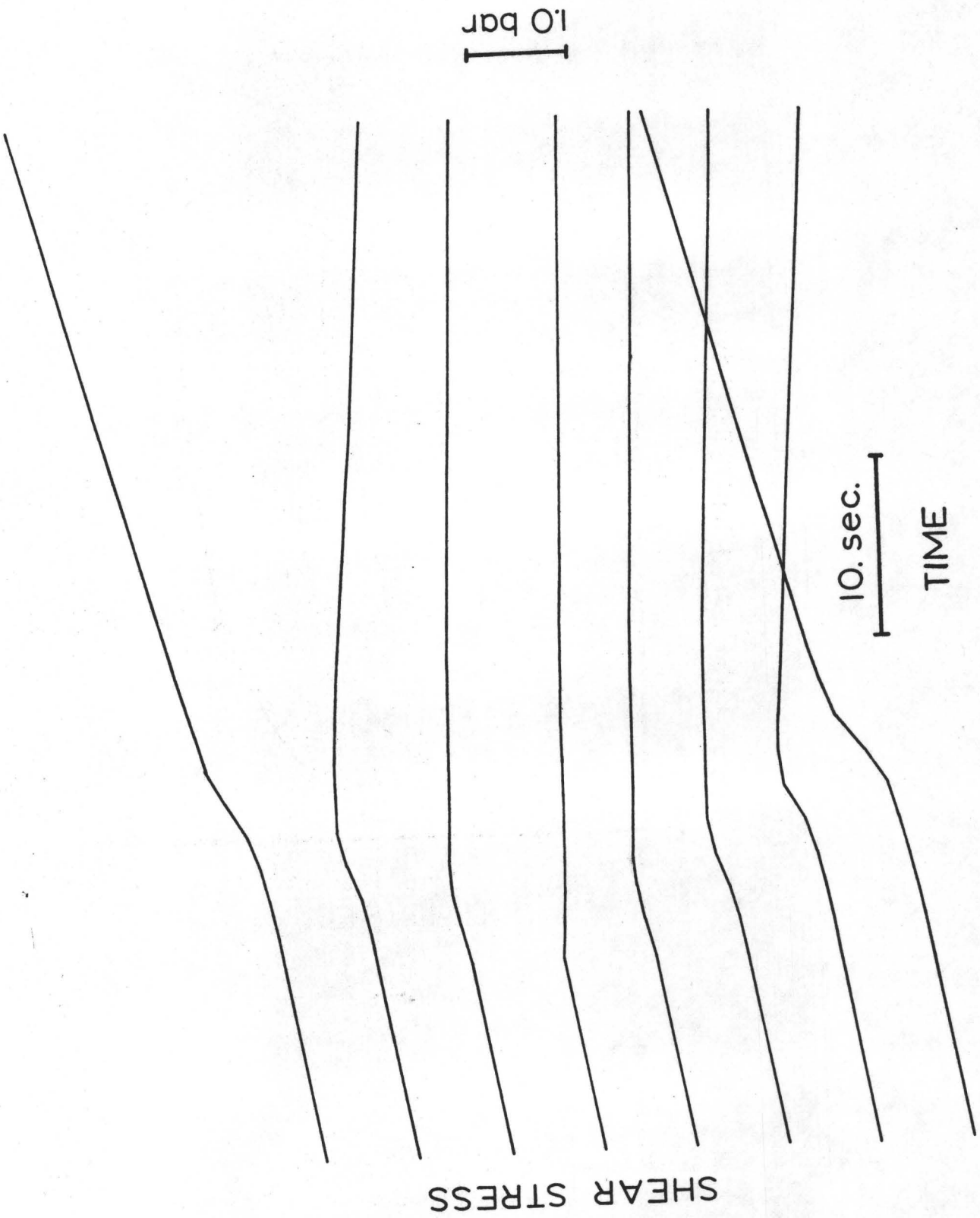


Figure 8

Figure 9

THE MECHANICS OF DYNAMIC SHEAR CRACK PROPAGATION

L. B. Freund
Division of Engineering, Brown University, Providence, RI 02912

INTRODUCTION

The mechanical modeling of an earthquake source has been a research objective in seismology for many years, and a recent trend in this area has been to view the source as a dynamically extending shear crack. In pre-cracked laboratory specimens, shear cracks often tend to extend in a direction oblique to the initial crack direction under stress. On the other hand, it is almost universally assumed that shear cracks serving as earthquake source models extend as planar cracks, and this assumption appears to be consistent with observation. Possible reasons for the planar growth are that a pre-existing fault provides a weakened path which is preferential for crack extension and that the confining pressure reduces the effect of tensile stresses near the crack tip which might otherwise lead to oblique crack growth.

The emphasis here is on basic concepts which seem to be important in dynamic shear crack analysis, and methods of analysis for specific problems are not considered in detail. A number of significant contributions which have concentrated on the dynamic shear fracture process have appeared in the literature in recent years, including those by Burridge [1], Burridge and Halliday [2], Fossum and Freund [3], Husseini, et al. [4], Kostrov [5,6], Kostrov and Nikitin [7], Richards [8], and Weertman [9] on analytical solutions to particular problems, and those by Andrews [10,11], Das and Aki [12], and Ida [13] on numerical solutions to particular problems. The present discussion is based primarily on these contributions, as well as on the results of research on dynamic tensile fracture which were recently reviewed by Freund [14]. It should also be noted that a number of signi-

ficant contributions which have concentrated on the stress wave radiation from a propagating shear crack have also appeared, including those by Brune [15], Husseini and Randall [16], Madariaga [17] and Richards [18].

SOME GENERAL RESULTS FOR PROPAGATING SHEAR CRACKS

Two general concepts which have played a major and fundamental role in the study of dynamic crack propagation are those of dynamic stress intensity factor and dynamic energy release rate. In this section, these concepts are given mathematical definitions in terms of the elastodynamic stress and deformation fields which prevail in a body of isotropic elastic material during crack propagation.

Stress and Particle Velocity Near a Crack Tip

A number of analytical solutions for the propagation of a sharp crack through a plane elastic solid are available. It has been observed for some time that the dependence of the stress field on spatial coordinates local to the crack tip is common to all solutions. For running cracks the spatial distribution is dependent on the speed of crack propagation, and it reduces to the appropriate expression for the stationary crack when the crack speed is set equal to zero. It can be demonstrated that this common spatial dependent of the near tip elastic field for running crack solutions is a general results, independent of the configuration of the body and the details of the loading system. The only quantity which varies from one specific problem to another is a time-dependent scalar multiple of the universal spatial dependence. The corresponding result for elastostatic fracture mechanics was first presented by Irwin [19] and Williams [20]. The general result for rapid propagation of a mode II shear crack may be derived by following the asymptotic analysis of Freund and Clifton [21], who presented the corresponding results for dynamic propagation of a mode I tensile crack at sub-sonic speeds.

A Cartesian coordinate system is oriented in the body in such a way that the particle displacement is in the x,y-plane; see Fig. 1. The region of the plane

occupied by the body is denoted by D and the outer boundary by S. The inner boundary of D consists of the crack faces, except near the crack ends where the inner boundary is augmented by small loops surrounding the crack tips. The loops have arbitrary shape, but they are fixed with respect to the moving crack tips. The outer boundary S is subjected to traction boundary conditions, displacement boundary conditions, or some suitable combination of both. Tractions may be acting on the crack faces, but they are assumed to be bounded in magnitude near the crack tips, and no body forces are acting. For simplicity, it is assumed that conditions are such that crack extension occurs in mode II, the plane strain shearing mode.

Attention is directed to the crack tip surrounded by the loop L, and a local Cartesian coordinate system (ξ,η) and a local polar coordinate system (r, θ), both of which move with the crack tip, are introduced as shown in Fig. 1. The crack is assumed to be extending in its own plane, and the instantaneous rate of extension v is any continuously varying function of time. With the requirement that the local energy density must be integrable, the following general result may be established by application of the asymptotic method described in [21]: For all plane elastodynamic solutions for running mode II cracks the shear stress component σ_{xy} and the particle velocity component $\dot{u}_x$ are given by

$$\sigma_{xy} = \frac{K_2(t)}{\sqrt{2\pi}\, R(v)} \left\{ 4\alpha_\ell \alpha_s \frac{\cos(\theta_\ell/2)}{r_\ell^{1/2}} - (1+\alpha_s^2)^2 \frac{\cos(\theta_s/2)}{r_s^{1/2}} \right\} \tag{1}$$

$$\dot{u}_x = \frac{v\alpha_s K_2(t)}{\mu R(v)\sqrt{2\pi}} \left\{ \frac{\sin(\theta_\ell/2)}{r_\ell^{1/2}} - (1+\alpha_s^2) \frac{\sin(\theta_s/2)}{r_s^{1/2}} \right\} \tag{2}$$

for $0 < v < v_s$, and by

$$\sigma_{xy} = \frac{K_2^*(t)}{\sqrt{2\pi}} \left\{ \frac{\cos(m\theta_\ell)}{r_\ell^m} + \frac{H(-\xi-\beta_s|\eta|)\sin(m\pi)}{(-\xi-\beta_s|\eta|)^m \tan m} \right\} \tag{3}$$

$$\dot{u}_x = \frac{vK_2^*(t)}{2\mu\alpha_\ell\sqrt{2\pi}} \left\{ \frac{\sin(m\theta_\ell)}{r_\ell^m} - 2\alpha_\ell\beta_s \frac{H(-\xi-\beta_s|\eta|)\sin(\pi m)}{(-\xi-\beta_s|\eta|)^m \tan m} \right\} \tag{4}$$

for $v_s < v < v_\ell$, as $r \to 0$. The subscripts ℓ and s refer the longitudinal and shear wave speeds v_ℓ and v_s, μ is the elastic shear modulus, and

$$\alpha_\ell = (1 - v^2/v_\ell^2)^{1/2} \quad , \quad r_\ell e^{i\theta\ell} = \xi + i\alpha_\ell\eta \tag{5}$$

$$\alpha_s = (1 - v^2/v_s^2)^{1/2} \quad , \quad r_s e^{i\theta s} = \xi + i\alpha_s\eta \tag{6}$$

$$\beta_s = (v^2/v_s^2 - 1)^{1/2} \quad , \quad \pi m = \tan^{-1}\left[4\alpha_\ell\beta_s/(1+\alpha_s^2)^2\right] \tag{7}$$

$$R(v) = 4\alpha_\ell\alpha_s - (1+\alpha_s^2)^2 \tag{8}$$

$$H(t) = \begin{cases} 1 & t > 0 \\ 0 & t < 0 \end{cases} \tag{9}$$

Similar expressions may be written for the local crack tip variation of other components of stress and particle velocity for both crack velocity ranges. However, the above expressions are typical and perhaps are the most important for discussing mode II shear fractures, and the expressions for other components are not included here.

The expression (1) has been normalized with respect to the relation

$$K_2(t) = \lim_{r\to 0} (2\pi r)^{1/2} \sigma_{xy}(r, 0, t) \tag{10}$$

The time dependent scalar $K_2(t)$ is then the mode II stress intensity factor of elastic fracture mechanics. For crack tip speeds in the range $0 < v < v_s$, the local stress and velocity vary as the inverse square root of distance from

the crack tip. The algebraic sign of the coefficient of the square root singular factor depends on whether the crack speed is less than or greater than the Rayleigh surface wave speed v_r. The function $R(v)$, which appears in (1, 2) and which is defined in (8), is the so-called Rayleigh wave function. This function has the properties that $R(\pm v_r) = 0$, $R(v) > 0$ for $0 < v < v_r$, and $R(v) < 0$ for $v_r < v < v_s$. Examination of (1, 2) leads to the conclusion that shear traction on the prospective fracture plane $\theta = 0^{\pm}$ and particle velocity on the fracture surface $\theta = \pm \pi$ have opposite algebraic sign if $0 < v < v_r$ and the same sign if $v_r < v < v_s$. This observation has important implications in considering energy fluxes associated with crack growth. It should also be noted that the expression (1) represents only the dominant singular term in the expansion of the elastic field about the crack tip, and the expansion could be continued. For example, if the crack faces are subjected to tractions such that $\sigma_{xy} \to \tau_o$ as $\xi \to 0^-$ and $\eta = \pm 0$, then the "order one" term in the expansion is simply τ_o. The next term in the expansion is proportional to $\sqrt{r}$, and so on. Of course, if $K_2(t) \equiv 0$ then the higher order terms in the expansion take on added importance.

The expression (3) has been normalized with respect to the relation

$$K_2^*(t) = \lim_{r \to 0} \sqrt{2\pi}\, r^m\, \sigma_{xy}\,(r, 0, t) \qquad (11)$$

so that the time dependent scalar $K_2^*(t)$ is also a mode II stress intensity factor, although the singularity in stress for $v_s < v < v_\ell$ is in general weaker then the inverse square root singularity which arises for $0 < v < v_s$. As can be seen from (7), the exponent m which appears in (3, 4) varies continuously from $m = 0$ at $v = v_s$, up to a maximum value $m = 1/2$ at $v = v_s \sqrt{2}$, and back to $m = 0$ at $v = v_\ell$. It should be noted that the shear wave contribution to the expressions (3, 4) is very different from that in (1, 2). Because the crack tip

speed is greater than the shear wave speed, no shear waves radiated from the crack can propagate ahead of the running crack tip. Instead, shear wave motion can exist only behind discrete wavefronts which appear in the form of Mach waves trailing from the extending crack. In the local (ξ, η) coordinate system, these Mach wavefronts coincide with the lines $\xi + \beta_s|\eta| = 0$. In the solution of a particular mode II crack propagation problem Burridge [1] noted that the stress singularity for crack speeds in the range $v_s < v < v_\ell$ has the form shown in (3). It is concluded from an application of the asymptotic method of [21] that this is a general result, common to all mode II crack propagation problems.

Finally, the local crack tip field for dynamic propagation of a crack in the anti-plane shear mode, or mode III, is included. Referring once again to Fig. 1, the only nonzero component of displacement is in the direction normal to the x,y-plane which is, say, the z-direction. The shear stress component σ_{yz} and the particle velocity $\dot{u}_z$ are

$$\sigma_{yz} = \frac{K_3(t)}{\sqrt{2\pi r_s}} \cos(\theta_s/2) \tag{12}$$

$$\dot{u}_z = \frac{vK_3(t)}{\mu\alpha_s\sqrt{2\pi r_s}} \sin(\theta_s/2) \tag{13}$$

as $r \to 0$ for any crack speed in the range $0 < v < v_s$. The notation is the same as that for mode II, except for the appearance of the mode III stress intensity factor,

$$K_3(t) = \lim_{r\to 0} (2\pi r)^{1/2} \sigma_{yz}(r, 0, t) \quad . \tag{14}$$

As can be seen from (12, 13) the local stress and velocity fields have the characteristic square root singular dependence on distance from the moving crack tip.

The Dynamic Energy Release Rate

Consider once again the two-dimensional body of linear elastic material con-

taining a crack shown in Fig. 1. For present purposes, it is not necessary that the crack be mathematically sharp but it is assumed that the loop L originates on one crack surface, that it terminates on the opposite crack surface, and that it completely surrounds the crack tip zone, whether this zone is a point or a diffuse region. As before, the loop is considered to be fixed with respect to the crack tip, which is moving at an arbitrary rate. The figure represents the body at a fixed instant of time. As was shown in [14], a general expression for energy absorption rate may be derived in terms of the local crack tip stress and velocity fields without specification of a particular mode of crack propagation.

As the crack tip surrounded by L moves under the action of the applied loading, mechanical energy flows through L at a rate which will be denoted by F. This energy flux F may be computed in terms of the elastic field of the body by application of an overall energy rate balance. Let P, T and U denote the rate of work of the applied tractions, the total kinetic energy of the material in D, and the total strain energy of the material in D, respectively. Then the flux of energy through L is equal to the difference between the rate of work of the applied loads and the rate of increase of internal energy in D, that is,

$$F = P - (\dot{T} + \dot{U}) \tag{15}$$

If the terms on the right side of (15) are expressed in terms of the instantaneous stress and velocity fields in the body according to standard definitions and the divergence theorem is applied, then the following main result is obtained

$$F = \int_L \left[\sigma_{ij} n_j \dot{u}_i + \frac{1}{2}\left(\sigma_{ij} u_{i,j} + \rho \dot{u}_i \dot{u}_i \right) v n_x \right] dL \tag{16}$$

where σ_{ij} and $\dot{u}_i$ are the stress and velocity components, n_i is the normal to L pointing away from the crack tip, and ρ is the material mass density.

The interpretation of the terms in (16) is straightforward. The first term is the rate of work of the material outside of L on the material inside L, that is, it is the sum along L of the inner product of traction and particle velocity and, if L were fixed with respect to the material, this would be the only contribution to the energy flux. Because L is moving through the material, however, material particles cross L. Associated with each material particle is an energy density, and the second term in (16) represents the contribution to the energy flux due to the flux of energy-bearing material particles.

Several remarks should perhaps be added here concerning previous work on the energy flux (16) or quantities related to it. First of all, in the original derivation of (16) in [22], only the limiting case in which L is shrunk onto the crack tip was considered. It seems that the expression (16) provides a basis for unification of several apparently diverse results on dynamic energy release rate, however, and consideration of limiting cases of L is postponed. Secondly, F does reduce to the appropriate form of the path-independent J-integral [23] for the special case of quasi-static deformation. In general, however, the value of F will indeed depend on the path employed to evaluate it [22,24]. The fact that F is path-independent under the special condition that the complete elastic field is constant with respect to an observer moving with the crack tip has been observed by Sih [25]. Finally, it is noted that the result (16) is also valid for three dimensional crack growth provided that L is understood to be a tube of fixed cross-section surrounding the crack edge and moving with it, and vn_x is the normal speed of the tube at each point on its surface.

Sharp crack tip. For extension of a sharp crack under plane conditions, the rate of energy being supplied to the growing crack is given by F in (16) in the limit as L is shrunk onto the moving crack tip. To perform this calculation, F may be evaluated for a path within the region where the elastic field is ade-

quately described by the results (1)-(4) and (12)-(13) prior to taking the limit. Because the near-tip fields have universal spatial dependence, the integral defining F may be evaluated and an expression for F in terms of the dynamic stress intensity factors may be obtained. Furthermore, the near-tip fields are steady with respect to an observer moving with the crack tip, so that the value of F which is calculated is independent of the shape of L within the crack tip region [14]. Some of the details of calculating F in terms of the dynamic stress intensity factor for extension of a mode I crack are included in [22]. For combined mode II and mode III crack growth at a speed in the range $0 < v < v_s$, the energy flux into the crack tip is given by

$$F = \frac{(1-\nu^2)}{E} \frac{v^3 \alpha_s}{v_s^2 R} K_2^2 + \frac{v}{2\mu\alpha_s} K_3^2 \tag{17}$$

where ν and E are Poisson's ratio and Young's modulus, respectively, of the material. The energy flux F is related to the energy released per unit crack advance, the dynamic energy release rate G, simply by $F = vG$. Equation (17) thus represents a relationship between the dynamic stress intensity factors and the dynamic energy release rate which is valid for all loading conditions and all geometrical configurations. In this sense, it is the analogue for dynamic shear cracks of Irwin's well-known relationship between energy release rate and the mode I stress intensity factor. It is noted from (17) that if $K_2 \neq 0$ then energy is absorbed by a mode II crack tip for speeds in the range $0 < v < v_r$ and energy is radiated from the tip for speeds in the range $v_r < v < v_s$. For a mode III crack with $K_3 \neq 0$, energy is absorbed into the tip for all speeds in the range $0 < v < v_s$.

For propagation of a mode II crack with speed in the range $v_s < v < v_\ell$, the crack tip singularity in stress and velocity is weaker than the inverse square root singularity, unless $v = v_s\sqrt{2}$, and therefore $F = 0$ for all speeds in this

range except for $v = v_s \sqrt{2}$. For this particular speed, the shear waves trailing the crack tip are absent and the energy flux into the tip due to longitudinal wave motion is

$$F = v(K_2^*)^2/4\mu\alpha_\ell \tag{18}$$

It must be noted that a stress distribution with an infinite singularity is clearly a mathematical idealization, in that no real material can actually support such a stress. The usual rationalization for admitting the singular stress distribution, the strength of which is measured by the stress intensity factor, is based on the concept of small scale yielding [23]. It is thus assumed that in the immediate vicinity of the crack tip the potentially large stresses are relieved by some nonlinear process in a region whose dimensions are small compared to crack length and body dimensions. It is assumed further that the stress distribution in the elastic material adjacent to the small zone is adequately described by the dominant singular term in the elasticity solution. Under small scale yielding conditions, the stress intensity factor may be considered to be a one-parameter measure of the amplitude of the stress which is being applied to the material in the crack tip region. The stress intensity factor approach circumvents consideration of how the material in the crack tip region actually responds to the applied stress.

Cohesive zone crack tip model. In order to avoid infinitely large stresses on the fracture plane, a number of models involving a one-dimensional cohesive zone extending ahead of the physical crack tip in the fracture plane have been proposed. In each case, the model is analyzed by making the crack longer by an amount R_c, the cohesive zone size (see Fig. 2a), with the cohesive stresses in this zone acting so as to restrain crack opening (in mode I) or crack face sliding (in modes II or III). The size of the cohesive zone is chosen so that the net stress intensity factor due to both the applied loads and the cohesive stress is zero. Cohesive zone models have been proposed for a variety of physical processes, for example,

for pure cleavage tensile fracture by Barenblatt, Salganik and Cherepanov [26] and for fully developed plane stress plastic yielding at a tensile crack tip by Dugdale [27]. A cohesive zone model which seems to be quite realistic for a number of geophysical applications involving frictional sliding is the so-called slip weakening model, according to which it is assumed that slip will commence at a point on a slip plane when the local shear stress on the slip plane is elevated to a certain level σ_s, that the shear stress required to sustain slip is reduced as the amount of slip δ is increased to some critical amount δ_c, and that the shear stress required to sustain slipping beyond the critical amount of slip is σ_d. The shear stress magnitudes σ_s and σ_d are usually associated with the static and dynamic frictional resistance of an interface to relative slip. The slip weakening model has been discussed by Palmer and Rice [28], Ida [29], and Andrews [10], among others.

In the slip weakening model for a mode II crack, the local cohesive stress σ depends on the local slip $\delta(x, t) = u_x(x, 0^+, t) - u_x(x, 0^-, t)$ as shown in Fig. 2b. The appropriate path L for computing energy flux into the cohesive zone during crack propagation is shown in Fig. 2a. Because $n_x = 0$ for all points on L, the expression for F in terms of the cohesive stress and the slip reduces to

$$F = \int_{\substack{\text{cohesive} \\ \text{zone}}} \sigma(\delta) \frac{\partial \delta}{\partial t} \, dx \quad . \tag{19}$$

The slip may be viewed as a function of the crack tip coordinate ξ and time, in which case $\partial\delta(x, t)/\partial t$ may be replaced by $- v\partial\delta(\xi, t)/\partial\xi + \partial\delta(\xi, t)/\partial t$. The energy flux (13) then becomes the sum of two terms

$$F = v \int_0^{\delta_c} \sigma(\delta) d\delta + \int_0^{R_c} \sigma(\delta) \frac{\partial \delta}{\partial t} \, d\xi \tag{20}$$

where $\delta_c = \delta(0, t)$ is the crack tip opening displacement. The expression (20)

is for the total energy flux into the cohesive zone during crack propagation. The first term represents the rate at which energy must be supplied per unit crack surface area to completely overcome the cohesive stress, i.e., to produce a total slip of an amount δ_c. The second term represents the rate of work which must be supplied to change the amount of slip within the cohesive zone and to change zone size during nonsteady crack propagation. If the elastic field of the propagating crack is constant as seen by an observer moving with the crack tip, then δ depends on t only through ξ, that is $\partial\delta(\xi, t)/\partial t = 0$ and the second term in (20) is zero.

Assuming such steady-state crack propagation and assuming a linear relationship between cohesive stress and slip, $\sigma(\delta) = \sigma_s - (\sigma_s - \sigma_d)\delta/\delta_c$, the energy flow into the cohesive zone per unit crack advance is $G = (\sigma_s + \sigma_d)\delta_c/2$. It should be noted that most authors view the excess of σ above σ_d as the cohesive stress and in this case the energy absorbed per unit crack advance in overcoming cohesion is $G = (\sigma_s - \sigma_d)\delta_c/2$. The energy release rate expressions for the slip weakening model applied to mode III crack extension are identical to those for mode II. It is noted that if conditions for small scale yielding are met, that is, if the cohesive zone size R_c is much less than all other physical dimensions, then a relationship between the stress intensity factor and the physical parameters of the cohesive zone is obtained by equating the energy flow into the singular crack tip to the energy absorbed in the cohesive zone.

The results reported in this section provide a number of general relationships among parameters which characterize the mechanical conditions which prevail near the tip of a shear crack during rapid crack propagation. The actual dependence of any of these parameters on applied loading and geometrical configuration in any specific model of a dynamic shear rupture process can only be established by analyzing the corresponding boundary value problem. Although the collection of

crack propagation problems which have been studied is rich in variety from the analytical, computational and physical points of view, only a few analytical models are considered in the following sections. These particular models have been chosen because they address questions which appear to be important at the present time in the modeling of the dynamic crustal faulting process, they can be analyzed without undue complexity, and the results seem to provide some insight into the process of dynamic shear crack propagation.

STEADY-STATE CRACK PROPAGATION

The simplest dynamic crack propagation problems which can be analyzed are those of the steady-state type, that is, the crack propagation speed is assumed to be constant and the complete stress and deformation fields are taken to be fixed as seen by an observer moving with the crack tip. This idealization is clearly inadequate when considering the abrupt initiation or arrest of a dynamic fracture, but it may be quite acceptable for describing processes for which the duration of the acceleration and deceleration phases is short compared to the total duration of the process. Furthermore, in the present context, the steady-state assumption makes it possible to demonstrate in a simple way some of the main features of dynamic shear crack propagation which may be relevant in earthquake source modeling. Attention is directed toward several specific interlated features in the steady-state problems to be discussed below. Among these features is the phenomenon of fault plane healing. That is, analysis of a simple steady state model allows consideration not only of the onset of slipping between the fault faces at the leading edge of the dynamically extending fault but also of the cessation of slipping between the faces, or fault healing, at some distance behind the leading edge. A second specific aspect concerns the concept of the stress drop associated with fault extension. The inherent inhomogeneity of the stress distribution in the vicinity of a crack leads to the identification of several

significant stress magnitudes in the description of the dynamic crack growth process. Thus, several stress differences can be identified as stress drops or, in other words, the definition of stress drop appears to be ambiguous. A third feature of the dynamic faulting process which can be considered by means of the analysis of steady-state problems is the influence on the process of a change in remote loading conditions. An almost universal assumption in dynamic crack propagation analysis is that the process is driven by a quasi-static, spatially uniform stress which is applied far from the process region. As will be shown below, some interesting results may be deduced quite simply from a steady-state crack model involving remote displacement conditions. In contrast to the remotely applied stress condition, which represents dead-weight or perfectly soft loading, the remotely applied displacement condition represents the opposite extreme of perfectly stiff loading. These features will be discussed in greater detail in the remainder of this section on the basis of the analysis of a few particular problems. An approach whereby complete stress and deformation fields are established is briefly described first, and then an indirect approach based on an energy integral is considered.

A stress analysis approach. As typical of the problems of the steady-state type which can be analyzed by standard methods is the anti-plane shear mode III problem represented in Fig. 3a. Under the action of a uniform remote shear stress $\sigma_{yz} = \sigma_\infty$, the mode III crack grows in the x-direction at speed v. The x,y coordinate system is fixed with respect to the moving crack. At the leading edge of the crack $x = a$, a singularity in $\sigma_{yz}(x, 0)$ will be admitted, and it is assumed that the strength of this singularity is governed by the cohesive strength of the fault plane $y = 0$. Slipping begins at $x = a - 0$, $y = 0$ and it continues throughout the interval $-a < x < a$, $y = 0$. Relative slipping of the crack faces is resisted by a uniform frictional stress $\sigma_{yz}(x, 0) = \sigma_d$ which is less than σ_∞ and which might be due to Coulomb friction arising from a uniform

compressive stress in the y-direction. At $x = -a, y = 0$ the relative particle velocity across the fault reduces to zero, and for $x < -a, y = 0$ no further slipping occurs. The total amount of slip between the fault surfaces will be denoted by δ_t, and the slipping process is assumed to terminate smoothly with bounded stresses at the healing point $x = -a, y = 0$.

An analytical technique of general applicability in problems of this type is that based on the theory of analytic functions of a complex variable. A general discussion of this method for steady-state plane strain elastodynamic problems has been given by Radok [30] and several specific applications to mode I fracture propagation are cited in [14]. An equivalent formulation for anti-plane shear problems is straight-forward. It is easily verified that the relevant field equations are satisfied if stress and particle velocity have the representations

$$\sigma_{xz} - i\sigma_{yz}/\alpha_s = g(\zeta) \quad , \quad \mu\dot{u}_z = -v \operatorname{Re} g(\zeta) \tag{21}$$

where $g(\zeta)$ is a function of the complex variable $\zeta = x + i\alpha_s y$ which is analytic in the complex ζ-plane cut along $-a < x < a, y = 0$. The determination of stress and particle velocity then reduces to the determination of $g(\zeta)$ according to the powerful methods of analytic function theory; cf. [23]. For the problem at hand, the result is

$$g(\zeta) = i(\sigma_\infty - \sigma_\alpha)\alpha_s^{-1}\left(1 - \zeta/\sqrt{\zeta^2-a^2}\right) - i\mu\delta_t/2\pi\sqrt{\zeta^2-a^2} - i\sigma_\infty\alpha_s^{-1} \tag{22}$$

from which the stress components and particle velocity may be extracted according to (21). Graphs of $\sigma_{yz}(x, 0)$ and $\dot{u}_z(x, 0)$ are shown in Fig. 3b. The length of the slipping region is determined from the condition that stresses be bounded at $x = -a, y = 0$ as $a = \mu\alpha_s\delta_t/2\pi(\sigma_\infty - \sigma_d)$ and the stress intensity factor for the leading edge of the slipping zone is $K_3 = 2(\sigma_\infty - \sigma_d)\sqrt{\pi a}$. From (12) the energy release rate at the leading edge is then

$$G = (\sigma_\infty - \sigma_d)\,\delta_t \quad . \tag{23}$$

A stress difference which appears naturally in this analysis is $(\sigma_\infty - \sigma_d)$, and (23) gives a value for this stress difference as the cohesive energy density of the fault surface divided by the total slip displacement across the fault. This cohesive energy is a measure of the static frictional strength of the fault and it is usually assumed to be several orders of magnitude larger than the actual surface energy of the material which is typically about $1 J/m^2$.

If small scale yielding conditions prevail and if the interface failure at the crack tip occurs in the slip weakening mode represented in Fig. 2b, then G in (23) may be replaced by $\delta_c(\sigma_s - \sigma_d)/2$, where $\delta_c < \delta_t$. In this context, the stress σ_s is viewed as a measure of fault resistance to onset of slipping, the resistance to slipping decreases from σ_s as the amount of slip increases to δ_c, and slipping is resisted by the stress σ_d for greater amounts of slip. In this case, the relationship (23) may be rewritten as

$$2\delta_t/\delta_c = (\sigma_s - \sigma_d)/(\sigma_\infty - \sigma_d) \tag{24}$$

It is clear from (24) that a second stress difference $(\sigma_s - \sigma_d)$ has been introduced. Whereas the previously introduced stress difference $(\sigma_\infty - \sigma_d)$ represents the stress drop from the uniform remote tectonic stress σ_∞ to the frictional stress on the slipping part of the fault, the difference $(\sigma_s - \sigma_d)$ represents the stress drop from a point just ahead of the propagating edge of the slipping region of the fault to a point just behind the edge within the slipping region. It can be seen from (24) that the magnitudes of these two stress drops can be very different. For example, if δ_t is about 1 m and δ_c about 1 mm then the stress drop $(\sigma_s - \sigma_d)$ is three orders of magnitude larger than $(\sigma_\infty - \sigma_d)$.

In the case of mode II crack propagation, the same general analytical procedure may be applied although the detailed formulation is somewhat more complicated.

However, it turns out that certain key results for problems of this type, such as those in (23) and (24), may be extracted without actually solving the boundary value problems, that is, without determining the complete stress and deformation fields, as is discussed in the following subsection.

An energy integral approach. The analysis is based on the energy flux integral (16) for the special case of steady-state crack propagation in the x-direction at speed v, for which $\partial/\partial t = - v\partial/\partial x$. If the ratio F/v for a given contour L surrounding the crack tip is denoted by $E(v; L)$, then it is a simple matter to show that

$$E = \int_L \left[\frac{1}{2}(\sigma_{ij}u_{i,j} + \rho v^2 u_{i,x}u_{i,x})n_x - \sigma_{ij}n_j u_{i,x} \right] dL \qquad (25)$$

If L were a closed contour containing no singularities of the elastic field, then $E = 0$. This may be demonstrated by application of the divergence theorem to the integral in (25). A direct consequence of this result is that E is a path-independent integral. That is, the value of E is the same for all simple paths which originate at a particular point on one face of the crack, which completely surround the crack tip region, and which terminate at a particular point on the opposite face of the crack. If the crack faces are traction-free, then the integrand of (25) vanishes on the crack faces and the value of E is independent of path even if the comparison contours do not originate and terminate at the same points on opposite faces of the crack; cf. [23].

As a first example of the application of the path-independent integral (25), consider the anti-plane shear problem which was discussed in the previous subsection and which is represented in Fig. 3a. Suppose that a uniform stress $\sigma_{yz} = \sigma_\infty$ is subtracted from the stress state shown. The stresses and displacement gradients then decay as $(x^2 + y^2)^{-1/2}$ at large distances from the crack tip which implies that the value of E along any remote contour will be zero, $E_{remote} = 0$. If a

second path L is considered which originates on the lower crack surface at $x \to -\infty$, $y = -0$, which runs along the lower crack face to $x = a$, which surrounds the crack tip at $x = a$ and then runs along the upper crack face to $x \to -\infty$, $y = +0$, then the value of E for this path is $G + (\sigma_d - \sigma_\infty)\delta_t = E_{crack}$. Because the integral is path independent, $E_{remote} = E_{crack} = 0$ and the result (23) is reproduced. It is noteworthy that a cohesive zone representing a slip weakening region could be included explicitly with this approach and the result (24) could be derived directly without recourse to the small scale yielding assumption. Similar results for a variety of other problems concerned with steady-state crack propagation in unbounded bodies, such as that considered by Weertman [9], can be directly extracted in much the same manner.

The energy integral approach becomes particularly useful in certain situations where the complete stress and deformation fields cannot be determined by any of the standard analytical methods, but where enough information is available to compute the value of E. As typical of such a situation, consider the steady-state propagation of a mode II shear crack in a strip of width $2h$ as shown in Fig. 4. Under the action of applied displacements $u_x(x, \pm h) = \pm u_o$, $u_y(x, \pm h) = 0$, the mode II crack grows in the x-direction at speed v. The x,y coordinate system is fixed with respect to the moving crack. Although a singularity in the stress component σ_{xy} at the leading edge of the slipping region could be admitted just as in the above mode III problem a linear slip weakening cohesive zone is introduced at the outset instead. Thus, $\sigma_{xy}(a, 0) = \sigma_s$ with $\delta(a) = 0$, $\sigma_{xy}(0, 0) = \sigma_d$ with $\delta(0) = \delta_c$, and $\sigma_{xy}(-b, 0) = \sigma_d$ with $\delta(-b) = \delta_t > \delta_c$. Just as before, slipping begins on the fault plane at $x = a$, it continues throughout the interval $-b < x < a$, and it terminates with net displacement off-set of δ_t at $x = -b$. Relative slipping is resisted by the cohesive/frictional stress $\sigma_{xy}(x, 0) = \sigma_s - (\sigma_s - \sigma_d)\delta/\delta_c$ in the interval $0 < x < a$ and by the uniform frictional stress $\sigma_{xy}(x, 0) = \sigma_d$ in the interval $-b < x < 0$. The frictional stress might

be due to Coulomb friction arising from a superimposed uniform compressive stress in the y-direction. If such a stress were actually introduced by specifying the alternate boundary condition $u_y(x, \pm h) = \pm \bar{u}_y$, however, it would have no influence on the main results to be obtained in the subsequent discussion.

The unique feature of this particular problem is that it represents crack propagation under displacement control or very stiff loading conditions, in contrast to the mode III problem of crack propagation under stress control or very soft loading conditions which was considered previously. Displacement control conditions make possible the consideration of stress relief by crack extension, that is, the shear stress on the fault plane at some point far behind the slipping region, say $\sigma_- = \mu(u_o - \delta_t/2)/h$, will be less than the shear stress on the fault plane at some distance ahead of the slipping region, say $\sigma_+ = \mu u_o/h$. The relationship among the various stress magnitudes and slip magnitudes is obtained by direct application of the energy integral (25).

The two choices of contour L for evaluation of E are shown in Fig. 4, and the value of E along the outer contour, say E_{remote}, will be considered first. For points far ahead/behind the crack tip, the only nonzero stress component is $\sigma_{xy} = \sigma_\pm$ and the only nonzero displacement gradient component is $u_{x,y} = \sigma_\pm/\mu$, and these components are essentially uniform across the strip. For the portion of the remote contour along $y = \pm h$, the displacement components are uniform and $n_x = 0$ so that the integrand of (25) vanishes identically and there is no contribution to the value of E. It is thus clear that

$$\begin{aligned} E_{remote} &= (\sigma_+^2 - \sigma_-^2)h/\mu \\ &= (\sigma_+ + \sigma_-)\delta_t/2 \end{aligned} \tag{26}$$

As shown in Fig. 4, the other choice of contour L embraces the crack and $n_x = 0$

at all points of the inner contour. Then, according to (25), the value of E for the inner path is

$$E_{crack} = -\int_{-b}^{0} \sigma_d \frac{\partial \delta}{\partial x}\, dx - \int_{0}^{a} \sigma(\delta) \frac{\partial \delta}{\partial x}\, dx$$

$$= (\sigma_s - \sigma_d)\delta_c/2 + \sigma_d \delta_t \tag{27}$$

In view of the path-independence of E, $E_{remote} = E_{crack}$ or

$$\frac{\delta_t}{\delta_c} = \frac{(\sigma_s - \sigma_d)}{(\sigma_+ + \sigma_- - 2\sigma_d)} \tag{28}$$

This analysis introduces two additional stress differences which might be identified as stress drops, that is, $(\sigma_+ - \sigma_d)$ and $(\sigma_+ - \sigma_-)$. The second of these is particularly interesting because it satisfies the relationship

$$\sigma_+ - \sigma_- = \mu\delta_t/2h\ . \tag{29}$$

If h can be identified in some way with the total distance of travel of the fault edge, which is not an entirely unreasonable identification, then the right side of (29) coincides with one of the standard definitions of stress drop [31]. It is noteworthy that σ_+ and σ_- do not represent frictional properties of the fault. The magnitudes of σ_+ and σ_- are arbitrary except that $\sigma_+ + \sigma_- > 2\sigma_d$ and $\sigma_+ > \sigma_-$ for the process to occur at all. Unfortunately, the energy integral approach provides no apparent means for computing the physical dimensions a and b in Fig. 4.

The steady-state crack propagation models considered in this section are representative of problems in this class. There are, of course, many other steady-state crack propagation models involving other fault plane strength characterizations or other friction laws which might be profitably studied. The mode III

problem represented in Fig. 3a was chosen here because it is typical of the simplest steady-state problems which can be analyzed, and the mode II strip problem represented in Fig. 4 was chosen because it is typical of problems for which useful information can be extracted by means of the energy integral approach, even though the stress and deformation fields cannot be determined by known analytical methods. If a crack tip is moving more or less steadily, then a steady-state analysis yields relationships among the various parameters used to characterize the process without specific reference to a crack propagation criterion. A study of how a crack tip or fault edge actually moves according to a particular crack propagation condition must be based on an analysis of transient crack propagation, and certain results of such analysis are summarized in the next section.

TRANSIENT CRACK PROPAGATION

The discussion of transient shear crack propagation is based on the model of a half-plane crack extending in mode II in an otherwise unbounded body. This model exhibits most of the conceptual features which have been considered in dynamic shear crack propagation analysis. In the geophysics literature, more emphasis has been placed on the analysis of the symmetrical expansion of a mode II shear crack of finite length and the three-dimensional expansion of an elliptical combined mode II and III shear crack than on the semi-infinite mode II or mode III cracks. The reason for this emphasis seems to be that actual faults have finite dimensions and if primary interest is on the details of seismic radiation due to fault motion, for example, then these fault dimensions must be included. If primary interest is on the fracture process, on the other hand, then the actual fault dimensions are of lesser importance and the semi-infinite crack models appear to be suitable. In general, for a specific characterization of fracture resistance, the main influence on crack tip motion is the increased area over which the stress drop acts as the crack increases in size, and this effect can be included

in the semi-infinite plane crack models. The influence on crack tip motion at some point on a fault edge due to stress waves generated at some other point on the fault edge seems to be minor by comparison, and this effect is automatically precluded in the semi-infinite plane crack models. In his study of the symmetrical extension of a mode I tensile crack of finite length, Rose [32] showed that the wave disturbance generated by either moving crack tip had a negligible effect on the motion of the other. Concerning three-dimensional shear cracks, the edge deformation can be resolved into a combination of mode II and mode III deformations and, with this point of view, the results obtained from analysis of plane models have bearing on three-dimensional plane crack propagation as well.

The specific mode II shear crack model on which this discussion is based is represented in Fig. 5. At the initial time $t = 0$ the half-plane crack begins to grow in the plane $y = 0$ and at time t the crack edge has advanced a distance $\ell(t)$. Equal and opposite shear tractions of magnitude σ_o act on the crack faces in the interval $0 < x < \ell(t)$ as shown. The stress σ_o may be viewed as the difference between a remotely applied stress of magnitude σ_∞ and a frictional stress which resists sliding σ_d. Thus, σ_o is a stress drop and it acts over a region of increasing size as the crack grows. There are no other crack face tractions or remote loads acting on the body. Posed in this way, the problem is not suitable for consideration of fracture initiation. Initial loading of the body is not included explicitly, and the role of the pre-existing crack as a stress concentrator in fracture initiation is not considered. These exclusions were made in order to keep the analysis as simple as possible, and loading conditions for this problem which might be more suitable for the study of fracture initiation are considered in [3].

The solution of this problem for arbitrary nonuniform motion of the crack tip with speed $\dot{\ell}(t)$ in the range $0 < \dot{\ell} < v_r$ has been given by Fossum and Freund [3]. A solution of this problem for crack speeds greater than the Rayleigh

wave speed v_r has not appeared in the literature, although solutions for crack growth at constant speed $\dot{\ell} = v > v_r$ can be obtained by following the analytical procedure of [3]. In terms of the notation established in (1)-(4), the shear stress on the prospective fracture surface $y = 0$ directly ahead of the crack tip at $x = \ell(t)$ is

$$\sigma_{xy} = \frac{K_2(t)}{\sqrt{2\pi(x-\ell)}} - \sigma_o \quad , \quad K_2(t) = 2\sigma_o\sqrt{2\ell/\pi}\, k(\dot{\ell}) \tag{30}$$

for arbitrary $\ell(t)$ within the range $0 < \dot{\ell} < v_r$,

$$\sigma_{xy} = -\sigma_o \quad , \quad K_2(t) = 0 \tag{31}$$

for constant speed v in the range $v_r < v < v_s$, and

$$\sigma_{xy} = \frac{K_2^*(t)}{\sqrt{2\pi}\,(x-vt)^m} - \sigma_o \quad , \quad K_2^*(t) = k^*(v)\,\sigma_o\,(vt)^m \tag{32}$$

for constant speed v in the range $v_s < v < v_\ell$. The function k in (30) is defined precisely in equation 12 of [3] and a rough approximation of this function is $k(\dot{\ell}) = 1 - \dot{\ell}/v_r$; in particular, $k(0) = 1$ and $k(v_r) = 0$. For the crack tip speed range $v_r < v < v_s$, a solution with $K_2 \neq 0$ can be found. However, as can be seen from (17), a mode II crack tip propagating at a speed within this range with $K_2 \neq 0$ acts as a point source of energy. Such a result must be ruled out on physical grounds, so that $K_2 = 0$ and the next term in the expansion of stress with distance from the crack tip is that shown in (31). The function $k^*(v)$ has not yet been evaluated in detail, but it seems to be bounded, nonzero and of order unity for speeds in the range $v_s < v < v_\ell$. When considered in conjunction with a fracture propagation criterion, the results (30)-(32) imply crack tip motions having certain characteristics. These will be briefly discussed for several different fracture criteria.

Critical stress intensity factor criterion. Suppose that the shear crack is required to propagate in such a way that the stress intensity factor always has a fixed critical value. This is a generalization to the case of extending cracks of Irwin's well-known fracture initiation criterion. After the initiation phase has passed, the expression for K_2 in (30) suggests that the fracture condition can be satisfied as ℓ increases indefinitely by having the crack tip gradually accelerate toward the Rayleigh wave speed so that the product $\ell k(\dot{\ell})$ is held constant. In this sense, the Rayleigh wave speed represents the terminal velocity for crack propagation. Furthermore, if the traditional definition or interpretation of the stress intensity factor is understood, then crack speeds greater than the Rayleigh wave speed are not considered with this criterion because the coefficient of the square-root singular contribution to the local stress field is zero (except for the special case $v = v_s\sqrt{2}$).

Critical energy release rate criterion. Suppose that the shear crack is required to propagate in such a way that the energy release rate G always has a fixed critical value. This is a generalization of the classical Griffith criterion. The resulting crack motion will be almost identical to that for the critical stress intensity factor criterion, differing only in minor details. The reason for this similarity is that the energy release rate is proportional to the square of the stress intensity factor, and the energy release rate is nonzero only if the local crack tip stress field is square-root singular. Stress fields with weaker singularities and nonsingular stress fields result in zero energy release rate for a sharp crack tip. Crack motion according to this criterion is discussed in detail in [3]. The main qualitative result is that an unstable sharp tipped mode II crack propagating with a constant energy release rate will accelerate toward the Rayleigh wave speed, but its speed will always be less than the Rayleigh wave speed. The energy release rate is also nonzero for the special crack speed $v_s\sqrt{2}$ but, according to (32), any specific value of energy release rate can be

achieved at this speed only for a particular fixed crack length. Thus, the fracture condition cannot be satisfied for a crack extending at the speed $v_s\sqrt{2}$.

Critical stress level criterion. Suppose that large, and possibly singular, shear stresses on the prospective fracture plane ahead of the crack tip are admitted, but that the interface can support such stresses without fracture only if the interval over which some critical stress level, say σ_s, is exceeded in less than some critical length, say λ. Further, suppose that fracture will begin if the fault plane shear stress is greater than σ_s over an interval of length λ, and that the fracture will proceed with this condition satisfied. The shear stress can never exceed σ_s over a length interval greater than λ. This notion of a critical stress acting over a certain distance was used by Irwin [33] in his early studies of crack tip plasticity. In this case the critical stress was proportional to the plastic flow stress of the material and the critical distance was the radius of the plastically deforming region around the crack tip. A similar viewpoint was adopted by Congleton and Petch [34] in modeling the initiation of microcrack growth ahead of the tip of a much larger crack. In this model, growth of the large crack was achieved through coalescence of numerous microcracks. In this case, the critical stress was the Griffith stress necessary to extend pre-existing microcracks and the corresponding critical distance was the mean spacing of these microcracks. A version of the same criterion was recently employed by Das and Aki [12] in a form particularly well-suited for numerical computation by a finite difference scheme. However, it seems that they used the crack tip stress distribution (30) for the entire velocity range $0 < v < v_\ell$.

After some fracture initiation phase has passed, the critical stress level fracture condition implies that the crack will accelerate according to

$$\frac{K_2(t)}{\sqrt{2\pi\lambda}} - (\sigma_\infty - \sigma_d) = (\sigma_s - \sigma_\infty) \tag{33}$$

The term $-\sigma_\infty$ has been inserted on the right side of (33) to account for the

fact that the remote loading contribution to the prospective fracture plane stress was not included in the formulation of the boundary value problem. With the explicit expression for $K_2(t)$ given in (30), the result (33) becomes (cf. equation 27 of [12])

$$\frac{2k(\dot{\ell})}{\pi}\left(\frac{\ell}{\lambda}\right)^{1/2} = \frac{(\sigma_s - \sigma_d)}{(\sigma_\infty - \sigma_d)} \tag{34}$$

After any appreciable amount of crack growth $\ell >> \lambda$. Therefore k must approach zero, or the wave speed must approach the Rayleigh wave speed.

During the acceleration phase, the shear stress on the fracture plane is usually small compared to σ_s at distances ahead of the crack tip which are large compared to λ. However, as the crack propagates steadily at speeds near the Rayleigh wave speed, the shear stress distribution on the fracture plane develops a sharp peak at a point traveling with the shear wave speed. This behavior was also observed for the symmetrically growing shear crack by Burridge [1]. The shear stress magnitude at this peak is $n(\sigma_\infty - \sigma_d)$. An estimate of the integral defining n yielded a value of about three, which is not too different from the numerically computed value $n = 1.63$ reported in [1]. If the interface is relatively weak, i.e., if σ_s is not much larger than σ_∞, then it is quite possible that $n(\sigma_\infty - \sigma_d) > (\sigma_s - \sigma_\infty)$. If this inequality is satisfied at the shear stress peak, then it will simply be a matter of time before it is satisfied over an interval of length λ. Thus, if the interface is relatively weak, then a secondary fracture will initiate at some distance ahead of the main fracture at some time.

A question then arises as to how this secondary fracture will grow. A detailed analysis of the growth process would be prohibitively complicated. There are not many possibilities, however, and the growth of the secondary fracture may be discussed in qualitative terms. The end of the secondary fracture nearest the main crack tip and the main crack tip will likely coalesce shortly after formation of

the secondary fracture, so that the other tip of the secondary fracture becomes the main crack tip. One possibility is that this tip will move at a speed just below the Rayleigh wave speed, another secondary fracture will be initiated some time later, and this process will be repeated over and over. A second possibility is that this tip will move at a speed greater than the Rayleigh wave speed. If $\sigma_s > \sigma_\infty$, then it can be seen from (31) that speeds in the range $v_r < v < v_s$ are ruled out, and possible speeds in the range $v_s < v < v_\ell$ must be sought. For this purpose, use is made of (32) with the interpretation of vt in (32) as the amount of crack growth since the formation of the secondary fracture, say $\Delta\ell$. The crack growth criterion then takes the form

$$k^*(v) \left(\frac{\Delta\ell}{\lambda}\right)^m = \frac{(\sigma_s - \sigma_d)}{(\sigma_\infty - \sigma_d)} \tag{35}$$

If k^* is insensitive to variations in speed v, then m must decrease as $\Delta\ell$ increases for fixed λ. As can be seen from its definition (7), the value of m decreases from $m = 1/2$ at $v = v_s\sqrt{2}$ to $m = 0$ at $v = v_\ell$. Thus the secondary crack would begin to grow at a speed equal to or greater than $v_s\sqrt{2}$ and it would accelerate toward the speed v_ℓ. The second of these two possibilities is qualitatively consistent with the numerical calculations of Andrews [11] and of Das and Aki [12].

In summary, it seems that for a relatively strong interface with $(\sigma_s - \sigma_\infty)/(\sigma_\infty - \sigma_d) > n$ the crack will accelerate rapidly to speeds approaching the Rayleigh wave speed, but will continue to propagate indefinitely at speeds below the Rayleigh wave speed. For a relatively weak interface with $(\sigma_s - \sigma_\infty)/(\sigma_\infty - \sigma_d) < n$ the crack tip will quickly accelerate to speeds approaching the Rayleigh wave speed and it will propagate for some shorttime at sub-Rayleigh wave speeds. Then, due to secondary fracturing, the crack tip speed will abruptly increase to $v_s\sqrt{2}$ or beyond and it will continue to accelerate toward the longitudinal wave speed v_ℓ. The qualitative picture of crack propagation based on the critical stress level fracture

criterion and the simple mode II crack problem represented in Fig. 5 is generally consistent with the numerical results of [11] and [12]. One fundamental difficulty with the critical stress level criterion is that the existence of a critical length λ seems to have no particular physical basis. However, as discussed by Das and Aki [12], the criterion appears to be well-suited for numerical computation by a finite difference approach where λ can be identified with the difference mesh spacing on the fracture plane. Furthermore, in such a numerical scheme the shear stress can be reduced from σ_s to σ_d according to some cohesive zone model to simulate energy uptake in overcoming an intrinsic material cohesion in the crack tip region. Unless σ_s is larger than σ_∞ and σ_d by two or three orders of magnitude, however, the rate of work against cohesion is negligible by comparison to the rate at which is energy absorbed through frictional sliding of the crack faces. A direct comparison of the results of the numerical calculations in [11] and [12] strongly suggests that the crack motion is quite insensitive to whether or not a finite, nonzero crack tip fracture energy exists. Andrews [11] uses a cohesive zone crack tip model to simulate nonzero energy uptake at the crack tip at all crack speeds. On the other hand, Das and Aki [12] simulate crack growth in their finite difference scheme by abruptly releasing nodes in the difference mesh whenever the nodal stress reaches a critical value. As observed by Rice [35], crack growth consisting of abrupt release of stress over a finite interval (the mesh spacing) occurs without energy absorption at the crack tip. The fact that the crack motion predictions in these two studies are qualitatively identical suggests that the influence on crack motion of crack tip energy absorption rates is small, probably because such rates are usually small compared to other energy rates involved as noted above.

Nonuniform stress drop and fracture resistance. This section is concluded with a brief discussion of crack arrest. As noted by Husseini, et al. [4], a running fracture can be arrested by either of two mechanisms, either a reduction

of the crack driving force to a subcritical level or an increase in the resistance of the material to a supercritical level. A simple illustration of these effects can be given in terms of the present mode II crack propagation model and the critical stress intensity fracture criterion, for example. The dynamic stress intensity factor for the shear crack shown in Fig. 5 for crack face traction representing a stress drop $\sigma_o(x)$ which varies in an arbitrary manner along the slip plane is given in equation 18 of [3] for sub-Rayleigh wave speeds. In terms of this arbitrary stress drop, the critical stress intensity factor fracture condition requires that the crack move in such a way that the dynamic stress intensity factor is always equal to some critical value, say $C(\ell)$, which may also vary in an arbitrary manner along the slip plane. The equation of motion of the crack tip, i.e., the equation governing $\ell(t)$, is then

$$k(\dot{\ell})\, K_{st}(\ell) = C(\ell) \tag{36}$$

where $K_{st}(\ell)$ is the equivalent static crack stress intensity factor

$$K_{st}(\ell) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^{\ell} \frac{\sigma_o(x)dx}{\sqrt{\ell - x}} \tag{37}$$

Some specific implications of (36) are shown schematically in Figs. 6 and 7. The case considered in Fig. 6 is that with constant critical stress intensity factor C and with continuously decreasing stress drop σ_o as shown. The equation (37) implies that the equivalent static stress intensity factor will first increase after initiation and then decrease, with a maximum at the point where $\sigma_o = 0$. The equation of motion (36) implies that the crack tip velocity will also increase and then decrease, with a maximum at the same point. The crack will arrest when K_{st} has been reduced to C. With reference to the sketch, the speed $\dot{\ell} = 0$ when $K_{st} = C$.

A similar example in which the stress drop σ_o is constant but the fracture

resistance C increases along the slip plane is shown schematically in Fig. 7. The equation (37) implies that K_{st} will continuously increase with crack length, and according to (36) the crack will continue to accelerate as long as the slope of K_{st} vs. ℓ is greater than the slope of C vs. ℓ. The crack tip reaches a maximum velocity when these slopes coincide, and the crack tip will come to rest if C reaches the value K_{st} at some point. With reference to the sketch, $\dot{\ell} = 0$ when $K_{st} = C$.

Finally, it is noted that the functions $\sigma_o(x)$ and $C(\ell)$ need not be monotonic, but they could vary in an arbitrary way along the slip plane in order to simulate variations in driving stress, frictional resistance and fracture resistance in the faulting process. It is clear from (36) that if these functions varied in an oscillatory manner, then crack motions $\ell(t)$ which solve (36) or other equations of motion similar to it could be very complicated indeed. The effects of nonuniform driving stress and fracture resistance are likely to be important in future development of dynamic shear crack models.

REFERENCES

1. Burridge, R., Admissible speeds for plane-strain self-similar shear cracks with friction but lacking cohesion, Geophys. J. R. astr. Soc. 35 (1973) 439-455.

2. Burridge, R. and G. S. Halliday, Dynamic Shear cracks with friction as models for shallow focus earthquakes, Geophys. J. R. astr. Soc. 25 (1971) 261-283.

3. Fossum, A. F. and L. B. Freund, Nonuniformly moving shear crack model of a shallow focus earthquake and mechanism, J. Geophys. Res. 80 (1975) 3343-3347.

4. Husseini, M. I., D. B. Jovanovich, M. J. Randall and L. B. Freund, The fracture energy of earthquakes, Geophys. J. R. astr. Soc. 43 (1975) 367-385.

5. Kostrov, B. V., Unsteady propagation of longitudinal shear cracks, Appl. Math. Mech. (English transl. of PMM) 30 (1966) 1241-1248.

6. Kostrov, B. V., On the crack propagation with variable velocity, Int. J. Fracture 11 (1975) 47-56.

7. Kostrov, B. V. and L. V. Nikitin, Some general problems of mechanics of brittle fracture, Arch. Mech. Stos. 22 (1970) 749-775.

8. Richards, P. G., The dynamic field of a growing plane elliptical shear crack, Int. J. Solids Structures 9 (1973) 843-861.

9. Weertman, J., Theory of velocity of earthquake dislocations, Geological Soc. Am. Mem. 142 (1975) 175-183

10. Andrews, D. J., Rupture propagation with finite stress in antiplane shear, J. Geophys. Res. 81 (1976) 3575-3582.

11. Andrews, D. J., Rupture velocity of plane strain shear cracks, J. Geophys. Res. 81 (1976) 5679-5687.

12. Das, S. and K. Aki, A numerical study of two-dimensional spontaneous rupture propagation, Geophys. J. R. astr. Soc. 50 (1977) 643-668.

13. Ida, Y., Stress concentration and unsteady propagation of longitudinal shear cracks, J. Geophys. Res. 78 (1973) 3418-3429.

14. Freund, L. B., Dynamic crack propagation, in The Mechanics of Fracture, AMD Vol. 19, edited by F. Erdogan, Amer. Soc. Mech. Engrs. (1976) 105-134.

15. Brune, J. N., Earthquake modeling by stick-slip along pre-cut surfaces in stressed foam rubber, Bull. Seis. Soc. Am. 63 (1973) 2105-2119.

16. Husseini, M. I. and M. J. Randall, Rupture velocity and radiation efficiency, Bull. Seis. Soc. Am. 66 (1976) 1173-1187.

17. Madariaga, R., Dynamics of an expanding circular fault, Bull. Seis. Soc. Am. 66 (1976) 639-666.

18. Richards, P. G., Dynamic motions near an earthquake fault: a three-dimensional solution, Bull. Seis. Soc. Am. 66 (1976) 1-32.

19. Irwin, G. R., Analysis of stresses and strains near the end of a crack traversing a plate, J. Appl. Mech. 24 (1957) 361.

20. Williams, M. L., On the stress distribution at the base of a stationary crack, J. Appl. Mech. 24 (1957) 109.

21. Freund, L. B. and R. J. Clifton, On the uniqueness of elastodynamic solutions for running cracks, J. Elasticity 4 (1974) 293-299.

22. Freund, L. B., Energy flux into the tip of an extending crack in an elastic solid, J. Elasticity 2 (1972) 341-349.

23. Rice, J. R., Mathematical analysis in the mechanics of fracture, in Fracture, Vol. 2, edited by H. Liebowitz, Academic Press (1968) 191-311.

24. Atkinson, C. and J. D. Eshelby, The flow of energy into the tip of a moving crack, Int. J. Frac. Mech. 4 (1968) 3-8.

25. Sih, G. C., Dynamic aspects of crack propagation, in Inelastic Behavior of Solids, edited by M. F. Kanninen, et al., McGraw-Hill (1970) 607-639.

26. Barenblatt, G. I., R. L. Salganik and G. P. Cherepanov, On the nonsteady motion of cracks, Appl. Math. Mech. (transl. of PMM) 26 (1962) 469.

27. Dugdale, D. S., Yielding of steel sheets containing slits, J. Mech. Phys. Solids 8 (1960) 100-104.

28. Palmer, A. C. and J. R. Rice, The growth of slip surfaces in the progressive failure of over-consolidated day, Proc. Roy. Soc. (London) A332 (1973) 527-548.

29. Ida, Y., Cohesive force across the tip of a longitudinal shear crack and Griffith's specific surface energy, J. Geophys. Res. 77 (1972) 3796-3805.

30. Radok, J. R. M., On the solution of problems of dynamic plane elasticity, Quart. Appl. Math. 14 (1956) 289-298.

31. Nur, A., Advances in rock mechanics, in Proc. Third Congress Int. Soc. Rock Mech., Denver (1974) 243-317.

32. Rose, L. R. F., On the initial motion of a Griffith crack, Int. J. Fracture I2 (1976) 829-841.

33. McClintock, F. A. and G. R. Irwin, Plasticity effects in fracture mechanics, in Fracture Toughness Testing and Its Applications, ASTM STP 381, Amer. Soc. Testing Matls. (1965) 84-113.

34. Congleton, J. and N. J. Petch, Crack branching, Phil. Mag. 16 (1967) 749-760.

35. Rice, J. R., Radiant energy dissipation in crack propagation, July 1977 Proc., ARPA Materials Research Council, Univ. of Mich., Dept. of Matls. Engrg., Technical Report (1978).

FIGURE CAPTIONS

Figure 1. Configuration of a body containing a crack at a fixed instant of time.

Figure 2. The cohesive zone crack tip model. The amount of slip between the crack faces is denoted by δ and the shear traction resisting slip within the cohesive zone is $\sigma(\delta)$. The contour L shown in (a) is employed in the derivation of (19). The cross-hatched area in (b) is the cohesive energy density of the interface.

Figure 3. A typical mode III steady-state crack propagation problem is shown in (a), where slip is resisted over a slipping region of length 2a by a shear traction σ_d. The variation of shear stress and particle velocity along the slip plane is sketched in (b).

Figure 4. A typical mode II steady-state crack propagation problem which can be analyzed by means of the energy integral (25).

Figure 5. Transient extension of a mode II semi-infinite shear crack due to a uniform stress drop σ_o acting over the interval $0 < x < \ell(t)$.

Figure 6. A sketch of the variation with position along the slip plane of the physical quantities appearing in the crack tip equation of motion (36) for the special case when the critical stress intensity factor C is a constant but the stress drop decreases along the slip plane.

Figure 7. Same as Fig. 6 for the special case when the stress drop σ_o is uniform but the critical stress intensity factor increases along the slip plane.

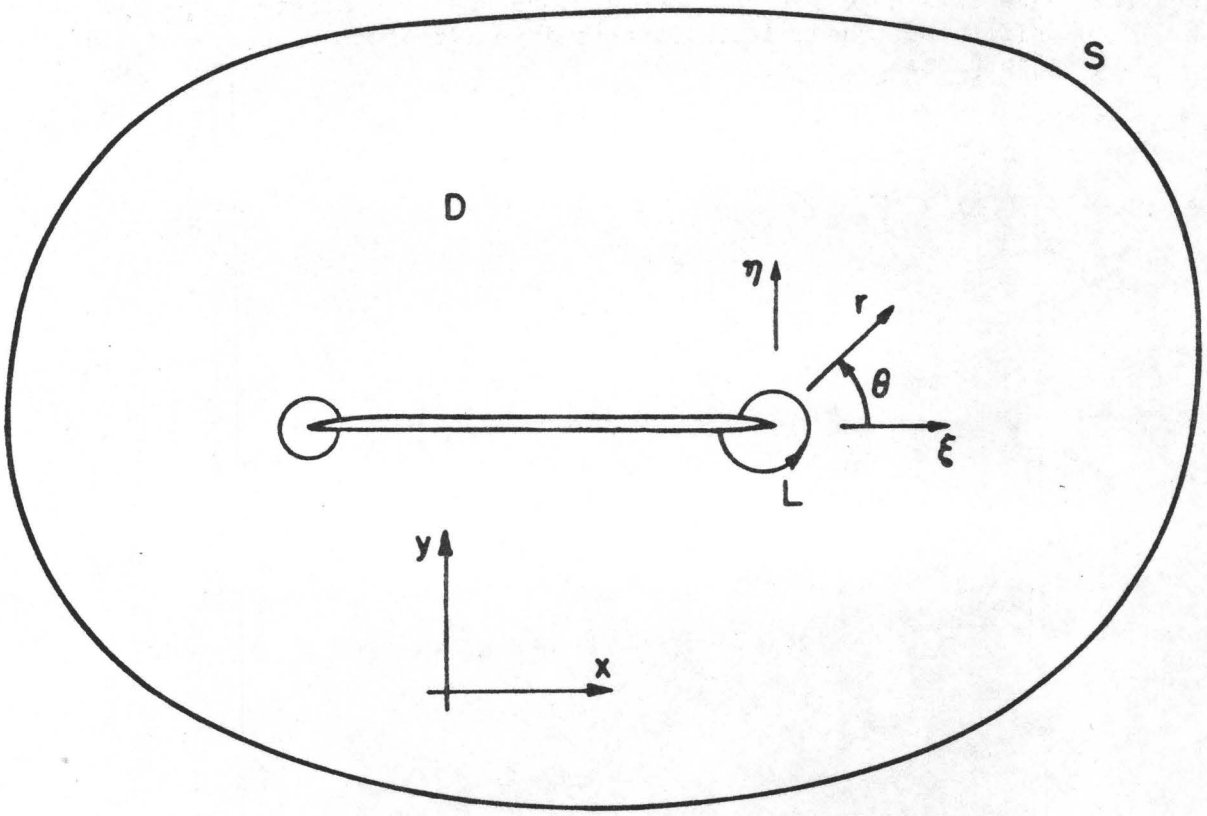


FIGURE 1

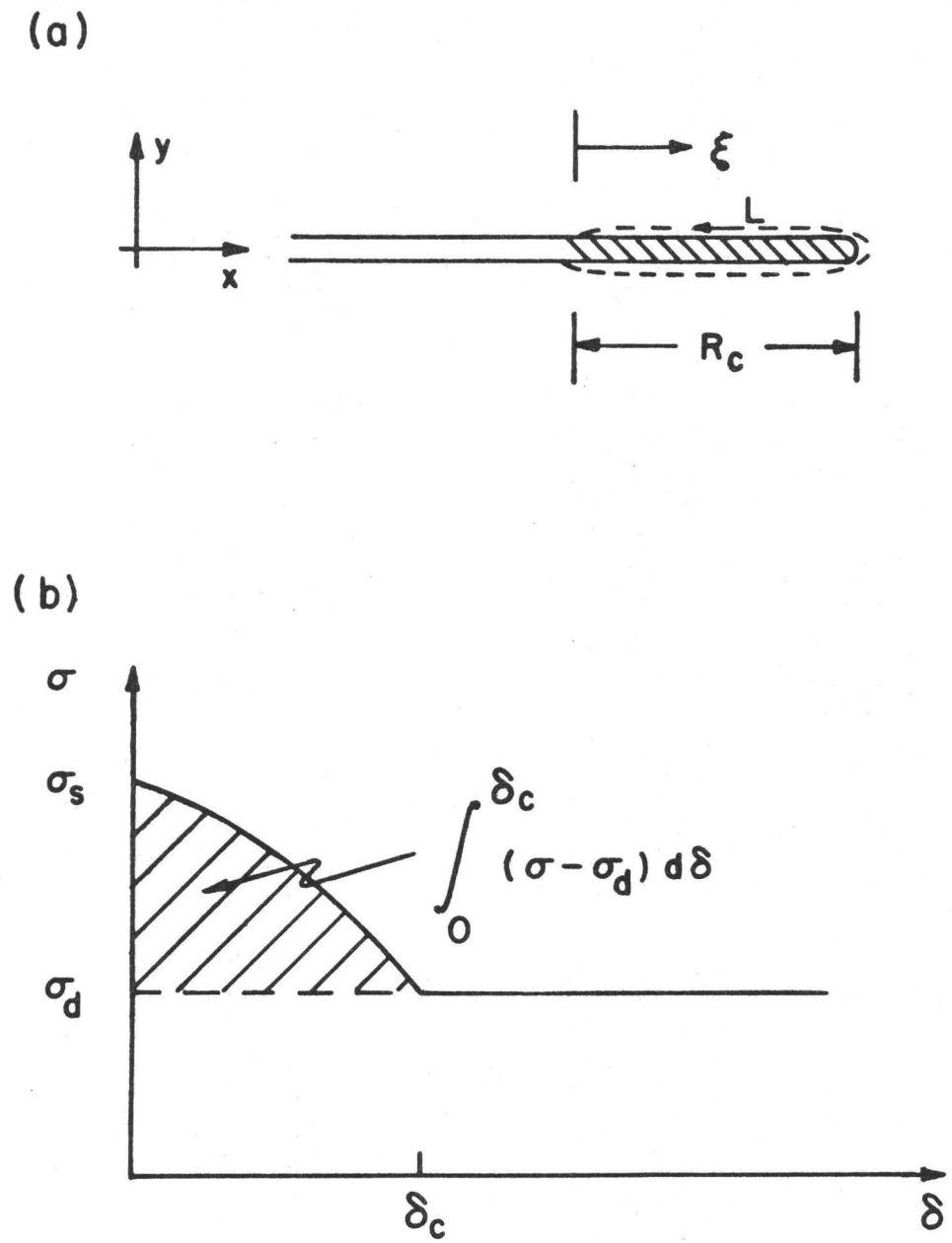

(a)
y
x
ξ
L
R_c
(b)
σ
σ_s
σ_d
δ_c
$\int_0^{\delta_c} (\sigma - \sigma_d)\, d\delta$
δ_c
δ

FIGURE 2

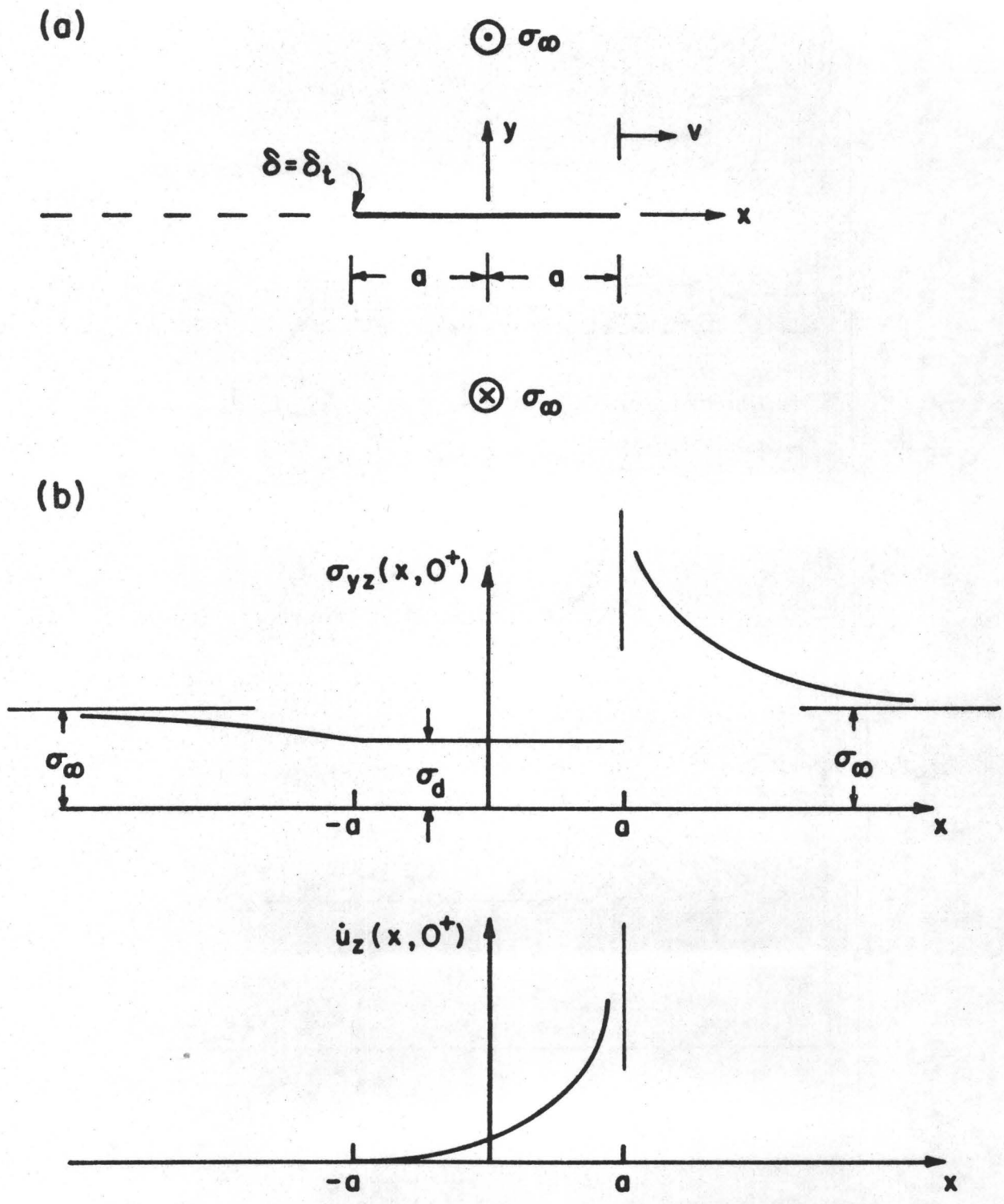

(a)
σ_∞
y
v
$\delta=\delta_t$
x
a
a
σ_∞
(b)
$\sigma_{yz}(x,0^+)$
σ_∞
σ_d
σ_∞
-a
a
x
$\dot{u}_z(x,0^+)$
-a
a
x

FIGURE 3

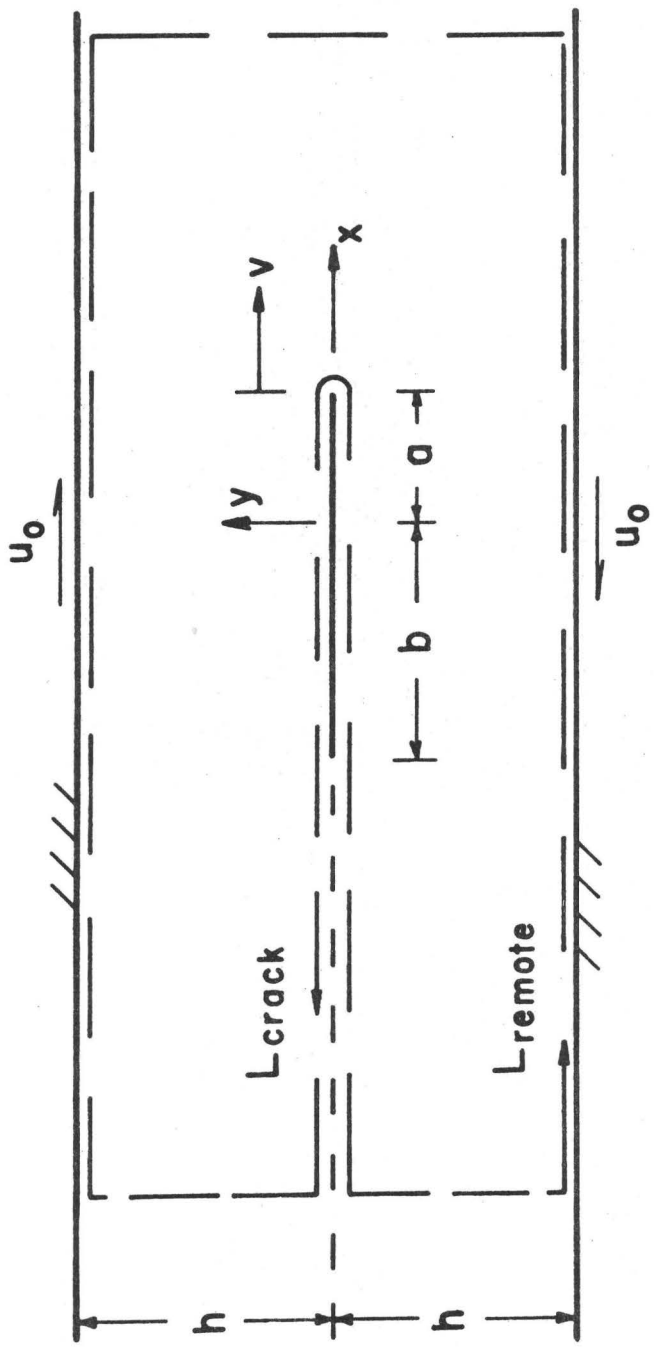


FIGURE 4

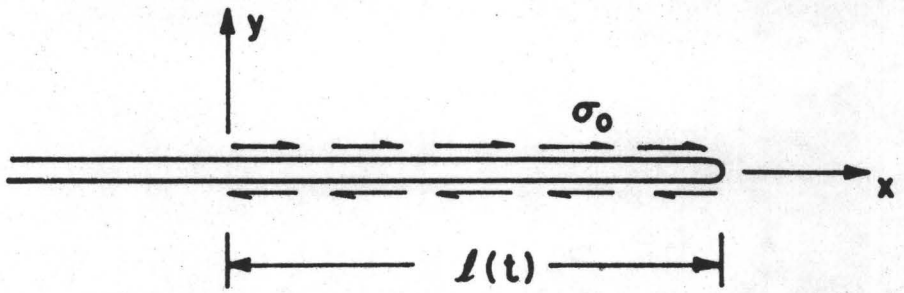


FIGURE 5

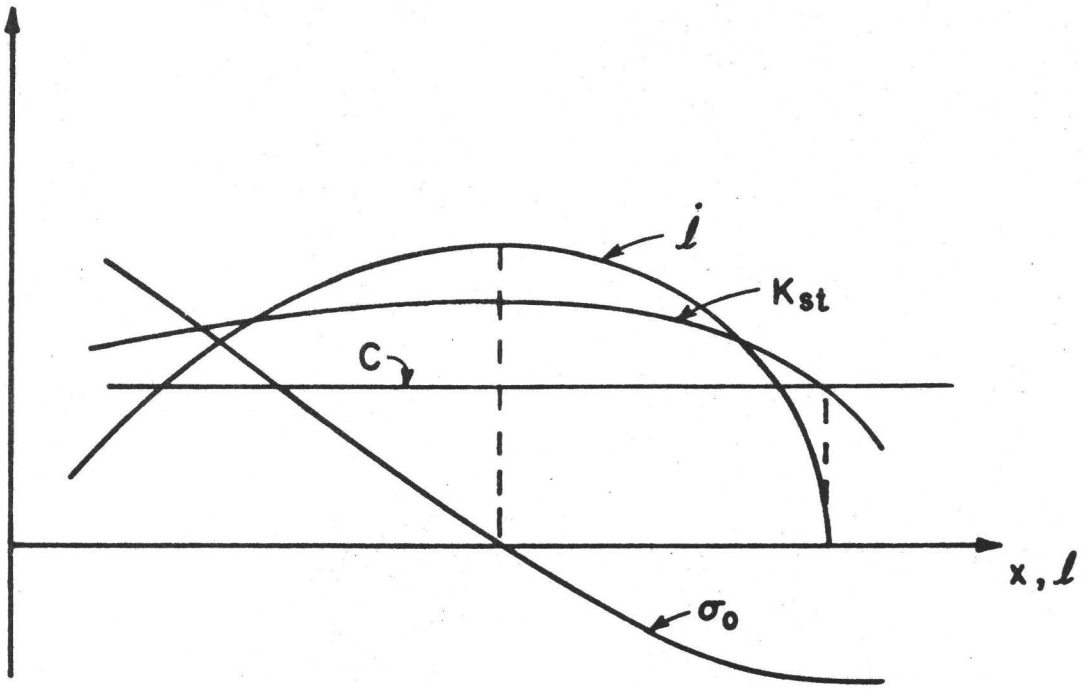


FIGURE 6

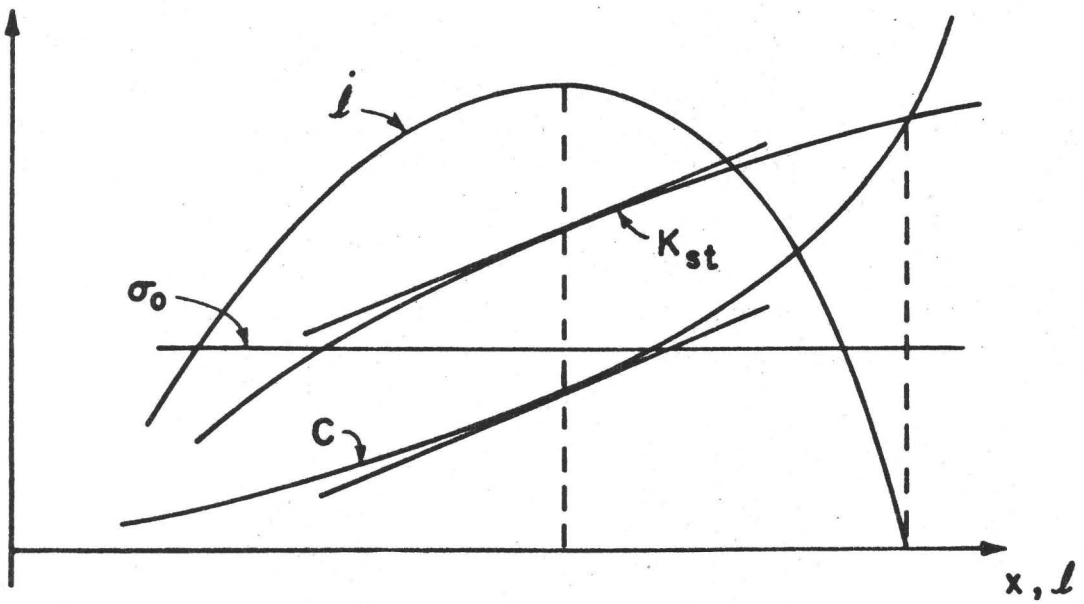


FIGURE 7

The June 10, 1975 Kurile Islands Tsunami Earthquake:

An Extended Abstract

Robert J. Geller
Department of Geophysics
Stanford University
Stanford, California 94305

and

Seismological Laboratory
California Institute of Technology
Pasadena, CA 91125

Kunihiko Shimazaki
Earthquake Research Institute
University of Tokyo
Tokyo, Japan

ABSTRACT

The source process of the June 10, 1975 Kurile Islands earthquake (m_b = 5.8, M_s = 7.0) is studied using WWSSN data, ultra-long period seismic records, and tsunami data. The observed tsunami amplitudes are much larger than are normally expected for an earthquake of this magnitude, suggesting that the rupture process may have a very long time duration. This is confirmed by a study of the WWSSN Rayleigh and Love waves, which show that the effective moment (the seismic moment of the step function source which would produce the same amplitudes at a particular period) increases rapidly from 5 x 10^{26} dyne cm at 50 sec. to 2 x 10^{27} dyne cm at 200 sec. The extremely long rise time causes relatively low amplitudes at periods of about 20 sec. (at which M_s is measured), while the tsunami amplitude (which depends on the moment at periods of several hundred seconds) is thus larger than expected for this M_s.

The mechanism inferred from body and surface waves is a shallow angle thrust fault, dipping 16° W and striking N 41° E. The main shock is relocated at T_o = 13:47:12.5, 42:97°N, 147.17°E, h = 2.1 km. On the basis of relocated after shocks within one day, the fault area is 50 km x 35 km.

This earthquake is classified as a tsunami earthquake (Kanamori, 1972) on the basis of its anamolously large tsunami amplitude and low dislocation velocity. Complex multiple events in the body waves and the unusually slow rupture process suggest that the fault zone may be unusually fractured and weak, with a very heterogeneous distribution of material strength.

APPENDIX

Figure 1: General Tectonic Setting of the June 10, 1975 Kuriles Earthquake (labeled 1975 Nemuro-Oki Earthquake in the figure), from Fukao (1978). Note that the aftershock zones are plotted using JMA locations.

Figure 2: Relocated aftershocks (within one day of main shock) of the June 10, 1975 tsunami earthquake.

Figure 3: RMS residual of observed-computed P arrival time when the epicenter of the mainshock is relocated with the depth constrained. h = 2.1 km is optimal. The circle in the upper left shows the number of stations in each quadrant.

Figure 4: Three long period vertical WWSSN seismograms of the June 10, 1975 earthquake. Note the extreme complexity of the waveforms, showing complex multiple events, with a duration of about two minutes. Also note the small foreshock about 10 seconds before the first arrival from the main shock. The earthquake waveforms have been traced over.

Figure 5: Difference in arrival time between the small foreshock and the first waves from the main shock, as a function of azimuth. There is no significant azimuthal variation, suggesting that the foreshock and mainshock have essentially the same location.

Figure 6: Focal mechanism determined from first motions on WWSSN stations and Rayleigh waves. The steeply dipping plane is constrained by the first motions and the conjugate plane, which represents the shallow angle thrust plane on which the earthquake actually occurred, is constrained by the Rayleigh wave radiation pattern.

APPENDIX

Figure 6 continued:

Solid circles represent compressions, and open circles dilatations.

Figure 7: Observed and theoretical Rayleigh wave spectral radiation patterns at periods of 200 sec, 100 sec and 50 sec. All theoretical spectra are for a moment of 2×10^{27} dyne cm. The observed spectral amplitudes are as large as the theoretical amplitudes at 200 sec, half as large at 100 sec and one fourth as large at 50 sec. Thus the moment is 2×10^{27} at a period of 200 sec, 1×10^{27} at 100 sec and 0.5×10^{27} at 50 sec. This is consistent with the slow slip shown by the body waves.

Figure 8: Observed Rayleigh wave spectra at several WWSSN stations. Note the holes (indicated by arrows) which vary in period as a function of azimuth.

Figure 9: If the spectral holes are interpreted as being caused by directivity, then by minimizing this misfit of the period of spectral holes as a function of azimuth we find that the rupture propagated N 40° E (along the fault strike), with characteristic rupture time L/V_R = 30 sec. This is difficult to reconcile with the observed source duration of about 2 min, and we are studying this question.

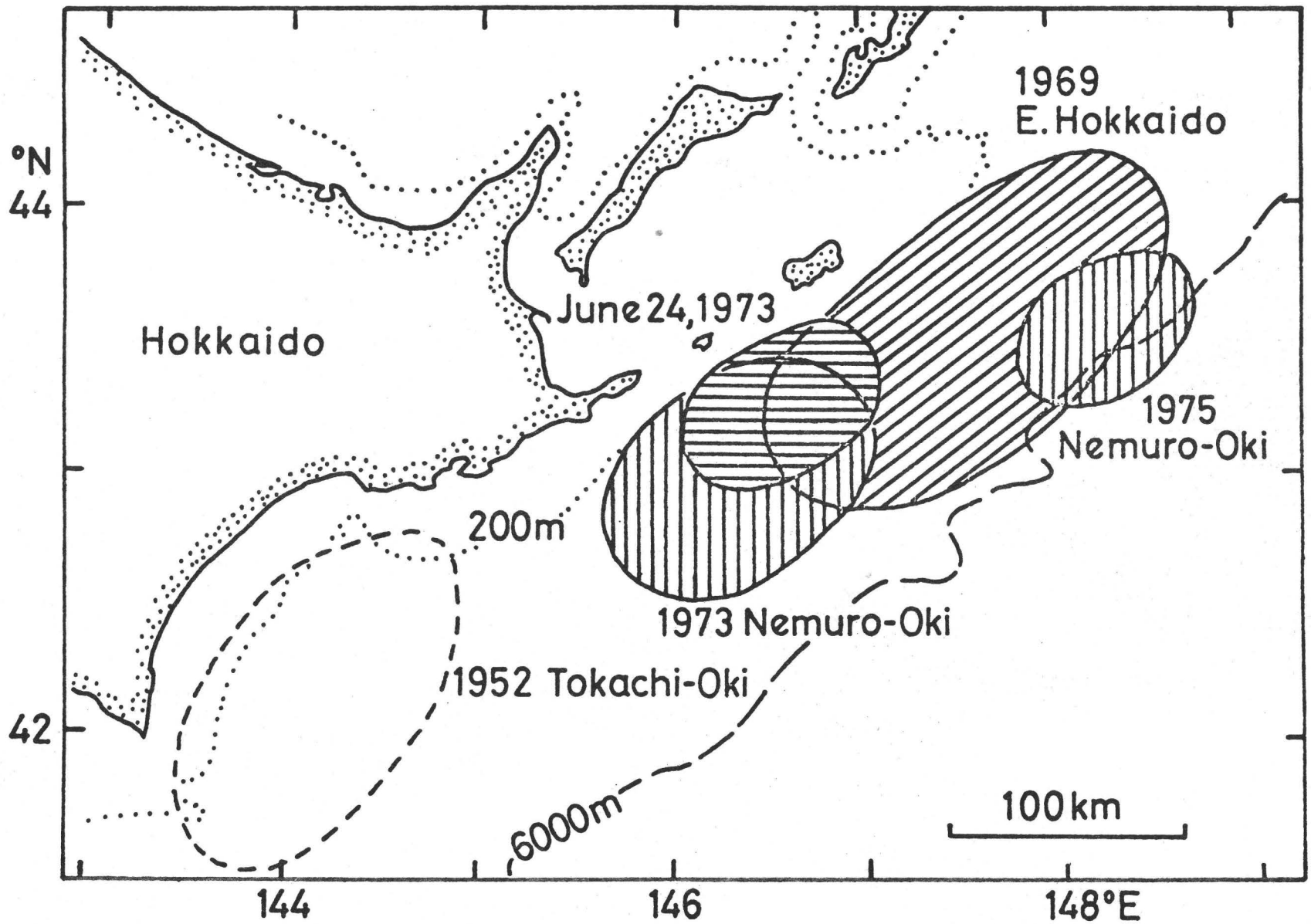

1969
E. Hokkaido
°N
44
June 24, 1973
Hokkaido
1975
Nemuro-Oki
200m
1973 Nemuro-Oki
1952 Tokachi-Oki
42
6000m
100km
144
146
148°E

Figure 1

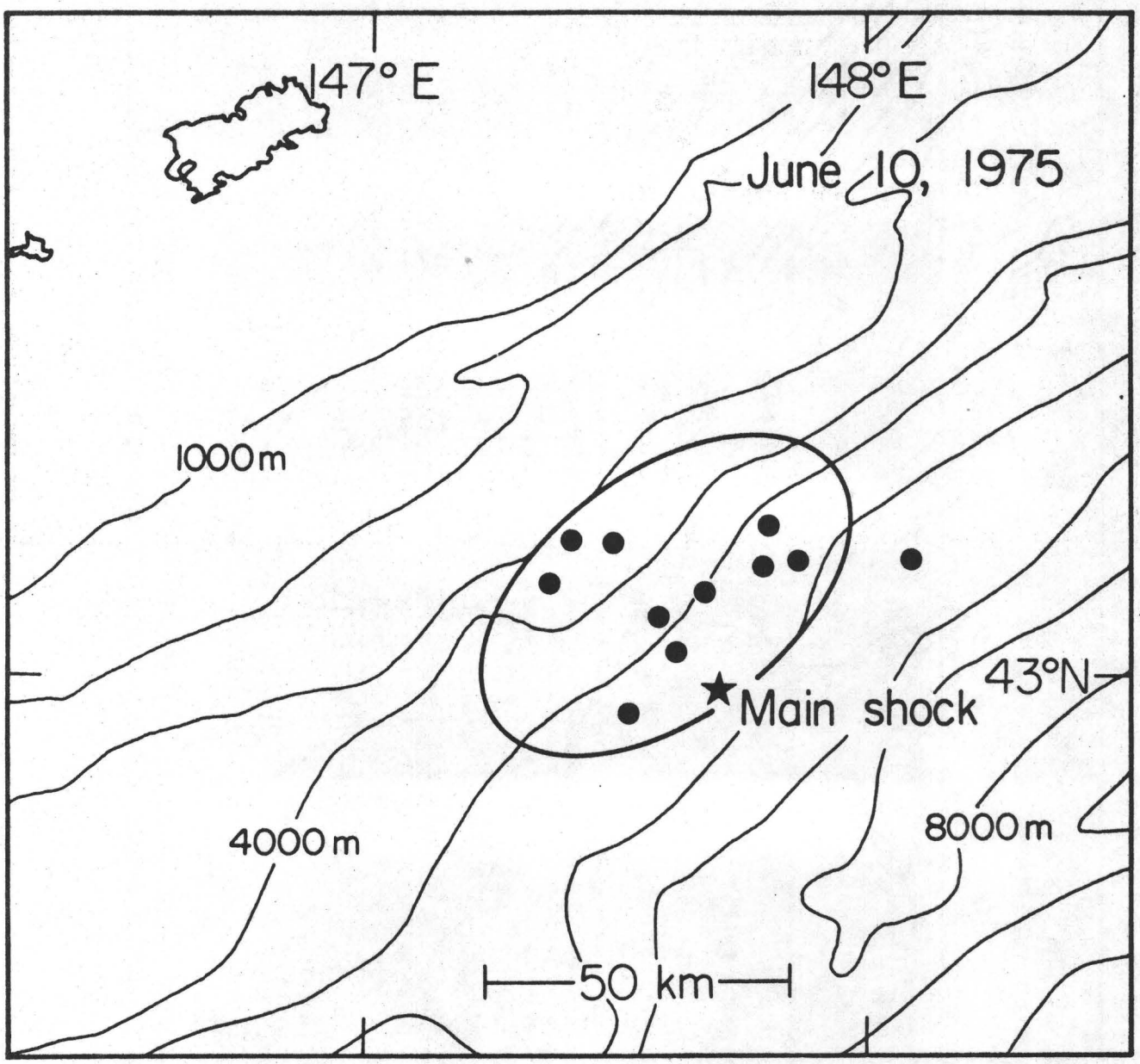


Figure 2

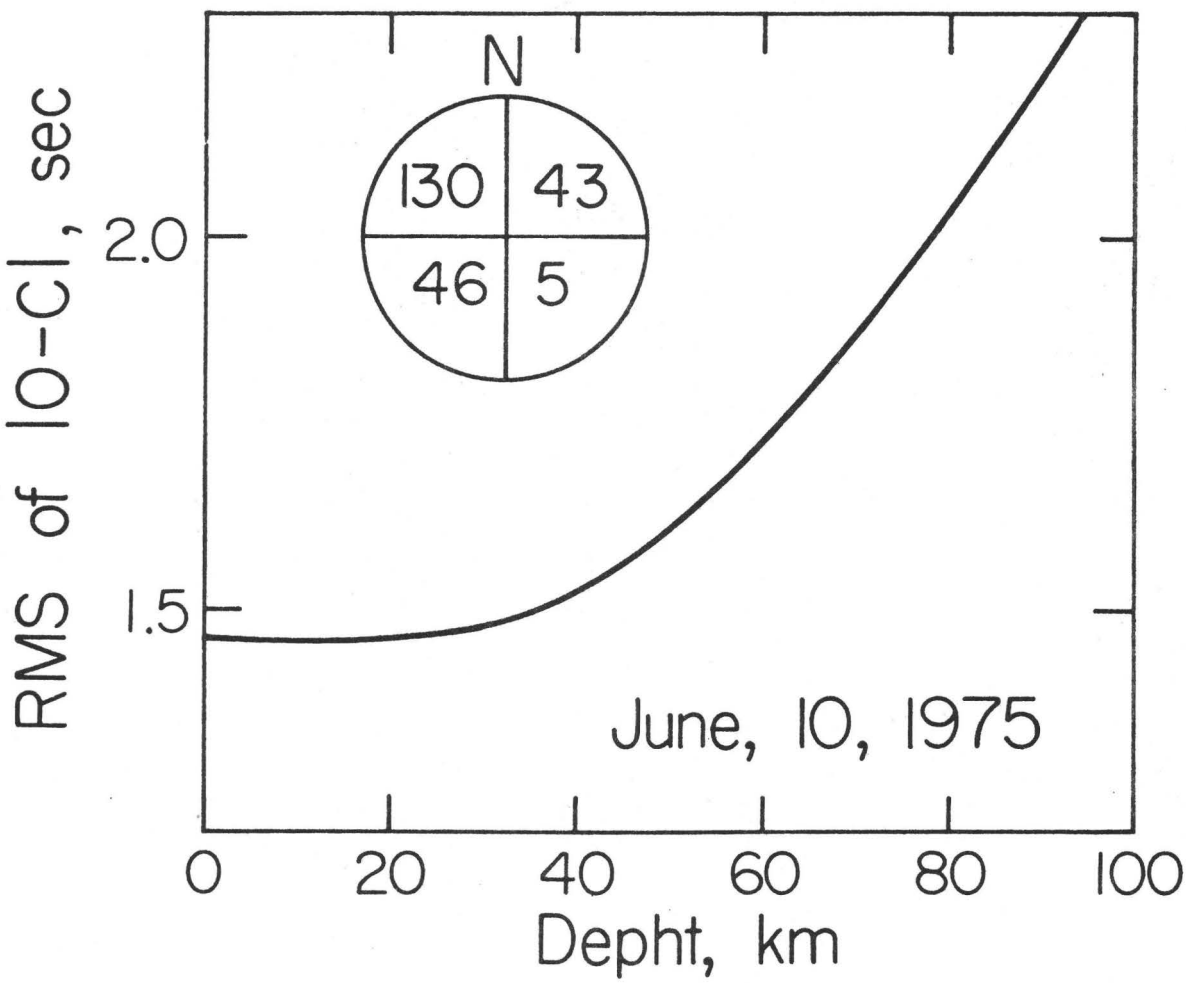

N
130
43
46
5
RMS of IO-CI, sec
2.0
1.5
June, IO, 1975
0
20
40
60
80
100
Depht, km

Figure 3

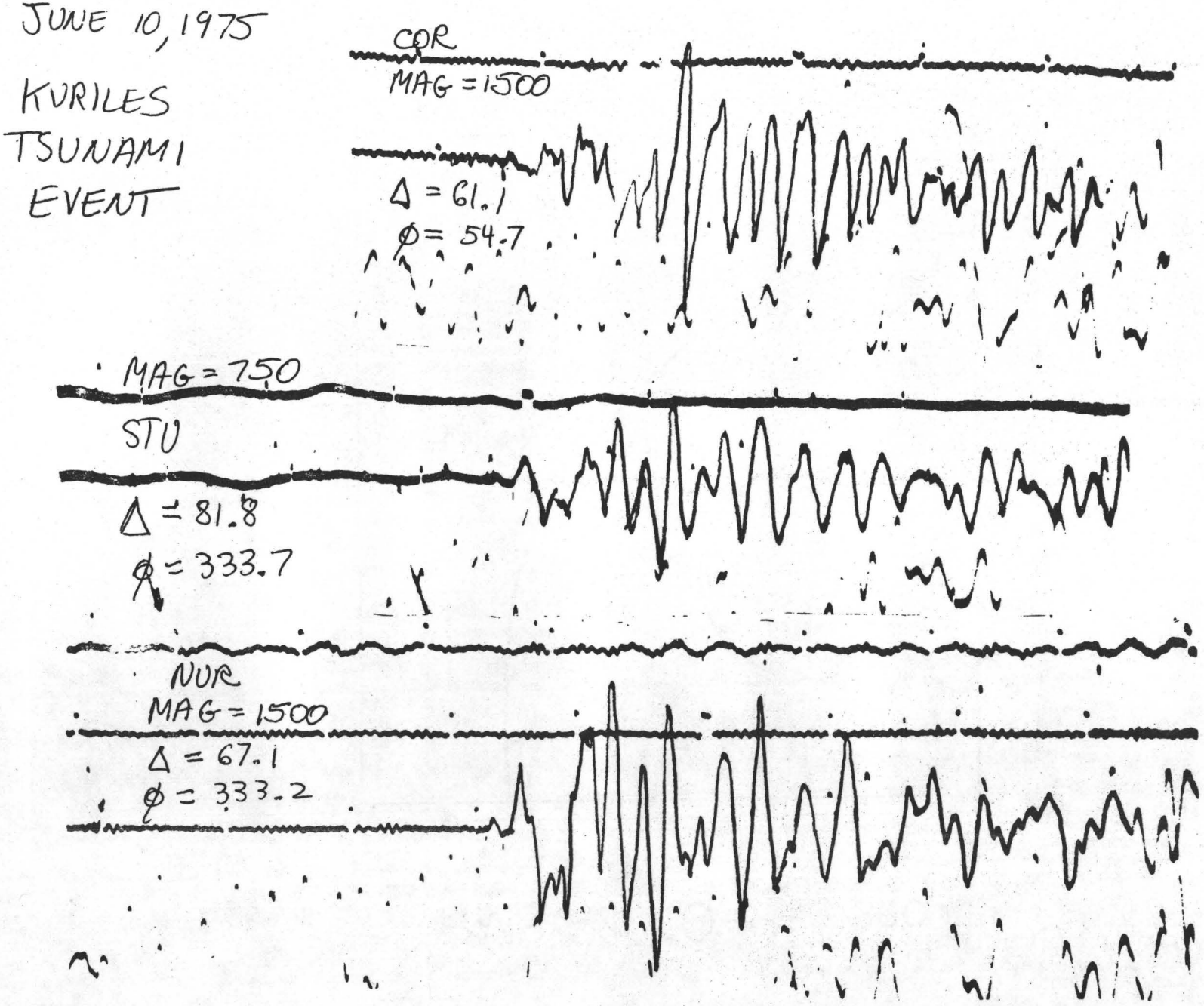

JUNE 10, 1975
KURILES
TSUNAMI
EVENT
COR
MAG = 1500
Δ = 61.1
Φ = 54.7
MAG = 750
STU
Δ = 81.8
Φ = 333.7
NUR
MAG = 1500
Δ = 67.1
Φ = 333.2

Figure 4

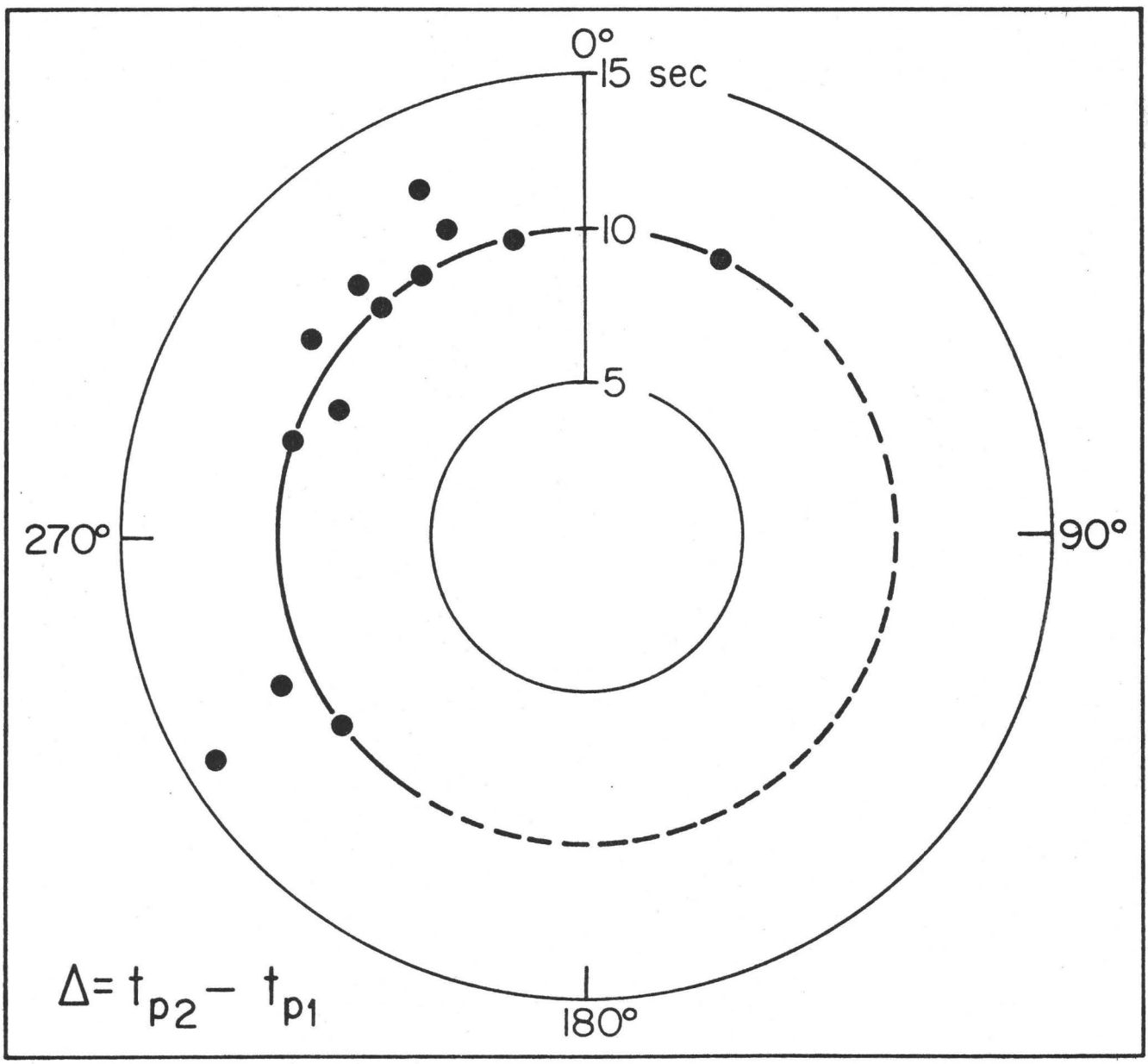


Figure 5

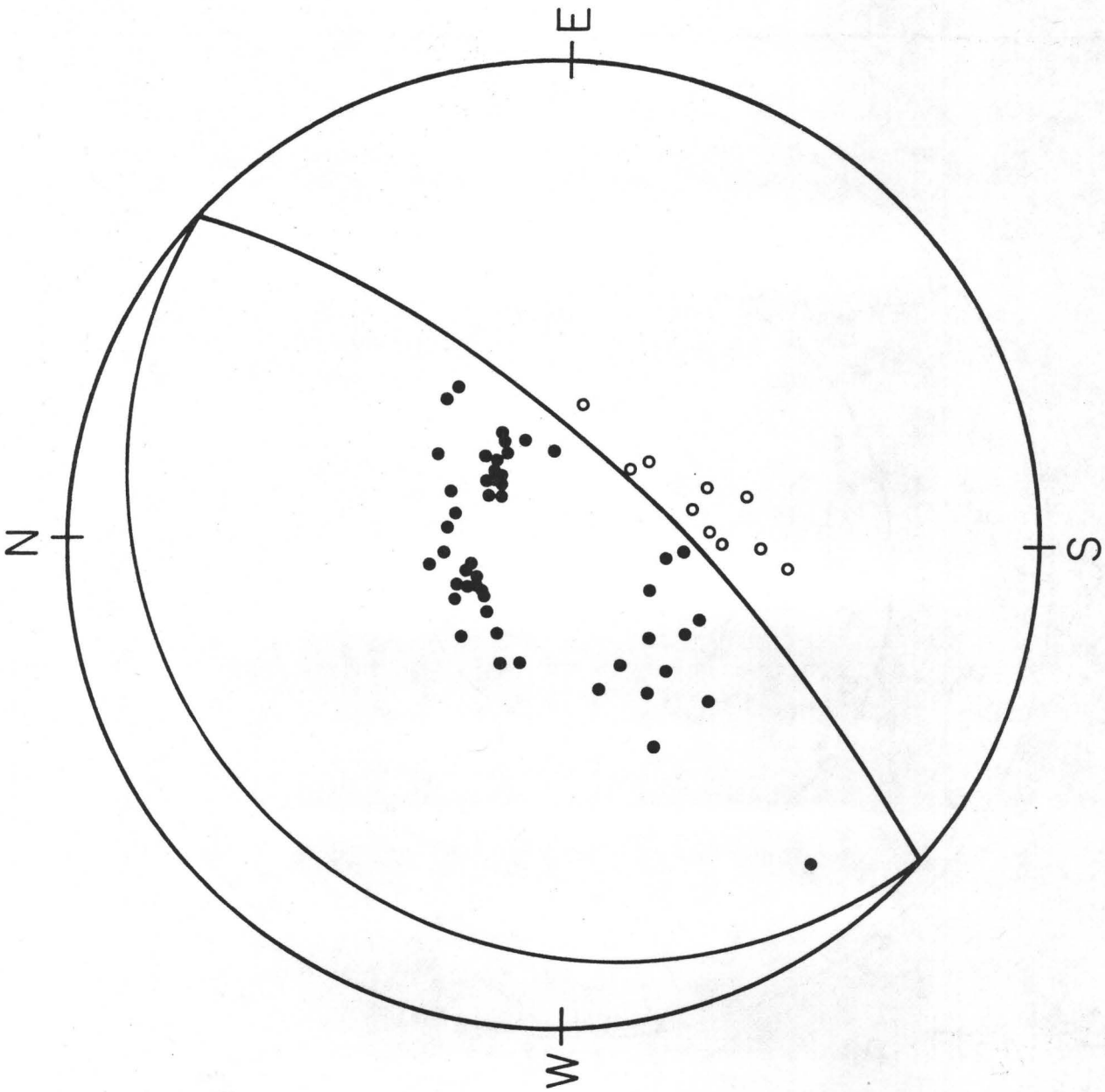


Figure 6

JUNE 10, 1975 TSUNAMI EARTHQUAKE

T=200. S T=100. S T= 50. S

DIP: 74. SLIP: 90. STRIKE: 41.

T=200. S
M0= 2.0 E27

T=100. S
M0= 2.0 E27

T= 50. S
M0= 2.0 E27

Figure 7

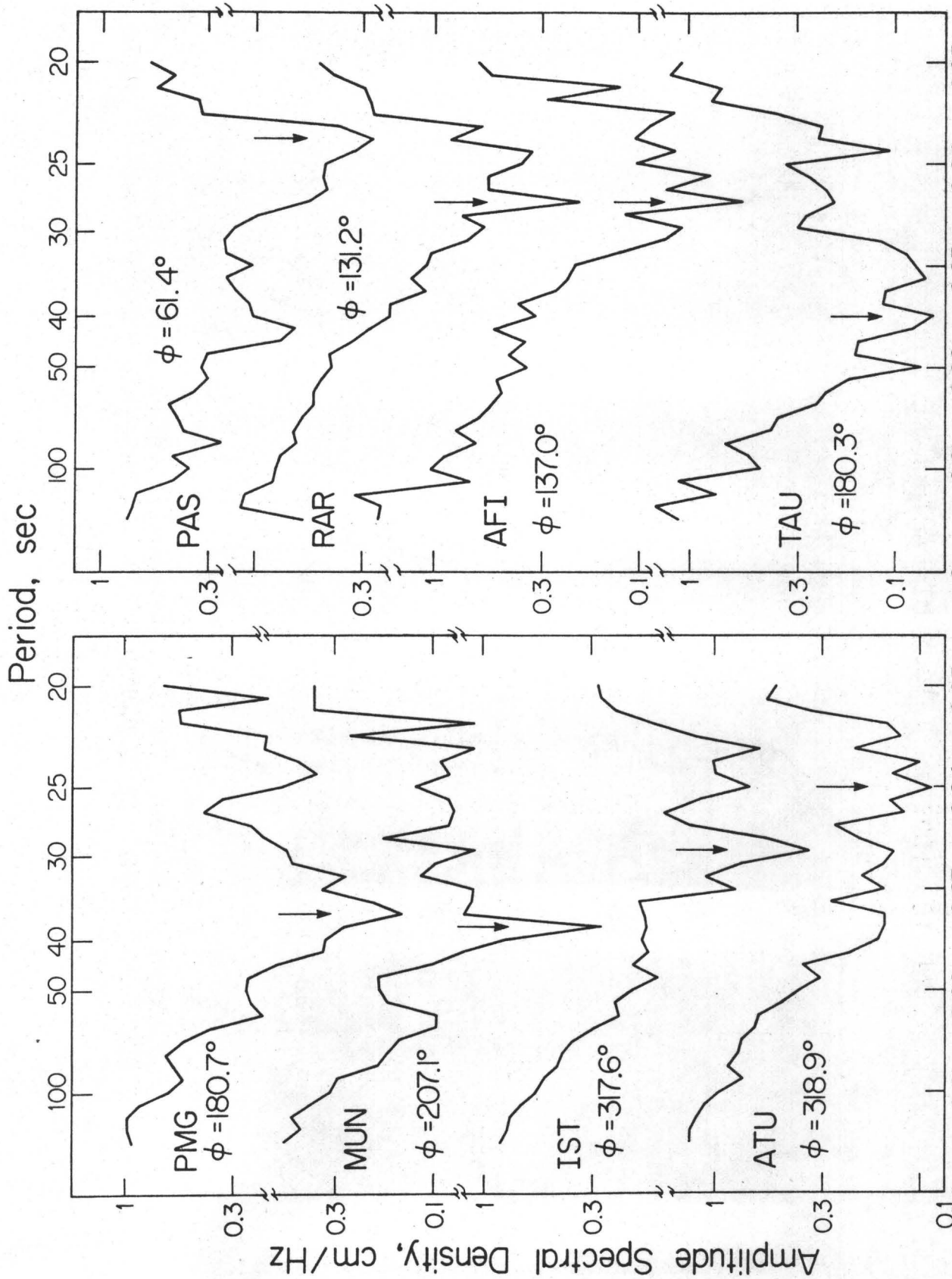

Period, sec
100
50
40
30
25
20
Frequency, Hz
0
0.01
0.02
0.03
0.04
0.05
Amplitude Spectral Density, cm/Hz
PMG
φ =180.7°
MUN
φ =207.1°
IST
φ =317.6°
ATU
φ =318.9°
PAS
φ =61.4°
RAR
φ =131.2°
AFI
φ =137.0°
TAU
φ =180.3°

Figure 8

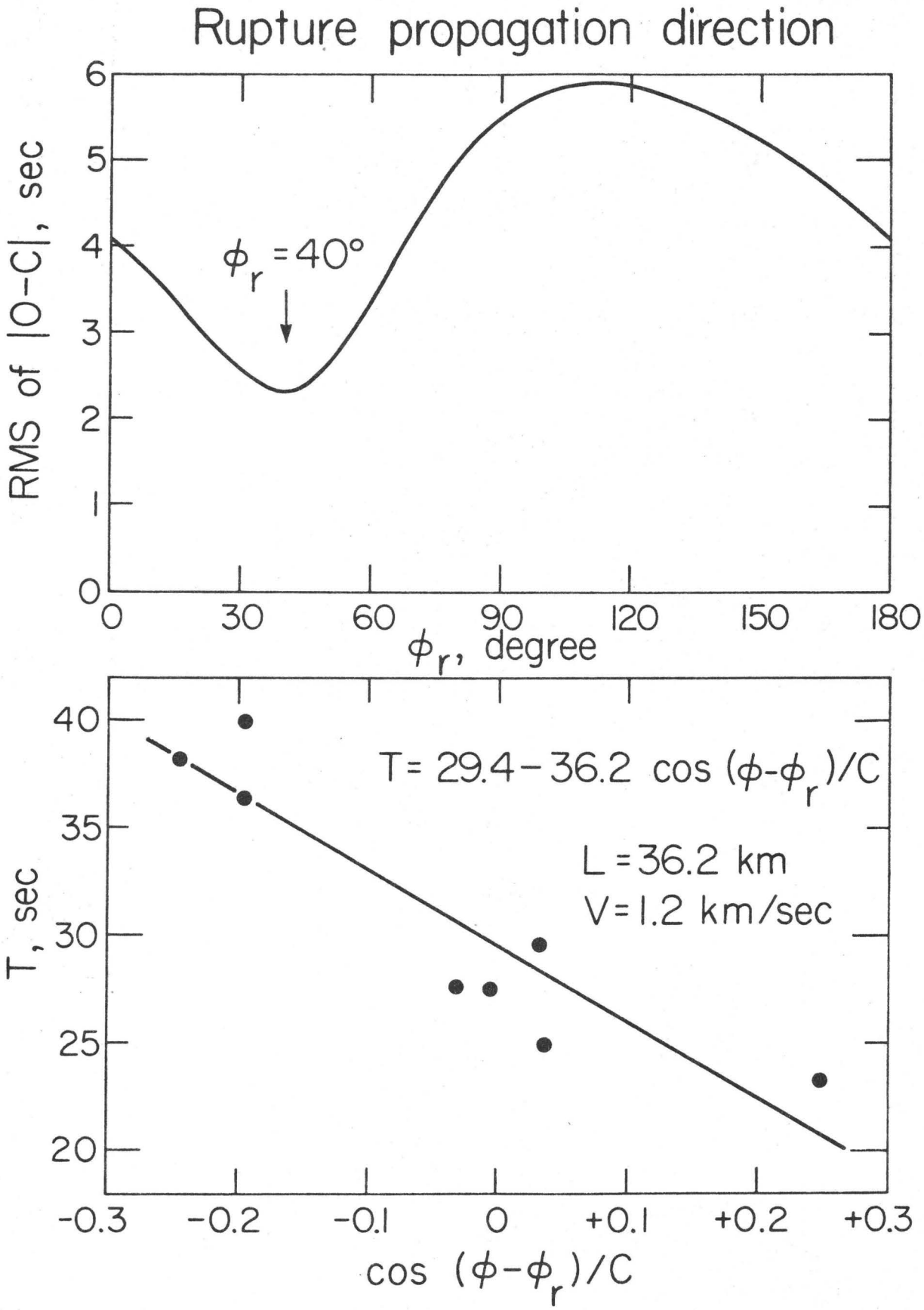

Rupture propagation direction
RMS of |O-C|, sec
ϕ_r = 40°
ϕ_r, degree
T = 29.4 − 36.2 cos (ϕ-ϕ_r)/C
L = 36.2 km
V = 1.2 km/sec
T, sec
cos (ϕ-ϕ_r)/C

Figure 9

b-VALUES AND $\omega^{-\gamma}$ SEISMIC SOURCE MODELS

Thomas C. Hanks

U. S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025

ABSTRACT

In this study, earthquake magnitude-frequency of occurrence statistics (b-values) and high-frequency spectral characteristics of crustal earthquakes ($\omega^{-\gamma}$ models) are the basis of a discussion of fault zone heterogeneity as it might be inferred from seismological observations. b-values of 1 for constant stress drop earthquakes imply the existence of a stress-drop potential function the spectral amplitude composition of which is white; its rms-value is determined by the average earthquake stress drop $\overline{\Delta\sigma}$. Several lines of evidence suggest that the $\omega^{-\gamma}$, $\gamma = 2$ seismic source model is the one generally applicable to describing the high-frequency radiation of crustal earthquakes in the far-field. This model is interpreted in terms of a white, random(?) distribution of dynamic stress differences $\tilde{\sigma}$ developing in the source region in the course of crustal faulting. The $\gamma = 2$ model predicts the rms acceleration for the San Fernando earthquake at Pacoima Dam to 50% and for the Kern County earthquake at Taft to 20%.

INTRODUCTION

The purpose of this discussion is to address the issue of how heterogeneity of crustal fault zones may be inferred from two classes of seismological observations, those pertaining to earthquake magnitude-frequency of occurrence relations (b-values) and high-frequency spectral characteristics of seismic sources ($\omega^{-\gamma}$ models). At the outset it is appropriate to emphasize that whatever inferences about fault-zone heterogeneity we may derive from these observations are indirect and non-unique. Fundamentally, these seismological observations can be related only to distributions of stress differences that develop along crustal fault zones (quasi-statically in the case of b-values, dynamically in the case of the high-frequency radiated field of an earthquake). These distributions of stress differences may arise from material heterogeneity, geometric irregularities, or dynamic propagation of rupture--or some combination of these and perhaps other variables along the fault zone.

Apart from the significance these seismological observations may play in unraveling fault-zone heterogeneity and its role in earthquake mechanics, the format of this session is especially well-suited for their discussion here. My colleagues who share the soap box with me today have separately considered b-values (Nur, 1977) and $\omega^{-\gamma}$ seismic source models (Geller, 1976) --and have reached distinctly different conclusions from the ones I shall present in this communication.

b-VALUES

In Hanks (1977), the relations between the frequency of occurrence N of earthquakes of magnitude M

$$\log N = a - bM, \qquad (1)$$

between seismic moment M_0 and M

$$\log M_0 = cM + d, \tag{2}$$

and between source radius r, earthquake stress drop $\Delta\sigma$, and M_0

$$M_0 = k\Delta\sigma r^3 \tag{3}$$

were algebraically combined to obtain

$$\log N = (a + \frac{bd}{c}) - \frac{b}{c} \log (k\Delta\sigma r^3). \tag{4}$$

In these relations a is a constant defined by the choice of region and time interval in which earthquakes are counted, d is an empirically determined constant, and k = 16/7 for a circular fault of radius r. It is empirically known that b is generally but not always very nearly equal to 1, irrespective of the choice of region and time interval in which earthquakes are counted. Also c is empirically known to be 1.5 whether local magnitude M_L (Thatcher and Hanks, 1973) or surface wave magnitude M_S (Kanamori and Anderson, 1975) is used in (2), although serious departures from (2) with c = 1.5 begin to develop for $M_S \simeq 7\frac{1}{2}$.

For b = 1 and c = 1.5, (4) reduces to

$$N = \frac{\text{const}}{(\Delta\sigma)^{2/3} r^2}. \tag{5}$$

If the earthquakes of the counted sample share the same $\Delta\sigma$, as they do on the average for all samples for which $\Delta\sigma$ have been determined (e.g., Hanks, 1977), earthquake magnitude-frequency of occurrence statistics reduce to a simple matter of geometrical scaling in terms of the reciprocal faulting area. In everyday English, the number of coins that fit in a ring of specified size depends on the area of the coin, and, provided that earthquake stress drops are

constant, (1) says nothing more than this.

For those earthquakes with $r \lesssim h' = h/\cos\delta$, where h is the seismogenic depth and δ is the fault dip, (5) can be interpreted in terms of a two-dimensional "stress-drop potential" function on the fault surface, the functional form of which is specified by a mean-square value determined by the average earthquake stress drop $\overline{\Delta\sigma}$ and a spectral composition with constant amplitudes at all wavelengths $\leq h'$; the stress-drop potential in a region of incipient faulting is realized as the earthquake stress drop at the time of faulting. On the average, such a stress-drop potential function will produce earthquakes with stress drop $\overline{\Delta\sigma}$ and their frequency of occurrence will scale as $1/r^2 (r \lesssim h')$, due to the constant spectral amplitudes at all wavelengths $\lesssim h'$ in two dimensions (Figure 1). Earthquakes with $\Delta\sigma$ both higher and lower than $\overline{\Delta\sigma}$ will occur, however, with certain probabilities. Whatever the origin of the stress differences recoverable in crustal faulting may be, then, they are more or less distributed in a white, possibly random fashion through the seismogenic zone. The stress-drop potential function, moreover, retains its mean-square value and white, random(?) characteristics until such time as the region of interest is faulted by a major earthquake, inasmuch as earthquakes with $r \lesssim h'$ cannot materially reduce the net stress on the fault surface.

When the fault length L becomes significantly greater than the fault width w, the two-dimensional character of the faulted surface collapses essentially to one, and it can be expected that the ideas presented above will no longer hold. For a seismogenic depth of 15 km and a vertical transform fault, we may estimate roughly that this will occur for $L \gtrsim 30$ km, or equivalently when $M_S \gtrsim 6\frac{1}{2}$. In particular (3) then takes the form

$$M_o = k'\Delta\sigma L w^2 . \qquad (6)$$

Moreover, (2) with c = 1.5 begins to fail at slightly larger M_s, 7 to 7½. Finally, as is well-known, M_s becomes an increasingly poorer measure of source strength for $M_o \gtrsim 10^{27}$ dyne-cm, or $M_s \gtrsim 7½$. As such, present uncertainties in estimating both "magnitude" and c at large magnitude preclude, at the present time, an extension of these results to the more nearly one-dimensional character of large and great earthquakes. But these difficulties in no way change the arguments given above for $M_s \lesssim 6½$ earthquakes for which $r \lesssim h'$.

In Nur (1977) $A(\lambda)$, where λ is wavelength, is a measure of the frictional heterogeneity and may be identified with the stress-drop potential function defined earlier and in Figure 1. Nur (1977) finds that $A(\lambda) \sim \lambda$, where we have found that the spectral amplitude composition of the stress-drop potential function is independent of wavelength. Had Nur (1977) used (3) to relate M_o, $\Delta\sigma$, and r, rather than the one-dimensional form analogous to (6) that he did use, he would have found $A(\lambda) \sim$ const -- as we have here.

$\omega^{-\gamma}$ SEISMIC SOURCE MODELS

Simple models of the seismic source are generally characterized, in spectral form, by a long-period level Ω_o proportional to M_o, a corner frequency f_o proportional to r^{-1}, and a high-frequency spectral decay of the form $(f/f_o)^{-\gamma}$. (In the following discussion, we denote frequency with f in hertz rather than ω in radians/sec; "high-frequency" means frequencies high with respect to f_o). The corner frequency f_o, fundamentally, is closely allied with the reciprocal duration of faulting T_d^{-1}, but it is well-known that several "faulting durations" can be defined, in particular those associated with the fault length, fault width, and the rise time of a propagating displacement discontinuity. Depending on the faulting geometry and rise time characteristics, the associated corner frequencies can be well separated,

leading to more complicated high-frequency spectral amplitude decay (that is, γ is a function of frequency). Moreover, by making the displacement discontinuity a smooth enough function of time, γ can become arbitrarily large at high enough frequencies. Whether, in fact, a generally applicable source representation of high-frequency spectral characteristics exists within the infinity of possibilities is, as yet, theoretically controversial and observationally unresolved.

More as a matter of convenience than a matter of hard fact, high-frequency spectral characteristics of seismic sources are generally discussed in terms of Ω_o and f_o related by the constant stress drop assumption ($\Omega_o f_o^3$ = const in the context of the Ω_o-f_o relations of Hanks and Thatcher, 1972) and $\gamma = 2$ (the ω-square model) or $\gamma = 3$ (the ω-cube model, in the terminology of Aki, 1967).

Figure 2 schematically illustrates the $\gamma = 2$ and $\gamma = 3$ seismic source models in terms of two idealized far-field shear-wave displacement spectra at the same distance R. In both the $\gamma = 2$ and $\gamma = 3$ cases, the two earthquakes have been assigned the same $\Delta\sigma$, so the corner frequencies lie on a line of slope -3 in these log-log plots. In both cases, the larger event (1) has Ω_o and M_o 3 orders of magnitude larger than the smaller event (2), and $f_o^{(1)}$ is ten times smaller than $f_o^{(2)}$.

At frequencies greater than $f_o^{(2)}$, spectral amplitudes are ten times greater for event (1) than event (2) in the $\gamma = 2$ case but are the same in the $\gamma = 3$ models. How do we interpret these models in terms of time-domain amplitudes, recognizing that $T_d^{(1)} \simeq 10\ T_d^{(2)}$? Figure 3 presents the extreme interpretations. Here, for purposes of illustration, we have taken $f_o^{(1)}$ = 0.05 Hz, $f_o^{(2)}$ = 0.5 Hz, $T_d^{(1)}$ = 20 sec., $T_d^{(2)}$ = 2 sec., and are investigating possible interpretations of 1-second time-domain amplitudes.

Figure 3a is the interpretation for the $\gamma = 2$ earthquakes if the 1-second energy arrives more or less continuously over the complete faulting duration. In this case, 1-second spectral amplitudes for the larger event are ten times greater than for the smaller event, but the 1-second time-domain amplitudes are the same for both events--they have the same m_b. If all the 1-second energy arrives at the same time, however, the 1-second time-domain amplitudes and m_b of the larger event are ten times larger (Figure 3b).

For the $\gamma = 3$ earthquakes, 1-second spectral amplitudes must be the same. In Figure 3c, this is achieved in a manner analogous to Figure 3a, but now 1-second time domain amplitudes for event (1) are ten times smaller than for event (2); that is m_b must decrease with M_o. Figure 3d is the analogue to Figure 3b; here 1-second time domain amplitudes for the two earthquakes are the same; they have the same m_b.

Interpretation c is certainly unacceptable: m_b does not systematically decrease with increasing M_o. Neither, however, does m_b increase beyond $m_b \simeq$ 6½ to 7, and interpretation b is also inappropriate, at least for $M_o \gtrsim 10^{26}$ dyne-cm. One's preference for interpretation a or d, and thus one's preference for $\gamma = 2$ or $\gamma = 3$ seismic source models, then depends on whether one believes that all (or most) of the 1-second energy arrives more or less continuously through T_d ($T_d > 1$ sec.) or arrives more or less impulsively in a $\sim$ 1-second window (and in the case of m_b, the first one or two such windows) no matter what the value of T_d. It is appropriate to recall, now, that both possibilities are extreme interpretations, and the truth, in most cases, should lie somewhere in between. Even so, when $T_d >> 1$ sec. in the case of the larger earthquakes ($M_s \gtrsim 6½$), it is clear that Figure 3d is much more the exception than the rule, as almost all short-period seismograms of large and great earthquakes reveal. Thus, I conclude, as Aki (1967) did ten years

ago, that $m_b - M_s$ data support the $\gamma = 2$ model, in the interpretation of Figure 3a, as the one generally (but certainly not always) applicable to the representation of high-frequency spectral characteristics.

With the assumptions that (1) fault propagation in both coordinates of the fault plane is equally phase coherent and (2) the source displacement time function is a propagating ramp of finite duration (with singular particle accelerations), Geller (1976) correctly followed Haskell (1964) to obtain $\gamma = 3$ at high frequencies. His justification of this model with existing $m_b - M_s$ data is not correct, however, because he assumed that m_b and M_s faithfully represent spectral amplitudes at 1- and 20-second periods, respectively, across the entire range of magnitudes observable at teleseismic distances. Geller (1976) notes that "it is not exactly correct" to do this; quite generally it is not at all correct to do this, except for the smaller earthquakes ($M \lesssim 5$) for which $f_o \gtrsim 1$ Hz. In the latter case, both m_b and M_s become long-period measurements, proportional to M_o, but then, of course, $m_b - M_s$ data carry no information at all about high-frequency spectral characteristics of earthquake sources.

There are, in addition, several other observations that are in general accord with the high-frequency spectral characteristics of the ω^{-2} model. First, the difference of a factor of 20 in the maximum m_b of ~ 7.0 and maximum M_s of ~ 8.3 is "exactly" predicted by the $\gamma = 2$ model (because of the period shift in the amplitude measurement from 1 to 20 seconds) provided that $m_b \simeq M_s$ at $M \simeq 7$. In the "latest" form of the linear relations between m_b and M_s, $m_b = M_s$ at 6.75 (Richter, 1958, p. 348). Secondly, the same arguments used above to justify the $\gamma = 2$ model in terms of m_b and M_s data, and the upper limit to each, may also be used to explain why peak acceleration data at fixed R are such a weak function of magnitude (Figure 4). The analogy is some-

what complicated by the fact that the peak acceleration may occur anywhere in the record, whereas the m_b amplitude cannot, and thus are susceptible to biasing by chance interference by two or more phases coincidentally arriving at the same time. At some fixed high frequency, the probability of this occurring plainly increases with increasing T_d and M, without reflecting intrinsic changes in high-frequency source excitation. Finally, the high-frequency spectral characteristics of the San Fernando earthquake are very well-known, even at frequencies two orders of magnitude greater than f_o, because of the large number of strong-motion accelerograms that recorded this earthquake at local distances. Figure 5 and 6 suggest that the simplest possible interpretation of these data is the $\gamma = 2$ model, although more complicated interpretations are possible (and perhaps warranted).

INTERPRETATION OF THE ω^{-2} MODEL

Figure 7 is the $\gamma = 2$ model for shear-wave displacement spectral amplitudes in terms of shear-wave <u>acceleration</u> spectral amplitudes. In the frequency band $f_o \leq f \lesssim f_{max}$, shear-wave acceleration spectral amplitudes are constant (f_{max} is determined by setting the argument of the exponent equal to 1 in the usual attenuation relation $e^{\frac{-\pi f R}{Q\beta}}$ for a given choice of R and Q). That is, the associated ground motion time history may be regarded as band-limited ($f_o \leq f \leq f_{max}$), finite duration ($0 \leq t - R/\beta \leq T_d$), white noise(?) in acceleration, where (again) the randomness has simply been assumed but not unreasonably so in view of the generally chaotic nature of strong-motion accelerograms of $M \gtrsim 5$ earthquakes.

Hanks and Johnson (1976) developed the following linear relation between the amplitude $\ddot{u}$ of any acceleration pulse at R and the dynamic stress difference $\tilde{\sigma}$ giving rise to it in the source region

$$\ddot{u} = \frac{1}{\rho} \frac{\tilde{\sigma}}{R} \quad , \tag{7}$$

where ρ is density. We may then interpret the constant spectral amplitudes of acceleration ($f > f_o$) for the $\gamma = 2$ model in terms of a white, random(?) distribution of dynamic stress differences ($\lambda \leq r$) arising in the course of faulting that generates the earthquake.

The coincidence of this result with the one inferred earlier from b-values for the distribution of quasi-static stress differences on crustal fault zones is interesting--but quite possibly meaningless. While the whiteness of both distributions seems to be a reasonably firm result, the random character assumed in both cases is simply a guess. Moreover, there is no justification for assuming, at the present time, that the whiteness of $\tilde{\sigma}$ and $\Delta\sigma_p$ arise from the same physical condition. If this should be the case, however, b-values different from 1 would be expected to occur in association with γ different from 2. In the case of the Oroville aftershocks b was significantly less than 1 (0.61 according to Morrison et al., 1976), and 13 of 21 small ($1 \lesssim M_L \lesssim 2\frac{1}{2}$) but well-recorded aftershocks had $\gamma \gtrsim 3$ (the other 8 had $\gamma \simeq 2$) (J. B. Fletcher, personal communication). For the mining-induced seismicity at the East Rand Proprietary Mines in South Africa "γ is 3 more often than 2 and $b \simeq 0.6$" (A. McGarr, personal communication). In both cases, the interpretation would be that the spectral composition of $\Delta\sigma_p$ and $\tilde{\sigma}$ is deficient in high-frequency amplitudes, relative to those for $b = 1$ and $\gamma = 2$. $\Delta\sigma_p$ deficient in high-frequency amplitude is seemingly a logical (if not always necessary) prelude to a larger dimension shock, and a pursuit of this matter in connection with the poorly documented claim that b is less than 1 prior to larger earthquakes in the same region as the counted sample would be a potentially interesting area of earthquake prediction research.

ESTIMATION OF HIGH-FREQUENCY STRONG GROUND MOTION

The interpretation of the $\gamma = 2$ model given above allows, through an application of Parseval's theorem, a determination of the rms acceleration

through T_d that is based only on the properties Ω_o, f_o, and f_{max}--or $\Delta\sigma$, f_o, and f_{max}. Parseval's theorem is

$$\int_{-\infty}^{\infty} |a(t)|^2 dt \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} |\tilde{a}(\omega)|^2 d\omega \tag{8}$$

where $a(t)$ is the acceleration time history and $\tilde{a}(\omega)$ is its amplitude spectrum. Ignoring contributions outside the ranges $0 \le t-R/\beta \le T_d$ and $f_o \le f \le f_{max}$, we write (8) as

$$\int_0^{T_d} |a(t)|^2 dt = \frac{2}{2\pi} \int_{2\pi f_o}^{2\pi f_{max}} |\tilde{a}(\omega)|^2 d\omega. \tag{9}$$

The rms acceleration is

$$a_{rms} = \left[\frac{1}{T_d} \int_0^{T_d} |a(t)|^2 dt \right]^{1/2}, \tag{10}$$

or with (9)

$$a_{rms} = \left[\frac{1}{T_d} \frac{1}{\pi} \int_{2\pi f_o}^{2\pi f_{max}} |\tilde{a}(\omega)|^2 d\omega \right]^{1/2}. \tag{11}$$

For $f_o \le f \le f_{max}$, $\tilde{a}(\omega) = \Omega_o (2\pi f_o)^2$ so (11) becomes

$$a_{rms} = \left[\frac{2(2\pi)^4}{T_d} \Omega_o^2 f_o^4 (f_{max} - f_o) \right]^{1/2} \tag{12}$$

For earthquakes with $M > 5$ observed at $R \lesssim 20$ km, $f_{max} >> f_o$. This together with the approximation

$$f_o = \frac{1}{T_d} \tag{13}$$

gives

$$a_{rms} = \left[2(2\pi)^4 \Omega_o^2 f_o^5 f_{max} \right]^{1/2} \tag{14}$$

which we write as

$$a_{rms} = \sqrt{2}\,(2\pi)^2 \Omega_o f_o^3 \sqrt{\frac{f_{max}}{f_o}} \tag{15}$$

to factor out the stress drop quantity $\Omega_o f_o^3$.

For the Brune (1970, 1971) model, Hanks and Thatcher (1972) give

$$\Omega_o f_o^3 = \frac{\Delta\sigma}{106\,\rho R} \tag{16}$$

so (15) becomes, in terms of $\Delta\sigma$,

$$a_{rms} = \frac{\sqrt{2}(2\pi)^2}{106} \frac{\Delta\sigma}{\rho R} \sqrt{\frac{f_{max}}{f_o}} . \tag{17}$$

Table 1 compares a_{rms} values estimated from (17) with those "observed" for the San Fernando earthquake at Pacoima Dam and the Kern County earthquake at Taft. The "observed" values are those given in Volume I, Part A, of the series "Strong Motion Earthquake Accelerograms", corrected by (record length/T_d)$^{\frac{1}{2}}$ to estimate the (larger) a_{rms} value that occurs in the time interval of the S-wave arrival through the S-wave arrival plus T_d. Because the accelerations are non-zero outside of this interval, the "observed" values are overestimates of the actual a_{rms} values, $0 \leq t - R/\beta \leq T_d$. In any case, however, the agreement between the estimated and "observed" values is remarkable, by conventional seismological standards in estimating high-frequency amplitudes, being approximately 50% in the case of the San Fernando earthquake and 20% in the case of the Kern County earthquake.

Equation (17) is built on the far-field representation of the $\gamma = 2$ model, and so it may be expected to fail close to a large enough shock, unless we can be clever enough to break big earthquakes into small enough analytical chunks. If the major part of the radiated energy comes from depths of several km or greater, however, we are always in the far-field for radiation at frequencies of several hz and greater, which make the principal contribution to (15) at R $\sim$ 10 km. It thus appears that (15), or a discrete sum based upon it, holds considerable promise for accurately and reliably estimating high-frequency

strong ground motion for those earthquakes for which the $\gamma = 2$ model is valid. With the additional assumption of constant stress drop, a_{rms} for any hypothetical earthquake can be estimated with knowledge of only the source property f_o and the path properties R and Q, exclusive of anomalous site effects.

References

Aki, K., Scaling law of seismic spectrum, J. Geophys. Res., 72, 1212-1231, 1967.

Berrill, J.B., A study of high-frequency strong ground motion from the San Fernando earthquake, Ph.D. Thesis, California Institute of Technology, Pasadena, 1975.

Brune, J.N., Tectonic stress and the spectra of seismic shear waves, J. Geophys. Res., 75, 4997-5009, 1970.

Brune, J.N., Correction, J. Geophys. Res., 76, 5002, 1971.

Geller, R.J., Scaling relations for earthquake source parameters and magnitudes, Bull. Seism. Soc. Am., 66, 1501-1523, 1976.

Hanks, T.C., Earthquake stress drops, ambient tectonic stresses, and stresses that drive plate motions, PAGEOPH, 115, 441-458, 1977.

Hanks, T.C., and D.A. Johnson, Geophysical assessment of peak accelerations, Bull. Seismol. Soc. Amer., 66, 959-968, 1976.

Hanks, T.C., and W.R. Thatcher, A graphical representation of seismic source parameters, J. Geophys. Res., 77, 4393-4405, 1972.

Haskell, N. A., Total energy and energy spectral density of elastic wave radiation from propagating faults, Bull. Seism. Soc. Am., 54, 1811-1841, 1964.

Kanamori, H., and D.L. Anderson, Theoretical basis of some empirical relations in seismology, Bull. Seismol. Soc. Amer., 65, 1073-1096, 1975.

Morrison, P.W., B.W. Stump, and R. Uhrhammer, The Oroville earthquake sequence of August 1975, Bull. Seism. Soc. Am., 66, 1065-1084, 1976.

Nur, A., Nonuniform friction as a physical basis for earthquake mechanics: a review, PAGEOPH, in press, 1977.

Richter, C.F., Elementary Seismology, pp. 768, W. H. Freeman, San Francisco, Calif., 1958.

Thatcher, W., and T.C. Hanks, Source parameters of southern California earthquakes: Jour. Geophys. Research, 78, 8547-8576.

Table 1

Comparisons of Estimated and Observed a_{rms} Values

San Fernando Earthquake Feb. 9, 1971; M_L = 6.4		Kern County Earthquake July 21, 1952; M = 7.7
10^{26} dyne-cm	M_o	2 x 10^{27} dyne-cm
10 km	r	25 km
50 bars	$\Delta\sigma$	60 bars
0.1 Hz	f_o	(0.04 Hz)
at Pacoima Dam		at Taft, California
~ 10 km	R	~ 40 km
25 Hz	f_{max}	7 Hz
	a_{rms}	
140 cm/sec^2	estimated	36 cm/sec^2
220, 240 cm/sec^2	"observed"	42, 42 cm/sec^2

Figure Captions

Fig. 1 Interpretation of the stress-drop potential function when b = 1. The stress-drop potential function is the areal distribution of all the stress drops (to be realized at the time of discrete faulting for individual events) of the incipient earthquake population.

Fig. 2 The ω^{-2} and ω^{-3} source models for constant stress drop earthquakes.

Fig. 3 Interpretation of the ω^{-2} and ω^{-3} models in the time domain. Event (2) (left-hand side) has a duration of 2 seconds, and event (1) (right-hand side) has a duration of 20 seconds, although only 10 of these are shown in b and d. The range of possible 1-second time-domain amplitudes for the ω^{-2} and ω^{-3} models are graphically illustrated here and discussed in the text. Relative 1-second amplitudes are given in two groups of four (one each for the ω^{-2} model, a and b, and the ω^{-3} model, c and d); the choice of 1 in the upper left corner of each square is arbitrary.

Fig. 4 Peak acceleration at $R \underline{\sim} 10$ km as a function of magnitude, as presented in Hanks and Johnson (1976).

Fig. 5 Fourier amplitudes of acceleration determined from 15-second windows, beginning with the S-wave arrival, of strong motion accelerograms of the San Fernando earthquake as a function of frequency (0.4, 1, 2, 4, 8, and 16 Hz) and R. Data from transverse components are indicated as x, from radial components as o. The curved lines are least-squares fitted to the data and are of the form $\tilde{a}(f,R) = \frac{\tilde{a}_o(f, 1\ \text{km})}{R} e^{-\frac{\pi f R}{Q \beta}}$ where β = 3.2 km/sec and Q = 330. Values

of $\tilde{a}_o$ are plotted as a function of frequency in Fig. 6. Modified from Berrill (1975).

Fig. 6 $\tilde{a}_o$ (f, 1 km) at 0.4, 1, 2, 4, 8, and 16 Hz at three azimuth ranges from the San Fernando earthquake compared to those for the $\gamma = 2$ model at an average azimuth (in the sense of a point source radiation pattern). Modified from Berrill (1975).

Fig. 7 Acceleration amplitude spectra at R for two constant stress drop earthquakes with attenuation explicitly shown.

$\Delta\sigma_p(x,y)$ is the stress-drop potential, a white, random(?) function of the fault plane coordinates x and y

$$|\widetilde{\Delta\sigma_p}(k,l)| = \left|\int_{-\infty}^{\infty}\int \Delta\sigma_p(x,y)\, e^{-ikx} e^{-ily}\, dx\, dy\right| = \text{const}$$

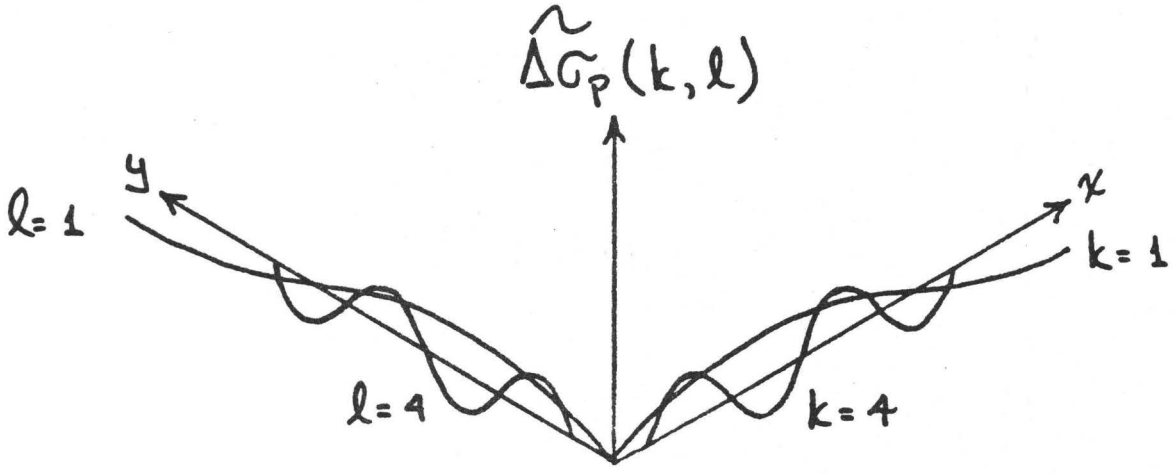


Figure 1

ω^{-3}
ω^{-2}
$\Omega_0^{(1)}$
$f_0^{(1)}$
$\Omega_0^{(2)}$
$f_0^{(2)}$
log frequency
log $\frac{\Omega}{2\pi}$

Figure 2

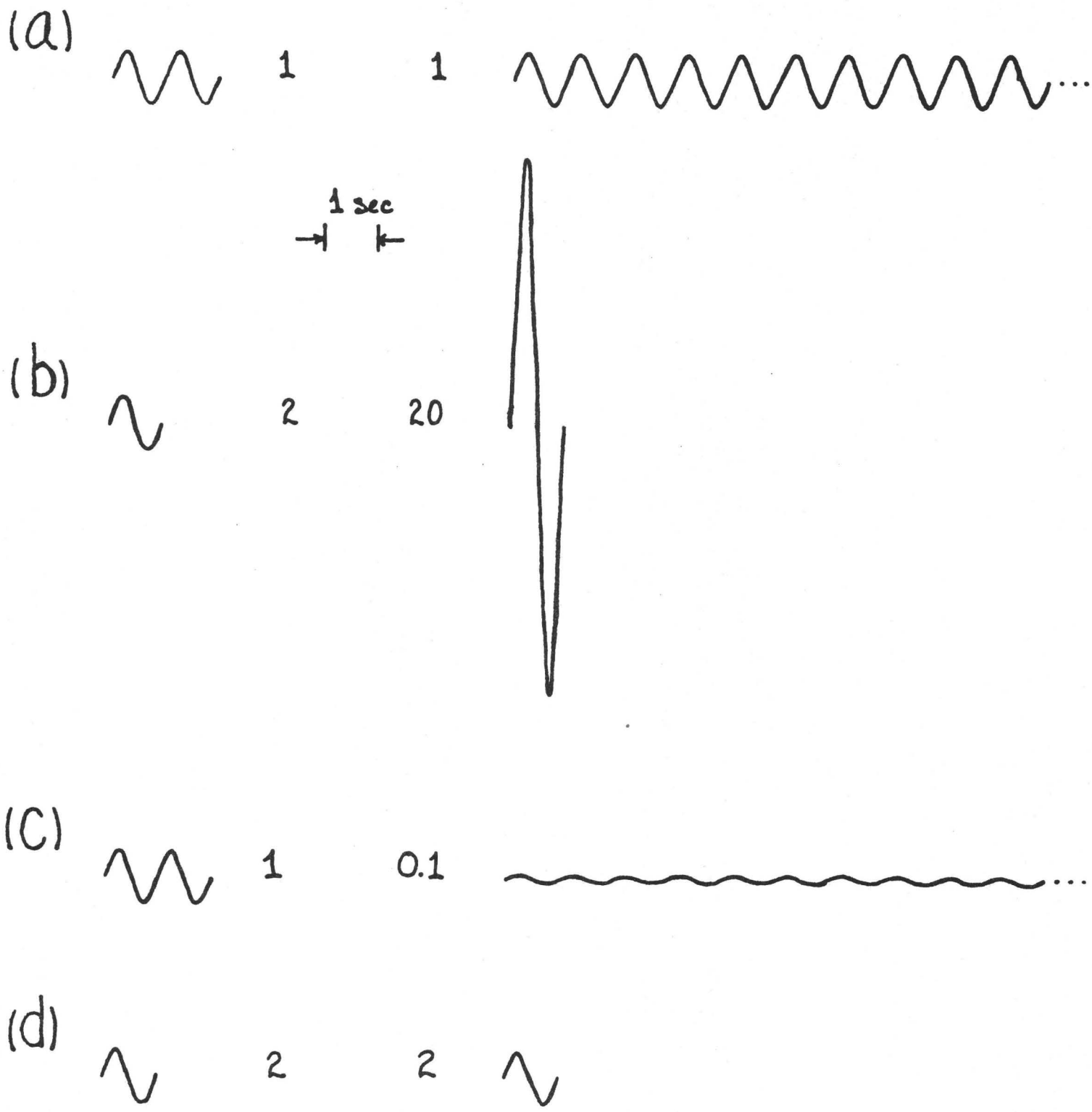


Figure 3

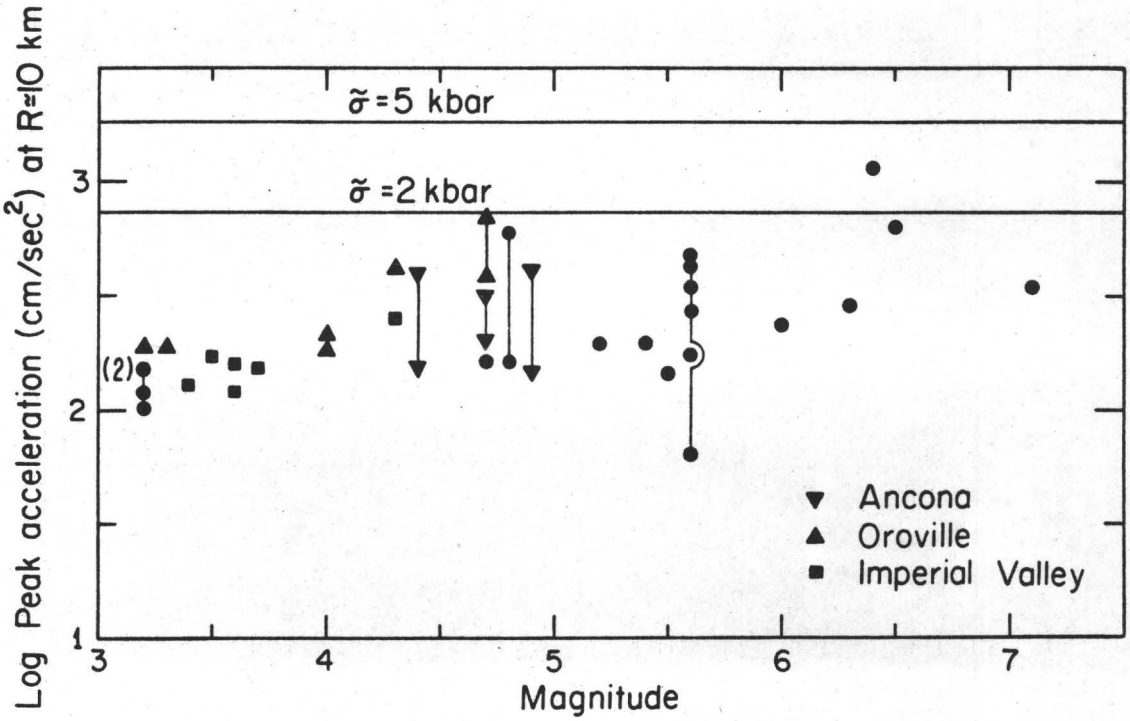

$\tilde{\sigma}$ =5 kbar
$\tilde{\sigma}$ =2 kbar
(2)
Ancona
Oroville
Imperial Valley
Log Peak acceleration (cm/sec^2) at R=10 km
Magnitude
1
2
3
3
4
5
6
7

Figure 4

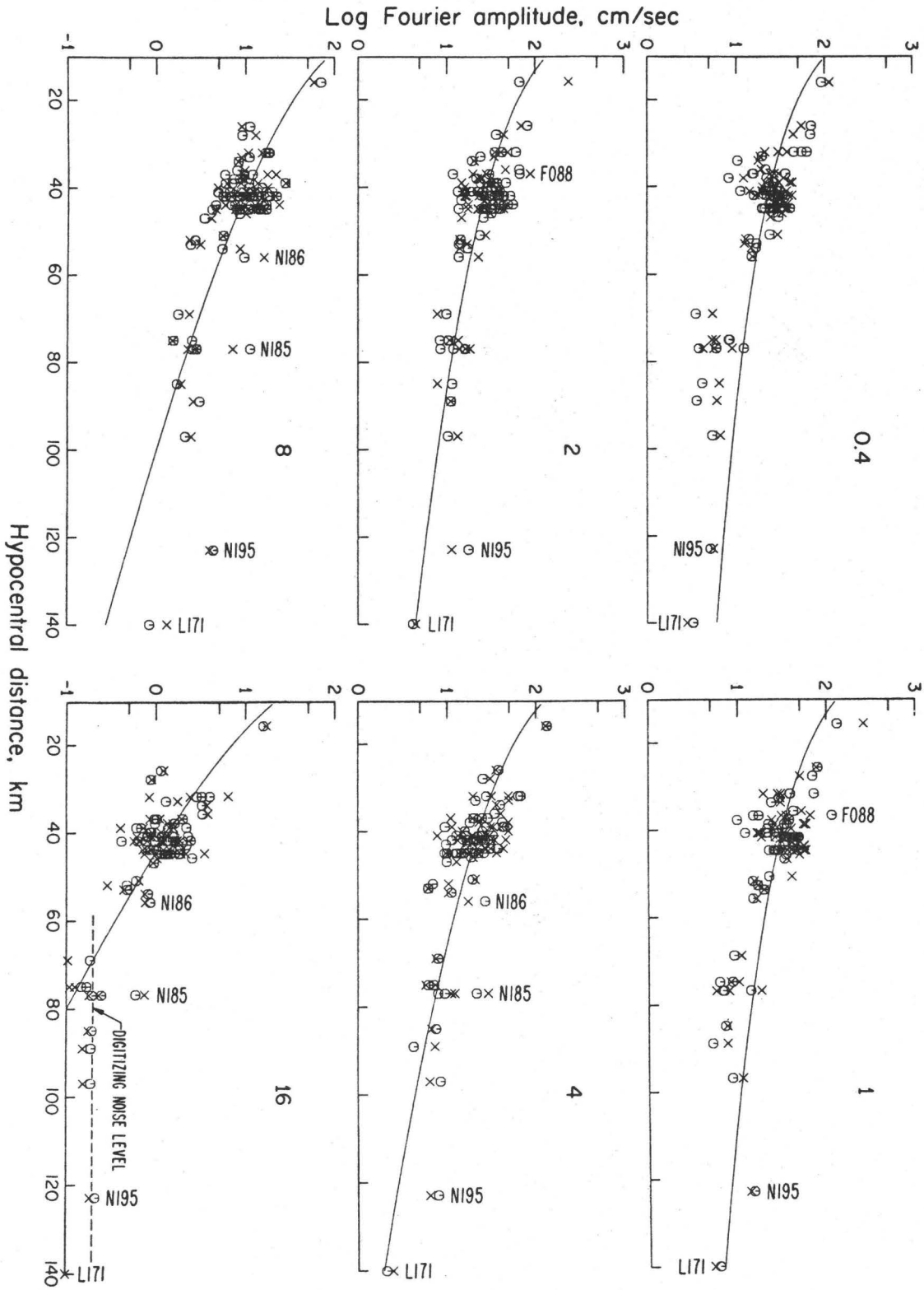

Log Fourier amplitude, cm/sec
Hypocentral distance, km
F088
N186
N185
N195
L171
DIGITIZING NOISE LEVEL
0.4
1
2
4
8
16

Figure 5

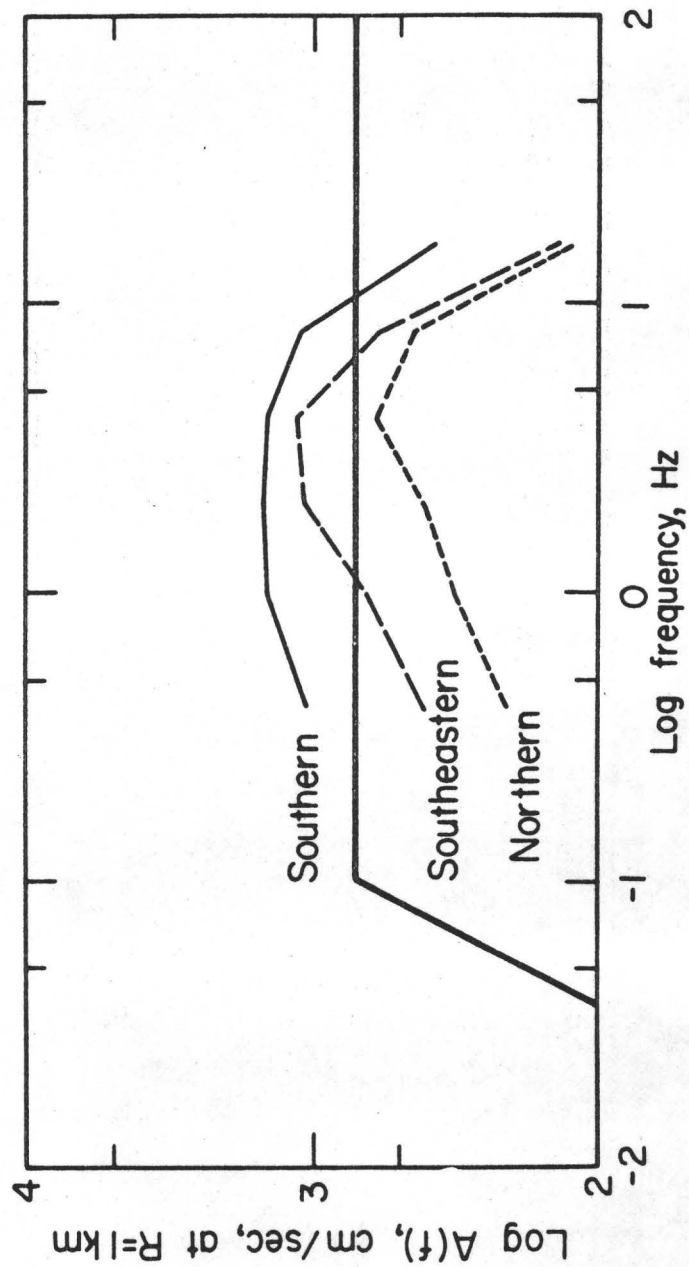

Log A(f), cm/sec, at R=1km
4
3
2
-2
-1
0
1
2
Log frequency, Hz
Southern
Southeastern
Northern

Figure 6

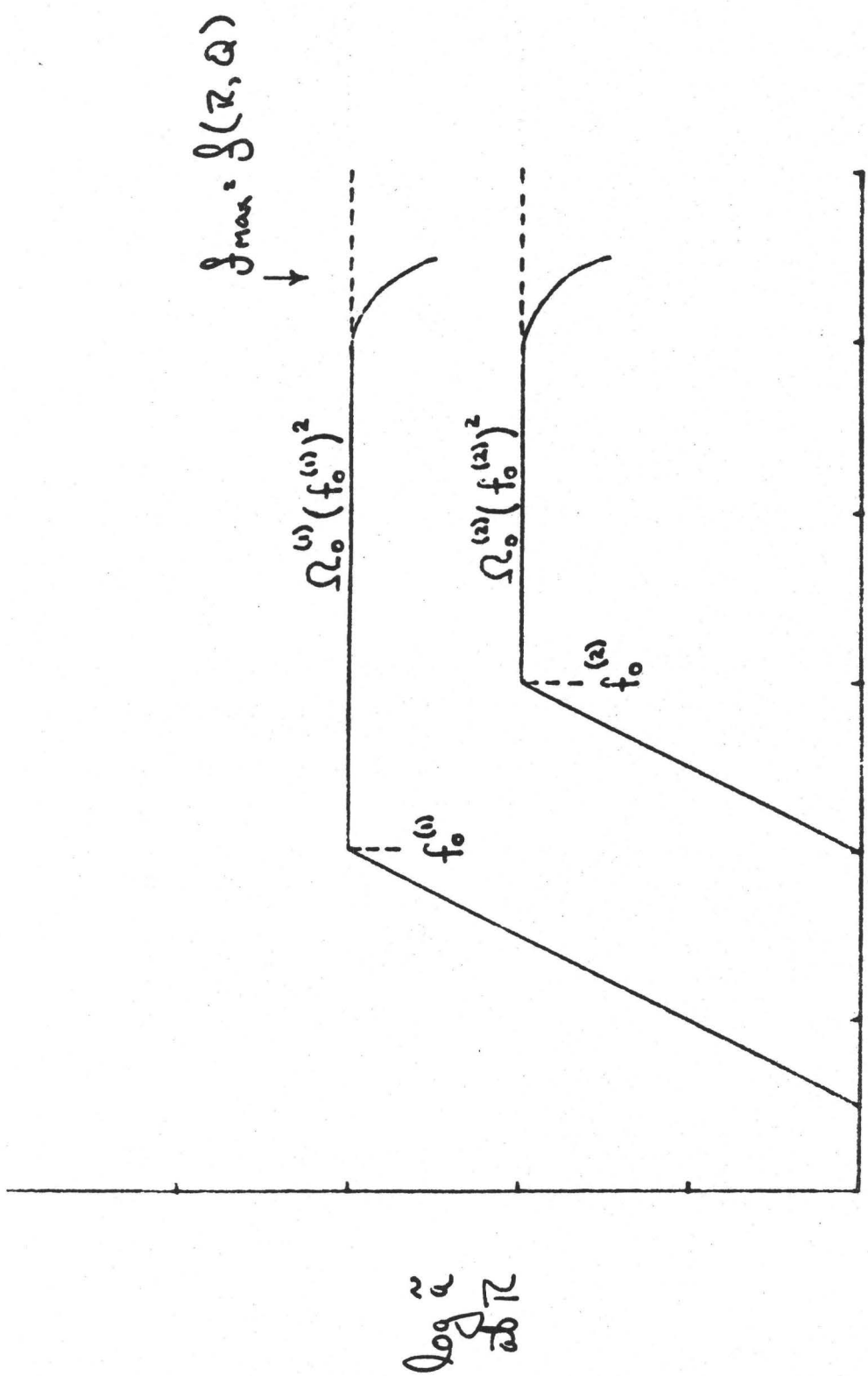

$f_{max}: f(r, Q)$
$\Omega_0^{(1)}(f_0^{(1)})^2$
$\Omega_0^{(2)}(f_0^{(2)})^2$
$f_0^{(1)}$
$f_0^{(2)}$
Log Frequency
$\log \frac{a^2}{4\pi r^2}$

Figure 7

VARIABLE FRICTION AND STRESS IN ONE DIMENSIONAL MODEL

A SUMMARY

by

Moshe Israel and Amos Nur

Department of Geophysics
Stanford University
Stanford, California 94305

THE MODEL

The one dimensional model (Fig.1), introduced by Knopoff et al. (1973)† is assumed. Using this model, we examine the effects of variable stress distribution and energy absorbed at the crack tip. Mathematically, one has to solve the equation:

$$\rho(x)\frac{\partial^2 u}{\partial t^2} = \mu(x)\frac{\partial^2 u}{\partial x^2} - A(x)u - B(x)\frac{\partial u}{\partial t} + F(x) \qquad (1)$$

where $\mu(x)$ is the rigidity, $\rho(x)$ the density, x is spatial coordinate and t the time. The terms $A(x)$ and $B(x)$ take into account the interaction of the fault with the surrounding medium, $F(x)$ is the stress drop, i.e., the difference between tectonic stress and dynamic stress.

Equation (1) may be solved if we assume that the rupture velocity, $\frac{d\xi}{dt}$, is known in advance ($\xi(t)$ is the position of the crack tip). However, from a physical point of view, the rupture velocity is determined by the stress drop $F(x)$ and the rate of energy absorption per unit distance at the crack, γ. Using some results of Knopoff et al. we find:

$$\gamma = \frac{\mu}{2} u_x^2 \left(1 - \frac{\dot{\xi}^2}{c^2}\right) \qquad (2)$$

where c is the wave velocity. We may solve (1) by specifying the rate of energy absorption, γ, or the rupture velocity, $\frac{d\xi}{dt}$.

---

†Knopoff, L., Mouton, J.U. and Burridge, R., 1973. The dynamic of a one-dimensional fault in the presence of friction, Geophys. J. R. astr. Soc., 35, 169-184.

We assume that static friction effects the motion in the following way: if for some $\bar{x}$ and some $\bar{t}$ the particle velocity $\frac{\partial u}{\partial t}$ vanishes, then $\frac{\partial u}{\partial t} = 0$ for $t > \bar{t}$ at $x = \bar{x}$. In other words, we assume that static friction is large enough to prevent negative velocity. Generally, the acceleration does not vanish when the velocity does, and therefore the solution involves discontinuity of the second partial derivatives.

THE METHOD

The characteristic method is used to solve the hyperbolic equation (1). The method is numerically stable, furthermore discontinuity in partial derivatives are treated in a natural way. We assume a constant wave velocity equal to 1. In this case, the grid points are given by the intersection of the characteristics:

$x = t + \alpha$ (α characteristics),

$x = -t + \beta$ (β characteristics).

The grid and the boundary conditions are described in Fig. 2.

Let us consider the healing process. Suppose that at some point the velocity of some particle vanishes. At this point there is usually a jump in the acceleration from some negative value to zero. This jump propagates at a speed equal to the wave velocity. Now, the solution is not effected by this jump while the healing velocity is larger than the wave velocity. At a point Q, where the healing process reaches the sound velocity, the discontinuity plays an important roll. From this point, the particles are accelerated and the motion continues in the remaining segment. Referring to fig. 2, we note that to the right of Q, a boundary line in which $\frac{\partial u}{\partial t} = 0$

is formed. The discontinuity at Q propagates along the α_1 characteristics, the β_1 characteristics and so on.

NUMERICAL RESULTS

We assume, for the purpose of numerical computation, that $\mu(x) = 1$, $\rho(x) = 1$. Different positive values of A(x) and B(x) were assumed ($0 \leq A(x) \leq 3$, $0 \leq B(x) \leq 3$). Since the conclusions drawn do not depend on the specific values of A(x) and B(x), we present solutions where A(x) = 0 and B(x) = 0. Let us examine first the case in which the stress drop is variable.

CASE A

The rate of energy absorption is given by

$$\gamma = \xi^3/8 \tag{3}$$

The stress drop is

$$F(x) = 1 - \varepsilon \sin(2\pi x), \tag{4}$$

where $\varepsilon = 0., 0.46, 0.5$.

The solutions are described in figures 3 to 12. We note the following features of the solutions:

(a) When $\varepsilon = 0$ and 0.46 the fault extends itself out to about x = 1.3. When $\varepsilon = 0.5$ the extension is only out to x = 0.5 (see figure 4 for rupture velocity). This is certainly a non linear effect associated with rupture propagation.

(b) We denote the difference between the initial and final stress by dynamic stress drop. From Figures 3 and 5, we conclude that overshooting

always exists. It is larger near the end of the fault and in the cases of variable stress drop.

(c) It may be shown that the stress drop of a subsequent earthquake is given by $\rho\frac{\partial^2 u}{\partial t^2} + T$, where $\frac{\partial^2 u}{\partial t^2}$ is conputed at the instant of healing. T is the tectonic stress applied on the system between the two successive earthquakes. If T is independent of x, then particle accelerations (Figure 6) describes the form of the stress drop function for a subsequent earthquake. We note that the stress drop of an expected earthquake is generally smoother than that of the actual earthquake (consider the cases $\varepsilon = 0.46$, $\varepsilon = 0.5$).

(d) The healing process is continuous in all cases (Fig. 7).

(e) Particle accelerations and velocities are significantly dependent on stress drop variation (Figures 9 to 12).

CASE B

We assume a constant stress drop given by

$$F(x) = 1 \tag{5}$$

The rupture velocity is given by

$$\dot{\xi}(t) = \eta(1 + \sin(7\pi t/4)) + (1 - 2\eta)(1 - t/2) \tag{6}$$

when $0 \leq t \leq 2$ and zero otherwise. η is a parameter which assumes the values: 0, 0.1, 0.2, 0.4.

The computed rate of energy absorbed at the crack is given in Figure 13. We may regard the energy absorbed as completely specifying the physical con-

dition of the system. The rupture velocity (equation 6) will result from this condition.

The numerical results are desrcibed in Figures 13 to 20. We note that:

(a) Variations in the absorbed energy cause discontinuities in the healing process (see Figures 13, 15 when $\eta = 0.2, 0.4$).

(b) When the healing process is discontinuous we observe irregularities in the dynamic stress drop and particle motion. Consequently, fluctuations in the stress drop of a subsequent earthquake occur.

Finally, we compute static solutions by the equation:

$$\frac{\partial^2 u}{\partial x^2} + \sigma = 0 \tag{7}$$

where σ is the mean of the dynamic stress drop. The boundary conditions are $u(0) = u(L) = 0$, L is the fault lenght. We denote by M_o the seismic moment, using (7) we obtain

$$\sigma = \frac{12M_o}{L^3} \tag{8}$$

In table 1 we summarize the mean of the dynamic stress drop computed by the dynamic and static methods (Equation 8). In the static case we use the values of M_o and L compted in the dynamic case. It may be seen that the static approximation (8) leads to an underestimation when the healing process is discontinuous.

CONCLUSIONS

It is shown that:

(a) The healing process is usually supersonic.

(b) Variation in friction at the crack tip is connected with discontinuity in the healing process. The discontinuity causes irregularities in the motion of the fault and the final stress.

(c) Dynamic and static computations of stress drop are in reasonable agreement when the healing process is continuous. When it is discontinuous, the static approximation leads to an underestimation.

TABLE 1

CASE	L	M_o	σ: Dynamic Solution	σ: Static Solution
A $\varepsilon = 0$	1.28	0.2	1.1	1.2
A $\varepsilon = 0.46$	1.32	0.25	1.1	1.3
A $\varepsilon = 0.5$	.46	0.06	0.68	0.75
B $\eta = 0.$	1.	0.1	1.15	1.2
B $\eta = 0.1$	1.018	0.13	1.31	1.5
B $\eta^+ = 0.2$	1.036	0.093	1.38	1.0
B $\eta^+ = 0.4$	1.072	0.057	1.47	0.56

$^+$Healing process discontinuous

FIGURE CAPTIONS

Fig. 1 The one dimensional model is composed by particles which slide on a rough surface. The particles are connected to each other by coil springs and to the moving support by leaf springs. The continuous case is obtained when the distance a, between the particles tends to zero.

Fig. 2 Grid points and boundary conditions.

Fig. 3 Case A, variation of stress drop (difference between tectonic and dynamic friction) as a function of the spatial coordinate x.

Fig. 4 Case A, computed rupture velocity.

Fig. 5 Case A, computed dynamic stress drop (difference between initial and final stress).

Fig. 6 Case A, particle accelerations at the time of healing.

Fig. 7 Case A, the healing time as a function of the spatial coordinate x.

Fig. 8 Case A, final displacement as a funciton of the spatial coordinate x.

Fig. 9 Case A, particle accelerations as a function of time, the case of constant stress drop ($\varepsilon = 0$).

Fig. 10 Case A, particle velocities as a function of time, the case of constant stress drop ($\varepsilon = 0$).

Fig. 11 Case A, particle accelerations as a function of time, the case of variable stress drop ($\varepsilon = 0.46$).

Fig. 12 Case A, particle velocities as a function of time, the case of variable stress drop ($\varepsilon = 0.46$).

Fig. 13 Case B, variations in the absorbed energy rate per unit length

Fig. 14 Case B, rupture velocity as a funciton of the spatial coordinate x.

Fig. 15 Case B, healing time as a function of x.

FIGURE CAPTIONS (cont'd)

Fig. 16 Case B, dynamic stress drop as a function of x.

Fig. 17 Case B, particle accelerations at the time of healing.

Fig. 18 Case B, final displacement as a function of x.

Fig. 19 Case B, particle accelerations as a function of the time, the case of discontinuous healing ($\eta = 0.2$).

Fig. 20 Case B, particle velocities as a function of the time, the case of discontinuous healing ($\eta = 0.2$).

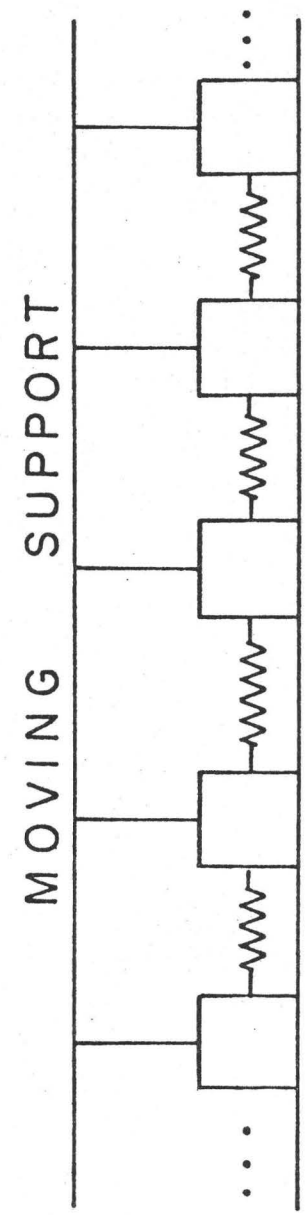


Fig. 1

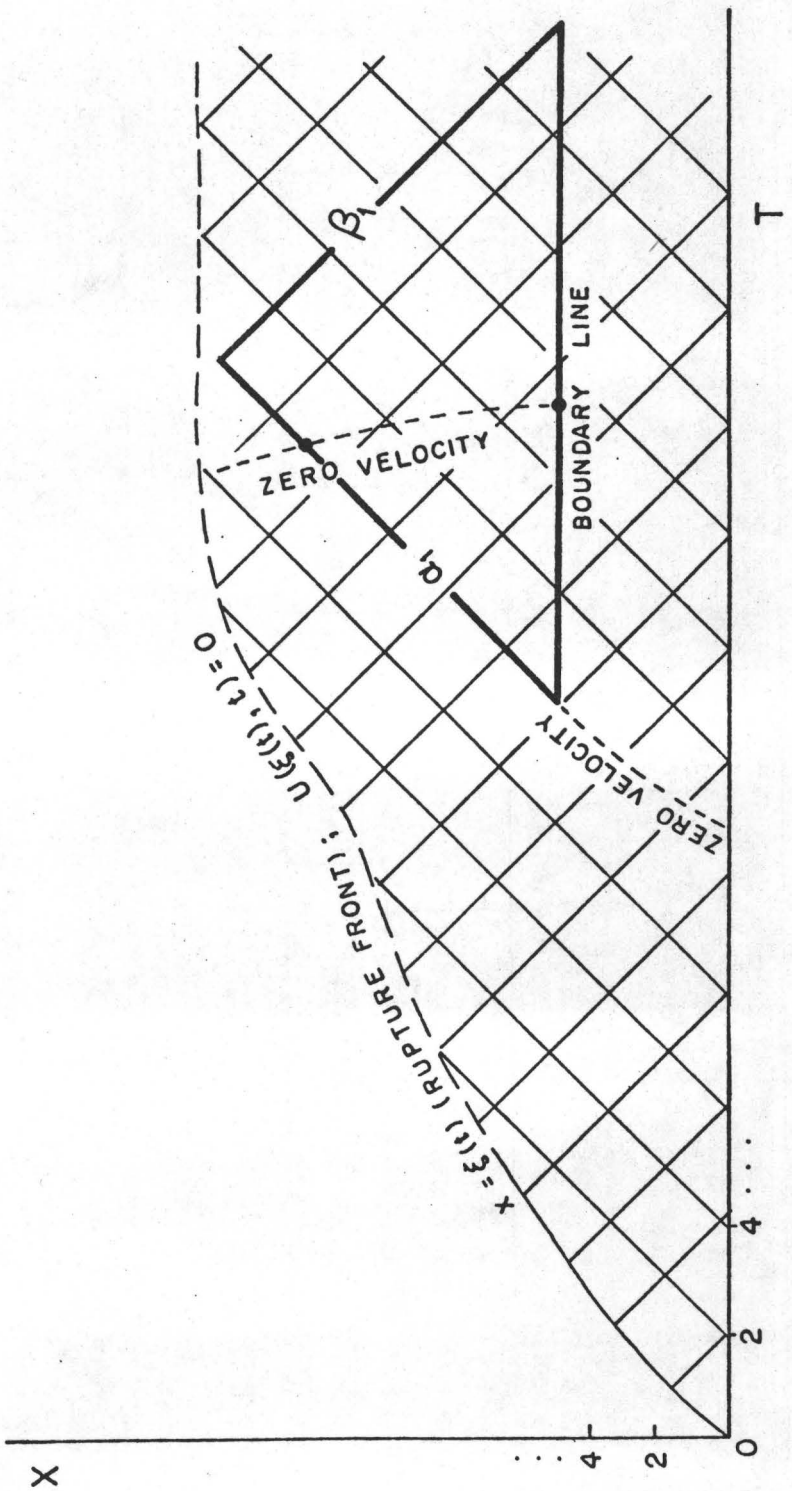

X
T
0
2
4
...
$x = \xi(t)$ (RUPTURE FRONT)
$u(\xi(t), t) = 0$
ZERO VELOCITY
BOUNDARY LINE
ZERO VELOCITY
β_1
β_2

Fig. 2

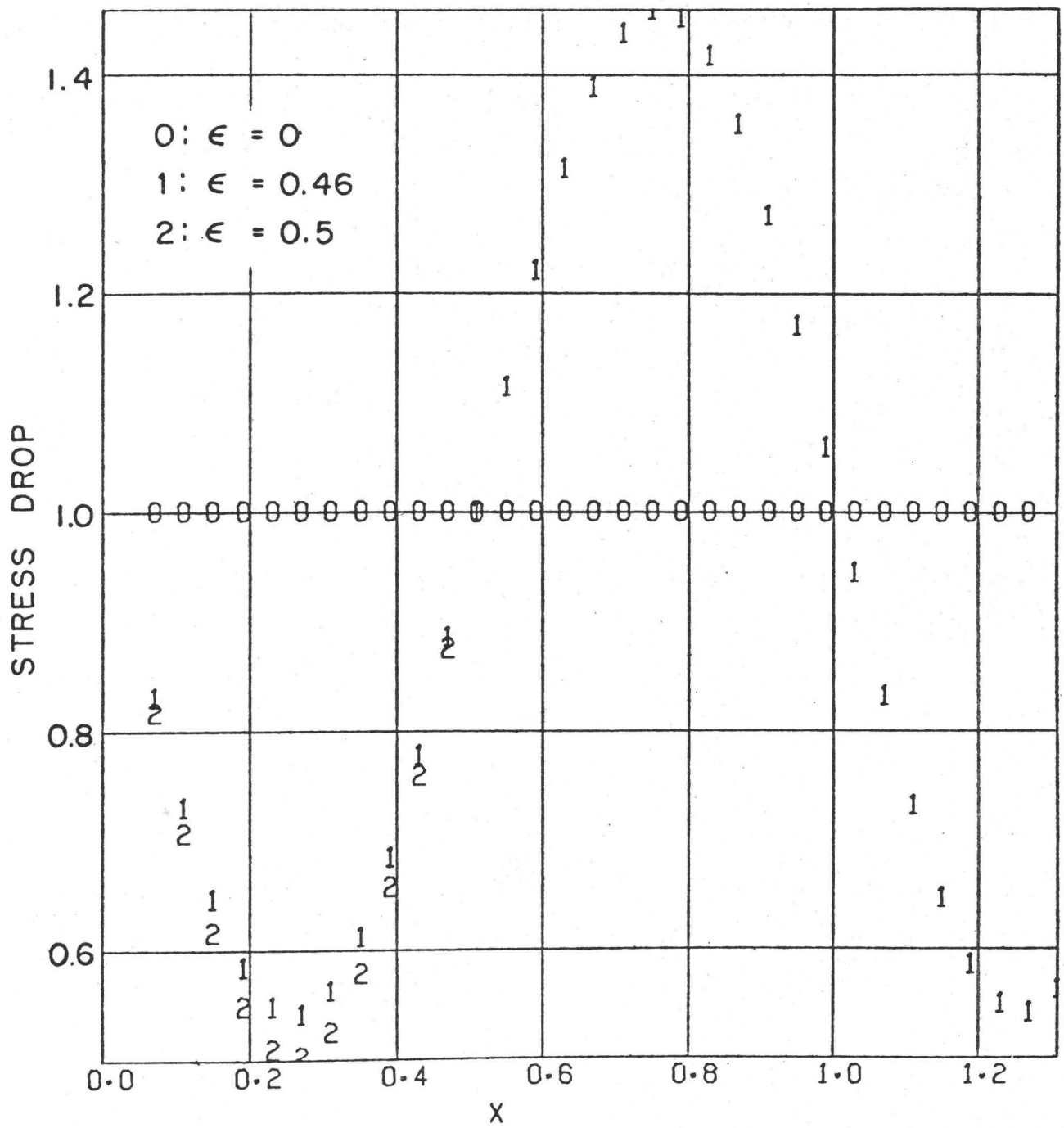


Fig. 3

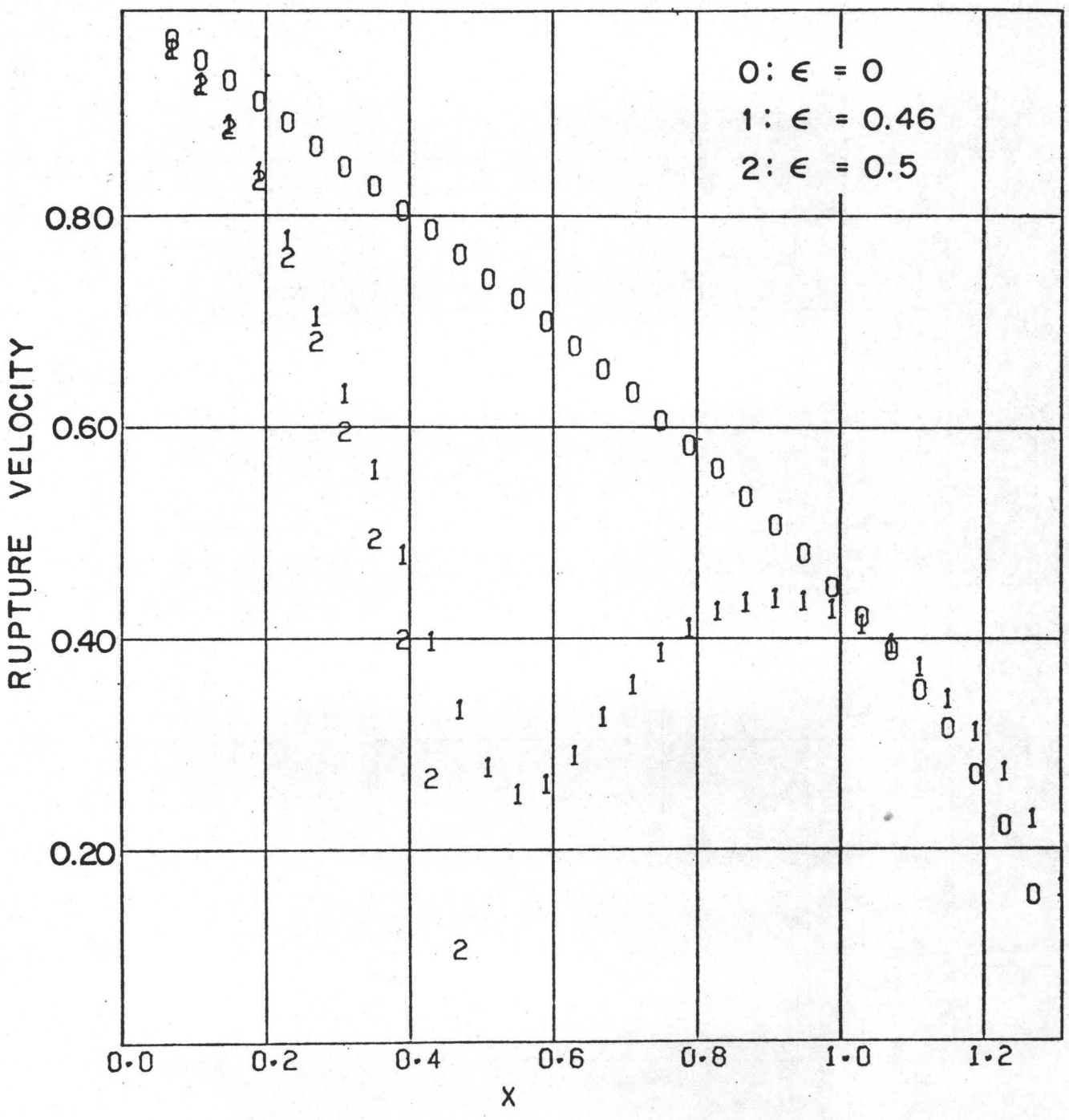

0: ϵ = 0
1: ϵ = 0.46
2: ϵ = 0.5
RUPTURE VELOCITY
0.80
0.60
0.40
0.20
0·0
0·2
0·4
0·6
0·8
1·0
1·2
X

Fig. 4

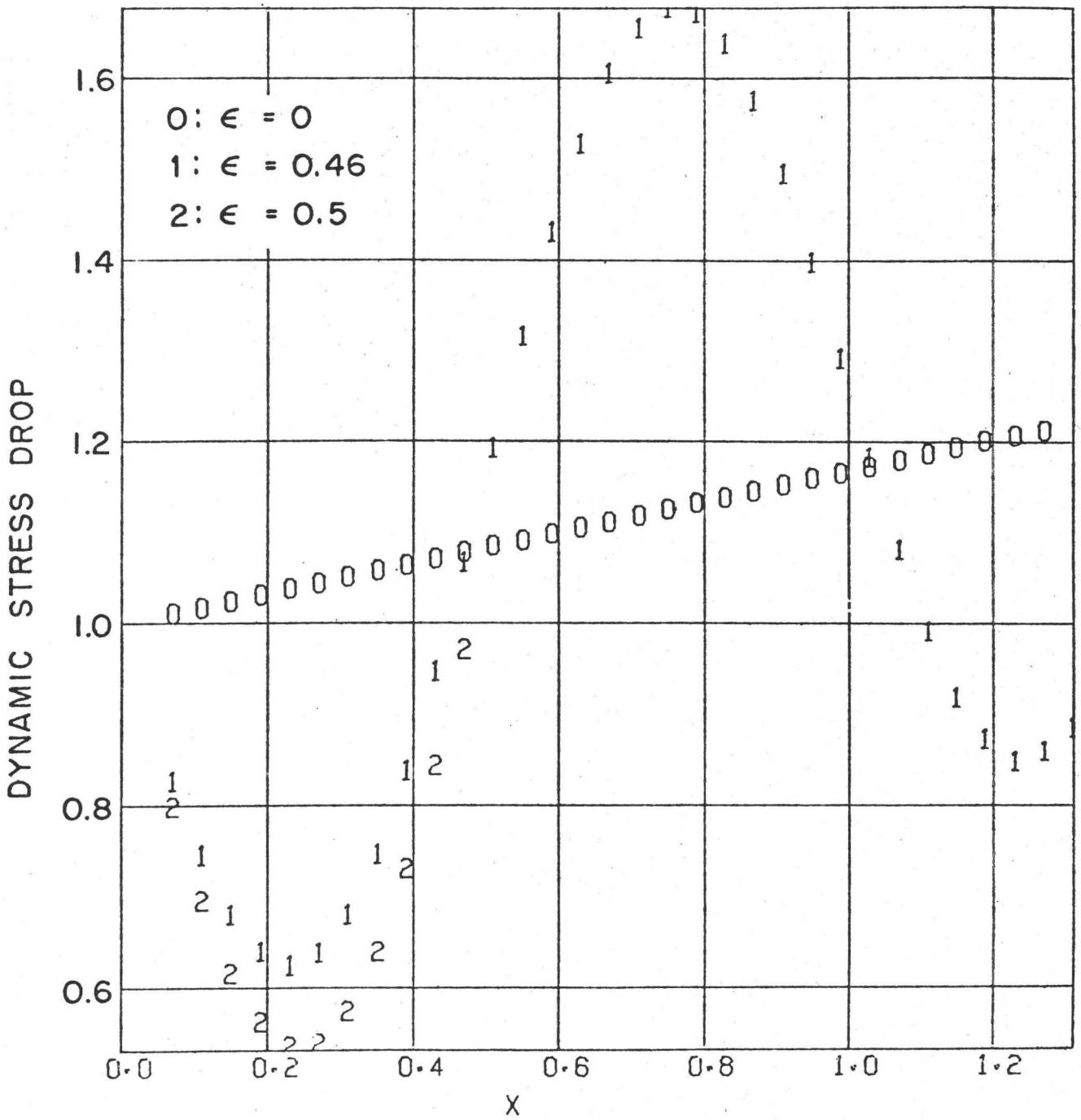

0: ϵ = 0
1: ϵ = 0.46
2: ϵ = 0.5
DYNAMIC STRESS DROP
1.6
1.4
1.2
1.0
0.8
0.6
0·0
0·2
0·4
0·6
0·8
1·0
1·2
X

Fig. 5

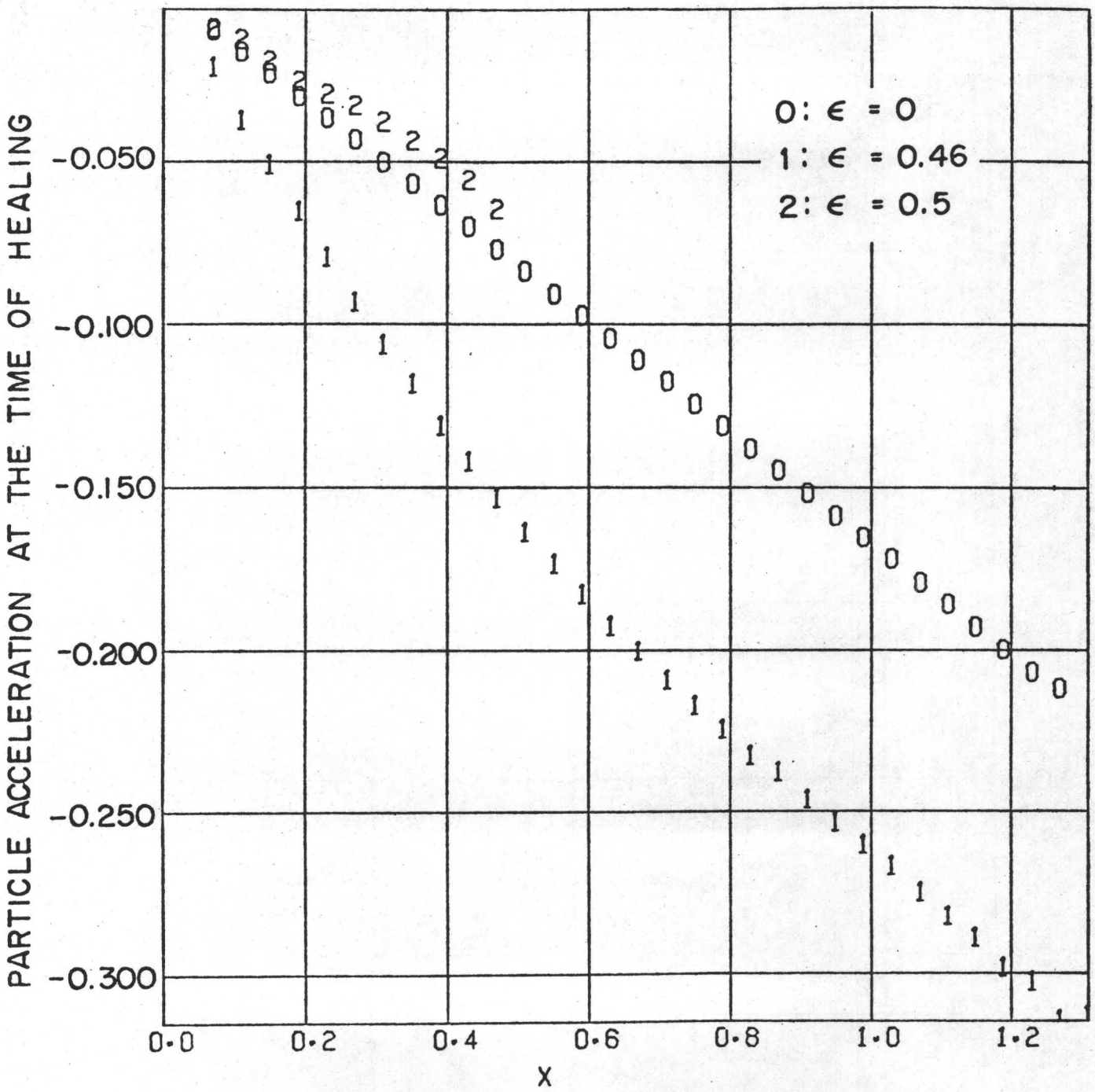

PARTICLE ACCELERATION AT THE TIME OF HEALING
-0.050
-0.100
-0.150
-0.200
-0.250
-0.300
0·0
0·2
0·4
0·6
0·8
1·0
1·2
X
0: ϵ = 0
1: ϵ = 0.46
2: ϵ = 0.5

Fig. 6

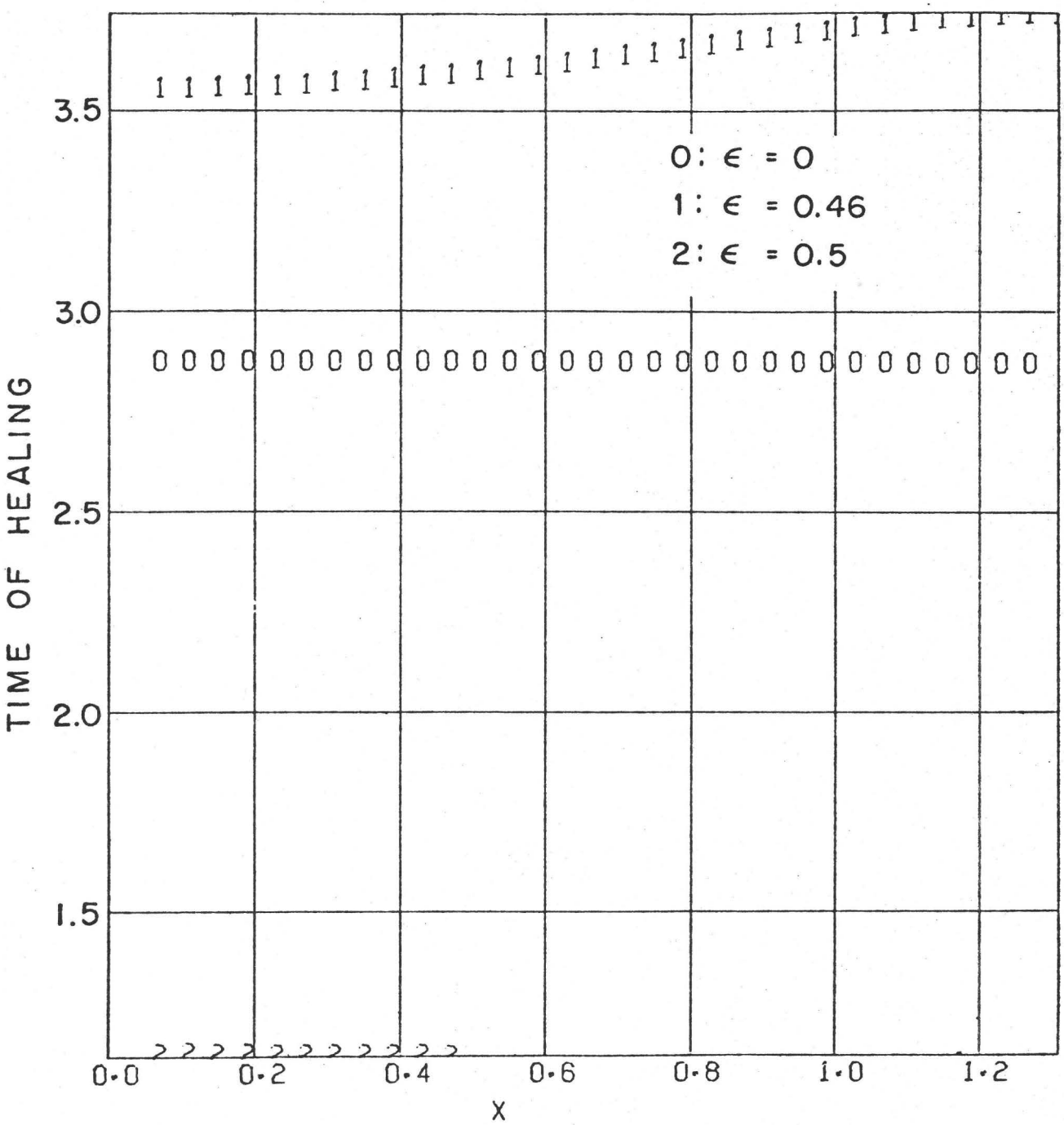

TIME OF HEALING
3.5
3.0
2.5
2.0
1.5
0·0
0·2
0·4
0·6
0·8
1·0
1·2
X
0: ϵ = 0
1: ϵ = 0.46
2: ϵ = 0.5

Fig. 7

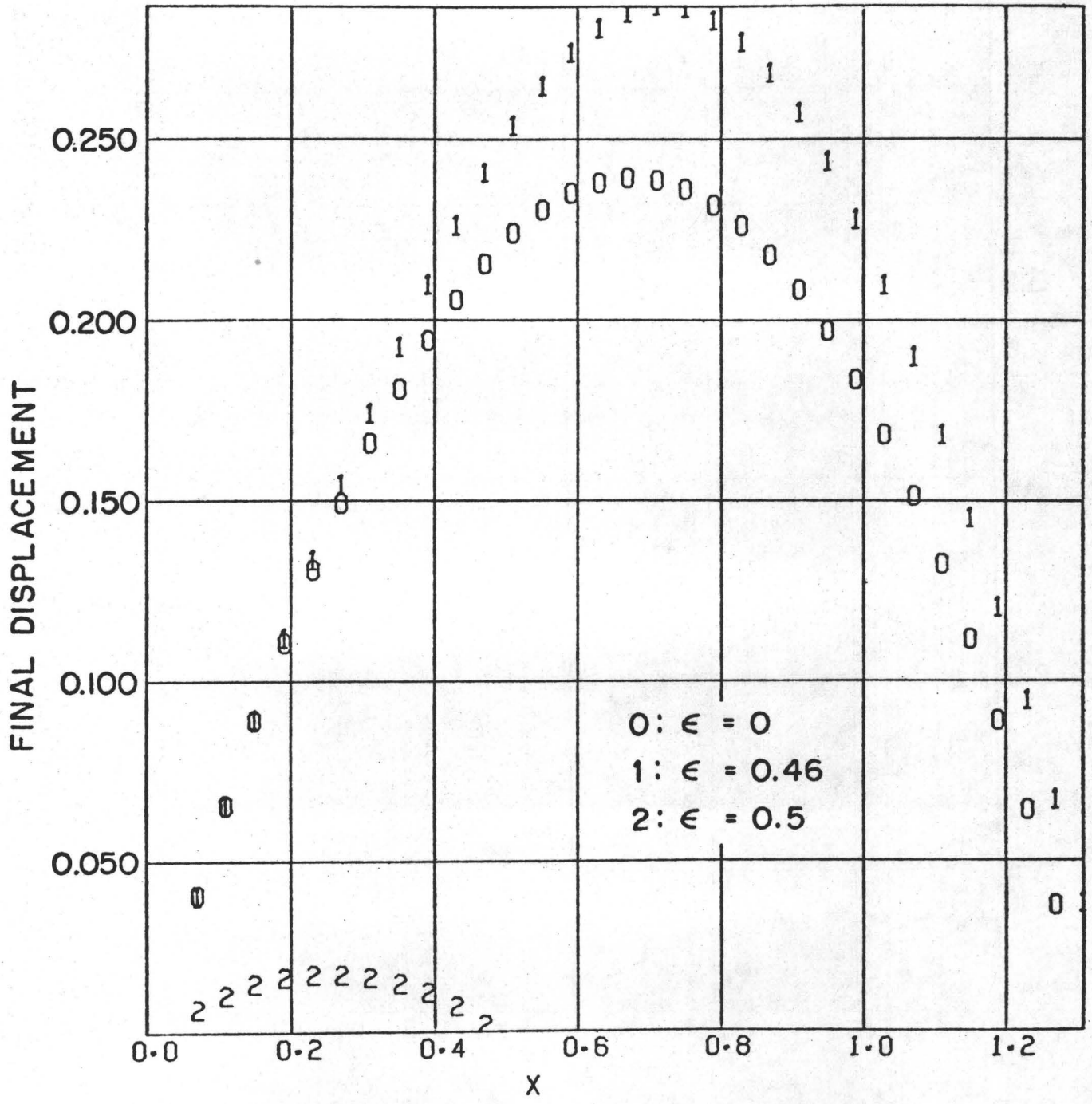

FINAL DISPLACEMENT
0.250
0.200
0.150
0.100
0.050
0·0
0·2
0·4
0·6
0·8
1·0
1·2
X
0: ∈ = 0
1: ∈ = 0.46
2: ∈ = 0.5

Fig. 8

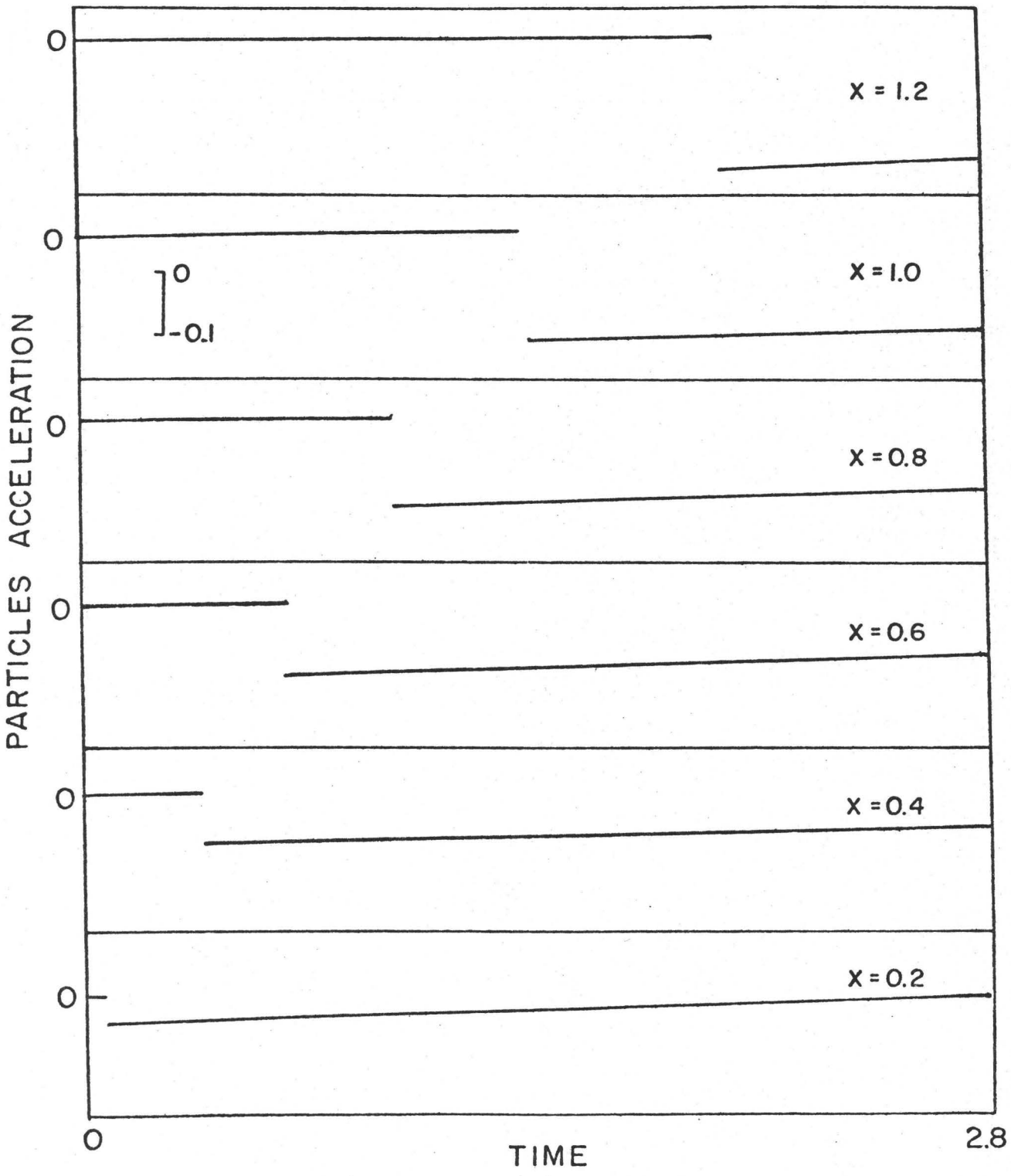

0
X = 1.2
0
0
-0.1
X = 1.0
0
X = 0.8
0
X = 0.6
0
X = 0.4
0
X = 0.2
PARTICLES ACCELERATION
0
TIME
2.8

Fig. 9

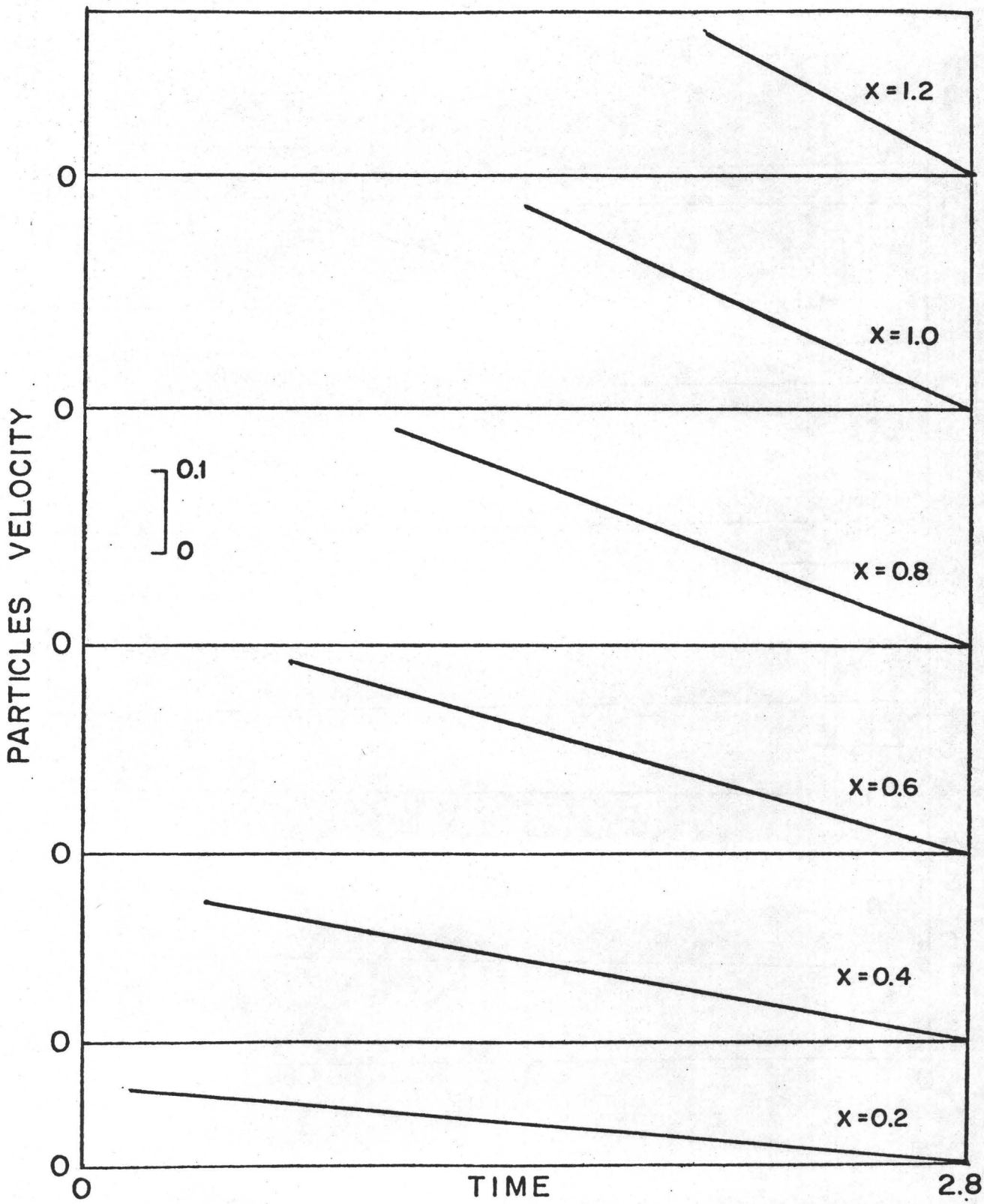

X = 1.2
X = 1.0
X = 0.8
X = 0.6
X = 0.4
X = 0.2
0.1
0
0
0
0
0
0
0
PARTICLES VELOCITY
TIME
2.8

Fig. 10

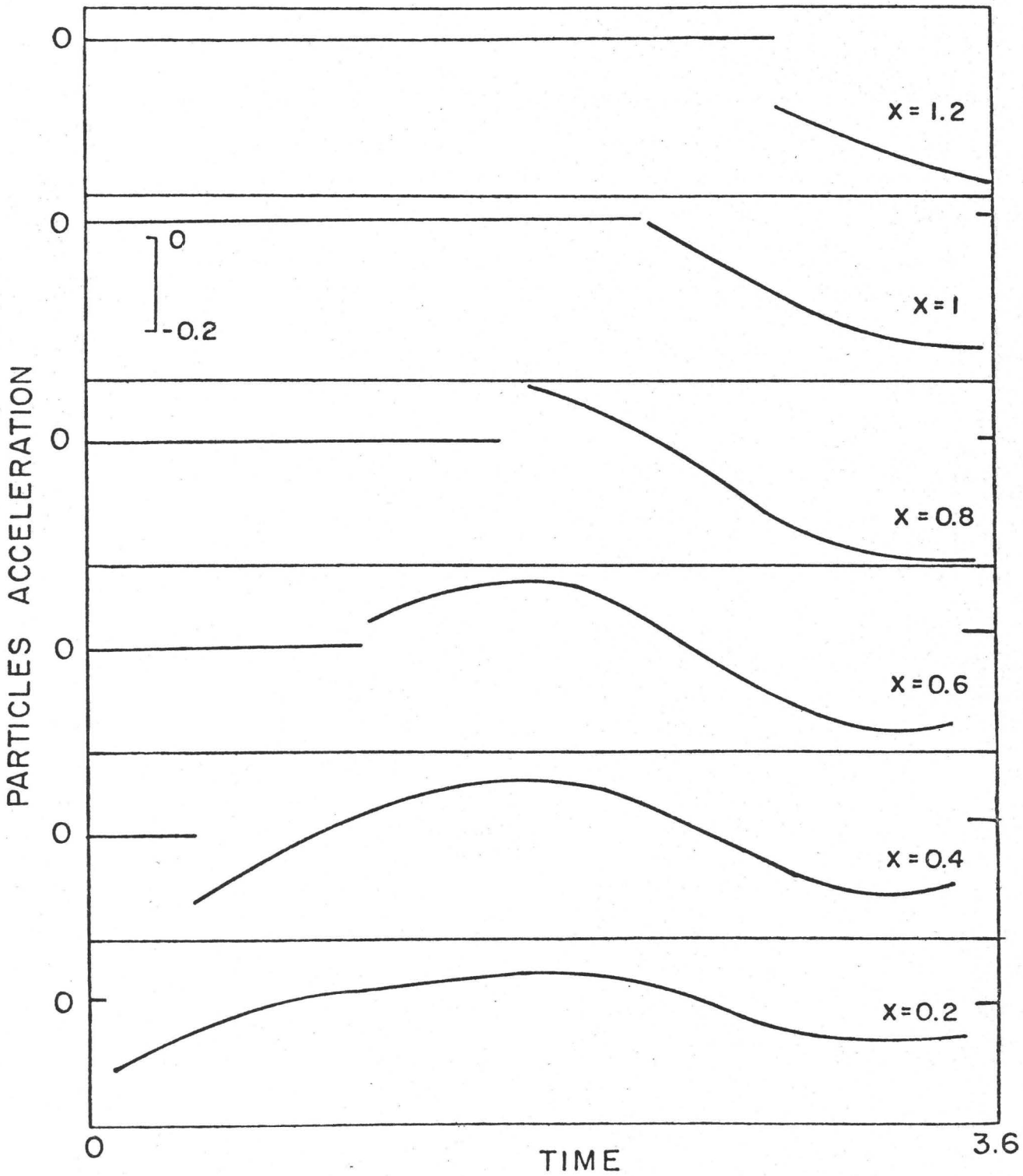

PARTICLES ACCELERATION
0
0
0
-0.2
0
0
0
0
X = 1.2
X = 1
X = 0.8
X = 0.6
X = 0.4
X = 0.2
0
3.6
TIME

Fig. 11

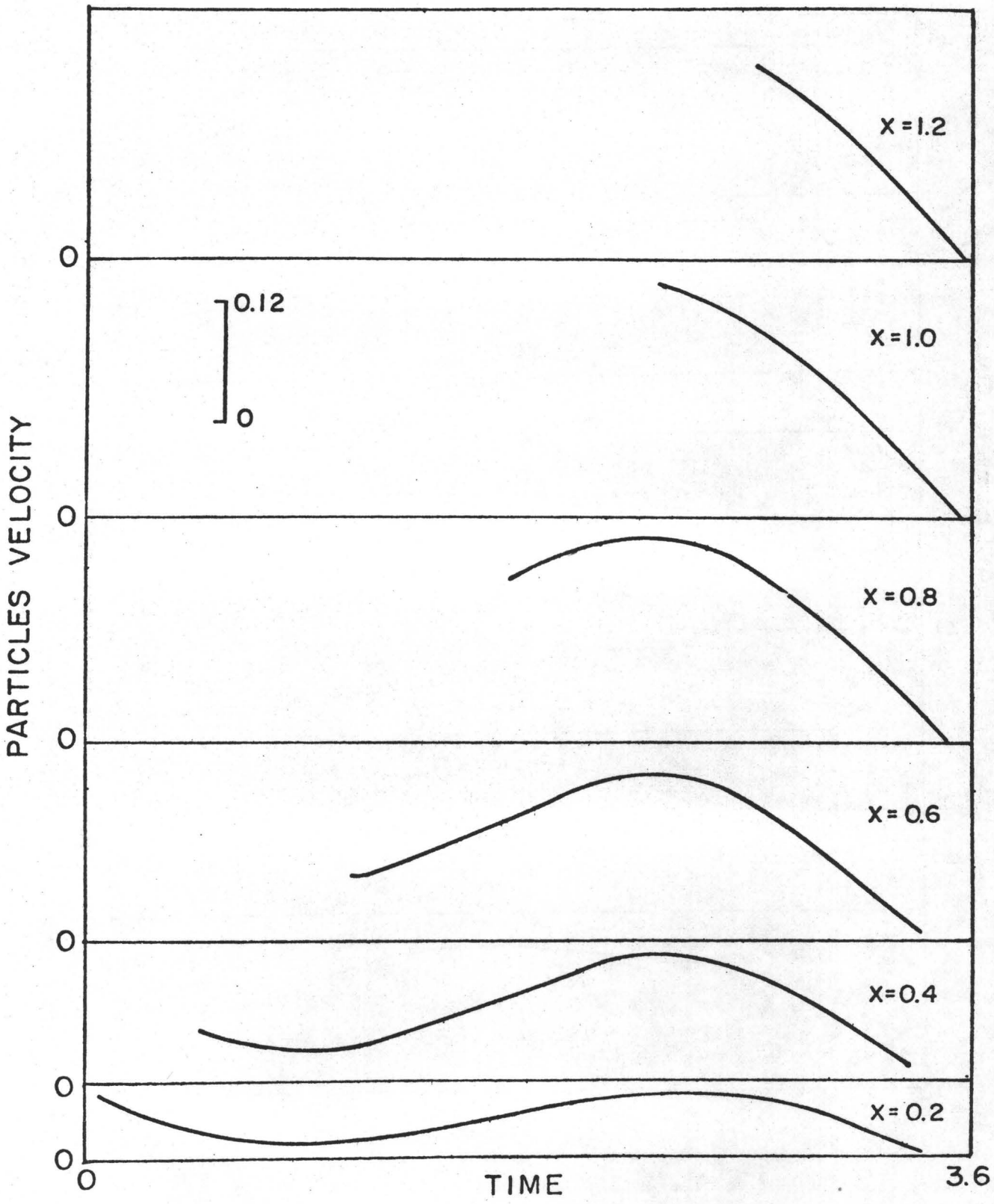

X = 1.2
0.12
0
X = 1.0
X = 0.8
X = 0.6
X = 0.4
X = 0.2
0
0
0
0
0
0
PARTICLES VELOCITY
TIME
3.6

Fig. 12

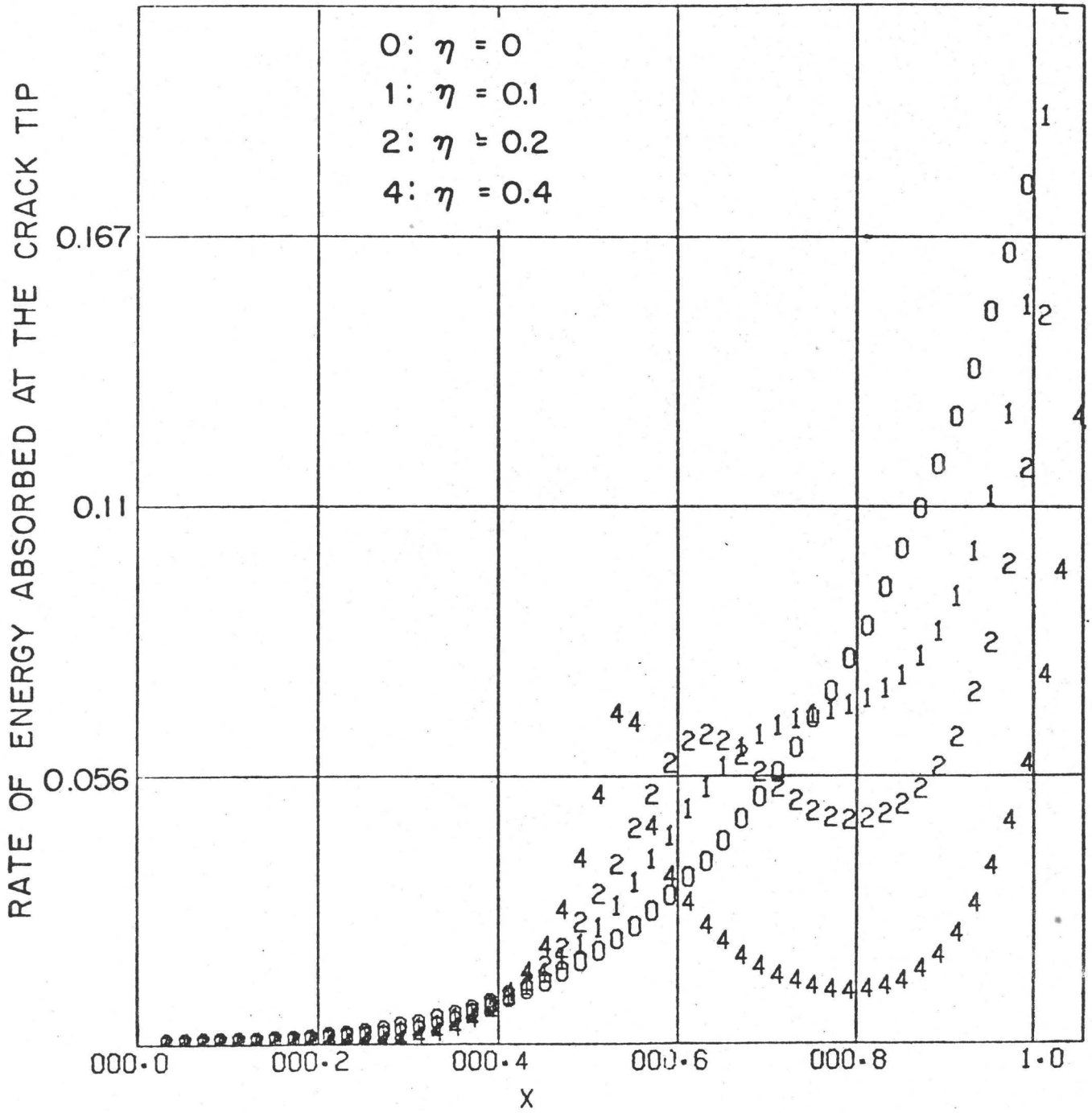

0: η = 0
1: η = 0.1
2: η = 0.2
4: η = 0.4
RATE OF ENERGY ABSORBED AT THE CRACK TIP
0.167
0.11
0.056
000·0
000·2
000·4
000·6
000·8
1·0
X

Fig. 13

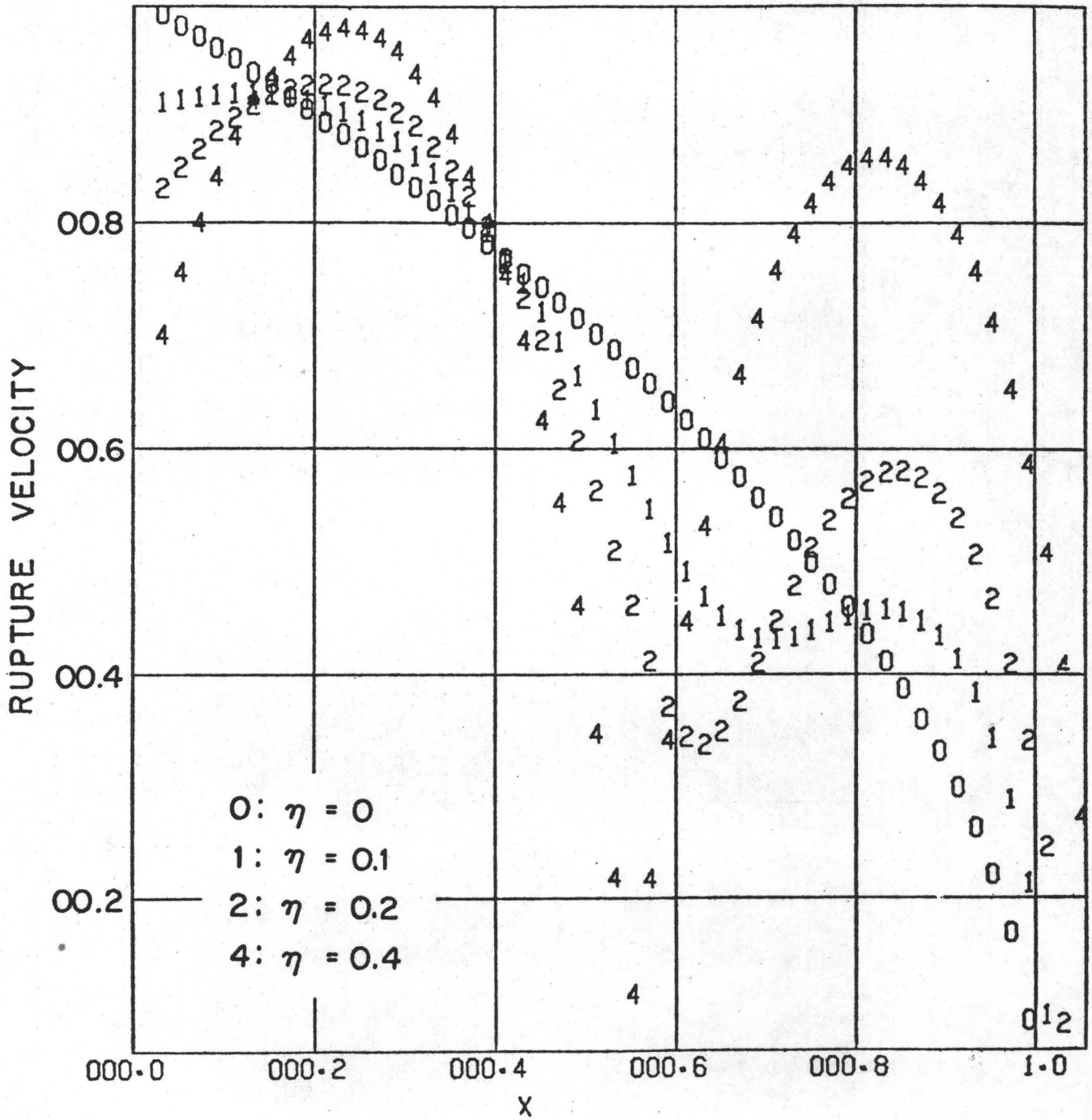

RUPTURE VELOCITY
00.8
00.6
00.4
00.2
000·0
000·2
000·4
000·6
000·8
1·0
X
0: η = 0
1: η = 0.1
2: η = 0.2
4: η = 0.4

Fig. 14

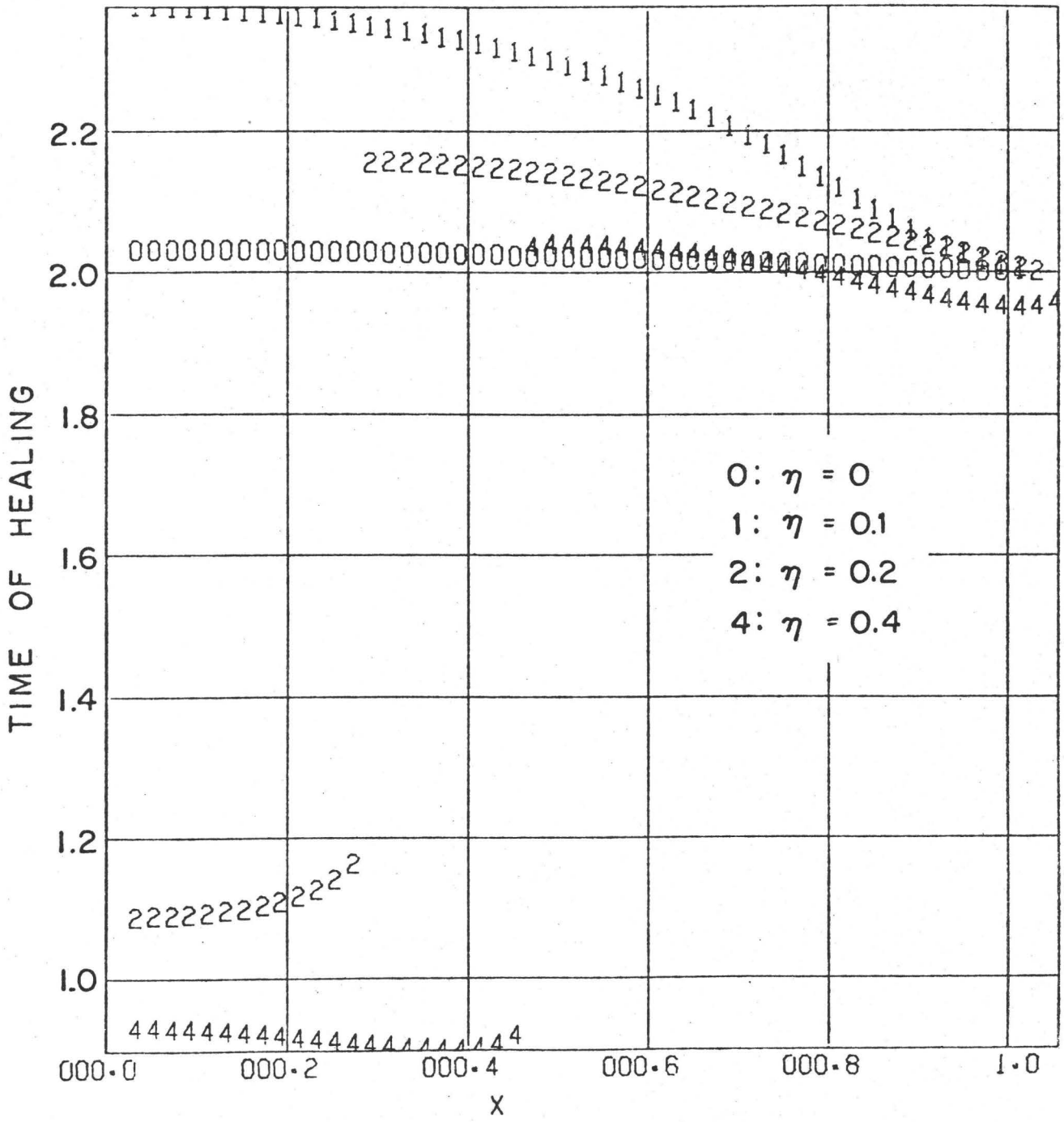

TIME OF HEALING
2.2
2.0
1.8
1.6
1.4
1.2
1.0
000·0
000·2
000·4
000·6
000·8
1·0
X
0: η = 0
1: η = 0.1
2: η = 0.2
4: η = 0.4

Fig. 15

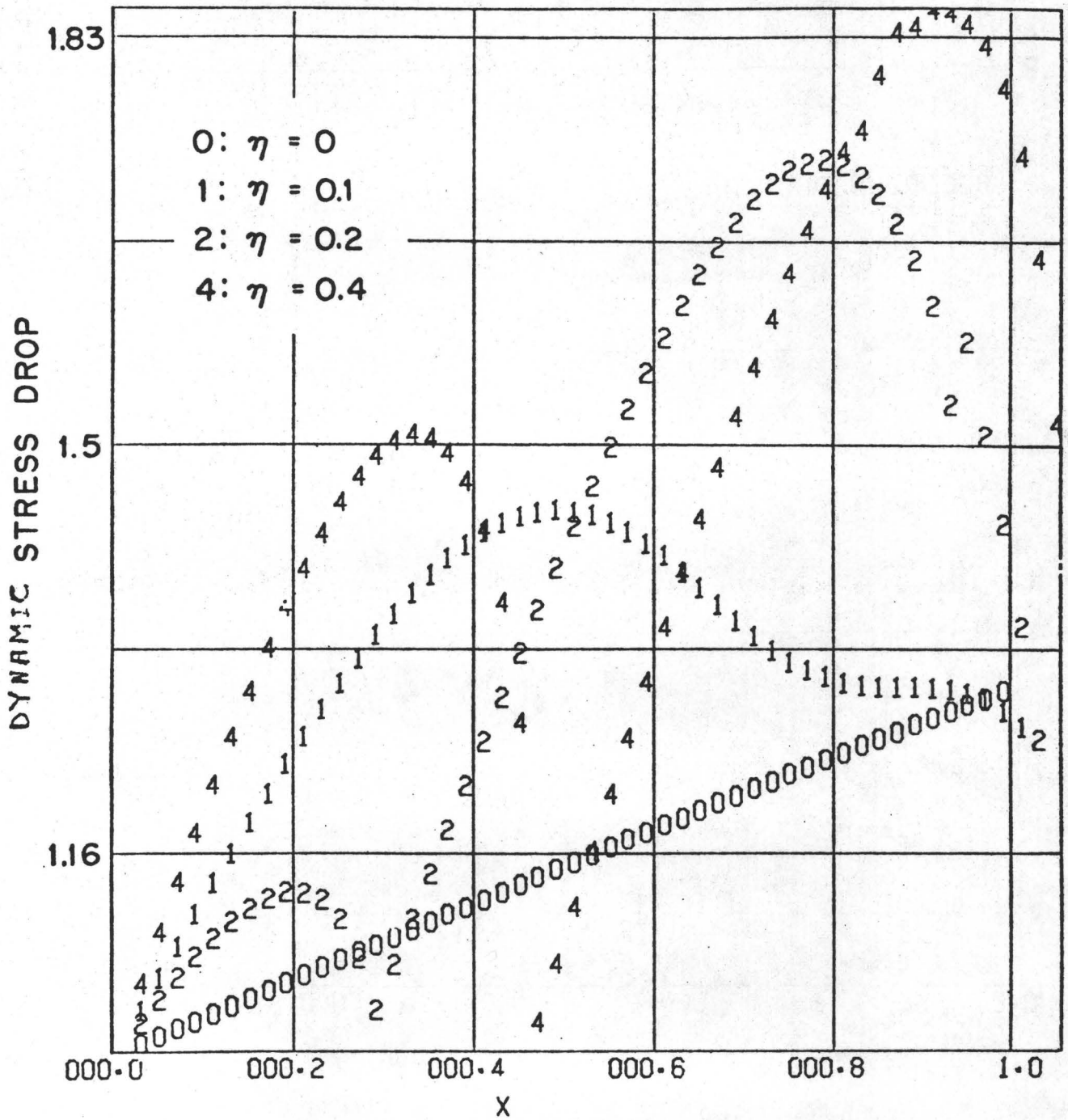

1.83
1.5
1.16
DYNAMIC STRESS DROP
0: η = 0
1: η = 0.1
2: η = 0.2
4: η = 0.4
000·0
000·2
000·4
000·6
000·8
1·0
X

Fig. 16

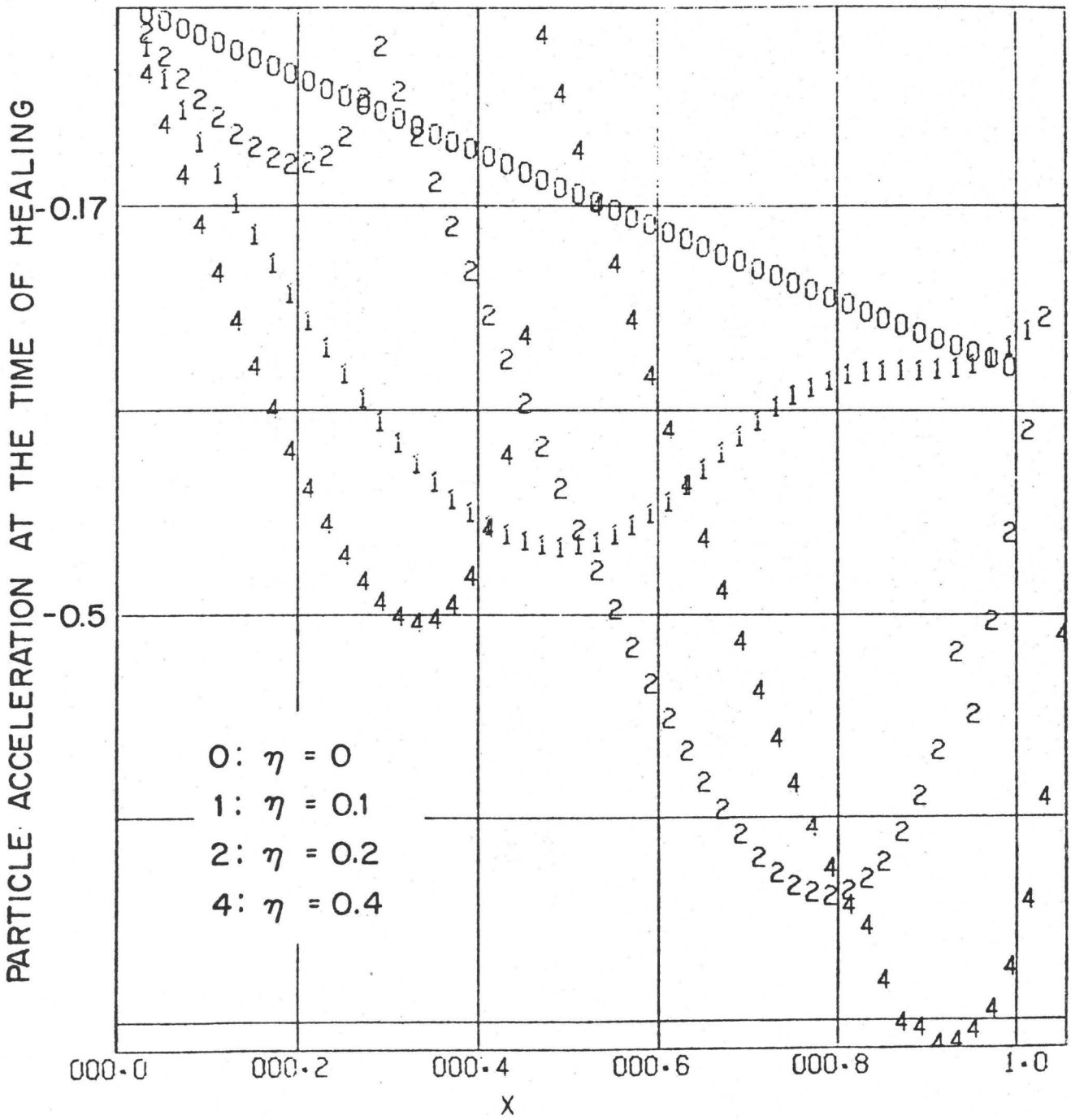

PARTICLE ACCELERATION AT THE TIME OF HEALING
-0.17
-0.5
0: η = 0
1: η = 0.1
2: η = 0.2
4: η = 0.4
000·0
000·2
000·4
000·6
000·8
1·0
X

Fig. 17

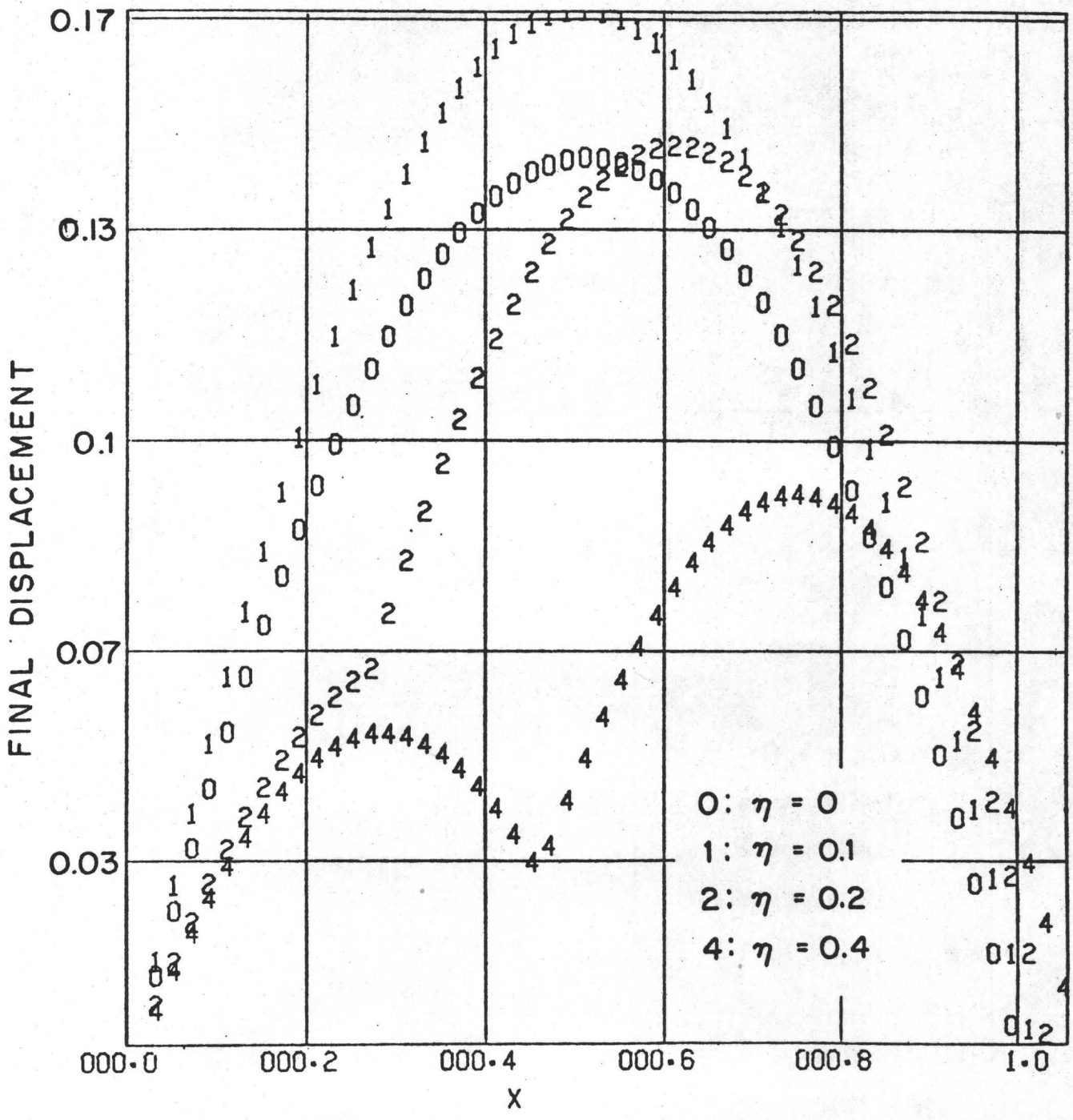

0.17
0.13
0.1
0.07
0.03
FINAL DISPLACEMENT
000·0
000·2
000·4
000·6
000·8
1·0
X
0: η = 0
1: η = 0.1
2: η = 0.2
4: η = 0.4

Fig. 18

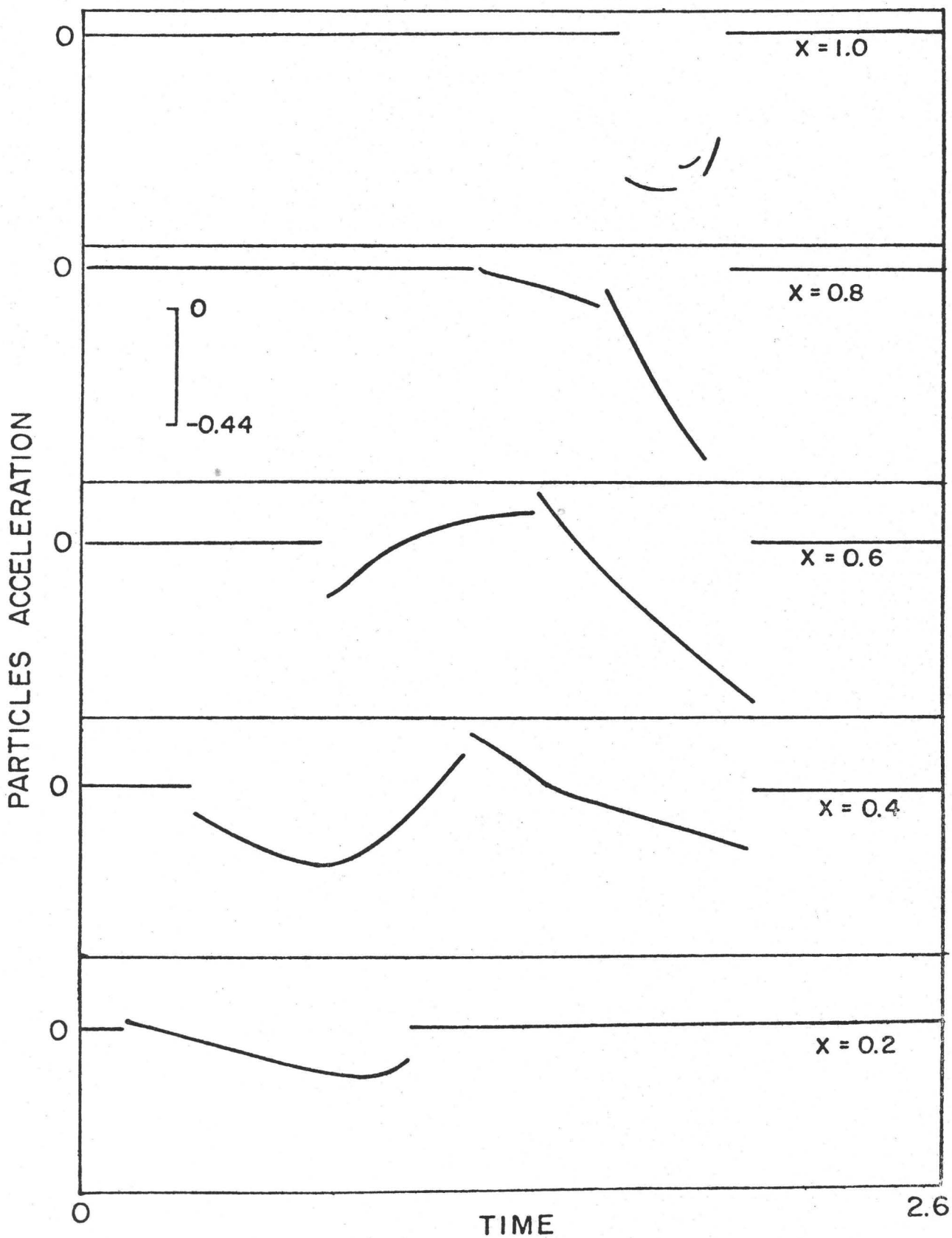

0
X = 1.0
0
X = 0.8
0
-0.44
0
X = 0.6
0
X = 0.4
0
X = 0.2
PARTICLES ACCELERATION
0
TIME
2.6

Fig. 19

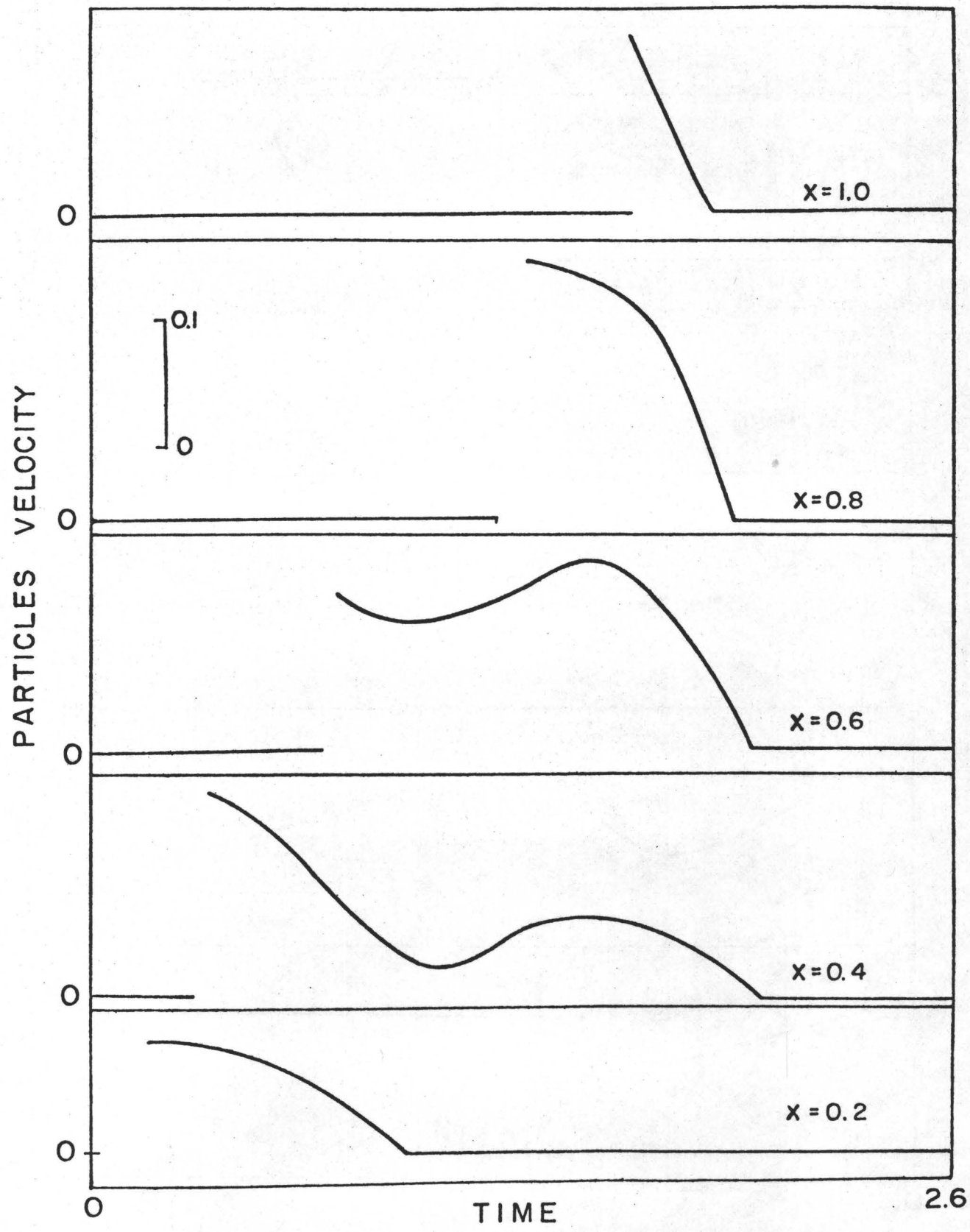


Fig. 20

USE OF SEISMIC RADIATION TO INFER SOURCE PARAMETERS

Hiroo Kanamori

Seismological Laboratory
California Institute of Technology
Pasadena, California 91125

Contribution No. 3026 of the Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125.

1. BRIEF REVIEW OF SEISMIC METHODS FOR DETERMINATION OF SOURCE PARAMETERS

In this section we list earthquake source parameters relevant to the subject and briefly summarize commonly used seismological methods for the determination of these parameters.

(1) Fault Geometry

The geometry of an earthquake fault can usually be defined by three parameters: the fault strike, the angle of the fault plane and the slip angle of the fault motion. Other representations and useful relations between them are given by Jarosch and Aboodi (1970). These parameters can be determined by 1) P-wave first-motion data (Stauder, 1962, Honda, 1962), 2) S-wave polarization angles, (Stauder, 1962; Honda, 1962; Hirasawa, 1966), 3) wave forms of body waves (Langston and Helmberger, 1975; Langston, 1976; Langston and Butler, 1976), 4) radiation pattern of surface-waves (Brune, 1961; Ben-Menahem and Harkrider, 1964; Aki, 1966; Kanamori, 1970; Ben-Menahem et al., 1970), 5) the excitation of normal modes (Saito, 1967; Abe, 1970; Ben-Menahem, et al., 1971; Dziewonski and Gilbert, 1975; Gilbert and Dziewonski, 1975), 6) geodetic data (e.g., Chinnery, 1964, 1969; Savage and Hastie, 1966; Ando, 1971), and 7) field observations.

(2) Fault Dimension

The dimension of an earthquake fault can be determined from 1) the size of the aftershock area, 2) geodetic data, 3) tsunami source area, 4) directivity and asymmetry of the radiation pattern of long-period surface waves, 5) pulse width of body waves, and 5) seismic corner frequency.

1) Aftershock Area

Although there is no standard definition of aftershocks and the aftershock area, the aftershock area defined by the somewhat subjective

judgment of the investigator often provides a very good estimate of the fault area of very large earthquakes (fault dimension $\geq$ 100 km), particularly for great shallow thrust earthquakes along subduction zones (Benioff et al., 1961; Press et al., 1961; Mogi, 1968a). For small earthquakes, errors in the epicenteral location of aftershocks and the temporal expansion of the aftershock area often cause a substantial error in the estimate of the size of the fault plane.

2) Geodetic Data

When geodetic data (leveling and triangulation) are available over the entire area of faulting, the size of the fault plane can be determined very well (Savage and Hastie, 1966; Plafker, 1972; Ando, 1971; Kasahara, 1957; Chinnery, 1964; Kanamori, 1973; Jungles and Frasier, 1973). The spatial decay rate of the displacement field can be used to infer the vertical extent of the faulting (Knopoff, 1958), although the resolution is often limited by the quality and quantity of the data.

3) Tsunami Source Area

When tide-gage data are available near the epicenter of a large tsunamigenic earthquake, the source area of tsunami can be estimated by using the inverse refraction diagram (Miyabe, 1934; Hatori, 1966; Abe, 1973). Usually, a good correlation between the size of tsunami source area and the aftershock area (Hatori, 1965; Abe, 1973) is found.

4) Directivity

From the directivity of very long-period (200 to 300 sec) surface waves, the rupture length can be estimated (Ben-Menahem, 1961). The asymmetry of the radiation pattern can also be used (Kanamori, 1970).

For very large earthquakes such as the 1960 Chilean earthquake, the 1952 Kamchatka earthquake, the 1964 Alaskan earthquake, these methods gave a reliable estimate of the fault length, perhaps accurate to $\pm 15\%$ (Benioff et al., 1961; Press et al., 1961; Ben-Menahem and Toksoz, 1963). However, for events whose dimension is smaller than 100 km, the resolution of these methods becomes very poor.

5) Pulse Width of Body Waves

The pulse width of body waves was used to infer the fault dimension of deep focus earthquakes (Fukao, 1970). In order to determine the source time function from observed body waves from shallow focus earthquakes, the effect of structure near the source, particularly the free surface effects, and propagation effects must be removed from the observed records. Techniques have been developed to correct for these effects (Fukao, 1971; Helmberger, 1974; Langston and Helmberger, 1975; Helmberger and Malone, 1975), and the source time function can be recovered very accurately for relatively simple events (Burdick and Mellman, 1976). For complex events, the analysis of body waves is more difficult but several attempts have been made (Fukao, 1972; Chung and Kanamori, 1976; Kanamori and Stewart, 1978; Chung and Kanamori, 1978) to recover the complex time history of the rupture process. The interpretation of the pulse width in terms of the source dimension involves assumptions on the geometry of the fault, mode of rupture and rise time of the local slip function, and is often nonunique.

6) Corner Frequency

The corner frequency of the spectrum of body-waves is a frequency domain representation of the pulse width. Brune (1970) proposed a

relation between the source dimension and the corner frequency. Although Brune's equation provides a useful average relation, the estimate for an individual event depends on the geometry of the fault, the mode of rupture and the rise time of the local slip function. When the wave form becomes complex due to source complexity and propagation effects, including reflections and refractions, interpretation of the corner frequency in terms of the source dimension becomes very difficult.

(3) Rupture Mode and Rupture Velocity

Whether the fault rupture is unilateral, bilateral or two-dimensional is usually determined from the spatial relation of the main shock to the aftershock area. For very large earthquakes, the rupture velocity can be determined from the directivity function (Ben-Menahem, 1961; Benioff et al., 1961). For multiple shocks, the apparent rupture velocity is given by the ratio of the spatial separation to the temporal separation of the individual events. The rupture velocity is sometimes determined from the wave forms of near field records (Aki, 1968; Kanamori, 1972; Abe, 1974a).

(4) Dislocation

The dislocation on the fault plane is in general a function of position and time. From geodetic data, the static value of the dislocation can be determined as a function of position on the fault. However, the details of the spatial distribution are usually very difficult to resolve (Chinnery, 1964; Kasahara, 1957; Savage and Hastie, 1966; Kanamori, 1973).

If the fault area is known, the dislocation can be estimated from the amplitude of seismic body waves, surface waves and free oscillations.

However, it is very difficult to resolve the details of the spatial distribution of the dislocation; usually, only a spatial average can be determined.

(5) Particle Velocity

The particle velocity at a point on the fault plane is directly related to the effective tectonic stress (Brune, 1970). In principle, the particle velocity can be determined from the frequency spectrum or the rise time of near field seismograms. However, it is difficult to remove the effect of rupture propagation and near source geometry from the observed seismogram. Only a few determinations of the particle velocity have been made (Kanamori, 1972; Abe, 1974a,b, 1975a).

(6) Complexity

The complexity of faulting process can be determined by the analysis of distinct arrivals on seismograms (Imamura, 1937; Miyamura et al., 1964; Wyss and Brune, 1967; Trifunac and Brune, 1970). More recently synthetic seismograms have been used to determine more details of the multiple shock sequence (Fukao, 1972; Chung and Kanamori, 1976; Kanamori and Stewart, 1978; Chung and Kanamori, 1978). Detailed study of complexity of faulting is important in understanding the stress state in the fault zone and also in predicting ground motions resulting from an earthquake.

2. SUMMARY OF RESULTS

(1) Geodetic Data

Geodetic data (both leveling and triangulation) are summarized by Rikitake (1974) and are shown in Figure 1. Figure 1 shows the

strain change in the immediate vicinity of the epicenters of various earthquakes as a function of magnitude. The strain seems to be almost constant from 2×10^{-5} to 2×10^{-4} regardless of the magnitude of the earthquake. Since the rigidity of crustal rocks averages 3.5×10^{11} dyne/cm^2, this strain change corresponds to a stress drop of 7 to 70 bars (Chinnery, 1964).

(2) Great and Large Earthquakes

The results for large and great earthquakes are summarized by Kanamori and Anderson (1975) and Geller (1976) (Table 1).

Figure 2 shows the relation between log S (S: fault area) and log M_o ($M_o = \mu \bar{D} S$: seismic moment) for great and large earthquakes. The remarkable linearity between log M_o and log S can be interpreted in terms of a constant average stress drop (30 to 60 bars) in earthquakes (Aki, 1972; Kanamori and Anderson, 1975; Abe, 1975b; Geller, 1976).

The effective stress $\sigma_{oe} = \sigma_o - \sigma_f$ (σ_o = initial tectonic stress on the fault plane; σ_f = dynamic friction during faulting) is the stress which drives the fault motin (Brune, 1970). The effective stress can be obtained from the frequency spectrum or the rise-time of near-field seismograms. Table 2 summarizes the results. Although these results are subject to large uncertainty, it is important that σ_{oe} is about the same order of magnitude as the stress drop.

(3) Small Earthquakes

Figure 3 shows the relation between log r (r: source dimension) and log M_o (Hanks, 1977). Although the trend is similar to that for large earthquakes, the stress drop varies over a larger range (0.5 to 100 bars) than for large earthquakes. Whether this large variation

is due to real variation of the stress drop or experimental uncertainty is not clear. For very small earthquakes, the source dimension r is estimated mainly from the corner frequency and the uncertainty of this measurement is very difficult to estimate.

For several earthquakes, the duration of the source time function has been determined from time-domain analyses (Figure 4, Helmberger and Johnson, 1977). These results again indicate a stress drop of 10 to 100 bars. In some cases, a very large stress drop (1 kbar or larger) has been reported (e.g., Brune et al., 1976). Although the absolute values of the stress drop are subject to large uncertainty due to the lack of information about the rupture mode and the source dimension, these results indicate a larger range of stress drop. for small earthquakes than for large earthquakes.

3. MULTIPLE SHOCKS

Many seismograms indicate that earthquake fault motion is extremely complex. This complexity exists at all scales. Figure 5 shows an example of the strong-motion seismogram of the 1971 San Fernando earthquake recorded at Pacoima Dam. The "displacement" trace shows the ground displacement recorded by an instrument whose response is given by curve 3. The displacement is relatively smooth and various theoretical methods can be used to explain this trace. The stress drop has been estimated by various analyses (Mikumo, 1973; Trifunac, 1974; Hanks, 1974). The "acceleration" trace shows the ground displacement recorded by an instrument with a response shown by curve 1. The displacement in this high frequency range is extremely complex, and simple theoretical

models fail to explain this irregularity.

Figure 6 shows an irregularity at a larger scale. It is obvious that very complex source models are necessary to explain these complex events. However, estimates of the average stress drop and effective stress have been obtained using simple dislocation models or simple crack models. These models can explain the long-period component of the seismograms but fail to explain the short-period component. How meaningful is the estimate of the stress drop and other source parameters obtained by using these incomplete models? Madariaga (1977) showed that the estimate of the average stress drop depends upon the distribution of the stress drop on the fault plane. However, it is probably unlikely that the estimate of the average stress drop obtained for earthquakes using a simple model is in error by a factor of five or so.

A more detailed analysis was made for the 1976 Guatemala earthquake (Kanamori and Stewart, 1978). Figure 7 shows wave forms of P waves at seven stations which exhibit remarkable complexity. These wave forms were matched by synthetic seismograms computed for a sequence of point sources (Figure 7). The resulting source time sequence is shown by Figure 8. This result sugests that the earthquake can be represented by a sequence of approximately ten distinct events, the seismic moment of which varies by a factor of about four. The rupture can be represented by a stop-and-go sequence with an average rupture velocity of 2 km/sec. The spatial separation of the individual events is 14 to 40 km suggesting that either stress, frictional characteristics or sliding characteristics on the fault plane vary with comparable

spatial scale along the fault plane. Although the average stress drop is about 30 bars, the local stress drop for the individual events may be significantly higher than this value, perhaps by a factor of two or three. Other multiple shocks which were studied earlier include the 1923 Kanto earthquake (Imamura, 1937), the 1964 Alaskan earthquake, (Wyss and Brune, 1967), and the 1940 Imperial earthquake (Trifunac and Brune, 1970). These multiple shocks will provide important clues to the understanding of the mechanics of faulting.

4. ASPERITY

The multiple shocks can be interpreted in terms of asperities on the fault plane. Here the asperities can be geometrical asperities, heterogeneities of the frictional strengh or a combination. A fault plane can probably be represented by a random distribution of stress concentrations of various scale lengths. This asperity model can be used to interpret seismicity patterns before large earthquakes.

Several investigators (e.g., Mogi, 1968b; Kelleher and Savino, et al., 1975), found that foreshocks tend to cluster near the epicenter of the main shock (Figure 9). For the 1971 San Fernando earthquake, Ishida and Kanamori (1977) found a clustering of activity for a two-year period before the main shock (Figure 10). As shown in Figure 11, these events show nearly identical wave forms at Pasadena indicating that they originated nearly at the same hypocenter. As shown by Figure 10 the distribution of small earthquakes in the epicentral area prior to 1965 was relatively random. During the period fom 1965 to 1968, the seismic activity was very low in the epicentral area.

These observations may be interpreted in terms of the asperity model. The distribution of asperities is initially uniform. As the tectonic stress builds up, weak asperities break in sequence resulting in small earthquakes distributed over the fault plane. As the weak asperities break, stress concentrations occur near the stronger asperities and eventually near the strongest one. This redistribution of stress results in clustering of earthquakes near the strongest asperity, the hypocenter of the impending main shock, and relative quiescence elsewhere on the fault plane. This stage corresponds to the "foreshock" activity. When the last asperity breaks, the entire fault plane ruptures resulting in the main shock. In this case, the stress drop in the beginning of the faulting process is substantially higher than the average stress. Hanks (1974) suggested that the 1971 San Fernando earthquake was initiated by an event with a very high stress drop.

Under certain conditions, a failure of one asperity may load up the neighboring asperities and cause failure resulting in a swarm-type earthquake activity.

From the point of view of earthquake prediction, it is important to distinguish foreshock activity from swarm activity or background activity. At the present time no established method exists, but the result for the San Fernando earthquake is encouraging in that the events just before the main shock are very distinct from the earlier events in terms of both wave form and clustering characteristics. Detailed study of temporal variations of wave forms, spectra, source mechanisms, and locations of small earthquakes may be very important for identifying foreshocks.

5. LOW EARTHQUAKE STRESS DROP VERSUS HIGH FRICTIONAL STRENGTH OF ROCKS

Byerlee (1977) conclude that, under laboratory conditions, the coefficient of friction does not depend on mineralogy, pressure, temperature, and texture of the sample. The coefficient of friction was found to be 0.6 ± 0.05 (Figure 12). Under mid-crustal conditions, the normal stress is about 5 kbars so that a frictional strength of about 3 kbars is suggested unless the pore pressure is very large. If the pore pressure is very large, the frictional strength may be about the same order of magnitude as the earthquake stress drop. In this case, earthquakes represent a complete release of the tectonic stress.

On the other hand, if the pore pressure is small compared with the lithostatic pressure, the frictional stress is nearly two orders of magnitude larger than the stress drops in large earthquakes (see Figure 2). In this case the stress drop in earthquakes is only partial (about 1%), and the frictional stress during faulting must be very high. As a result, high heat flow may be expected along the fault zone. The lack of high heat flow along the San Andreas fault has been used as evidence against this idea (Brune et al., 1969). However, if the frictional heat is transferred by mechanisms other than conduction, the lack of a heat flow anomaly may not be compelling evidence against high frictional stress on the fault (Hanks, 1977). Since the earthquake stress drop is very uniform (about 30 to 60 bars), a mechanism which provides a uniform fractional stress drop is necessary. One such mechanism may be suggested from the results of friction experiments (J. Rudnicki, personal communication, 1977; see Figure 12). At high

pressures the shear stress (here interpreted as the frictional stress) fluctuates very little about the average value. This small fluctuation is due to difference in the mineralology, grain size and texture of rocks. Since the fault zone is nonuniform in composition, a small amount of fluctuation in the frictional strength $\Delta\sigma = \sigma_M - \sigma_m$ would be expected where σ_M and σ_m are the maximum and the minimum strengths respectively. In this case, when the tectonic stress exceeds the maximum frictional stress σ_M, a sudden failure takes place and the stress on the fault plane drops. When the stress on the fault plane drops to σ_m then the fault is locked again. Then the average stress drop would be of the order of $\Delta\sigma$. In this model it is the range of the frictional stress that controls the stress drop in earthquakes. Since experimental data suggest that, at pressures corresponding to the mid-crustal depth, the fluctuation of the frictional strength is very small, only a few percent of the frictional strength itself, the stress drop in earthquakes can be a very small fraction of the frictional strength. This model suggests that, if the fault zone is completely homogeneous, the stress drop is zero and stable sliding rather than earthquakes occurs. In the above discussion, dynamic loading effects and the stiffness of the crust are ignored for simplicity. If these effects are included the process would become more complex.

REFERENCES

Abe, K., Determination of seismic moment and energy from the earth's free oscillation, Phys. Earth Planet. Interiors, 4, 49-61, 1970.

Abe, K., Tsunami and mechanism of great earthquakes, Phys. Earth Planet. Interiors, 7, 143-153, 1973.

Abe, K., Seismic displacement and ground motion near a fault: The Saitama earthquake of September 21, 1931, J. Geophys. Res., 79, 4393-4399, 1974a.

Abe, K., Fault parameters determined by near- and far-field data: The Wakasa bay earthquake of March 26, 1963, Bull. Seismol. Soc. Am., 64, 1369-1382, 1974b.

Abe, K., Static and dynamic parameters of the Saitama earthquake of July 1, 1968, Tectonophysics, in press, 1975a.

Abe, K., Reliable estimation of the seismic moment of large earthquakes J. Phys. Earth, 23, 381-390, 1975b.

Aki, K., Generation and propagation of G waves from the Niigata earthquake of June 16, 1964, Bull. Earthquake Res. Inst. Tokyo Univ., 44, 23-88, 1966.

Aki, K., Seismic displacement near a fault, J. Geophys. Res., 73, 5359-5376, 1968.

Aki, K., Earthquake mechanism, Tectonophysics, 13, 423-446, 1972.

Ando, M., A fault-origin model of the great Kanto earthquake of 1923 as deduced from geodetic data, Bull. Earthquake Res. Inst. Tokyo Univ.., 49, 19-32, 1971.

Ben-Menahem, Ari, Radiation of seismic surface-waves from finite moving sources, Bull. Seismol. Soc. Am., 51, 401-435, 1961.

Ben-Menahem, A., and D. G. Harkrider, Radiation patterns of seismic surface waves from buried dipolar point sources in a flat stratified earth, J. Geophys. Res., 69, 2605-2620, 1964.

Ben-Menahem, A., M. Rosenman, and D. G. Harkrider, Fast evaluation of source parameters from isolated surface-wave signals, 1, Universal tables, Bull. Seismol. Soc. Am., 60, 1337-1387, 1970.

Ben-Menahem, A., and M. N. Toksoz, Source mechanism from the spectra of long-period seismic surface waves, 2, The Kamchatka earthquake of November 4, 1952, J. Geophys. Res., 68, 5207-5222, 1963.

Ben-Menahem, A., M. Israel, Uri Levite, Theory and computation of amplitudes of terrestrial line spectra, Geophys. J. R. Astr. Soc., 25, 307-406, 1971.

Benioff, H., F. Press, and S. Smith, Excitation of the free oscillations of the earth by earthquakes, J. Geophys. Res., 66, 605-619, 1961.

Brune, J. N., Radiation pattern of Rayleigh waves from the southeast Alaska earthquake of July 10, 1958, Publ. Dom. Obs. Ottawa, 24, 373-383, 1961.

Brune, J. N., Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res., 75, 4997-5009, 1970.

Brune, J. N., T. L. Henyey, and R. F. Roy, Heat flow, stress, and the rate of slip along the San Andreas fault, California, J. Geophys. Res., 74, 3821-3827, 1969.

Brune, J. N., J. Prince, S. Hartzell, L. Mungia, A. Reyes, Acapulco strong motion record of October 6, 1974, abstract, Seismo. Soc. Soc. Am., meeting program, 394, 1977.

Burdick, L. J., and G. R. Melleman, Inversion of the body waves from the Borrego mountain earthquake to the source mechanism, Bull. Seismol. Soc. Am., 66, 1485-1499, 1976.

Byerlee, J., Friction of Rocks, Proceedings of Conference II: Experimental studies of rock friction with application to earthquake prediction, edited by J. F. Evernen, U. S. Geological Survey, Menlo Park, pp. 55-77, 1977.

Chinnery, M. A., The strength of the earth's crust under horizontal shear stress, J. Geophys. Res., 69, 2085-2089, 1964.

Chinnery, Michael A., Theoretical Fault Models, Publications of the Dominion Observatory, Ottawa, XXXVII, No. 7, 211-223, 1969.

Chung, Wai-Ying and H. Kanamori, Source process and tectonic implications of the Spanish deep-focus earthquake of March 29, 1954, Phys. Earth Planet. Interiors, 13, 85-96, 1976.

Chung, Wai-Ying and H. Kanamori, Subduction process of a fracture-zone and aseismic ridges - focal mechanism and source characteristics of the New Hebrides earthquake of January 19, 1969 and some related events, Geophys. J. R. Astr. Soc., in press, 1978.

Dziewonski, A. M., and F. Gilbert, Temporal variations of the seismic moment tensor and the evidence of precursive compressin for two deep earthquakes, Nature, 247, 185-188, 1974.

Fukao, Y., Focal process of a deep focus earthquake as deduced from long period P and S waves, Bull. Earthquake Res. Inst. Tokyo Univ., 48, 707-727, 1970.

Fukao, Y., Seismic body waves from surface faults, J. Phys. Earth, 19, 271-281, 1971.

Fukao, Y., Source process of a large deep-focus earthquake and its tectonic implications - The western Brazil earthquake of 1963, Phys. Earth Planet. Interiors, 5, 61-76, 1972.

Geller, R. J., Scaling relations for earthquake source parameters and magnitudes, Bull. Seism. Soc. Am., 66, 1501-1523, 1976.

Gilbert, F., and A. M. Dziewonski, An application of normal mode theory to the retrieval of structural parameters and source mechanisms from seismic spectra, Philosophical Transactions of the Royal Society of London, A, 278, 187-269, 1975.

Hanks, T., The faulting mechanism of the San Fernando earthquake, J. Geophys. Res., 79, 1215-1229, 1974.

Hanks, T. C., Earthquake stress drops, ambient tectonic stresses and stresses that derive plate motions, Pure Appl. Geophys., 115, 441-458, 1977.

Hatori, T., On the tsunami which accompanied the Niigata earthquake of June 16, 1964 - Source deformation, propagation and tsunami run-up, Bull. Earthquake Res. Inst. Tokyo Univ., 43, 129-148, 1965.

Hatori, T., Vertical displacement in a tsunami source area and the topography of the sea bottom, Bull. Earthquake Res. Inst. Tokyo Univ., 44, 1449-1464, 1966.

Helmberger, D. V., Generalized ray theory for shear dislocations, Bull. Seismol. Soc. Am., 64, 45-64, 1974.

Helmberger, D. V., and L. R. Johnson, Source parameters of moderate size earthquakes and the importance of receiver crustal structure in interpreting observations of local earthquakes, Bull. Seismol. Soc. Am., 67, 301-313, 1977.

Helmberger, D. V., and S. D. Malone, Modeling local earthquakes as shear dislocations in a layered half space, J. Geophys. Res., 80, 4881-4888, 1975.

Hirasawa, T., A least squares method for the focal mechanism determination from S wave data; Part I, Bull. Earthquake Res. Inst. Tokyo Univ., 44, 901-918, 1966.

Honda, H., Earthquake mechanism and seismic waves, J. Physics of Earth, 19, No. 2, 1-97, 1962.

Imamura, A., Theoretical and Applied Seismology, Maruzen, Tokyo, 358 pp. 1937.

Ishida, M., and H. Kanamori, The spatio-temporal variation of seismicity before the 1971 San Fernando earthquake, California, Geophys. Res. Let., 4, 345-346, 1977.

Ishida, M., and H. Kanamori, The foreshock activity of the 1971 San Fernando earthquake, California, submitted to Bull. Seismol. Soc. Am., 1978.

Jarosch and Aboodi, Towards a unified notation of source parameters, Geophys. J. R. Astr. Soc., 21, 513-529, 1970.

Jungels, P. H., and G. A. Frazier, Finite element analysis of the residual displacements for an earthquake rupture: Source parameters for the San Fernando earthquake, J. Geophys. Res., 78, 5062-5083, 1973.

Kanamori, H., Synthesis of long-period surface waves and its application to earthquake source studies - Kurile Islands earthquake of October 13, 1963, J. Geophys. Res., 75, 5011-5027, 1970.

Kanamori, H., Determination of effective tectonic stress associated with earthquake faulting - Tottori earthquake of 1943, Phys. Earth Planet. Interiors, 5, 426-434, 1972.

Kanamori, H., Mode of strain release associated with major earthquakes in Japan, in Annual Review of Earth and Planetary Science, Vol. 1, edited by F. A. Donath, 213-239, 1973.

Kanamori, H., and D. L. Anderson, Theoretical basis of some empirical relations in seismology, Bull. Seismol. Soc. Am., 65, 1073-1095, 1975.

Kanamori, H., and G. S. Stewart, Seismological aspects of the Guatemala earthquake of February 4, 1976, submitted to J. Geophys. Res., 1978.

Kasahara, K., The nature of seismic origins as inferred from seismological and geodetic observations, 1, Bull. Earthquake Res. Inst. Tokyo Univ., 35, 473-532, 1957.

Kelleher, J., and J. Savino, Distribution of seismicity before large strike-slip and thrust-type earthquakes, J. Geophys. Res., 80, 260-271, 1975.

Knopoff, L., Energy release in earthquakes, Geophys. J., 1, 44-52, 1958.

Langston, C. A., A body wave inversion of the Koyna, India earthquake of December 10, 1967, and some implications for body wave focal mechanisms, J. Geophys. Res., 81, 2517-2529, 1976.

Langston, C. A., and R. Butler, Focal mechanism of the August 1, 1975, Oroville earthquake, Bull. Seismol. Soc. Am., 66, 1111-1120, 1976.

Langston, C. A., and D. Helmberger, A procedure for modelling shallow dislocation sources, Geophys. J., 42, 117-130, 1975.

Madariaga, R., Implications of stress-drop models of earthquakes for the inversion of stress drop from seismic observations. Pure Appl. Geophys., 115, 301-316, 1977.

Mikumo, T., Faulting process of the San Fernando earthquake of Feb. 9, 1971, inferred from static and dynamic near-field displacements, Bull. Seismol. Soc. Am., 63, 249-269, 1973.

Miyabe, N., An investigation of the Sanriku tsunami based on mareogram data, Bull. Earthquake Res. Inst. Tokyo Univ., Suppl., 1, 112-126, 1934.

Miyamura, S., S. Omote, R. Teisseyre, and E. Vesanen, Multiple shocks and earthquake series pattern, Int. Inst. Seismol. Earthquake Eng. Bull., 2, 71-92, 1964.

Mogi, K., Development of aftershock areas of great earthquakes, Bull. Earthquake Res. Inst. Tokyo Univ., 46, 175-203, 1968a.

Mogi, K., Source locations of elastic shocks in the fracturing process in rocks (1), Bull. Earthquake Res. Inst. Tokyo Univ., 46, 1103-1125, 1968b.

Plafker, G., Alaskan earthquake of 1964 and Chilean earthquake of 1960: Implications for arc tectonics, J. Geophys. Res., 77, 901-925, 1972.

Press, F., A. Ben-Menahem, and N. Toksoz, Experimental determination of earthquake fault length and rupture velocity, J. Geophys. Res., 66, 3471-3485, 1961.

Rikitake, T., Probability of earthquake occurrence as estimated from crustal strain, Tectonophysics, 23, 299-312, 1974.

Saito, M., Excitation of free oscillations and surface waves by a point source in avertically heteorogeneous earth, J. Geophys. Res., 72, 3689-3699, 1967.

Savage, J. C., and L. M. Hastie, Surface deformation associated with dip-slip faulting, J. Geophys. Res., 71, 4897, 1966.

Stauder, W. S. J., The focal mechanism of earthquakes, in Advances in Geophysics, 9, 1-76, 1962.

Trifunac, M. D., Stress estimates for the San Fernando, California, earthquake of February 9,1971; Main event and thirteen aftershocks, Bull. Seismol. Soc. Am., 62, 721-750, 1972.

Trifunac, M. D., A three-dimensional dislocation model for the San Fernando California, earthquake of February 9, 1971, Bull. Seismol. Soc. Am., 64, 149-172, 1974.

Trifunac, M. D., and J. N. Brune, Complexity of energy release during the Imperial Valley, California, earthquake of 1940, Bull. Seismol. Soc. Am., 60, 137-160, 1970.

Wyss, M., and J. Brune, The Alaska earthquake of 28 March 1964: A complex multiple rupture, Bull. Seismol. Soc. Am., 57, 1017-1023, 1967.

TABLE 1

EARTHQUAKE SOURCE PARAMETERS

Event	Date	M_s	m_b	M_o × 10^{27} dyne-cm	L (km)	W (km)	D (m)	τ (sec)	τ^* (sec)	V_R (km/sec)	$\Delta\sigma$ (bars)
1. Kanto	1 Sep. 1923	8.2	—	7.6	130	70	2.1	7	10	--	21
2. Tango	27 Mar. 1927	7.75	—	0.46	35	13	3	6	2.5	2.3	115
3. North Izu	25 Nov. 1930	7.1	—	0.2	20	11	3	—	1.7	-	150
4. Saitama	21 Sep. 1933	6.75	—	0.068	20	10	1	2	1.6	2.3	59
5. Sanriku	2 Mar. 1933	8.3	—	43	185	100	3.3	7	12	3.2	42
6. Long Beach	11 Mar. 1933	6.25	—	0.028	30	15	0.2	2	2.5	2.3	7
7. Imperial Valley	19 May 1940	7.1	—	0.48	70	11	2	—	3.2	—	55
8. Tottori	10 Sep. 1943	7.4	—	0.36	33	13	2.5	3	4.0	2.3	99
9. Tonankai	7 Dec. 1944	8.2	—	15	120	80	3.1	—	9.2	--	39
10. Mikawa	12 Jan. 1945	7.1	—	0.087	12	11	2.2	—	1.3	....	140
11. Nankaido	20 Dec. 1946	8.2	—	15	120	80	3.1	—	9.2	---	39
12. Fukui	28 Jun. 1948	7.3	—	0.33	30	13	2	2	1.9	2.3	100
13. Tokachi-Oki	4 Mar. 1952	8.3	—	17	180	100	1.9	—	14	—	17
14. Kern County	21 Jul. 1952	7.7	--	2	60	18	4.6	1	3.6	-	140
15. Fairview	16 Dec. 1954	7.1	—	0.13	36	6	2	--	1.7	--	100
16. Chile	22 May 1960	8.3	--	2400	800	200	21	—	36	3.5	91
17. Kitamino	19 Aug. 1961	7.0	—	0.09	12	10	2.5	2	1.3	3.0	170
18. Wasaka Bay	27 Mar. 1963	6.9	--	0.033	20	8	0.6	2	1.5	2.3	40
19. North Atlantic I	3 Aug. 1963	6.7	6.1	0.12	32	11	1	—	2.2	—	44
20. Kurile Islands	13 Oct. 1963	8.2	5.7	75	250	140	3	—	17	3.5	28
21. North Atlantic II	17 Nov. 1963	6.5	5.9	0.038	27	9	0.48	—	1.8	—	24
22. Spain	15 Mar. 1964	7.1	6.2	0.13	95	10	0.42	—	3.6	1.4	11
23. Alaska	28 Mar. 1964	8.5	6.2	520	500	300	7	—	35	3.5	22
24. Niigata	16 Jun. 1964	7.4	6.1	3.2	80	30	3.3	—	5.3	—	66
25. Rat Island I	4 Feb. 1965	7.9	6.0	140	500	150	2.5	—	25	4.0	17
26. Rat Island II	30 Mar. 1965	7.5	5.7	3.4	50	80	1.2	—	5.8	--	33
27. Parkfield	28 Jun. 1966	6.4	5.3	0.032	26	7	0.6	0.7	1.6	2.7	32
28. Aleutian	4 Jul. 1966	7.2	6.2	0.226	35	12	1.6	—	2.4	—	64
29. Truckee	12 Sep. 1966	5.9	5.4	0.0083	10	10	0.3	—	1.2	—	20
30. Peru	17 Oct. 1966	7.5	6.3	20	80	140	2.6	—	9.6	—	41
31. Borrego	9 Apr. 1968	6.7	6.1	0.063	33	11	0.58	—	2.2	—	22
32. Tokachi-Oki	16 May 1968	8.0	5.9	28	150	100	4.1	—	12	3.5	37
33. Saitama	1 Jul. 1968	5.8	5.9	0.019	10	6	0.92	1	0.9	3.4	100
34. Portuguese	28 Feb. 1969	8.0	7.3	5.5	80	50	2.5	—	6.1	—	53
35. Kurile Islands	11 Aug. 1969	7.8	7.1	22	180	85	2.9	—	12	3.5	28
36. Gifu	9 Sep. 1969	6.6	5.5	0.035	18	10	0.6	1	1.7	2.5	35
37. Peru	31 May 1970	7.8	6.6	10	130	70	1.6	—	8.7	2.5	28
38. San Fernando	9 Feb. 1971	6.6	6.2	0.12	20	14	1.4	1	2.0	2.4	62
39. Nemuro-Oki	17 Jun. 1973	7.7	6.5	6.7	60	100	1.6	—	7.5	—	35
40. Turkey	22 Jul. 1967	7.1	6.0	0.83	80	20	1.7	—	4.7	—	32
41. Iran	31 Aug. 1968	7.3	5.9	1	80	20	2.1	—	4.7	—	38

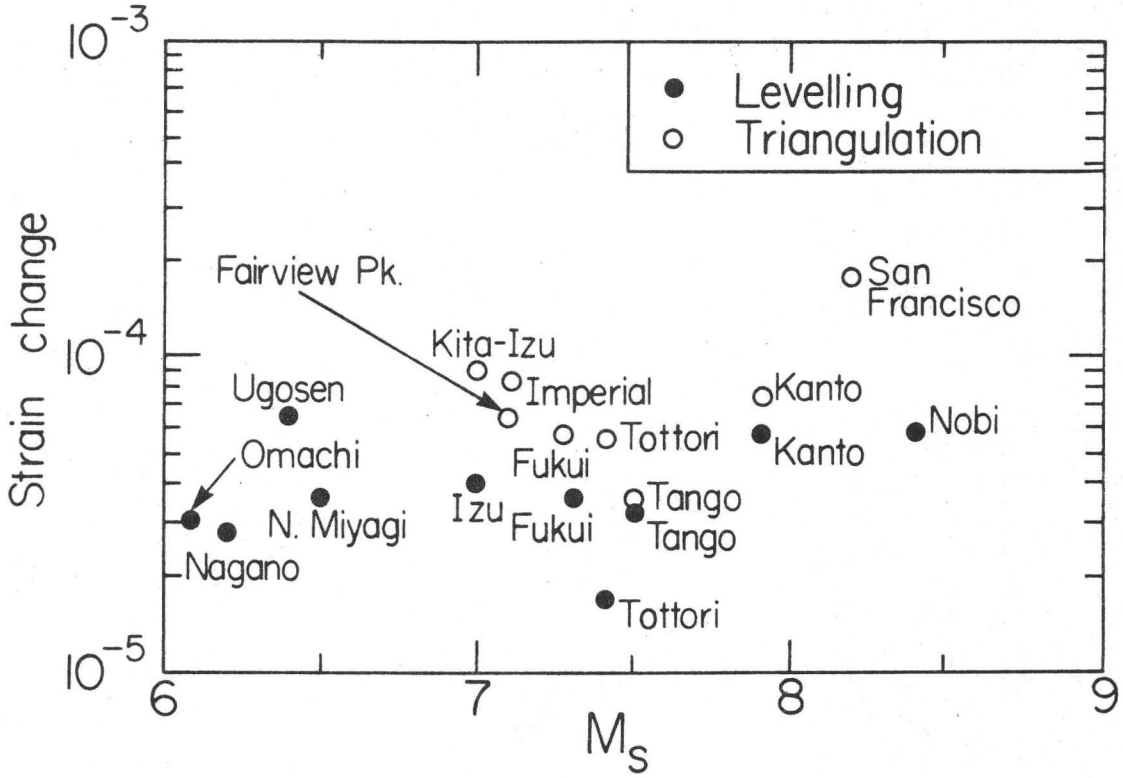

Levelling
Triangulation
Strain change
10^{-3}
10^{-4}
10^{-5}
6
7
8
9
M_S
Fairview Pk.
Kita-Izu
Imperial
Ugosen
Omachi
N. Miyagi
Nagano
Izu
Fukui
Fukui
Tottori
Tottori
Tango
Tango
Kanto
Kanto
San Francisco
Nobi

Fig. 1

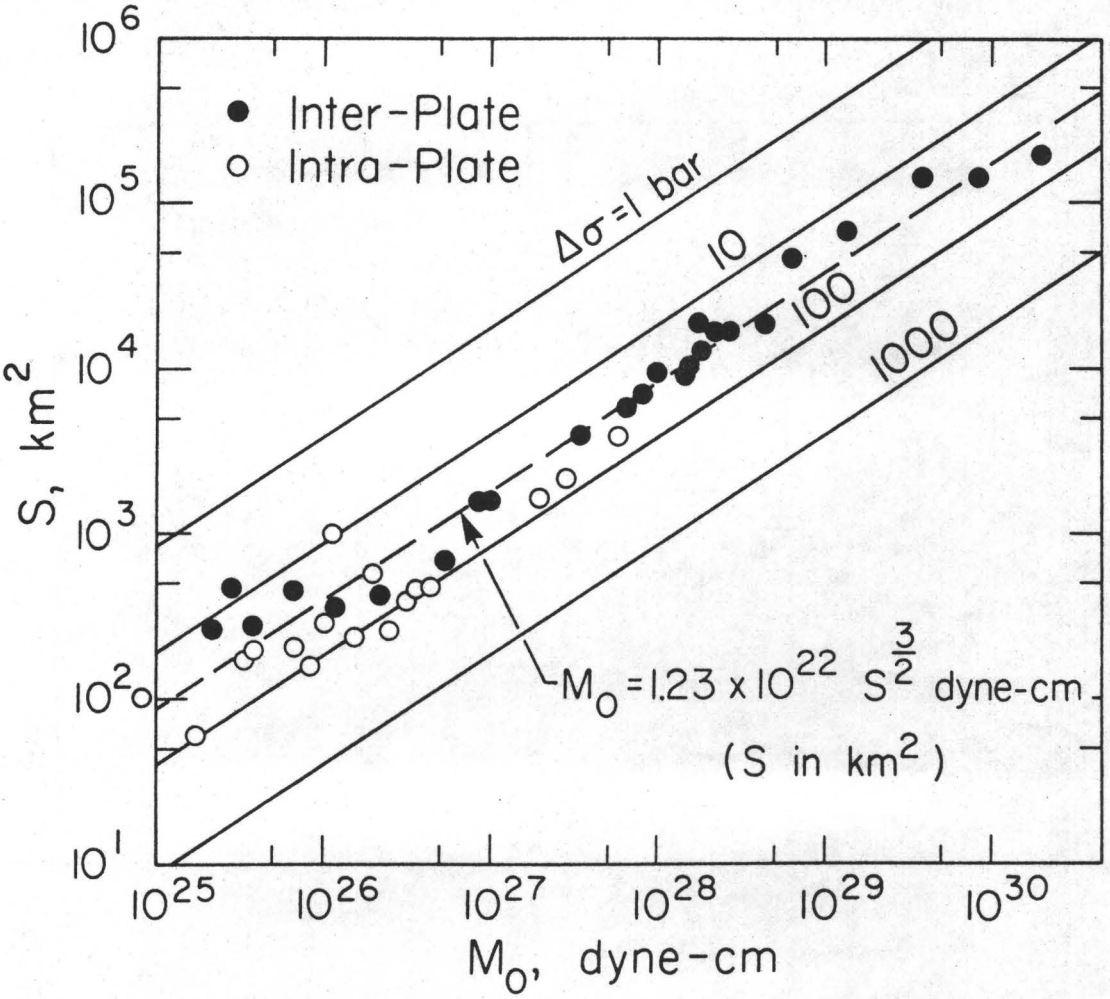

Inter-Plate
Intra-Plate
Δσ=1 bar
10
100
1000
$M_0 = 1.23 \times 10^{22}\ S^{\frac{3}{2}}$ dyne-cm
(S in km²)
S, km²
M_0, dyne-cm
10^1
10^2
10^3
10^4
10^5
10^6
10^{25}
10^{26}
10^{27}
10^{28}
10^{29}
10^{30}

Fig. 2

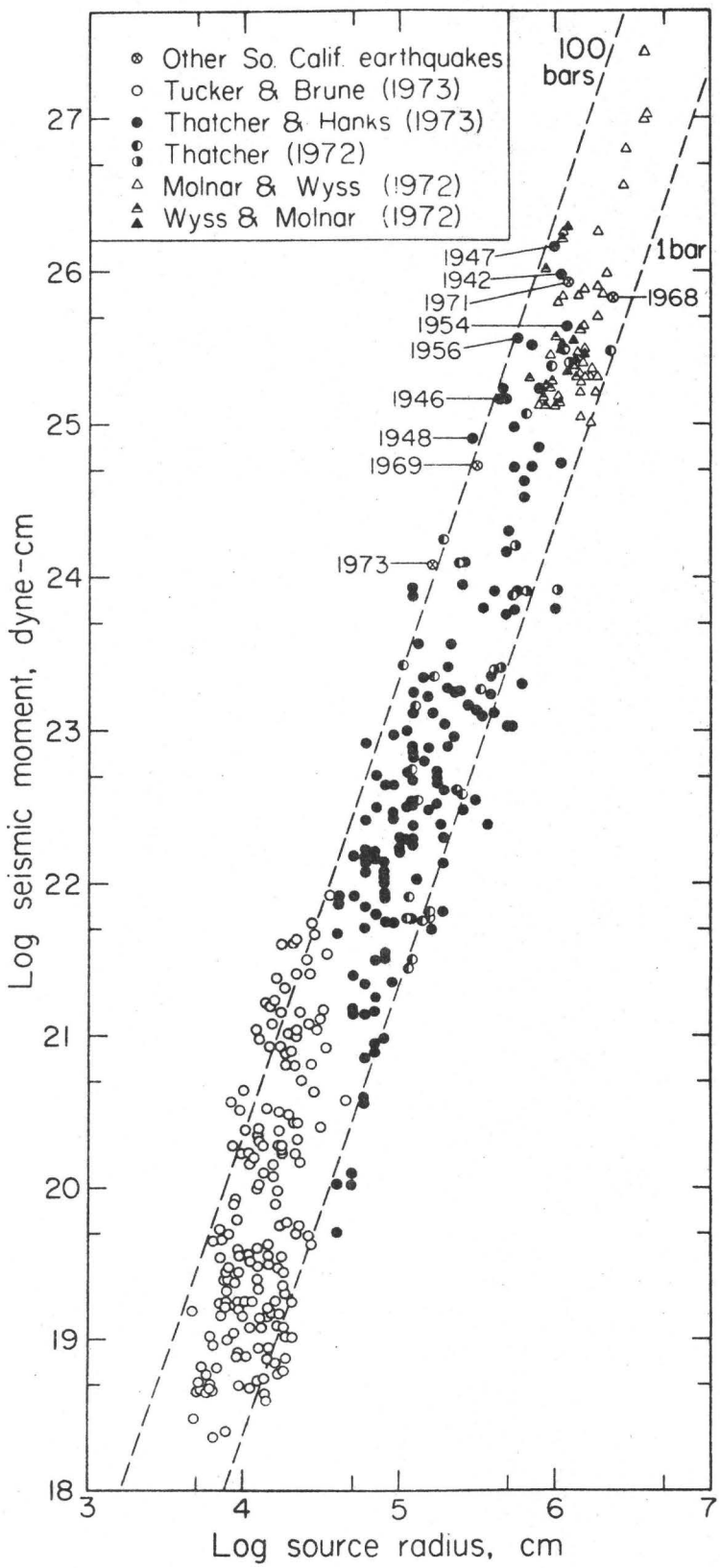

Other So. Calif. earthquakes
Tucker & Brune (1973)
Thatcher & Hanks (1973)
Thatcher (1972)
Molnar & Wyss (1972)
Wyss & Molnar (1972)
100 bars
1 bar
1947
1942
1971
1954
1956
1946
1948
1969
1973
1968
Log seismic moment, dyne-cm
Log source radius, cm
18
19
20
21
22
23
24
25
26
27
3
4
5
6
7

Fig. 3

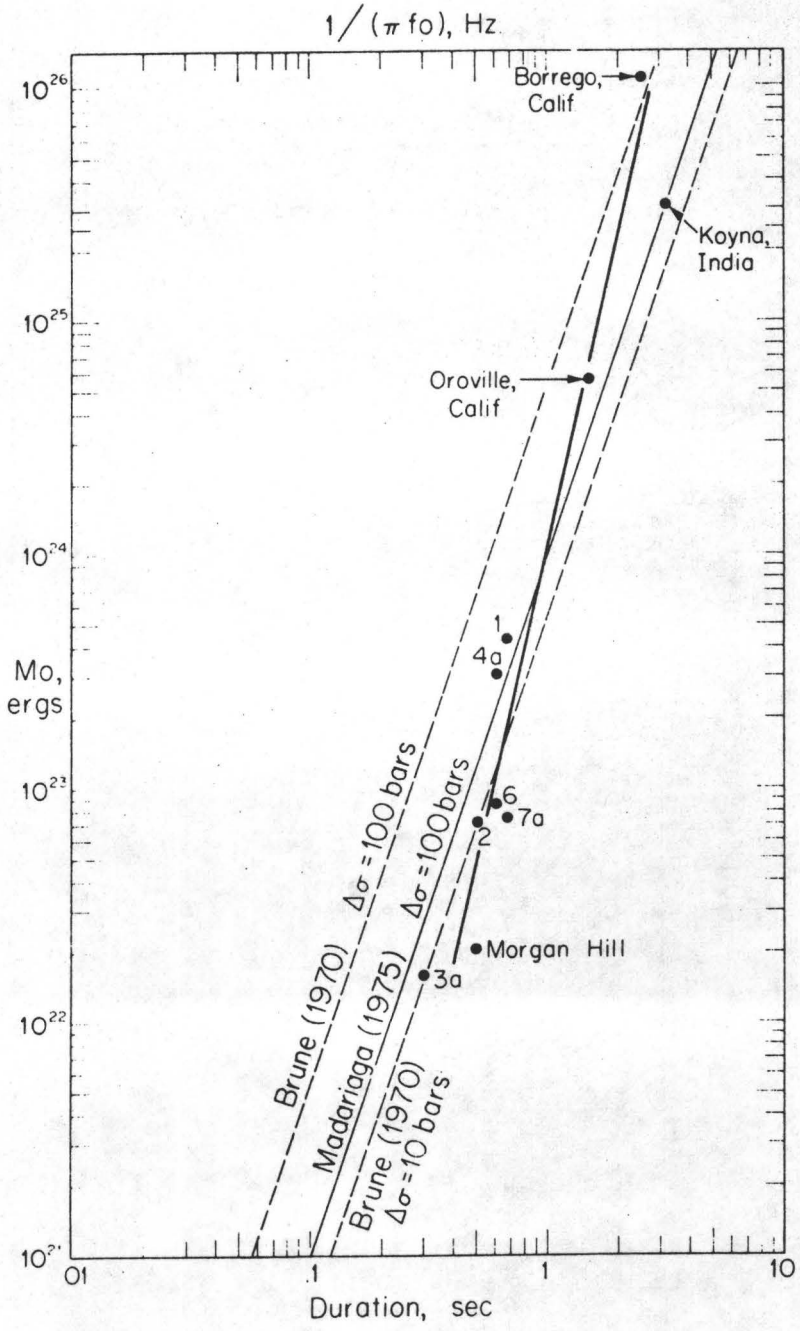
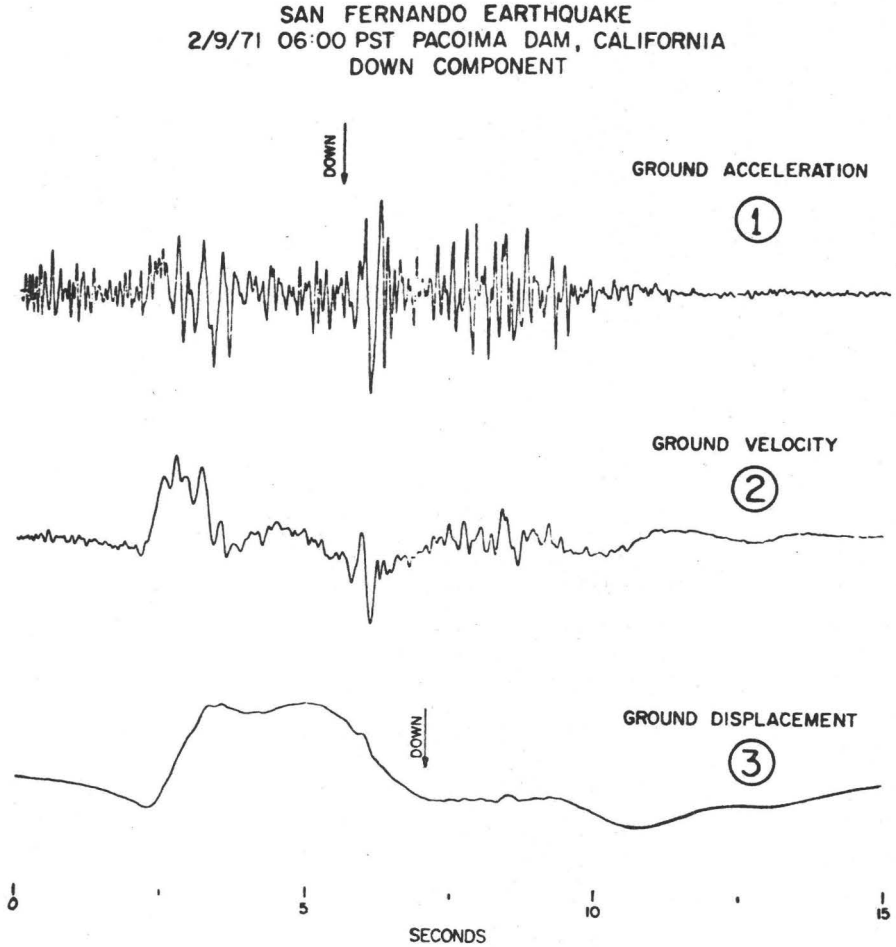

1/(π fo), Hz
Borrego, Calif
Koyna, India
Oroville, Calif
Morgan Hill
1
4a
6
7a
2
3a
Brune (1970) Δσ = 100 bars
Madariaga (1975) Δσ = 100 bars
Brune (1970) Δσ = 10 bars
Mo, ergs
Duration, sec
10^26
10^25
10^24
10^23
10^22
10^21
.01
.1
1
10

Fig. 4


SAN FERNANDO EARTHQUAKE
2/9/71 06:00 PST PACOIMA DAM, CALIFORNIA
DOWN COMPONENT
DOWN
GROUND ACCELERATION
1
GROUND VELOCITY
2
DOWN
GROUND DISPLACEMENT
3
0
5
10
15
SECONDS

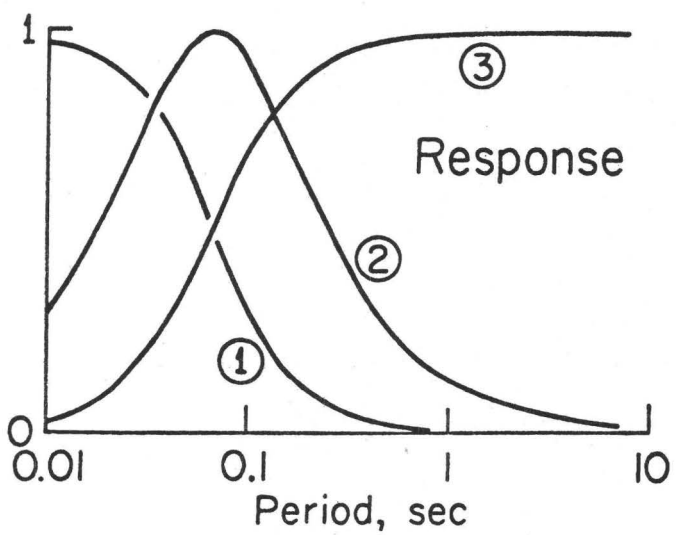

1
3
Response
2
1
0
0.01
0.1
1
10
Period, sec

Fig. 5

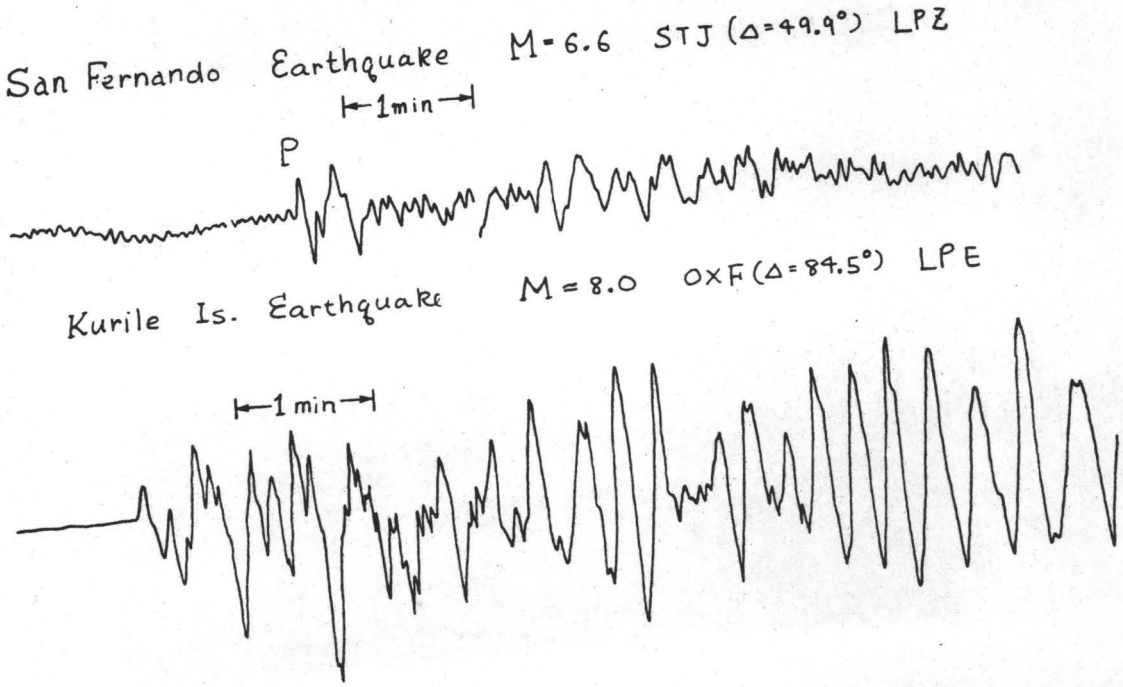

San Fernando Earthquake M-6.6 STJ (Δ=49.9°) LPZ
1min
P
Kurile Is. Earthquake M = 8.0 OXF (Δ=84.5°) LPE
1 min

Fig. 6

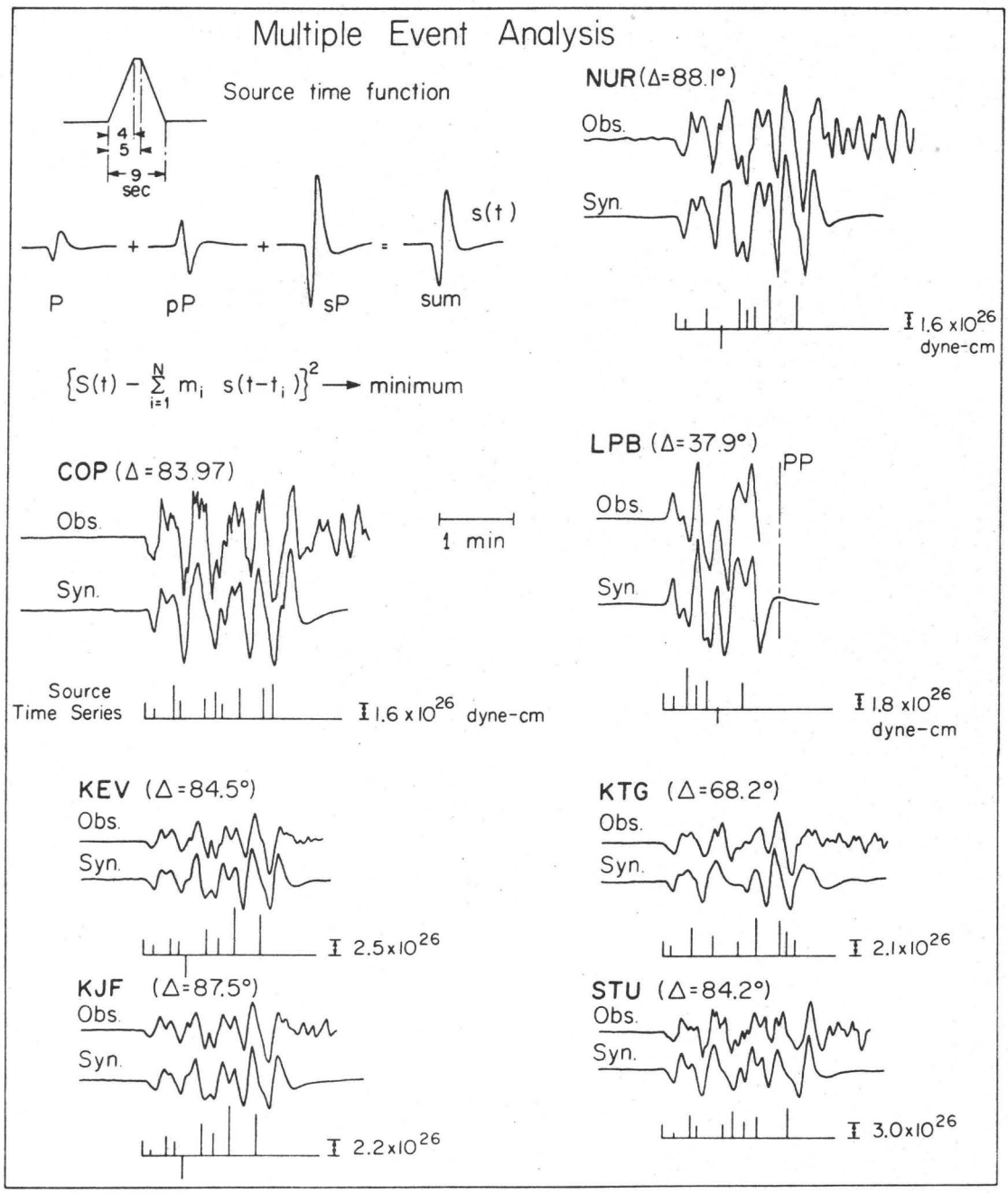

Multiple Event Analysis
Source time function
4
5
9
sec
s(t)
P
pP
sP
sum
$\left\{S(t) - \sum_{i=1}^{N} m_i\ s(t-t_i)\right\}^2 \rightarrow$ minimum
NUR (Δ=88.1°)
Obs.
Syn.
I 1.6 x10^26 dyne-cm
COP (Δ=83.97)
Obs.
Syn.
1 min
Source Time Series
I 1.6 x10^26 dyne-cm
LPB (Δ=37.9°)
PP
Obs.
Syn.
I 1.8 x10^26 dyne-cm
KEV (Δ=84.5°)
Obs.
Syn.
I 2.5x10^26
KTG (Δ=68.2°)
Obs.
Syn.
I 2.1x10^26
KJF (Δ=87.5°)
Obs.
Syn.
I 2.2x10^26
STU (Δ=84.2°)
Obs.
Syn.
I 3.0x10^26

Fig. 7

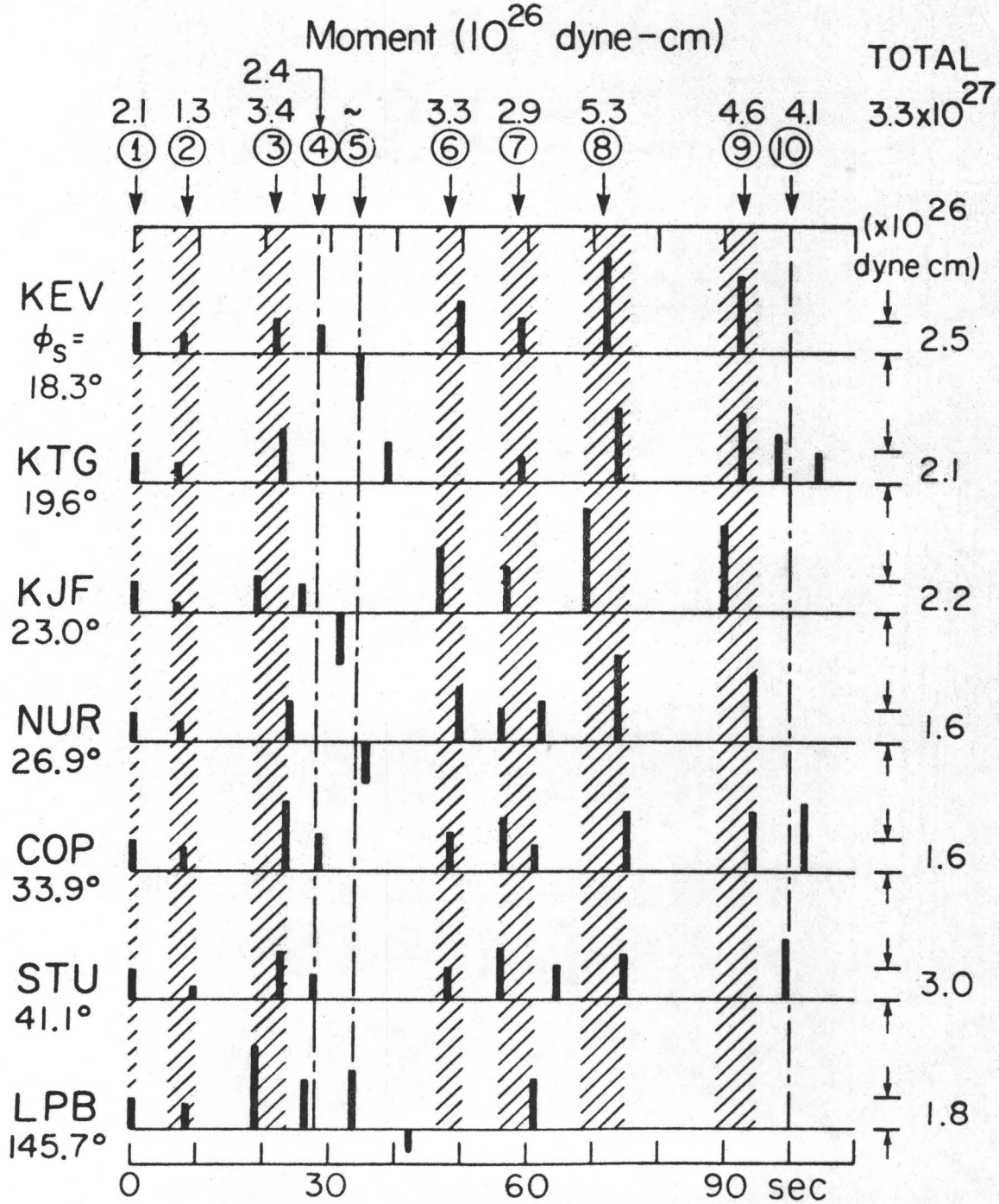
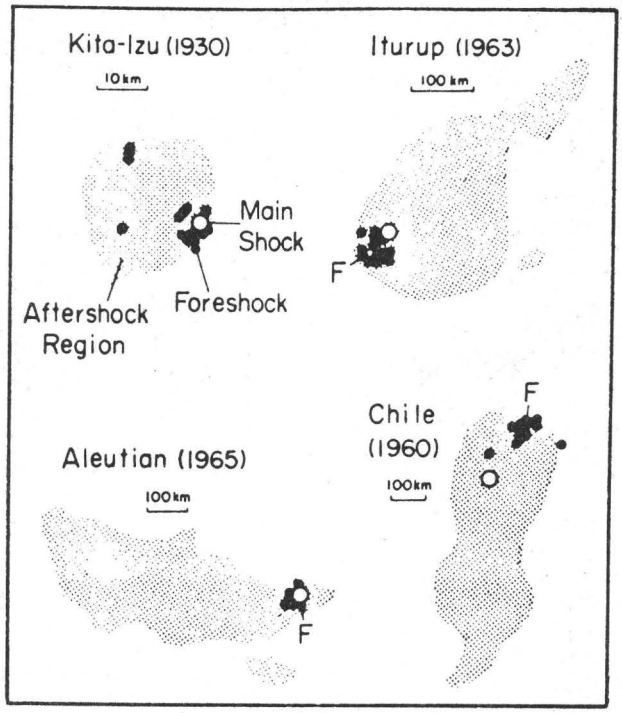
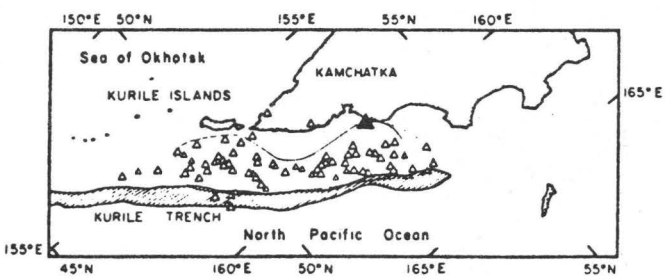
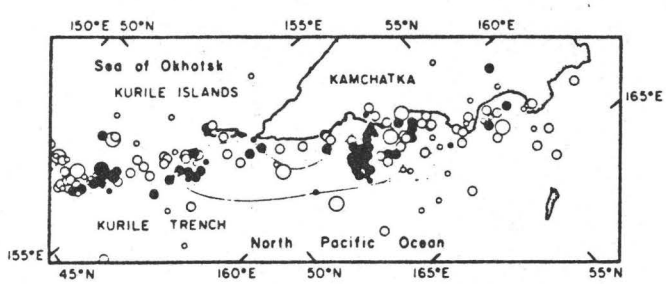

Moment (10^{26} dyne-cm)
TOTAL
3.3×10^{27}
2.4
2.1 1.3 3.4 ~ 3.3 2.9 5.3 4.6 4.1
① ② ③ ④ ⑤ ⑥ ⑦ ⑧ ⑨ ⑩
($\times 10^{26}$ dyne cm)
KEV
ϕ_s = 18.3°
2.5
KTG
19.6°
2.1
KJF
23.0°
2.2
NUR
26.9°
1.6
COP
33.9°
1.6
STU
41.1°
3.0
LPB
145.7°
1.8
0 30 60 90 sec

Fig. 8


Kita-Izu (1930)
10km
Main
Shock
Foreshock
Aftershock
Region
Iturup (1963)
100 km
F
Chile
(1960)
100km
F
Aleutian (1965)
100km
F


150°E 50°N
155°E
55°N
160°E
Sea of Okhotsk
KAMCHATKA
KURILE ISLANDS
165°E
KURILE TRENCH
North Pacific Ocean
155°E
45°N
160°E
50°N
165°E
55°N


150°E 50°N
155°E
55°N
160°E
Sea of Okhotsk
KAMCHATKA
KURILE ISLANDS
165°E
KURILE TRENCH
North Pacific Ocean
155°E
45°N
160°E
50°N
165°E
55°N

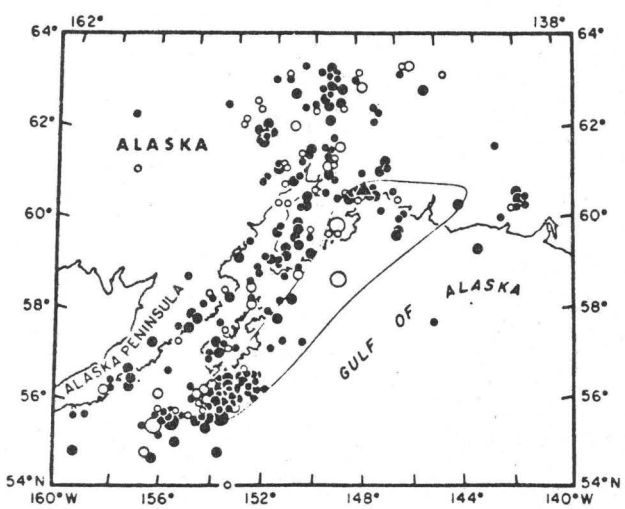

64°
162°
138°
62°
ALASKA
60°
58°
ALASKA PENINSULA
GULF OF ALASKA
56°
54°N
160°W
156°
152°
148°
144°
140°W

Fig. 9

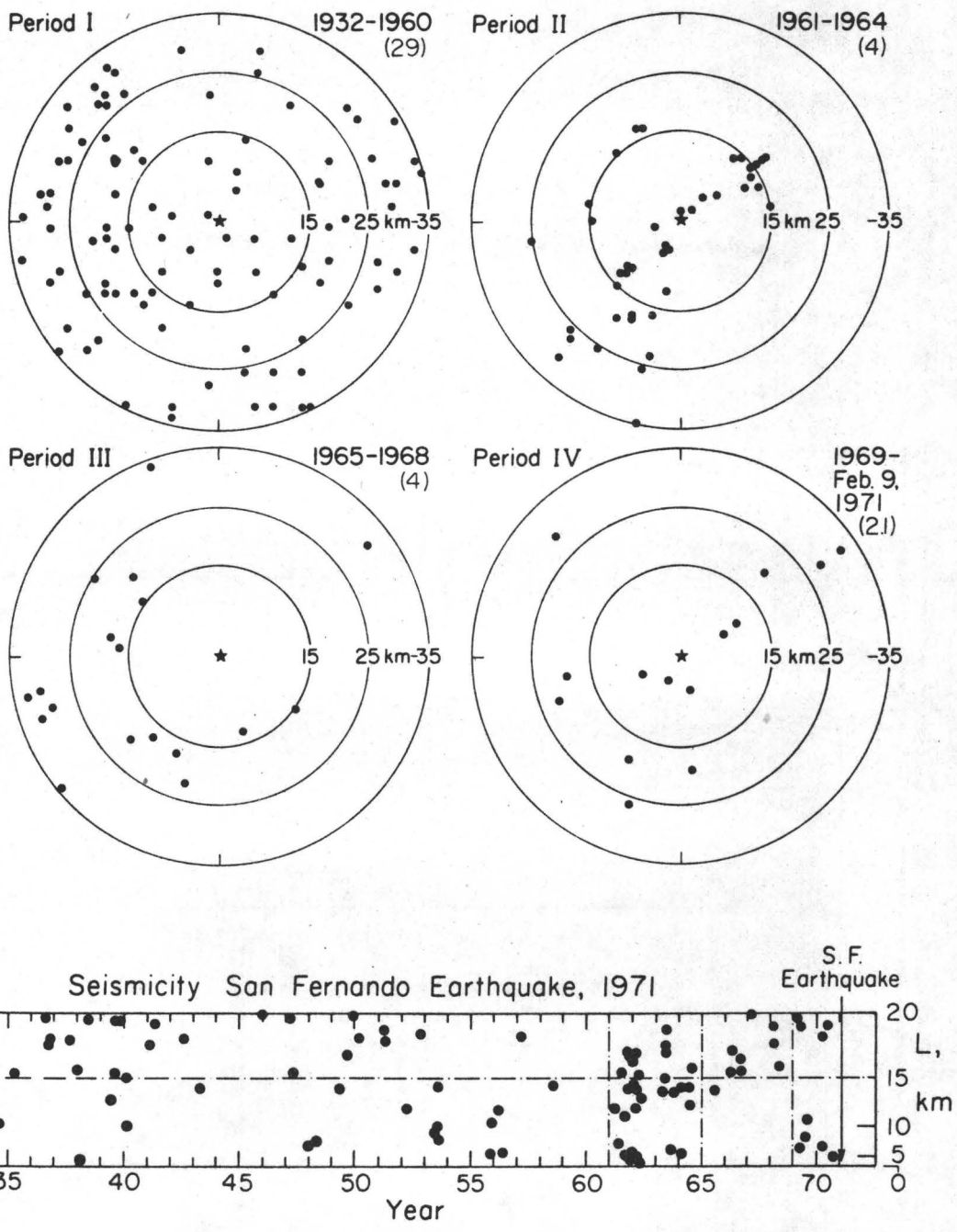

Period I
1932-1960
(29)
15
25 km-35
Period II
1961-1964
(4)
15 km 25
-35
Period III
1965-1968
(4)
15
25 km-35
Period IV
1969-
Feb. 9,
1971
(2.1)
15 km 25
-35
Seismicity San Fernando Earthquake, 1971
S.F.
Earthquake
20
15
10
5
L,
km
35
40
45
50
55
60
65
70
0
Year

Fig. 10

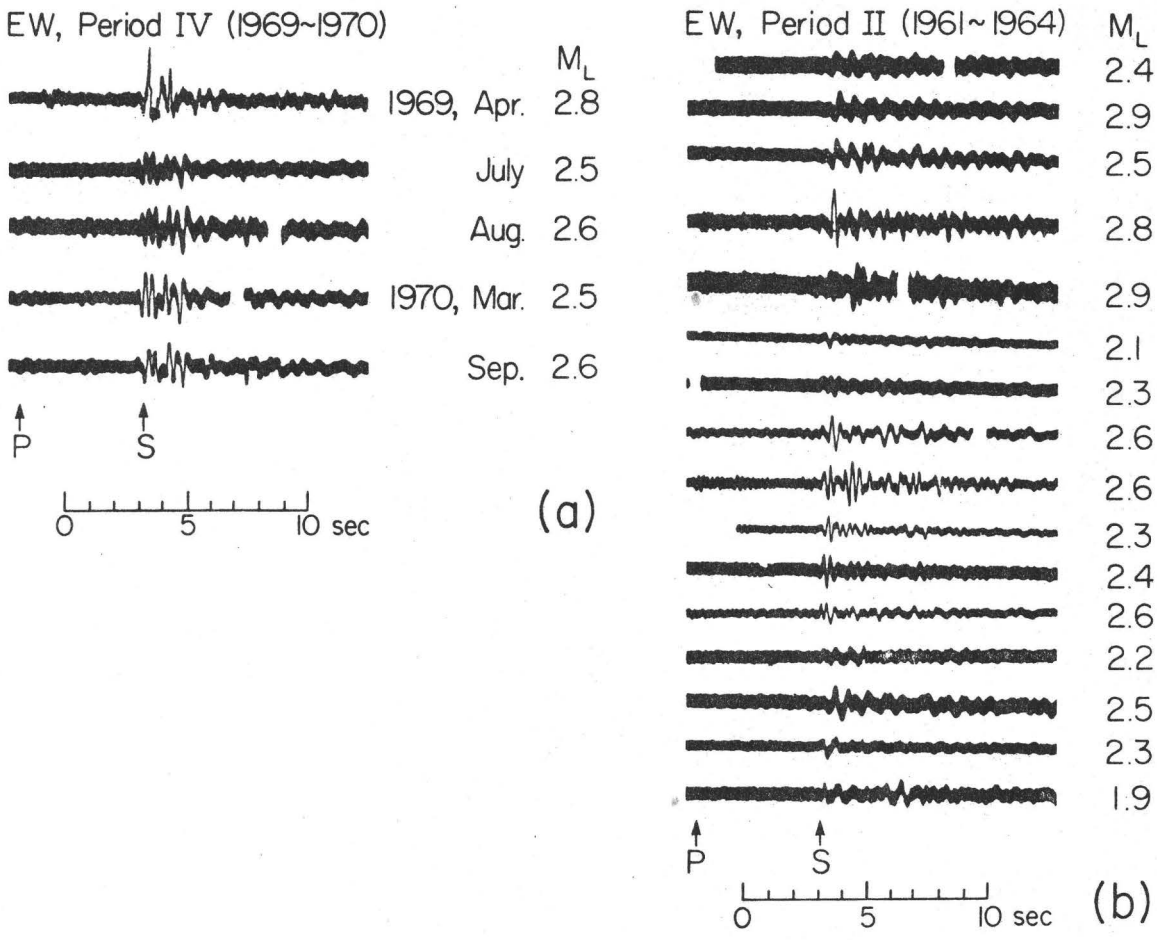

EW, Period IV (1969~1970)
M_L
1969, Apr. 2.8
July 2.5
Aug. 2.6
1970, Mar. 2.5
Sep. 2.6
P
S
0
5
10 sec
(a)
EW, Period II (1961~1964)
M_L
2.4
2.9
2.5
2.8
2.9
2.1
2.3
2.6
2.6
2.3
2.4
2.6
2.2
2.5
2.3
1.9
P
S
0
5
10 sec
(b)

Fig. 11a, b

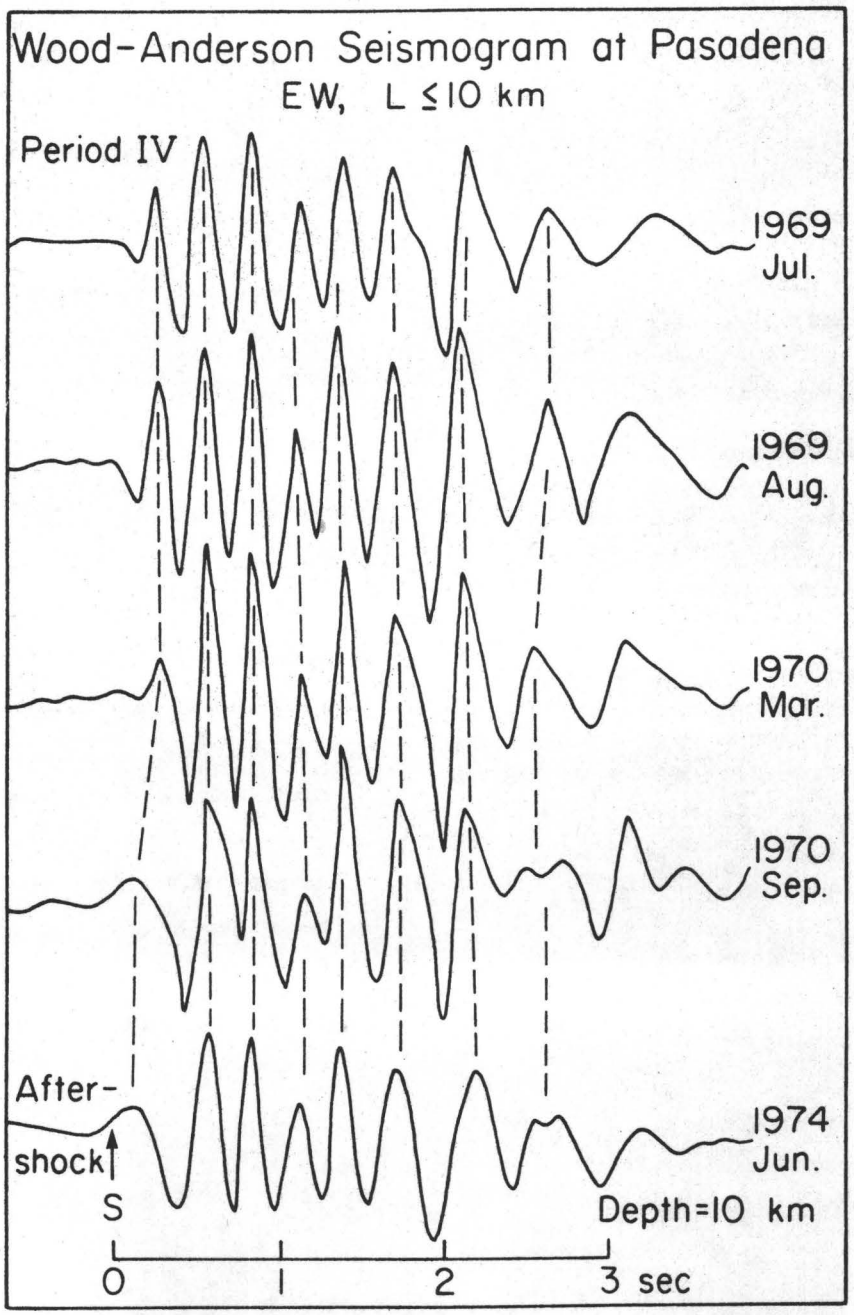

Wood-Anderson Seismogram at Pasadena
EW, L ≤ 10 km
Period IV
1969 Jul.
1969 Aug.
1970 Mar.
1970 Sep.
After-shock
S
1974 Jun.
Depth=10 km
0
1
2
3 sec

Fig. 11c

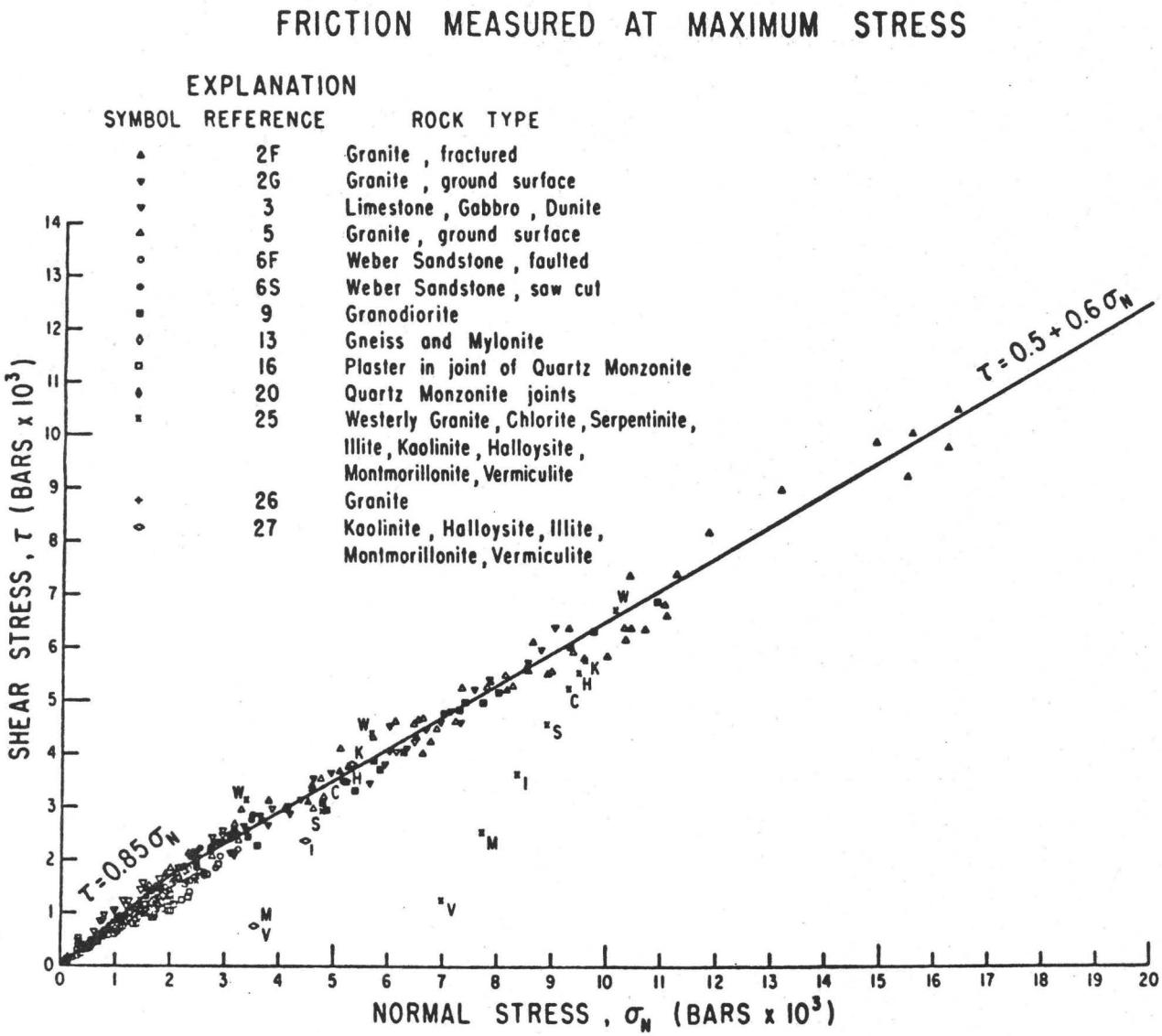

FRICTION MEASURED AT MAXIMUM STRESS
EXPLANATION
SYMBOL REFERENCE ROCK TYPE
2F Granite , fractured
2G Granite , ground surface
3 Limestone , Gabbro , Dunite
5 Granite , ground surface
6F Weber Sandstone , faulted
6S Weber Sandstone , saw cut
9 Granodiorite
13 Gneiss and Mylonite
16 Plaster in joint of Quartz Monzonite
20 Quartz Monzonite joints
25 Westerly Granite , Chlorite , Serpentinite , Illite , Kaolinite , Halloysite , Montmorillonite , Vermiculite
26 Granite
27 Kaolinite , Halloysite , Illite , Montmorillonite , Vermiculite
τ = 0.5 + 0.6 σN
τ = 0.85 σN
SHEAR STRESS , τ (BARS x 10³)
NORMAL STRESS , σN (BARS x 10³)
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Fig. 12

A RELATION BETWEEN SEISMIC MOMENT AND STRESS DROP OF COMPLEX EARTHQUAKES(§)

Raul MADARIAGA
Institut de Physique du Globe
Université de Paris VI
75230 Paris Cedex 05
FRANCE

§ Contribution IPG N° 275

ABSTRACT

We study the relation between seismic moment and the stress drop on the fault, or subfaults, of multiple events. By means of the reciprocity theorem we show that the seismic moment is an integral of the stress drop on the fault weighted by a slip function for the same fault geometry but uniform stress drop equal to μ , the rigidity. This relation shows that the seismic moment is strongly dependent on the geometry of the fault. In particular, for a given stress drop, a multiple fault has a smaller seismic moment than a simple fault of the same area. Conversely, for the same seismic moment and total area the stress drop estimate obtained by the usual smooth circular fault approximation underestimates the actual stress drop on the subfaults. The stress drop is subestimated by factors of ρ /R where ρ is the typical radius of a subfault and R is the characteristic radius of the total fault area.

Approximations for multiple circular faults demonstrate that the absolute high frequency radiation for multiple faults and simple faults of the same stress drop and total area are approximately the same. However in relative terms, for the same low frequency seismic moment, multiple faults radiate more high frequency energy than simple faults.

INTRODUCTION

Since the determination of the seismic moment of the Niigata Earthquake by Aki (1966) the number of earthquakes for which this source parameter has been accurately determined has steadily increased. Kanamori and Anderson (1975) and Kanamori (1977) have presented a list of seismic moments for about 50 large events. For these events, these authors estimate that the seismic is probably measured with a precision of a factor of two. Many determinations of seismic moments exist also for smaller events (Thatcher and Hanks, 1973). The seismic moment was derived from the representations of dislocations as equivalent double couple sources. At very long periods a plane fault appears essentially as a double couple. The seismic moment M_o is defined as the moment of one of these couples and the representation theorem (Maruyama, 1963 ; Burridge and Knopoff, 1964) yields a simple relation between M_o and the average slip $\bar{D}$ on the fault (see eq.1). This is essentially a kinematic description of the source. We would also like to express the seismic moment in terms of dynamic source parameters like the stress drop. This has been done using some specific models of faults: long thin faults (Knopoff, 1958), circular faults (Keilis Borok, 1959) with constant stress drop. The use of these models has led to the remarkable result that the calculated stress drops are almost constant for large earthquakes. A question naturally arises as to the exact meaning of these calculated stress drops. What is the meaning of the estimated stress drop when the actual stress drop is not constant on the fault, the geometry is not a simple circle and the source is complex ? We will show in the following that a general expression for the seismic moment and the stress drop may be found. The relation actually depends on the geometry of the source. For many simple geometries this expression may be given a precise meaning. Finally we obtain expressions for the seismic moment of complex multiple events which indicates that the stress drops determined from simple circular models may substantially understimate the actual stress drop of each of the subearthquakes.

THE GENERAL RELATION OF M_o, STRESS DROP AND GEOMETRY

The seismic moment is the most reliable source parameter determined by seismologists. It is usually defined in the form

$$M_o = \mu \int_S dS\, D(x,y) = \mu \bar{D} S \qquad (1)$$

where μ is the rigidity, D is the offset of the fault, S its total area and $\bar{D}$ the average slip. As defined in (1) the seismic moment is a scalar : it is the moment of one of the couples in the double couple representation of a point dislocation (Burridge and Knopoff, 1964). Implicit in this definition is the assumption that the fault is plane. For curved fault surfaces, or for non-coplanar multiple sources, the seismic moment is defined as a tensor whose components are (Kostrov 1974)

$$M_{ij} = \mu \int_S (D_i(\vec{r})\, n_j + D_j(\vec{r})\, n_i) dS \qquad (2)$$

where $\vec{r}$ is the position vector on the fault, $\hat{n} = (n_x, n_y, n_z)$ is the unit normal and $\vec{D}$ is the slip vector at the point $\vec{r}$. In (2) it was assumed that the faults do not open (i.e. $\vec{D}.\hat{n} = 0$). The surface S may now be curved or a collection of smaller faults in the case of multiple earthquakes. At long periods the excitation of body waves, surface waves and free oscillations is directly proportional to the moment tensor components M_{ij} (Gilbert and Dziewonski 1975).

The definitions (1) or (2) express the seismic moment in terms of the slip at the fault. From observations of the seismic moment and the surface area S, the average slip $\bar{D}$ has been determined for many large earthquakes. In some cases $\bar{D}$ has also been determined directly form geodetic observations. Stress drop $\Delta\sigma$ is usually determined by assuming that the fault was circular and the stress drop was constant on the fault. The stress drops thus obtained are remarkably constant and limited for most events to the range 10 to 100 bars (Aki, 1972, Kanamori and Anderson, 1975 and Kanamori, 1977). The physical significance of this stress drop in terms of the fault geometry and the actual, probably non uniform stress drop on the fault does not seem to have been explored. In the following we shall derive a relation between the seismic moment and stress drop that is valid for general source geometries and stress drop variations.

The basis of this work is Betti's reciprocity theorem for static elastic fields. Let us consider two general elastic static states of the same elastic, possibly non uniform body of volume V and surface S. The surface S includes the external surface S_∞, which will be taken at infinity and a collection of internal surfaces S^i which we shall eventually close to form cracks. A first

elastic state of displacement u_i and stresses σ_{ij} and a second one of displacements v_i and stresses τ_{ij} are considered. We shall assume that there are no body forces in V. Then the elastic reciprocity theorem may be written in the form

$$\int_S \tau_{ij}\, u_i\, n_j\, dS = \int_S \sigma_{ij}\, v_i\, n_j\, dS \qquad (3)$$

where n_i are the components of the external normal to S. Let us now close the internal surfaces S^i to form cracks. We choose on each of these cracks a negative and a positive side. We define the normal to the negative side as the reference normal. Then the identity (3) may be rewritten as :

$$\int_{S_\infty} \tau_{ij}\, u_i\, n_j\, dS + \int_\Sigma \tau_{ij}\, \Delta u_i\, n_j\, dS = \int_{S_\infty} \sigma_{ij}\, v_i n_j dS + \int_\Sigma \sigma_{ij}\, \Delta v_i\, n_j\, dS \qquad (4)$$

where Σ is the collection of all internal surfaces S^i, $\Delta v_i = v_i^+ - v_i^-$ and $\Delta u_i = u_i^+ - u_i^-$ are the slips at the cracks for each elastic field. We let now the surface S_∞ tend to infinity. If we take zero stress as the reference stress level, the integrals on S_∞ will be finite in general. But the absolute stress level does not affect seismic radiation, only stress changes do. For this reason we take the initial stress σ^o_{ij} or τ^o_{ij} before the earthquake as the reference stress. In this case τ_{ij} and σ_{ij} are the stress changes due to slip at the faults. They, and the displacements fields u_i, v_i, decrease like R^{-2} at infinity and the integrals on S_∞ may be dropped from (4). Let us remark that, as noticed by Savage (1969), in the earth σ^o_{ij} and τ^o_{ij} are of internal origin so that the elastic reciprocity theorem (3) does not apply to them. A second remark is that S_∞ may be taken as the surface of the Earth. For shallow earthquakes this may require modifications of the fields v_i, τ_{ij} used below.

Let us derive now an expression for the seismic moment (2) in terms of the stress drops. For that purpose we take u_i, σ_{ij} to be the actual elastic field due to an earthquake, then :

$$\begin{aligned} \Delta u_i &= D_i \\ \sigma_{ij} n_j &= \Delta\sigma_i \end{aligned} \qquad \text{on } S^i \qquad (5)$$

where $\bar{D}$ is the slip and $\overline{\Delta\sigma}$ the stress drop at the fault S^i. For v_i, τ_{ij} we take the following boundary conditions

$$\Delta v_i = E_i^{\ell} \qquad \text{on } S^i \qquad (6)$$

$$\tau_{ij} n_j = \mu n_{\ell}\, \delta_{ik}$$

where E_i^{ℓ} will be the slip determined from the solution of the crack problem with stress drop $\mu\, n_{\ell} \delta_{ik}$ and δ_{ik} is Kröneker's symbol. Replacing (5) and (6) in the reprocity theorem (4) we find

$$\mu \int_{\Sigma} n_{\ell} D_k \, dS = \int_{\Sigma} \Delta\sigma_i \, E_i^{\ell} \, dS$$

Repeating the process with another solution of (6), where $\tau_{ij} n_j = \mu\, n_k \delta_{i\ell}$ adding we finally find

$$M_{\ell k} = \int_{\Sigma} \Delta\sigma_i (E_i^{\ell} + E_i^{k}) \, dS \qquad (7)$$

Where Σ is the collection of all fault segments S^i. The expression (7) defines the seismic moment tensor in terms of the stress drop $\Delta\sigma_i$ in a completely general case. But this definition uses weighting functions E_i^{ℓ} which themselves have to be obtained from the solution of crack problems. This may seem not to be very useful. However in most problems of interest in seismology the slips E_i^{ℓ} are already known from standard solutions in fracture mechanics. For example, for plane faults the stress drop 6b is constant and solutions are known for many simple fault shapes and for some multiple cracks. We shall use these solutions to define the seismic moment in several seismologically interesting problems.

SIMPLE PLANE FAULTS

For a plane fault the seismic moment tensor may be written in the simpler form

$$M_{\ell k} = M_o(n_\ell \nu_k + n_k \nu_\ell) \qquad (8)$$

where n_ℓ and ν_k are the directions of the fault normal and the slip vector respectively, M_o is given by (1). From (7) we find now that

$$M_o = \int_S \Delta\sigma_i \, E_i dS \qquad (9)$$

where E_i is the slip calculated for a crack of the same shape S as the studied fault, but with a uniform stress drop $\Delta\tau_i = \mu \nu_i$. For simple faults, solutions are known for a multiplicity of source shaı . For most earthquakes the after-shocks areas are approximately elliptical so that approximating the source area by an ellipse appears to be a natural choice. The solution for elliptical faults with uniform stress drop may be derived from Eshelby's (1957) study of ellipsoidal inclusions. The slip is found as

$$E_i = A_i \, W\left(1 - \frac{x^2}{L^2} - \frac{y^2}{W^2}\right)^{1/2} \qquad (10)$$

where we choose the axis x in the semi major axis direction ; L and W are the semi major and semi minor axii, respectively. The non dimensional constants A_i are a function of the ellipticity $\varepsilon = W/L$ and are shown in fig 1 for longitudinal A_x or transversal A_y slip. The constants A_i are very smoothly varying functions of the ellipticity which vary at most by a factor of about 2. In the case of a circular fault L = W = R, and (10) reduces to the well-known result

$$E_i = \frac{24}{7\pi} R \, (1 - r^2/R^2)^{1/2} \qquad (11)$$

where r is the radial variable (Eshelby, 1957 ; Keilis Borok, 1959).

The expression (9) is valid for a general stress drop distribution on the fault and yields a precise relation between seismic moment and stress drop for elliptical faults. We may give it a more conventional form.

$$M_o = C_i \langle\Delta\sigma_i\rangle \, W \, S \qquad (12)$$

where now $C_i = \frac{2}{3} A_i$ is a non dimensional constant which depends on the direction of slip at the fault and the ellipticity of the fault. It varies from 0.728 for circular faults to a maximum of

1.33 for longitudinal slip on a very long fault. The estimated stress drop

$$\langle \Delta\sigma \rangle = \frac{3}{2S} \int_{\Sigma} \Delta\sigma \left(1 - \frac{x^2}{L^2} - \frac{y^2}{W^2}\right)^{1/2} dS \qquad (13)$$

is an average of the stress drop with a weighting function that emphasizes the stress drop near the center of the fault. The fault appears to be stiffer near the ends and softer towards the center. If stress drop is uniform then $\langle \Delta\sigma \rangle \equiv \Delta\sigma$. For general non-uniform stress drop $\langle \Delta\sigma \rangle$ will in general differ from $\overline{\Delta\sigma}$, the average stress drop. This difference will be large if stress drops are large near the edges of the fault. This is the case for instance of some dislocation models (like Haskell's) where stress drop is infinite at the edges but the seismic moment is finite. In actual earthquakes little is known about the true stress drops. For bounded stress drops we expect that $\langle \Delta\sigma \rangle$ will not be too different from $\overline{\Delta\sigma}$ since the ellipsoidal averaging function in (13) is very smooth and close to one over much of the fault plane.

In many observational work, however, the shape of the fault is not well known ; in those cases a usual assumption is that (12) may be approximated by

$$M_o = C' \langle \Delta\sigma \rangle S^{3/2} \qquad (14)$$

where for the elliptical fault $C' = C_i (\pi W/L)^{1/2}$. But in most practical work the circular fault approximation $C' = 16/7\, \pi^{3/2} = .41$ is used. The effect in general is to underestimate the true average stress drop.

Several recent works have shown that many large seismic events are in fact complex suites of smaller events clustered in the source region. Observations of this sort have been reported, among others, by Wyss and Brune (1968), Wu and Kanamori (1975), Fukao and Furumoto (1975), Aki (personal communication, 1977). Evidence for multiple events usually comes from the complexity of the P-waves radiated by most large events. Direct evidence of the segmentation of faults has been reported by Spottinswoode and Mc Garr (1975) for shocks in deep gold mines of South Africa. Complex events have also been found in the numerical modelling of faulting with variable strength by Das (1976) and Mikumo and Miyatake (1977). The complex events may be of two types: first, connected sources where fault patches remain unbroken (Mikumo et al., 1977) or isolated multiple faults (Aki and Das, 1977). Second, simple fault but very inhomogeneous stress drop (Kanamori, 1978 this conference). The latter case may be analysed with the solution of the previous section. In this case $\Delta\sigma$ in (13) will be very non-uniform and $\langle\Delta\sigma\rangle$ is the average stress drop estimated by the usual methods.Here we shall concentrate on the case of a fault that is made of several isolated sub-faults.

In order to discuss the expression for the seismic moment we shall make several additional simplifying assumptions. First, we shall assume that all the multiple faults are coplanar. This assumption is also generally made in observational work. If this is not assumed the moment tensor may not be written in terms of a single scalar M_o as in (8). For multiple coplanar faults the seismic moment is again given by the expression (9). But now E_i is the slip produced by a constant stress drop $\Delta\tau_i = \mu\ \nu_i$ in each of the subfaults of the complex event. This slip is necessarly smaller than the slip calculated for a single source of the same total area. Thus, for a given stress drop on the subfaults, the seismic moment will be larger for a single fault than for a multiple fault of the same equivalent area. When using (12) or (15) to estimate $\langle\Delta\sigma\rangle$ in the usual way, the actual stress drop in every subfault will be underestimated. To illustrate this problem we shall assume that the actual event is made of N smaller fractures of area S^i. We shall also assume that each of these subfaults is of circular shape. E_i then the slip in a coplanar array of circular cracks with identical, uniform stress drop. The solution of this problem depends on the actual separation between faults and their spatial arrangement. For simplicity we shall neglect the interaction between the faults and assume that each of them is idependent. The slip in E^i in each fault is then of the form (11) with $R = R_i$ the radius of each of the subfaults.

The seismic moment for this complex event may be written as

$$M_o = \frac{24}{7\pi} \sum_i R_i \int_{S_i} \Delta\sigma \, (1 - r^2/R_i^2)^{1/2} dS \qquad (15)$$

where the sum is over all the subfaults. We may rewrite this equation in the slightly different form

$$M_o = \frac{16}{7\pi} \sum_i R_i \, \langle \Delta\sigma \rangle_i \, S_i \qquad (16)$$

where $\langle \Delta\sigma \rangle_i$ is a weighted average of the form (13) of the stress drop in each of the subfaults. Compared to the simple crack model of stress drop $\langle \Delta\sigma \rangle = \langle \Delta\sigma \rangle_i$ and total surface $S = \sum_i S_i$, the seismic moment (16) is smaller by a factor of R_i/R than that of the simple fault. We may see this more clearly if we assume that all the subfaults are identical, with radius $R_i = \rho$, stress drop $\langle \Delta\sigma \rangle = \hat{\sigma}$ and that the total area $S = \sum_i S_i$. In this case the seismic moment (16) has the simpler expression

$$M_o = \frac{16}{7\pi} \rho \, \hat{\sigma} \, S \qquad (17)$$

We may compare this to the seismic moment of a simple circular fault of area S (and radius R) with apparent stress drop $\langle \Delta\sigma \rangle$

$$M_o = \frac{16}{7\pi} R \, \langle \Delta\sigma \rangle \, S \qquad (18)$$

For the same stress drop on the broken patches of the fault $\hat{\sigma} = \langle \Delta\sigma \rangle'$, the seismic moments differ by the ratio of the radius of the total fault to the radius of a typical subevent. Similar results would be obtained for faults where isolated patches of the fault remain unbroken. It appears as if a simple fault is much more efficient in generating low frequency waves than a complex source. In the latter case the fault is effectively "pinned down" at the unbroken patches. But the problem of interest in seismology is usually an inverse one : given a certain seismic moment and a certain source area, what is the stress drop ? This may be seen again from (17) and (18). Using (18) for a complex event instead of (17) will yield a stress drop estimate $\langle \Delta\sigma \rangle$ which is smaller than $\hat{\sigma}$. In conclusion, the stress drops obtained for complex events may severely underestimate the actual stress drops in the broken patches of the fault.

We may view this result in still another perspective. We think now that the fault is simple but that in the unbroken segments a stress concentration (a stress "rise") exists that is just enough to make the slip in this region equal to zero. Then the stress change in the fault has stress drops in the subfaults and stress rises in the unbroken patches. The average stress drop over the fault will be lower than the stress drop on the subfaults: it is biased to low stress drops because of the stress concentrations in the unbroken sections.

THE SPECTRA OF SIMPLE AND COMPLEX EVENTS

At low frequencies the radiation of body waves is directly proportional to the seismic moment. At high frequencies, however, the radiation depends on the details of rupture. Thus, it is very likely that the spectra of complex events be quite different from that of simple events. This has already been observed in two-dimensional models by Das (1976). Here we shall study this problem with the help of the array of non interacting circular faults used in the previous section. The radiation from a simple circular fault was studied by Madariaga (1976) with numerical methods. In fig.2, we show a typical far field spectra and pulses in a direction that makes an angle $\theta = 60°$ with the normal to the fault. In this example the fault grows self-similarly with a constant rupture velocity $v_R = .9\ v_S$ until it stops abruptly at a These spectra show the typical constant low frequency level proportional to the seismic moment. At the high frequency end, on the other hand, the spectra decay like f^{-2}, where f is the frequency in Hertz. There are also intermediate frequency trends which depend on the details of the rupture process. In order to discuss the radiation it is convenient to write the far field displacement spectrum in the form

$$|u(\vec{R},f)| = \frac{1}{4\pi\rho c^3} \frac{\mathcal{R}_{\theta\phi}}{R} |M_o(\hat{R},f)| \qquad (19)$$

where ρ is the density, $\vec{R}$ is the position of the observation point, $R = |\vec{R}|$, $\mathcal{R}_{\theta\phi}$ is the P or S radiation pattern, and c is the velocity of P or S waves. $M_o(\hat{R},f)$ is then a moment rate spectral density. It depends on $\hat{R}$ the direction of radiation, so that at high frequencies a fault does not really radiate like a pure double couple. This is due to effects like focusing and directivity. At low frequencies $M_o(\hat{R},f) \to M_o$, the static seismic moment. At high frequencies the spectrum decays like f^{-2} and the moment rate spectrum has the general form

$$|M_o(\hat{R},f)| \simeq m(\hat{R})\ \Delta\sigma\ a\ v_S^2\ f^{-2} \qquad (20)$$

where $\Delta\sigma$ is the stress drop, a is the source radius, v_S is the shear wave velocity and $m(\hat{R})$ is a function that depends on the direction of radiation. From Madariaga(1977) we find that it varies strongly with rupture velocity v_R. For instance it is m=.25 at v_R=.9 v_S and m=.17 at v_R=.6v_S. These are values of m at $\theta \simeq 60°$, a mean value of the angle of radiation with respect to the normal to the fault. Let us remark that the seismic moment of a circular fault is

$$M_o = 16/7\ \Delta\sigma\ a^3 \qquad (21)$$

Thus m is a measure of the corner frequency, $f_o = (7m/16)^{1/2} V_s/a$. The most important consequence of (20) is that high frequencies scale like a linear dimension of the fault while the low frequencies scale like the cube of this linear dimension.

Consider now, as in the previous section, an array of N circular cracks of identical radii ρ . We shall assume for simplicity that the rupture velocities in every fault is the same. Then the total radiation from the collection of faults may be obtained superposing the high frequency energies. The energy is proportional to $|M_o|^2$ so that the total high frequency seismic moment of a complex fault is

$$M_o(f) = m\, \overline{\Delta\sigma}\, \rho N\, v_S^2\, f^{-2} \qquad (23)$$

where $\overline{\Delta\sigma} = N^{-1} \Sigma\, \Delta\sigma_i^2$ is the mean square stress drop over the collection of faults. If the collection of circular faults covers an area of approximate radius R, then $N\rho$ will not be very different from R. As a consequence, the high frequency radiation from a simple and a complex fault of the same stress drop $\overline{\Delta\sigma}$ and total radius R are approximately the same. This is a surprising result because the usual assumption is that complex faulting enriches the high frequency radiation. Instead we find that what really happens is that the low frequency radiation is diminished while the high frequency part remains approximately the same (fig.3a). This effect may in fact be observed in the non normalized spectra of Das and Aki(1977) for two dimensional faults with barriers.

In general, spectra are normalized by the seismic moment at long periods. In this case, as seen in fig 3b, a multiple event appears enriched in high frequency energy with respect to a simple event of the same seismic moment. Thus complex events scale differently than simple events and stress drop estimates from the high frequency trend of the spectra should yield a stress drop which is much closer to the actual stress drop on the subfaults. The scaling of corner frequencies is more complex. As discussed by Das (1976) complex events usually show intermediate frequency trends and at least two more or less clear corner frequencies. One of the corner frequencies would be associated naturally with the characteristic size of the subevents (f_o^1 in fig.3) and the other f_o^2 with the total duration of the rupture process. If the subfaults occur simultaneously, or separated by times comparable to the time of propagation of a wave accross the total source area, the corner frequency f_o^2 will reflect the total size of the fault as for simple source models. If the multiple rupture develops more slowly then f_o^2 may be much lower and the low frequency spectrum may appear to grow at very long periods. Intermediate frequency trends should appear between these two corner frequencies. This

region of the spectrum contains all the information about the time development of the complex rupture.

DISCUSSION

The seismic moment was introduced as a kinematic parameter describing the radiation of low frequency seismic waves by a dislocation source. From the representation theorem it was found that it is a measure of the average slip discontinuity on the fault multiplied by the source area. In this sense it is a measure of the distortion of the source region due to faulting. We have shown in this paper that it may also be given a dynamic interpretation in terms of the stress drop of the fractured areas at the source. The relation is very general but not as simple as that between moment and slip at the source. In fact we showed that the geometry of the source affects the seismic moment as much as the stress drop. In particular, for multiple sources the seismic moment is smaller than for simple faults of the same area and stress drop. A simple interpretation of this result is that the unbroken sections of the fault pin down the fault and reduce the distorsion of the source region, reducing in consequence the seismic moment which measures this distorsion. Consequences for the inverse problem are severe : the stress drops for complex events, estimated by the usual circular fault formula (Keilis-Borok, 1959), are only a fraction of the actual stress drop on the broken sections of the faults. The actual and estimated stress drop are in the ratio of the total source size to the size of a typical subevent.

Observations for many large events have led to the conclusion that the stress drops are almost constant ranging from about 10 to a 100 bars (Aki, 1972, Kanamori and Anderson ,1975). These stress drops have been invariably obtained from formulas for simple smooth faults. There is abundant evidence that many of these large events are not really simple events but complex multiple events. In this case our previous discussion implies that the actual stress drops are very probably larger that those estimated from observations. What is the meaning then of the almost constancy of stress drops ? If we admit that large events, for which the seismic moment has been estimated, include both simple and complex faults we are led to the conclusion that what is really constant is the product of actual stress drop by the ratio of subfault size to total size. It appears then that the more complex the source area, the larger is the stress drop on each of the subfaults. This

in fact may not be surprising if we consider that the resistance of a source with closely spaced barriers is certainly larger than that of a fault which has a relatively uniform strength. We may interpret observations to imply that the actual stress drops are inversely proportional to the distance between high stress patches on the fault.

ACKNOWLEDGEMENTS

This work was done at the Laboratoire d'Etude Géophysique de Structures Profondes associé au C.N.R.S. n° 195.

Discussions and preprints by K. Aki were very usefull. Discussions with A.F.Dahlen helped to clarify several points.

REFERENCES

AKI, K.(1966). Generation and propagation of G waves from the Niigata earthquake of June 16, 1964, 2, Estimation of earthquake moment, released energy, and stress-strain drop from G-waves spectrum, Bull.Earthquake Res. Inst., Tokyo Univ.,44, 73-88.

AKI, K. (1967). Scaling law of seismic spectrum, J.Geophys. Res., 72, 1217-1231.

AKI,K. (1972) Earthquake Mechanism, Tectonophysics, 13, 423-446.

BURRIDGE, R. and L.KNOPOFF (1964). Body force equivalents for seismic dislocations, Bull.Seism. Soc. Am.,54, 1875-1888.

DAS, S. (1976); A numerical study of rupture propagation and earthquake source mechanism Thesis. Massachusetts Institute of Technology, Cambridge. Mass.

DAS, S. and K. AKI (1977) Fault plane with barriers: A versatile Earthquake model, J. Geophys. Res., 82, 5658-5670.

ES HELBY J.D. (1957). The determination of the elastic field of an ellipsoidal inclusion, and related problems. Proc. Roy. Soc.,A 241, 376-391.

FUKAO, Y. and M.FURUMOTO (1975). Foreshocks and multiple shocks of large earthquakes. Phys. Earth Planet. Interiors, 7, 143-153.

KANAMORI, H. (1977). The energy release in great earthquakes J.Geophys. Res, 82, 2981-2987.

KANAMORI, H. and D.L. ANDERSON (1975). Theoretical basis of some empirical relations in Seismology Bull. Seism. Soc. Am.,65, 1073-1095.

KEILIS-BOROK, V.I. (1959). On the estimation of the displacement in an earthquake source and of source dimensions Ann. Geofis.,12, 205-214.

KNOPOFF, L. (1958). Energy release in earthquakes Geophys. J. R. astr. Soc., 1, 44-52.

KOSTROV. B.V. (1974). Seismic moment and energy of earthquakes, and seismic flow of rock, Izv. Fizika Zemli,1, 23-40.

MADARIAGA, R. (1976). Dynamics of an expanding circular fault.
Bull. Seism. Soc. Am., 65, 163-182.

MADARIAGA, R. (1977). High frequency radiation from crack (stress drop) models of earthquake faulting
Geophys. J.R. astr. Soc., 51, in press.

MIKUMO, T. and T.MIYATAKE (1977). Dynamical rupture process on a three-dimensional fault with non uniform friction and near-field seismic waves.
Preprint. Disaster prevention Research Institute, Kyoto University, Uji, Kyoto, Japan.

SAVAGE, J.C. (1969). Steketee's paradox
Bull. Seismol Soc. Am., 59, 381-384.

SPOTTINSWOODE, S.M. and A.Mc GARR (1975). Source parameters of tremors in a deep-level gold mine.
Bull. Seismol. Soc. Am., 65, 93-112.

THATCHER, W. and T.HANKS (1973). Source parameters of southern California earthquakes
J.Geophys. Res, 78, 8547-8576.

WU, F.T and H.KANAMORI (1975). Source mechanism of february 4, 1965, Rat Islands Earthquake
J.Geophys. Res., 78, 6082-6092.

WYSS, M. and J.B. BRUNE (1967). The Alaskan earthquake of March 28 1964, a complex multiple event
Bull. Seismol. Soc. Am., 57, 1017-1023

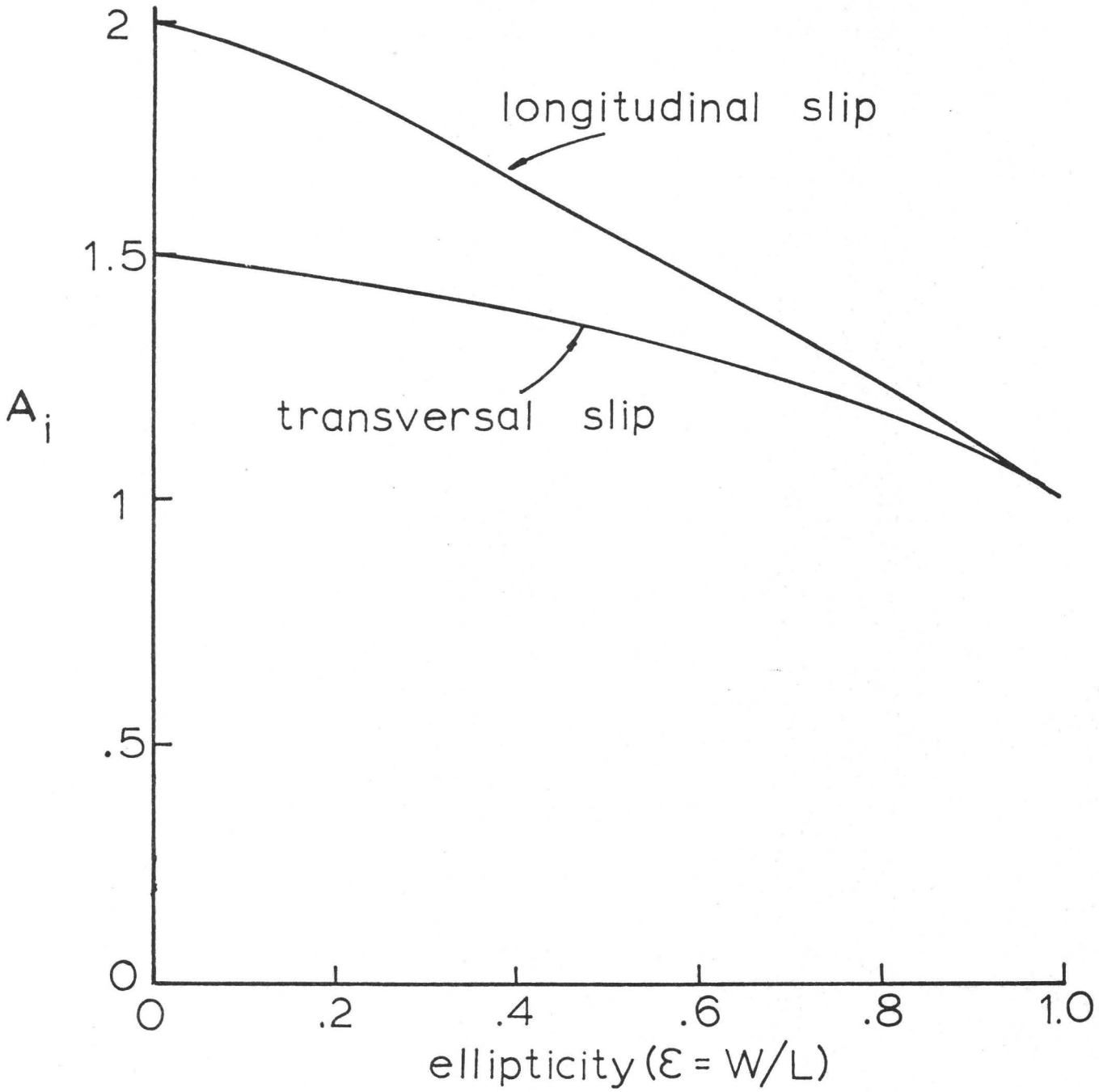
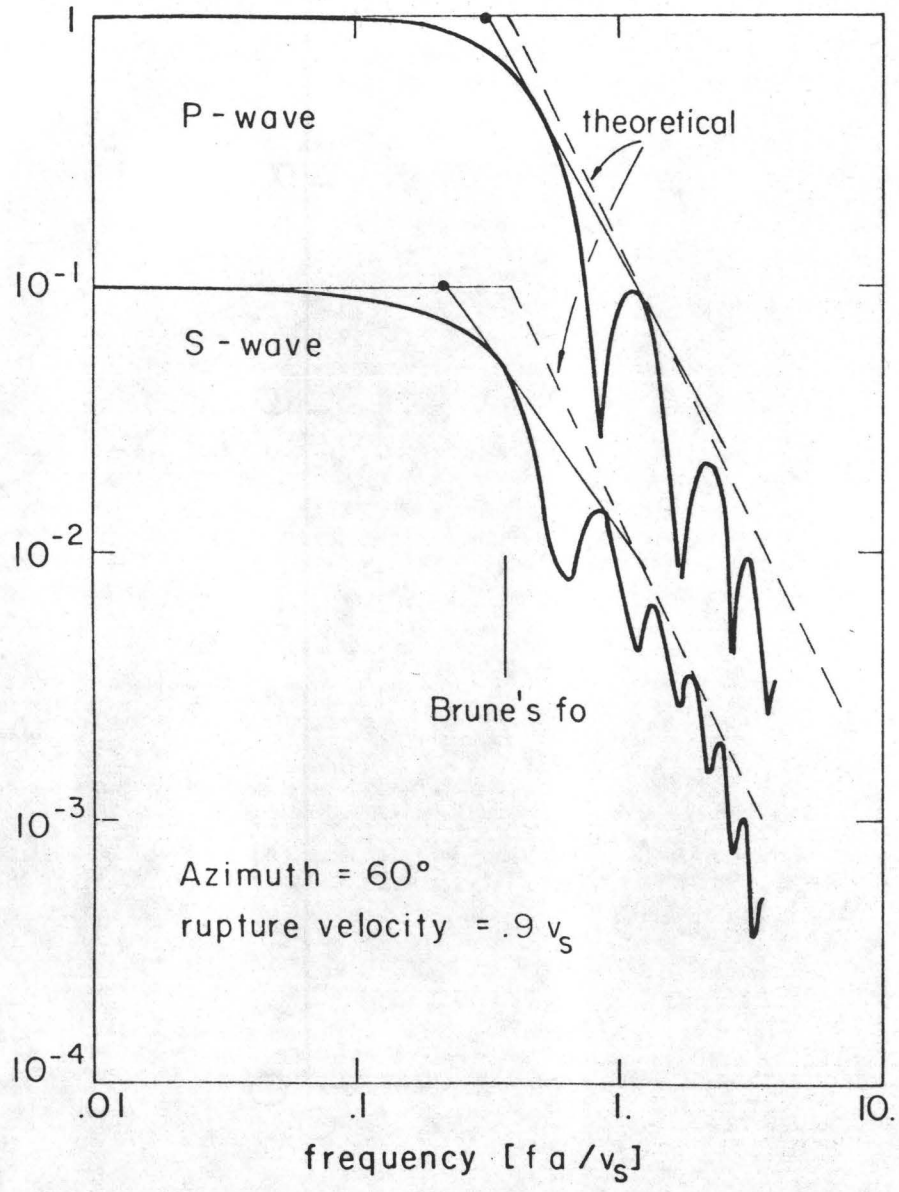

2
1.5
1
.5
0
A_i
longitudinal slip
transversal slip
0
.2
.4
.6
.8
1.0
ellipticity (ε = W/L)

Figure 1


1
10^{-1}
10^{-2}
10^{-3}
10^{-4}
P - wave
theoretical
S - wave
Brune's fo
Azimuth = 60°
rupture velocity = .9 v_s
.01
.1
1.
10.
frequency [f a / v_s]

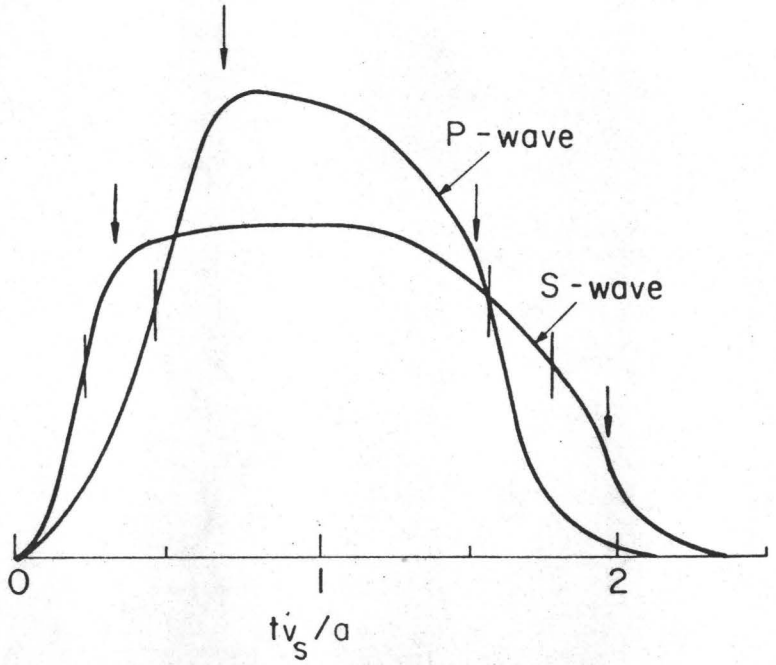
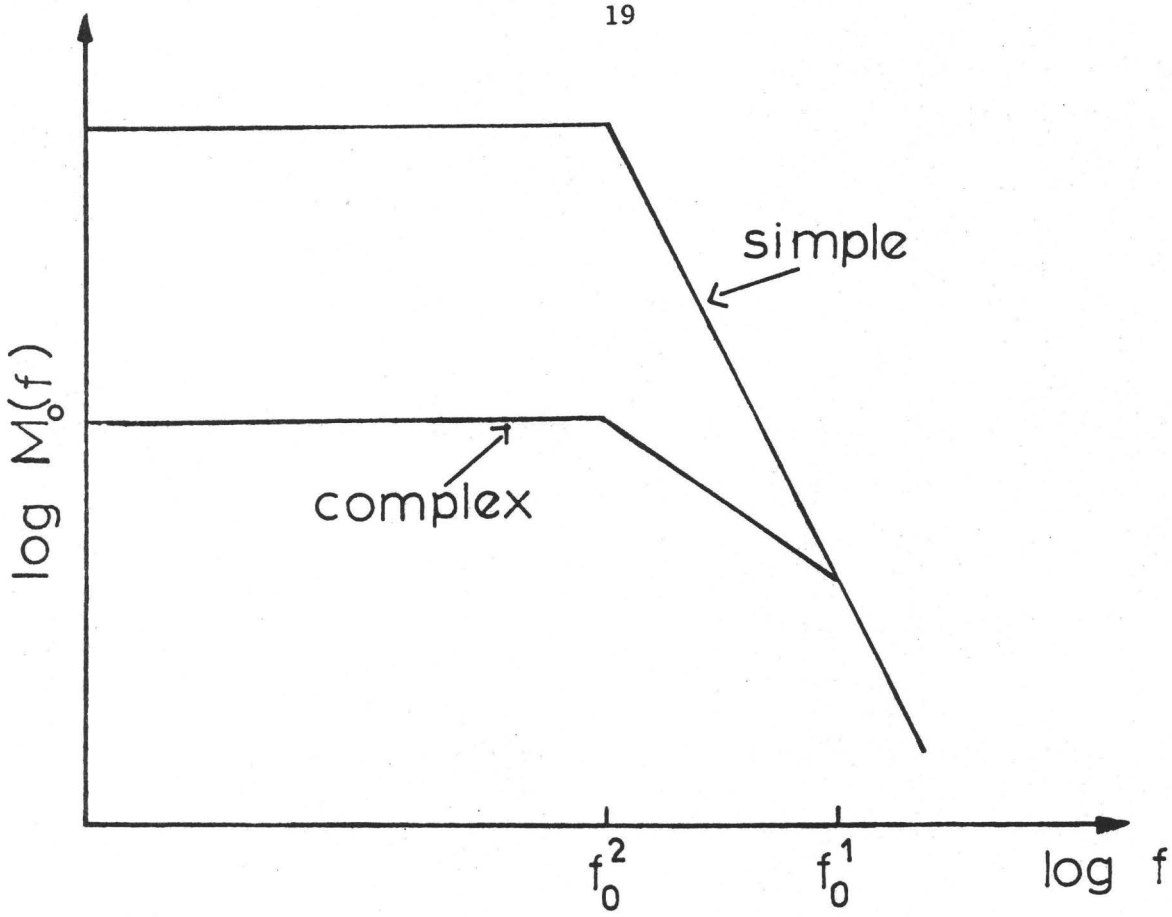
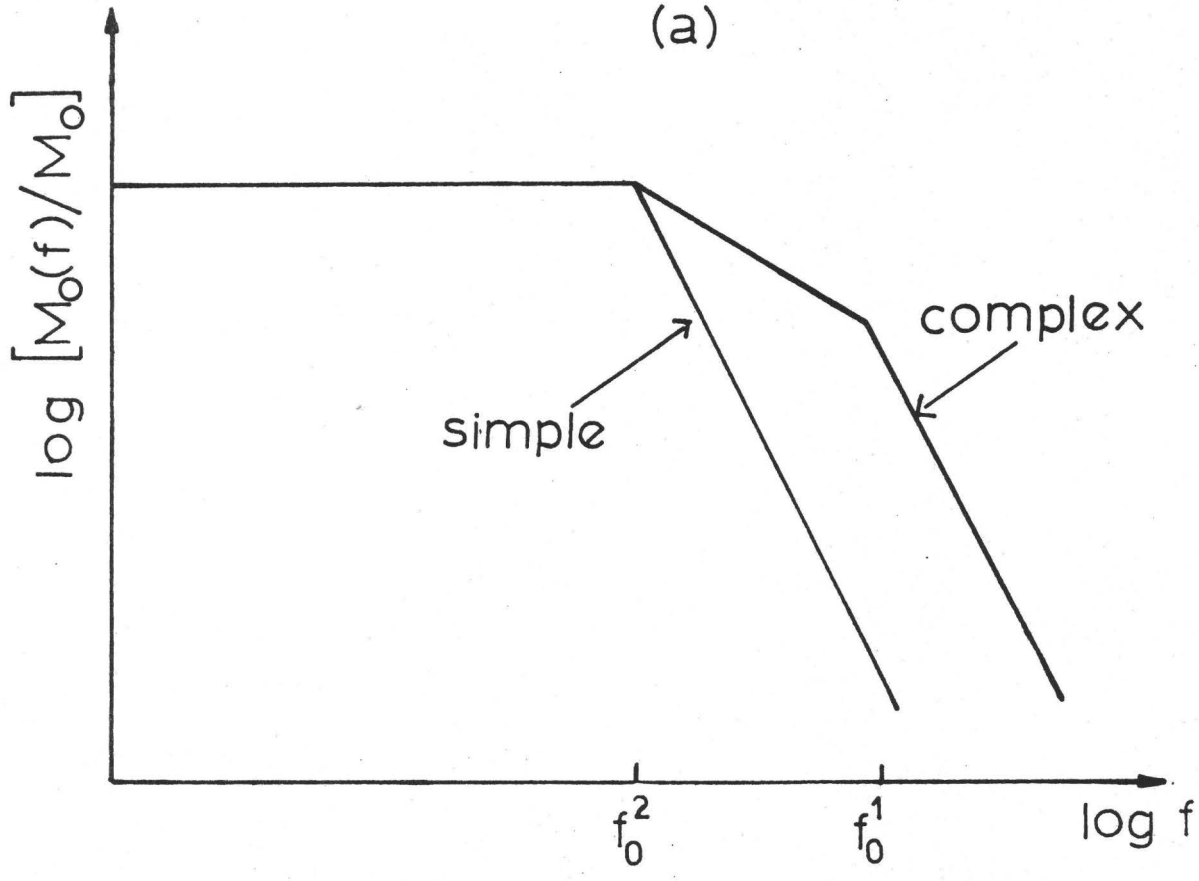

P - wave
S - wave
0
1
2
tv_s/a
far - field source function

Figure 2.


simple
complex
log $M_o(f)$
f_0^2
f_0^1
log f

(a)


log $[M_o(f)/M_o]$
complex
simple
f_0^2
f_0^1
log f

(b)

Figure 3

LARGE SCALE QUASI-STATIC FAULT MODELS

Gerald M. Mavko

Department of Geophysics

Stanford University

Stanford, California 94305

Most theories concerning earthquakes are based on elastic rebound- the idea that elastic strain energy is gradually stored in the earth and is abruptly released during episodes of failure known as earthquakes. Comparisons of the accumulation of deformation at the earth's surface before large earthquakes with the instantaneous deformation during earthquakes show that they often approximately cancel. This led to the idea of a rebound [Reid, 1910]. In the context of plate tectonics, the process of strain build up and release at major plate boundaries repeats itself in a roughly cyclic fashion. The driving mechanism or source of energy for the earthquake cycle is the relative plate motion across the common boundary. Whether or not strain accumulates, and the way it is released depends on the nature of slip on the boundary.

In a very general way we can describe a major cycle in terms of four time phases relative to the earthquake [Mescherikov, 1968; Lensen, 1970; Scholz, 1972]: In the strain accumulation phase, the average fault slip on the plate boundary is slower than the long term average plate rate far from the fault. A simple geometric deficiency of slip accumulates causing strain energy to be stored in the plate. The coseismic phase is the period of several seconds to minutes during which rapid fault slip occurs, generating seismic waves. All or part of the slip deficiency is recovered; stored elastic strain energy is converted into heat and waves (kinetic energy). The coseismic phase may or may not be preceded by a preseismic phase. This is a period of incipient strain release characterized by higher strain rates than occur during the strain accumulation phase. Rapid changes of any sort during this period might

be interpreted as precursors. Finally the postseismic phase is a period of transient readjustment following the rapid earthquake movement. This may take place through aseismic creep, aftershocks, or viscoelastic relaxation.

In this paper we review some recent models for the earthquake cycle. Our emphasis will be on the observable, quasi-static deformation associated with the preseismic, postseismic, and strain accumulation phases of major earthquakes at plate boundaries. To begin, we summarize some observations of transient deformation ranging over periods of several minutes to several years. These define a relaxation spectrum that is much broader than might be expected from simple elastic rebound, and may, therefore, contain information about the physical processes important to earthquakes. A number of possible relaxation mechanisms are also presented. Next, we discuss models for postseismic deformation. These are perhaps the most useful of the quasistatic models for interpreting rheology since postseismic data is more abundant and reliable than, for example, preseismic transients. Finally we focus on strain accumulation and the earthquake cycle as a whole. This is the most critical since it encompasses predictive information, yet is most difficult to measure without knowing when and where events will occur. An appendix is added to summarize some early models for coseismic deformation. Although these largely neglect the aseismic portion of the cycle they often form the computational basis for more recent models. In addition, the nature of seismic strain release obviously sets the conditions for postseismic transient deformation.

THE NATURE OF THE OBSERVATIONS

In a strictly elastic earth, complete elastic rebound would take place in a few seconds, with the characteristic time of strain release determined by the seismic source rise time, fault dimensions, and rupture velocity. The only slow deformation would be the accumulation of tectonic strain. However, pre- and postseismic transients are observed which indicate a much broader relaxation spectrum. For example, following the 1966 Parkfield, California, earthquake (M = 5.5; right-lateral strike slip) near surface fault creep continued, at a decaying rate, for several years. Figure 1 [Smith and Wyss, 1968; Scholz, et al., 1969] shows the postseismic fault slip inferred from five small-scale geodetic stations established in the epicentral region. Although little or no surface breakage occurred during the main event, as much as 25 cm of slip accumulated in the 3 years immediately thereafter. In addition, road damage occurring within several years before, and en-echelon cracks formed within a month before the earthquake [Allen and Smith, 1966] suggest a preseismic transient.

Even the great 1906 San Francisco earthquake, which led H.F. Reid to propose the elastic rebound mechanism, was followed by transient deformation. Thatcher [1975a] suggests that substantial postseismic crustal strains, continuing for at least 30 years following the earthquake, can be inferred from geodetic resurveys since 1906. These strains can be explained (though not uniquely) by $\sim$4 m of aseismic fault slip from 10 to 30 km depth, without additional surface slip. Thatcher [1975a] also suggests anomalously rapid strain accumulation during the 50 years prior to 1906, although the evidence is weak.

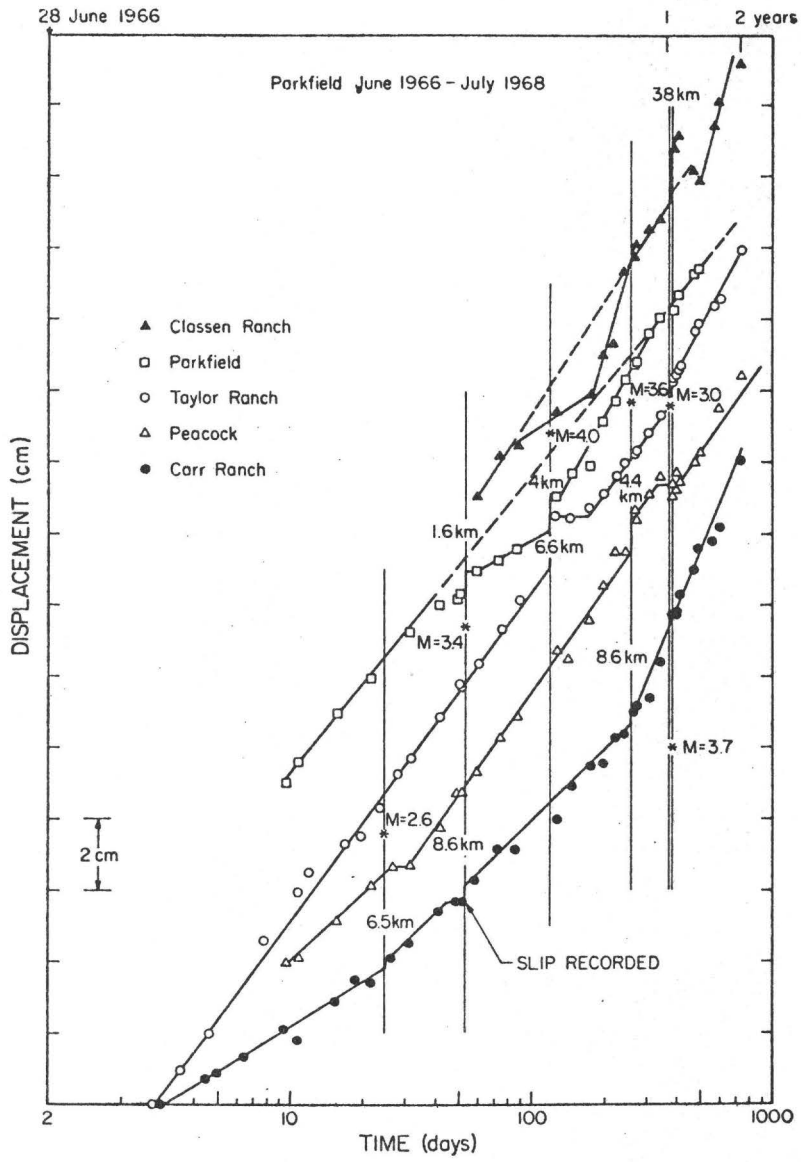
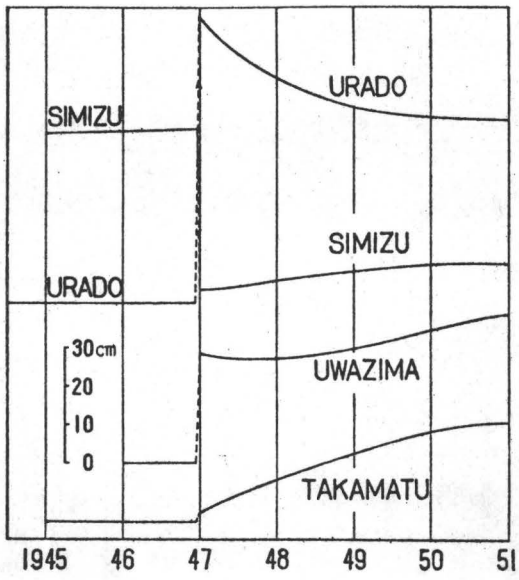
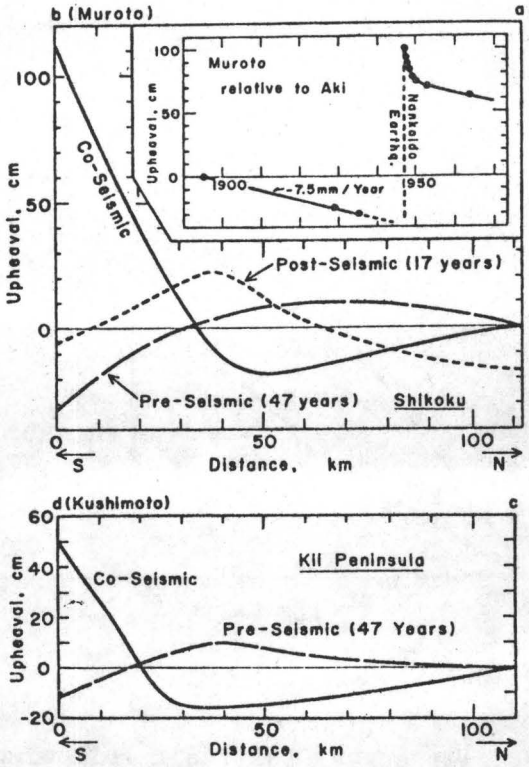


Figure 1. Cumulative slip across the fault following the Parkfield earthquake measured with five small-scale geodetic networks. The ordinate intercepts are arranged arbitrarily. The mainshock was at t = 1. [Scholz, et al., 1969]



(a)



(b)

Figure 2. Deformation accompanying the 1946 Nankaido earthquake.
(a) Smoothed sea level at four tide guage stations [Matuzawa, 1964].
(b) Pre-, co-, and postseismic vertical deformation along profiles roughly normal to the fault trace [Kanamori, 1973].

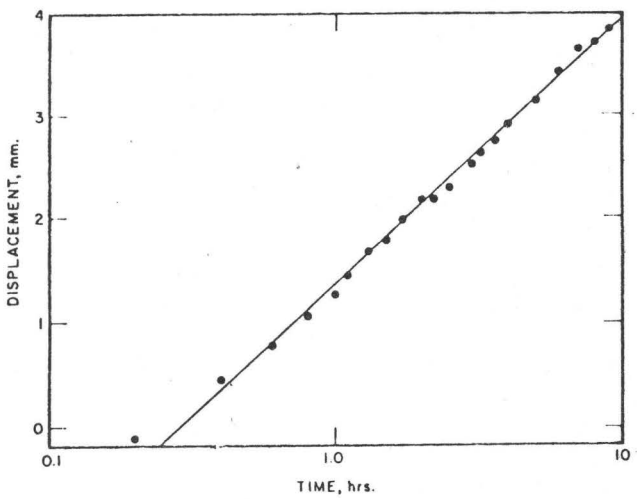


Figure 3. Fault slippage following the magnitude 5.0 earthquake of September 6, 1966 at Matsushiro. [Nakamura and Tsuneishi, 1967; Scholz, 1972].

Perhaps the most spectacular example of measureable postseismic deformation was observed following the 1946 Nankaido, Japan, earthquake (M = 8.2; thrust type) where upheavals of as much as 2 m occurred over a 1 to 3 year period. Figure 2 [Matuzawa, 1964; Kanamori, 1973] shows the rather complicated nature, in space and time, of the vertical displacement. The transient mechanism has been interpeted either as aseismic fault slip at depth [Fitch and Scholz, 1971] or viscoelastic rebound of the asthenosphere without fault slip [Nur and Mavko, 1974].

A similar analysis of postseismic uplift following the 1964 Alaskan (thrust type) earthquake is reported by Brown, et al. [1977]. In this case the transient decay lasted about 1 to 8 years depending on distance from the fault trace, and was attributed to fault slip.

Particularly short-lived transients have also been observed. Figure 3 shows fault slip lasting only several hours, recorded after a Matsushiro shock on September 6, 1966 [Nakamura and Tsuneishi, 1967; Scholz, 1972]. A precursory aseismic slip with time constant of 300 to 600 sec, starting about 1000 sec before the main shock of the 1960 Chilean earthquake has been inferred from long-period surface waves and body waves [Kanamori and Cipar, 1974] and from free oscillations [Kanamori and Anderson, 1975].

Further examples of transient deformation are reviewed by Scholz [1972], Kanamori [1973] and Dunbar [1977].

THE VISCOUS ELEMENT

These observations suggest an obvious deviation from simple elasticity. Episodes of accelerated strain or creep that occur without an earthquake require a localized viscous instability which can initiate and arrest each strain event. Furthermore, seismic precursors and postseismic transients imply a strain or strain rate-dependent damping associated with the otherwise brittle release of energy.

What are these viscous elements? A simple mechanical model for the earth's crust and upper mantle, suggested by plate tectonics, consists of a relatively elastic, brittle lithosphere overlying a ductile asthenosphere. Within this framework we can distinguish geometrically three general sources of relaxation:

Relaxation in the asthenosphere

The asthenosphere is characterized by high temperature relaxation mechanisms. Solid mineral grains can flow plastically by atomic diffusion and the motion of lattice dislocations [Gordon, 1965; Weertman and Weertman, 1975; Heard, 1976; Carter, 1976]. This makes the polycrystalline composite fluidlike over long time scales and can account for the large-scale, finite deformation implied by plate motion and the low strength implied by isostatic equilibrium. In addition, enhanced deformation at grain boundaries can occur resulting from dislocation motion and diffusion [Ke, 1947; Zener, 1948; Anderson, 1967] or the viscous flow of melt [Walsh, 1969; Mavko and Nur, 1975]. Other loss mechanisms which are relevant at seismic frequencies include thermoelastic, dislocation damping, point defect diffusion, and grain boundary effects [Anderson, 1967]. We will discuss in more detail the possible role of asthenospheric relaxation

during major earthquake faulting in the next section.

Relaxation in the lithosphere

The lithosphere by definition is a relatively strong, rigid layer with a long term strength. This means that it can resist permanent deformation or plastic flow for millions of years under stress differences of several hundred to a thousand bars, whereas the asthenosphere cannot [Le Pichon, et al., 1973]. This is consistent with analysis of glacial rebound and lithospheric flexure [Walcott, 1973], as well as our concept of continental drift. It is by definition, then, that we can rule out large scale, large strain, plastic flow or solid state creep as an important relaxation mechanism in the lithosphere.

The important question then becomes: how thick is the lithosphere? Or, at least, if we are to construct simple mechanical models for the earthquake cycle, what thickness is appropriate for the elastic layer? Estimates of lithospheric thickness can be made in several ways [Le Pichon, et al., 1973]: the depth of the solidus can be obtained from a computation of the thermal structure; seismic studies give the depth of the top of the low velocity zone; studies of rebound and flexure give an elastic plate thickness. Altough it is not completely obvious that these various depths based on different physical parameters should agree, they are all generally consistent with the existence of a lithosphere about 75 km thick under the deep ocean basins and 110-130 km thick under continental shields. If we consider that the mechanical properties of strength and elastic stiffness are both effected by temperature, in particular by the presence of partial melt, then perhaps the comparison, made over many orders of magnitude of time, makes sense. In contrast,

Anderson [1971] and Hadley and Kanamori [1977] suggest that in parts of southern California the shallow crust is mechanically decoupled from the lower crust, so that the moving surface plate is much thinner than is commonly assumed in plate tectonic theory (as discussed above). Lachenbruch and Sass [1973] suggest a similar decoupling between the shallow crust (15-20 km) around the San Andreas Fault and the more ductile material below in order to explain a low broad heat flow anomaly. However in this case, the crustal plate is also undergoing permanent shear flow, generating heat. We shall see that this uncertainty in plate thickness can greatly affect interpretation of surface strain.

Aside from large scale flow which distinguishes the asthenosphere from the lithosphere, a smaller viscoelastic response to changes in the stress field can occur within the lithosphere. A material is viscoelastic when its response to abrupt changes is initially elastic, followed by a longer term relaxation or flow (not necessarily linear). The relaxed configuration may also be essentially elastic, distinguished from the unrelaxed state only by a smaller effective rigidity. Hence, a viscoelastic lithosphere exhibiting transient relaxation times on the order of several years would look elastic at seismic frequencies as well as over the longer periods of flexure and isostatic rebound.

A number of relaxation mechanisms can be considered to account for a viscoelastic response. Concentrated plastic flow at grain boundaries is reasonable in much of the lower lithosphere (below, say, 20-30 km) where the ratio of absolute temperature T to the melting temperature Tm is greater than one half ($T/Tm > 1/2$). Presumably, motion at grain boundaries could occur while the grains themselves remained essentially

elastic, giving to the polycrystalline composite a long term finite strength, yet a short term viscoelastic strain.

In the shallow crust stress induced viscous shearing and local squirt of pore fluids [Mavko and Nur, 1975, Mavko and Nur 1977b] as well as large scale, regional diffusion [Biot, 1941; Nur and Booker, 1972] can give a time dependent deformation qualitatively similar to a visco-elastic response. The regional diffusion might also be enhanced by dilatancy [Nur, 1973].

Fault Creep

In addition to direct observations of surface fault creep, aseismic fault slip has been invoked at depth in the lithosphere to explain pre- and postseismic surface deformation [Fitch and Scholz, 1971; Thatcher, 1975a; Brown, et al., 1977]. However, very little is known about the detailed stress-strain behavior of the fault zone, and hence the physical mechanism of creep, at any depth. Nason and Weertman [1973] conclude little more than the existence of an upper yield phenomenon from observations of shallow creep events. In the laboratory transient stable sliding sometimes precedes stick slip on frictional surfaces [Scholz, et al., 1969] at conditions corresponding to several kilometers depth. At higher temperatures and pressures Stesky [1974] observes a nonlinear stress-strain rate sliding law similar to that expected for solid-state creep. Laboratory measurements on fault gouge and clay have also been made. The main problem lies in determining what kind of material is representative of a fault zone at depth.

In addition to creep on the primary fault being studied, creep on nearby faults can have an effect on relaxation. Even though the bulk of the crustal material is elastic, slip on secondary faults makes the crust

effectively more compliant. If the slip is creep-like, the change in compliance is gradual, and the overall effect may not be distinguishable from viscoelastic relaxation.

MODELS FOR POSTSEISMIC TRANSIENTS

Despite the wide range of relaxation mechanisms several general observations can be made: For short times and small strains the mechanical behavior of the crust and upper mantle is approximately linear elastic. This holds at least out to periods of many minutes, as demonstrated by wave propagation and free oscillations. As a result when rapid fault slip occurs the initial response of the earth is elastic. The earthquake rupture superimposes a stress perturbation on the preseismic state -- decreasing the stress over much of the fault and increasing it elsewhere. Under this new stress field any number of viscous elements may immediately begin to relax. Those that cause a large enough change to be measurable and which fit within a reasonable time window may turn out to be observed as postseismic transients.

Dip Slip Earthquakes

One major distinction to be made when discussing important mechanisms for dip slip and strike slip faulting is the depth of rupture in the lithosphere. Major thrust type earthquakes at subducting plate margins often rupture through a substantial fraction of the lithosphere. Thus, in addition to adjustments in slip at seismically loaded boundaries of the rupture surface, it is reasonable to expect a large interaction with the asthenosphere.

The earliest studies of relaxation following major thrust-type earthquakes ignored completely the details of fault geometry and examined the nature of the interaction of the elastic lithosphere with the ductile asthenosphere. Elasser [1969, 1971] introduced the stress guide hypothesis of a strong elastic lithospheric plate over a fluid asthenospheric channel.

If the scale of lateral variations in stress is large compared to the thickness of the plate and asthenosphere, the horizontal displacement u takes the form of the diffusion equation:

$$\frac{\partial u}{\partial t} = \frac{h_1 h_2 E}{\eta} \nabla^2 u$$

where h_1 and E are the thickness and Young's modulus of the lithosphere and h_2 and η are the thickness and linear viscosity of the asthenosphere. The obvious interpretation pointed out by Bott and Dean [1973] is that localized disturbances (stress drops) associated with earthquakes on plate margins will diffuse away. Quasi-static stress waves might be measurable near the plate boundaries, but changes will be damped out before reaching the interior region of the plate.

The mechanical effect of a major earthquake which ruptures through the lithosphere is to locally decouple the two plates. Stresses are relieved over much of the slip zone and transferred laterally along the plate margin, and vertically downward to the asthenosphere. As the asthenosphere relaxes more and more load is applied to the unslipped plate margin. Anderson [1975] suggests that this diffusive transfer of load along the plate margin can explain observed migrations of earthquakes [Mogi, 1968a, b]. For example, a systemic migration of activity from Japan to Alaska has occurred during the last several decades.

Similar patterns have been noted from Central America to Chile and on the Anatolian fault in Turkey.

The utility of the stress guide model arises from its simplicity. Only two mechanical elements are present. The elastic plate supplies the driving force for diffusion; the viscous asthenosphere balanced against the elastic stiffness determines the rate. Hence the nature of

diffusion is insensitive to geometric detail at the plate margin, but is diagnostic of upper mantle rheology. The earliest models of the stress guide concept assumed a linear viscous asthenosphere which results in the linear diffusion equation. The results are consistent with an effective viscosity of 5×10^{19} poise [Anderson, 1975]. Melosh [1976] has modified the model to account for a nonlinear stress-strain behavior that appears to be more appropriate for the asthenosphere [Weertman and Weertman, 1975; Post and Griggs, 1973]. The result is, naturally, a nonlinear diffusion which strongly restricts the postseismic stress and displacement fields close to the edge of the plate. The typical behavior is for stress and displacement contours to spread rapidly from the plate edge, and then quickly slow down beyond some characteristic distance. Comparison of predicted stress diffusion with the observed pattern of aftershock migration from the 1965 Rat Island earthquake strongly suggests a nonlinear asthenosphere.

A second class of postseismic thrust models concentrates on the near field plate interaction. In most cases the data of interest are confined to a very small region on the upthrown block (ie., the island arc). These allow some insight into the relaxation process, but yield little information on the large scale diffusion of the earthquake perturbation away from the fault.

One of the largest and most studied episodes of postseismic deformation followed the 1946 Nankaido, Japan earthquake (see fig. 2). Fitch and Scholz [1971] obtain reasonable fits to the postseismic data with additional forward slip on the down-dip extension of the fault plane and back slip of several meters on portions of the fault that slipped during the earthquake.

The model is based on the solution for slip on a rectangular fault plane in an elastic half space. The requirement of backslip has been criticized by Nur and Mavko [1974] who suggest the alternative explanation that viscoelastic decoupling of the lithosphere from the asthenosphere caused the deformation. Comparison of theoretical and observed deformation suggests an effective viscosity of 10^{19}- 10^{20} poise for the asthenosphere. This is an order of magnitude or more lower than viscosities based on rebound from crustal loading [Cathles, 1975; Walcott, 1973], but may be appropriate for the anomalously high temperature, low Q Nankaido region [Kanamori, 1970; Barazangi, et al., 1975]. This is similar to the model presented by Smith [1974].

As with many strike slip earthquakes it is possible, because of limited geodetic data, to fit the Nankaido observations with a combination of mechanisms [Smith, 1976]. Furthermore it is reasonable that anomalously high temperature will tend to enhance fault creep as well as asthenospheric relaxation. However, based on the evidence for diffusive stress guide behavior and the problem of backslip it is probably not reasonable to neglect asthenospheric interaction.

One of the largest events with measurable postseismic deformation was the 1964 Alaskan earthquake [Brown, et al., 1977]. A series of first order leveling surveys was made in 1964, 1965, 1968 and 1975 along a route between Anchorage and Whittier, Alaska, roughly normal to the surface trace of the fault. A gradual uplift of as much as .55 m was revealed with a pattern roughly parabolic in shape, convex upward. The motion occurred at an approximate exponentially decaying rate, with time constants ranging from 1 to 8 years depending on position. There is some

indication that the position of maximum uplift has migrated toward the trench. Brown, et al. [1977] explain the observations with postseismic creep along a deep extension of the fault zone, with the possibility of some superimposed strain accumulation and viscoelastic rebound.

Strike Slip Earthquakes

As with the case of dip slip earthquakes, a major uncertainty exists in resolving the source location of postseismic strain following strike slip events. The two most likely sources appear to be viscoelastic decoupling at the base of the surface elastic plate and aseismic fault slip within the plate.

In order to determine the characteristics of deformation due to viscoelastic relaxation at the base of the lithosphere, Nur and Mavko [1974] and Mavko [1977a] consider an earth model consisting of an elastic layer (lithosphere) overlying a linear viscoelastic half space (asthenosphere) as shown in figure 4. The half space is chosen to be a standard linear solid which allows for a nonzero relaxed shear modulus μ_r (always less than the initial or unrelaxed modulus: $\mu_u > \mu_r \geq 0$). The special case $\mu_r = 0$ is a Maxwell solid which behaves like a fluid over long times.

Assuming that at time t = 0 sufficient tectonic stress has accumulated to cause sudden faulting, a screw dislocation is introduced instantaneously into the layer at depth D (D <H where H is the layer thickness). The slip Δu occurs uniformly over a vertical fault surface as a step function in time.

The resulting surface displacement is shown in figures 5 and 6. The sudden appearance of the screw dislocation produces horizontal deformation (at t = 0) identical to values in an elastic layered half-space. But as

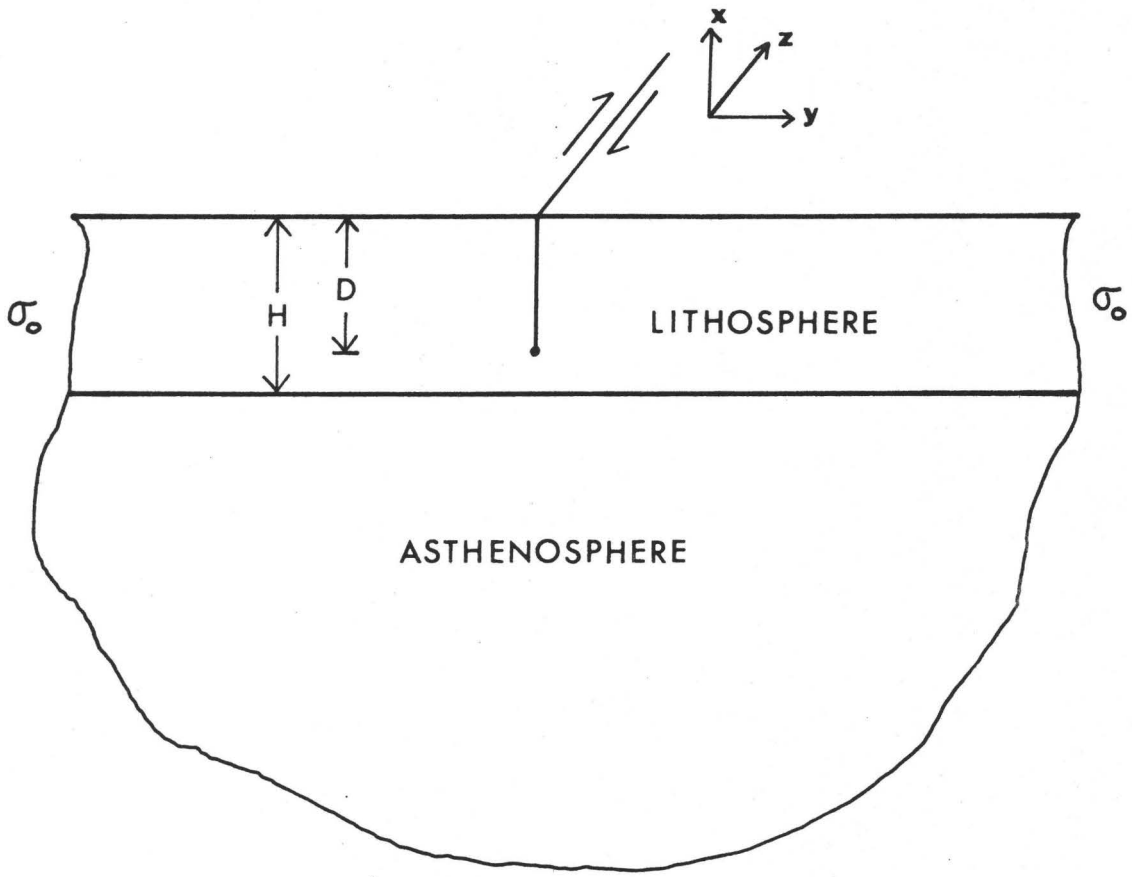


Figure 4. Simple model for earthquakes in an elastic lithosphere over a viscoelastic asthenosphere.

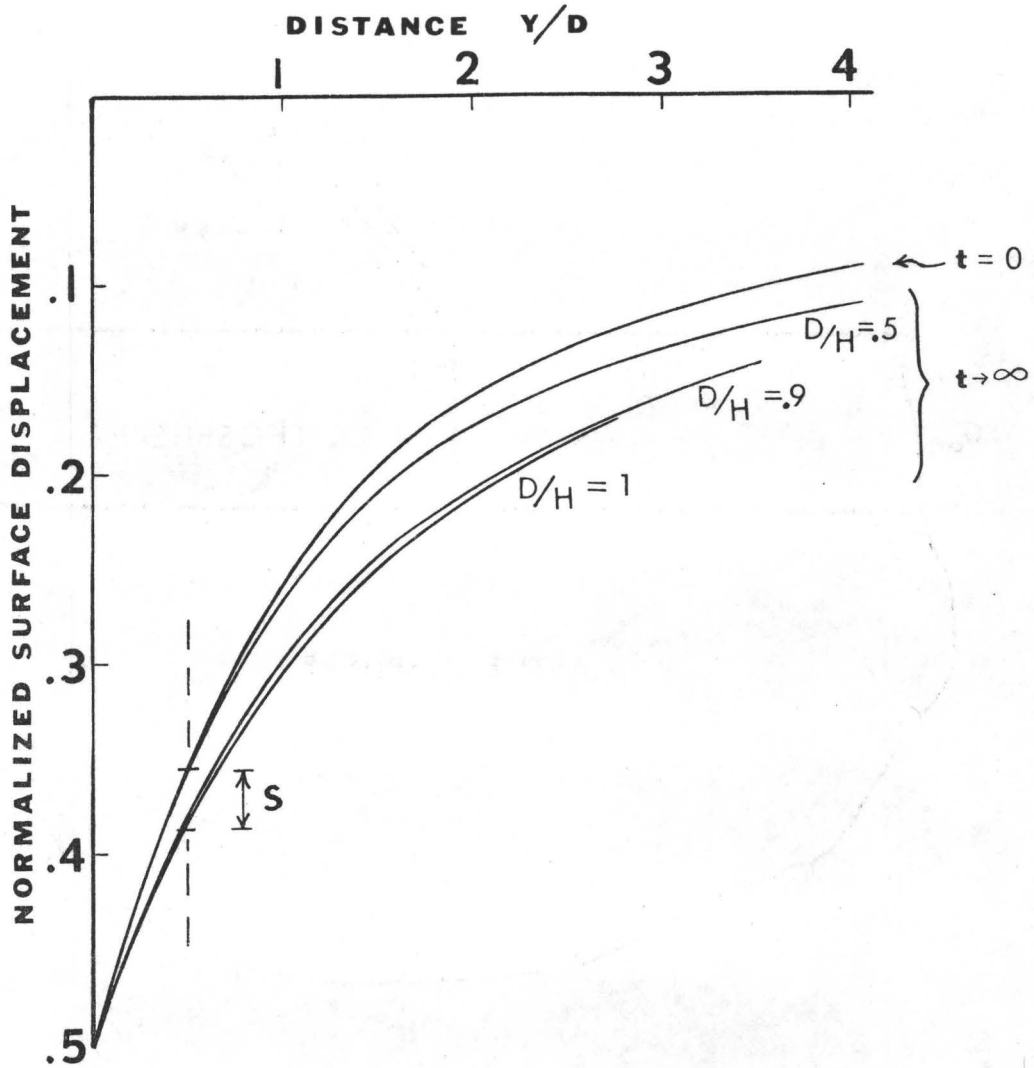


Figure 5. The postseismic surface displacements vs. distance from the fault trace resulting from partial viscoelastic decoupling at the base of the lithosphere. The curve t = 0 gives the coseismic surface displacement for uniform slip to depth D in an elastic earth (halfspace). The curves $t \to \infty$ give the subsequent relaxed displacements (for $\mu_r/\mu_u = .5$) for several values of D/H. Displacement is normalized to unit slip and horizontal distance is in units of fault depth.

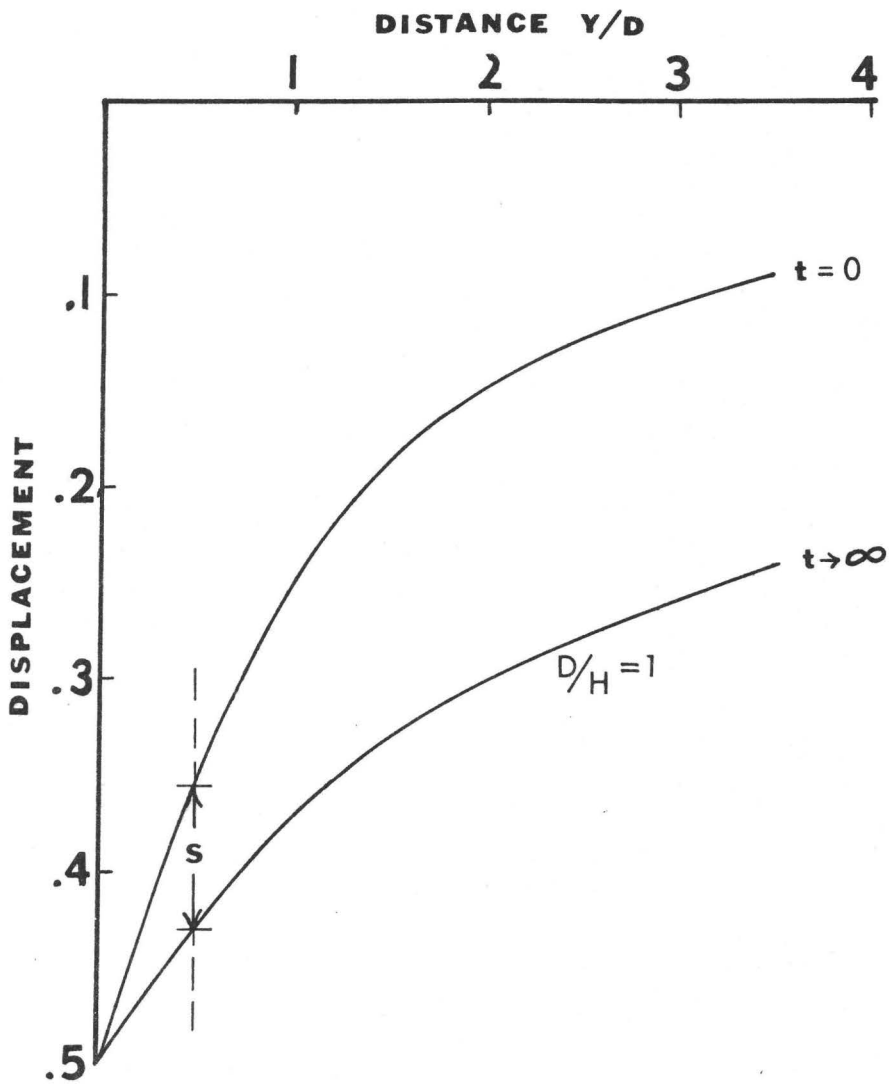


Figure 6. Postseismic surface displacement as in figure 5 for $\mu_r/\mu_u = .2$.

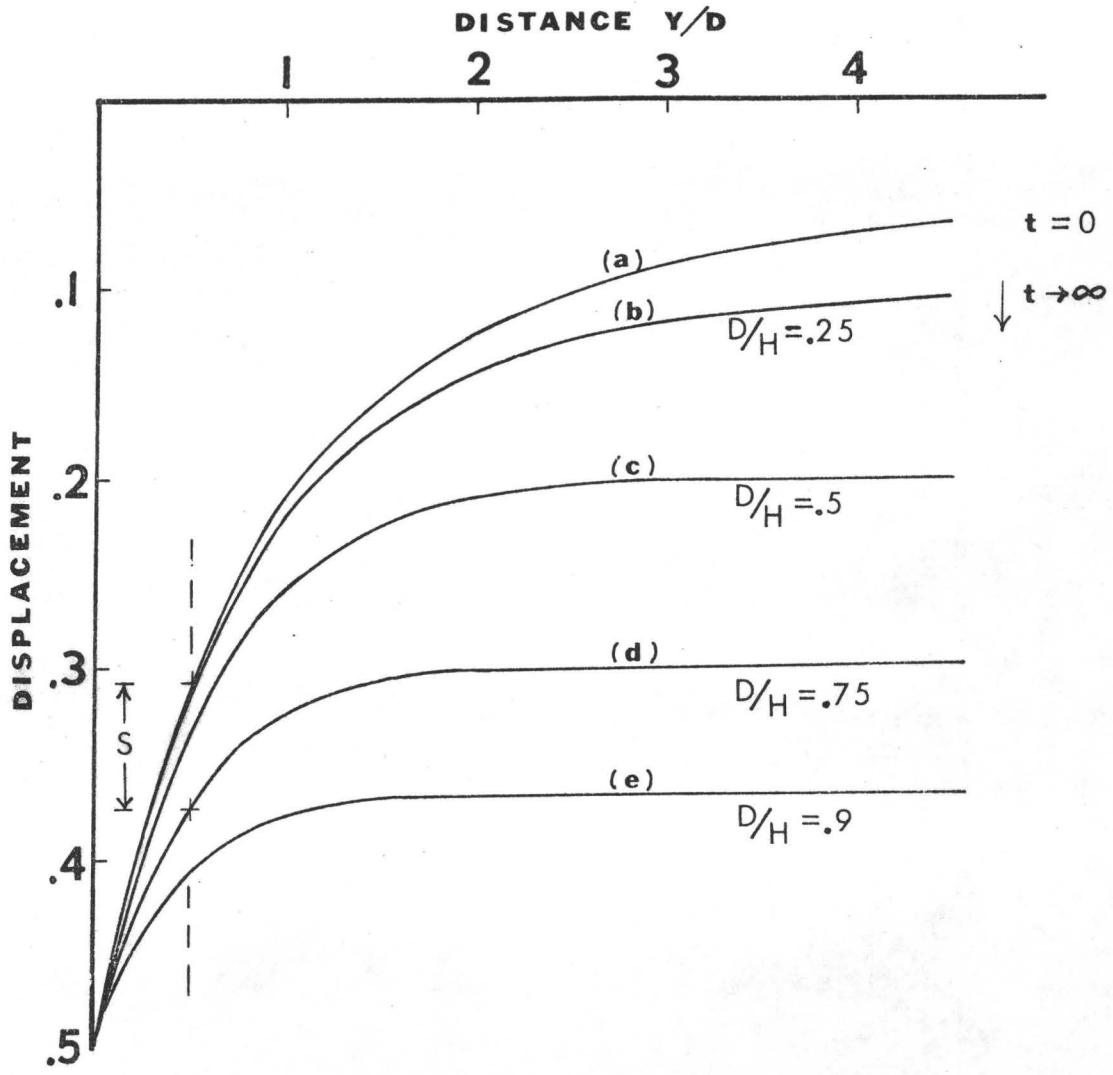


Figure 7a. Postseismic relaxation curves similar to figures 5 and 6. Here the stress drop is uniform over the fault instead of slip, and $\mu_r/\mu_u = 0$.

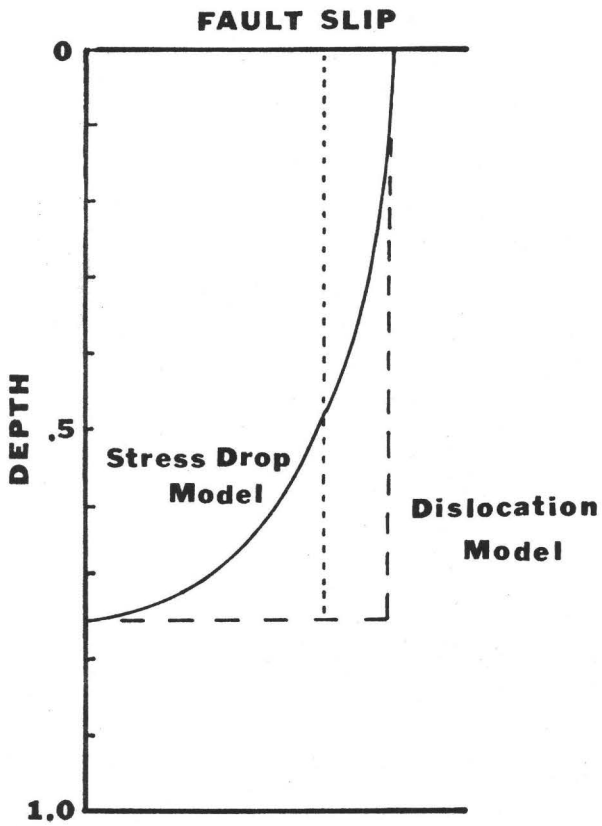


Figure 7b. Comparison of fault slip vs. depth for the dislocation model(broken curves)and uniform stress drop model (solid curve). The vertical axis is in units of plate thickness (This example is for D/H = .75). The dashed dislocation curve has the same surface slip as the stress drop curve. The dotted dislocation curve has the same average slip as the stress drop curve.

time progresses, the deformation changes, due to the relaxation of stresses in the lower half-space. Curves for the limiting case $t \to \infty$ are shown for several values of D/H and μ_r/μ_u. The difference between the curves $t \to \infty$ and $t = 0$ gives the total postseismic deformation.

The net change in deformation between initial and relaxed states plus some estimate of the relaxation time are often all that we need from a viscoelastic model. Observations of deformation are usually not frequency enough to reveal a detailed time history, so that the time dependence predicted by the model doesn't add much to the analysis. In cases like this it is convenient to note that the initial and relaxed states for a viscoelastic problem can be completely described by the initial and relaxed elastic moduli. Hence we can always supplement our collection of relaxation curves like those in figures 5 and 6 by constructing the appropriate static elastic solutions.

For the special case of an elastic layer over a Maxwell viscoelastic half-space, the relaxation process following strike slip faulting gradually transforms the initial solution in an elastic layered half-space to a final solution in an elastic plate with a traction-free base. The "relaxation" curves in figure 7 are obtained from the static elastic solution for a plate with a uniform stress drop fault, given Turcotte and Spence [1974]. To plot the change corresponding to a constant applied stress, the shear field has been subtracted out. Each curve is normalized to unit offset at the fault trace. Although the slip is not uniform with depth as in figures 5 and 6 this is a more convenient representation since the screw dislocation solution is very slow to converge for $\mu_r / \mu_u \to 0$. It is also interesting to note that these relaxed curves describe an

elastic layer over any decoupling material. Since we only examine the completely relaxed state, it only matters that the base tractions eventually become very small.

Certain features of the deformation shown in figures 5, 6 and 7 are immediately evident. The overall amplitude of surface displacement is proportional to fault slip. For fixed material constants, variations in fault depth simply stretch or shrink the horizontal scale without changing the basic shape of deformation. Shallower faults have a much more localized effect; deeper faults, much broader. Deeper faults (larger D/H) are also accompanied by a larger and, hence, more measurable change in displacement between initial and relaxed states. This occurs because the closer the dislocation is to the boundary, the greater is the coseismic stress superimposed on the asthenosphere -- hence, the greater the relaxation that can occur. Similarly, a greater change in displacement between initial and relaxed states without change in horizontal scale, is accomplished with a lower ratio μ_r/μ_u. This permits a higher degree of relaxation in the half space.

It is tempting to apply these results to observations. Consider, for example the 1906 San Francisco earthquake. Thatcher [1975a] reports that a (tensor) shear strain accumulated parallel to the San Andres fault at a rate of $\sim 1.2 \times 10^{-6} yr^{-1}$ for about 30 years after the earthquake or $\sim 3.6 \times 10^{-5}$ for the entire period. This represents an average over a region of roughly $\pm$ 5 km on either side of the fault. The corresponding right lateral displacement at 5 km relative to the fault trace is $\sim$.36 m.

Assuming a coseismic slip of $\sim$ 5 m from the surface to 10 km depth we can scale the curves in figures 5, 6, and 7 to find the postseismic rebound that might be expected from viscoelastic relaxation below the seismic layer. For this example a horizontal distance of 5 km corresponds to y/D = .5, shown by the vertical dashed lines in figures 5, 6 and 7. The values of displacement in each of the curves should be multiplied by 5 m. For a partial relaxation of (μ_r/μ_u) = .5 the greatest possible postseismic displacement at 5 km is $\sim$.15 km (labeled S in figure 5) for an elastic layer thickness of 10 km (D/H = 1). This is a factor of 2 - 3 smaller than inferred from the observed strain. For a greater viscoelastic relaxation, (μ_r/μ_u) = .2 (figure 6), a correspondingly greater displacement of $\sim$.4 m is estimated for a 10 km elastic layer (D/H = 1). This is on the order of the observed value. Other fits to the near field strain data can be found using a range of thicker layers and greater viscoelastic relaxation. For the Maxwell solid (figure 7) a displacement of $\sim$.3 m is estimated for a 13 km thick layer (D/H = 3/4). Although the surface slip here is 5 m, the average slip at depth is $\sim$.8 what it would be for a dislocation model since figure 7 gives the uniform stress-drop solution. Hence, a slightly thicker plate might accomplish the same deformation if relatively more slip occurred near the lower edge of the break. The subsurface fault slip for the dislocation and uniform stress drop models is compared in figure 7b.

We conclude from this simple-minded comparison that all of the observed postseismic strain change can be explained by viscoelastic relaxation alone if the elastic layer is only slight deeper than the 10 km depth of

faulting, and if it becomes almost decoupled during relaxation.

In contrast, Thatcher [1975a] explains the observed postseismic strain change with additional fault slip at depth, occurring over a period of years following the earthquake. He dismisses viscoelastic relaxation of the asthenosphere by assuming that the lithosphere is much thicker than the depth of faulting. This is reasonable for a "typical" lithosphere of 100 km. But in view of an anomalously low P_n velocity [Thompson and Burke, 1974; Herrin, 1972] and the suggestion that the crust in parts of California may be decoupled [Anderson, 1971; Hadley and Kanamori, 1977], a discussion of thin plates is warranted.

A systematic comparison of the mechanisms of postseismic straining for the 1906 San Francisco earthquake is attempted by Rundle and Jackson [1977]. Their model consists of an elastic layer over a (Maxwell) viscoelastic half-space. They prescribe on the lower portion of the plate a steady fault creep as well as an episodic postseismic creep as suggested by Thatcher [1975a]. A fairly arbitrary loading strain proportional to the far field plate velocity is also included. The strain data (angle changes) within a few kilometers of the fault do not resolve the question of the dominant mechanism. However good fits are obtained without episodic fault creep using layer thicknesses of 16-50 km with a poorly constrained range of viscoelastic decay times and loading strain fields. Because steady straining is assumed to account for a portion of the observed angle changes, the inferred plate thickness is systematically greater than found in our purely viscoelastic analysis given above. Rundle and Jackson are probably the first to quantitatively suggest the viscoelastic mechanism for the 1906 event.

Barker [1976] makes a similar comparison of displacements from transient fault creep and viscoelastic relaxation, though not specifically for any one earthquake. His viscoelastic model consists of an elastic layer over a Maxwell viscoelastic half-space. At time t = 0 a strike slip fault of finite length is introduced, breaking through the entire lithosphere. At subsequent time the surface deformation along the perpendicular bisector of the trace is qualitatively similar to the 2-D results in figures 5, 6 and 7. Barker compares this deformation with that obtained from a purely elastic model where the original fault gradually extends deeper into the (elastic) half-space. He concludes that with the limited quality and spatial extent of data generally available we cannot expect to distinguish between the two mechanisms.

A second piece of evidence which may help to distinguish the source of postseismic strain is the net relative displacement of points further from the trace -- the "rigid block" offset. To establish a framework for looking at more distance plate deformation, let us consider again the lithosphere-asthenosphere model shown in figure 4. The upper plate surface is stress free. On the lower surface we assume rate-dependent viscous shear tractions that are negligibly small during much of the earthquake cycle, but which may be substantial during episodes of rapid strain release. This assumption is clearly subject to debate. Melosh [1977] argues that the average basal stress must be at least one or two orders of magnitude smaller than shear stresses acting at plate boundaries. In the following discussion we describe the forces acting on a fairly local part of plates making up the earthquake system. The basal area of interest will be in general less than a factor of 10 times the corresponding area of the plate edge.

If Melosh is correct then we may neglect the basal tractions. In the context of our viscoelastic discussion we must, however, acknowledge significant basal stresses during periods of rapid coseismic and post-seismic change. If the viscoelastic relaxation time is short compared to the recurrence time of major earthquakes then we might expect these tractions to eventually die away.

Driving forces on the plates include a gravitational push at ocean ridges and a pull at subduction zones. In addition asthenospheric tractions integrated over the entire plate area might contribute a net drive or drag even though we might neglect them on a local scale. For the local system in figure 4 we assume that the net effect of the distant forces is to apply a nearly uniform shear stress $\sigma_{yz} = \sigma_o$ to the lithosphere. At the fault plane, y = 0, the shear stress σ_{yz} can vary signficantly with x and z resulting from variations in fault displacement and frictional resistance stresses. However, for a system in equilibrium, the mean of σ_{yz} at y = 0 must equal σ_o during periods when the viscous shear tractions from below are negligible.

To explore the role of boundary conditions on a system like this Mavko [1977b] derives a simple relation between the far field applied stress σ_o, the far field displacement u(x,y,z) and the fault slip $\Delta u(x,z) = u(x,0^+,z) - u(x,0^-,z)$. In two dimensions the result for negligible asthenospheric tractions is

$$W \frac{\sigma_o}{\mu} = 2U_o - \frac{1}{H} \int_o^H \Delta u(x) dx \qquad (1)$$

where μ is the average shear modulus of the lithosphere and U_o is the fairly

uniform plate displacement relative to the fault trace at a distance y = W from the fault. W must be large enough (W>H) so that elastic perturbations due to variable fault have died away. Defining $\Delta U_o = 2U_o$ and $\Delta\bar{u} = \frac{1}{H}\int_o^H \Delta u dx$, (1) becomes

$$\sigma_o = \frac{\mu}{W}(\Delta U_o - \Delta\bar{u}) \qquad (2)$$

Equation (2) is an extremely simple relation between the relative plate displacement at $y = \pm W$, the average fault slip $\Delta\bar{u}$ and the mean or applied stress σ_o. Note that the average slip is defined in terms of the plate dimension H. Nothing has been assumed about fault material properties or the details of seismic or aseismic creep. It is valid during portions of the earthquake cycle when asthenospheric tractions are small. Differentiating with respect to time we obtain:

$$\dot{\sigma}_o = \frac{\mu}{W}(\Delta\dot{U}_o - \Delta\dot{\bar{u}}) \qquad (3)$$

The simple interpretation of (3) is that if the gross plate speed $\Delta\dot{U}_o$ exceeds the fault slip rate $\Delta\dot{\bar{u}}$ because of frictional resistance or locked portions of the fault, then the mean stress in the plate grows. The mean stress is relieved when the fault slip rate $\Delta\dot{\bar{u}}$ exceeds $\Delta\dot{U}_o$.

Using equation (2) we can examine the far field displacement and strain accompanying the 1906 earthquake. Prior to the earthquake the mean stress, plate displacement and fault slip are given by σ_o^i, ΔU_o^i, $\Delta\bar{u}^i$. If the asthenospheric tractions are small these are related by equation (2):

$$\Delta U_o^i = W\left(\frac{\sigma_o^i}{\mu}\right) + \Delta\bar{u}^i \qquad (4)$$

When the earthquake occurs, a substantial amount of fault slip takes place to a depth D. Since the short term response of the earth is essentially elastic the abrupt surface strain change should be confined to within ~D of the fault. This is the result predicted by elastic half-space models of faulting. The coseismic changes in σ_o and ΔU_o are therefore ~0.

During the postseismic period asthenospheric stresses gradually relax. from an elastic half-space to a plate-like response.
In the case of a Maxwell asthenosphere the system evolves from an elastic half-space to a plate-like response, as discussed earlier.

In addition, postseismic slip may simultaneously occur at depth on the fault. When readjustment is complete the final mean stress and far field displacement σ_o^f and ΔU_o^f are once again related to the fault slip $\Delta u^f(x)$ by equation (2):

$$\Delta U_o^f = W\left(\frac{\sigma_o^f}{\mu}\right) + \Delta \bar{u}^f \qquad (5)$$

Subtracting (4) from (5) we can write the postseismic far field displacement change as:

$$\Delta U_o^f - \Delta U_o^i = \frac{W}{\mu}(\sigma_o^f - \sigma_o^i) + \Delta \bar{u}^f - \Delta \bar{u}^i \qquad (6)$$

The first term on the right is linear with distance, W, from the fault. This corresponds to a homogeneous strain change proportional to the mean stress change in the plate. The second term is a uniform rigid block displacement equal to the average coseismic plus postseismic slip. Neither far field term is predicted by elastic half-space models of faulting where disturbances -- stress, strain and displacement -- tend to zero far from the fault.

If the mean or applied stress in the plate is the same after relaxation as before the earthquake ($\sigma_o^f = \sigma_o^i$) as suggested by Savage [1975] then the relative block offset that occurs during the postseismic period is equal to:

$$\Delta U_o^f - \Delta U_o^i = \frac{1}{H} \int_o^H [\Delta u_c(x) + \Delta u_p(x)]\, dx \qquad (7)$$

where Δu_c is the coseismic slip and Δu_p is the postseismic slip. Hence, the far field offset measures the sum of coseismic and postseismic slip, and it is not correct to interpret it as due to postseismic slip alone.

As an illustration of the effect of viscoelastic decoupling, consider the faulting parameters determined by Thatcher [1975a] for the 1906 earthquake. The average coseismic slip is $\sim$4 m from the surface to a depth of 10 km. The postseismic slip is $\sim$4 m between 10 and 30 km. We simplify calculations by using a uniform stress drop model in place of uniform slip. The results are shown in figure 8a. Curve (a) shows the coseismic surface displacement in a half-space for a 10 km deep fault with maximum surface slip of 4 m. Curve (b) gives the displacement after the slip has deepened during the postseismic period to 30 km, while keeping the surface slip fixed at 4 m. For both (a) and (b) the displacement tends to zero far from the fault, though rather slowly for the deeper case. Curve (c) shows the displacement for the same 30 km deep slip but now in a layer decoupled at 40 km. This value of layer thickness is arbitrary, chosen for illustration to be only slightly greater than the depth of faulting. Far from the fault the displacement is asymptotic to 1.2 m.

Figure 8b shows the postseismic displacement obtained by subtracting curve (a) from (b) and (c) in figure 8a. The curve for a 100 km thick plate is also shown for comparison. For the first 5 km the curves are within about 15% of each other. This suggests that the near fault strains are fairly insensitive to the details of deep model structure. This is the same conclusion arrived at earlier [Rundle and Jackson, 1977,

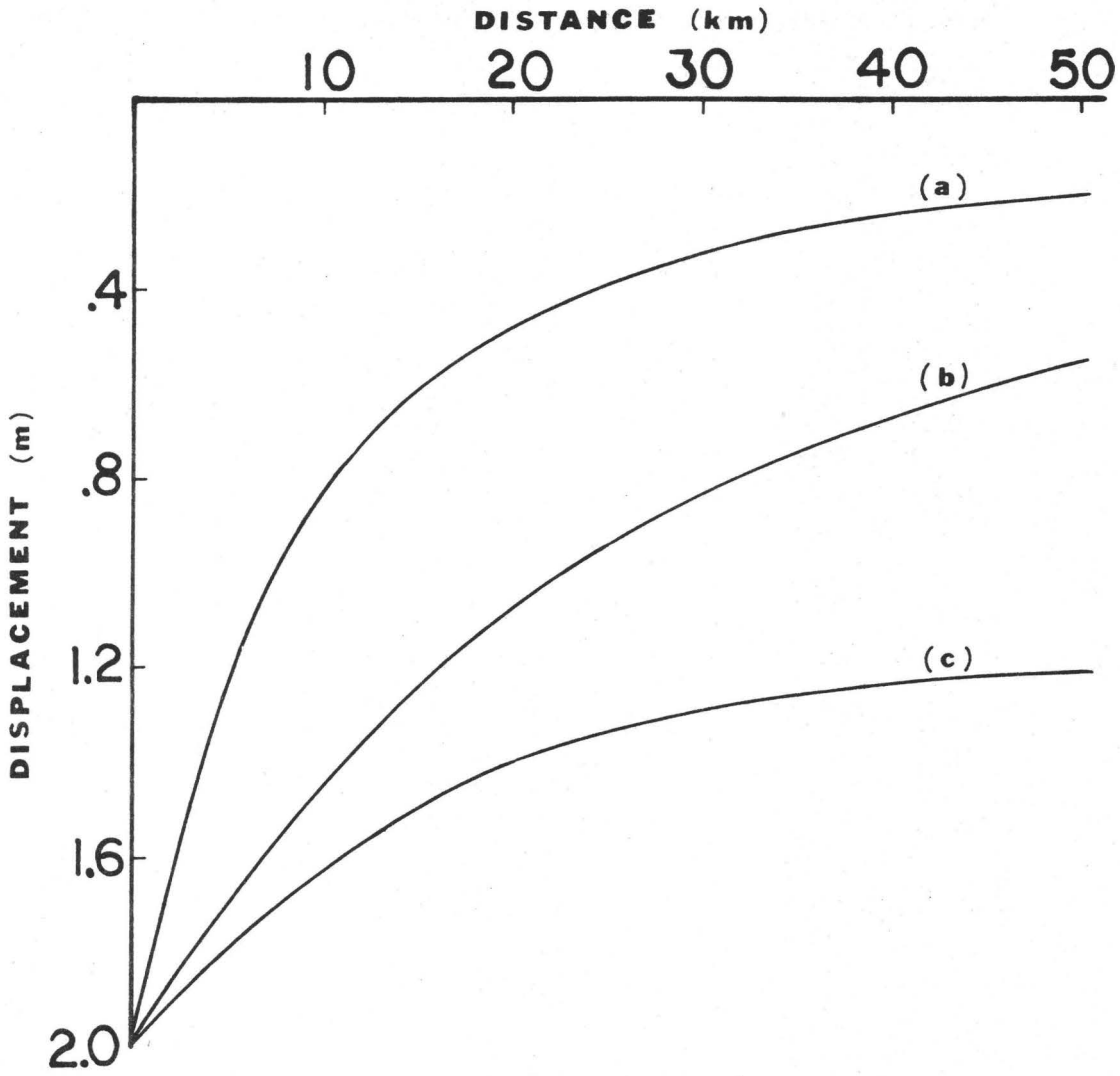


Figure 8a. Postseismic displacement models for the 1906 San Francisco earthquake. Curve (a): the coseismic displacement for a constant stress drop crack 10 km deep in an elastic half-space. Curve (b): the displacement after the original fault deepens to 30 km in an elastic half-space, with no viscoelastic decoupling. Curve (c): the displacement for 30 km deep fault in a layer decoupled at 40 km.

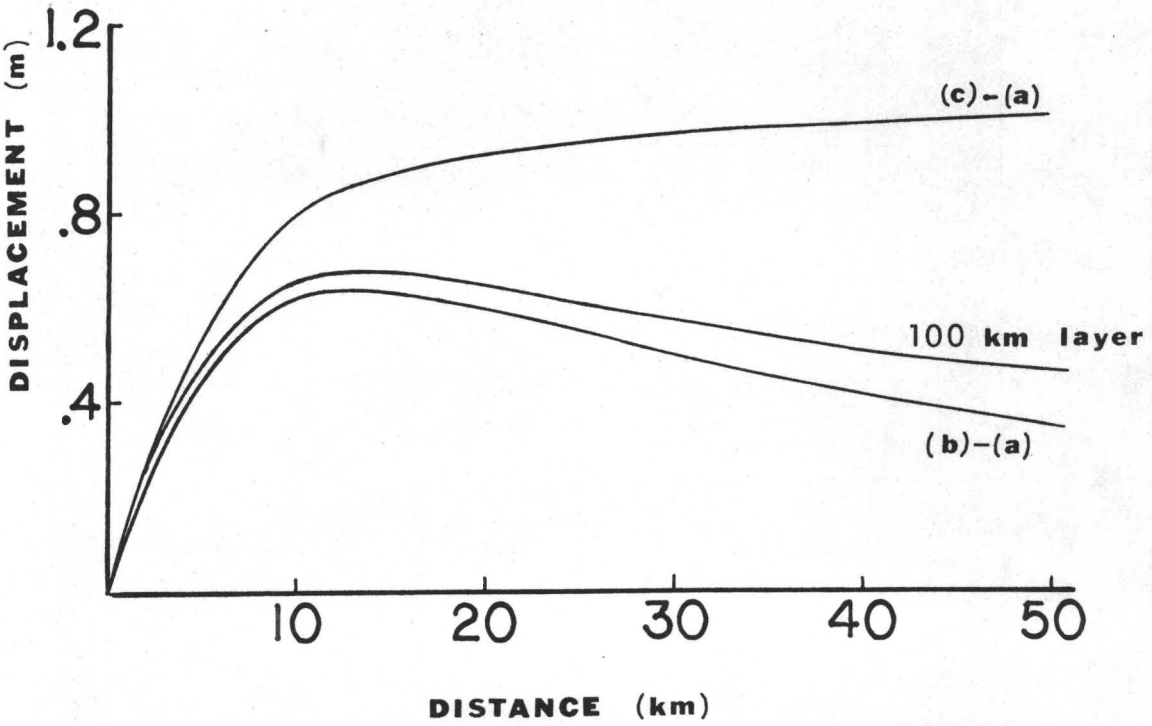


Figure 8b. Postseismic displacement given by the difference of curves in part a. Curve for a 100 km decoupled layer is also shown.

Barker, 1976]. Beyond 5 km the curves diverge. At 50 km the thin layer model predicts 2.5 times the displacement given by the half space. Hence for a given set of data both the halfspace and plate models may predict similar near-fault strain, but halfspace models demand significantly greater slip at depth to explain distance offsets. If the plate is thick, say 100 km, the curves are within 30% for the first 40-50 km

We can draw several conclusions from the above results. Postseismic changes in strain near a strike slip fault can result from a number of sources -- fault slip, viscous decoupling at depth, and changes in the applied or mean stress field. If we constrain all but one mechanism, then we can generally resolve features of the remaining one. Though without other information the source of strain cannot be resolved from near field data alone. In contrast, the far field data may contain separable information if measurements span most of the period of postseismic adjustment. The far field strain changes reflect mean stress changes in the lithospheric plate. The rigid block offset gives the average fault slip. If the plate thickness is known, the total of coseismic and postseismic slip can be found. If only the coseismic slip is known, an upper bound on the effective plate thickness can be determined.

Discussion

Some generalities of postseismic relaxation are evident. Earthquakes always tend to decrease stress over most of the slip zone and increase it elsewhere, with the greatest stress concentration usually near the margins of the slip. Following an earthquake any number of viscous elements may begin to relax.

The largest postseismic changes in stress and strain probably result

from aseismic fault slip and nonlinear viscoelastic decoupling of the surface plate. The nature of these changes are largely the subject of this section. Other mechanisms may be more subtle. For example, large shallow earthquakes can induce coseismic changes in fluid pore pressure that are comparable to the stress drop on faults. The subsequent redistribution of pore pressure [Nur and Booker, 1972] via diffusion slowly decreases the strength of rock and may result in delayed fracture, i.e. the characteristic 1/t decay in aftershock frequency. Other effects of fluid relaxation include transients in seismic velocity, gravity, water well levels, and electrical resistivity.

The relative importance of postseismic slip and viscoelastic decoupling will depend on a number of factors. Major strike-slip earthquakes are usually confined to the shallow, brittle portion of the lithosphere (the upper 20 km or so of the crust). Depending on the steady or accelerated creep prior to the earthquake a substantial amount of postseismic slip adjustment may be required near the bottom of the rupture to relieve the coseismic stress concentration and to accomodate relative plate offset everywhere on the fault [Mavko, 1978]. Deeper ruptures that break through the entire lithosphere might in principle require less slip adjustment. A similar statement would apply to major thrust type earthquakes. In contrast the viscous component of adjustment occurs as the earthquake stress field relaxes in the decoupling layer (asthenosphere) and the surface strain evolves from a half space to plate response to faulting. For shallow earthquakes a smaller viscous adjustment is expected since the stresses fall off quickly with distance below the rupture. Deep earthquakes interact much more with the asthenosphere and produce a larger relaxation. Once this geometric condition for interaction is achieved, we obviously need a low enough viscosity for the

relaxation to be recognized as a discrete phase in the earthquake cycle.

Kanamori [1971] outlines several stages in the evolution of subducting plate boundaries that we can try to identify with different modes of relaxation (see figure 9). A young subduction zone, of the Alaska-Aleutian type is characterized by a low angle thrust of thin, stiff oceanic lithosphere beneath the much thicker continental lithosphere. Temperatures are low because frictional (viscous) heating has just recently begun. The corresponding high effective viscosity (perhaps 10^{21}- 10^{22} poise, representative of postglacial uplift under more stable areas) and the large plate thickness will tend to reduce the effectiveness of viscoelastic relaxation and emphasize the importance of postseismic creep. This is consistent with the view of Brown, et al. [1977] that postseismic deformation following the relatively shallow 1964 Alaska earthquake was due primarily to creep.

An older subduction zone (Figure 9b) is characterized by higher temperatures. This results in a thinner lithosphere due to partial melting and a lower viscosity asthenosphere on the continental side of the trench. The dip angle also tends to be greater, so that a given down-dip rupture length will break through a relatively greater portion of the lithospheric thickness compared with the low angle thrust. In general, conditions are favorable for enhanced viscoelastic relaxation.

Most of Japan has the features of such an older subduction zone [Kanamori, 1971], although signfiicant local variations in structure are evident (see figure 10). Much of northern Japan along the Japan trench and Kurile trench is underlain by a relatively high Q asthenosphere [Barazangi, et al., 1975]. The 1973 Nemuro-Oki earthquake in this region had a large postseismic deformation which Kasahara [1975] attributes to

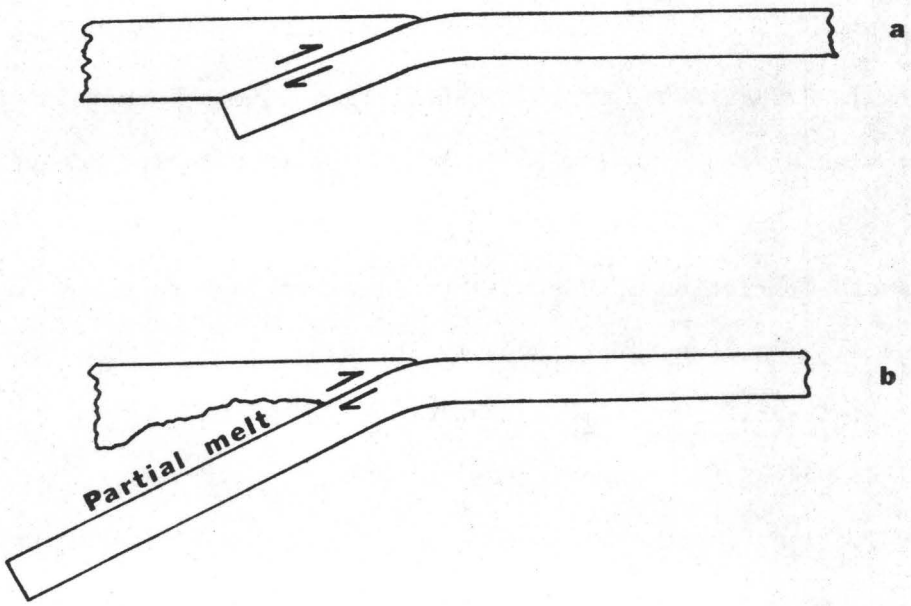


Figure 9. Two stages in the model of Kanamori [1971] for the evolution of subduction zones.

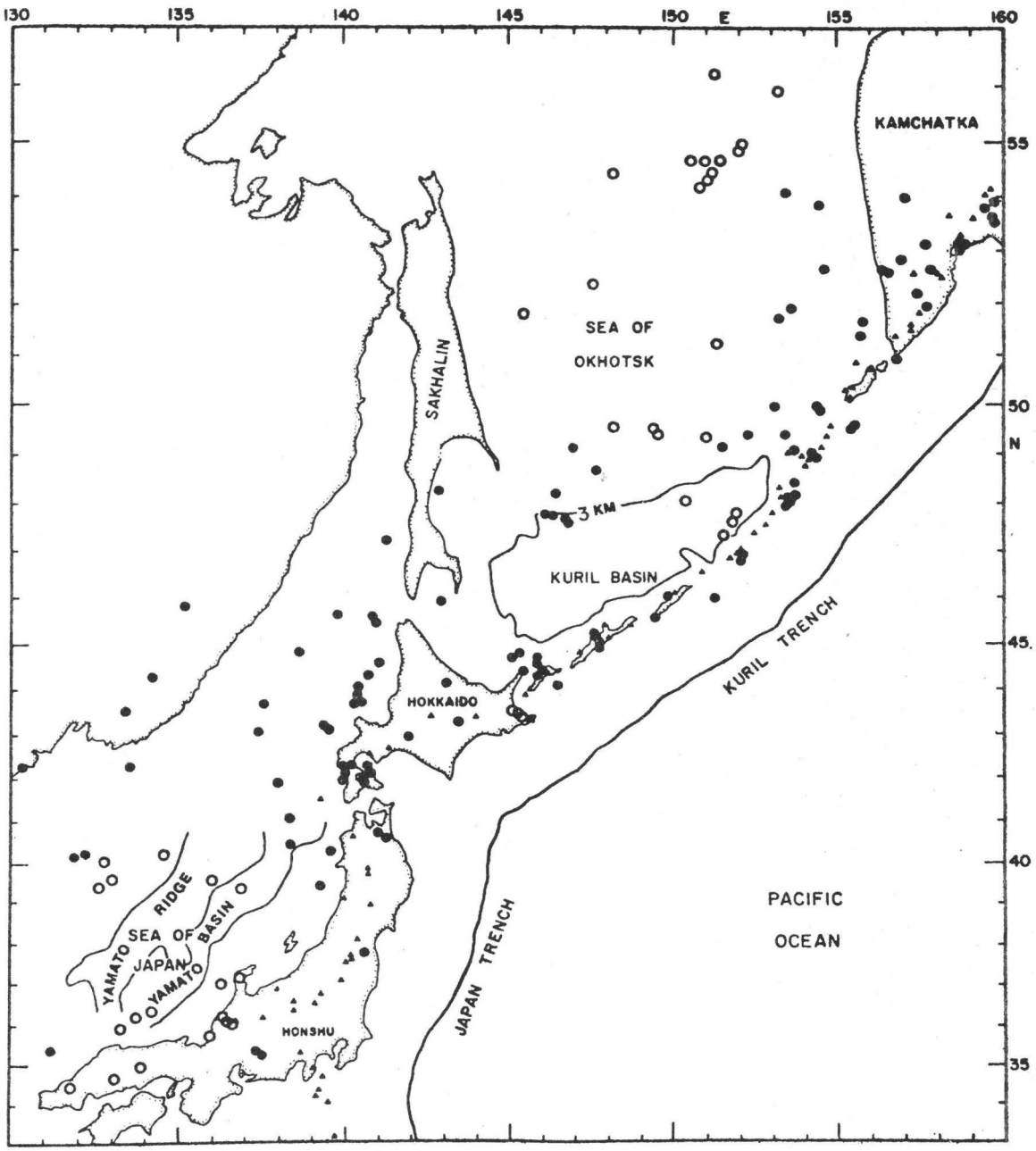


Figure 10. Map showing distribution of high and low Q regions behind Island arcs [Barazangi, et al., 1975]. Open circles - low Q; Closed circles - high Q.

deep aseismic slip. This interpretation is consistent with the high Q (low temperature) asthenosphere and the fact that the coseismic slip ruptured less than half the thickness of the lithosphere. Several other thrust type earthquakes (1968 Tokachi-Oki, M = 7.9; 1976 Kurile Islands, M = 7.8) showed no observable postseismic deformation [Kanamori, 1973].

Farther to the south, under Shikoku and southern Honshu there is evidence for an anomalously high temperature, low viscosity asthenosphere [Barazangi, et al. 1975; Kanamori, 1970]. The 1946 Nankaido earthquake (thrust) and 1923 Kanto (substantial strike slip) each broke through a substantial portion of the lithosphere and each had very large postseismic deformation. The interpretation of viscoelastic relaxation at Nankaido [Nur and Mavko, 1974; Smith 1974] or a combination of aseismic slip and viscoelastic relaxation [Smith 1976] is consistent with the deep rupture and low Q (high temperature creep might also be enhanced). The deformation at Kanto has been described as a difference in seismic and geodetic "slip" [Kanamori, 1973] but it is likely that viscoelastic relaxation also played an important role.

We can extend this analogy to transform faults. On continents, where postseismic measurements are possible, the lithosphere is in general very thick compared to the brittle seismic layer. Relaxation, if it occurs, is probably due to shallow effects like creep (1966 Parkfield earthquake, [Scholz, et al., 1969]),compaction of sediments (1927 Tango earthquake, [Kanamori, 1973]), fluid diffusion, etc. Transform faults with a complex history like the San Andreas fault, might have a thinner lithosphere, due to old subduction or some shallow effect which decouples the upper crust. The asymmetry in postseismic surface strain across the

San Andreas near San Francisco [Thatcher, 1975a] could be due to a non-vertical fault [Walsh, 1969], the variation in geology across the fault, or more interesting, some remnant asymmetry in the deep structure due to the old subduction.

A complication in interpreting postseismic transients is the possible coupling of contrasting mechanisms. Consider, for example the 1966 Parkfield earthquake. Seismic and geodetic evidence suggest that coseismic slip on the order of 30 cm occurred between 4 and 10 km depth. Little or no coseismic slip occurred at the surface, but as much as 25 cm of surface slip accumulated at a decaying rate in the 3 years following the earthquake. One simple interpretation of the postseismic slip is that the upper 4 km of the fault zone has a viscous stress-slip relation. For example,

$$\dot{\Delta u} = A \sigma^n$$

where $\dot{\Delta u}$ is the creep rate, σ is the applied stress, and A and n are positive constants. With this type of damped motion, very abrupt motion is resisted, but gradual motion occurs with relative ease. (A similar behavior is suggested by Kasahara [1975] for aseismic slip in the lower lithosphere). Hence, the abrupt concentration of stress near the surface is gradually relieved following the earthquake. Any sliding law with slip rate monotonic with stress can qualitatively explain postseismic slip at a decaying rate. The rate of slip is determined by the constants of the slip law. A second interpretation is that the upper 4 km of the fault are characterized by a single value of creep strength S, independent of strain or strain rate. Whenever the stress meets S, sliding occurs to guarantee that the stress does not exceed S (eg. plastic fault zone).

In this case we assume that the sum of preseismic plus coseismic stress is at or slightly below S. Near the fault, viscoelastic relaxation of some sort (perhaps a decoupling from below) gradually transfers stress back into the fault plane. Slip occurs at a rate proportional to the rate of reloading. Hence the slip rate is determined by the viscoelastic law somewhere off the fault and not by the sliding law. Obviously other combinations are possible.

STRAIN ACCUMULATION AND THE WHOLE CYCLE

Almost all earthquake models are designed to be compatible with the concept of relative plate motion. Over periods longer than at least several earthquake cycles the plates remain essentially elastic, while at narrow zones along plate boundaries the motion is accommodated as frictional sliding or plastic flow. Hence, a simple condition that we can impose is that over several earthquake cycles, the time averaged slip everywhere on the fault must equal the relative plate offset; the stress and strain must everywhere in the plate be roughly periodic without a net accumulation. Of course, deviations from this do occur, resulting in permanent deformation and additional faulting, but for our purposes we will assume them to be small.

In this section we discuss some recent models for strain accumulation and its role in the earthquake cycle. The main emphasis will be on strike-slip faults.

The only reasonably constrained portion of the cycle is the coseismic part. In general we can determine from seismic and geodetic data the approximate area and orientation of the fault plane, and the average slip. This is simplified by the short term elastic response of the earth which is insensitive to uncertainties in rheology and plate thickness. Heterogeneities in elastic response can also complicate interpretation, but these can in principle be corrected for using known velocity structure. The common feature of all coseismic strain fields is that they rapidly die out away from the fault, within several 10's of kilometers for strike slip faults and 100 km or so for major thrusts. This difference in scale results mainly from the depth of faulting and dip angle. Contrary to what some authors have suggested [Brune, 1974, Thatcher, 1975a], the fact that

the localized coseismic strain field depends on fault depth instead of fault length is not an indication of a large plate thickness. It indicates only that the short term response of the crust and upper mantle is essentially elastic to considerable depth.

A strict interpretation of elastic rebound would suggest that strain accumulation must match coseismic strain release. For example, Scholz and Fitch [1969] have argued that along the San Andreas fault in California accumulated strain must be concentrated near the fault on a scale comparable to the eventual earthquake rupture depth. If strains were not localized near the fault then earthquakes would not continue to occur along the same fault, but would migrate laterally where the unrelieved strain had accumulated.

There are a number of weaknesses in this argument. For example, following the large (M = 7.1) Imperial Valley, California earthquake of 1940, geodetic measurements revealed a broad zone (100 km wide) of right lateral shear strain that accumulated over the period 1941-1954 (see figure 11). Strict elastic rebound, as outlined above, would require an earthquake 25-40 km deep [Scholz and Fitch, 1969] in order to relieve this broad zone of strain. This depth is unreasonably large, since in California strike-slip earthquakes almost never occur below 15 km. The most widely accepted explanation for this discrepancy is that aseismic slip will take place on the fault, while a combination of creep and coseismic slip will occur shallower [Scholz and Fitch, 1969; Chinnery, 1970]. Savage and Burford [1970] point out that the broad zone of strain accumulation spans both the Imperial and San Andreas faults. There is a suggestion in the data that the single broad zone might, in fact, be resolved into two narrower zones straddling the separate faults.

Although a number of authors have tried to associate the depth of faulting with crustal thickness [Benioff, 1951; Byerly and DeNoyer, 1958; Scholz and Fitch, 1969] relatively few [Turcotte and Spence, 1974; Spence and Turcotte, 1976; Mavko, 1977b, 1978] have accounted for the effect of finite plate thickness in analysis of strain accumulation. For the case of the 1940 Imperial Valley earthquake, consider first the effect of postseismic viscoelastic decoupling that may have occurred at the base of the elastic plate. The average coseismic slip was about 1.7 m [Brune and Allen, 1967] and the fault depth was 8 km [Kasahara, 1958]. The expected change in surface displacement can be obtained from figure 7. The coseismic displacement is given approximately by curve (a) with the amplitude multiplied by 2.3 m (which gives an average slip at depth of 1.7 m) and the horizontal scale multiplied by 8 km (depth of faulting). If the plate is gradually decoupled, for example, at the base of the crust (32 km) the relaxed displacement is given by curve (b), scaled similarly. The net change is given by the difference between (a) and (b) and is shown in figure 11 (curve a), superimposed on the data. An additional surface offset of .17 m during the survey period has been suggested [Scholz and Fitch, 1969]. If we simply add this to the post-seismic displacement in curve (a) we get the larger net change shown as curve (b) in figure 11. Curve (b) shows that more than half of the accumulated displacement can be explained with simple viscoelastic decoupling.

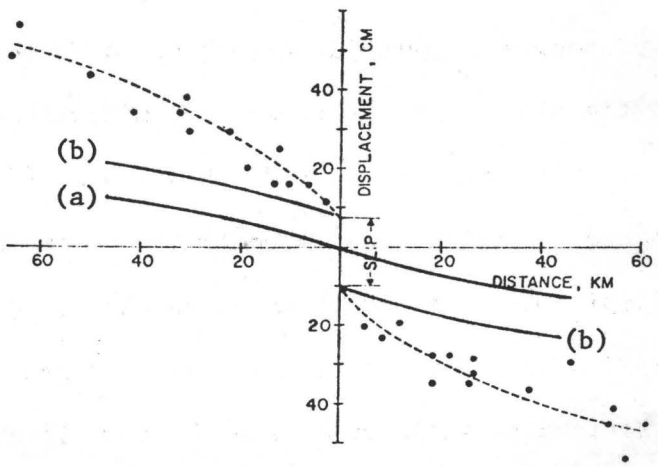


Figure 11. Displacements parallel to the Imperial fault for the period 1941-1954 [Scholz and Fitch, 1969]. Curve (a): Postseismic displacements due to viscoelastic decoupling at base of crust. Curve (b): Postseismic displacements due to decoupling plus 0.17 m of surface creep. Dashed line is for virtual dislocation model proposed by Scholz and Fitch.

Viscoelastic relaxation is just one of several modes of strain accumulation. If the average stress in the lithosphere far from the fault stays constant [Savage, 1975] an earthquake will tend to transfer stress downward by loading the asthenosphere and the lower portion of the lithosphere. As the asthenosphere relaxes the stress is transferred back to the lithosphere. The resulting strain will be roughly as shown in figures 5-8. This process has been discussed in some detail by Nur and Mavko [1974], Spence and Turcotte [1977] and Rundle and Jackson [1977]. Normally we would classify this as the postseismic phase, but it is also obviously a mechanism of strain accumulation. If the relaxation time becomes comparable to the earthquake recurrence time [Spence and Turcotte, 1977] the distinction completely disappears.

Steady aseismic slip on the lower portion of the lithosphere will also tend to concentrate strain near the upper locked portion of the fault. This is the most commonly quoted mechanism, but it can only be correctly modeled by accounting for the nature of plate motion at the lithosphere-asthenosphere boundary. Turcotte and Spence [1974] assume a two dimensional plate model for the lithosphere and calculate the surface strain rate due to steady slip on the lower portion of a long straight transform fault. The shear tractions at the base of the lithosphere and on the creeping fault are assumed to be zero. (The solution is conveniently found for the strike-slip problem by using the solution for mode III deformation of a periodic array of cracks, which automatically gives regularly spaced stress-free planes.) Computed strain rates are consistently lower (by a factor of 2) than observed values along the San Andreas fault south of San Francisco reported by Savage and Burford [1973]. This suggests that either the locked portion of the San Andreas

fault is shallower than the 5 km chosen by Turcotte and Spence or the strain measurements are too close to the actively slipping portion of the San Andreas to be explained with a two dimensional model. Thatcher [1975b] has suggested a further complication in the San Francisco area. Since about 1950, shear straining has been uniform across an 80 km wide zone east of the San Andreas near San Francisco and indicates significant slip at depth on the Calaveras fault. Farther to the northwest the strain is more concentrated near the San Andreas and is somewhat smaller. This latter region is probably more suitable for a two dimensional description.

Turcotte and Spence [1974] extend their model results qualitatively to discuss an earthquake cycle in three dimensions. Because the lower portion of the fault is stress free, the eventual earthquake rupture of the upper locked portion essentially reduces the stress on the entire vertical section of the plate boundary. There is no means for transferring stress downward to the asthenosphere or lower lithosphere. Stress can only be transferred horizontally along the fault beyond the locked ends. With this analysis the fault length q (third dimension) is included only approximately. The resulting coseismic change is similar to suddenly introducing a cut of length q through the entire plate thickness. The most important conclusion is that strain variations throughout the earthquake cycle penetrate into the plate to distances comparable to the fault length, rather than the fault depth.

Savage [1975] criticizes the model based on comparisons of preseismic, coseismic and postseismic slip. He points out that the zero stress condition on the lower fault requires an enormous amount of postseismic

slip both at the surface and at depth. He proposes instead a constant stress condition which confines strain changes close to the fault. Mavko [1977b, 1978] has shown that the differences between the Turcotte and Spence and the Savage versions of the model can be summarized in terms of boundary conditions. In two dimensions a constant stress boundary condition confines all strain changes during an earthquake cycle to within a few tens of kilometers of the fault. In contrast, a constant plate speed boundary condition results in a small but varying component of uniform shear out to great distances from the fault. Along the major active and locked regions of the San Andreas fault a two-dimensional analysis has some validity. However, the three dimensionality of the system determines the boundary conditions for the two-dimensional approximations. It appears that when adjacent portions of the fault vary from smoothly slipping to locked behavior, each section alone is best described by constant plate rate boundary condition. This suggests that shear strains on the order of 4-9 x 10^{-6} accumulated and were released out to distances of hundreds of kilometers during the 1906 San Francisco earthquake cycle.

Using this type of analysis Mavko has proposed a modification to the Turcotte and Spence [1974] model of the San Andreas. By adding a nonzero creep strength at depth on the fault a larger stress accumulates on the locked portions, allowing a greater coseismic surface slip. This can eliminate the need for postseismic surface slip to accommodate the net plate offset. The magnitude of creep strength depends on a number of factors. If the seismic surface slip is assumed to just equal the accumulated preseismic slip at the bottom of the plate, then a minimum

preseismic creep strength of about 37 bars is needed to account for 5 m of surface slip during an event like the 1906 San Francisco earthquake. However, this implies that the creep strength at the beginning of strain accumulation is greater than it is just before an earthquake. On the other hand, if the creep strength is constant during the entire period of strain accumulation, a creep strength of at least 68 bars is required, and the seismic slip is greater than the preseismic slip at the bottom of the plate.

FINAL COMMENT

Nearly all fault models are consistent with the concepts of plate tectonics and elastic rebound. Through some combination of remotely applied forces the elastic plates move relative to each other. Whether or not strain accumulates and the way it is released depends on the nature of slip at the common plate boundaries.

The most serious and fundamental problem in developing fault models is our general inability to resolve the location of sources of aseismic strain. For example the postseismic strain following the 1906 San Francisco earthquake can be explained by either aseismic slip or visco-elastic rebound. A combination of these can give an uncertainty of several 10's of kilometers in effective plate thickness. The main reason for the uncertainty is the lack of strain data at large distances from the fault.

It is probably fair to say that we are not limited by our computational ability to construct models, but instead by how clear we are about using them. Model assumptions about geometry and rheology are always necessary in lieu of reliable data. However it is essential that we always identify which model outputs result from the various assumptions. An example is the question of the scale of strain accumulation and release around strike-slip faults. While the scale clearly results from the type of stress or displacement boundary conditions built into a model, some authors have tried to associate the problem with plate thickness etc. This confusion, in part, has resulted from our extensive experience with elastic dislocation models which result in strictly local effects.

APPENDIX - THE EVOLUTION OF STATIC FAULT MODELS

Since Reid introduced the concept of elastic rebound in 1910, the greatest effort to understand and model the earthquake mechanism has focused on the coseismic or strain release portion of the earthquake cycle. There are a number of reasons for this. It is the most noticeable part and it presents the greatest hazard. The coseismic period also provides the most accessible data, conveniently fit into time windows short enough for study. Of particular interest in understanding the tectonic role of earthquakes is the very low frequency data — specifically, the net change in elastic fields that result from fault slip. In this section we summarize briefly the evolution of models for the static stress change and displacement that accompany faulting. Although the primary emphasis of this paper is on aseismic motion, the coseismic slip is by far the best known, while the distribution of slip at depth between earthquakes is very poorly resolved. Hence, static models provide a useful constraint. The following summary is roughly chronological reflecting the gradual increase in sophistication of fault models.

Early insights into the faulting process followed from the study of seismic wave arrivals. Shida [1929] made the first observation of a systematic azimuthal distribution of push-pull first motions on seismograms, today interpreted as the quadrantal nature of the source mechanism. This led to the consensus that most earthquakes (or at least their seismic radiation) are caused by abrupt stress changes at their origins in the form of shear force couples or perpendicular pairs of tensile and compressional stresses [Honda, et al., 1957]. Source models were first treated

as points in space and were later extended to include stress changes distributed on the surface of finite size spheres [Sezawa, 1936; Kawasumi and Yosiyama, 1935; Honda, 1934; Inouye, 1936]. Kasahara [1957], noting the success of the stress-release model for explaining the pattern of seismic wave arrivals, observed that permanent crustal deformations [C. Tsuboi, 1933] sometimes resembled the same push-pull pattern. Combining this with the concept of stress release by faulting or fracture, suggested by elastic rebound, he computed the displacement field corresponding to a prescribed stress drop on a vertical plane in an elastic half space. His model approximated an infinitely long strike slip fault, although the slip was only approximately zero below the prescribed depth of faulting. Perhaps the most important result was the conclusion that the width of the zone of greatest seismic strain release straddling the fault is roughly proportional to the depth of faulting, an idea earlier dismissed by Reid. With this model Kasahara inferred depths of 15, 8, and 6 km respectively for the Tango (1927), North-Izu (1930) and San Francisco (1906) earthquakes.

In a rather different approach Byerley and DeNoyer [1958] attribute the concentration of strain near the fault zone to an anomalously low rigidity. They assume that faulting penetrates a decoupled crustal plate and that the observed displacements are uniform with depth. The strain energy release, therefore, becomes proportional to the plate thickness. Comparing this with the observed distribution of wave energy they infer the plate thickness and, hence, the fault depth. The notion of a soft (low rigidity) fault zone extending to the Moho has also been suggested by Benioff [1951, 1955]. Although it is generally believed that major faults (or perhaps concentrated shear zones at depth) do penetrate the

lithosphere, seismic slip is often confined to the brittle portion of the upper crust(15-20 km). Furthermore, the short term mechanical response of the earth tends to be elastic, even in the presence of horizontal viscous zones that may gradually decouple the upper plate. Hence the approximation of uniform strain with depth can only be very rough.

Further refinements in computing static stress and strain fields have fallen into two types: crack models and dislocation models. The crack model is based on a prescribed stress change during faulting. This is the case studied by Kasahara [1957]. In contrast, the dislocation model is based on a prescribed fault slip. An advantage of the dislocation approach over the prescribed stress approach is the ability to compute the stress and displacement fields due to well defined, arbitrarily shaped faults with arbitrary slip distributions. (Recall that Kasahara [1957] treated an infinitely long fault with slip only approximately zero outside the slip zone.) The crack problem, on the other hand involves mixed stress and slip boundary conditions in the plane of the fault which is generally more difficult to solve. Of course, both stress and slip changes accompany faulting and the two descriptions are equivalent, but the dislocation approach has been much more popular because of the computational flexibility it allows.

Steketee [1958a, b] has shown that if we approximate a fault as a discrete surface of discontinuity (or dislocation) in an otherwise elastic half space, the resulting displacement field is given by an integral over the fault surface of point nuclei or strain [Love, 1944] weighted with the local value of slip. He presents the method for obtaining the six kernals corresponding to the general dislocation surface in a half space, and derives the one corresponding to strike slip motion. In an

earlier study Rochester [1956] had used dislocation theory to solve for three dimensional slip distributions, but only discussed in detail the infinitely long fault. Other authors who had studied or suggested internal nuclei of strain to describe the static problem were Sezawa [1929], Whipple [1936], Soeda [1944], Mindlin and Cheng [1950], Yamakawa [1955] and Vvedenskaya [1956]. A similar problem due to surface loading was treated by Honda and Muira [1935].

Chinnery [1961] integrated Steketee's expressions to find analytic expressions for the surface displacements due to a finite size, rectangular, vertical strike-slip fault. In this and a number of subsequent papers the integration was simplified by assuming uniform slip on the fault. He refined the relation between the scale of surface deformation (both horizontal and vertical components) and fault depth, showing the effect of finite length. Although the qualitative features of deformation agreed with Kasahara's [1957] Chinnery's inferred depths for the Tango (1927), North Izu (1930) and San Francisco (1906) earthquakes were systematically shallower, probably because Kasahara's slip tapered to zero at depth while Chinnery's went discontinuously to zero.

To calculate the stress field around a rectangular strike slip fault Chinnery [1963] differentiated his earlier expressions for displacement. He later [Chinnery, 1964] used these results to evaluate the stresses at the center of rectangular strike slip faults. Analyzing the displacements of the San Francisco (1906), Tango (1927), Idu (1930), Imperial Valley (1940) and Fairview Peak (1954) earthquakes, he concluded that the stress drop and hence fault strength prior to rupture is on the order of a few tens of bars. (It was assumed that complete stress release occurred.) This value was several orders of magnitude less than the known typical strength of rock. Chinnery [1965] also discussed the vertical component

of surface displacement accompanying strike slip faulting, suggesting that permanent topographic features along faults (eg. mountains, basins) could be the accumulation over many earthquakes (and many km of offset) of the vertical deformation suggested by the (infinitesimal strain) elastic model.

Maruyama [1964] extended the work of Steketee and Chinnery, giving the expressions for all six of the Green's functions for the displacement and stress field due to a general dislocation surface in a half space. He integrated these to obtain analytic expressions for the displacement fields for vertical and horizontal faults, each with three choices of uniform Burgers vector. He also solved for the deformation due to an arbitrarily dipping rectangular fault.

Press [1965] used the method of Steketee [1958a], Maruyama [1964] and Chinnery [1961] combined with catalogs of nuclei of strain by Mindlin and Cheng [1950] to study the displacement, strains, and tilts resulting from vertical strike-slip and dip-slip rectangular faults. He was mostly concerned with the deformation at teleseismic distances (at distances 10 times the fault length). Comparisons of computed strains with observed values in Hawaii and California from the 1964 Alaskan earthquake were reasonable and suggested that far field studies of static source parameters are feasible. It should be noted, however, that Press used the complimentary fault plane (he used a vertical fault to approximate a near vertical reverse fault) to the one generally accepted today (low angle thrust).

Savage and Hastie [1966] integrated Maruyama's expressions to find the displacement field due to a rectangular dip slip fault with arbitrary dip. They compare the near field vertical displacement for a low angle thrust model of the Alaskan (1964) earthquake as suggested by Plafker

[1965] with the near vertical reverse fault studied by Press [1965]. Though many qualitative features of the two models are the same, the case for thrust faulting is stronger. Similar studies of the Fairview Peak, Nevada (1954) and Hebgen Lake, Montana (1959) earthquakes were made.

The next step in sophistication of elastic fault models was to explore the effect of a variable distribution of slip over the fault surface. The formalism of Steketee [1958a] and Maruyama [1964] allows for general slip in a half space, but numerical evaluation of the integral is usually necessary. Chinnery and Petrak [1967] explored the effect of smoothly tapering the slip at the edges of strike slip faults. Tapering removes the singularities present in earlier uniform slip models. They concluded that the pattern of surface displacement is very similar to that for uniform slip — hence insensitive to details of slip deep on the fault. Weertman [1965] makes a similar argument that the surface displacements are insensitive to the details of stress drop on the fault.

A more thorough study of the variation of fault slip is possible using two dimensional models. Starr [1928], studying the mechanisms of fracture in crystals, examined the plane strain deformation of a flat, two dimensional elliptical crack in shear. Because the crack surfaces are stress free, introducing the crack into a uniform stress field results in a uniform stress drop over the crack face with an increase in stress (actually a stress singularity) outside the crack tips. Knopoff [1958] studied the same crack in antiplane shear. Many features of the two models are essentially the same — an elliptical distribution of slip, with stress singularities outside the crack tips. (For a given stress the in plane slip is diminished by the factor $(1 - \nu)$ where ν is Poisson's

ratio, from the antiplane case.) A convenient feature of the antiplane problem is that the perpendicular bisecting plane of the crack is stress free. This is true for any symmetric distribution of antiplane slip. Hence the same solution holds for a very long vertical strike-slip fault intersecting the free surface in an elastic half space. The solution of Starr applied rigorously only to a fault in an infinite medium, though may be used to approximate a deeply buried fault in a half space.

In order to decrease the arbitrariness of specifying slip or stress drop Weertman [1964, 1965], Walsh [1968], and Berg [1967] study the infinitely long strike slip fault with fault friction increasing linearly with depth, proportional to the lithostatic stress. By specifying the applied tectonic stress, the friction law, and fault dip angle a unique equilibrium depth of faulting and slip distribution can be found. However, the details of faulting become dependent on previous slip episodes since each earthquake represents a slip relative to the previous equilibrium configuration.

A natural approach to solving two dimensional faulting problems is the use of elastic screw and edge dislocations [Bilby and Eshelby, 1968] or line dislocations. These represent spatial step functions in slip — or slip on a half plane — in an elastic medium. (Here "slip" refers to a discontinuity in the displacement — exactly in the sense of the usual meaning of "slip" on a fault.) A continuous distribution of slip corresponds simply to a continuous distribution of infinitesimal dislocations, the density of dislocations given by the derivative of the slip function. Hence, the elastic stress (and strain) field of a single dislocation is the Green's function or spatial impulse response for the derivative of slip, the fields due to complex patterns of slip, obtained by superposition. Weertman [1964] studies the effect of applied

and frictional stresses on faults on the resulting slip, using dislocation theory as a means to solve the elastic problem. Mavko and Nur (1977a) discuss the use of dislocation theory to solve for the deformation of cracks (or faults) in elastic solids. They discuss a method for obtaining analytic expressions for slip resulting from arbitrary continuous stress distributions using polynomials.

Both the dislocation and crack models of faulting have been extended to elastic earth models more complex than homogeneous half spaces. Kasahara [1964] solved for the deformation resulting from a prescribed stress drop on an infinitely long strike slip fault at the top of a half space underlying a surface layer. Oda and Hirasawa [1976] solve a similar problem for a crack at an arbitrary distance from the interface between two half spaces, analyzing the slip and strain energy change as a function of shear moduli and fault dimension.

The stress and strain solutions for a screw dislocation near planar boundaries can be solved using the method of images. Chou [1966] gives the stress field resulting from a screw dislocation in or near a layer imbedded between two half-spaces. By giving one half-space zero rigidity the model simulates a layer with a free surface overlying a half-space. Rybicki [1971] solves for the stress and displacements from a pair of screw dislocations (simulating a long strike slip fault with constant slip) in and near a layer overlying a half-space. Modeling the Parkfield (1966) earthquake as a buried fault in a half space underlying a soft surface layer Rybicki finds the surface displacements to be more localized than for an identical fault in a half space. In addition, the stress concentration near the top edge of the fault is enhanced by the soft layer relative to the lower edge, suggesting a mechanism for the greater concentration of

observed aftershocks near the top edge. Chinney and Jovanovich [1972] solve the screw dislocation problem for two layers over a half-space. They find that low rigidity layers beneath a rigid crust or lithosphere will enhance the surface displacements by decoupling the upper layer from the half-space.

Barnett [1972], Canales [1975], and Mahrer and Nur [1977] have studied the case of a strike slip fault in a half space with a continuously varying shear modulus.

The problem of an edge dislocation near an interface cannot be treated using simple images. Hence the modeling of very long dip-slip faults can be more cumbersome. The problem of an edge dislocation near a circular inclusion was studied by Dundurs and Mura [1964]. Weeks et al. [1968] solve for the edge dislocation near a surface layer and Lee and Dundurs [1973[treat an edge in thin layer.

Singh [1970] presents a method for finding the stress and displacement due to arbitrary nuclei of strain in an arbitrary multilayered half space.

ACKNOWLEDGEMENTS

This work was supported by a National Science Foundation Postdoctoral fellowship as well as grants/contracts GA 36135X and DES 75-04874 from the National Science Foundation (Division of Earth Sciences); 14-08-0001-15885 from the U.S. Geological Survey; and by E(04-3) 326-PA#45 from the U.S. Energy Rescurce and Development Administration.

References

1. Allen, C.R., and S.W. Smith, Parkfield earthquakes of June 27-29, Monterey and San Luis Obispo Counties, California. Pre-earthquake and post-earthquake surficial displacements, Bull. Seismol. Soc. Am., 56, 966-967, 1966.
2. Anderson, D.L., The Anelasticity of the Mantle, Geophys. J. R. astr. Soc., 14, 135-164, 1967.
3. ____, The San Andreas Fault, Scientific American, 225, 5, 52-66, 1971.
4. ____, Accelerated Plate Tectonics, Science, 187, 1077-1079, 1975.
5. Barazangi, M., W. Pennington, and B. Isacks, Global Study of Seismic Wave Attenuation in the Upper Mantle Behind Island Arcs Using pP Waves, J. Geophys. Res., 80, 1079-1092, 1975.
6. Barker, T., Quasi-Static Motions Near the San Andreas Fault Zone, Geophys, J.R. astr. Soc., 45, 689-706, 1976.
7. Barnett, D.M., On the Screw Dislocation in an Inhomogeneous elastic Medium: The Case of Continuously Varying Elastic Moduli, Int. J. Solids Structures, 8, 651-660, 1972.
8. _____, and L.B. Freund, An Estimate of Strike Slip Fault Friction Stress and Fault Depth from Surface Displacement Data, BSSA, 65, 5, 1259-1266, 1975.
9. Benioff, H., Mechanism of Earthquake Generation, Bull. Geol. Soc. Amer., 62, 1526, 1951.
10. _____, Mechanism and Strain characteristics of the White Wolf Fault as Indicated by the Aftershock Sequence, Cal. Div. Mines, Bull. 171, 199-202, 1955.

11. Berg, C. A., A Note on the Mechanics of Seismic Faulting, Geophys. J., 14, 89-99, 1967.

12. Bilby, B.A. and J.D. Eshelby, Dislocations and the Theory of Fracture, In: Fracture, An Advanced Treatise, edited by H. Liebowitz, Vol. I, p.99, Academic Press, New York, 1968.

13. Biot, M.A., General Theory of Three Dimensional Consolidation, J. Appl. Phys., 12, 155-164, 1941.

14. Bott, M.H.P. and D.S. Dean, Stress Diffusion from Plate Boundaries, Nature, 243, 339-341, 1973.

15. Brown, L.D., R.E. Reilinger, S.R. Holdahl, and E.I. Balazs, Post-Seismic Crustal uplift Near Anchorage Alaska, J. Geophys. Res., 82, 23, 3369-3378, 1977.

16. Brune, J.N., Current Status of Understanding Quasi-Permanent Fields Associated with Earthquakes, EOS, 55, 9, 820-827, 1974.

17. Brune, J.N. and C.R. Allen, A Low Stress-Drop, Low-Magnitude Earthquake with Surface Faulting: The Imperial, California, Earthquake of March 4, 1966, BSSA, 57, 501-514, 1967.

18. Byerly, P. and J. DeNoyer, Energy in Earthquakes as Computed from Geodetic Observations, In: Contributions in Geophysics in Honor of Beno Gutenberg, H. Benioff, et.al. (eds.) New York, Pergamon Press, 17-35, 1958.

19. Canales, L., Inversion of Realistic Fault Models, Ph.D. Thesis, Stanford University, 1975.

20. Carter, N.L. Steady State Flow of Rocks, Rev. Geophys. and Space Phys. 14, 3, 301-360, 1976.

21. Cathles, L.M., The Viscosity of the Earth's Mantle, p. 156, Princeton Univ. Press, Princeton, N.J., 1975.

22. Chinnery, M.A., The Vertical Displacements Associated with Transcurrent Faulting, J. Geophys. Res., 70, 18, 4627-4632, 1965.

23. ____, The Deformation of the Ground Around Surface Faults, BSSA, 51, 3, 355-372, 1961.

24. ____, The Stress Changes that Accompany Strike Slip Faulting, BSSA, 53, 5, 921-932, 1963.

25. ____, The Strength of the Earth's Crust Under Horizontal Shear Stress, J. Geophys. Res., 69, 10, 2085-2089, 1964.

26. ____, Earthquake Displacement Fields, In: Earthquake Displacement Fields and the Rotation of the Earth, Reidel Company, Dordrecht-Holland, L. Mansinha et al. editors, 17-38, 1970.

27. ____, and D.B. Jovanovich, Effect of Earth Layering on Earthquake Displacement Fields, BSSA, 62, 6, 1629-1639, 1972.

28. _____,and J.A. Petrak, The Dislocation Fault Model with a Variable Discontinuity, Tectonophysics, 5, 6, 513-529, 1967.

29. Chou, Y.T., Screw Dislocations in and Near Lamellar Inclusions, Bain Laboratory, U.S. Steel Corp. Res. Center, No. 1276, May 1966.

30. Dunbar, W.S., The Determination of Fault Models from Geodetic Data, Ph.D. Thesis, Stanford University, 1977.

31. Dundurs, J. and T. Mura, Interaction Between an Edge Dislocation and a Circular Inclusion, J. Mech. Phys. Solids, 12,177-189, 1964.

32. Elasser, W.M.,Convection and Stress Propagation in the Upper Mantle In: The Application of Modern Physics to the Earth and Planetary Interiors, edited by S.K. Runcorn, pp..223-246, John Wiley, New York, 1969.

33. _____, Two-Layer Model of Upper-Mantle Circulation, J. Geophys. Res., 76, 20, 4744-4753, 1971.

34. Fitch, T. and C.H. Scholz, Mechanism of Underthrusting in Southwest Japan: A Model of Convergent Plate Interactions, J. Geophys. Res., 80, 11, 1444-1447, 1971.

35. Freund, L.B. and D.M. Barnett, A Two Dimensional Analysis of Surface Deformation Due to Dip Slip Faulting, BSSA, 66, 3, 667-676, 1976.

36. Gordon, R.B., Diffusion Creep in the Earth's Mantle, J. Geophys. Res., 70, 10, 2413-2418, 1965.

37. Hadley, D. and H. Kanamori, Seismic Structure of the Transverse Ranges, California, G.S.A. Bull, 88, 10, 1469-1478, 1977.

38. Heard, H.C. Comparison of the Flow Properties of Rocks at Crustal Condition, Phil. Trans. R. Soc. Lond. A. 283, 173-186, 1976.

39. Herrin, E., A Comparative Study of Upper Mantle Models: Canadian Shield and Basin and Range Province, In: The Nature of the Solid Earth, ed., E. Robertson, 216-34, New York: Mc Graw, 1972.

40. Honda, H., Geophys. Mag., 8, 153, 1934.

41. _____, A. Masatsuka, and K. Emura, Sci. Rep., Tohoku Univ., Ser. 5, 8, 30, 1957.

42. _____, and T. Muira, On the Strain Produced in a Semi-Infinite Elastic Solid by Statical Surface Force, with Some Applications to Seismology, Geophys. Mag., 9, 61-81, 1935.

43. Inouye, W., Bull. Earthq. Res. Inst., 14, 582, 1936.

44. Kanamori, H., Mantle Beneath the Japan Arc, Phys. Earth Planet. Interiors, 3, 475-483, 1970.

45. ____, Great Earthquakes at Island Arcs and the Lithosphere, Tectonophysics, 12, 187-198, 1971.

46. ____,Mode of Strain Release Associated with Major Earthquakes in Japan In: Annual Review of Earth and Planetary Sciences, 1, ed. by F.A. Donath, F.G. Stehli, and G.W. Wetherill, Annual Reviews, Palo Alto, Ca., 1973.

47. _____, and D.L. Anderson, Amplitude of the Earth's Free Oscillations and Long Period Characteristics of the Earthquake Source, J. Geophys. Res., 80, 8, 1075-1078, 1975.

48. ____, and J. J. Cipar, Focal Processes of the Great Chilean Earthquake, May 22, 1960, Phys. Earth Planet. Interiors, 9, 128-136, 1974.

49. Kasahara, K., The Nature of Seismic Origins as Inferred from Seismological and Geodetic Observations (1), Bull. Earthq. Res. Inst., Tokyo Univ., 35, 473-532, 1957.

50. ____, Physical Condition of Earthquake Faults as Deduced from Geodetic Data, Bull. Earthquake Res. Inst. Tokyo Univ., 36, 455-464, 1958.

51. ____,A Strike Slip Fault buried in a Layered Medium, Bull. Earthq. Res. Inst., 42, 609-619, 1964.

52. _____, Aseismic Faulting Following the 1973 Nemuro-Oki Earthquake, Hokkaido, Japan (a possibility), Pageoph, 113, 127-139, 1975.

53. Kawasumi, H. and R. Yosiyama, Bull. Earthq. Res. Insti. 13, 496, 1935.

54. Ke, T.S., Experimental Evidence of the Viscous Behavior of Grain Boundaries in Metals, Phys. Rev., 71, 533, 1947.

55. Knopoff, L., Energy Release in Earthquakes, Geophys. J., 1, 44-52, 1958.

56. Lachenbruch, A.H. and J.H. Sass, Thermo-Mechanical Aspects of the San Andreas Fault System In: Proceedings of the Conference on Tectonic Problems of the San Andreas Fault System, eds. R.L. Kovach and A. Nur, Stanford Univ. Publications, Geological Sciences, 13, Stanford, Ca., 1973.

57. Lee, M.-S. and J. Dundurs, Edge Dislocation in a Surface Layer Int. J. Engng. Sci., 11, 87-94, 1973.

58. Lensen, G., Elastic and non-elastic surface deformation in New Zealand, Bull. N.Z. Soc. Earthquake Eng., 3, 131-143, 1970.

59. LePichon, X, J. Francheteau, and J. Bonnin, Plate Tectonics, Elsevier, New York, 1973.

60. Love, A.E.H., A Treatise on the Mathematical Theory of Elasticity, Dover, Pub., New York, 1944.

61. Mahrer, D.K. and A. Nur, Two-Dimensional Strike-slip Model in a Variable Modulus Crust, (Abstract) EOS, 58, 12, 1227, 1977.

62. Maruyama, T.,Statical Elastic Dislocations in an Infinite and Semi-infinite Medium, Bull. Earthq. Res. Inst., 42, 289-368, 1964.

63. Matuzawa, T., Study of Earthquakes, Uno Shoten, Tokyo, 1964.

64. Mavko, G., Time Dependent Fault Mechanics and Wave Propagation in Rocks, Ph.D. Thesis, Stanford Univ., 1977a.

65. ____, Mechanical Constraints on Models of Strain Accumulation Along the San Andreas Fault, abstract, in EOS, 58, 6, 494, 1977b.

66. ____, The Mechanics of Strain Accumulation and Release on a Strike-slip Fault, submitted to J. Geophys. Res., 1978.

67. ____, and A. Nur, Melt Squirt in the Asthenosphere, J. Geophys. Res., 80, 11, 1444-1447, 1975.

68. _____, The Effect of Non-elliptical Cracks on the Compressibility of Rocks, submitted to J. Geophys. Res. 1977a.

69. _____, Wave Attenuation in Partially Saturated Rocks, Submitted to Geophysics, 1977b.

70. Melosh, H.J., Nonlinear Stress propagation in the Earth's Upper Mantle, J. Geophys. Res., 81, 32, 5621-5632, 1976.

71. _____, Shear Stress on the Base of a Lithospheric Plate, Pageoph, 115, 1/2, 429-439, 1977.

72. Mescherikov, J.A., Recent Crustal Movements in Seismic Regions: Geodetic and Geomorphic Data, Tectonophysics, 6, 29-39, 1968.

73. Mindlin, R.D., and D.H. Cheng, Nuclei of Strain in the Semi-Infinite Solid, J. Applied Phys. 21, 926-930, 1950.

74. Mogi, K., J. Phys. Earth, 16, 30, 1968a.

75. ____, Bull. Earthq. Res. Inst., 46, 53, 1968b.

76. Nakamura, K. and Tsuneishi, Ground Cracks at Matsushiro Probably of Strike-slip Fault Origin, Bull. Earthquake Res. Inst. Univ. Tokyo, 45, 417-472, 1967.

77. Nason, R. and J. Weertman, A Dislocation Theory Analysis of Fault Creep Events, J. Geophys. Res., 78, 32, 7745-7751, 1973.

78. Nur, A., Role of Pore Fluids in Faulting, Phil. Trans. R. Soc. Lond. A., 274, 297-304, 1973.

79. ____, and J. R. Booker, Aftershocks Caused by Pore Fluid Flow?, Science, 175, 885-887, 1972.

80. ____, and G. Mavko, Postseismic Viscoelastic Rebound, Science, 183, 4121, 204-206, 1974.

81. Oda, H. and T. Hirasawa, The effect of medium heterogeneity upon the static deformation due to a longitudinal shear crack, Sci. Rep. Tohoku Univ., Ser. 5, Geophysics, 23, 3-4, pp. 115-132, 1976.

82. Plafker, G., Tectonic deformation associated with the 1964 Alaska earthquake, Science, 148, 1675-1687, 1965.

83. Post, R.L. and D.T. Griggs, The earth's mantle: evidence of non-Newtonian flow, Science, 181, 1242-1244, 1973.

84. Press, F., Displacements, strains and tilts at teleseismic distances, J. Geophys. Res., 70, 10, 2395-2412, 1965.

85. Reid, H.F., The mechanics of the earthquake, in: The California Earthquake of April 18, 1906. Rept. State Earthquake Invest. Comm., Carnegie Inst., Washington, D.C., 1910.

86. Rochester, M.G., The application of dislocation theory to fracture of the earth's crust, M.A. Thesis, Univ. of Toronto, 1956.

87. Rundle, J.B. and D.D. Jackson, A viscoelastic relaxation model for post-seismic deformation from the San Francisco earthquake of 1906, Pageoph., 115, 401-411, 1977.

88. Rybicki, K., The elastic residual field of a very long strike-slip fault in the presence of a discontinuity, BSSA, 61, 1, 79-92, 1971.

89. Savage, J.C., Comment on 'an analysis of strain accumulation on a strike slip fault' by D.L. Turcotte and D.A. Spence, J. Geophys. Res., 80, 29, 4111-4114, 1975.

90. Savage, J.C. and R.O. Burford, Accumulation of tectonic strain in California, BSSA, 60,6, 1877-1896, 1970.

91. Savage, J.C., and R.O. Burford, Geodetic determination of relative

plate motion in central California, J. Geophys. Res., 78, 832-845, 1973.

92. Savage, J.C. and L.M. Hastie, Surface deformation associated with dip-slip faulting, J. Geophys. Res., 71, 20, 4897-4904, 1966.

93. Scholz, C.H., Crustal movements in tectonic areas, in: E.F. Savarensky and T. Rikitake (editors) Forerunners of Strong Earthquakes, Tectonophysics, 14 (3/4), 201-217, 1972.

94. Scholz, C.H. and T.J. Fitch, Strain accumulation along the San Andreas fault, J. Geophys. Res., 74, 27, 6649-6666, 1969.

95. Scholz, C.H., M. Wyss and S.W. Smith, Seismic and aseismic slip on the San Andreas fault, J. Geophys. Res., 74, 8, 2049-2069, 1969.

96. Sezawa, K., The tilting of the surface of a semi-infinite solid due to internal nuclei of strain, Bull. Earthq. Res. Inst., 7, 1-14, 1929.

97. Sezawa, K., Bull.Earthq. Res. Inst. 14, 10, 1936.

98. Shida, J., Tôyo-gakugei Zasshi, 45, 275, 1929 (in Japanese).

99. Singh, S.J., Static deformation of a multilayered half space by internal sources, J. Geophys. Res., 75, 17, 3257-3263, 1970.

100. Smith, A.T., Time-dependent strain accumulation and release at island arcs, abstract in EOS, 55, 4, 1974.

101. Smith, A.T., Post-seismic fault creep on stress-relaxation following the 1946 Nankaido earthquake?,abstract in EOS, 57, 4, 1976.

102. Smith, S.W., and M. Wyss, Displacement on the San Andreas fault initiated by the 1966 Parfield earthquake, BSSA, 58, 1955-1974, 1968.

103. Soeda, K., On the deformation produced in a semi-infinite elastic

solid by an internal source of stress, Kenshin Ziho, 13, 263-291, 1942 (in Japanese).

104. Spence, D.A. and D.L. Turcotte, Stress relaxation at plate margins, abstract in EOS, 58, 6, 494, 1977.

105. Starr, A.T., Slip in a crystal and rupture in a solid due to shear, Proc. Camb. Phil. Soc., 24, 489-500, 1928.

106. Steketee, J.A. On Volterra's dislocations in a semi-infinite medium, Can. J. Phys., 36, 192-205, 1958a.

107. --- , Some geophysical applications of the elasticity theory of dislocations, Can. J. Phys., 36, 1168-1197, 1958b.

108. Stesky, R.M., Steady-state creep law for frictional sliding at high temperature and pressure, abstract in EOS, 55, 4, 1974.

109. Thatcher, W., Strain accumulation and release mechanism of the 1906 San Francisco earthquake, J. Geophys. Res., 80, 35, 4862-4872, 1975a.

110. Thatcher, W., Strain accumulation on the northern San Andreas fault zone since 1906, J. Geophys. Res., 80, 35, 4873-4880, 1975b.

111. Thompson, G.A. and D.B. Burke, Regional geophysics of the Basin and Range province, in: Annual Review of Earth and Planetary Sciences, vol. 2, ed. by F.A. Donath, F.G. Stehli, and G.W. Wetherill, Annual Reviews, Palo Alto, Ca., 1974.

112. Tsuboi, C., Jap. Journ. Astro. Geophys., 10, 93, 1933.

113. Turcotte, D.L. and D.A. Spence, An analysis of strain accumulation on a strike slip fault, J. Geophys. Res., 79, 29, 4407-4412, 1974.

114. Vvedenskaya, A.V., The determination of displacement fields by means of dislocation theory, Bull. Acad. Sci. USSR, Geophys. Ser., English Transl., 3, 277-284, 1956.

115. Walcott, R.I., Structure of the earth from glacioisostatic rebound, in Annual Review of Earth and Planetary Sciences, vol. 1, edited by F.A. Donath, F.G. Stehli, and G.W. Wetherill, Annual Reviews, Palo Alto, Ca., 1973.

116. Walsh, J.B., The effect of cracks on the compressibility of rock, J. Geophys. Res., 70, 2, 381, 1965a.

117. Walsh, J.B., The effect of cracks on the uniaxial elastic compression of rocks, J. Geophys. Res., 70, 2, 399, 1965b.

118. Walsh, J.B., Seismic wave attenuation in rock due to friction, J. Geophys. Res., 71, 2591-2599, 1966.

119. Walsh, J.B., Mechanics of strike-slip faulting with fri-tion, J. Geophys. Res., 73, 761-776, 1968.

120. Walsh, J.B., A new analysis of attenuation in partially melted rock, J. Geophys. Res., 74, 4333, 1969.

121. Walsh, J.B., Dip angle of faults as calculated from surface deformation, J. Geophys. Res., 74, 8, 2070-2080, 1969.

122. Weeks, R., J. Dundurs, and M. Stippes, Exact analysis of an edge dislocation near a surface layer, Int. J. Engr. Sci., 6, 365-372, 1968.

123. Weertman, J., Continuum distribution of dislocations on faults with finite friction, BSSA, 54, 1035-1058, 1964.

124. Weertman, J., Relationship between displacements on a free surface and the stress on a fault, BSSA, 55, 945-953, 1965.

solid by an internal source of stress, Kenshin Ziho, 13, 263-291, 1942 (in Japanese).

104. Spence, D.A. and D.L. Turcotte, Stress relaxation at plate margins, abstract in EOS, 58, 6, 494, 1977.

105. Starr, A.T., Slip in a crystal and rupture in a solid due to shear, Proc. Camb. Phil. Soc., 24, 489-500, 1928.

106. Steketee, J.A. On Volterra's dislocations in a semi-infinite medium, Can. J. Phys., 36, 192-205, 1958a.

107. --- , Some geophysical applications of the elasticity theory of dislocations, Can. J. Phys., 36, 1168-1197, 1958b.

108. Stesky, R.M., Steady-state creep law for frictional sliding at high temperature and pressure, abstract in EOS, 55, 4, 1974.

109. Thatcher, W., Strain accumulation and release mechanism of the 1906 San Francisco earthquake, J. Geophys. Res., 80, 35, 4862-4872, 1975a.

110. Thatcher, W., Strain accumulation on the northern San Andreas fault zone since 1906, J. Geophys. Res., 80, 35, 4873-4880, 1975b.

111. Thompson, G.A. and D.B. Burke, Regional geophysics of the Basin and Range province, in: Annual Review of Earth and Planetary Sciences, vol. 2, ed. by F.A. Donath, F.G. Stehli, and G.W. Wetherill, Annual Reviews, Palo Alto, Ca., 1974.

112. Tsuboi, C., Jap. Journ. Astro. Geophys., 10, 93, 1933.

113. Turcotte, D.L. and D.A. Spence, An analysis of strain accumulation on a strike slip fault, J. Geophys. Res., 79, 29, 4407-4412, 1974.

114. Vvedenskaya, A.V., The determination of displacement fields by means of dislocation theory, Bull. Acad. Sci. USSR, Geophys. Ser., English Transl., 3, 277-284, 1956.

115. Walcott, R.I., Structure of the earth from glacioisostatic rebound, in Annual Review of Earth and Planetary Sciences, vol. 1, edited by F.A. Donath, F.G. Stehli, and G.W. Wetherill, Annual Reviews, Palo Alto, Ca., 1973.

116. Walsh, J.B., The effect of cracks on the compressibility of rock, J. Geophys. Res., 70, 2, 381, 1965a.

117. Walsh, J.B., The effect of cracks on the uniaxial elastic compression of rocks, J. Geophys. Res., 70, 2, 399, 1965b.

118. Walsh, J.B., Seismic wave attenuation in rock due to friction, J. Geophys. Res., 71, 2591-2599, 1966.

119. Walsh, J.B., Mechanics of strike-slip faulting with fri-tion, J. Geophys. Res., 73, 761-776, 1968.

120. Walsh, J.B., A new analysis of attenuation in partially melted rock, J. Geophys. Res., 74, 4333, 1969.

121. Walsh, J.B., Dip angle of faults as calculated from surface deformation, J. Geophys. Res., 74, 8, 2070-2080, 1969.

122. Weeks, R., J. Dundurs, and M. Stippes, Exact analysis of an edge dislocation near a surface layer, Int. J. Engr. Sci., 6, 365-372, 1968.

123. Weertman, J., Continuum distribution of dislocations on faults with finite friction, BSSA, 54, 1035-1058, 1964.

124. Weertman, J., Relationship between displacements on a free surface and the stress on a fault, BSSA, 55, 945-953, 1965.

125. Weertman, J. and J.R. Weertman, High temperature creep of rock and mantle viscosity,In:Annual Review of Earth and Planetary Sciences, vol. 3, edited by F.A. Donath, F.G. Stehli, and G.W. Wetherill, Annual Reviews, Palo Alto, Ca., 1975.

126. Whipple, F.J.W. , On the theory of the strain in an elastic solid bounded by a plane where there is a nucleus of strain at an internal point; and on the relation of the theory to seismology, Mon. Not. Roy. Astron. Soc. Geophys. Supple., 3, 380-388, 1936.

127. Yamakawa, N., On the strain produced in a semi-infinite elastic solid by an interior source of stress, Zisin, 8, 84-98, 1955 (in Japanese)

128. Zener, C., Elasticity and Anelasticity of Metals, Univ. of Chicago Press, Chicago, 1948.

OBSERVATIONS RELEVANT TO SEISMIC DRIVING STRESS, STRESS DROP, AND RADIATION EFFICIENCY

BY

Art McGarr
U. S. Geological Survey
345 Middlefield Road
Menlo Park, CA 94025

S. M. Spottiswoode and N. C. Gay
Bernard Price Institute of Geophysical Research
University of the Witwatersrand
Johannesburg, South Africa

W. D. Ortlepp
Rand Mines, Ltd.
Corner House
Johannesburg, South Africa

ABSTRACT

Observations of the effects of tremors in deep gold mines suggest that (1) seismic stress drops are typically at least an order of magnitude less than the ambient shear stress driving the faulting, (2) the seismic failure associated with a tremor is very inhomogeneous within the source region of the tremor, (3) nearly all of the elastic energy released during a tremor is consumed in creating the comminuted fault zone with generally less than 1% of the available energy accounted for by the seismic radiation, and (4) the available data on shear stress in the crust indicates that natural tectonic shocks are probably similar in these respects to mine tremors.

INTRODUCTION

Currently, the most important outstanding questions in seismology concern the relationship of earthquake source properties, deduced from the analysis of seismograms, to the actual faliure process and the conditions in the source region. For example, how is the stress drop inferred from the seismic radiation related to the stresses driving the seismic failure? Does failure associated with earthquakes bear any resemblance to brittle failure in laboratory samples? These questions must be answered if any genuine progress is to be made in understanding the mechanics governing earthquakes and faults.

Observation since the 1940's of tremors in the vicinity of the deep gold mines of the Witwatersrand in South Africa have indicated that the tremors are identical to small to medium sized earthquakes and yet are clearly related to the mining activity (e.g. Gane et al., 1946; Cook, 1963). The high level of seismicity and the localization of this seismicity to regions of active mining make the deep mines an excellent laboratory to study earthquakes in detail at

close hand. The tremors occur in rock with known mechanical properties and in response to well-determined stress and deformation fields (Cook, 1963). The high level of seismicity and the localization of this seismicity to regions of active mining make the deep mines an excellent laboratory to study earthquakes in detail at close hand. The tremors occur in rock with known mechanical properties and in response to well-determined stress and deformation fields (Cook et al., 1966; McGarr et al., 1975; McGarr, 1976). Perhaps most importantly, the effects of some of the tremors are accessible to direct underground observations so that the failure process can be observed and compared in detail to laboratory measurements.

This report first reviews instrumental observations of tremors and rock conditions in the E.R.P.M. (East Rand Proprietary Mines) Mine, near Johannesburg, with emphasis on results related to stress and then attempts to relate these observations to recent underground observations and laboratory analyses of faulting associated with medium to large tremors. The primary intent here is to try to relate in as much detail as possible, the failure observed in the laboratory to that generating mine tremors or earthquakes. Perhaps the most important result to be inferred from the observations, to be described, is that there is no essential difference between the failure of dry brittle rock in the laboratory and rock within the source region of a tremor. This finding reduces the class of possible models of earthquake mechanics considerably because we know that much of the seismogenic region of E.R.P.M. has no ground water (e.g. McGarr et al., 1975) and so the ambiguities associated with possible pore pressure effects do not arise.

SOURCE PARAMETERS AND INFERRED STRESS DROPS

Spottiswoode and McGarr (1975) reported on seismic moments and source dimen-

sions for a suite of tremors with magnitudes ranging from 0 to 3. the tremors occurred at typical depths of about 3 km and were recorded on magnetic tape at a three-component seismograph station almost directly above the source region. Both the P and the S waves were analyzed in terms of Brune's (1970,) 1971) model, or modifications of it for P waves (Hanks and Wyss, 1972) and the results are summarized in Figure 1. We see that nearly all of the data fall between the lines of constant stress drop (Hanks and Thatcher, 1972) of 5 and 50 bars. Thus, there does not appear to be any substantial difference between stress drops observed for tremors and natural earthquakes, which have stress drops almost invariably in the range of 1 to 100 bars regardless of moment or magnitude (e.g. Hanks, 1977). Whatever problem there is in relating the apparently low earthquake stress-drops to the high stress drops observed in the laboratory also exists for mine tremors.

AMBIENT STATE OF STRESS

In situ stress measurements made in the deep mines of the Witwatersrand were reported by Gay (1972, 1975) and McGarr and Gay (1977) and are summarized in Figure 2. The measurements were all made using overcoring stress-relief techniques in boreholes extending far enough into the solid rock that the stress field was not significantly influenced by the proximity of the tunnel. The maximum principal stress is generally close to the stress corresponding to the weight of the overburden and oriented in a near-vertical direction, consistent with the "basin" tectonics of the region. The near-horizontal principal stresses are typically of the order of half the overburden stress. On the basis of these data we have chosen

$$\frac{\partial S_V}{\partial z} = 265 \text{ bars/km}$$

$$\frac{\partial S_H}{\partial z} = 132.5 \text{ bars/km} \qquad (1)$$

as representative stress gradients in the Witwatersrand, where S_V is the vertical stress, S_H is the horizontal stress and z is the depth coordinate (Figure 2).

All of the available evidence indicates that in the vicinity of the deep mines near Johannesburg the ambient state of stress illustrated in Figure 2 corresponds to a very low level of seismicity. Tremors only began to be noticed in Johannesburg after the mines had attained depths of about a hundred meters of so (Cook et al., 1966). Furthermore, all of the careful studies of seismicity in these deep mines have shown that tremors only occur where the stress field has been substantially changed by mining, as will be seen.

LOCATIONS OF MINE TREMORS AND INDUCED STRESS CHANGES

In 1971 the Bernard Price Institute, in cooperation with the Chamber of Mines of South Africa, installed an underground array of 10 geophones at depths ranging from 1.8 to 3.3 km in the E.R.P.M. gold mine. The array was about 10 km across and was centered about the central section of the mine. Shortly after the seismic array became operational we decided to focus our efforts on studying a small region about 300 m in extent near the center of the array. It was an area where the gold was concentrated at unusually high levels in the ore body with the result that the mine faces advanced at considerably greater rates than in other parts of the mine. The seismicity in this region, termed region A, was correspondingly high.

Figure 3 is a plan view of the seismicity in region A during a 46-day period in the first quarter of 1973. The mining results in a broad tabular excavation about 1.25 to 1.5 m in thickness which causes very high stress concentrations in the abut-

ting rock. The excavation parallels the stratigraphic layers of quartzite which dip southwestward at between 25° and 30°; the average depth of mining in region A is about 3.2 km. We see that most of the seismicity is concentrated in an area within 100 m of the advancing mine face.

A cross section view of the seismicity along line A - A' (Figure 3) is shown in Figure 4, where the hypocenters, which were located with an uncertainty in position of 35 m, have been plotted relative to the instantaneous face position. The events shown in Figure 4 occurred during two 50-day periods, one in the last quarter of 1972 and the other during the first quarter of 1973. We see that most of the hypocenters were located in the rock above the mine excavation, which is an exceptionally siliceous and thickly-bedded quartzite, in contrast with the quartzite below the mining which has a much lower quartz content and closely spaced bedding planes.

As the ambient stress fields and the changes in these fields caused by mining are essentially the same above and below the mining horizon we would expect the seismicity to be symmetrically distributed about the level of mining. It seems likely that the asymmetry arises because of difference is in the strengths. The rock above has a typical strength in uniaxial compression of 3.5 kb whereas that below normally fails at about 2.1 kb (McGarr et al., 1975), a contrast of 1.4 kb. Thus, the tremors tend to occur in especially strong rock that is capable of storing considerable elastic strain energy before failure.

The contours in Figure 4 indicate the degree to which the stress field has been altered from its ambient state (McGarr et al., 1975). Before discussing the stress changes, however, consider the ambient state. From Figure 2 and equation (1) we predict vertical and horizontal components of stress of about 850 b and 425 b, respectively, at a depth of 3.2 km, the approximate depth at the center of

Figure 4.

The contours are meant to show the extent to which the changes in the stress field from its ambient state caused the quartzite to fail. The failure criterion in compression is approximated by

$$s = \sigma_1 - 6\sigma_c \qquad (2)$$

where s is the strength in uniaxial compression, σ_1 is the axial or maximum principal stress, and σ_c is about 3.5 kb.

On the basis of equation (2) we chose to represent the change in the stress field as contours of the function

$$F = \Delta o_1 \quad 6\Delta\sigma_3 \qquad (3)$$

where Δ_1 and Δ_3 are the maximum and minimum principal stress changes induced by mining. F was calculated from Mushkehishvili's (1953) relationships for the stress field of a slit in a compressed medium. $\Delta\sigma$ and $\Delta\sigma_3$ arise because of the interaction between the component of stress normal to the plane of the mining and the collapsing excavation.

Most of the seismicity occurs between the contours F= 1kb and 4kb (Figure 4) which supports the hypothesis that the changes in the stresses induced by mining are triggering the events. It is of interest, however, to analyze the stress changes in somewhat more detail for comparison with the strength of the rock. Along the dashed line in Figure 4 $|\Delta\sigma_1| = |6\Delta\sigma_3|$. To the right of the dashed line $\Delta\sigma_1$ is the more important term and it is oriented more or less parallel to S_1, the maximum principal ambient stress. Thus, in this region the tremors are presumably triggered by the addition of between 1 and 4kb to S_1. To the left of the dashed line the more

change in the stress field is an added tension stress of between 160 b and 670 b.

It is important to note that even though the stress changes induced by mining seem quite substantial the overall stress level in the quartzite is considerably below a failure state according to inferences from laboratory data. As will be discussed, one of the key underground observations of failure associated with tremors is that they are the result of shear failure in previously intact rock and that the fault zone resembles those created on a smaller scale in the laboratory even in microscopic detail. Thus, one might expect a failure criterion for tremors to be similar to equation (2). At the center of the distribution of seismicity, about 25 m to the right of the fact and about 70 m above the slit, the total stress field (ambient+induced) has maximum and minimum principal stresses of about 1.8 kb and 400 b, respectively. Inserting these values into equation (2) gives = -600 b whereas the observed values of s for laboratory failure generally range from about 2 kb to 4 kb. thus, although the mine-induced changes in the stress field seem to trigger the tremors, the tremors occur at levels of stress far below a failure state according to criteria developed from laboratory measurements. Even if pre-existing faults are more important in the failure process than we currently think they are, the quartzite would be far from a state of failure. The only region expected to fail on the basis of laboratory results in the shaded region in Figure 4, which, in fact, does fail in a continuous manner as mining proceeds. This continuous failure results in fracture planes spaced typically at about 10 cm extending 10 to 20 m above and below the slit and forming as far as 10 m ahead of the face (McGarr, 1971a).

In summary, the tremors appear to be triggered by change (equation (3)) in the ambient stress field of the order of 1 kb. The typical shear stress in the seismogenic region (Figure 4) is roughly 700 b which is more than an order of magnitude greater than the stress drops, which range from 5 to 50b. Thus, from

this evidence the average stress driving the seismic failure is of the order of 700 b and the stress drop is insignificant by comparison.

SEISMIC EFFICIENTY IN REGION A

From the observation that the typical driving stress in Region A is about 700 b in conjunction with the information on stress drops, the seismic radiation efficienty, n, can be estimated. If the stress drips to the frictional stress on the fault then (e.g. Wyss and Molnar, 1972)

$$n = \Delta\tau/2\bar{\tau} \qquad (4)$$

UNDERGROUND OBSERVATIONS OF FAULTING ASSOCIATED WITH TREMORS

As seen in Figure 4 most of the seismic failure occurs in the solid rock well away from the mine excavations. Occasionally, however, the tremors, especially in the magnitude range of 1.5 and greater, result in damage to the workings due either to a very high levels of vibration or, more seriously, to sudden closure of the excavation. Sometimes a zone of highly-sheared rock actually outcrops in the workings and when this happens, shear displacements across these freshly-formed faults can be measured if distinct stratigraphic horizons are displaced (McGarr, 1971b).

Figure 5 shows a shear zone associated with underground damage in the eastern section of E.R.P.M. at a depth of about 3 km. The pencil indicates where the hangingwall side of the fault has been displaced downward about 3 cm relative to the footwall side. The occurrence of these normal faults is consistent with the state of stress that is believed to give rise to these failures.

There are several features about these shear zones that are of importance to the present discussion. First, these zones most frequently involve the failure of intact rock. Although there are abundant geological faults and joints in the Witwatersrand mines the faulting associated with the tremors does not show any marked tendency to follow pre-existing faults. The dark mylonite of the geological faults forms a very obvious contrast to the quartzite and if the contemporary faulting, which forms a white zone of finely crushed rock, followed a pre-existing fault, the dark mylonite would be easily visible.

The second key feature of the shear zones is that the failure has an appearance very similar to brittle shear failure in laboratory samples subjected to axial loades in conditions of confining stresses of 100 bars to several kilobars. Thus, it seems entirely reasonable to compare in detail the failure corresponding to mine tremors with laboratory results.

In 1974 we found one of the best examples of an outcropping shear zone intersecting a tunnel in the western part of E.R.P.M. The fault outcropped along the sidewalls and roof of a bay, which had been excavated in the side of the tunnel to house a ventilation fan. Figure 6 shows a shear displacement of 4½ cm being measured in the sidewall of the fan chamber.

From May 1974 until October 1975, a "rabbit warren" of inclined tunnels was developed to follow the complex fracture system and from these tunnels the system was mapped and photographed in detail by Gay and Ortlepp (1978) with help from others of us. In essence, the mapping showed that there were two distinct fracture systems, which intersected, termed shear zones A and B. Each shear zone is approximately planar with zone A dipping at 75° toward the NE and shear zone B dipping 56° to the north (Gay and Ortlepp, 1978); the faulting shown in Figure 6 is part of shear zone A. The morphology of the faulting shown in the vicinity of their intersection indicated that shear zone B formed before A and from detailed comparison

of the positions and attitudes of the shear zones with the mine geometry and the reports of the underground damage it appears that both zones formed in September 1970.

The tremors associated with the two fracture systems unfortunately occurred between two periods of time when seismic data were monitored at E.R.P.M. and so the only seismic information available was recorded at the WWSSN station in Pretoria, 55 km due north. From the reports of underground damage shear zones A and B were correlated with events having local magnitudes of 2.1 and 1.7 (or 3.4) respectively; admittedly the correlations are somewhat uncertain especially for zone B, as will be discussed.

Returning to the mophology of the shear zones, the observations in the inclined tunnels showed that each shear zone consisted of a comples series of sub-parallel shear planes sometimes offset in an en-echelon manner, as seen in Figure 7, or a network of planes that bifurcate and join along their lengths as seen in the upper half of the exposure of Figure 8.

The maximum observed length of a single, continuous shear plane was 4 to 5 m and at the end points the movement died out and was taken up by an adjacent parallel fracture as seen in Figure 7. Closely-spaced, conjugate shear fractures connect the end regions of the primary shear fractures and presumably form as a result of the interaction of the ends of the offset fractures. A close inspection of Figure 7 indicates that the conjugate fractures tend to offset the primary fractures and do not extend beyond the region between the en-echelon fractures. Thus the conjugate shear fractures must have occurred after the primary fractures. Finally, we note that the amount of offset of the end points of en-echelon fractures varies from a few cm to roughly 1 m.

Extension cracks were commonly observed to branch off from the main shear

fractures forming angles of about 30°; one such example is seen in the upper left-hand portion of Figure 7.

Generally, each of the major shear zones consisted of planar systems of fractures distributed over a region several metres in thickness. The typical shear displacement across the planes was 5 - 6 cm with a maximum observed displacement of 10 cm distributed between three parallel fractures. The mapping showed that the fracture system extended from the mine stope downward for at least 20 to 30 m to the level of the tunnel, where we first discovered it. Shear zones A and B had total observed areas of 500 m^2, respectively.

ANALYSIS OF FAULT GOUGE

On the basis of the similarity between the fractures observed undergound and those formed in the laboratory, Spottiswoode (1978) was motivated to analyze the distributions of particle sizes of the comminuted rock in the fault zones in terms of the energy required to create a given amount of free-surface area. To establish the energy necessary to produce free surfaces during failure in triaxial compression he performed a series of laboratory experiments which involved testing specimens of E.R.P.M. quartzite to beyond failure at confining stresses similar to those expected underground. The work W, done on the specimens by the testing apparatus was then compared to the total surface area of the crushed rock in the shear zone.

The crushing energy, G, is defined as

$$G = W/MS \qquad (5)$$

where M is the total mass of the crushed rock and S is the surface area per unit mass.

The most difficult measurements in Spottiswoode's study were the estimates of S. First the size distribution of the quartz particles was determined by

measuring numbers of particles in various size ranges from 5 to 320μ. The observed distributions could be described quite well by log-normal curves and from these empirically determined distributions, S could be estimated from (Herdan, 1953)

$$S = (K/\rho)\sum m_i/x_i \qquad (6)$$

where m_i is the proportioned by mass of particles of size x_i, K is a geometrical shape factor, taken as 11, and ρ = 2.7 gm/cm3, the density of quartz. The summation was taken over the 12 measured size intervals as well as for sizes less than 5μ. The surface area from particles larger than 320μ, of the order of the grain sizes, is of little consequence.

The results of the laboratory studies are summarized in Table 1. The average driving stresses, $\bar{\tau}$, were calculated from

$$\bar{\tau} = W/AD \qquad (7)$$

where A is the fault area and D is the fault displacement. The specimens were tested at two levels of confining stress of 138 b and 276 b (Table 1), which, from equations (1), correspond very roughly to depths of 1 km and 2 km, respectively. Because of some experimental difficulties in the post-failure region of sample 21, Spottiswoode (1978) gave considerably greater emphasis to the results of samples 20 and 22. These analyses yielded values of G of 2.7×10^5 and 4.5×10^5 ergs/cm, respectively, values which are very high in comparison to determinations from failure in tension (e.g. Brace and Walsh, 1962). Another point of possible significance is the apparent dependence of G on confining stress from these three determinations.

Spottiswoode (1978) then applied the laboratory results to the analysis of gouge particles from the fracture system in the western part of E.R.P.M. Gouge particles were extracted from three regions of the portion of the fracture system shown in Figure 7 and from a fourth region nearby and the results of analyzing these samples of gouge are summarized in Table 2.

To estimate the driving stress, $\bar{\tau}$, Spottiswoode combined equations (5) and (7) to obtain

$$\bar{\tau} = SG\rho w/D \qquad (8)$$

where w is the thickness of a particular zone of gouge and the gouge density, ρ, was estimated as 2 gm/cm^3. The average depth of the fracture system is about 2 km and so, from equation (1), the value of G of 4.5×10^5 ergs/cm^2, corresponding to the higher confining stress of 276 bars (Table 1) is more appropriate for the present analysis. This departs from Spottiswoode's (1978) analysis in that he averaged the determinations of G from samples 20 and 22 to obtain G = 3.6×10^5 ergs/cm^3 and thus, somewhat lower estimates of $\bar{\tau}$ than those listed in Table 2. The average value of $\bar{\tau}$ of 400 bars is in reasonable agreement with the value of 700 b determined in region A from the distribution of hypocenters.

The analysis of the fault gouge indicates that nearly all of the elastic energy released during faulting is consumed in creating the zone of crushed rock and this is consistent with the estimates of seismic efficiency of less than 1%, discussed earlier.

DISCUSSION

To what extent can the results discussed here be applied to the analysis of naturally occurring earthquakes? Are the driving stresses of natural earthquakes also an order of magnitude or so greater than the observed stress drops as suggested by Hanks (1977), among others? The available data on the levels of shear stress within the crust indicate a generally increasing level of τ with depth as seen in Figure 9. At depths of 5 to 10 km τ must almost certainly be in excess of 300 bars in hard rock such as granite. Unfortunately, only one of the measurements represented in Figure 9 was made in the vicinity of a currently-active tectonic

fault and so it is conceivable that shear stresses in a major fault zone are much lower than those suggested by the data of Figure 9. As indicated in Figure 9, the measurement made near the San Andreas fault (Zoback et al., 1977) is somewhat higher than the average of the shallow measurements and so there are no data so far that indicate remarkably low shear stresses in the San Andreas fault zone or along any other active fault.

From the present evidence, then, it seems likely that natural tectonic earthquakes result from faulting driven by shear stresses of about 300 bars or more with the necessary implication that crustal earthquakes have very low seismic efficiencies of the order of 1% or less. Nearly all of the elastic strain energy may be consumed in creating fault gouge.

It is of prime interest to compare the various aspects of shear zones A and B to the seismic source parameters normally measured for tremors with magnitudes close to 2 and this comparison is summarized in Table 3.

At the outset we should emphasize that most of the measurements listed in Table 3 have a considerable amount of uncertainty and are primarily intended to illustrate the comparisons between the conventional modes of faulting and what was observed. The limitations on the estimates of the areas of the shear zones have been discussed already. The average shear displacements, D, were based on a suite of about 10 measurements of offsets of marker horizons or geological joints such as shown in Figures 5,6 and 7. The values ranged from about 4 to 10 cms and 6 cm is a reasonable average displacement for each of the zones. Because the fracture system did not appear substantially different away from the positions where displacements could be measured we are confident that the measured displacements are representative within the general region of observation.

As indicated in Table 3 the moments inferred from the underground measurements are lower bound estimates only as the entire fracture system was not encom-

passed by the exploratory tunnels. Similarly the radii of the shear zones, idealized as circles, are lower limits and tabulated only for the purpose of comparison with a circular model of seismic faulting.

The stress drops, $\Delta\tau = (7/16)\ {}^{M_o}/_{r_o^3}$ (9)

and are almost certainly over-estimated, as will be seen.

As mentioned before, the assignment of particular tremors to the two shear zones was somewhat uncertain, especially the correlation between the event of magnitude 1.7 and shear zone B, the first of the shear zones to be formed. On 5 September 1970 (Saturday) a tremor of magnitude 3.4 occurred at 21.53 GMT and then on 6 September at 21.58 GMT there was an event of magnitude 1.7; as both of these events occurred during the weekend break when there was no production, the corresponding underground damage was reported about 300 m to the west of the area under discussion and lesser damage was reported in the mine stope almost coincident with shear zone B. We associated the smaller event with shear zone B on the basis of the relative amounts of damage but there is a fair possibility that this shear zone was part of the failure that occurred at the time of the event of magnitude 3.4. We have considered both possibilities in Table 3 as indicated by the additional set of values in parenthesis for zone B. Fortunately, the correlation of shear zone A with an event of magnitude 2.1 is more certain.

All of the seismically inferred results listed in Table 3 were based on Brune's (1970, 1971) model of faulting and relations between source parameters and magnitude developed by Spottiswoode and McGarr (1975). We see that there are considerable discrepancies between the fracture parameters observed underground and those inferred seismically. The seismically inferred moments are significantly larger than the moments measured underground which means that the exploratory tunnels circumscribed only a small fraction of the total fracture surface. This is not too surprising because at the lower and souther extremities of the tunnels (e.g.

Figure 6), the fracture system was still showing large displacements with no indications of attenuation.

The seismically determined source radii are also an order of magnitude, or more greater than the approximate radii of the region explored underground, again suggesting that only a small fraction of the total system was actually mapped.

Perhaps the most interesting comparison is between the observed displacements and those inferred using Brune's (1970, 1971) model. The observed displacements are more than an order of magnitude greater than displacements expected for a circular fault with a uniform stress drop. As seen in Table 3 the seismically inferred displacements are only weakly dependent on the magnitude.

Furthermore, according to the model, the maximum displacement, at the center of the circular fault, should only be 33% larger than the average displacement. Thus, the large shear displacements of 4 to 10 cm measured in the tunnel, are clearly inconsistent by an order of magnitude, or so, with a model involving a homogeneous stress drop of 5 to 50 bars over the fault surface.

Finally we note that the stress drops estimated from the underground observations are about two orders of magnitude greater than those typically calculated from the seismic data. The calculated from the seismic data. The calculation of $\Delta\tau$ (underground) was based on the assumption that all of the faulting was included within the tunnel network, which is almost certainly not the case, especially in view of the discrepancies between underground and seismic moments. The only reason for listing $\Delta\tau$ (underground) is to suggest the extent to which locally high values of stress drop occur. It is reasonably certain that the stress drops are nowhere as high as the apparent values listed in Table 3 because if the driving stress is, in fact, close to 400 b then $\Delta\tau > 800$ bars implies that the fault overshoots in relaxing the shear stress.

Although we know that shear zones A and B must extend well beyond the limits

of the tunnel system we neither know how far they extend nor the nature of the extension. One possibility is that the faulting extends outward from the explored region for more than 100 m but with shear-displacements that are only a small fraction of 1 cm on the average. Another possibility is the the faulting extends over a region with perhaps 10 times the area of the explored region, with shear displacements comparable to those that were measured.

The analysis of our underground observations of shear zones has not solved any of the fundamental problems about the relationship between failure associated with earthquakes and failure in laboratory samples but the observations and comparisons presented here may serve to define the problem more closely.

ACKNOWLEDGEMENT

We are grateful to the management and staff of the East Rand Proprietary Mines for their cooperation and support which made the underground observations possible. Mr. W. Lagrange and Mr. M. Spengler were particularly helpful in the development of the inclined tunnels. The Chamber of Mines of South Africa provided financial support for this investigation.

REFERENCES

Brace, W.F. and J.B. Walsh, Some direct measurements of the fracture energy of quartz and orthoclase, Am. Miner, 47, 1111-1122, 1962.

Brune, J.N., Tectonic stress and the spectra of seismic shear waves from earthquakes, J. Geophys. Res., 75, 4997-5009, 1970.

Brune, J.N., Tectonic stress and the spectra of seismic shear waves from earthquakes; correction, J. Geophys. Res., 76, 5002, 1971:

Cook, N.G.W., The seismic location of rock bursts, Proc. Fifth Rock Mechanics Symposium, Pergamon Press, Oxford, 493-516, 1963.

Cook, N.G.W., E. Hoek, J.P.G. Pretorius, W.D. Ortlepp, and M.D.G. Salamon, Rock mechancis applied to the study of rockbursts, J.S. African Inst. Mining Met., 66, 435-528, 1966.

Gane, P.G., A.L. Hales, and H.A. Oliver, A seismic investigation of the Witwatersrand earth tremors, Bull. Seismol. Soc. Am., 36, 49-80, 1946.

Gay, N.C., Virgin rock stresses at Doornfontein Gold Mine, Carletonville, South Africa, J. Geol., 80, 61-80, 1972.

Gay, N.C., In situ stress measurements in Southern Africa, Tectonophysics, 29, 447-459, 1975.

Gay, N.C., and W.D. Ortlepp, The anatomy of a mining-induced fault zone, Bull. Geol. Soc. Am., in press, 1978.

Hanks, T.C., Earthquake stress drops, ambient tectonic stresses and stresses that drive plate motions, Pure & Applied Geophys., 115, 441-458, 1977.

Hanks, T.C., and W. Thatcher, A graphical representation of seismic source parameters, J. Geophys. Res., 77, 4393-4405, 1972.

Hanks, T.C. and M. Wyss, The use of body-wave spectra in the determination of

seismic-source parameters, Bull. Seismol. Soc. Am., 62, 561-589, 1972.

Herdan, G., Small Particle Statistics, Elsevier, Amsterdam, pp 50 and 113, 1953.

McGarr, A., Stable deformation of rock near deep-level tabular excavations, J. Geophys. Res., 76, 7088-7106, 1971a.

McGarr, A., Violent deformation of rock near deep-level tabular excavations--seismic events, Bull. Seismol. Soc. Am., 61, 1453-1466, 1971b.

McGarr, A., S.M. Spottiswoode, and N.C. Gay, Relationship of mine tremors to induced stresses and to rock properties in the focal region, Bull. Seismol. Soc. Am., 65, 981-993, 1975.

McGarr, A. and N.C. Gay, State of stress in the earth's crust, Annual Review of Earth and Planetary Sci., 6, 1978.

Muskhelishvili, N.I., Some Basic Problems on the Mathematical Theory of Elasticity, Noordhoff, Groningen, 1953.

Spottiswoode, S.M., Fault gouge, driving stress and seismic efficiency, submitted to J. Geophys. Res., 1978.

Spottiswoode, S.M. and A. McGarr, Source parameters of tremors in a deep-level gold mine, Bull. Seismol. Soc. Am., 65, 93-112, 1975.

Wyss, M. and P. Molnar, Efficiency, stress drop, apparent stress, effective stress, and frictional stress of Denver, Colorado, earthquakes, J. Geophys. Res., 77 1433-1438, 1972.

Zoback, M.D., J.H. Healy, and J.C. Roller, Preliminary stress measurements in central California using the hydraulic fracturing technique, Pure & Appl. Geophys., 115, 135-152, 1977.

TABLE 1

Failure and gouge data of E.R.P.M. quartzite tested in triaxial compression

Sample	Max. axial stress (bars)	Confining stress (bars)	W (10^8 ergs)	$\bar{\tau}$ (bars)	M (gm)	S (cm^2/gm)	G (10^5 ergs/cm^2)
20	3400	138	7.3	800	3.50	780	2.7
21	4320	276	5.4	1550	1.81	600	5.0
22	4240	276	11.0	930	3.04	800	4.5

TABLE 2

Data on samples of burst zone gouge

Gouge	W (cm)	D (cm)	S (cm^2/gm)		(bars)
A+B	2.1	6.	1500		470
C	0.75	5.	2700		360
D	1.2	6.	2200		400
				Average	410

TABLE 3

Comparison of observed and inferred source parameters

	Shear zone A	Shear zone B
Area, m^2	$\geq$500	$\geq$600
D, cm	6	6
M_o (underground), [a]dyne-cm	$\geq 1.9 \times 10^{19}$	$\geq 1.4 \times 10^{19}$
r_o (underground), m	$\geq$13	$\geq$14
$\Delta\tau$ (underground), b	$\geq 3.8 \times 10^{3}$	$\geq 2.2 \times 10^{3}$
Magnitude	2.1	1.7 (3.4)
M_o (seismic), [b]dyne-cm	1.7×10^{20}	5.5×10^{19} (6×10^{21})
r_o (seismic), [c]m	215	130 (1030)
$\Delta\tau$ (seismic)[c], b	5 to 50	5 to 50

[a] $Mo = \mu \cdot Area \cdot D$, $\mu = 4 \times 10^{11}$ dynes/cm^2

[b] log Mo = 17.7 + 1.2·Magnitude (Spottiswoode and McGarr, 1975)

[c] From results of Spottiswoode and McGarr (1975)

FIGURE CAPTIONS

Figure 1. The long period asymptotic level of the seismic spectrum as a function of the corner frequency of the seismic moment as a function of the source radius. (From Spottiswoode and McGarr, 1975)

Figure 2. Vertical and horizontal components of stress as a function of depth. The vertical stresses were measured in Southern Africa, North America and Australia. The horizontal stresses were all measured in Southern Africa and all determinations at depths below 1 km are from the deep mines of Witwatersrand. S_{Hmin} and S_{Hmax} denote the minimum and maximum near-horizontal principal stresses, respectively.

Figure 3. Seismicity in region a from February 13 to March 31, 1973. The tremors, represented as dots, have been classified into five ranges of local magnitude, M_L: $M_L < 0$, $0 \leq M_L < 0.75$, $0.75 \leq M_L < 1.50$, $1.50 \leq M_L < 2.25$, $M_L \leq 2.25$. The cross-hatched region indicates the advance of the mine face during this period. (From McGarr et al., 1975)

Figure 4. Cross section view of seismicity relative to the instantaneous position of the mine face. The mined-out area is the slit extending leftward from the middle of the picture and has a thickness of 1.25 to 1.5 m. The tremor sizes were explained in the previous figure caption. (From McGarr et al., 1975.)

Figure 5. A burst fracture discovered in the Eastern Section of E.R.P.M. The pencil indicates where the fracture has offset the conglomerate horizon about 3 cm in a sense appropriate for normal faulting.

Figure 6. Part of the outcrop of shear zone A where it outcrops in the southern sidewall of the fan chamber. A shear displacement of 4½ cm is being measured.

Figure 7. An en-echelon offset in shear zone A where it outcrops in the face of the initial inclined tunnel. A prominent geological joint has been displaced

about 6 cm across the offset region of the fault.

Figure 8. Shear zone A where it outcrops in the sidewall of the fan chamber. In the lower left portion of the fault is an en-echelon offset and in the upper right the fault bifurcates and joins several times.

Figure 9. Measurements of shear stress as a function of depth made in North America, Southern Africa and Australia. The circles indicate measurements made in soft rock such as shales, sandstones and limestones. The triangles represent measurements in hard rocks such as granite and quartzite.

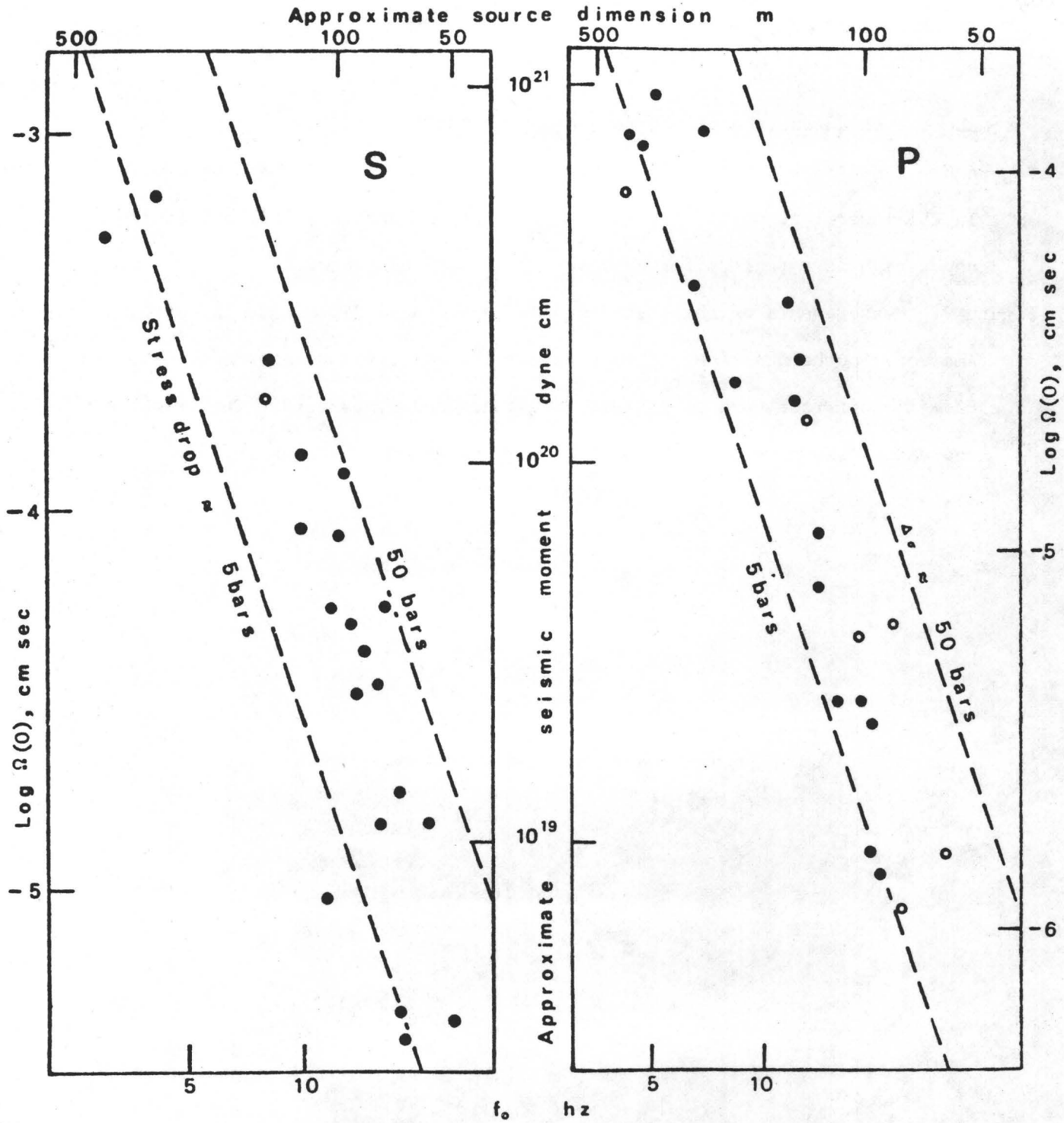

Approximate source dimension m
500
100
50
S
P
Stress drop = 5 bars
50 bars
Δσ = 50 bars
5 bars
Log Ω(0), cm sec
Approximate seismic moment dyne cm
10^{21}
10^{20}
10^{19}
-3
-4
-5
-6
5
10
f_o hz

Figure 1

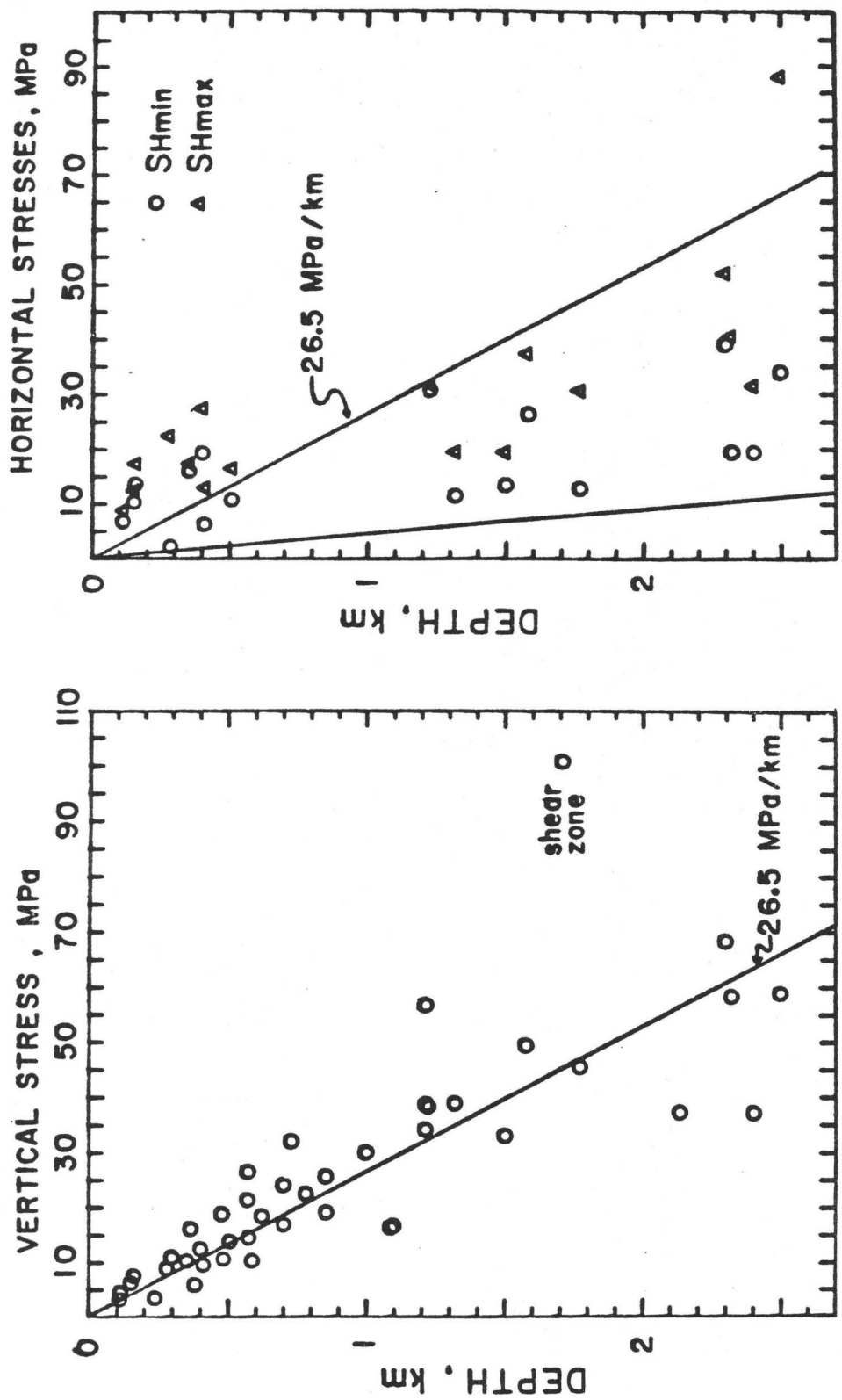

HORIZONTAL STRESSES, MPa
SHmin
SHmax
26.5 MPa/km
DEPTH, km
VERTICAL STRESS, MPa
shear zone
26.5 MPa/km
DEPTH, km

Figure 2

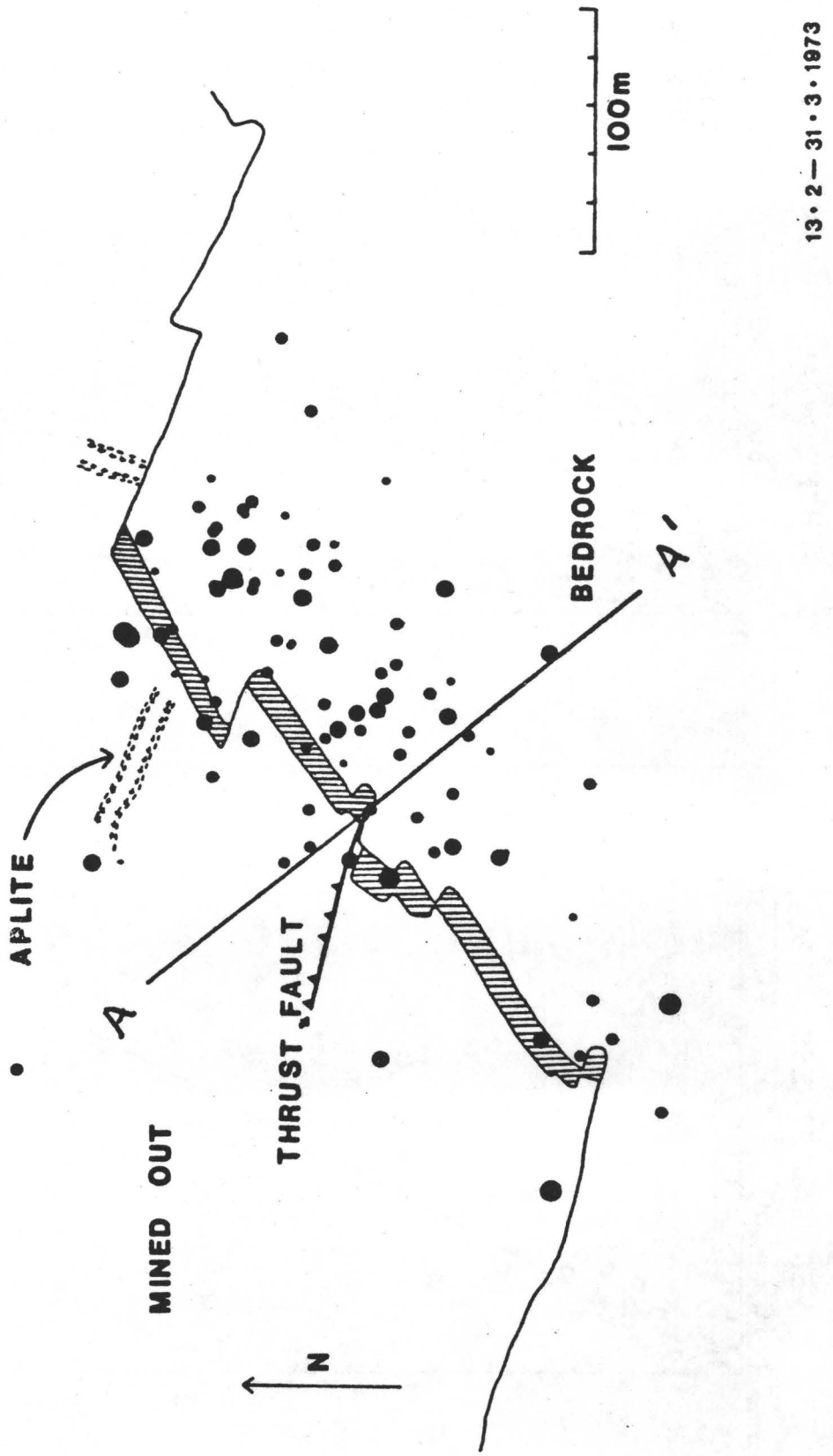


Figure 3

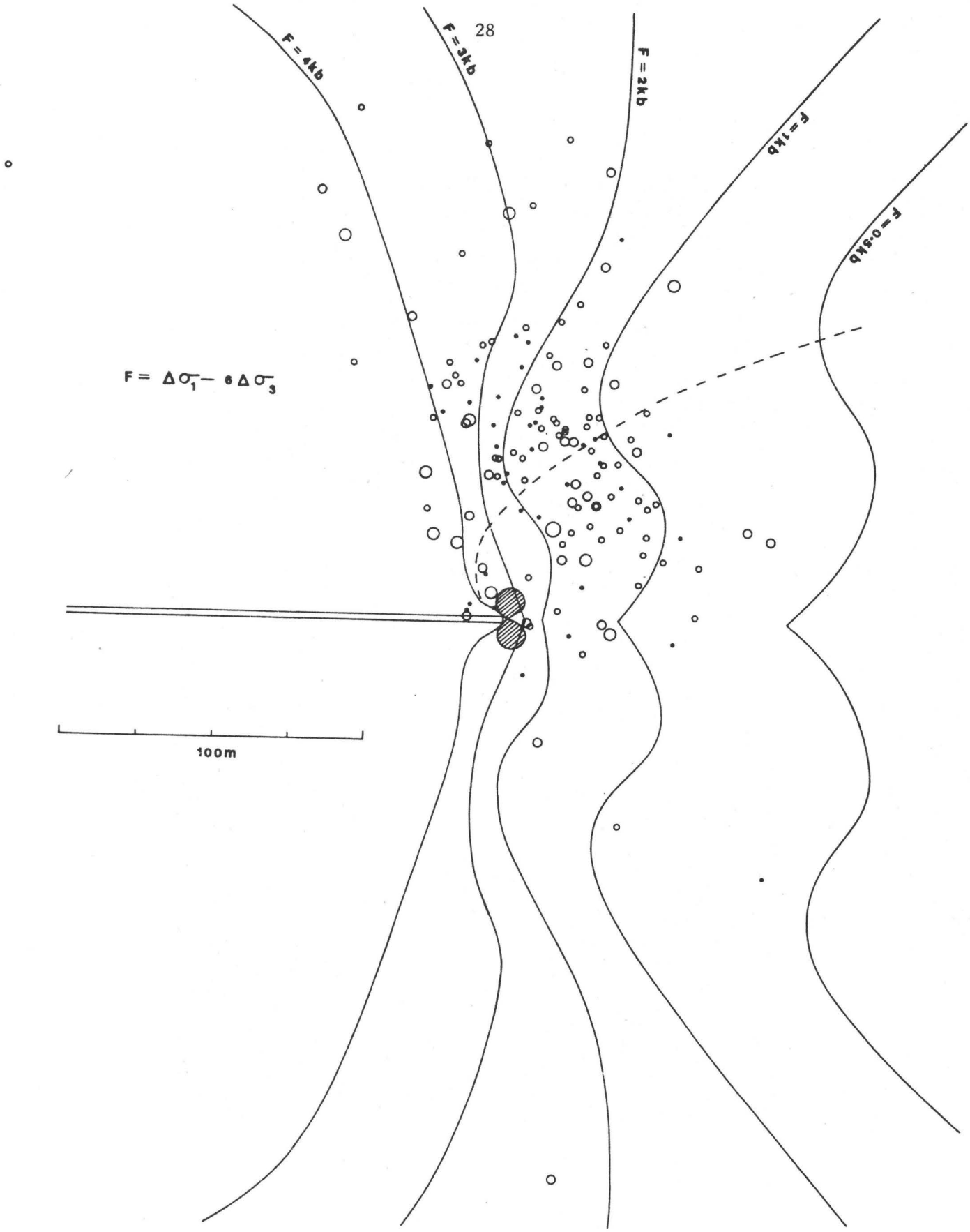


Figure 4

Figure 5

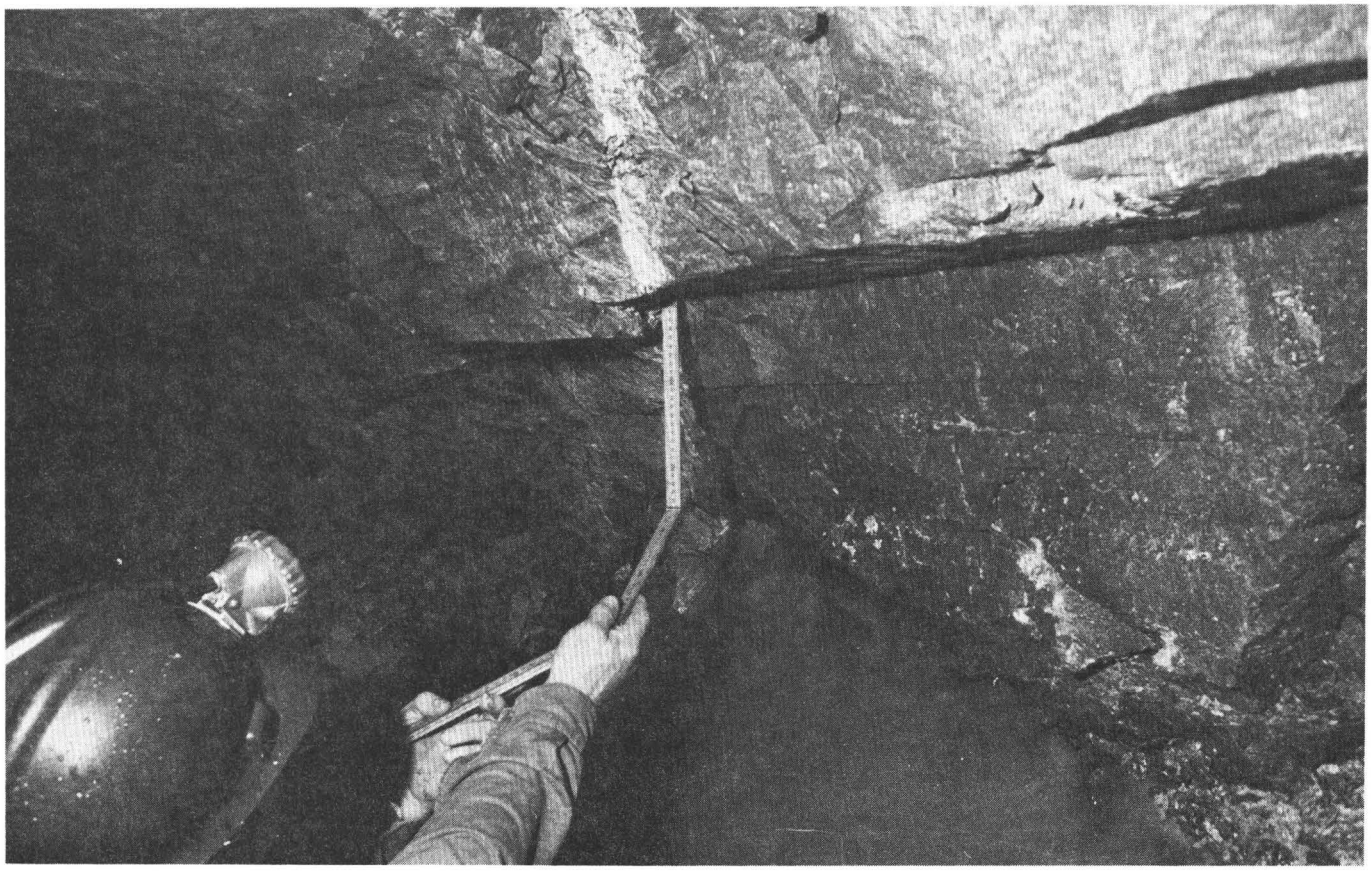


Figure 6

Figure 7

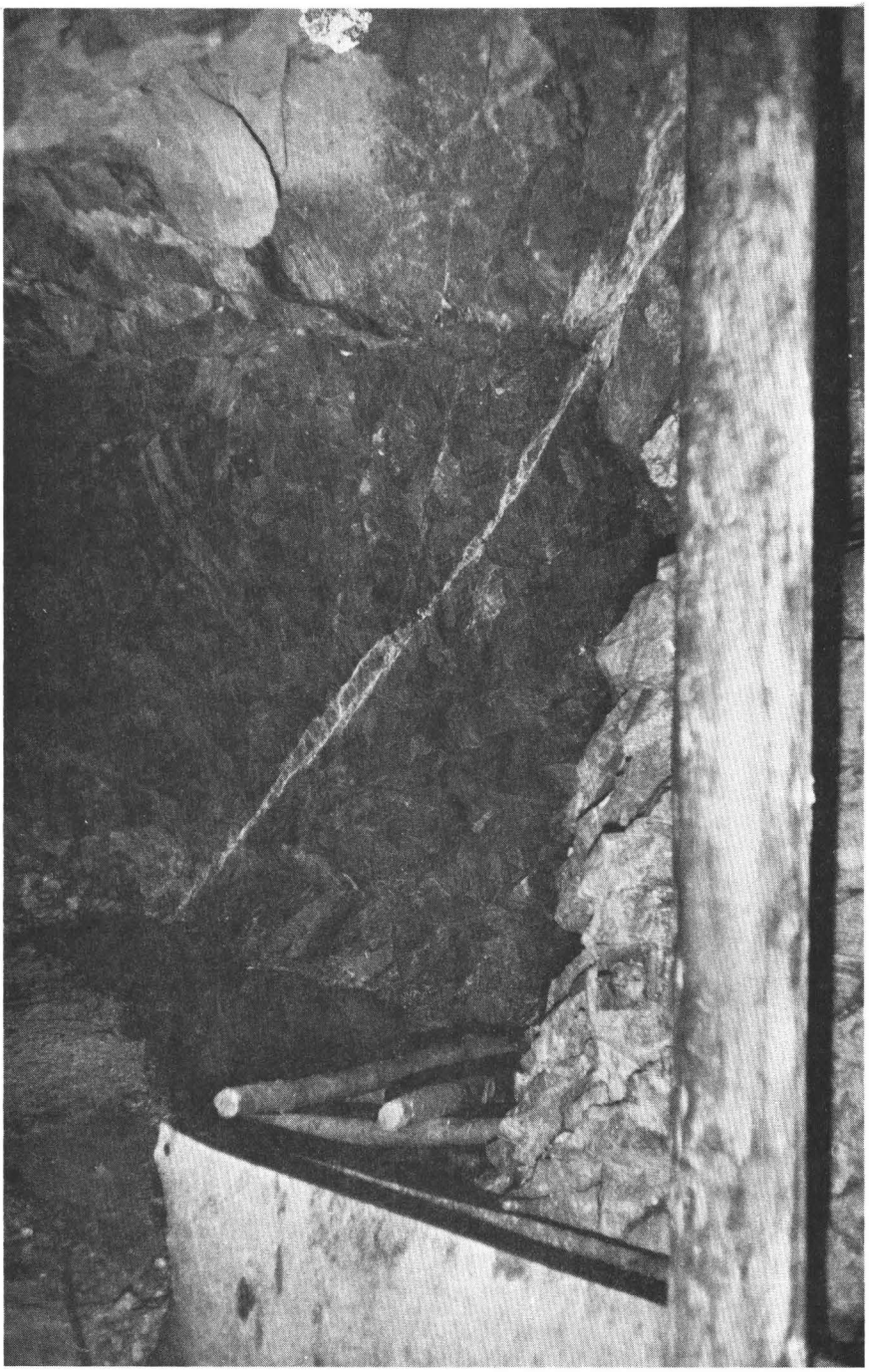


Figure 8

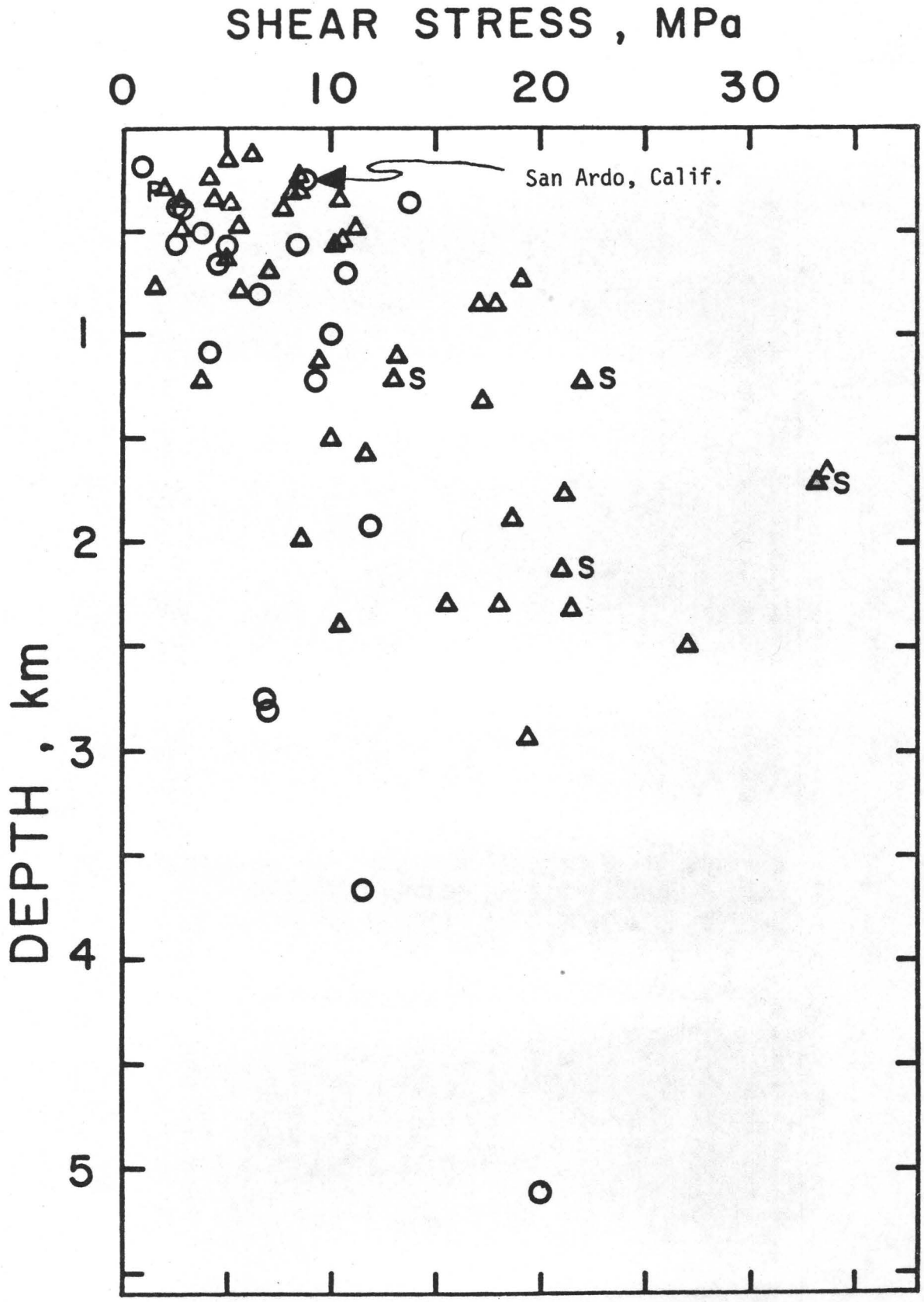


Figure 9

Slow Earthquakes
and Very Slow Earthquakes

by J. H. Pfluke

U.S. Geological Survey

Menlo Park, California

Introduction

There has recently been a resurgence of interest in some slow deformations that oftentimes accompany or slightly precede regular earthquakes. These deformations are detected on strain meters or on ultra-long period seismometers. Sacks and others (1977) reported detecting several such events on a borehole strainmeter in Japan. The duration of these events was of the order of tens of minutes. Sacks named these events "slow earthquakes", since they were observed to have very slow rise times on the record. The slow rise time was interpreted to be the result of abnormally low velocity of rupture on the fault plane. Sacks also made a comparison of two events with similar magnitude and location that occurred off Hokaido in June of 1975. One was a normal earthquake and the other was accompanied by a large coseismic slow deformation. The latter generated a tsunami whereas the former did not. He offered that pair as evidence that slow earthquakes and normal earthquakes can occur at the same locality.

R. J. Geller reported at this conference (paper included in this volume) of discovering independently the same "slow earthquake" off the coast of Hokaido by observing that its seismic moment, when computed from 200 second surface waves, was much larger than what would be computed from the standard magnitude vs. moment relationship. J. H. Pfluke (comments to Geller's paper) pointed out in the ensuing discussion that there are several documented cases in the literature of the early 1970's of another type of slow fault slip,

some with duration times up to 30 hours. Such prolonged fault movements might well be called "very slow earthquakes". It is the purpose of this paper to review certain of these cases with which I am familiar and to emphasize two investigations in which I was a participant.

Evidence of coseismic slow deformation from tsunamis

Sacks and others (1977) and Geller (this volume) give evidence of earthquakes whose total moment was considerably larger than that which was calculated from conventional seismic methods. In both of these cases, tsunamis resulted because, although the time constant of the fault slippage was long compared to the response band of seismographs, it was short compared to response characteristics of the sea whereunder the deformation took place.

Ando (1975) presented a well-documented study revealing a large component of slow deformation that occurred coseismically with the Nankaido earthquake of 1946. He showed that the fault moment ($\mu \bar{A} \bar{u}$) necessary to produce the deformations associated with the observed geodetic data is some five times that which is calculated from conventional seismic observations. This earthquake also produced a tsunami which was too large to be consistent with the source area calculated by seismic means.

The observations of Sacks and Ando might lead us to believe that information on coseismically occurring slow deformations might be obtainable from a catalog of tsunamis. With this in mind, I have

extracted from the Catalog of Tsunamis in Hawaii (Pararas-Carayannis and Calebaugh, 1977) some statistics on the largest tsunamis that have occurred in the Pacific since 1930. The list (Table I) summarizes all of the tsunamis that had a tsunami magnitude, M_t, greater than 2. M_t is a quantitative measure of the size of a tsunami (Iida et al., 1967). It is calculated by the relationship

$$M_t = \log_2 H$$

where H is the maximum runup or amplitude on a coastline near the origin.

The tsunamis are arranged by date. The other information in the table is the location of the origin, the earthquake magnitude, the tsunami magnitude, and the ratio of these quantities. The ratio M_t/M_s (M_s is the seismic magnitude from 20 second surface waves) is taken to be a measure of the seismic efficiency of tsunami generation.

Three generalizations that can be made from the table and from the originating catalog are as follows:

1) All of the tsunamis of Table I, with the exception of two (1935, Hawaii; and 1975, Hawaii) originated in a zone of lithospheric subduction (a well-known fact of tsunami generation).

2) The majority of tsunamis generated by earthquake magnitudes less than 8.0 have tsunami magnitudes less than 1.

3) Earthquakes with $M_s > 8.25$ display the largest ratios--M_t/M_s. This phenomenon is probably the result of the fact that the magnitudes scale, M_s, is known to be deficient in reflecting the true seismic moment of very large earthquakes (Geller and Kanamori, 1977).

There are two notable exceptions (1946, E. Aleutian Islands; 1963, Kurile Islands) to generalization number 2, above, which, although seismically rather small, generated large tsunamis and exhibited high tsunami generating efficiency.

It should be pointed out that neither the 1946 Nankaido earthquake nor the 1975 Nankaido earthquake had unusually high tsunami generating efficiency (M_t/M_s = .30), although in each case good evidence has been presented (Ando, 1975; Shimazaki and Geller, 1977) that these earthquakes had associated with them substantial components of slow deformation, i.e., aseismic fault slip. This fact leads to the speculation that several other earthquakes listed in Table I had associated components of aseismic fault slip as part of their mechanism.

It is interesting to focus on those events of Table I for which $M_s < 8.0$. For these events, the conventional means of computing earthquake magnitude should render a good understanding of the "size" of the earthquake, i.e., the difficulties of consistency that are known to exist for $M_s > 8.0$ do not pertain. Eight events fall into this category and two of these (1946, E. Aleutian Islands; 1963, S. Kuril Islands, M_s = 6.8) stand out as having been highly efficient at tsunami generation. In fact, the 1946 Aleutian earthquake generated the most destructive tsunami ever to strike the Hawaiian Islands (Pararas-Caraymis, and Calebaugh). It is intriguing to speculate that large components of coseismic aseismic fault slip may have been operative for these two events. Unfortunately, neither the geodetic

data nor the long period seismic data exist to permit investigations such as those done on the Nankaido earthquake and the 1975 Hokaido earthquake and the 1975 Hokaido earthquake. Assuming that the high tsunami generating efficiency of the Kurile, $M_s = 6.8$, earthquake was due to a large component of aseismic slip, and comparing its parameters with those of the 1963 Kurile event ($M_s = 8.25$, $M_t/M_s = .24$) supports a statement made by Sacks and others (1977) that "normal" earthquakes and "slow" earthquakes can occur in the same region.

The 1975 Hawaiian earthquake-tsunami stands out as an exceptional event in Table I. This event and the Hawaiian event of 1935 are the only ones in the list that do not fall in a zone of lithospheric subduction. The ratio M_t/M_s is .41 for the 1975 event (no value for M_s has been published for the 1935 event). A fault plane solution (Ando, to be published) along with tsunami data (Hatori, 1976) indicate a shallow dipping fault plane (dip $< 5^o$) extending in a southerly direction away from the epicenter. The tsunami and geodetic data indicate that the slip plane is much larger than what can be deduced from seismic data alone (R. S. Crosson, personal communication). For example, that area of the sea floor attributed by Hatori to have generated the tsunami extends some 30 kilometers away from earthquake epicentral and aftershock zones (R. S. C.). These observations indicate that this event was indeed a slow earthquake.

There are two notable exceptions (1946, E. Aleutian Islands; 1963, Kurile Islands) to generalization number 2, above, which, although seismically rather small, generated large tsunamis and exhibited high tsunami generating efficiency.

It should be pointed out that neither the 1946 Nankaido earthquake nor the 1975 Nankaido earthquake had unusually high tsunami generating efficiency (M_t/M_s = .30), although in each case good evidence has been presented (Ando, 1975; Shimazaki and Geller, 1977) that these earthquakes had associated with them substantial components of slow deformation, i.e., aseismic fault slip. This fact leads to the speculation that several other earthquakes listed in Table I had associated components of aseismic fault slip as part of their mechanism.

It is interesting to focus on those events of Table I for which $M_s < 8.0$. For these events, the conventional means of computing earthquake magnitude should render a good understanding of the "size" of the earthquake, i.e., the difficulties of consistency that are known to exist for $M_s > 8.0$ do not pertain. Eight events fall into this category and two of these (1946, E. Aleutian Islands; 1963, S. Kuril Islands, M_s = 6.8) stand out as having been highly efficient at tsunami generation. In fact, the 1946 Aleutian earthquake generated the most destructive tsunami ever to strike the Hawaiian Islands (Pararas-Caraymis, and Calebaugh). It is intriguing to speculate that large components of coseismic aseismic fault slip may have been operative for these two events. Unfortunately, neither the geodetic

data nor the long period seismic data exist to permit investigations such as those done on the Nankaido earthquake and the 1975 Hokaido earthquake and the 1975 Hokaido earthquake. Assuming that the high tsunami generating efficiency of the Kurile, $M_s = 6.8$, earthquake was due to a large component of aseismic slip, and comparing its parameters with those of the 1963 Kurile event ($M_s = 8.25$, $M_t/M_s = .24$) supports a statement made by Sacks and others (1977) that "normal" earthquakes and "slow" earthquakes can occur in the same region.

The 1975 Hawaiian earthquake-tsunami stands out as an exceptional event in Table I. This event and the Hawaiian event of 1935 are the only ones in the list that do not fall in a zone of lithospheric subduction. The ratio M_t/M_s is .41 for the 1975 event (no value for M_s has been published for the 1935 event). A fault plane solution (Ando, to be published) along with tsunami data (Hatori, 1976) indicate a shallow dipping fault plane (dip $< 5^o$) extending in a southerly direction away from the epicenter. The tsunami and geodetic data indicate that the slip plane is much larger than what can be deduced from seismic data alone (R. S. Crosson, personal communication). For example, that area of the sea floor attributed by Hatori to have generated the tsunami extends some 30 kilometers away from earthquake epicentral and aftershock zones (R. S. C.). These observations indicate that this event was indeed a slow earthquake.

Of the remaining five small events ($M_s < 8.0$) of Table I, three (occurring in 1965, 1969 and 1971) were characterized by the ratio M_s/M_t being .4 or greater. This together with the statistics on the large earthquakes listed in Table I ($M_s > 8.0$), whose ratios are generally greater than .4, suggest that perhaps a considerable component of aseismic fault slip is required before an earthquake can generate a destructive tsunami.

Evidence of very slow earthquakes from regional strain data

From 1968 to 1971 a large number of strain seismometers (of a modified Benioff type) were in operation in Nevada and the Aleutian Islands. Most of these instruments were operated continuously for about two years as part of a national effort to discern whether large nuclear explosions were likely to trigger large earthquakes. During this period, some interesting observations of earth strain were made in these regions which would not have been possible without the increased number of strainmeters. Major and Tocher (1971) reported to have observed large transients on the Aleutian strainmeters. The transients were as large as 3-4 parts in 10^6 and were correlatable across the entire array of instruments whose locations spanned some 350 km along the Aleutian arc. Smith and Kind (1972) reported similar correlatable transients in Nevada over an array of six strainmeters covering an area 60 x 150 km.

The regional strain variations in Nevada were such that a secular trend would be observed for a month or so, followed by an

overall field reversal that would take place over one or two days. Smith and Kind performed an analysis on one such sequence spanning two months of data. They assumed that the strain field was uniform in space and varied quasi-statically in time. Since there were six strainmeters, the problem of solving for the principal axes was overdetermined. They then determined the strain field at daily intervals using an iterative least-squares technique. They found that the principal axes, thus determined, remained more or less statically oriented but built up gradually to values as large as 10^{-5} and subsided spasmotically as described above. The direction of the principal axes agreed closely with those determined from earthquake focal plane solutions and from geomorphological evidence.

Smith and Kind attempted to interpret the strain transient as aseismic slip occurring on geologic faults in the area. Using numerical techniques and the displacement field equations of Mansinha and Smylie (1971) for uniform slip on a rectangular fault did not provide a satisfactory fit of the model to the data. Their final interpretation was that the strains were direct results of relative movement between the crust and the mantle.

Examples of the types of strain variation observed in the Aleutian Islands are shown in Figures 1 and 2. Figure 1 depicts tracings of two components of the strain- and tilt-meter array that operated in the Aleutians in 1970-71. The records are characterized by a large annual variation which approaches 1×10^{-5} in magnitude. Superposed on this are abrupt variations, the largest of

which are 1 to 3 parts in 10^6. Following the abrupt variation, the strain trace would typically return back to the long term mean trend for the particular season. The return phase typically lasted from 1 to 4 weeks. During the period of operation of the Aleutian array, about 6-8 similar such phases could be correlated across the array each year.

The abrupt variations can be better examined in the expanded scale tracing shown in Figure 2. The figure depicts the strain and tilt tracings of the 14 components that were operating at that time. The geographic locations and azimuthal orientations of the respective components are indicated in Figure 3. Two of the above-mentioned abrupt strain episodes occurred during the ten day period shown in Figure 2. The vertical dashed lines mark points of common absolute time. The episode occurring on 11 October can be clearly seen on the Amchitka records and on the Adak records. There were no records for Rat Island at that time and the records from the tilt components on Amatignak were quite noisy. An even larger episode occurred on 16-17 October that was visible on all records. The duration of each of these episodes was about 30 hours.

It was suggested by Major and Tocher (1971) that these large strain episodes might be caused by aseismic slip on fault surfaces at depth. They were motivated by the similarity of form of these episodes with the well-documented episodes of aseismic slip that take place at the surface along some of the active faults in California (e.g. Nason, 1971). Pfluke and Stewart (1973) carried

this hypothesis further by solving the inverse problem of finding the location, orientation and dimensions of the fault plane of one such event, assuming that that the episodes were indeed caused by aseismic fault slip.

Using the displacement field equations of Mansinha and Smylie (1971) for a finite dimensioned fault in a half-space, the solution for the fault plane assumed to cause the strain episode was found by an iterative least squares technique (Marquardt, 1963). The formulations of Mansinha and Smylie has nine parameters and there were 14 elements of data, thus allowing a degree of overdetermination that is seldom available for this type of problem. The parameters of the best fitting solution to the 16 October data are as follows:

strike azimuth	90°
dip	13°
depth to top edge	27 km
width	29 km
length	113 km
dip component of slip	5.1 meters
lateral component of slip	7.2 meters

A surface projection of this hypothetical fault appears on the map of Figure 3. The location and orientation of the solution fault agree well with our knowledge of the upper surface of the crust undergoing subduction, hence inferring that the 16 October strain episode was caused by an episode of aseismic slip in the process of tectonic crustal subduction.

Evidence of very slow earthquakes from local strain data

It is well-known that very slow earthquakes take place repeatedly along the San Andreas fault in California. Evidence for these events are customarily recorded on displacement meters that span the fault zones at localities where aseismic fault slip is known to take place. The time duration of these events varies typically from the order of 30 minutes to 3-4 days (Nason, 1971).

Although it is commonplace to observe these events on displacement meters (often called creepmeters), these events are rarely observed on strain measuring instruments. I know of only one such event that was recorded both on a creepmeter and on a type of strain sensing instrument. Johnson et al. (1973) documented a moving front of aseismic slip that was observed on a creepmeter and on a water well level recording device nearby. On the other hand, a small number of strain episodes have been observed that were assumed to have been caused by aseismic slip that did not propagate past a creepmeter (Bufe and Tocher, 1972; Bufe et al., 1973; Stewart et al., 1973).

Beginning on 15 March 1972, a series of strain events were recorded over a six-month period at Stone Canyon Geophysical Observatory, south of Hollister, California. One of these could be classified as a slow earthquake as evidenced by the strain record shown in Fig. 4 (taken from Bufe and Tocher, 1972). A creep event recorded the previous day on a creepmeter seven kilometers distant from the epicenter is shown for trace form comparison. The strain

record was taken at the time of an earthquake of magnitude 4.7 which occurred in the immediate vicinity of the observatory. The record shows the seismic oscillation superposed upon a transient deformation of about one hour duration. The slow transient was interpreted by Bufe and Tocher to be the result of aseismic slip very near the instrument. Data loss on the other two strain components prevented the solution of the inverse problem as done by Pfluke and Stewart.

One of the strain transients observed in this sequence was well-recorded on all three strain components. Tracings of that event, which occurred on 15 March 1972 are shown in Figure 5 (taken from Bufe et al.). By constraining a number of the parameters using geologic evidence, it was possible to solve the inverse problem at different increments of time for the sequence of dimensions of the growing plane of aseismic slip. The first and last of the incremental solutions are shown in Figure 6 (also taken from Stewart et al.).

Summary and Discussion

Two earthquakes (1946 Nankaido, 1975 Hokaido) were found to have seismic moments that were much larger than those which would normally have been expected from the conventional magnitude-moment relationship. An hypothesis that much of the fault slip occurred at a rate that is slow compared to the response of our seismographs, but yet fast when compared to the response time of tsunami generation is solidly supported by the work of Ando, Geller, and Sacks and associates. An examination of the catalog of large tsunamis in the

but yet fast when compared to the response time of tsunami generation is solidly supported by the work of Ando, Geller, and Sacks and associates. An examination of the catalog of large tsunamis in the Pacific further indicated that a few of these tsunamis could well have been generated by similar such slow earthquakes. An indicator that supports such a hypothesis is an abnormally high ratio of tsunami magnitude to earthquake magnitude. From the evidence reviewed here, it would appear that slow earthquakes are probably commnplace events.

Case histories of large, truly aseismic, strain transients that occurred in the Aleutian Islands and Nevada give evidence that large-scale aseismic fault slip may also be a commonplace occurrence. But if these phenomena are commonplace, then why have so few been documented in the literature? I think the answer lay in the general nature of the instrumentation that is in widespread use. Many short and medium period seismographs are in operation, but comparatively few strain measuring devices capable of detecting very slow deformation are in use. The ultra-long period seismic records of the type used by Geller have been deployed only in recent years. Furthermore, the rapid fall-off of the static strain field with distance coupled with the high noise customarily experienced by strain meters require that a natural event occur very close to an instrument (probably within 20 faults lengths) before it can be detected. With the deployment of larger numbers of strain sensitive devices that is taking place in the Earthquake Hazards Reduction Program, we may find that the occurrence of slow earthquakes and very slow earthquakes occur far more frequently than had heretofore been suspected.

Table I

Date	Origin	M	M_t	M_t/M_s
1932	Jalisco, Mexico	8.1	2	.25
1933	Sanriku, Japan	8.3	4.8	.57
1935	Hawaii	--	2	--
1946	E. Aleutian Is.	7.4	5	.68
1946	Nankaido, Japan	8.1	2.4	.30
1952	Hokaido, Japan	8.1	2.0	.25
1952	E. Kamchatka	8.25	4	.48
1957	Andreanof Is.	8.3	3.5	.42
1958	S. Kuril Is.	8.7	2	.23
1960	S. Chile	8.5	4.5	.53
1963	S. Kuril Is.	8.25	2	.24
1963	S. Kuril Is.	6.8	3.5	.51
1964	Gulf of Alaska	8.4	4.5	.54
1965	Rat Island	7.75	3	.40
1969	Kamchatka	7.1	3	.42
1971	New Ireland	7.5	2.5	.33
1971	Kamchatka	7.3	3	.41
1975	Hokaido, Japan	6.9	2.2	.31
1975	Hawaii	7.2	3	.41

REFERENCES

Ando, M., 1975, Source mechanisms and tectonic significance of historical earthquakes along the Nankai Trough, Japan, Tectonophysics, 27, 119-140.

Bufe, C. G., W. H. Bakun and D. Tocher, 1973, Geophysical studies in the San Andreas fault zone at the Stone Canyon Observatory, California, in Prodeedings of the Conference on Tectonic Problems of the San Andreas Fault System, Stanford U. Pubs., Geol. Sci., XIII, R. L. Kovach and A. Nur, eds.

Bufe, C. G. and D. Tocher, 1972, Strain and tilt changes associated with the Stone Canyon Earthquake of September 4, 1972, Earthquake Notes, 43 (3), 21-24.

Hatori, T., 1976, Wave source of the Hawaii tsunami of 1975 and the tsunami behavior in Japan (in Japanese), J. Seism. Soc. Japan, Ser. 2, 29, 355-364.

Iida, K., D. C. Cox, and G. Pararas-Carayannis, 1967, Preliminary catalog of tsunamis occurring in the Pacific Ocean, data report no. 5, Hawaii Institute of Geophysics, Univ. Hawaii, Honolulu.

Johnson, A. G., R. L. Kovach, A. Nur and J. R. Booker, 1973, Pore pressure changes during creep events on the San Andreas Fault, J. Geophys. Res. 78 (5), 851-857.

Kanamori, H., 1977, The energy release in great earthquakes, J. Geophys. Res. 82 (20), 2981-2988.

Major, M. W. and D. Tocher, 1971, Deformation without earthquakes, the Central Aleutians, Geol. Soc. Am. Abst. 3 (2), 154.

Mansinha, L., and D. E. Smylie, 1971, The displacement fields of inclined faults, Bull. Seism. Soc. Am. 61, 1433-1440.

Marquardt, D. W., 1963, An algorithm for least-squares estimation of nonlinear parameters, J. Soc. Indus. Appl. Math. 11, 431-441.

Nason, R. D., 1971, Investigation of fault creep slippage in northern and central California, Ph. D. dissertation, Univ. Calif. San Diego, 232 pp.

Pararas-Carayannis, G. and J. P. Calebaugh, 1977, Catalog of tsunamis in Hawaii, NOAA Rep. SE-4, U.S. Dept. Commerce, Boulder, Colo.

Pfluke, J. H. and R. M. Stewart, 1973, Aleutian strain events: observations and interpretation, Geophys. J. R. Astr. Soc. 35, 229-241.

Press, F., 1965, Displacements, strains, and tilts at teleseismic distances, J. Geophys. Res. 70 (10), 2395-2412.

Sacks, I. S., S. Suyehiro, A. T. Linde and J. A. Snoke, 1977, The existence of slow earthquakes and the redistribution of stress in seismically active regions (abs.), Trans. Am. Geophys. Un. 58 (6), 437.

Shimazaki, K., and R. J. Geller, 1977, Source process of the Kurile Islands tsunami earthquake of June 10, 1975 (abs.), Trans. Am. Geophys. Un. 58 (6), 437.

Smith, S. W. and R. Kind, 1972, Regional secular strain fields in southern Nevada, Tectonophysics 14, 57-69.

Stewart, R. M., C. G. Bufe, and J. H. Pfluke, 1973, Creep-caused strain events at Stone Canyon, California, Geol. Sci. XIII, R. L. Kovach and A. Nur, eds., Stanford University.

Tocher, D., J. H. Pfluke, and M. E. Blackford, 1971, Strain measurements in the Aleutian Islands, NOAA Technical Report ERL 215-ESL 17, U.S. Government Printing Office.

Wideman, C. J. and Major, M. W., 1967, Strain steps associated with earthquakes, Bull. Seism. Soc. Am. 57 (6), 1429-1444.

CAPTIONS FOR TABLE AND FIGURES

Table 1 List of tsunami generating earthquakes occurring since 1932.

Fig. 1 Examples of secular strain recorded in the Aleutian Islands at Adak (ADK) and Amchitka (ANW).

Fig. 2 Record tracings of the Aleutian strain episodes of October 1970.

Fig. 3 Map of the central Aleutian Islands, showing locations of strainmeters and a surface projection of the fault plane of the 16 October 70 very slow earthquake

Fig. 4 Creep event and slow earthquake recorded at Stone Canyon Geophysical Observatory, California.

Fig. 5 Record tracings of strain event of 15 March 72.

Fig. 6 First and final fault plane solutions of 15 March 72 slow earthquake as derived from strain data.

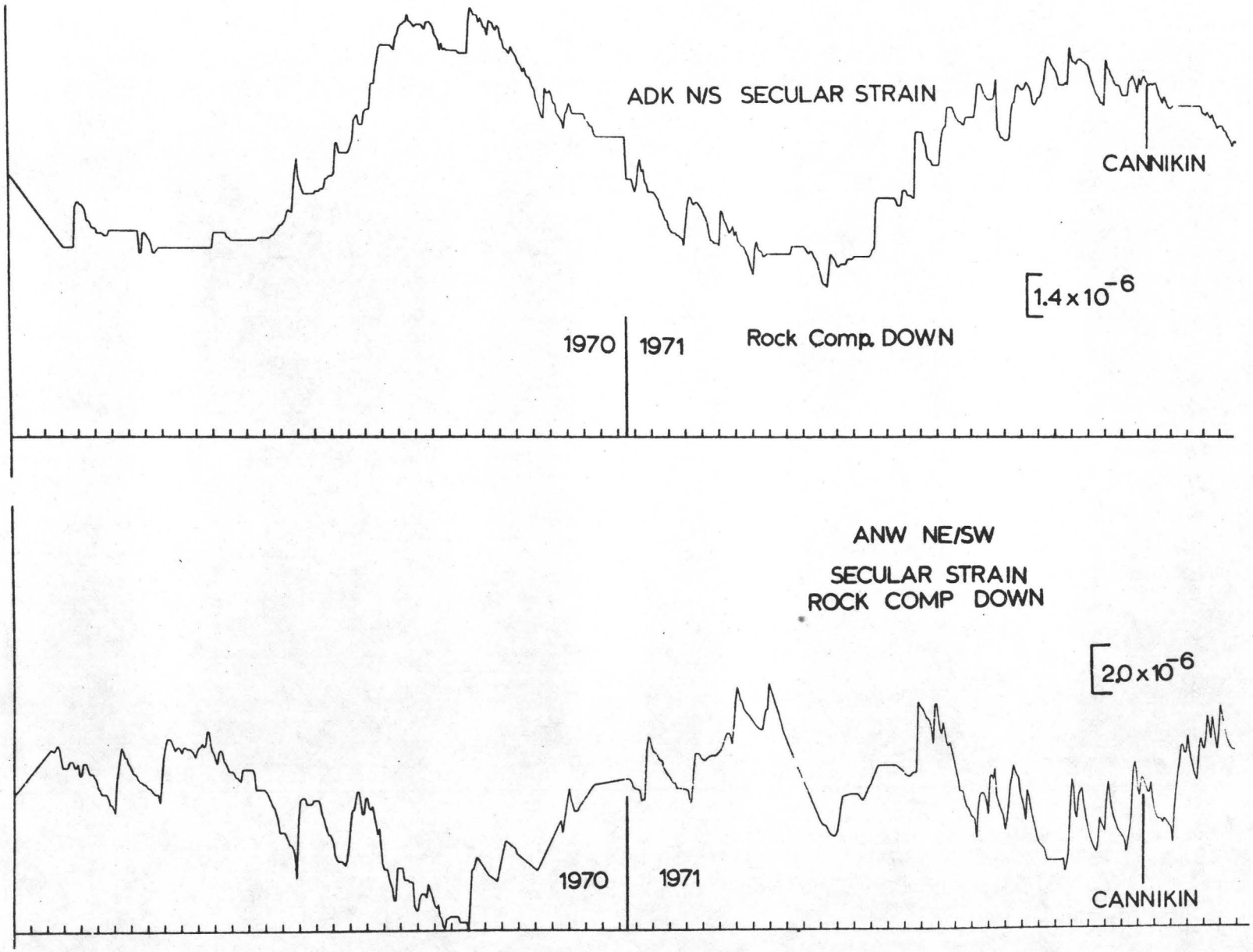

ADK N/S SECULAR STRAIN
CANNIKIN
1.4 x 10⁻⁶
1970
1971
Rock Comp. DOWN
ANW NE/SW
SECULAR STRAIN
ROCK COMP DOWN
2.0 x 10⁻⁶
1970
1971
CANNIKIN

Figure 1

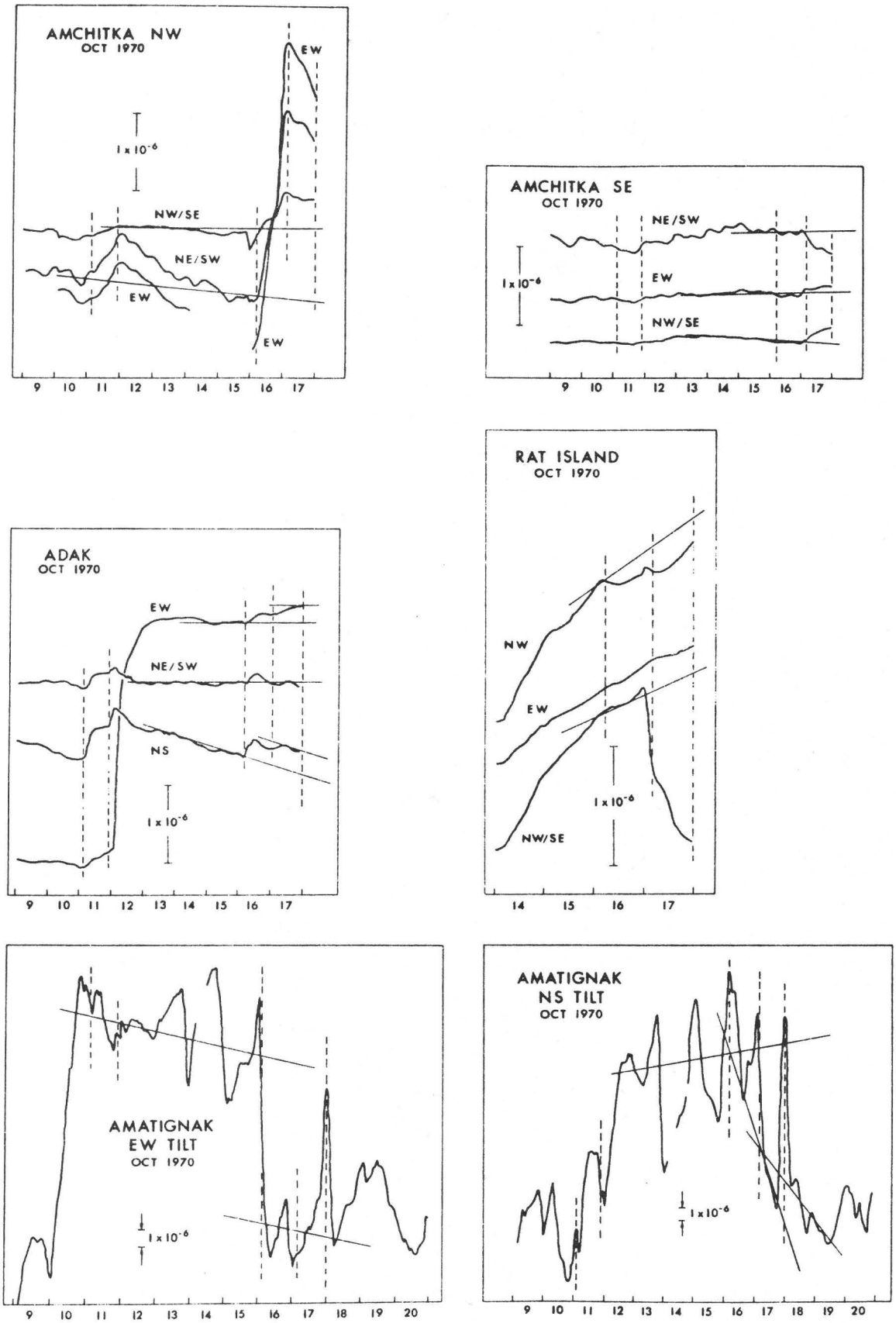

AMCHITKA NW
OCT 1970
EW
1 x 10⁻⁶
NW/SE
NE/SW
EW
EW
9 10 11 12 13 14 15 16 17
AMCHITKA SE
OCT 1970
NE/SW
1 x 10⁻⁶
EW
NW/SE
9 10 11 12 13 14 15 16 17
RAT ISLAND
OCT 1970
NW
EW
1 x 10⁻⁶
NW/SE
14 15 16 17
ADAK
OCT 1970
EW
NE/SW
NS
1 x 10⁻⁶
9 10 11 12 13 14 15 16 17
AMATIGNAK
EW TILT
OCT 1970
1 x 10⁻⁶
9 10 11 12 13 14 15 16 17 18 19 20
AMATIGNAK
NS TILT
OCT 1970
1 x 10⁻⁶
9 10 11 12 13 14 15 16 17 18 19 20

Figure 2

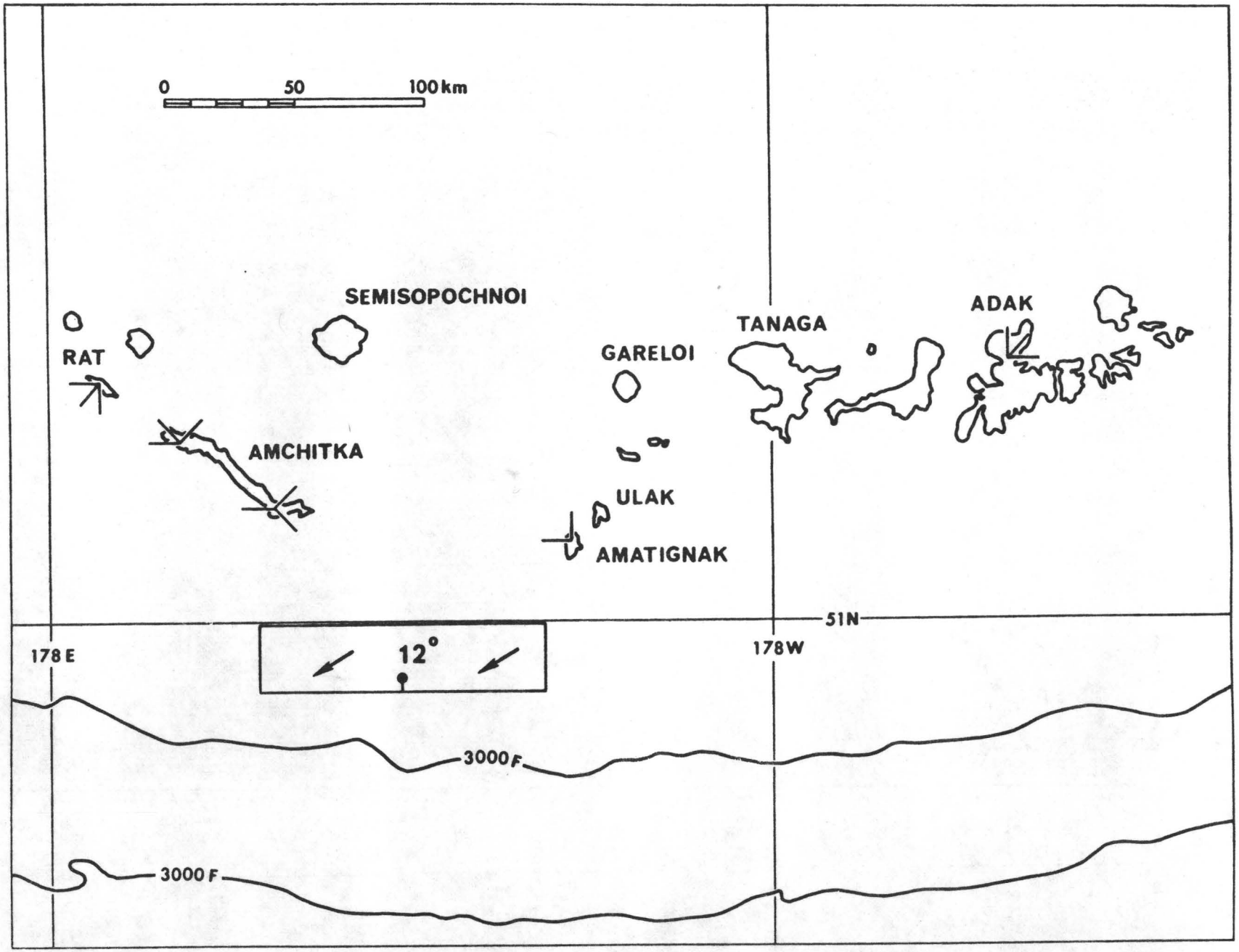

0
50
100 km
SEMISOPOCHNOI
RAT
AMCHITKA
GARELOI
TANAGA
ADAK
ULAK
AMATIGNAK
51N
178 E
178W
12°
3000 F
3000 F

Figure 3

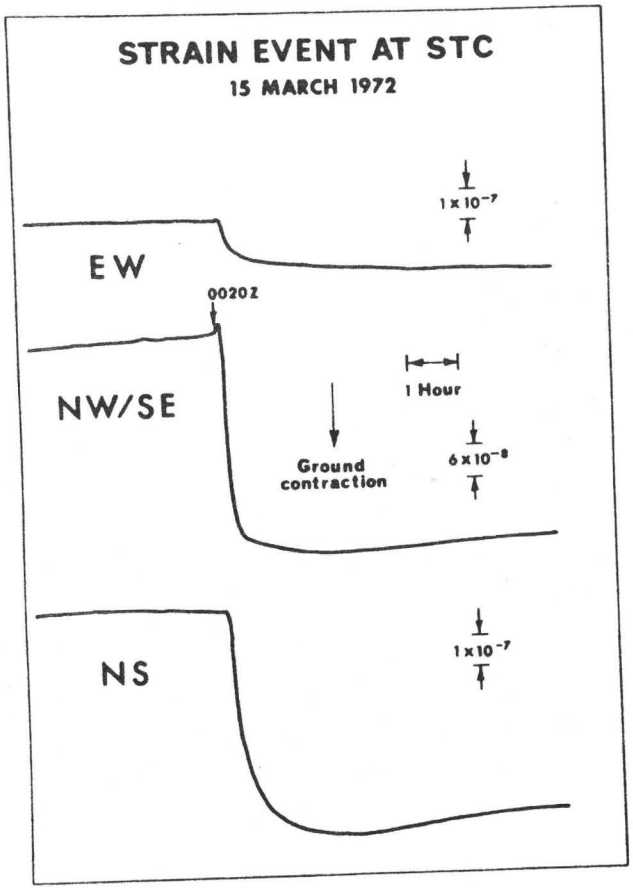
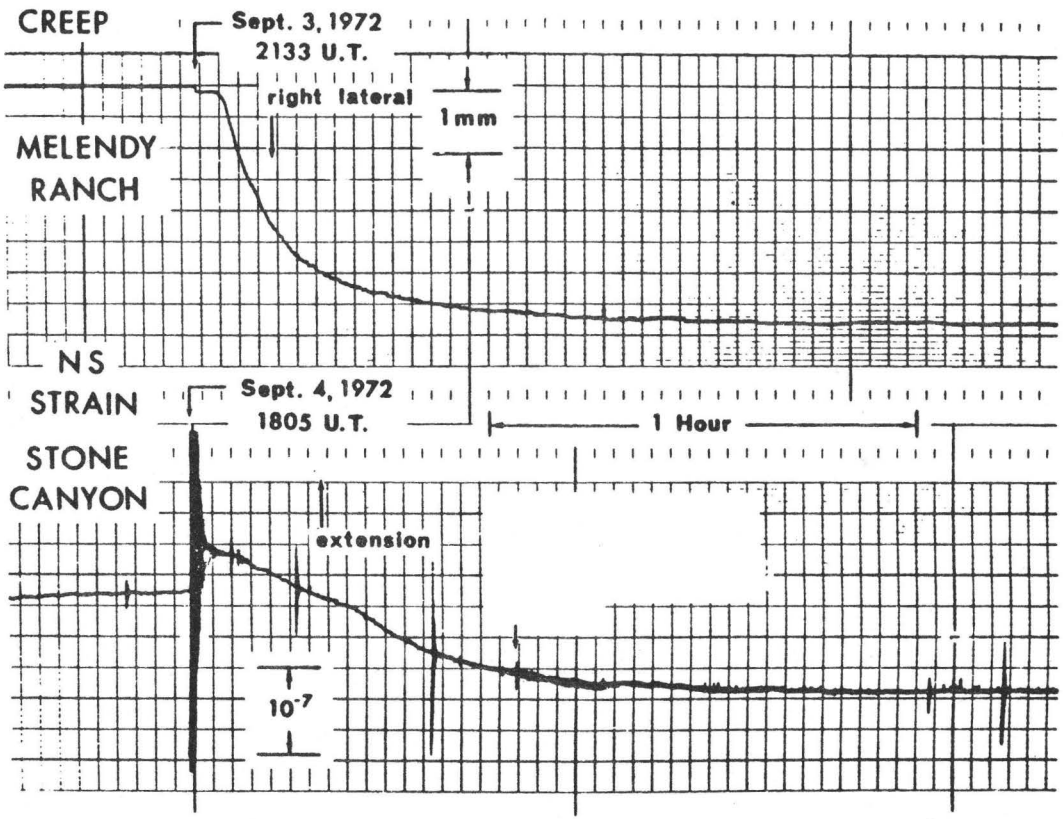


Figure 4



Top: Creep event recorded at Melendy Ranch on September 3, 1972

Bottom: Stone Canyon strainmeter recording of September 4 earthquake and episode

Figure 5

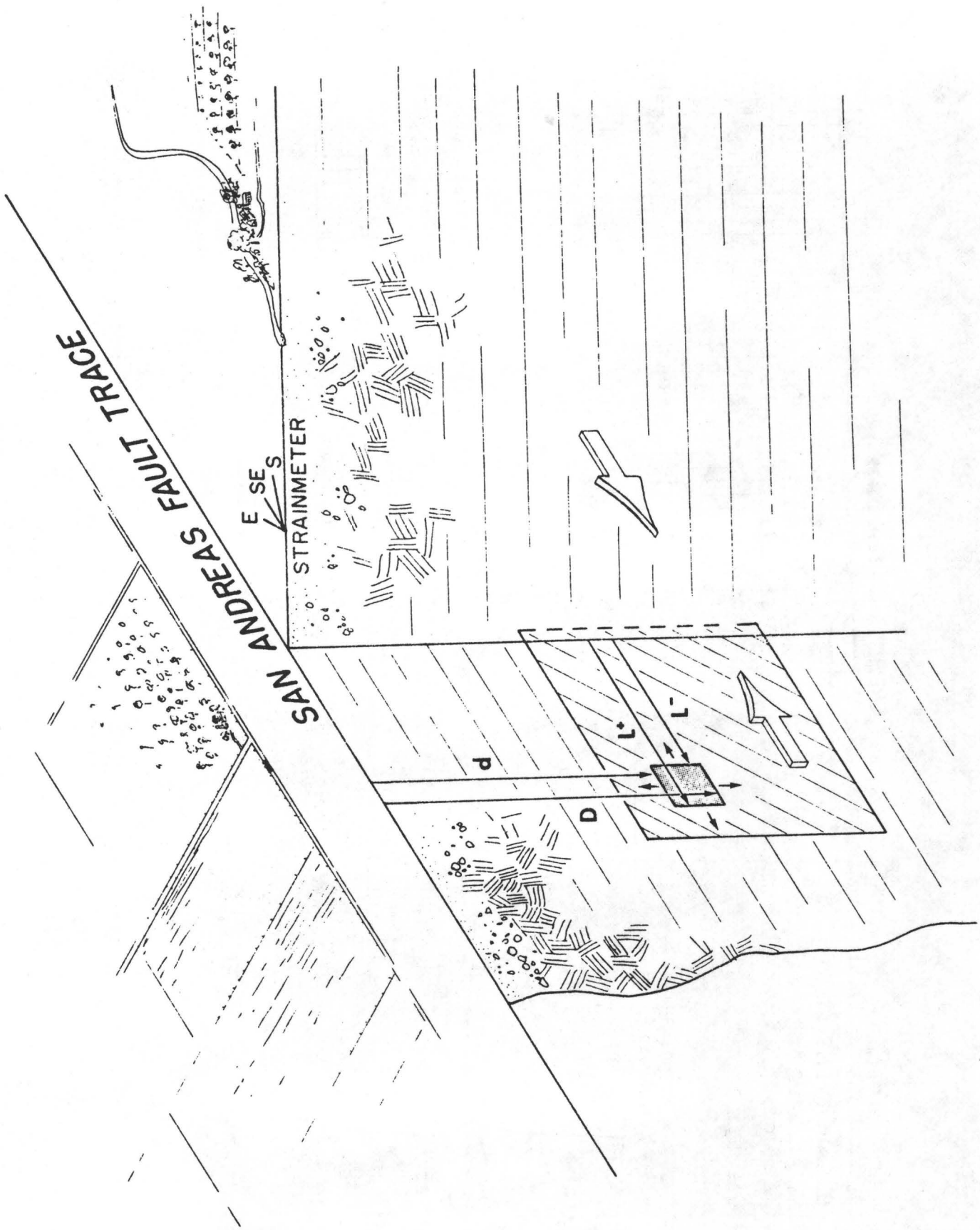


Figure 6

EARTHQUAKE PRECURSORY EFFECTS DUE TO PORE FLUID STABILIZATION OF A WEAKENING FAULT ZONE

by

James R. Rice
Division of Engineering
Brown University
Providence, Rhode Island 02912

and

John W. Rudnicki
Seismological Laboratory
California Institute of Technology
Pasadena, California 91125

ABSTRACT

We report the analysis of two mechanisms by which pore fluids could partially stabilize the earthquake rupture process in natural rock masses. These mechanisms are based on dilatancy strengthening and on the increase of elastic stiffness for undrained as opposed to drained conditions. Both are studied in relation to an inclusion model in which a zone of strain weakening material, possibly representing a highly stressed seismic gap zone, is embedded in nominally elastic surroundings subjected to steadily increasing tectonic stress. Due to the coupling between deformation and pore fluid diffusion the inclusion does not exhibit an abrupt rupture instability but, rather, a period of self-driven precursory creep occurs which ultimately accelerates to dynamic instability. The precursory time scale is reported for a wide range of constitutive parameters, including fluid diffusivity, ratio of undrained to drained stiffness, and factors expressive of strain softening and dilatancy. Our conclusions are that the precursory times for a spherical inclusion of 1 km radius are of the order of 15 to 240 days for a range of constitutive parameters that we suggest are representative. The predicted times are shorter by a factor of approximately 10 for a flattened ellipsoidal inclusion that we analyze with an 18:1 aspect ratio. It is suggested that perhaps only toward the latter part of the precursory period are the effects of accelerating inclusion strain detectable in terms of surface deformation or alteration of transport or seismic properties.

INTRODUCTION

We consider the possibility that rock near sites of shallow earthquakes is infiltrated with groundwater, and examine quantitatively processes by which mechanical coupling of rock deformation and pore fluid diffusion could transiently stabilize the rock against rapid failure. The stabilization results from dilatant strengthening within a fault region and from time-dependent relaxation of the effective elastic unloading stiffness of the surroundings from undrained to drained conditions. Both mechanisms are shown to allow failure to occur in a less abrupt manner than predicted without consideration of fluid coupling. Instead of an instantaneous dynamic instability, there is a period of initially quasi-static deformation that proceeds on a time scale governed by pore-fluid diffusion and non-elastic deformation characteristics of the failing material, and that ultimately accelerates to dynamic instability.

The time-scale and character of this quasi-stable deformation has interest as a possible basis for discernible precursors to earthquakes, in the form of accelerating strain or tilt and local pore pressure alterations. Our aim in the present work is to develop quantitative estimates of these precursory processes on the basis of mechanically consistent models of the failure process. In doing so we do not, of course, claim that all time-dependent precursory phenomena is traceable to such mechanical effects of pore fluids in an, essentially, deformation-rate-insensitive rock skeleton. Indeed, some precursory effects arise inevitably from the amplification of local fault region strain rates over remote tectonic strain rates as instability conditions are approached (e.g., Rice, 1977a). Also, time-dependence from corrosive microcracking of stressed rock, creep processes at depth, time-dependent

adherence at frictional contacts, and perhaps other mechanisms could be important in different settings.

The first mechanism of fluid coupling that we consider, namely dilatant hardening, is well known in the mechanics of granular materials (Reynolds, 1886). Because rock masses and fault zones typically dilate when deformed inelastically, the local pore fluid pressure will decrease unless the deformation is sufficiently slow to permit alleviation of the induced suctions by diffusion. This decrease in pore fluid pressure causes the effective stress (that is, the total stress minus the pore fluid pressure) to increase in compression, thereby inhibiting inelastic deformation mechanisms such as frictional sliding and microcracking. The relevance of dilatant hardening to seismic processes was first suggested by Frank (1965) and it has been studied in the laboratory deformation of rock by Brace and Martin (1968). A preliminary analysis of the role of dilatancy in stabilization of faulting, in the spirit of that to be discussed here, has been given by Rudnicki (1977b).

The second mechanism by which the interaction of deformation and pore fluid diffusion can stabilize against rapid failure arises because of the dependence of fault instability conditions on the effective elastic stiffness of the fault's surroundings. Such stiffness dependencies are well known in laboratory studies of failure (Jaeger and Cook, 1976); the stiffer the loading apparatus, the more rock can be deformed stably into the post-peak range. For fluid infiltrated rock, the surroundings are elastically stiffer for stress alterations which are rapid compared to diffusive relaxation times than for long term stress alterations. If the time scale of deformation is rapid compared to that for diffusion so that the fluid mass in material elements remains constant, the response is termed "undrained," and is elastically stiffer than

the long-time or "drained" response for which the local pore fluid pressure is constant. This time dependence provides a mechanism of transient stabilization as discussed by Rice et al. (1977) and Rice (1977a) on the basis of analyses using the linear elastic theory of Biot (1941) for porous solids. Similar effects have been studied in problems related to aftershock mechanisms by Booker (1974).

Our analysis of both the foregoing pore fluid coupling effects is based on a model for the inception of faulting introduced by Rudnicki (1977a). This model considers the deformation of a rock mass which contains an ellipsoidal inclusion in which the properties are uniform but different from those of the surroundings (Fig. 1a). Rudnicki (1977a) interpreted this inclusion as deforming non-linearly and exhibiting a peak stress as a consequence of weakening by faulting, whereas the surrounding material remains nominally elastic. Rice (1977a), however, has indicated that the model applies as well and, indeed, may be more pertinent if the inclusion is interpreted as sustaining higher stresses than its surroundings. This would be the case if the inclusion had undergone less strain due to past faulting than the surroundings and is, in effect, a "seismic gap" zone. Because of the large local stress, the inclusion, as in Rudnicki (1977a), exhibits a peak stress whereas the surrounding material remains elastic.

In either interpretation, the slow increase of tectonic stress drives the inclusion material past peak stress. At some point, the slope of the descending stress-strain curve for the inclusion material becomes sufficiently negative so that no further increase of tectonic stress can be accommodated quasi-statically. This dynamic "runaway" is interpreted as the occurrence of an earthquake. As we shall see, and as suggested by prior discussions of

the inclusion model by Rudnicki (1977a,b), Rice et al. (1977), and Rice (1977a), the effect of fluid coupling mechanisms of the type discussed earlier is to cause this instability to occur not abruptly, but rather in a more gradual manner with a time scale controlled by fluid diffusion, and, possibly, with discernible precursors.

Despite the simplicity of this model, it is motivated by the idea that pre-failure processes are likely to be dominated by large scale heterogeneities of mechanical properties which remain as a result of the past history of faulting. Indeed, because the material properties and the geometry of such fault zones are at present so poorly known, it seems unwise to concoct overly-detailed models, although we do believe that simple analyses founded upon consistent mechanical principles are useful for identifying the salient features of precursory processes.

Although our discussion here is organized primarily with reference to the inclusion model, another model which is relevant and may be more appropriate in many circumstances is one in which non-elastic deformations are assumed to be localized along a narrow fault from the outset, and in which the slipping zone along the fault can propagate in a shear-crack-like mode (Fig. 1b). In the presence of an infiltrating pore fluid, both the stabilizing effects of dilatant hardening and of time-dependent stiffness of the surroundings are applicable here as well, as has been remarked by Rice and Cleary (1976). Rice and Simons (1976) examined further the stabilization which results from time-dependent elasticity by solving for the stress-concentration near the tip of a quasi-statically propagating shear crack in a fluid-infiltrated porous elastic solid. They found that for a range of crack speeds comparable

to propagation rates inferred for episodic creep events in central California (King et al., 1973), the crack extended stably in the sense that with increasing propagation speed an increase in far-field driving stress, relative to frictional resistance, was needed to maintain the same stress intensity near the tip. Also, by adapting an earlier analysis developed for slip surface propagation in overconsolidated clay soils, Rice (1977a) developed an expression for the increase in far-field driving stress that is required, with increasing propagation speed, to overcome the augmented frictional resistance due to dilatantly induced suctions from the shear "breakdown" process near a fault tip. The result suggests that the maximum induced suction and required driving stress increase approximately in proportion to $V^{1/2}$ at low speeds, where V is the propagation speed. These studies of pore fluid effects on the criterion for propagation of a shear crack suggest precursory effects similar to those that we discuss here based on the inclusion model. In the inclusion model as we present it, an entire zone deforms into the inelastic range and ultimately becomes unstable. The shear crack model is more complicated because it involves the gradual enlargement of the non-elastic (slipping) region. Fluid effects of the kind discussed above provide a possible mechanism for control of the time-scale and extent of enlargement of the slip region before an unstable, dynamic spreading of the slip zone occurs, just as they may control the time-scale and extent of non-elastic strain in the inclusion model. We leave as a goal for future work the quantification of the precursory time-scale predicted on the basis of the shear crack model, and concentrate here on the inclusion model.

In addition to the mechanical effects of the pore fluid, there are also surface-chemical effects. These seem to be important in time-dependent crack

growth in quartz-based rocks (e.g., Martin, 1972; Scholz, 1972; Swolfs, 1972; Martin and Durham, 1975; Anderson and Grew, 1977) and they may be significant in determining the strength of rocks on a time scale comparable to that for tectonic alterations in stress. The constitutive description of these effects, however, is as yet insufficient to permit incorporation into the rupture models discussed here. Their presence, as well as presence of time-dependence of the frictional resistance of rock (Dieterich, 1972, 1977), would not invalidate the mechanical effects of pore fluids that we discuss here, but would add additional components to the overall precursory time-dependent deformation.

We will begin by reviewing in more detail the inclusion model for instability based on the deformation of an inhomogeneous zone. Then we will analyze the stabilizing effects of the pore fluid by employing the solution of Rice et al. (1977), discussing dilatant hardening on the basis of the work of Rudnicki (1977b) and time-dependent stiffness on the basis of that by Rice (1977a). In particular, we will demonstrate that both the mechanisms of dilatant hardening of the inclusion material and the time-dependence of elastic stiffness of the surrounding material can stabilize the rock mass against dynamic failure at a point where runaway would occur if the infiltrating fluid were not present. The subsequent deformation is initially slow but is self-driving and ultimately accelerates to dynamic instability. Because this self-driven accelerating deformation may display itself in observable precursors, such as accelerated creep or anomalous tilting of the ground surface, we estimate its time duration based on the models presented here and the incorporation of plausible material properties.

INSTABILITY OF A ROCK MASS WITH AN INHOMOGENEOUS ZONE

As discussed in the Introduction, we consider a rock mass containing an inclusion in which the mechanical properties differ from those of the surroundings (Fig. 1a) and for convenience of analysis, we assume that the inclusion is ellipsoidal in shape. For the present, we neglect pore fluid effects and consider the material surrounding the inclusion to be linearly elastic. If, in addition, the inclusion material is homogeneous, the inclusion deforms homogeneously (Eshelby, 1957) even if its material is nonlinear. As a result, the difference between the uniform strain of the inclusion $(\varepsilon_{ij})_{inc}$ and that applied in the far-field $(\varepsilon_{ij})_{\infty}$ is related linearly to the corresponding difference in stresses:

$$(\varepsilon_{ij})_{inc} - (\varepsilon_{ij})_{\infty} = Q_{ijk\ell}[(\sigma_{k\ell})_{\infty} - (\sigma_{k\ell})_{inc}] \quad , \tag{1}$$

where $(\varepsilon_{ij})_{\infty}$ is related to $(\sigma_{ij})_{\infty}$ by the elastic constants of the surrounding material. The tensor $Q_{ijk\ell}$ depends on the geometry of the inclusion and the elastic constants of the surrounding material, but not on the inclusion properties (as long as they are homogeneous). In particular, eq. (1) pertains whether the inclusion material deforms linearly or nonlinearly. Expressions for $Q_{ijk\ell}$ in specific cases have been given by Rudnicki (1977a,b).

The state of the inclusion at a particular level of far-field stress may be deduced from a simple graphical construction based on (1). Following Rice et al. (1977), we consider the component of (1) for a single shear stress as illustrated in Fig. 1a. Then, if γ denotes the "engineering" shear strain,

$$\gamma_{inc} - \gamma_{\infty} = \frac{\xi}{G}(\tau_{\infty} - \tau_{inc}) \tag{2}$$

where G is the shear modulus of the surroundings and ξ a shape factor.

This equation is plotted in Fig. 2 as the "Eshelby" line of negative slope G/ξ, along with the stress-strain relation for the far field, $\tau_\infty = G\gamma_\infty$, and the non-linear τ_{inc} vs. γ_{inc} relation for the inclusion. If τ_∞ has its value at point A in Fig. 2, the stress in the inclusion is given by the value at point A'; that is, the intersection of the τ_{inc} vs. γ_{inc} curve with Eshelby line through point A. As τ_∞ is increased, the successive states in the inclusion traverse the τ_{inc} vs. γ_{inc} curve as illustrated schematically in fig. 2b. When τ_∞ has reached its value at point B, the Eshelby line is tangent to the inclusion stress strain curve at B'. No further increase of τ_∞ can be sustained quasi-statically and a dynamic runaway of the inclusion shear strain occurs.

The parameter ξ in (2) is given by the following expressions (Rudnicki, 1977a):

$$\xi = \frac{2(4-5\nu_e)}{7-5\nu_e} \quad \text{, for a spherical inclusion} \tag{3}$$

$$\xi = \frac{4(1-\nu_e)}{\pi(2-\nu_e)} \frac{a}{b} \quad \text{, for a narrow axisymmetrical ellipsoidal inclusion} \tag{4}$$

$$\xi = \frac{(1-\nu_e)a}{\pi b} \quad \text{, for a narrow ellipsoidal inclusion in plane strain,} \tag{5}$$

where ν_e is Poisson's ratio in the surrounding material, and a and b are the semi-major and semi-minor axes of the ellipsoid. (The second expression is a correction by Rice (1977a) of a formula given by Rudnicki (1977a)). An examination of eqs. (3-5) reveals that the slope of the Eshelby line is less negative for narrow inclusions than for nearly spherical ones and, consequently, runaway occurs nearer to peak stress for the more narrow zones.

Using the construction of Fig. 2b, we see that an increase of local shear strain rate is predicted as the instability point is approached. In Fig. 2b

τ_∞ is increased in equal increments but it is evident that the corresponding increments of γ_{inc} increase in size. This acceleration of the local strain rate is a general precursory effect but whether it is sufficient to be observable in terms of surface deformation is unclear.

The graphical construction corresponding to the seismic gap interpretation is shown in Fig. 2c. Because the inclusion has undergone less strain than the surroundings, the peak of the τ_{inc} vs. γ_{inc} curve is drawn to the left of the line $\tau_\infty = G\gamma_\infty$. Therefore, the stress sustained by the "seismic gap" is always greater than τ_∞. That γ_{inc} is initially negative is inconsequential and results because the strains are measured relative to those corresponding to the unloaded state in the far-field. Otherwise, the analysis is the same as that for the "weakened zone" interpretation of Figs. 2a,b.

The τ_{inc} vs. γ_{inc} curve in Fig. 2 is to be regarded as the relation appropriate to in situ conditions. If the inclusion material dilates in response to shear, as is typical for brittle rock near failure, it must do so against the constraint of the surrounding material. The induced compressive stresses will then inhibit further inelastic deformation and effectively elevate the τ_{inc} vs. γ_{inc} curve over that for constant mean stress. These effects have been addressed by Rudnicki (1977a) for general deformation states, but it suffices for the purpose of the constructions in Fig. 2 simply to regard the τ_{inc} vs. γ_{inc} curve as drawn so as to incorporate the effects of dilatancy induced compression. Rudnicki (1977a) has also shown that prior to "runaway" instability critical conditions of the type discussed by Rudnicki and Rice (1975) will be met, beyond which the deformation pattern in the inclusion need not remain homogeneous but can begin to bifurcate into localized shear zones. Thus, it is appropriate to regard the stress-strain

relation of the inclusion as representing the overall response of a non-elastically deforming region even if the deformation is not locally homogeneous.

Effects of Pore Fluid Coupling

Before examining mathematically the stabilizing effects of coupling of the deformation with pore fluid diffusion, we discuss these effects qualitatively in terms of the schematic illustration of instability in Fig. 2b. It is evident that Fig. 2b may be altered in two ways to prevent instability at B': The local slope of the τ_{inc} vs. γ_{inc} curve may be increased; or the slope of the Eshelby line may be steepened. These correspond to the two mechanisms of stabilization which were discussed in the INTRODUCTION, namely dilatant hardening of the inclusion, and stiffening of the elastic response of the surroundings due to undrained conditions.

We consider separately each of the stabilizing mechanisms. First, we neglect the time-dependent stiffness of the surroundings and concentrate on dilatant hardening of the inclusion material. The relatively rapid deformation induced near B' will cause the slope $(d\tau/d\gamma)_{inc}$ to have its elevated, undrained value which is denoted by "u" in Fig. 3, reproduced from Fig. 4 of Rice (1977a). (We shall later show the relation of this undrained slope to the drained slope and to other constitutive parameters). Consequently, runaway instability will not occur at B'. If τ_∞ deviates only slightly from its value at B' (assumed to be the critical value predicted on the basis of drained response), the subsequent deformation will follow the Eshelby line rather than the continuation of the stress-strain curve for drained deformation (shown as a dashed curve in Fig. 3). Continued softening, however, will diminish the value of $(d\tau/d\gamma)_{inc}$ for both drained ("d") and undrained

("u") response (e.g., point C' in Fig. 3). Ultimately, the value of the slope for undrained response falls to the value of that for the Eshelby line and instability occurs at D'.

Similarly, we can neglect the changes in the τ_{inc} vs. γ_{inc} curve which are induced by dilatant hardening and focus on the stiffness changes due to time-dependent response of the surroundings (Fig. 4, reproduced from Fig. 5 of Rice, 1977a). Because of rapid deformation near B', the response of the surroundings will not be drained, as indicated by the solid line at B' in Fig. 4, but rather increments of inclusion and far-field deformation will be related according to the stiffer response appropriate to undrained conditions as indicated by the dashed line in Fig. 4. Although the τ_{inc} vs. γ_{inc} curve is not tangent to the dashed line at B', continued softening will reduce the value of $(d\tau/d\gamma)_{inc}$ until it equals the slope of the Eshelby line for undrained response at D'. Thus, instability is delayed beyond B' but occurs at D'.

In the remainder of the paper, we will examine these processes more precisely and concentrate on the time evolution of the system from point B' to point D' in Figures 3 and 4. For this purpose we require the extension of the Eshelby realtions (1) for application to a fluid-infiltrated solid. This generalization has been accomplished for spherical inclusions by Rice et al. (1977) and in the next subsection we will review their results. Although these relations apply rigorously only for spherical inclusions, the results for a more narrow zone can be approximated by an appropriate modification of parameters.

Eshelby Relations for a Fluid Infiltrated Elastic Solid

Rice et al. (1977) observe that when an unbounded fluid infiltrated elastic solid, containing a spherical cavity, is subjected to a sudden alter-

ation of pore pressure and surface traction (derivable from a homogeneous stress tensor) at the cavity wall, the wall displaces as if the cavity interior underwent a homogeneous deformation. This enabled them to generalize the Eshelby relations of eqs. (1,2) to spherical inclusions in fluid infiltrated solids, on the assumption that the inclusion is sufficiently permeable by comparison to its surroundings that pore pressure is effectively uniform within it. This assumption seems reasonable in the present context since the inclusion material is in the dilatant, strain softening range in the time scale over which we make use of the analysis.

Consequently, the strain and pore pressure fields are homogeneous within a spherical inclusion and, following Rice et al. (1977), the Eshelby relation of eq. (2) connecting the mismatch of shear stress and strain between the inclusion and the far-field is generalized to

$$\gamma_{inc}(t) - \gamma_{\infty}(t) = \frac{1}{G}\int_{-\infty}^{t} \{\xi_u + (\xi - \xi_u) f\left[\frac{c(t-t')}{a^2}\right]\}[\dot{\tau}_{\infty}(t') - \dot{\tau}_{inc}(t')]dt' \qquad (6)$$

where t is time, the dots denote time differentiation, G is the elastic shear modulus of the material surrounding the inclusion, a is the radius of the inclusion, and c is the diffusivity which appears in the porous medium equations (e.g., Biot, 1941; Rice and Cleary, 1976). The parameters ξ and ξ_u are given by (3) with the value of Poisson's ratio of the surroundings appropriate for drained (ν) and undrained (ν_u) deformation, respectively. (Of course, $\xi > \xi_u$ since $\nu < \nu_u$; some numerical estimates are given in a subsequent Table). Consequently, (6) reduces to the forms of (2) which are appropriate for undrained and drained response when the nondimensional function $f(ct/a^2)$ takes on its limiting values for short and long times, respectively,

$$f(0) = 0 \quad \text{and} \quad f(\infty) = 1 \ .$$

The complete variation of $f(\theta)$, where $\theta = ct/a^2$, is shown in Fig. 5 from Rice et al. (1977), for their $\eta = 0.8$ and 1.0. (The dependence of $f(\theta)$ on η is very weak and η itself varies only from 0.8 to 1.0 for representative material parameters).

It is convenient to have a version of (6) which involves only finite times t. Specifically, we assume that prior to $t = 0$ there are no excess pore pressures and the inclusion is in equilibrium with its surroundings based on the fully drained elastic properties of the surroundings. That is, prior to $t = 0$ the deformations are assumed to take place slowly enough so that (6) reduces to (2) with ξ based on the drained elastic properties, whereas deformation for $t > 0$ involves the coupling with diffusion discussed above. In that case it is straightforward to rewrite (6) as

$$\gamma_{inc}(t) - \gamma_{\infty}(t) = \frac{\xi}{G}[\tau_{\infty}(o) - \tau_{inc}(o)] + \frac{1}{G}\int_o^t \{\xi_u + (\xi-\xi_u) f\left[\frac{c(t-t')}{a^2}\right]\}[\dot{\tau}_{\infty}(t') - \dot{\tau}_{inc}(t')]dt' \qquad (7)$$

As mentioned earlier, the response of the inclusion to shear may be coupled to that for compression by dilatancy induced compressive stresses. Thus, for a complete analysis, the Eshelby relation for the hydrostatic component of deformation is needed. The appropriate special case of (1) is (Rudnicki, 1977a)

$$G[\varepsilon_{inc}(t) - \varepsilon_{\infty}(t)] = (3/4)[\sigma_{inc}(t) - \sigma_{\infty}(t)] \qquad (8)$$

where the strains ε and stresses σ (positive in compression) are the hydrostatic components, $\varepsilon = \varepsilon_{kk}$ and $\sigma = -\sigma_{kk}/3$. As shown by Rice et al. (1977), this same relation applies for fluid infiltrated solids whenever σ_{∞} and ε_{∞} change with time in a manner consistent with a constant fluid mass content at infinity. This includes the special case which

we will consider, for which σ_∞ and ϵ_∞ are taken to be constant as τ_∞ is increased. It is remarkable that the dilational Eshelby relation of (8) involves only the total hydrostatic stress σ_{inc} within the inclusion and, for a given σ_{inc}, is independent of the pore pressure p_{inc}. This feature follows from the solution to the Biot equations as developed by Rice and Cleary (1976) for simultaneous application of a total stress and pore pressure alteration to the wall of a spherical cavity.

One additional equation is needed to relate the pore fluid pressure $p_{inc}(t)$ in the inclusion to the fluid mass $m_{inc}(t)$ per unit volume within the inclusion. Rice et al. (1977) have obtained this equation by again taking advantage of the solution by Rice and Cleary (1976) for the pore pressure distribution induced outside a spherical cavity by sudden application of a pore pressure to the wall of the cavity. The mass flux rate into the cavity, as computed from the pore fluid pressure distribution outside the cavity and Darcy's law, is required to balance the rate of fluid mass accumulation at the cavity wall. The result, after extension by superposition to arbitrary time variation is (Rice et al., 1977)

$$m_{inc}(t) = -\frac{3\rho k}{a^2}\{p_{inc}(t) - p_\infty(t) + \int_{-\infty}^{t} \frac{a}{[\pi c(t-t')]^{\frac{1}{2}}}[\dot{p}_{inc}(t') - \dot{p}_\infty(t')]dt'\} \qquad (9)$$

where ρ is the density of the pore fluid in the material surrounding the inclusion (ρ is assumed to be spatially uniform except for small pressure dependent changes due to fluid compressibility), and where k is the permeability coefficient of the surroundings, defined so that Darcy's expression for the mass flux rate of pore fluid crossing unit area has the form $\rho k \nabla p$.

Sometimes k is written as k' (with units of area, measured in Darcies) divided by viscosity of the pore fluid; see Rice and Cleary (1976). Again, p_∞ and σ_∞ are assumed to vary in a manner that m_∞ is constant; for our calculation we will take all to be constant.

Thus, we wish to use eqs. (7,8,9) to analyze the inclusion model, assuming that $\dot{\sigma}_\infty = \dot{p}_\infty = 0$ and that some uniform tectonic shear stressing rate $\dot{\tau}_\infty$ is given. In order to complete the analysis of time-dependent response it is necessary to specify three constitutive relations for the inclusion material which relate its "strain" quantities γ_{inc}, ε_{inc}, and m_{inc} to its "stress" quantities τ_{inc}, σ_{inc}, and p_{inc}. For simplicity of analysis and clarity of presentation, the effects of dilatant hardening of the inclusion and of the time-dependent stiffness of the surroundings will be considered separately. Constitutive relations for the dilatant hardening analysis are given in the next section, and for that analysis we neglect the time-dependent stiffness of the surroundings implicit in eq. (7), replacing (7) by eq. (2). For the analysis of time-dependent stiffness effects it suffices to simply view τ_{inc} as a function of γ_{inc} without dilatancy effects, as in Figs. 2 and 4, and the analysis can then be based on eq. (7) without explicit consideration of eqs. (8,9).

Our inclusion model as based on eqs. (2) or (7) and (8,9) obviously regards the fault zone as being of small extent by comparison to other relevant dimensions. As such, its proximity to the earth's surface is neglected as is also the non-uniformity of rock properties with depth. These are important limitations (but removeable by more elaborate modelling) on the application of our results to large crustal earthquakes.

STABILIZATION BY DILATANT HARDENING

In this section, we will examine in detail the manner by which dilatant hardening can delay the onset of rapid failure and give rise to a period of initially slow, but accelerating deformation. For convenience the complementary mechanisms due to time-dependent stiffness of the surroundings will be neglected. This simplification corresponds to using (2), which we will assume relates rates of stress and strain, rather than (7) as the Eshelby relation for shear. We first however, will introduce constitutive relations for the inclusion which are intended to model the frictional, dilatant response of brittle rock.

Constitutive Relations

The constitutive relations employed to describe the response of the inclusion material are analogous to those which were introduced by Rice (1975), generalized to arbitrary deformation states by Rudnicki and Rice (1975), and used by Rudnicki (1977a,b). These relations are intended to describe both elastic response and frictional, dilatant inelastic response of compressed rock due to slip on nominally closed fissure surfaces and to microcrack growth from local tensile stress concentrations.

Consider a material element which is subjected to a hydrostatic stress σ (positive in compression), a shear stress τ, and a pore fluid pressure p. If the material responds elastically, the shear strain rate $\dot{\gamma}$ and volume strain rate $\dot{\varepsilon}$ may be written

$$\dot{\gamma} = \dot{\tau}/G \tag{10}$$

$$\dot{\varepsilon} = -(\dot{\sigma}-\dot{p})/K - \dot{p}/K_s \tag{11}$$

where G is the incremental shear modulus, K is the incremental bulk modulus for drained response, and K_s is the bulk modulus of the solid constituents. The combination $\sigma-p(1-K/K_s)$ is the form of the "effective

stress" which has been shown by Nur and Byerlee (1971) to be appropriate for elastic response. In general, however, an increment of deformation involves inelastic response and these contributions to the strain rate must be added to those of (10,11). Following Rice (1977a), we write the complete incremental stress-strain relations as

$$\dot{\gamma} = \dot{\tau}/G + [\dot{\tau} - \mu(\dot{\sigma} - \dot{p})]/h \tag{12}$$

$$\dot{\epsilon} = -(\dot{\sigma} - \dot{p})/K - \dot{p}/K_s + \beta[\dot{\tau} - \mu(\dot{\sigma} - \dot{p})]/h \tag{13}$$

where μ is a friction coefficient and β is a dilatancy factor which expresses the ratio of inelastic increments of volume strain to inelastic increments of shear strain. Estimates from experimental results for values of μ and β by Rice (1975), Rudnicki and Rice (1975) and Rudnicki (1977a,b) lie in the ranges 0.5 to 1.0 and 0.2 to 0.5, respectively. The "hardening" modulus (or "softening" if $h < o$) is related to the slope of the τ vs. γ curve for drained response at constant σ by

$$\left(\frac{\partial \tau}{\partial \gamma}\right)_{\text{drained}} = \frac{h}{1 + h/G} \qquad (\dot{\sigma} = o) \quad . \tag{14}$$

In general, all of the constitutive parameters may vary with the deformation, although the variation in h is typically most substantial.

The form of the effective stress which enters the inelastic contributions in (12) and (13) is $\sigma - p$. Recently, Rice (1977b) has shown rigorously that this is the appropriate form in describing the inelastic response which arises from slip at isolated asperity contacts and/or from local cracking at the tips of sharp micro-fissures.

One additional constitutive equation is needed for the fluid mass content m per unit volume, which is related to the apparent volume fraction

v of pore space by

$$m = \rho v$$

where ρ is the mass density of the homogeneous pore fluid. It is convenient to express m in terms of the stresses by using reciprocity relations (Biot, 1973; Rice, 1975) to deduce the form for the elastic portion of dv and assuming $d^p v = d^p \varepsilon$ for the inelastic portion (Rice, 1975). The latter has been shown (Rice, 1977b) to follow rigorously in the same circumstances for which the use of $\sigma - p$ for the inelastic effective stress measure is justified. The result, written in rate form, is

$$\frac{\dot{m}}{\rho} = \frac{v\dot{p}}{K_f} - \left[\frac{1}{K} - \frac{1}{K_s}\right](\dot{\sigma} - \dot{p}) - \frac{v}{K_s}\dot{p} + \beta\frac{[\dot{\tau} - \mu(\dot{\sigma} - \dot{p})]}{h} \tag{15}$$

where K_f is the bulk modulus of the pore fluid ($\dot{\rho} = \rho\dot{p}/K_f$) and the remaining quantities have been defined previously.

The stiffness of response for shearing under undrained conditions ($\dot{m} = 0$) may now be calculated and compared with the corresponding stiffness for drained conditions, (14). Assuming $\dot{\sigma} = 0$ and setting $\dot{m} = 0$ in (15) yields

$$\dot{p} = \frac{-\beta K'}{h + \mu\beta K'}\dot{\tau} \tag{16}$$

where

$$\frac{1}{K'} = \frac{1}{K} + \frac{v}{K_f} - \frac{1+v}{K_s} \quad . \tag{17}$$

By substituting (16) into (13), we obtain the stiffness of response for shear at constant hydrostatic stress under completely undrained conditions:

$$\left(\frac{d\tau}{d\gamma}\right)_{\text{undrained}} = \frac{h + \mu\beta K'}{1 + (h+\mu\beta K')/G} \qquad (\dot{\sigma} = 0) \quad . \tag{18}$$

Comparison of (18) with the analogous expression for drained response, (14), reveals that the effective slope of the τ vs. γ curve has been augmented

by an amount corresponding to the replacement of h by $h + \mu\beta K'$. The stiffened response under undrained conditions is depicted schematically in Fig. 3 by the arrow labelled "u". For representative values of μ , β , and K' this effect is substantial and may be sufficient to cause the slope for undrained response to be positive while that for drained response is negative.

For heavily fissured rock $K_s >> K$, K_f/v and therefore,

$$K' \simeq \frac{KK_f}{K_f + vK} \quad . \tag{19}$$

If K_f has a value appropriate for liquid water ($K_f \simeq 22$ kbar) and $v < 0.10$, as is typical of brittle rock, $K' \simeq K$. High temperatures, low pore pressure, or the presence of entrapped gas may, however, cause K_f to be reduced well below vK so that

$$K' \simeq K_f/v$$

and the dilatant hardening effect disappears in the limit as $K_f \to 0$.

Of course, the foregoing constitutive description is based on the assumption that the state of the pore fluid can be characterized by a single parameter, namely a pore pressure p , which is valid only when deformations are sufficiently slow that there is local pressure equilibrium between all fissure and pore spaces occupying what is regarded as a "point" in the continuum model of the material. This may not be the case in the presence of rapid deformations; some estimates of equilibrium times and generalizations of the pore pressure concept are considered by Cleary (1977) and O'Connell and Budiansky (1977). Indeed, some discussions of dilatancy (e.g., Nur, 1972) consider

the possibility that the effect may be so strong as to open substantial vapor-filled or "dry" crack space in rock that is otherwise liquid saturated. Such a concept seems to be widely associated with the term "dilatancy", and is consistent with alterations of seismic wave speeds. But it is important to realize that far milder dilatancy, insufficient to cause the opening of "dry" crack space and thus to affect seismic transmission, may nevertheless be present during a failure process and, possibly, be a major factor in controlling the time scale of that failure through the processes to be described.

Dilatant Hardening and Instability

In order to apply the results of the last section to the inclusion problem, we adopt (12), (13), and (15) as the constitutive laws for the inclusion material and employ them in conjunction with the Eshelby relations (2), (8), and (9). The analysis follows Chapter II of Rudnicki (1977b). Thus the relations between rates of stress and strain are, from (2) and (8),

$$\dot{\gamma}_{inc}(t) - \dot{\gamma}_{\infty}(t) = \frac{\xi}{G}\left(\dot{\tau}_{\infty}(t) - \dot{\tau}_{inc}(t)\right) \tag{20}$$

and

$$\dot{\varepsilon}_{inc}(t) = (3/4G)\dot{\sigma}_{inc}(t) \tag{21}$$

where we have assumed in (21) that $\dot{\sigma}_{\infty} = 0$. If (12) and (13) are combined with (20) and (21) to eliminate $\dot{\tau}_{inc}$ and $\dot{\sigma}_{inc}$, the results are

$$\dot{\gamma}_{inc} = \dot{\gamma}_{\infty}\left[1 + \frac{\kappa G}{h-h_d}\right] + \frac{\mu \dot{p}_{inc}\kappa\alpha(1+4G/3K_s)}{h-h_d} \tag{22}$$

and

$$\dot{\varepsilon}_{inc} = \frac{\beta G\alpha}{h-h_d}\dot{\gamma}_{\infty} + \frac{\alpha \dot{p}_{inc}}{K}\left[1 - \frac{K}{K_s} + \frac{\mu\beta K\alpha(1+4G/3K_s)}{h-h_d}\right] \tag{23}$$

where $\kappa = \xi/(1+\xi)$, $\alpha = 1/(1+4G/3K) = (1+\nu)/3(1-\nu)$, and we have assumed for simplicity that the elastic moduli of the inclusion are identical with those of the surroundings. (Rudnicki (1977b) has given the corresponding expressions if there is elastic mismatch between the inclusion and the surroundings but assuming effectively incompressible solid constituents for the inclusion $(K_s \to \infty)$). The parameter h_d is the value of h for which the ratio $\dot{\gamma}_{inc}/\dot{\gamma}_\infty$ in (22) becomes unbounded under completely drained conditions ($\dot{p} = 0$) and corresponds to the onset of dynamic runaway (point B' in Fig. 3):

$$h_d = -G/(1+\xi) - \mu\beta K/(1+3K/4G) \quad . \tag{24}$$

The first term on the right in (24) would result from equating the right hand side of (14) to the slope of the Eshelby line $(-G/\xi)$ whereas the second term reflects the inhibiting effect of dilatancy induced compressive stresses. In other words, the hardening modulus corresponding to the <u>in situ</u> stress strain curve of the inclusion is

$$H = h + \mu\beta K/(1+3K/4G) \quad .$$

In order to demonstrate that the stiffened response to undrained deformation (compare (18) with (14) or "u" with "d" in Fig. 3) can in fact stabilize against runaway at $h = h_d$, we again consider completely undrained conditions. First, using (13) we can rewrite (15) as

$$\frac{\dot{m}}{\rho} = \frac{v\dot{p}}{K_f} + \frac{(\dot{\sigma}_{inc} - v\dot{p})}{K_s} + \dot{\varepsilon}_{inc} \tag{25}$$

where we here and subsequently drop the designation "inc" from $\dot{m}$ and $\dot{p}$. Eliminating $\dot{\varepsilon}_{inc}$ and $\dot{\sigma}_{inc}$ from (25) by means of (21) and (23) yields the rate of pore fluid pressure decrease in the inclusion

$$(\dot{p})_{undrained} = \frac{-\beta G K_f' \dot{\gamma}_\infty}{h - h_d + \alpha\mu\beta K_f'(1+4G/3K_s)} \tag{26}$$

where

$$\frac{1}{K_f'} = \frac{v}{\alpha K_f}\,\frac{(1-K_f/K_s)}{(1+4G/3K_s)} + \frac{1}{K}\left(1 - \frac{K}{K_s}\right) \quad .$$

Substituting (26) into (22) yields

$$\dot{\gamma}_{inc} = \dot{\gamma}_{\infty}\left[1 + \frac{\kappa G}{h-h_d + \alpha\mu\beta K_f'(1+4G/3K_s)}\right] \tag{27}$$

and demonstrates that the ratio $\dot{\gamma}_{inc}/\dot{\gamma}_{\infty}$ remains finite at $h = h_d$. Even if the rock mass is constrained to deform in a completely undrained fashion, however, $\dot{\gamma}_{inc}/\dot{\gamma}_{\infty}$ does become unbounded at

$$h = h_u \equiv h_d - \alpha\mu\beta K_f'(1+4G/3K_s) \tag{28}$$

which corresponds to point D' in Fig. 3. The discussion of (28) in cases of limiting behavior for K_f' follows the earlier remarks for K'.

Although dilatant hardening can stabilize against the onset of rapid failure at $h = h_d$, the subsequent deformation is self-driving (Rudnicki, 1977b). More precisely, Rudnicki (1977b) showed that in a segment of constant h (and other constitutive parameters), any perturbation of γ_{inc} from its equilibrium value corresponding to some fixed τ_{∞} is stable if $h > h_d$ (i.e., prior to point B' in Fig. 3), in the sense that the perturbation decays with increasing time. However, when $h_d > h > h_u$ (i.e., for states between B' and D' in Fig. 3) perturbations in γ_{inc} grow exponentially with increasing time and the rate of growth becomes unbounded when $h = h_u$. We define the precursor time as the time which elapses from the onset of this period of self-driven deformation at B' to dynamic instability at D'.

In order to determine the evolution of the inclusion state and, thus, the precursor time, it is necessary to solve the system of equations consisting of the constitutive laws (12), (13), and (15) and the Eshelby relations

(20), (21), and (9). The number of governing equations may, however, be reduced to two: eq. (22) and a single equation for the alteration in pore fluid pressure $p(t)$. The latter can be obtained by using (21) and (23) to eliminate $\dot{\sigma}_{inc}$ and $\dot{\varepsilon}_{inc}$ from (25) which then becomes

$$\frac{\dot{m}}{\rho} = \frac{\alpha(1+4G/3K_s)}{K_f'(h-h_d)}\left[\beta G\dot{\gamma}_\infty + \dot{p}(h-h_u)\right] , \tag{29}$$

where K_f' and h_u have been defined in (26) and (28). Combining (9) with (29) yields

$$t_D\dot{p}(t) = -\frac{\beta K_f' G\dot{\gamma}_\infty t_D}{h-h_u} - \left[\frac{3K_f'}{\alpha(1+4G/3K_s)N}\right]\frac{h-h_d}{h-h_u} \times$$

$$\left\{p(t) + \int_o^t \frac{a}{\sqrt{\pi c(t-t')}}\dot{p}(t')dt'\right\} \tag{30}$$

where $t_D = a^2/c$ is the diffusion time, $p(t)$ is now measured relative to the ambient value which exists in response to the constant hydrostatic stress at infinity, and the lower limit of the integral has been chosen by taking $p(t) = o$ for $-\infty < t \leq o$. The parameter $N = c/k$ is an additional elastic modulus which Rice and Cleary (1976) have shown may be written as

$$N = \frac{2G(1-\nu)B^2(1+\nu_u)^2}{9(1-\nu_u)(\nu_u - \nu)}$$

where ν_u is the Poisson's ratio for undrained response and B is the ratio of pore fluid pressure decrease to mean normal stress increase for an increment of undrained deformation.

The dependence of h and, in general, all of the constitutive parameters, on the deformation couples (30) to (22). In solving these equations, we will assume, however, that the variation in all the parameters except h is small

enough that they may be treated as constant. Although this may be a poor approximation in the case of β and K_f, so little is known about the details of their variation that assuming they are constant seems justified in the interests of simplicity. The simple numerical procedure which was used to solve (30) and (22) is outlined in Appendix I. Results of the calculations and a discussion of the constitutive parameters which were used are presented in the next section.

Numerical Results for Precursor Time

As mentioned earlier, the precursor time is defined as the time which elapses from the onset of self-driven deformation at B' (Fig. 3) to instability at D'. For the calculation, we assume that the in situ τ_{inc} vs. γ_{inc} curve has the form of a linear segment of slope G up to an elastic limit and that this segment is connected smoothly to the following parabola (Fig. 6):

$$\tau_{inc} = \tau_p - \frac{G}{2\lambda}(\gamma_{inc} - \gamma_p)^2 \qquad (31)$$

where γ_p is the strain at peak stress, λ is the difference between γ_p and the strain at the elastic limit, and G is the slope at the elastic limit. Because this represents the in situ stress-strain curve, the value of $H = h + \mu\beta K/(1+3K/4G)$ (see following (24)) is related to $d\tau_{inc}/d\gamma_{inc}$ by

$$\frac{d\tau_{inc}}{d\gamma_{inc}} = \frac{H}{1+H/G} \ .$$

In particular, we choose $\tau_p = 1\text{kb}$, $\gamma_p = 6.25 \times 10^{-3}$, $\lambda = 2.5 \times 10^{-3}$, and $G = 200$ kb. For the remaining material paremeters we assume $\nu = 0.2$, $\mu = 0.6$, $\beta = 0.3$, $N = 2.0G$, $K_f' = 1.45G$, and $K_s = 5G$. The

value of N corresponds to using $\nu_u = 0.37$ and $B = 0.8$, and is in the range, though toward the lower end, of values inferred from Table 1 of Rice and Cleary (1976). The value of K_f' is based on an initial porosity of $v = 0.01$ and assuming $K_f = 22\text{kb}$. The calculation was begun at peak stress where it was assumed the alteration in pore fluid pressure (taken to be zero at the elastic limit) was given by the steady state term in (30)

$$p(H = 0) = -\frac{\beta\alpha(1+4G/3K_s)N}{3(-H_d)}\ \dot{\tau}_\infty t_D$$

where $H_d = -G/(1+\xi)$.

The dimensionless precursor time $\theta_{prec} = t_{prec}/t_D$ is shown in Fig. 7 as a function of the dimensionless far-field strain-rate $\dot{\gamma}_\infty t_D$. It is evident that the predicted precursor time does not simply scale with $\dot{\gamma}_\infty t_D$. In particular, although t_{prec} increases with $t_D = a^2/c$, the increase is not proportional.

The time history of post-peak straining is shown in Fig. 8 and Fig. 9 for $\dot{\gamma}_\infty t_D$ equal to 1.4×10^{-5} and 1.4×10^{-6} respectively. The strain is given in terms of the parameter $(\gamma_{inc} - \gamma_p)/(\gamma_{D'} - \gamma_p)$, where γ_p is the strain at peak stress and $\gamma_{D'}$ is the strain at point D' in Fig. 3. Point B' corresponds to the onset of the period of self-driven deformation. The dashed portion of the solid curve indicates where the numerical calculation was truncated (see Appendix I). The fully-dashed curves in Figs. 8b, 9b incidate for comparison the strain history in the absence of pore-fluid effects. Those were calculated by setting $\dot{p}_{inc} = 0$ in eq. (22) and integrating for γ_{inc} as a function of time. It is evident that dilatant hardening not only delays the onset of instability, but also gives rise to a more dramatic acceleration of precursory strain. In Figs. 8c, 9c, the

decrease in the pore fluid pressure of the inclusion is shown. It is noteworthy that the decrease is extremely rapid but occurs very close to final instability. This suggests that precursory phenomena which may be associated with the rapid decrease in pore fluid pressure may occur over a time period which is substantially less than the period of self-driving deformation.

The results are shown in dimensional form in Table 1. We consider three values for the radius of the sphere, namely, $a = 1$ km , 3 km , 5 km and two values of the diffusivity $c = 1m^2/sec$, $0.1m^2/sec$. The larger value of the diffusivity was suggested by Anderson and Whitcomb (1975) as reasonable for shallow earthquake zones whereas $c = 0.1m^2/sec$ is more in accord with well-head measurements near the San Andreas Fault (Kovach et al., 1975). The value of $\dot{\gamma}_\infty$ was chosen to correspond to $\dot{\tau}_\infty = 1$ bar/year, i.e., $\dot{\gamma}_\infty = \dot{\tau}_\infty/G$, which is consistent with the centennial occurrence of a large earthquake having a stress drop of 100 bars. Strain accumulation measurements along the San Andreas Fault near Palmdale (Prescott and Savage, 1976) suggest a value of $\dot{\tau}_\infty$ which is roughly an order of magnitude smaller so that values of $\dot{\gamma}_\infty t_D$ which are smaller than those plotted in Fig 7 may be relevant. However, because of the expense of computation for small values of $\dot{\gamma}_\infty t_D$, the precursor time for $a = 1$km and $c = 1m^2/sec$ was simply estimated by extrapolating the curve in Fig. 7. The values in Table 1 again indicate that although t_{prec} increases with the diffusion time a^2/c , the increase is not proportional. A tenfold increase in a^2/c corresponds to an increase of roughly four to five times t_{prec}.

The last row of Table 1 gives the "best fit" correlation by Scholz et al. (1973) of precursory times (based on V_p/V_s anomalies, radon emission, and crustal movements) with the length of aftershock zone. Individual data

points may differ by a factor of approximately two from this time. The values in Table 1 identify 2a with the length of the aftershock zone. Comparison of the calculated with the observed precursor times indicates that the calculated values are generally larger than but perhaps not inconsistent with those observed. The observation of seismic anomalies, for example, may be possible only toward the later stages of what is referred to here as the precursory period. This interpretation is consistent with the earlier remark concerning the prediction that rapid decrease of pore fluid pressure occurs relatively near final instability.

The values of the material parameters which were used in the calculations were chosen to be consistent with the existing laboratory and field data. In order to assess the magnitude of the effect that variations in the more uncertain of these parameters might have, we have performed a few additional calculations for alternative values of the dilatancy factor β, the peak strain parameter λ and the fluid bulk modulus K_f. The calculations were carried out for $\dot{\gamma}_\infty t_D = 1.4 \times 10^{-5}$ which corresponds to $\dot{\tau}_\infty$ = 1bar/year, a = 3km, and c = 0.1 m^2/sec. The results are summarized in Table 2.

The values of β which were quoted earlier were estimated from laboratory data for tests on intact rock. Because laboratory tests have indicated that dilatancy may diminish somewhat with cyclic loading (e.g., Scholz and Kranz, 1974; Zoback and Byerlee, 1975) and with increased confining stress (Brace et al., 1966), smaller values of β may be more representative of in situ conditions. Table 2 shows the calculated dimensionless precursor time for values of β equal to one-half and one-quarter of the value used for other computations (β = 0.3). Reduction of β by half apparently reduces t_{prec}/t_D by slightly more than half. Note, however, that even

for $\beta = 0.075$, $\theta_{prec} = 0.16$ corresponds to approximately 170 days in real time for $a = 3km$ and $c = 0.1m^2/sec$. This observation emphasizes that even a very small amount of dilatancy may have quite substantial effects of the kind described here.

The value of the peak strain parameter λ is one of the most uncertain in the analysis since laboratory investigations of postpeak behavior have been relatively sparse. Fortunately, the calculation does not appear to be especially sensitive to the value of λ . The entries in Table 2 suggest that decreasing λ by half reduces θ_{prec} by about 30%.

In the calculations for Table 1 and Fig. 7, the bulk modulus of the fluid K_f was assumed to have a value appropriate for liquid water, i.e., $K_f \simeq 22kb$. We remarked earlier, however, that high temperatures, low pressures or the presence of dissolved gases may reduce K_f . Table 2 shows the calculated precursor times for $K_f = 11$ kb and 2.2 kb. These correspond to values of K_f'/G (K_f' is defined following (26)) of 1.2 and 0.5 by comparison with $K_f'/G = 1.45$ for $K_f = 22$ kb. We remark that an equivalent reduction of K_f' to $0.5G$ can be accomplished by an increase of porosity from $v = 0.01$ to $v = 0.1$. The reduction of K_f by half reduces θ_{prec} by only about 6% and for $K_f = 2.2$ kb , the predicted precursor time is decreased by slightly less than half.

In all of the calculations, it has been assumed that the hardening modulus was the only material parameter which varied substantially with the deformation. A more detailed calculation should consider the variation of other material parameters. In particular, appreciable alteration of the dilantancy factor β or of the bulk modulus of the pore fluid K_f will probably be important. It seems evident that there must exist a limit to

the amount of dilantancy which a rock may undergo. Thus, it is likely that β may decrease at large strains and that the stabilizing effect of dilatant hardening may be limited by the attainment of a least dense state corresponding to no further dilatancy. This is consistent with the measurements by Crouch (1970) which indicated a decrease in the rate of dilatancy relative to the axial strain after some amount of deformation past peak stress although dilatancy did continue throughout the postpeak regime.

A second limiting effect, which was mentioned earlier, is the increase in fluid compressibility. For the calculations performed here, the decrease in the pore fluid pressure from the ambient level which was assumed to exist when the inclusion material was at the elastic limit did not approach 100 bars until very near instability. The variations shown in Figs. 8c and 9c are typical though the decreases became larger more rapidly for larger values of $\dot{\gamma}_\infty t_D$. Because the lithostatic pressure at 5km. depth is about 500 bars the calculations suggest that substantial decreases of K_f may be difficult to achieve by pore pressure decreases alone, at least until very near instability. It should be emphasized that "very near instability" in terms of the dimensionless precursor time may be on the order of days or even weeks in real time, as can be seen from Table 1. Furthermore, elevated temperatures and exsoultion of trapped gas may contribute significantly to reducing K_f .

The calculations also apply rigorously only for spherical inclusions. A preliminary analysis by Rudnicki (1977b) for a flat elliptical inclusion in plane strain indicates that the time scale of stabilizing effects due to dilatant hardening is very much shorter for narrow zones. In particular,

Rudnicki (1977b) derived an equation whose asymptotic long-time characteristics were similar to those of (30), except that the appropriate length which entered the diffusion time L^2/c was the short axis of the ellipse. This suggests that for an axi-symmetric ellipsoid of aspect ratio $a/b = 20$, the diffusion time is 1/400 of that for a sphere of radius a. A more complete analysis is needed to determine the extent to which dilatant hardening effects are influenced by geometry.

STABILIZATION BY TIME-DEPENDENT STIFFNESS OF SURROUNDINGS

For the analysis of this mechanism we neglect any suctions induced in the pore fluid within the inclusion and assume that the inclusion response to shearing can be represented by a plot of τ_{inc} vs. γ_{inc} as in Figs. 2 and 4. As remarked earlier, this plot may be thought of as representing in-situ conditions and hence to include effects of an increasing mean stress σ induced by the constraint of the surroundings against inelastic dilation of the inclusion material. Thus, letting

$$\tau_{inc} = F[\gamma_{inc}] \tag{32}$$

describe the stress-strain relation of the inclusion material, the mathematical problem is to solve eq. (7) subject to this relation for a given history $\tau_\infty(t)$. Since we assume in deriving (7) that there is equilibrium under fully drained conditions for $t \leq o$, the state at $t = o$ must satisfy eq. (2), namely

$$\gamma_{inc}(o) - \gamma_\infty(o) = \frac{\xi}{G}[\tau_\infty(o) - \tau_{inc}(o)] \quad ,$$

and this may be used to simplify the right side of (7). Hence, by using (32), and $\tau_\infty = G\gamma_\infty$, (7) becomes the non-linear integral equation

$$\gamma_{inc}(t) = \gamma_{inc}(o) + \frac{1}{G}[\tau_\infty(t) - \tau_\infty(o)] + \frac{1}{G}\int_o^t \left\{\xi_u + (\xi-\xi_u) f\left[\frac{c(t-t')}{a^2}\right]\right\}\left\{\dot{\tau}_\infty(t) - F'[\gamma_{inc}(t')]\dot{\gamma}_{inc}(t')\right\}dt' \tag{33}$$

where $F'[\gamma_{inc}] = dF[\gamma_{inc}]/d\gamma_{inc}$. This equation applies for $t \geq o$;

$\tau_\infty(t)$ is regarded as given and $\gamma_{inc}(t)$ is to be determined.

This equation is rigorous for a spherical inclusion with ξ and ξ_u calculated from (3) using the drained and undrained Poisson ratios, ν and ν_u respectively, for ν_e. As an approximation, however, we shall also assume that (33) applies to other types of inhomogeneous zones, specifically to the flattened axi-symmetric ellipsoidal zone of semi-major axis a, for which ξ and ξ_u are calculated from (4). For the function $F[\gamma_{inc}]$ of (32) we use the parabolic stress-strain relation in Fig. 6. We assume that the inclusion strain is at the peak of the curve when $t = o$, $\gamma_{inc}(o) = \gamma_p$, and the remote tectonic stress is increased at a constant rate $\dot{\tau}_\infty$ for all subsequent time.

Eq. (33) must be solved numerically and the details are explained in Appendix II. Essentially, we find that the solution can be put in the dimensionless form

$$[\gamma_{inc}(t) - \gamma_p]/[\gamma_{D'} - \gamma_p] = g(\theta; R, \xi/\xi_u, \xi) \quad . \tag{34}$$

Here $\gamma_{D'}$ is the inclusion strain at the point corresponding to D' in Fig. 4 (i.e., the undrained instability point, at which dynamic instability occurs), $\theta = ct/a^2 = t/t_D$ as before, and R is dimensionless measure of the tectonic stressing rate, namely

$$R = \xi(1+\xi)(a^2/c)(\dot{\tau}_\infty/G\lambda) \quad . \tag{35}$$

Apart from the factor $\xi(1+\xi)$, which depends on the shape of the inclusion, R can be interpreted as the ratio of the characteristic diffusion time $t_D(=a^2/c)$ to the time for the remote tectonic strain to increase by amount λ, where λ is defined in Fig. 6.

Some specific plots of the results represented by (34) will be shown

subsequently. First we discuss the choice of parameters. We have examined the cases $\xi = 1$ and $\xi = 10$. The first corresponds to a near-spherical inhomogeneous zone (e.g., (3) yields $\xi = 1$ when a drained Poisson ratio $\nu = 0.2$ is assumed). The second corresponds to a flattened, slit-like zone; from (4), $\xi = 10$ is consistent with an axi-symmetric ellipsoid with aspect ratio $a/b \approx 18$. Remarkably, our numerical results indicate that for given values of R and ξ/ξ_u, the function g of (34) is very nearly independent of ξ. Results for the time θ to go from B' to D' in Fig. 4, for example, typically differ by 1% or less for $\xi = 1$ versus $\xi = 10$, except at the largest values of R that we considered ($R \approx 3$), at which the differences were still only on the order of 10%.

The ratio ξ/ξ_u in (34) can be interpreted via (2) as the ratio of the elastic unloading stiffness of the inclusion surroundings under undrained conditions to the same under drained conditions. From (3), (4) and (5) the ratio is given by

$$\frac{\xi}{\xi_u} = \frac{(4-5\nu)(7-5\nu_u)}{(4-5\nu_u)(7-5\nu)} \quad , \quad \frac{(1-\nu)(2-\nu_u)}{(1-\nu_u)(2-\nu)} \quad , \text{ and } \quad \frac{1-\nu}{1-\nu_u} \tag{36}$$

for the respective cases of spherical, narrow axi-symmetric ellipsoidal, and narrow elliptical inclusions. The ratio evidently depends on ν and ν_u and there is no direct source of in situ values known to us. Neither corresponds to the Poisson ratio inferred from seismic wave speed ratios, as noted by Rice and Cleary (1976) and O'Connell and Budiansky (1977). Further, in situ values will almost certainly be dominated by the presence of joints and fractures, and will thus differ significantly from values inferred for intact laboratory specimens. For example, Rice and Cleary (1976) summarize data on intact specimens and report values of ν and ν_u

of .27 and .30 for Charcoal Granite, .25 and .34 for Westerly Granite, .12 and .31 for Ruhr Sandstone, and .20 and .33 for Berea Sandstone. Large, partially opened joints or fractures reduce the drained volumetric stiffness of rock as much as a far larger volume fraction of equi-axed pores, yet have almost no effect on the undrained volumetric stiffness, assuming full saturation of the flat pore space by liquid water. The effect is to reduce ν and increase ν_u so that the presence of such joints could make the ν and ν_u values for a granite rock resemble much more closely the values for intact, porous sandstones than intact granites.

An alternate approach to the effects of joints and fractures is to use the theoretical estimates of elastic properties of cracked rocks by O'Connell and Budiansky (1974). What those authors refer to as the Poisson ratio $\bar{\nu}$ for "dry" conditions corresponds to our drained ratio ν, and in Table 3 we summarize their results for ν as a function of their crack density parameter $Nr^3 (=\varepsilon)$ for a rock having a Poisson ratio of 0.25 when crack-free. Here N is the number of penny-shaped cracks per unit volume having radius r. The undrained Poisson ratio ν_u can be calculated from the values of the "dry" shear modulus G reported by O'Connell and Budiansky (1974) and from the undrained bulk modulus K_u which, as we have noted, is essentially equal to the bulk modulus of the solid for flat, crack-like pore spaces. The resulting values of ν_u are also listed in Table 3. We note that for crack densities Nr^3 greater than approximately 0.1 the crack-interactions considered by O'Connell and Budiansky lead to results for ν and ν_u that differ substantially from estimates made on the basis of dilute concentration formulae for G and K (Rice, 1977a).

With the results in Table 3 we can calculate the undrained to drained stiffness ratios, ξ/ξ_u, of the surroundings from (36) and the results are shown in the table for different shapes of the inhomogeneous zone. Now, from observations on wave speed ratios V_p/V_s prior to the 1971 San Fernando earthquake, O'Connell and Budiansky (1974) suggest that their crack density parameter should have a range from 0.1 to 0.3 to fit one set of seismic data and from 0.2 to 0.4 to fit another. Pursuing Table 3, this suggests that stiffness ratios

$$\xi/\xi_u = 1.10 \quad \text{and} \quad 1.25$$

might be taken as representative, and we have used these two ratios in our numerical evaluations of the function in (34).

Results of Numerical Solutions and Precursory Predictions

As will be seen, a value of the tectonic loading rate parameter $R = 10^{-2}$ is representative of the middle of the range considered in some subsequent numerical evaluations, and we show in Figs 10 and 11 the solution of the integral equation (33) for stiffness ratios of 1.25 and 1.10, respectively. The points marked B' correspond to the drained instability point of Fig. 4 beyond which the system is self-driving, and D' to the undrained instability point at which dynamic failure occurs. The lower graphs in Figs. 10 and 11 have an enlarged time scale and show details in the neighborhood of the instability. For comparison, the dashed line curves show the corresponding progression to instability when pore fluid effects are neglected (this corresponds to the construction in Fig. 2b and the details are explained in Appendix II).

The effects of the pore fluid are evident in leading to the more prolonged period of accelerating strain before the instability and, of course, the effect is more pronounced for the larger stiffness ratio (Fig. 10) than

for the smaller (Fig. 11). The choice of a time interval which could be defined as a precursory period is, of course, somewhat arbitrary. As with our dilatant hardening analysis, we define

$$\theta_{prec} = ct_{prec}/a^2 = t_{prec}/t_D$$

as the time for traversal of the self-driven range between B' and D', and this time interval is shown as a function of R in Fig. 12. It is, evidently, a period over which local strains in the soon to be ruptured zone accelerate significantly over those accumulated in previous periods of comparable duration. Deformations at ground surface vary in proportion to the local strain within the inclusion (which can be regarded as an isolated, time-dependent dislocation) and will show a similar time history. Also, the rapid deformations could conceivably lead to discernible variations in seismic and/or transport properties within the failing zone.

Table 4 shows some specific predictions of precurory times in days based on the results in Fig. 12. We consider in this Table three values, 1, 3, and 5 km , for the radius a of the inclusion and two inclusion shapes (i.e., values of ξ) , a sphere ($\xi = 1$) and a slit-like ellipsoid ($\xi = 10$; 18:1 aspect ratio for axi-symmetric geometry). Also, two values of the fluid diffusivity are considered, $c = 1\ m^2/s$ and $c = 0.1\ m^2/s$, which appear to be consistent with various field (Anderson and Whitcomb, 1975) and well-head (Kovach et al., 1975) measurements. The first results shown for t_{prec} are based on a stiffness ratio of 1.10, and following these we show in square brackets the results for a ratio of 1.25. In preparing Table 4 it is necessary to associate a value of R , eq. (35), with each value of ξ , a , and c corresponding to an entry in the Table.

This is done by choosing a tectonic stress rate $\dot{\tau}_\infty$ = 1 bar/yr and shear modulus G = 200 kbar as previously. Evidently, from Fig. 12, a tenfold reduction in $\dot{\tau}_\infty$ would lengthen all precursory times shown by a factor of 5 or so. The parameter λ of (35) is defined with reference to the stress-strain curve of Fig. 6 and, as remarked earlier, if the curve in Fig. 6 is to have a peak strength of 1 kbar and to be linear up to 3/4 of peak strength, with continuous slope where the linear portion joins the parabola, then $\lambda = 0.0025$. We use this value in the Table.

The results in Table 4 reveal that while t_{prec} increases with inclusion size, it is again not directly proportional to the diffusion time a^2/c. Indeed, a tenfold decrease in diffusivity increases t_{prec} by a factor of 2 to 3, and a 5-fold increase in inclusion size (hence 25-fold in a^2) increases t_{prec} by a factor ranging from approximately 2 to 4.

There is a significant effect of the shape of the inhomogeneous zone. Assuming, as we have, the same τ_{inc} vs. γ_{inc} relation in each case, the spherical zone has precursory times that are 10 to 20 times longer than for the flat zone considered. There is also a significant dependence on the stiffness ratio; t_{prec} for $\xi/\xi_u = 1.25$ is 2 to 4 times longer than for $\xi/\xi_u = 1.10$.

Comparing Tables 1 and 4 it is seen that t_{prec} for spherical zones due to dilatant hardening is 2 to 5 times longer than the mean (for the two ξ/ξ_u ratios) of that due to time-dependent stiffness. Of course, if smaller values of the dilatancy factor β are considered, as in Table 2, the t_{prec} values become more nearly comparable. We have commented that t_{prec} due to dilatancy should be shorter for flat zones than for spherical

zones, but there are no comparisons to be made between the two mechanisms for flat zones.

It is interesting that the t_{prec} values predicted by solving the integral equation (33) are not very different from those estimated by Rice (1977a), on the basis of an approximation to the response function $f(\theta)$ of Fig. 5. Rice's procedure amounts to replacing $f(\theta)$ by a simple exponential form appropriate to a "standard linear model" with relaxation time $t_D/10$. This has the effect of converting (33) to a first order non-linear differential equation. The resulting approximation to $f(\theta)$ is not very close but predicted t_{prec} values agree within typically 25% or so with the more exact results in Table 4. The cost associated with solving (33) is much larger than for solutions based on the standard linear model, mainly because of the necessity of computing error functions in the complex plane to determine $f(\theta)$ in (33); Appendix II. Hence, in future calculations of the kind reported here, it may suffice to use the approximation based on the standard linear model.

CONCLUDING DISCUSSION

We have demonstrated that the coupling of pore fluid diffusion with deformation can delay the onset of rapid failure and give rise to precursory periods of quasi-static, but accelerating, deformation. Our calculations for the duration of this precursory period indicate that the effects can be significant for values of the parameters which are consistent with existing experimental and observational data. More specifically, Figs. 8 to 11 demonstrate that the precursory effects are much more dramatic for fluid-infiltrated solids than those for which the pore fluid is absent. Further, we have presented results for the precursory time so that its dependence on constitutive properties will be evident; see Fig. 7 and Tables 1 and 2 for the dilatant hardening mechanism, and Fig. 12 and Table 4 for that based on time-dependent elastic stiffness. Tables 1 and 4 are based on a range of material properties which we think may be appropriate to crustal rocks at depths on the order of 5 km. The consequences of other choices for the constitutive parameters and loading rate may be estimated from Table 2 and Figs. 7 and 12.

Because much of the discussion of precursory pore fluid effects has concerned their role in connection with the possibility that dilatancy may cause alterations in seismic wave speeds, we emphasized that the effects described here can be important even if conditions are not suitable for wave speed alterations. Indeed, the mechanism of time-dependent stiffness of the surroundings is not contingent upon the dilatant inelastic opening of pores or cracks. Moreover, our calculations for the dilatant hardening mechanism suggest that an amount of dilatancy much smaller than that necessary to affect wave speeds can have a considerable effect of the kind described here in stabilizing the rupture process, and that the magnitude of the pore pressure

decrease induced by dilatancy alone becomes large only in close proximity to instability. For example, reductions of seismic wave speeds are sometimes postulated to result from the opening of vapor filled or dry crack space in extensive regions of rock, and this process would be accompanied by large suctions in the pore fluid. But results such as those in Figs. 8 and 9 suggest that large suctions can result only very late in what we identify as the precursory period, and it is not clear as to whether they will typically be large enough to significantly alter wave speeds. Further, there is nothing in our analysis which suggests a source of the sometimes suggested return of seismic properties to normal levels just before rupture. These conclusions could, however, be a consequence of oversimplifications in our model. For example, the weakening zone is regarded as being spatially uniform up to dynamic instability, and no provision is made for a gradual concentration of deformations into a narrow fault zone.

Because of the uncertainty of precursor time estimates based on observations prior to earthquakes it is difficult to draw definitive conclusions by comparison of our predicted precursor times with observations. In addition, while our definition of precursor time is unambiguous within the context of our model, it is likely to be an upper limit for precursory time as detected by surface observations. This is because the strain within the inclusion continuously accelerates, Figs. 8 to 11, and it may only be toward the latter portion of the precursory period that effects are significant enough to be observed at ground surface. Of course, strain and tilt at the surface will have a similar time history to that of the inclusion strain, although the magnitude will be attenuated approximately in proportion to the inverse square of distance from the source. In addition, the accelerating strain near failure may also

have a discernible effect on transport properties in the source region, for example, on electrical resistivity due to the progressive microfracturing that accompanies inelastic straining. We note that transport properties seem likely to be more affected than seismic properties, at least to the extent that for significant alterations the latter requires suctions that are large enough to deplete pore spaces of liquid.

To establish minimum estimates of the precursory time interval, we adopt the more conservative estimates of the size of constitutive parameters expressive of coupling between the rock and its pore fluid. Then it would appear that precursory times would be in the neighborhood of 10 to 50 days for an approximately spherical weakening zone of 1 km radius. For example, if we modify the entries in Table 1 by use of the lowest dilatancy factor in Table 2 ($\beta = 0.075$, one-quarter of the value that we suggest as representative of laboratory triaxial tests on coherent rock), we obtain $t_{prec} \approx$ 11 to 47 days for a range of fluid diffusivities between 1 and 0.1 m^2/s . Similarly, from Table 4 and using the smaller ratio of undrained to drained elastic stiffness, $\xi/\xi_u = 1.10$, we find $t_{prec} \approx$ 14 to 37 days for the same range of diffusivities. If, for example, the last fifth of the precursory time interval is regarded as being "readily" detectable, in view of the rapid acceleration of strain near instability, this minimal estimate of the precursory period for a spherical zone of 1 km radius is in the range of approximately 2 to 10 days. We would suggest this sort of period as most appropriate in searching for precursory effects of the kind we describe.

Larger spherical zones lead to longer precursory periods, but the predicted effect does not scale directly with the characteristic diffusion time, $t_D = a^2/c$. Instead, a much less rapid variation occurs, more nearly

proportional to a to the power 1 or lower, although no single power can fit all the size dependencies documented in Tables 1 and 4.

An analysis of the dilatant hardening mechanism has not been possible for the non-spherically shaped zones, although an approximate analysis of the time-dependent stiffness mechanism has been possible for ellipsoidal zones. The results in Table 4 correspond to a flattened axi-symmetric ellipsoid of approximately 18 to 1 aspect ratio and in this case the precursory times are reduced by a factor of approximately 10. The results for other aspect ratios can be obtained from Fig. 12, using the appropriate value of ξ. When precursory deformations are concentrated from the start on a single fault plane, it is probably inappropriate to model the process by a narrow inclusion as in the present work. Instead, a model like that suggested in Fig. 1b seems more appropriate. This involves the spread of a slipping region along an existing fault under the driving force of the stress concentrations at the edges of the slipping region. We leave the fuller analysis of this case to subsequent work.

The model which we have employed here is of course idealized and we have made many simplifications. We have considered separately each of the two stabilizing mechanisms of the pore fluid, but it seems evident that the effects will be more pronounced when both act together. For a given τ_{inc} vs. γ_{inc} curve undrained runaway generally occurs at a larger strain for the dilatant hardening mechanism than for that of the time-dependent stiffness, although the strains at final instability are comparable for smaller values of the dilatancy factor β or the modulus K_f' (see eq. 28). Although we have considered only parabolic τ_{inc} vs. γ_{inc} curves, the computations do not appear

to be strongly sensitive to their shape as expressed by λ . Nevertheless, the form which we adopted has a continuously decreasing slope so that dynamic rupture is inevitable. There does, however, exist the possibility that the inclusion stress-strain relation reverses curvature before point D' is reached in Figs. 3 and 4 (e.g., see Fig. 2c). In such cases it is possible that the pore fluid effects allow the inclusion to undergo strain in the form of a wholly stable creep episode.

More generally, we have shown that if fissured rock masses are fluid infiltrated, then the coupling of the deformation with the diffusion of pore fluid will be important in processes preparatory to faulting. As we remarked earlier, consideration of pore fluid effects has been primarily limited to aspects which we would regard as secondary to their role as setting the time scale of the failure process. We would argue that pore fluid stabilization of faulting merits more attention, but at the same time we stress that these processes are sensitive to values of material parameters and transport properties and there is a need for better data in this regard.

APPENDIX I

Equations (22) and (30) were solved by treating them as two coupled ordinary differential equations for $p(t)$ and $\gamma_{inc}(t)$. Equation (22) is already in the appropriate form and needs no further comment. However, in order to write (30) in the form of an ordinary differential equation, we first write the convolution integral as a sum of integrals over each time step $\Delta\theta = \theta/n$

$$I = \int_0^{\theta} \frac{1}{\sqrt{\pi(\theta-\theta')}} \frac{dp}{d\theta}(\theta')d\theta'$$

$$= \sum_{k=1}^{n} \int_{(k-1)\Delta\theta}^{k\Delta\theta} \frac{1}{\sqrt{\pi(n\Delta\theta-\theta')}} \frac{dp}{d\theta}(\theta')d\theta'$$

where $\theta = t/t_D$. For each integral, $dp/d\theta$ was assumed to vary linearly between its values at the limits. The remaining integrations can be performed analytically and the result may be rearranged to yield

$$I \approx (4/3)(\Delta\theta/\pi)^{\frac{1}{2}}\left[\dot{p}_n + b_n\dot{p}_o + \sum_{k=1}^{n-1} K_k\dot{p}_{n-k}\right] \tag{I.1}$$

where $\dot{p}_k = \frac{dp}{d\theta}(k\Delta\theta)$,

$$b_n = (3/2)n^{1/2} - n^{3/2} + (n-1)^{3/2}$$

$$K_k = (k+1)^{3/2} - 2k^{3/2} + (k-1)^{3/2} \quad .$$

Using (I.1) in (30) yields

$$\dot{p}_n = -\frac{\beta K_f' G\dot{\gamma}_\infty + (h+h_d)A\left[p_n + \frac{4}{3}\left(\frac{\Delta\theta}{\pi}\right)^{1/2}\left[b_n\dot{p}_o + \sum_{k=1}^{n-1} K_k\dot{p}_{n-k}\right]\right]}{h - h_u + A(h-h_d)\frac{4}{3}\left(\frac{\Delta\theta}{\pi}\right)^{1/2}} \tag{I.2}$$

where

$$A = \frac{3K_f'}{\alpha(1+4G/3K_s)N} \quad .$$

Standard fourth order Runge-Kutta integration formulae were used to solve (I.2) and (22).

It is clear from the form of (I.2) and (27), and from the asymptotic analysis of Rudnicki (1977b) that $\dot{\gamma}_{inc}$ and $\dot{p}$ become unbounded as the point D' in Fig. 3 is approached. Consequently, the numerical calculation must be truncated before $h = h_u$. Nevertheless, the contribution to the precursor time which was omitted by the truncation was negligible within the accuracy needed and no special effort was made to resolve the final instability. Typical examples are shown in Fig. 8 and Fig. 9 where, although the calculation was truncated well before the strain had reached its value at final runaway, the inclusion obviously would have reached this strain in an extremely short amount of additional time, at least as measured by θ.

APPENDIX II

To solve the integral equation (33) subject to the stress-strain relation (31), we introduce the notations

$$g = (\gamma_{inc} - \gamma_p)/(\gamma_{D'} - \gamma_p)$$

as in (34), noting that $\gamma_{D'} - \gamma_p = \lambda/\xi_u$, and $\alpha = \xi/\xi_u$. Then (33) becomes

$$g(\theta) - g^2(\theta)/2 = [(\alpha+\xi)/(1+\xi)]R\theta/\alpha^2 + (\alpha-1)\int_0^\theta f(\theta-\theta')[\xi R/\alpha^2(1+\xi) + g(\theta')dg(\theta')/d\theta']d\theta' \quad . \tag{II.1}$$

From Rice et al. (1977),

$$f(\theta) = 1+2[3\eta(4-3\eta)]^{-1/2}\,\mathrm{Im}[\beta \exp(\beta^2\theta)\,\mathrm{erfc}(\beta\theta^{1/2})] \quad , \tag{II.2}$$

where $2\beta = 3\eta - i[3\eta(4-3\eta)]^{1/2}$ and Im means "imaginary part of." The function is shown in Fig. 5 and is not very sensitive to η, over the allowable range; we use $\eta = 0.9$ in our calculations.

The time interval in (II.1) is divided into a series of steps. A small initial step is assumed and each new step is chosen so that the increment in g, predicted from the last calculated value of $dg/d\theta$, is smaller than some limiting size chosen to guarantee numerical convergence (the necessary step sizes decrease in approximate proportion to the size of R). The right side of (II.1) is evaluated by assuming that $dg/d\theta$ is constant $(\Delta g/\Delta\theta)$ in each step and by treating $f(\theta-\theta')$ as constant in each step, using the mid-step value of θ'. Consequently, at the end of each step (II.1) becomes a quadratic equation which is solved for $g(\theta)$ so that calculations for a new step can begin. The results for $R = 10^{-2}$ are shown as the solid curves for g in Figs. 10 and 11, and the resulting

precursory times are shown in Fig. 12 for a wide range of R values. As remarked, for given values of R and α there are negligible differences between results for $\xi = 1$ and 10.

The dashed curves of Figs. 10 and 11 correspond to neglecting pore fluid effects on stiffness, and are equivalent to replacing f by its long-time value of unity in (II.1). Thus by elementary calculations one shows that the equation of the dashed curves is

$$\alpha g \equiv \xi(\gamma_{inc} - \gamma_p)/\lambda = 1 - (1-2R\theta)^{1/2} \quad ,$$

and the inclusion strain rate becomes unbounded at $\theta = 1/2R$, corresponding to the attainment of point B' in Figs. 2 and 4. This shows that R may additionally be interpreted as half the ratio of the diffusion time t_D to the time for remote loadings to bring the inclusion strain from that at peak strength to that at instability.

ACKNOWLEDGEMENT

This research was supported by the USGS Earthquake Hazards Reduction Program and the NSF Geophysics Program, through grants to Brown University and the California Institute of Technology. We are grateful to V.C. Li for programming the numerical calculations outlined in Appendix II and leading to Figs. 10-12.

REFERENCES

Abramowitz, M., and I. A. Stegun (ed.), Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series 55, Washington, 1964.

Anderson, D. L., and J. H. Whitcomb, Time dependent seismology, J. Geophys. Res., 80, 1497-1503, 1975.

Anderson, O. L., and P. C. Grew, Stress corrosion theory of crack propagation with applications to geophysics, Rev. Geophys. Space Phys., 15, 77-104, 1977.

Biot, M. A., General theory of three dimensional consolidation, J. Appl. Phys., 12, 155-164, 1941.

Biot, M. A., Nonlinear and semilinear rheology of porous solids, J. Geophys. Res., 78, 4924-4937, 1973.

Booker, J. R., Time dependent strain following faulting of a porous medium, J. Geophys. Res., 79, 2037-2044, 1974.

Brace, W. F., and R. J. Martin, A test of the law of effective stress for crystalline rocks of low porosity, Int. J. Rock Mech. Min. Sci., 5, 415-436, 1968.

Budiansky, B., and R. J. O'Connell, Elastic moduli of cracked solids, Int. J. Solids Structures, 12, 81-97, 1976.

Carrier, G. F., M. Krook, and C. E. Pearson, Functions of a Complex Variable, McGraw-Hill, New York, 1966.

Cleary, M. P., Heterogeneity and porosity effects on the response of geomaterials, in Advances in Civil Engineering through Engineering Mechanics (Proc. of the Second ASCE Engrg. Mech. Specialty Conf., N. C. State Univ., Raleigh, N. C., 1977), published by ASCE, New York, 360-363, 1977.

Crouch, S. L., Experimental determination of volumetric strains in failed rock, Int. J. Rock Mech. Min. Sci., 7, 589-603, 1970.

Dieterich, J. H., Time-dependent friction in rocks, J. Geophys. Res., 77, 3690-3697, 1972.

Dieterich, J. H., Time-dependent friction and mechanics of stick-slip, Pure and Appl. Geophys., in press, 1977.

Eshelby, J. D., The determination of the elastic field of an ellipsoidal inclusion and related problems, Proc. Roy. Soc. London, Ser. A 241, 376-396, 1957.

Frank, F. C., On dilatancy in relation to seismic sources, Rev. Geophys. Space Phys., 3, 485-503, 1965.

Jaeger, J. C., and N. G. W. Cook, Fundamentals of Rock Mechanics, Methuen, London, 2nd Edition, 1976.

King, C-Y, R. D. Nason and D. Tocher, Kinematics of fault creep, Phil. Trans. Roy. Soc. London, A274, 355-360, 1973.

Kovach, R. L., A. Nur, R. L. Wesson, and R. Robinson, Water-level fluctuations and earthquakes on the San Andreas fault zone, Geology, 3, 437-440, 1975.

Martin, R. J., Time-dependent crack growth in quartz and its application to creep in rocks, J. Geophys. Res., 77, 1406-1419, 1972.

Martin, R. J., and W. B. Durham, Mechanisms of crack growth in quartz, J. Geophys. Res., 80, 287-293, 1975.

Nur, A., Dilatancy, pore fluids, and premonitory variations of t_s/t_p travel times, Bull. Seismol. Soc. Am., 78, 1217-1222, 1972.

Nur, A., and J. D. Byerlee, An exact effective stress law for elastic deformation of rock with fluids, J. Geophys. Res., 76, 6414-6419, 1971.

O'Connell, R. J. and B. Budiansky, Seismic velocities in dry and saturated cracked solids, J. Geophys. Res., 79, 5412- , 1974.

O'Connell, R. J., and B. Budiansky, Viscoelastic properties of fluid saturated cracked solids, J. Geophys. Res., in press, 1977.

Prescott, W. H., and J. C. Savage, Strain accumulation on the San Andreas fault near Palmdale, California, J. Geophys. Res., 81, 4901-4908, 1976.

Reynolds, O., On the dilatancy of media composed of rigid particles in contact, with experimental illustrations, Phil. Mag., 1885.

Rice, J. R., The initiation and growth of shear bands, in Plasticity and Soil Mechanics, edited by A. C. Palmer, Cambridge University Engineering Dept., Cambridge, England, 263-274, 1973.

Rice, J. R., On the stability of dilatant hardening for saturated rock masses, J. Geophys. Res., 80, 1531-1536, 1975.

Rice, J. R., Theory of precursory processes in the inception of earthquake rupture, in Proc. of the Symp. on Phys. of Earthquake Sources (IASPEI General Assembly, Durham, England, 1977), special publication by the German Democratic Republic (East German) Academy of Sciences, 1977a.

Rice, J. R., Pore pressure effects in inelastic constitutive formulations for fissured rock masses, in Advances in Civil Engineering through Engineering Mechanics (Proc. of the Second ASCE Engr. Mech. Specialty Conf., N. C. State Univ., Raleigh, N. C., 1977), published by ASCE, New York, 360-363, 1977b.

Rice, J. R., and M. P. Cleary, Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents, Rev. Geophys. and Space Phys., 227-241, 1976.

Rice, J. R., and D. A. Simons, The stabilization of spreading shear faults by coupled deformation-diffusion effects in fluid-infiltrated porous materials, J. Geophys. Res., 81, 5322-5334, 1976.

Rice, J. R., J. W. Rudnicki, and D. A. Simons, Deformation of spherical cavities and inclusions in fluid-infiltrated elastic solids, Int. J. Solids Structures, accepted for publication, 1977.

Rudnicki, J. W., The inception of faulting in a rock mass with a weakened zone, J. Geophys. Res., 82, 844-854, 1977a.

Rudnicki, J. W., Localization of deformation, brittle rock failure, and a model for the inception of earth faulting, Ph. D. Thesis at Brown University, 1977b.

Rudnicki, J. W., and J. R. Rice, Conditions for the localization of deformation in pressure-sensitive dilatant materials, J. Mech. Phys. Solids, 23, 371-394, 1975.

Scholz, C. H., Static fatigue of quartz, J. Geophys. Res., 77, 2104-2114, 1972.

Scholz, C. H., and R. Kranz, Notes on dilatancy recovery, J. Geophys. Res., 79, 2132-2135, 1974.

Scholz, C. H., L. R. Sykes, and Y. P. Aggarwal, Earthquake prediction: a physical basis, Science, 181, 803-810, 1973.

Sirovich, L., Techniques of Asymptotic Analysis, Springer-Verlag, New York, 182-185, 1971.

Swolfs, H., Chemical effects of pore fluids on rock properties, in Underground Waste Management and Environmental Implications, edited by T. D. Cook, American Assoc. of Petroleum Geologists, Memoir 18, Tulsa, 224-234, 1972.

Whitcomb, J. H., J. D. Garmany, and D. L. Anderson, Earthquake prediction: variation of seismic velocities before the San Fernando earthquake, Science, 180, 632-635, 1973.

Zoback, M. D., and J. D. Byerlee, The effect of cyclic differential stress on dilatancy in Westerly granite under uniaxial and triaxial conditions, J. Geophys. Res., 80, 1526-1530, 1975.

	a = 1 km	a = 3 km	a = 5 km
$c = 1m^2/sec$	(t_D = 11.6d) t_{prec} = 55.4d	(t_D = 104d) t_{prec} = 234d	(t_D = 289d) t_{prec} = 411d
$c = 0.1m^2/sec$	(t_D = 116d) t_{prec} = 238d	(t_D = 1042d) t_{prec} = 844d	(t_D = 2894d) t_{prec} = 1418d
Scholz et al. (1973)	$t_{prec} \approx$ 8d		$t_{prec} \approx$ 200d

Table 1. Predicted precursor times based on dilatant hardening of a spherical zone (t_{prec} is the time for the transition from B' to D' in Fig. 3). The precursor time for $a = 1$ km, $c = 1m^2/sec$ was estimated by extrapolation of the curve in Fig. 7. Data used were the following: tectonic stress rate $\dot{\tau}_\infty = 1$ bar/year; peak strain parameter (Fig. 6) $\lambda = 0.0025$; frictional coefficient $\mu = 0.6$; dilatancy factor $\beta = 0.3$; initial porosity $\nu = 0.01$; $N = c/k = 2.0G$; $K_s = 5.0G$; $K_f = 22$ kbar; $G = 200$ kbar; $\nu = 0.2$.

β	0.075	0.15	0.30
θ_{prec}	0.16	0.39	0.81
λ	0.00125	0.0025	0.0050
θ_{prec}	0.56	0.81	1.13
K_f	2.2 kb	11 kb	22 kb
θ_{prec}	0.44	0.76	0.81

Table 2. Effects of variations in the dilatancy factor β, the peak strain parameter λ, and the fluid bulk modulus K_f on θ_{prec}. All calculations were for $\dot{\gamma}_\infty t_D = 1.4 \times 10^{-5}$ and values of other parameters as in Table 1.

Crack density parameter Nr^3	Poisson ratios		(undrained stiffness)/(drained stiffness), ξ/ξ_u		
	drained ν	undrained ν_u	narrow elliptical cylinder	narrow axi-sym. ellipsoid	sphere
0	.25	.25	1.00	1.00	1.00
.1	.21	.28	1.11	1.06	1.07
.2	.17	.32	1.22	1.12	1.15
.3	.12	.36	1.37	1.20	1.25
.4	.08	.41	1.56	1.29	1.39

Table 3. Estimates of drained and undrained Poisson ratios for fluid-infiltrated solids with crack-like pore spaces, based on self-consistent calculations of O'Connell and Budiansky (1974). Also shown: ratio of elastic unloading stiffness of surroundings for undrained conditions to same for drained conditions, for various shapes of the inhomogeneous zone.

	a = 1 km	a = 3 km	a = 5 km
$c = 1\ m^2/s$	(t_D = 11.6d)	(t_D = 104d)	(t_D = 289d)
spherical (ξ=1)	$R = 1.27 \times 10^{-4}$ t_{prec} = 14d [29d]	$R = 1.14 \times 10^{-3}$ t_{prec} = 36d [92d]	$R = 3.16 \times 10^{-3}$ t_{prec} = 55d [150d]
flat (ξ=10)	$R = 6.96 \times 10^{-3}$ t_{prec} = 1.3d [3.9d]	$R = 6.27 \times 10^{-2}$ t_{prec} =2.6d [8.9d]	$R = 1.74 \times 10^{-1}$ t_{prec} = 3.5d [12d]
$c = 0.1\ m^2/s$	(t_D = 116d)	(t_D = 1042d)	(t_D = 2894d)
spherical (ξ=1)	$R = 1.27 \times 10^{-3}$ t_{prec} = 37d [96d]	$R = 1.14 \times 10^{-2}$ t_{prec} = 83d [261d]	$R = 3.16 \times 10^{-2}$ t_{prec} = 122d [399d]
flat (ξ=10)	$R = 6.96 \times 10^{-2}$ t_{prec} = 2.8d [9.5d]	$R = 6.27 \times 10^{-1}$ t_{prec} = 4.5d [18d]	R = 1.74 t_{prec} = 5.2d [22d]
Scholz et al. (1973)	$t_{prec} \approx$ 8d		$t_{prec} \approx$ 200d

Table 4. Predicted precursor times (for transition from B' to D' in Fig. 4) in days, based on time dependent elastic stiffness of surroundings. First values shown are for undrained/drained stiffness ratio, ξ/ξ_u, of 1.10; values which follow in brackets are for ratio of 1.25. Other data used: $\dot{\tau}_\infty$ = 1 bar/yr, G = 200 kbar, peak strain parameter λ (Fig. 6) = 0.0025. "Flat" zone corresponds to axi-symmetric ellipsoid with 18:1 aspect ratio.

LIST OF FIGURE CAPTIONS

Fig. 1 Types of failure models. (a) Deformation of "inclusion of different mechanical properties. "Runaway" instability based on strain softening of inclusion and elastic unloading stiffness. (b) Isolated region of slippage on pre-existing fault, spreading quasi-statically at small speed V.

Fig. 2 Deformation and instability of inhomogeneous zones. (a) Weakened zone. (b) Approach to instability. (c) Seismic gap zone.

Fig. 3 Stabilization of inhomogeneous zone by dilatant hardening; dynamic instability delayed to D'. (Rice, 1977a)

Fig. 4 Stabilization due to time-dependent elastic stiffness. The solid straight line represents the unloading stiffness for drained (d) conditions, the dashed line for undrained (u) conditions. Actual stiffness changes, d to u, will be smaller than shown. (Rice, 1977a)

Fig. 5 Function $f(\theta)$ arising in response of spherical cavity to a step shear loading, from Rice et al. (1977).

Fig. 6 Inclusion stress-strain relation.

Fig. 7 Predicted precursor time t_{prec}/t_D as a function of the non-dimensional far-field strain-rate $\dot{\gamma}_\infty t_D$. For values of material parameters, see text.

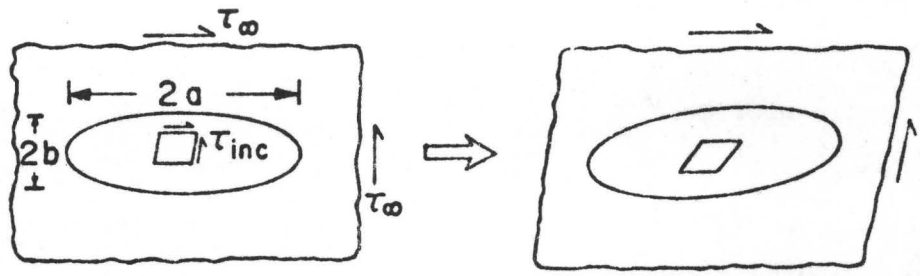
Fig. 8 History of post-peak straining of the inclusion (a,b) and the decrease of pore fluid pressure (c) in the inclusion. Computed for $\dot{\gamma}_\infty t_D = 1.4 \times 10^{-5}$. B' and D' are as in Fig. 3 and $\gamma_{D'}$ is the value of the strain at D'.

Fig. 9 Same as Fig. 8 except $\dot{\gamma}_\infty t_D = 1.4 \times 10^{-6}$.

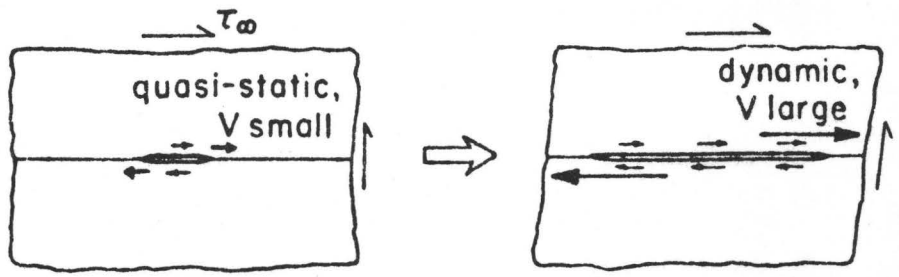
Fig. 10 Post-peak strain history for constant tectonic loading rate, based on undrained to drained stiffness ratio of 1.25. Dashed curve neglects pore fluid effects. Points B' and D' correspond to those in Fig. 4.

Fig. 11 Same as previous figure, but for undrained to drained stiffness ratio of 1.10.

Fig. 12 Precursory times as function of tectonic loading rate parameter R , based on time-dependent relaxation of elastic stiffness of material surrounding inclusion.

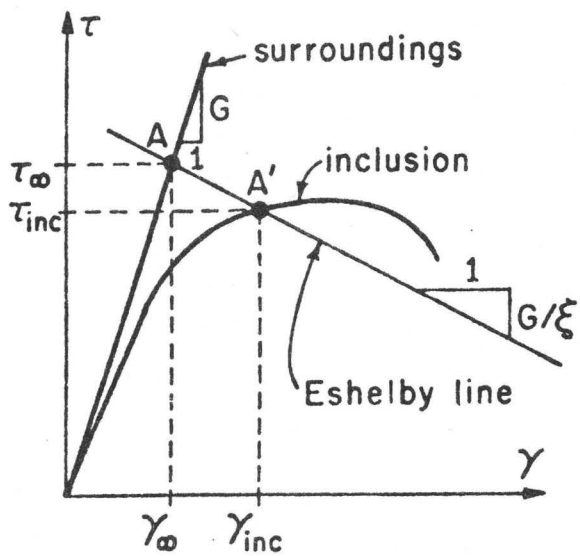


(a) Deformation of "inclusion" of different mechanical properties. "Runaway" instability based on strain softening of inclusion and elastic unloading stiffness.

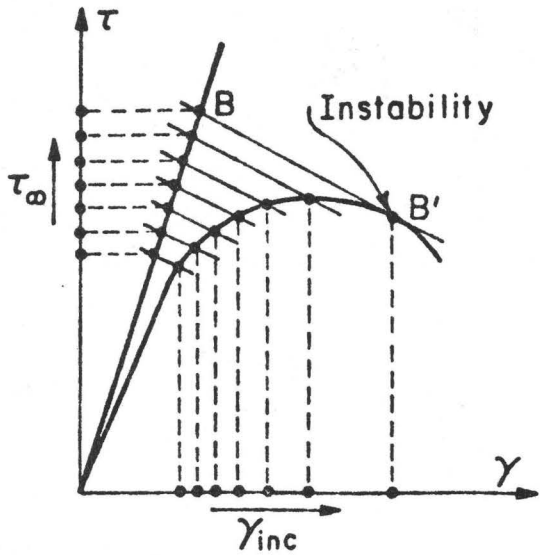


(b) Isolated region of slippage on pre-existing fault, spreading quasi-statically at small speed V.

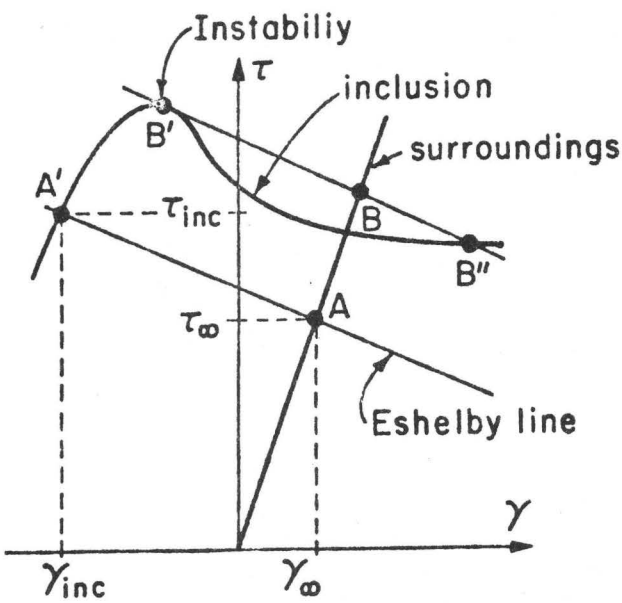
Fig. 1 Types of failure models.



(a) Weakened zone



(b) Approach to instability



(c) Seismic gap zone

Fig 2 Deformation and instability of inhomogeneous zones. (Rice, 1977a)

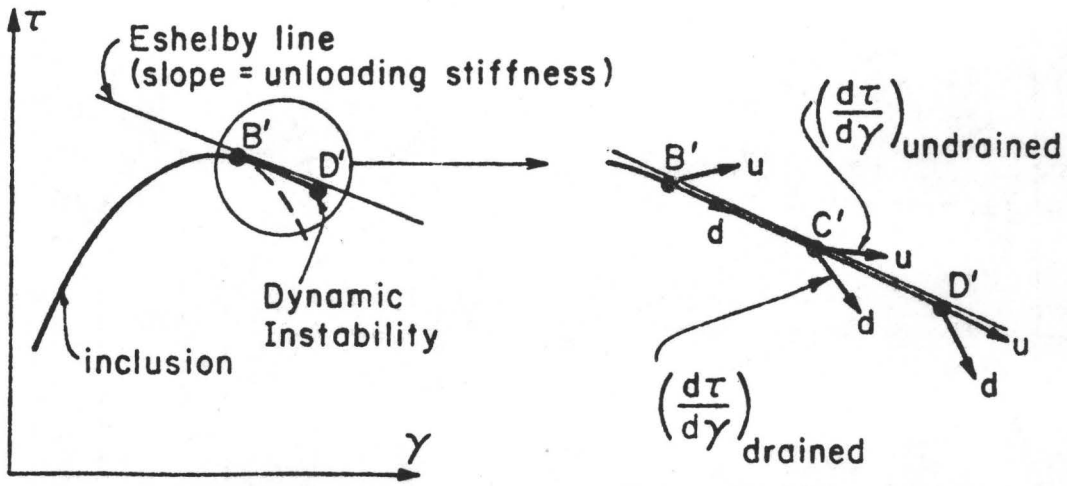


Fig. 3 Stabilization of inhomogeneous zone by dilatant hardening; dynamic instability delayed to D'. (Rice, 1977a)

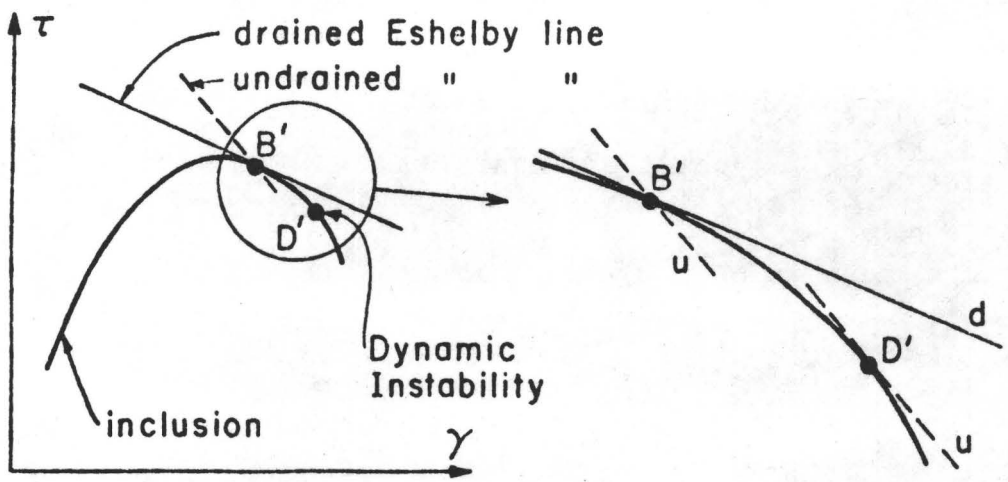


Fig. 4 Stabilization due to time-dependent elastic stiffness. The solid straight line represents the unloading stiffness for drained (d) conditions, the dashed line for undrained (u) conditions. Actual stiffness changes, d to u, will be smaller than shown. (Rice, 1977 a)

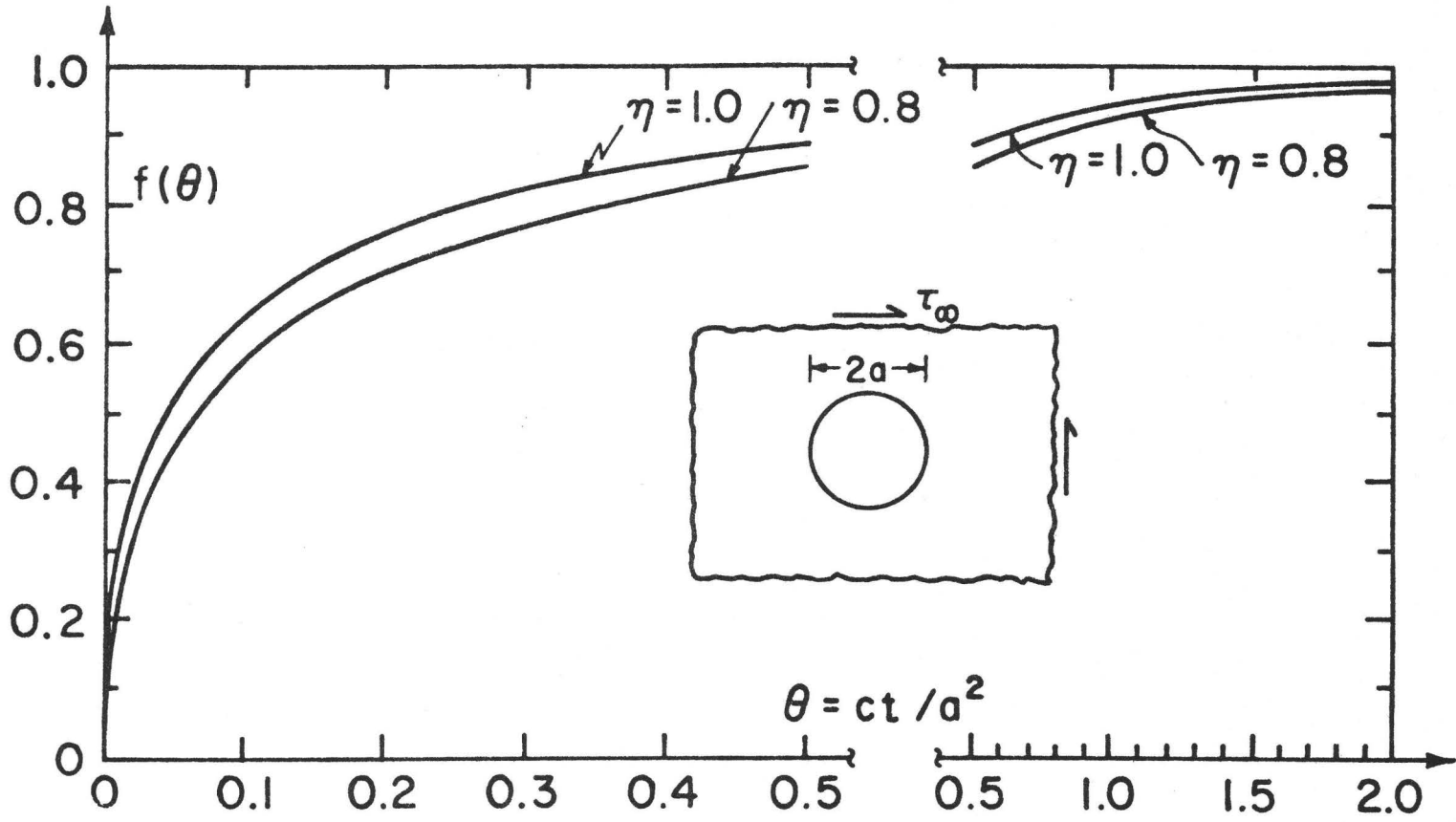


FIG. 5 Function $f(\theta)$ arising in response of sperical cavity to a step shear loading, Rice et al. (1977)

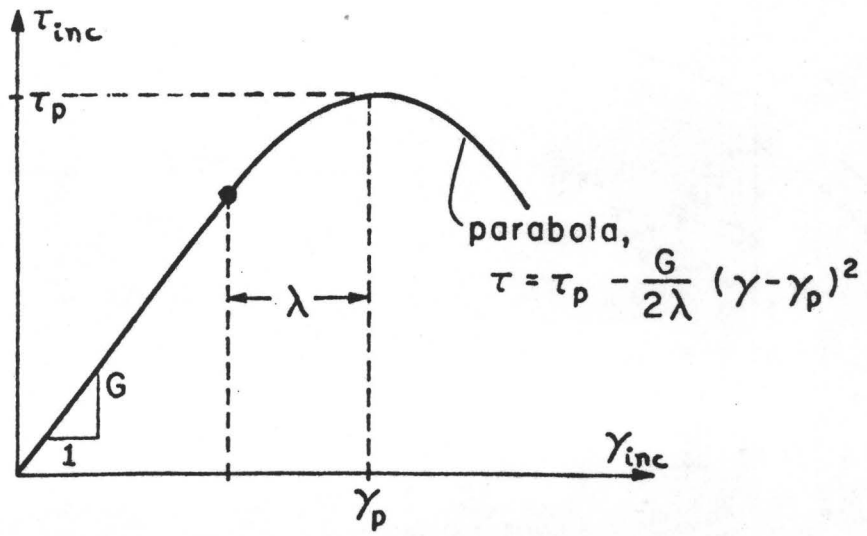


Fig. 6 Inclusion stress-strain relation

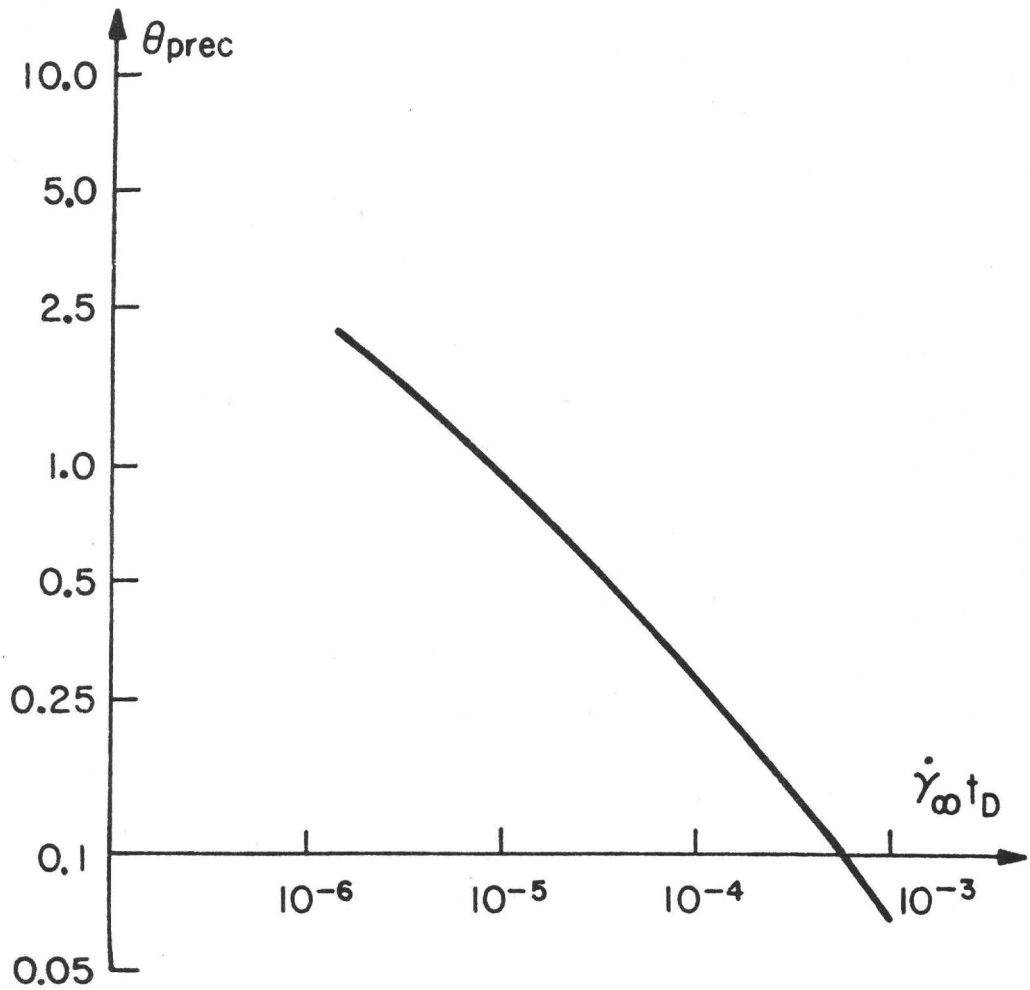


Fig. 7 Predicted precursor time t_{prec}/t_D as a function of the non-dimensional far-field strain-rate $\dot{\gamma}_\infty t_D$. For values of material parameters, see text.

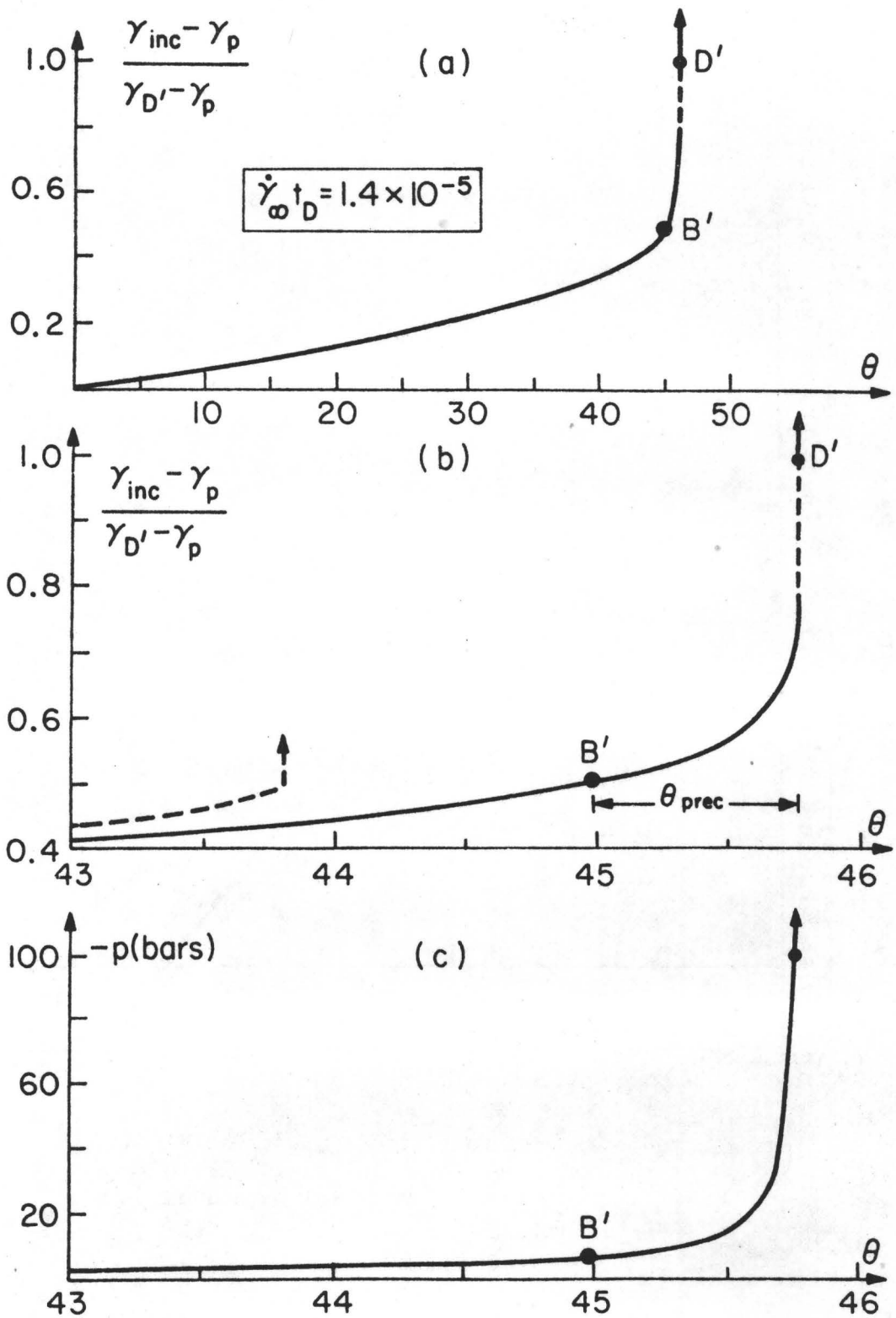


Fig. 8 History of post-peak inclusion strain (a,b) and of the decrease of inclusion pore fluid pressure (c). Computed for $\dot{\gamma}_\infty t_D = 1.4 \times 10^{-5}$. B' and D' are as in Fig. 3 and $\gamma_{D'}$ is the value of the strain at D'.

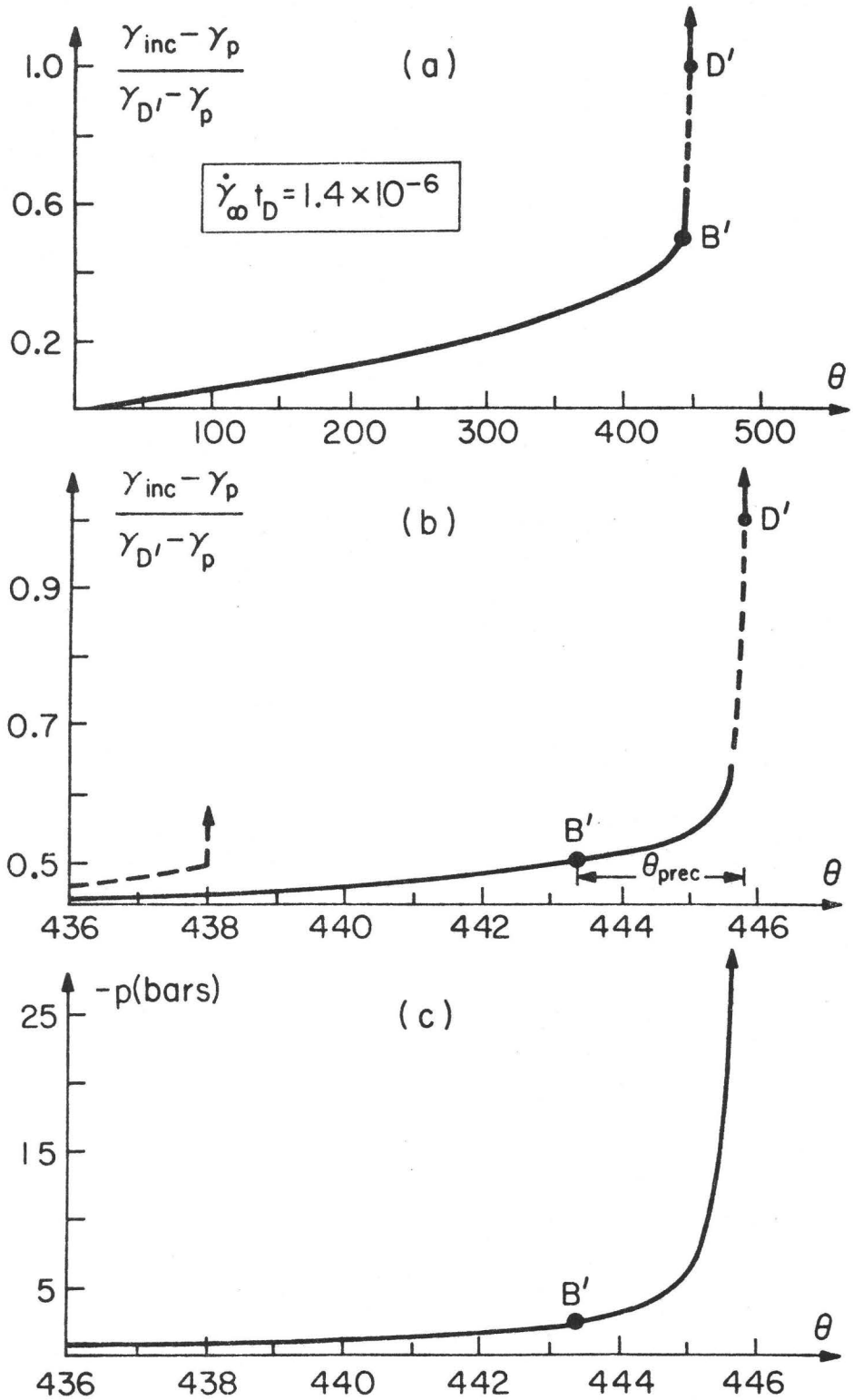


Fig. 9 Same as Fig. 8 except $\dot{\gamma}_\infty t_D = 1.4 \times 10^{-6}$

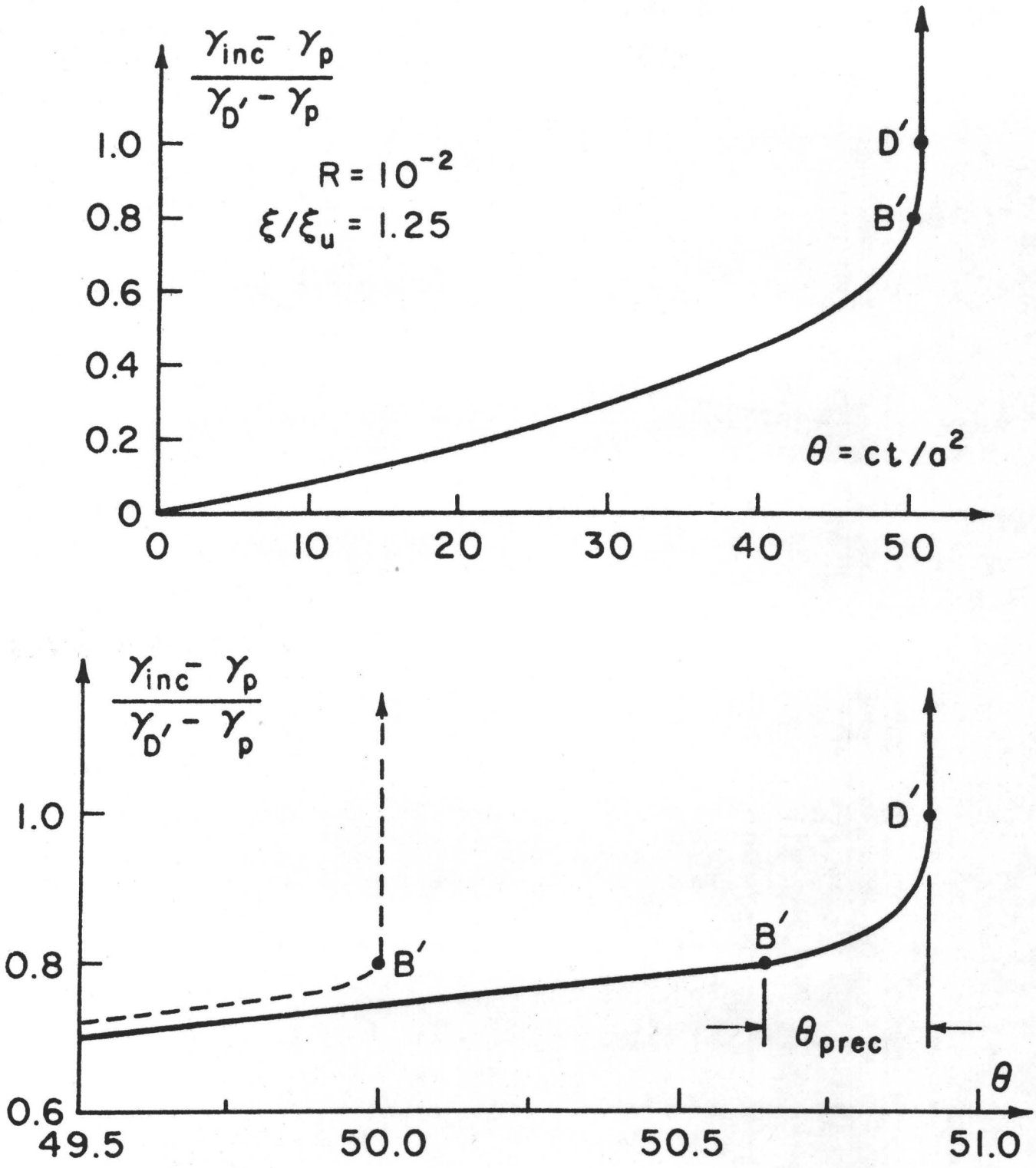


FIG. 10 Post-peak strain history for constant tectonic loading rate, based on undrained to drained stiffness ratio of 1.25. Dashed curve neglects pore fluid effects. Points B′ and D′ correspond to those in Fig. 4.

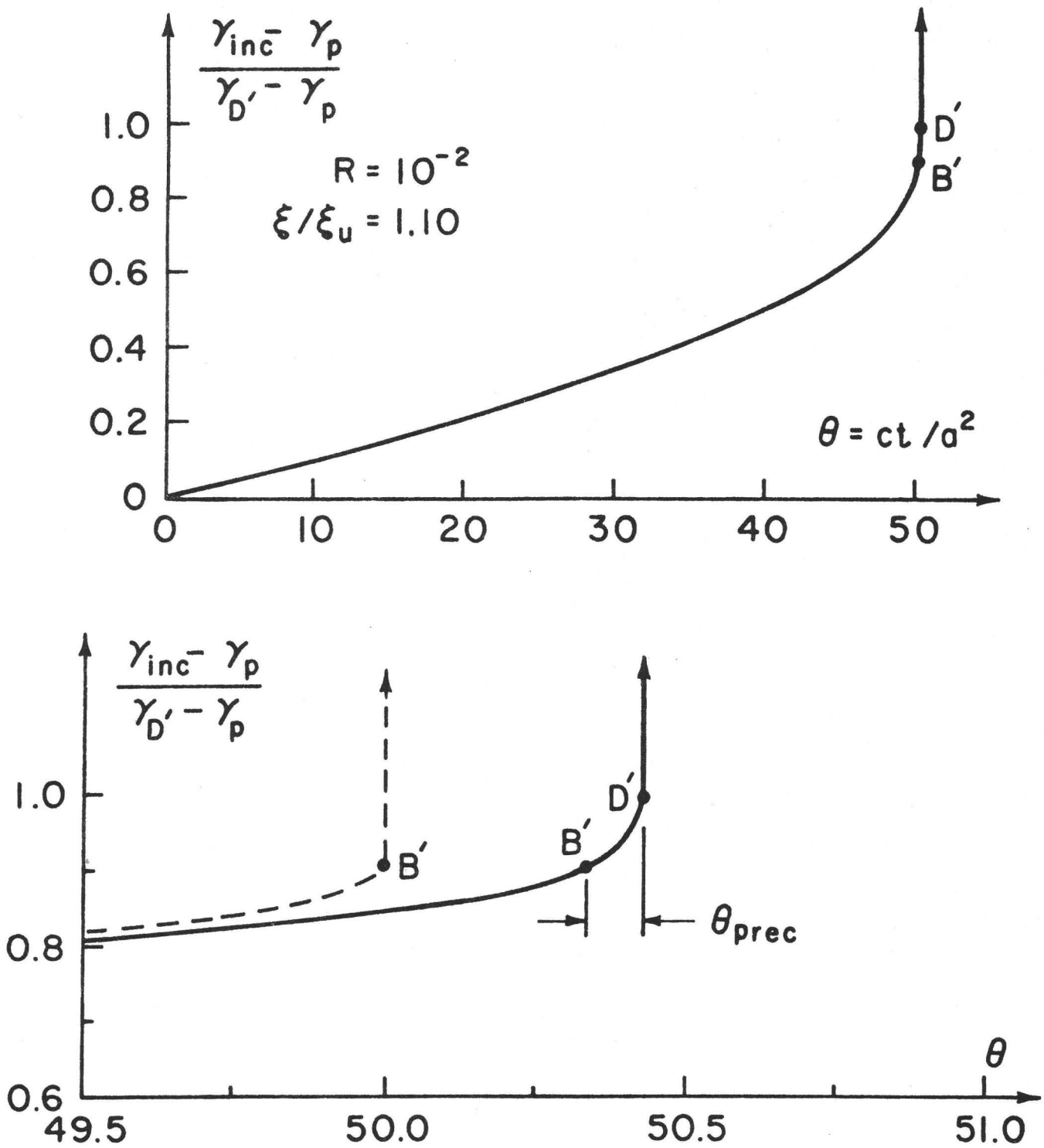


FIG. 11 Same as previous figure, but for undrained to drained stiffness ratio of 1.10.

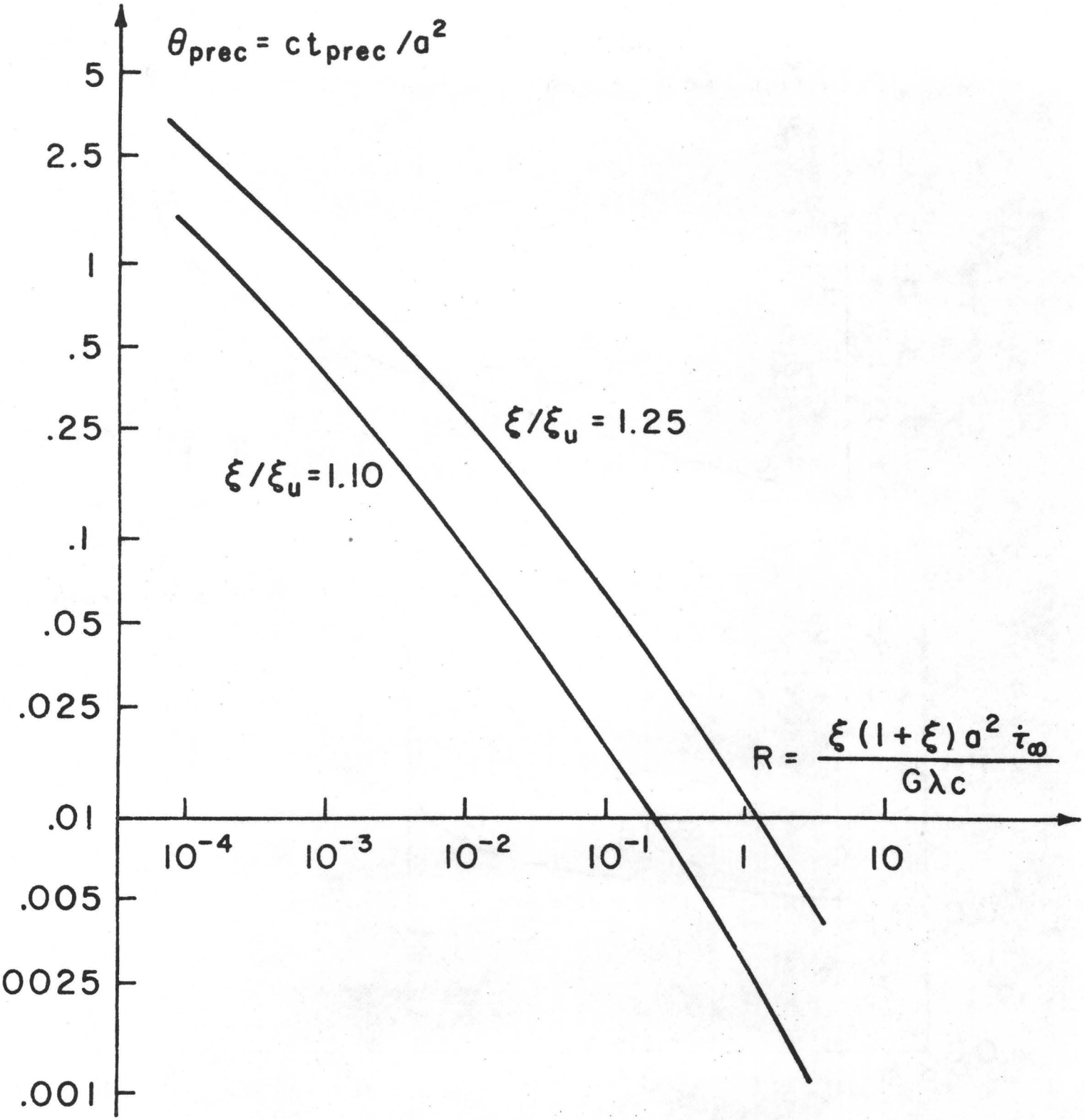


FIG. 12 Precursory times as function of tectonic loading rate parameter R, based on time-dependent relaxation of elastic stiffness of material surrounding inclusion.

REVIEW OF THEORIES FOR EARTHQUAKE INSTABILITIES

BY

William D. Stuart

U. S. Geological Survey

345 Midlefield Road

Menlo Park, California 94025

ABSTRACT

Theoretical mechanical models for earthquake instabilities are reviewed and compared on the basis of their postulated fault geometry, constitutive properties, and remote boundary conditions. All models invoke a fundamentally post-peak strain or strain rate weakening fault zone surrounded by elastic material and are specializations of Reid's elastic rebound hypothesis. Models other than the most elementary case of simple shear displacement in two dimensions differ principally according to whether the fault zone is approximated by: (1) an embedded homogeneous inclusion, (2) a propagating crack, or (3) a throughgoing crack with smooth peak stress variation along its length. All models whose formulation allows study of pre-instability deformation show growing deformation rates prior to instability, thus suggesting the existence of similar accelerating deformation prior to earthquakes. Instability may be delayed by pore fluid effects or physico-chemical healing, in which case another deformation time scale is introduced.

INTRODUCTION

The essential qualitative features of earthquake mechanics are well known. The observed suddenness and short duration of earthquakes indicate that they are mechanical instabilities because locally large strain and strain rate changes result from apparently small or slow changes of regional stress. In addition, direct observation of recent earthquake ruptures and exhumed ancient faults, plus first motion studies, imply that rocks fail during earthquakes by a shear mode and that the large deformations are restricted to thin tabular bodies, viz. fault zones. Occasionally earthquakes are attributed to volumetric sources, for example volcanic explosions or solid-solid phase transformations, or even to subterranean magma flowage as with harmonic tremors, but such earthquakes are not considered here.

An earthquake by definition involves strain rates large enough to generate stress waves in the nominally elastic crust and thus inertia must be considered in theoretical analyses. In contrast, if small or slow regional stress changes induce temporarily large but sub-inertial strain rates, a quasi-static formulation is sufficient. The so-called fault creep events and aseismic creep episodes fall into the latter category and their properties suggest that their and an earthquake's physical origin are fundamentally similar.

The elastic rebound hypothesis advocated by Reid (1910) seems to explain the main characteristics of crustal earthquakes and, with modification, deep focus earthquakes (< 700 km). Reid's proposal and later refinements, however, provide little insight into the precise mechanical conditions that might permit or prevent an earthquake instability, or for that matter, into deformation

preceding the instabilty. In general terms at least one thing is clear: in the context of a boundary value problem representation of Reid's hypothesis, the earthquake instability marks the transition from slow stress buildup to rapid propagation of an earthquake rupture surface. Being a boundary between sub-inertial and inertial deformation modes, an instability occupies a point in time and is a mathematical singularity in strictly quasi-static analyses. Notwithstanding the central role of the instability, only recently have theoretical fault models been formulated so as to reveal precise requirements for instability. Both quasi-static models for recurrent earthquakes and dynamic rupture propagation frequently impose rather than compute an instability by invoking a discontinuous fault stress drop according to some failure criterion.

The formal definition of instability is usually expressed as the derivative of fault strain or displacement with respect to remotely applied strain or displacement becoming unbounded. An example is $\partial u/\partial U \rightarrow \infty$ where u is fault slip and U is the remotely applied boundary condtion displacement. Substitution of stress or strain in the derivative may be more convenient in a specific problem. A corresponding definition employing potential energy derivatives may also be constructed. Finally, it is always possible to express the instability condition by comparing suitably defined stiffnesses of the fault zone and surrounding elastic medium. The stiffness criterion provides a particularly succinct graphical representation on a force-displacement plot where instability occurs when the weakening fault first becomes stiffer than the surroundings. In other words instability occurs when the fault weakens faster than the nearby elastic stress can decrease. A special case of instability is the sudden failure of brittle rock in a laboratory compression test with a compliant machine (Jaeger and Cook, 1976, p. 177).

Thus field observations, instability definition, and laboratory data suggest partitioning the earth into two coupled bodies of distinct geometry, constitutive properties, and strain history, namely: (1) a laminar fault zone containing initially brittle material experiencing finite and monotonic shear strains embedded in (2) an expansive elastic medium subjected to infinitesimal and reversible strains. Boundary conditions of force or displacement are applied to the exterior of the latter body. Modern instability theories retain Reid's basic postulates, but are distinguishable by and more explicit in their assumptions of geometry, constitutive properties, and remote forcing.

There is a practical reason for the theoretical study of instabilities. A successful theory may reveal that a restricted class of crustal deformations in space-time precedes earthquake instabilities. Thus a field monitoring program might detect rapid, even large, ground deformation, but the theory would determine if the deformation field were precursory to an instability or merely a reflection of a fault creep episode. The same theory might suggest a technological procedure for replacing a single forthcoming earthquake with an aseismic creep episode or several smaller earthquakes, for moving the earthquake elsewhere, or even for delaying the earthquake for posterity.

This paper reviews mathematical theoretical models for earthquake instabilities. Where possible, an attempt is made to identify pre-instability deformation that may lead to discovery of reliable earthquake precursors. The next section outlines the simplest strain softening (weakening) analytic model whose attributes and consequences underlie the more elaborate models. After exposition of the simple model, more complex strain or displacement softening models are considered. Next, stabilizing phenomena that can delay

instabilities are covered, followed by a section on strain rate softening models. The last section is a general discussion of earthquake instabilites with suggestions for future research.

THE SIMPLEST INSTABILITY MODEL

First consider the simple shear strain geometry of Figure 1a, where the constitutive response of the uniform fault zone f between h and -h is $\tau_f = \tau_f(\varepsilon_f)$, independent of time. τ_f is shear stress, $\varepsilon_f = u/2h$ is shear strain, and 2u is fault slip. Let the surrounding elastic material c bounded by h,H and -h,-H represent portions of the earth's crust forced by antisymmetric boundary conditions applied at $y = \pm H$. Assuming stress continuity at $y = \pm h$ (welded contact), smoothly increasing boundary displacement U at H will cause the shear stress to increase initially. U might, for example, approximate the inexorable movement of tectonic plates. If, as in Figure 1b, $\tau_f(\varepsilon_f)$ reaches a peak stress, shear stress will thereafter decline during strain softening of the fault zone. If the softening slope is steep enough, an instability will occur in the sense that an arbitrarily small increase of U will cause a large increase of u, i.e. $\partial u/\partial U \rightarrow \infty$. A quasi-static development will be followed here, but it is easily seen that if momentum terms were retained in the equations of motion acoustic waves would be generated and the instability might be characterized as "inertia limited"; if the fault constitutive law had a viscous component the instability might be "viscous limited". In any case, the instability, which is due to elastic rebound of c, resembles an earthquake in primitive respects. This elementary instability model, since it illuminates more sophisticated and realistic models discussed below, will be analyzed assuming that τ_f has a parabolic form.

Region c in Figure 1a is assumed to satisfy Hooke's law

$$\tau_c = 2\mu\varepsilon_c = \mu(U-u)/(H-h)$$

where μ is the rigidity modulus and ε_c is infinitesimal shear strain. Region f, the fault zone, surely is non-elastic and undergoes finite strain in situ, but neither laboratory experiments nor constitutive law theory establish its form or material coefficients well. If it has a power series representation in ε_f in the present geometry, the simplest expression allowing an instability is quadratic, namely

$$\tau_f = S\left[1 - (\varepsilon_f - \varepsilon_o)^2/\varepsilon_o^2\right]$$

where ε_o is the strain at the peak stress S (Figure 1b). Triaxial compression tests on sandstones and cohesionless aggregates by Griggs and Handin (1960), Rummel (1975), and Zoback and Byerlee (1976) show a roughly parabolic shape near the peak stress. Actual fault zones defined by large shear strain are probably thin compared to the space scale of smooth regional forcing, suggesting H>>h in the model. In fact for reasonable rock properties instability is impossible unless H>>h. As usual it is convenient to convert $\tau_f(\varepsilon_f)$ to a friction law

$$\tau_f = S\left[1 - (u-u_o)^2/u_o^2\right] \qquad (1)$$

thereby avoiding the ambiguity of strain and fault zone width. u_o is the fault slip at peak stress S. Equating $\tau_f = \tau_c$ gives the dependence of u on U

$$\frac{u}{u_o} = \tfrac{1}{2}\left\{\left[2+\mu'\right] - \left[(2+\mu')^2 + \frac{4\mu' U}{u_o}\right]^{\frac{1}{2}}\right\}$$
$$\mu' = \mu u_o/S(H-h) = \frac{\mu}{H-h} \cdot \frac{u_o}{S} \qquad (2)$$

where μ' can be interpreted as a dimensionless stiffness of the system in

response to U. The instability commences when $\partial u/\partial U \rightarrow \infty$, which performed on (2) gives the condition

$$(2+\mu')^2 - \frac{4\mu' U}{u_o} = 0 \qquad (3)$$

The value of fault displacement at the instability is

$$\frac{u_i}{u_o} = 1 + \mu'/2 \qquad (4)$$

Beyond u_i no quasi-static solution for u exists.

As shown at position 1 in Figure 1b, the instability occurs when the tangent to the friction law has a value $\partial\tau_f/\partial u = -S\mu'/u_o = -\mu/(H-h)$. Thus the point of instability is determined by a comparison of two slopes or stiffnesses. Observe in (2) that μ' is the ratio of elastic stiffness $\mu/(H-h)$ to a characteristic fault stiffness S/u_o. Walsh (1971), contemplating the relevance of laboratory stick slip events to earthquakes, pointed out that the earth is several orders of magnitude more compliant than laboratory apparatus. Since instability depends on stiffnesses of both the fault zone (gouge zone or friction surface) and surrounding crust (laboratory machine), both gouge and fault zone properties are needed to resolve the issue.

Equation (4) says that increasing μ' delays the instability. The earliest the instability can start is at $u_i = u_o$, i.e. at the peak stress when $\mu' = 0$; $\mu' = 0$ implies a perfectly flexible system. The latest realistic initiation of the instability with a quadratic law is at $u_i = 2u_o$ in which case $\mu' = 2$. Hence no instability is possible for $\mu' > 2$. To the extent that this model is relevant to the earth, an earthquake instability can

be prevented by increasing μ', for example by increasing the crustal rigidity or by modifying fault properties S and u_o such that the slip softening slope ($\sim S/u_o$) is gentler. In this and all other instability models considered here, stress appears scaled by a material constant. For this reason knowledge of the absolute stress level is not required to test instability models with field data.

Figure 2 shows how the fault slip u and stress $\tau = \tau_c = \tau_f$ vary with remote displacement U for both unstable deformation ($\mu' = 1.5$) and stable deformation ($\mu' = 3$). Before the instability and while the shear stress is decreasing from its maximum, u shows a rapid rate of increase. Both phenomena occur more gradually during the stable case too, which is a possible analog of a rapid aseismic fault slip episode. Enhanced slip or strain rates appear before instabilities in most models and seem to be a general consequence of a continuous strain softening law.

Similar results with a more realistic but more intricate friction law are not so amenable to concise analysis. A point worth making, however, is that unstable slip may be stabilized if the fault law slope becomes less steep, say at position 2 in Figure 1b. If this is so, the slip between positions 1 and 2 may be estimated with two quasi-static solutions for u, both for the same value of U.

A qualitative version of the above analysis was given by Stuart (1974). The principal conclusions of that paper, as here, are that average fault stress should decrease and fault slip rate increase before an earthquake instability.

An increasing stress boundary condition is a common alternate to increasing displacement, notably in dislocation theories. The instability

11

criterion in this instance becomes $\partial u/\partial \tau \rightarrow \infty$ where τ is now the stress boundary condition replacing U in Figure 1a . From (1), the condition is met only at the peak stress where $u = u_o$ if τ is monotonically increasing. The instabilty is equivalent to the $\mu' = 0$ situation with a displacement boundary condition, and consequently occurs earlier than all $\mu' > 0$ cases. A stress boundary condition in simple shear geometry does not provide much insight.

Strain localization is a phenomenon that invariably accompanies shear instability in continuous brittle media, though its physics is poorly understood. Localization is the concentration of deformation into thin bands of large and therefore non-elastic strain within a larger region of otherwise relatively homogeneous deformation. A complete theory for deformation in Figure 1a would involve localization and would likely show that the fault width 2h decreases in time. In practice, localization may be regarded as implicit in a softening friction or constitutive law. Shear bands in failed laboratory samples and narrow fault zones (meters or less wide) in the field appear to be examples of a localization process. Localization need not lead to an inertial instability, however, but the converse is probably false.

Even though it seems likely that shear failure in the crust is fundamentally associated with a localization mechanism, certain progress is possible without knowing details. In structural geology, for instance, numerous studies of the orientation and direction of faulting derived from elasticity theory (e.g. Hafner, 1951; Sanford, 1959) and plasticity theory (Odé, 1960) give solutions which agree well with both mapped faults and partially scaled simulations on clay and sand. The origins of localization in elastic-plastic media admitting a coupling between shear and normal strains are discussed below.

Figure 1a suggests four additional ways of making the simple shear model more realistic: generalize to three dimensions, possibly with a free surface; vary the shape of the fault zone; modify the fault constitutive law and let it vary with position; and alter boundary conditions. The models of the next section are distinguished on these grounds.

STRAIN SOFTENING INSTABILITY MODELS

Table 1 compares modern instability models whose mechanism is fundamentally strain or strain rate weakening. The models are best contrasted according to their assumption of fault zone geometry, or equivalently, in their postulated fault heterogeneity structure. The three variations to the simplest model just discussed (Figure 1a) are shown in Figure 3. The first two are also elements in the similar classification by Rice (1977).

In the first, Figure 3a, the fault zone is taken to be an elastic-plastic ellipsoidal inclusion embedded in an elastic full space subjected to a remotely applied shear stress τ_∞ or strain $\dot{\gamma}_\infty$. In the limit the ellipsoid may become a thin disk or a sphere. The ellipsoid may be regarded either as having greater or lesser "strength" than the surrounding medium. This work is due to Rudnicki (1977a) and Rice (1977).

The second and oldest fault geometry (Figure 3b), which draws on the early work of Griffith (1921) and subsequent development, is a zero-thickness crack whose ends propagate into the surrounding elastic medium under favorable friction strength and applied shear stress. A variety of external boundaries

can be treated, including full and half spaces. Fault constitutive properties are simulated by a stress-displacement weakening law: stress decreases from a peak value appropriate for unbroken pristine material at the crack tip to a residual frictional sliding value a short distance away. In models of this genre, the crack length actively experiencing failure during propagation is often assumed small compared with the total crack length (sharp tipped crack), and then fracture mechanics methods based on energy release rates are applicable. The fault is assumed to propagate in its own or another predetermined plane. Rigorous two-dimensional analyses (one-dimensional cracks) are due to Weertman (1967, 1978), Palmer and Rice (1973), and Cleary (1976).

A third type of fault zone geometry, Figure 3c, is a throughgoing crack over which a continuous position- and slip-dependent friction law is imposed. This strategy avoids consideration of crack tip propagation because fault slip never vanishes, though it may be arbitrarily small. Two-dimensional boundary value problems may be attacked with continuous distributions of dislocations (Stuart and Mavko, 1978), or with numerical discretization of the elastic continuum (Stuart, 1977). Faults where frictional strength varies only with position do not admit instability (Weertman, 1978), and the problem of a uniformly slip weakening fault is straightforward.

Theories invoking a throughgoing fault plane may themselves be further set apart according to the space scale of fault property variation relative to the fault slip required for complete softening. The last three workers postulate that fault slip is small compared to the scale of peak stress variation. Semi-quantitative one-dimensional models prompted by laboratory stick slip observations (e.g. Byerlee, 1970, Nur and Schultz, 1973) assume the softening

and peak stress are of the same scale inasmuch as friction depends on slip only. The same is true for phenomenological constructs employing rider masses moving across irregular frictional surfaces. Spacial property variations smaller than the slip softening scale are absorbed into a macroscopic continuum formulation.

No attempts yet exist where constitutive properties vary with position within a throughgoing, finite-width fault zone. Nor do any models consider a curving fault plane or multiple interacting faults.

Ellipsoidal Inclusion

Rudnicki (1977a) departs from the conventional fault plane geometry by considering the possibility of inertial instabilities within an ellipsoidal inclusion of different material embedded in an infinite, homogeneous, elastic space. Forcing is by a remotely applied shear strain γ_∞ or stress τ_∞. The inclusion, shown in Figure 3a, need not be elastic for Eshelby (1957) has shown that deformation within such an elastic ellipsoid is homogeneous. Thus any non-elastic constitutive law for which the stress and strain state satisfies displacements and tractions at the inclusion surface may be substituted. Special inclusion geometries recoverable from an ellipsoid are a flat disc, presumably representing a thin fault zone or plane, and a sphere. The physical interpretation of roughly equant inclusions is not so clear as that of a thin ellipsoid, but they could conceivably be appropriate models of the first throughgoing failure of a solidified igneous body emplaced in the crust. As Rice (1977) and Rice and Rudnicki (1978) mention, Rudnicki (1977a) regarded the inclusion to be a zone weakened compared to surrounding material, but the

theory is not particular and the inclusion may equally well be thought of as a body of greater failure strength. The latter situation corresponds roughly to a relatively locked "seismic gap" whose origins in the two dimensional case are traceable to Orowan (1960).

If the constitutive law for the inclusion has a peak stress followed by smoothly continuous strain softening, a stress level may be reached with increasing γ_∞ when the inclusion will weaken faster than the rate of stress decrease in the nearby surrounding material. The condition for this inertial instability, or "runaway instability" in Rudnicki's (1977a) terms, is $\partial\gamma_{inc}/\partial\gamma_\infty \rightarrow \infty$, where γ_{inc} denotes the inclusion strain. Rudnicki (1977a) gives an analytic expression for the instability in terms of geometric and constitutive parameters.

Rice (1977) provides a simplified graphical presentation of Rudnicki's (1977a) analysis by considering one component of the general Eshelby equation relating inclusion and remote stresses and strains

$$(\varepsilon_{ij})_{inc} - (\varepsilon_{ij})_\infty = Q_{ijkl}[(\tau_{kl})_\infty - (\tau_{kl})_{inc}] \qquad (5)$$

where ε is strain, τ is stress, and subscripts "inc" and "∞" refer to the inclusion and great distances from the inclusion respectively. The Q_{ijkl} are functions depending on the ellipsoid shape given in Rudnicki (1977a). The one component of (5) needed here is

$$\gamma_{inc} - \gamma_\infty = \frac{\xi}{\mu}(\tau_\infty - \tau_{inc}) \qquad (6)$$

$\xi = \frac{(1-\nu)a}{b}$ thin ellipsoidal disk in plane strain

$\xi = \frac{2(4-5\nu)}{7-5\nu}$ sphere

where ν is Poisson's ratio; 2a and 2b are major and minor axes of the ellipsoid. The two ξ are approximately 3a/4b and 1 for $\nu = .25$. Figures 4a (weak inclusion) and 4b (strong inclusion) show the non-elastic stress-strain curve for the inclusion and the elastic curve of slope G (rigidity modulus) for the surrounding region. As τ_∞ increases, say at a uniform rate, the intersection of the "Eshelby line" of slope $-\mu/\xi$ (see equation (6)) with the two constitutive laws determines the current stress and strain states. Points A and A' mark a possible quasi-static deformation state prior to the instability, but points B and B' mark the onset of the instability--at B,B' no quasi-static solution is possible for even an arbitrarily small increase of τ_∞. Positions B' appear where the Eshelby line is tangent to the inclusion constitutive law.

As with the simplest model outlined in the preceding section, the strong inclusion may distort to a new quasi-static configuration B" in Figure 4b if the inclusion law has a residual stress level (and if inclusion strains are infinitesimal). The Eshelby lines in Figure 4 indicate that the weak inclusion always has a stress lower than the surroundings; the situation is reversed for the strong inclusion (except at B"). By drawing lines parallel to A-A' it is also easy to see that the increment of inclusion strain for equal increments of τ_∞ increases as the instability approaches. As mentioned above, such accelerating strain (or slip) rates seem to be congenital properties of strain softening instability models.

Rudnicki (1977a,b), Rice (1977), and Rice and Rudnicki (1978) consider the influences on stability, including special stress states; inclusion geometries and constitutive laws; localization; and pore fluid stabilization. The last two effects will be discussed below.

Finite Fault

Two classes of methods have been used to study instability of semi-infinte and finite fault segments where the fault friction law may depend on position or slip. No closed solutions exist where the friction law depends on both position and slip, however.

The first method, which is also suitable for infinite faults, employs continuous distributions of straight edge or screw dislocations. For a full space and a two-dimensional vertical strike-slip fault, fault stress and slip derivative are related through a Hilbert transform. Weertman (1967) seems to have been the first to apply this approach to inertial instability on a slipping segment in a full space. Employing slip weakening friction laws having step or ramp stress drops, he showed that the remotely applied shear stress can be double valued for a certain range of plastic strain due to fault slip.

This explanation for discontinuous yielding in crystalline material may also be interpreted as a mechanism for inertial instability if one imagines that the slowly increasing applied stress is a similar regional stress. At a certain stress level, a plastic strain level is reached which is also an equilibrium solution for a lower applied stress level. Weertman shows that the instability vanishes if the rate of slip weakening is sufficiently gradual.

Linear fracture mechanics supplies the second general method for dealing with fault instability. The method, initially directed toward tensile crack propagation, dates back to Griffith (1921). Reviews are given by Irwin (1958) and Rice (1968). When the friction stress is constant in a shear crack, it is

well known that an $r^{-\frac{1}{2}}$ stress singularity exists at the crack tips where r is distance from the tip. By postulating that fault friction decreases from a value τ_p appropriate for intact material just ahead of the crack tip to a residual value τ_r after large slippage $\delta < \delta_r$ as in Figure 5, Palmer and Rice (1973) were able to remove the singularity. Furthermore, assuming also that the near-tip fault length over which $\tau_r < \tau < \tau_p$ (the breakdown region) is small compared to any other characteristic length such as the fault length or body dimension allows use of the J line intergral (Rice, 1968). The propagation criterion for such a sharp-tipped crack becomes

$$J_p - \tau_r \delta_p = \int (\tau - \tau_r) d\delta = \frac{1-\nu}{2G} k^2 \qquad (7)$$

Subscript P denotes that evaluation of the quantity is to include a fault point P outside the breakdown zone. Other symbols are Poisson's ratio ν, rigidity G, and stress intensity factor K. Palmer and Rice (1973) give the K formulae for various geometries including a semi-infinite crack in simple shear (Figure 5a) and a finite fault segment in an infinite two-dimensional space. The integral in (7) is interpreted as the crack driving force and is equal to the <u>supra</u> τ_r hachured area in Figure 5b. Any larger value of the left hand side of (7) will result in an inertial instability. As expected, the fault may propagate even though the applied stress τ_a is less than the peak fault strength τ_p. When the unfaulted medium is partially plastic, Palmer and Rice (1973) recommend using the unloading portion of the constitutive law. Equation (7) asserts that instability depends on the stress in excess of the residual level, implying that in the earth the absolute stress level is of no consequence to the model. (The absolute stress may,

nonetheless, enter implicitly into the fault slip law.) Andrews (1976) has derived a form of (7) without invoking the J integral, but founded on comparable energy arguments. The second equality in (7) holds only when the width of the breakdown zone is small compared to the overall fault length. The breakdown zone width is of order $G\bar{\delta}/(\tau_p-\tau_r)$ where $\bar{\delta}$ is the characteristic slip for breakdown given by (8).

$$\bar{\delta} = \frac{1}{\tau_p-\tau_r}\int(\tau-\tau_r)d\delta \qquad (8)$$

To maintain infinitesimal strains, $\bar{\delta}$ in (8) must be much smaller than the width of the breakdown region over which δ becomes comparable to $\bar{\delta}$.

Cleary (1976), whose work like that of Palmer and Rice (1973) is applied to slip surfaces in soil mechanics studies, especially slope failures, examined the propagation of two-dimensional faults with numerical methods. He combined a discretization of continuously distributed dislocations and fault stress-displacement relations like Figure 5b and of hyperbolic form to compute the onset of propagation. The friction law does not vary in space. Perturbations arising from proximity to a free surface were considered by allowing the distance of the fault from the free surface to be comparable to fault length.

Although fracture mechanics models for propagation of sharp-tipped faults or cracks have enjoyed success in technological and materials science applications, they would appear to benefit from elaboration when descibing earthquake instability. In particular, the models just outlined do not obviously yield phenomena which might be associated with reported earthquake precursors without the assumption of an additional mechanism for time dependence, such as pore fluid diffusion, rate-dependent constitutive properties, or geometry change. Such precursors might arise, however, from inhomogeneity of stress-displacement properties on the fault surface.

Throughgoing Fault

Using a finite element method, Stuart (1977) was able to determine with moderate precision the condition distinguishing stable and unstable slip on a two-dimensional strike-slip fault where the friction stress varied with both depth and slip. Referring to the geometry and variables depicted in Figure 6a, the friction law was postulated to be Gaussian in depth and fault slip according to

$$\tau_f' = \exp[-(2w')^2] \exp[-(\frac{y'-1}{b'})^2] \qquad (9)$$

where τ_f' is fault stress, $2w'$ is fault slip, x' is horizontal distance from the fault, y' is depth, and b' measures the broadness of the peak stress in depth. Primes denote dimensionless variables scaled according to $\tau' = \tau/S$, $w' = w/a$, $(x',y') = (x,y)/y_o$, $\mu' = \mu a/Sy_o$, and $b' = b/y_o$ (this definition of μ' differs from (2) by the substitution of y_o for $(H-h)$). S is the greatest peak stress (located at $y = y_o$), a measures the slip broadness of the friction law, and μ is the rigidity modulus. Friction law (9) is intended to approximate brittle failure of the fault zone and may be regarded as superimposed on an unknown stationary background stress. Thus once the boundaries of the rectangular region are chosen, here $x' = 0, 12.5$ and $y' = 0, 7.5$, μ' and b' sufficiently characterize a dimensionless solution. As with the simplest model in Figure 1 and equation (2), μ' is a stiffness; b' essentially determines the influence of the free surface. Although little laboratory data or theory exists for comparison with (9), it has the virtue of continuous first derivatives, slip softening to a residual

stress level, and an isolated peak stress maximum in depth. Increasing boundary displacement W' applied at $x' = 12.5$ and $y' = 7.5$ generates elastic stresses which are balanced by frictional resistance at the fault. Consequently, deformation time scales are determined by the time dependence of W'. The qualitative response can be seen by inspection. Since the fault becomes slip softening at increasingly shallow depths below $y' = 1$ as W' grows, stress becomes concentrated progressively closer to $y' = 1$. Finally, and if in addition $\mu' < 4$, a discontinuous displacement jump occurs which is considered to be an earthquake analog. The instability criterion is formally $\partial w'/\partial W' \rightarrow \infty$. No such jump occurs for $\mu' > 4$.

Figure 6b shows the displacement history for the case $b' = 1$, $\mu' = .25$. As anticipated, fault slip rates just below $y' = 1$ increase dramatically prior to the instability time $\zeta' = W'/12.5 = .48$, starting roughly at about $\zeta' = .3$. Figure 7a shows that these enhanced slip rates cause fairly distinctive increases in shear strain ε' rate starting at $\zeta' = .3$ at the free surface near the fault trace, $x' << 1$. (Oscillations at large ζ' are of numerical origin.) Because similar strain anomalies for stable cases $\mu' > 4$ are lacking, e.g. Figure 7b, such anomalies may point the way to a method for recognizing precursory anomalies in field data.

Except near the depth $y' = 1$, the fault has slipped beyond the peak stress position and resides at diminished levels before the instability. Consequently, the average fault stress also declines from the prior maximum by about one half. Furthermore, the average fault stress change during the instability jump is a small fraction of the stress change transpiring near the depth $y' = 1$.

A nearly identical problem has been solved numerically by Stuart and Mavko (1978) using a more elegant procedure involving continuous distributions of screw dislocations. Friction law (9) was used but the bottom plate boundary $y' = 8.$ was stress-free and remote displacement W' was applied at $x' = 20.$ Satisfying the nonlinear boundary condition at the fault and estimating τ_a', the applied stress at $x' = 20.$, was accomplished by an algorithm containing (9), a reciprocity relation between stresses and displacements at $x' = 0$ and $x' = 20.$ (Mavko, 1977),

$$\tau_a' = \mu'(W'-\bar{w}')/20$$

stress equilibrium at the fault

$$\tau_d' = -(\tau_a'-\tau_f')$$

and the Hilbert transform relating fault slip to fault stress (e.g. Weertman, 1965)

$$\frac{\partial w_d'}{\partial y'} = -\frac{2}{\mu'\pi}\int \frac{\tau_d'(\eta)d\eta}{y'-\eta}$$

$\bar{w}'$ is average fault slip, w_d' is dislocation slip after removal of average fault slip and τ_d' is dislocation stress.

Figures 8a and 8b contain computed displacements for the unstable case $\mu' = 2.5$ and the stable case $\mu' = 10.$ respectively, both for $b' = 1.$ The fault slip discontinuity is apparent in Figure 8a, but missing in Figure 8b. The latter case is taken to be the analog of a rapid aseismic slip episode. The displacements are qualitatively similar to those in Figure 6b, but somewhat smoother (disregarding numerical perturbations). The principle difference is that the stable-unstable boundary is near $\mu' = 10.$ Thus freeing the plate

bottom promotes instability (the small geometry difference has little effect on the critical μ'). In the finite element approach, the boundary between stable and unstable modes in the complete μ'-b' space remains to be mapped but the Hilbert transform solutions show that for a constant μ', raising b' promotes instability.

Preliminary calculations for a long thrust fault using the finite element method, displacement boundary conditions, and friction law (9) show a response analogous to the strike slip case. As in Figure 6b, the position of greatest fault slip rate migrates updip toward the depth of greatest peak stress prior to the instability. The magnitude of surface uplift rate grows in time as the position of the maximum moves toward the epicenter. Furthermore, fault displacement during the instability is roughtly ten times the nearby uplift preceding the instability. These results are in fair agreement with corresponding elevation changes associated with the $M_L = 6.4$ San Fernando earthquake (Castle et al., 1975).

STABILIZATION AND LOCALIZATION

It must be expected that the preceding instability models, even if substantially correct, will submit to modification by additional mechanical processes in the earth. Stabilization of rupture propagation by pore fluid effects and strain localization are two possible perturbing phenomena which have achieved theoretical development.

Mechanical consequences of pore fluids have received considerable attention, initially as a means of explaining time dependent consolidation in soil mechanics and large horizontal displacements of gently dipping thrust faults. More recently, the success in triggering seismicity by artificial fluid injection (Raleigh et al., 1976), in addition to the well known inverse relation between failure strength of brittle rock and pore fluid pressure, suggests the importance of pore fluid diffusion. Since boundary conditions associated with injection-stimulated seismicity, both in the laboratory and in the field, have no obvious direct counterpart with naturally occuring seismicity near faults, it is unclear whether pore fluids play an essential role in natural earthquake processes. Earthquakes in volcanic regions may constitute an exception. If the earth is sufficiently porous, permeable, and saturated, however, pore fluids have the potential of delaying an instability and thereby of adding another time scale to the deformation. The mechanical analysis, as a result, must consider the coupled equations of elasticity and fluid diffusion in a porous medium.

In the simplest case, the descriptive equations given by Biot (1941) and restructured by Rice and Cleary (1976) suffice as long as the solid skeleton

obeys linear elasticity and the interstitial fluid has constant compressibility and diffuses according to Darcy's law. Then the porous medium will appear temporarily stiffer during rapid deformation and exhibit time dependent elasticity during diffusion. Rice and Cleary (1976) prefer to specify the elastic properties of the medium in terms of an effective Poisson's ratio ν_e. Limiting values of ν_e are: (1) ν for completely drained response or constant pore pressure achievable when the pore fluid diffuses quickly compared to the external loading rate, and (2) ν_u, the undrained value appropriate for no pore fluid flow during rapid external loading. Clearly $\nu_u > \nu$ and rapid deformation implies a stiffer medium. In the instability models so far discussed, a steeper slope on the weakening part of the fault law was possible at instability when the stiffness of the surrounding medium was increased. Therefore, qualitatively one can see that rapid loading of the fluid-impregnated surroundings may allow overshoot of the instability point while retaining quasi-static deformation. Instability is then inevitable as the drained condition is approached, because instability under drained conditions occurs at gentler post-peak slopes and smaller fault slip. It is also evident that such an overshoot introduces an additional time scale associated with accelerating precursory deformation.

Booker (1974) and Rice and Cleary (1976) present analysis of such fluid diffusion effects near faults in plane strain distortion. Rudnicki (1977a), Rice (1977), and Rice and Rudnicki (1978) discuss similar effects in the inclusion model, and Rice and Simons (1976) examine the stabilization of propagating sharp-tipped faults. The latter authors regard migrating fault creep events as examples of fluid stabilization.

A more complicated fluid stabilization problem emerges when shear and volumetric strains are coupled, as during dilatancy. Not only do additional material coefficients arise, but judging from laboratory work, the constitutive law is nonlinear and partially non-elastic. As with the linearly elastic Biot (1941) description, when the medium is fluid filled, rapid deformation may cause a temporary stiffening of the medium and delaying of the instability (Rice, 1975). Coupled dilatant-diffusion fault theories are less well developed than the linearly elastic theories. Rudnicki (1977a,b), Rice (1977), and Rice and Rudnicki (1978) describe stabilization and precursory phenomena arising from dilatancy hardening. Rice (1977), drawing on the work of Rice (1973), shows that dilatancy-induced suctions at a spreading shear crack tip can raise the maximum quasi-static spreading velocity, thus causing stabilization.

Dilatancy-induced stabilization is an essential assumption of the dilatancy fluid diffusion hypotheses for earthquake precursors and instabilities (Nur, 1972; Scholz et al., 1973; Whitcomb et al., 1973). These models focus on the suction created during shear stress buildup near the impending earthquake focus. Pore fluid diffusion from surrounding regions is assumed to then raise the fluid pressure near the focus enough to permit an instability.

Another frequently mentioned stabilizing mechanism is healing or aging of recently disrupted fault zone material by a chemical or viscoelastic interlocking of mineral grains. Apparently no laboratory observations on bulk samples exist to demonstrate the importance or form of this phenomenon. Dieterich (1972) gives empirical evidence that, just as with metals and

ceramics, the coefficient of friction between smooth rock surfaces increases proportional to the logarithm of the contact time.

Localization of deformation into narrow shear bands, traditionally called faults, may replace more uniform deformation prior to or during an inertial instabilty. Failure of laboratory samples is in fact often due to localization if the testing machine is stiff enough to preclude instability, although examples of homogeneous post-peak softening exist.

Despite the obvious importance of localization, the subject has only recently received much theoretical attention, principally from Rice (1973), Rudnicki and Rice (1975), Rudnicki (1977a,b) and Rice (1976, 1977). They take the view that a single constitutive law is sufficient to describe all future deformation in a material. At some deformation state that depends on geometric, boundary condition, and constitutive parameters, the currently homogeneous strain state may either continue to be homogeneous or tend to generate a localization band within the otherwise uniformly strained body. This amounts to a bifurcation of admissible solutions to the boundary value problem. It is assumed that localization occurs over a planar zone in three dimensions or a band in two. The spread of a localization band from a point imperfection or crack tip is not well studied. Microscopically, localization in compact brittle rocks is apparently associated with creation of porosity, grain comminution, and relative displacement and rotation of grains.

With such dilatant frictional mechanisms in mind, Rudnicki and Rice (1975) developed by a rather intricate analysis the conditions for localization in an infinitely long shear band. Rudnicki (1977a,b) applied the theory to localization within the ellipsoidal inclusion considered previously by

postulating the incremental elastic-plastic constitutive law

$$2d\varepsilon'_{ij} = \frac{d\tau'_{ij}}{G} + \frac{\tau'_{ij}}{h\bar{\tau}}\left[\frac{\tau'_{kl}d\tau_{kl}}{2\bar{\tau}} + \mu\frac{d\tau_{kk}}{3}\right]$$

$$d\varepsilon_{kk} = \frac{d\tau_{kk}}{3K} + \frac{\beta}{h}\left[\frac{\tau'_{kl}d\tau_{kl}}{2\bar{\tau}} + \mu\frac{d\tau_{kk}}{3}\right] \qquad (10)$$

$$\bar{\tau} = (\tau'_{ij}\tau'_{ij}/2)^{\frac{1}{2}}$$

where G, h, μ, K, and β are the rigidity modulus, plastic hardening modulus, internal friction coefficient, bulk modulus, and dilatancy coefficient respectively, all functions of the state of stress. Primed variables are the deviatoric components and primeless variables are volumetric components. The β term connects shear stress with volumetric strain. h is the slope at a point on the suitably defined stress strain curve. Rudnicki (1977a) gives geometric interpretations of the coefficients for the far field stress states of axisymmetric compression (approximately the situation for uniaxial compression of laboratory samples under confining pressure) and estimates $\mu \sim 1.$ and $\beta \sim .5$ from experimental data. The condition for localization can be translated into finding the critical value h_{cr} of the hardening modulus h in terms of geometric and material coefficients as well as the remote stress state.

Rudnicki (1977a) finds that when using (10) for both the spherical and narrow inclusions (the latter simulating a thin fault zone) localization will generally precede inertial instability, though localization and instability will be closer together and to the peak stress for narrow inclusions. According to the theory, localization also depends on the stress state symmetry and will occur earliest and near peak stress in pure shear (e.g. a long strike slip fault). Inertial instability, of course, occurs only in the post peak

region where slopes are negative, for the sphere far more negative than for the narrow zone. Localization implies large strains in the localized zone, hence a weakened post-peak stress state, which in turn favors instability. In this view, localization is the most fundamental cause of earthquake instabilities.

Both localization and inertial instability depend delicately on the material coefficients, none of which are well known in the laboratory and essentially unknown in the field. Ayden (1977) has attempted to reconcile observations of exposed faults and shear bands in porous sandstone (Entrada and Navajo) with the preceding localization theory.

STRAIN RATE SOFTENING INSTABILITY MODELS

An inertia-limited instability appears possible in certain situations if the fault weakens with strain rate or velocity instead of strain or slip. At least two types of weakening mechanisms are possible. In the first a temperature- and activation energy-dependent creep law allows an increasing strain rate with a rise in temperature at constant stress. Heat generated during deformation of the shear zone raises the temperature, hence the strain rate even more, and an instabiltiy may ensue. A variant of this idea is the increase of pore pressure due to dehydration of hydrous minerals during the temperature rise (Raleigh, 1977). Thermal softening mechanisms are most attractive for explaining intermediate and deep focus earthquakes where conventional Coulomb frictional failure is considered to be unlikely because of high confining pressure.

To explain deep focus earthquakes, Griggs and Baker (1969) postulated that material within the thin, infinite fault zone of Figure 1a obeys the empirical creep law

$$\dot{\gamma} = \alpha\tau^n \exp(-E/RT) \qquad (11)$$

where $\dot{\gamma}$ is shear strain rate across the fault width $2h$, α is a material constant, τ is shear stress, n is a number between 2 and 10, E is the activation energy, R the gas constant, and T temperature. Although Griggs and Baker (1969) did not say so, the strain rate softening character of the fault zone emerges readily by combining (11) with the equation converting plastic work to thermal energy assuming slow diffusive heat loss. Strictly adiabatic deformation yields a strain softening law. The result is that strain rate varies inversely with stress, the opposite of both forms like (11) and viscous laws.

By combining (11) with an energy balance equation and the heat flow equation, Griggs and Baker (1969) computed the onset of inertial instabilities upon fault melting for increasing displacement boundary conditions (U in Figure 1a). Unreasonably large values of $\dot{U}$ and H/μ, μ being the rigidity, were required to provide instabilities, however. Similar calculations seemed to adequately account for the stick-slip sliding reported by Bridgman (1936) in samples deformed by torsion at pressures up to 50 kbar. The hot creep mechanism for instabilty proposed by Orowan (1960) appears to be essentially equivalent to a strain rate softening constitutive law.

Another mechanism, perhaps more appropriate for crustal depths where rocks are thought to be brittle, is indicated by friction experiments of Dieterich (1978a). He noted that frictional resistance between rock sample

surfaces decreased about 5 percent with increasing sliding velocity over the range 10^{-5} to 10^{-2} cm/sec. The microscopic processes responsible are not firmly established, but Dieterich (1978a) conjectures that a competition between a time-dependent asperity adhesion and a displacement-dependent tearing apart may transpire. This type of mechanism is easy to conceptually generalize to the continuum where one might envision fault zone rocks as strain weakening during high strain rates before an earthquake, but healing or aging during low post-earthquake strain rates. Virtually no relevant laboratory or theoretical work on this subject exists, however. Dieterich (1978b,c) proposes an empirical friction relation depending on slip and slip rate. With the friction law, numerical simulations of the laboratory experiment are able to reproduce the main features of observed stick slip events.

Weertman (1967, 1978) shows that inertial instability is possible on a fault segment when the friction decreases rapidly enough with sliding velocity. As in his similar study with a slip weakening law, dislocation distributions and a position independent friction law are used.

CONCLUSION

All of the well developed instability models rely on a strain (displacement) or strain rate (velocity) weakening fault zone. In fact Weertman (1978) argues that earthquakes cannot occur unless the fault weakens with slippage. Models differ primarily in their assumptions of fault zone geometry and spatial extent of actively failing fault material, the latter being controlled by heterogeneity of constitutive properties. Increasing fault slip rate prior to instabilities appears to be a universal feature of models formulated to include also precursory deformation. Thus models have potential of being distinguishable by field observations (when they become available) on the basis of theoretical predictions. Only a few models are in a form that is readily translatable into a field experiment, mainly because of geometry assumptions. Moreover, no single model connecting precursory deformation directly with earthquake parameters exists; such a model would simulate the pre-earthquake loading, onset of instability, dynamic propagation, and cessation of propagation. On the other hand, field testing is likely to be difficult because several mechanisms are able to stabilize or delay instabilities, including Biot-like stiffening of fluid-filled porous media, dilatancy-pore fluid coupling, and fault zone healing. These mechanisms may not exist in all regions, however, whereas some form of strain softening probably does.

Considerable spacial variation of fault zone properties is expected in the earth, and some heterogeneity is essential to all but the most elementary models. So far, only the simplest heterogeneities have been studied, and only for a single instability.

Knowledge of the precise form and coefficient values for fault constitutive laws is sparse. This shortcoming is in principle surmountable in two ways. The first is to assume that laboratory data represents in situ behavior and is appropriate for large scale boundary value problems. Both tests on bulk samples and instability simulations can provide constitutive data. Unfortunately, few experiments have been performed to elucidate the key constitutive behavior required by theoretical models, namely post-peak weakening, localization, and healing. Another approach is to postulate plausible constitutive laws and compare field data with the subset of theoretical calculations yielding instabilities. Limits on in situ constitutive properties can be thereby estimated. Successful theories might then be applied in the future to recognize anomalous crustal deformation precursory to earthquakes.

ACKNOWLEDGMENTS

Discussions with J. Andrews and G. Mavko have been helpful.

REFERENCES

Andrews, D. J., Rupture propagation with finite stress in antiplane strain, J. Geophys. Res., 81, 3575-3582, 1976.

Aydin, A., Faulting in sandstone, Ph.D. Thesis, Stanford University, Stanford, California, 1977.

Biot, M. A., General theory of three-dimensional consolidation, J. Appl. Phys., 12, 155-164, 1941.

Booker, J. R., Time-dependent strain following faulting of a porous medium, J. Geophys. Res., 79, 2037-2044, 1974.

Bridgman, P. W., Shearing phenomena at high pressure of possible importance to geology, Jour. Geology, 44, 653-669, 1936.

Byerlee, J. D., The mechanics of stick-slip, Tectonophysics, 9, 475-486, 1970.

Castle, R. O., J. P. Church, M. R. Elliott, and N. L. Morrison, Vertical crustal movements preceding and accompanying the San Fernando earthquake of February 9, 1971: A summary, Tectonophysics, 29, 127-140, 1975.

Cleary, M. P., Continuously distributed dislocation model for shear-bands in softening materials, Int. Jour. Numer. Meth. Eng., 10, 679-702, 1976.

Dieterich, J. H., Time-dependent friction in rocks, J. Geophys. Res., 77, 3690-3697, 1972.

Dieterich, J. H., Time-dependent friction and the mechanics of stick-slip, Pure and Applied Geophysics, in press, 1978a.

Dieterich, J. H., Modeling of rock friction, Part 1: Experimental results and constitutive equations, this volume, 1978b.

Dieterich, J. H., Modeling of rock friction;, Part 2: Simulation of pre-seismic slip, this volume, 1978c.

Eshelby, J. D., The determination of the elastic field of an ellipsoidal inclusion and related problems, Proc. Roy. Soc. London, Ser. A, 241, 376-396, 1957.

Griffith, A. A., The phenomena of rupture and flow in solids, Roy. Soc. London Philos. Trans., Ser. A., 221, 163-198, 1921.

Griggs, D. T., and D. W. Baker, The origin of deep-focus-earthquakes, in Properties of Matter under Unusual Conditions, ed. H. Mark and S. Fernbach, 23-42, Interscience, New York, 1969.

Griggs, D., and J. Handin, Observations on fracture and a hypothesis of earthquakes, in Rock Deformation, Geol. Soc. Amer. Mem. 79, 347-364, 1960.

Hafner, W., Stress distributions and faulting, Bull. Geol. Soc. Amer, 62, 373-398, 1951.

Irwin, G. R., Fracture I, in Handbuch der Physik, vol. 6,551-590, Springer, Berlin, 1958.

Jaeger, J. C., and N. G. W. Cook, Fundamentals of Rock Mechanics, 2nd ed., Chapman and Hall, London, 1976.

Mavko, G. M., Time dependent fault mechanics and wave propagation in rocks, Ph.D Thesis, Stanford University, Stanford, California, 1977.

Nur, A., Dilatancy, pore fluids, and premonitory variations of t /t travel times, Bull. Seismol. Soc. Amer., 62, 1217-1222, 1972.

Nur, A., and P. Schultz, Fluid flow and faulting, 2: A stiffness model for seismicity, in Proceedings of the Conference on Tectonic Problems of the San Andreas Fault System, vol. 13, 405-416, 1973.

Odé, H., Faulting as a velocity discontinuity in plastic deformation, in Rock Deformation, Geol. Soc. Amer. Mem. 79, 293-321, 1960.

Orowan, E., Mechanism of seismic faulting, in Rock Deformation, Geol. Soc. Amer. Mem. 79, 323-345, 1960.

Palmer, A. C., and J. R. Rice, The growth of slip surfaces in the progressive failure of over-consolidated clay, Proc. Roy. Soc. London A, 332, 527-548, 1973.

Raleigh, C. B., Frictional heating, dehydration, and earthquake stress drops, Proceedings of Confernce II, Experimental studies of rock friction with application to earthquake predicton, USGS, 291-304, Palo Alto, 1977.

Raleigh, C. B., J. H. Healy, and J. D. Bredehoeft, An experiment in earthquake control at Rangely, Colorado, Science, 191, 1230-1237, 1976.

Reid, H. F., Permanent displacements of the grounds, in The California Earthquake of April 18, 1906, Report of the State Earthquake Investigation Commission, vol. 2, 16-28, Carnegie Institution of Washington, Washington, D. C., 1910.

Rice, J. R., Mathematical analysis in the mechanics of fracture, in Fracture, An Advanced Treatise, vol. 2, 191-311, Academic Press, 1968.

Rice, J. R., The initiation and growth of shear bands, in Plasticity and Soil Mechanics, ed. A. C. Palmer, Cambridge University Engineering Department, 263-274, Cambridge, England, 1973.

Rice, J. R., On the stability of dilatant hardening for saturated rock masses, J. Geophys. Res., 80, 1531-1536, 1975.

Rice, J. R., The localization of plastic deformation, Theoretical and Applied Mechanics, vol. 1, Proc. 14th IUTAM Cong., Delft, 1976, 207-220, North Holland Publ. Co., Amsterdam, 1976.

Rice, J. R., Theory of precursory processes in the inception of earthquake rupture, Proceedings of the Symposium on Physics of Earthquake Sources, Academy of Sciences, German Democratic Republic, in press, 1977.

Rice, J. R., and M. P. Cleary, Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents, J. Geophys. Res., 14, 227-241, 1976.

Rice, J. R., and J. W. Rudnicki, Earthquake precursory effects due to pore fluid stabilization of a weakening fault zone, this volume, 1978.

Rice, J. R., and D. A. Simons, The stabilization of spreading shear faults by coupled deformation-diffusion effects in fluid-infiltrated porous materials, J. Geophys. Res., 81, 5322-5334, 1976.

Rudnicki, J. W., The inception of faulting in a rock mass with a weakened zone, J. Geophys. Res., 82, 844-854, 1977a.

Rudnicki, J. W., Localization of deformation, brittle rock failure, and a model for the inception of earth faulting, Ph. D. thesis, Brown University, Providence, Rhode Island, 1977b.

Rudnicki, J. W., and J. R. Rice, Conditions for the localization of deformation in pressure-sensitive dilatant materials, J. Mech. Phys. Solids, 23, 371-394, 1975.

Rummel, F. J. Experimentelle Untersuchungen zum Bruchvorgang in Gesteinen, Berichte des Institutes für Geophysik der Ruhr-Universität Bochum, Nr. 4, 1975.

Sanford, A. R., Analytical and experimental study of simple geological structures, Bull. Geol. Soc. Amer., 70, 19-51, 1959.

Scholz, C. H., L. R. Sykes, and Y. P. Aggarwal, Earthquake prediction: A physical basis, Science, 181, 803-810, 1973.

Stuart, W. D., Diffusionless dilatancy model for earthquake precursors, Geoph. Res. Let., 1, 261-264, 1974.

Stuart, W. D., Strain softening prior to two-dimensional strike-slip earth quakes, submitted to J. Geophys. Res., 1977.

Stuart, W. D., and G. M. Mavko, Earthquake instability on a strike slip fault, submitted to Jour. Geophys. Res., 1978.

Walsh, J. B., Stiffness in faulting and in friction experiments, J. Geophys. Res., 76, 8597-8598, 1971.

Weertman, J., Relationship between displacements on a free surface and the stress on a fault, Bull. Seismol. Soc. Amer., 55, 945-953, 1965.

Weertman, J., Theory of infinitesimal dislocations distributed on a plane applied to discontinuous yield phenomena, Can. J. Physics, 45, 797-807, 1967.

Weertman, J., Inherent instability of quasistatic creep slippage on a fault, this volume, 1978.

Whitcomb, J. H., J. D. Garmany, and D. L. Anderson, Earthquake prediction: Variation of seismic velocities before the San Fernando Earthquake, Science, 180, 632-635, 1973.

Zoback, M. D., and J. D. Byerlee, A note on the deformational behavior and permeability of crushed granite, Int. J. Rock Mech. Min. Sci. & Geomech. Abstr., 13, 291-294, 1976.

TABLE 1. Classification of instability models

Reference	Model Geometry	Remote Boundary Conditon	Fault Geometry	Fault Properties
This paper	two-dimensional simple shear	displacement or stress	constant width band	quadratic in strain with peak stress
Weertman (1967)	two-dimensional	stress	finite crack	displacement and velocity softening
Rudnicki (1977)	full space	stress or strain	homogeneous ellipsoidal inclusion	elastic-plastic, dilatant
Palmer and Rice (1973)	two-dimensional	displacement or stress	sharp tipped crack	linear displacement softening from peak stress to residual stress
Cleary (1976)	a. full space b. free surface inclined to gravity vector	a. stress b. gravity	a. sharp tipped finite crack b. sharp tipped crack intersecting free surface	linear and hyperbolic displacement softening from peak to residual stress
Stuart (1977)	two-dimensional, strike-slip; free surface	displacement	throughgoing fault plane	Gaussian in fault slip and depth
Stuart and Mavko (1978)	two-dimensional, strike-slip; free surface	displacement	throughgoing fault plane	Gaussian in fault slip and depth
Griggs and Baker (1969)	two-dimensional simple shear	velocity	constant width band	strain rate softening from creep law and plastic work
Weertman (1978)	two-dimensional	stress	transform fault with spreading centers	velocity softening

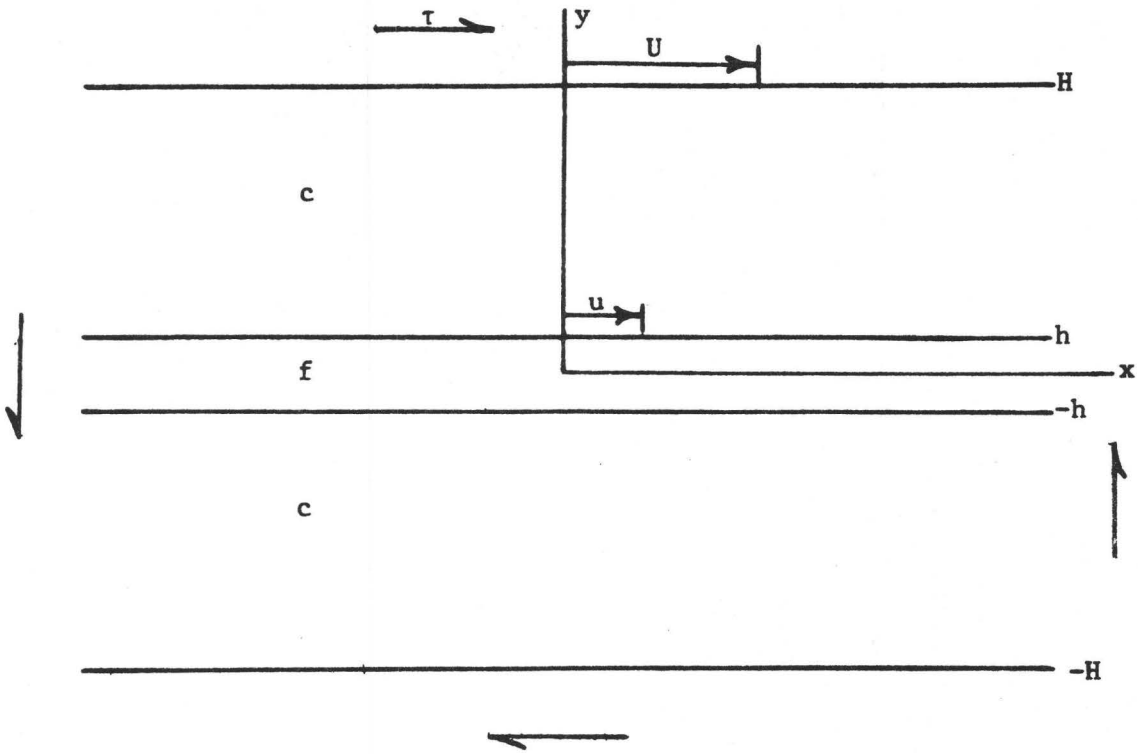
TABLE 1. (cont.) Classification of instability models

Reference	Instability Criterion[1]	Precursory Deformation	Solution Method	Pore Fluid Stabilization
This paper	$\partial u/\partial U \rightarrow \infty$ $\partial u/\partial \tau \rightarrow \infty$	increasing fault strain rate	analytic	Footnote 2
Weertman (1967)	$\partial \varepsilon/\partial \tau \rightarrow \infty$	---	analytic	---
Rudnicki (1977)	$\partial \varepsilon/\partial \varepsilon_\infty \rightarrow \infty$	increasing inclusion strain rate	analytic	3
Palmer and Rice (1973)	critical elastic energy release rate	---	analytic	3
Cleary (1976)	critical elastic energy release rate	---	distributed dislocations, numerical solution	---
Stuart (1977)	$\partial w'/\partial W' \rightarrow \infty$	increasing fault slip rate	finite element	---
Stuart and Mavko (1978)	$\partial w'/\partial W' \rightarrow \infty$	increasing fault slip rate	distributed dislocations, numerical solution	---
Griggs and Baker (1969)	$\partial \dot{u}/\partial \dot{U} \rightarrow \infty$	increasing fault strain rate	finite difference	---
Weertman (1978)	$\partial u/\partial \tau_\infty \rightarrow \infty$	---	analytic	---

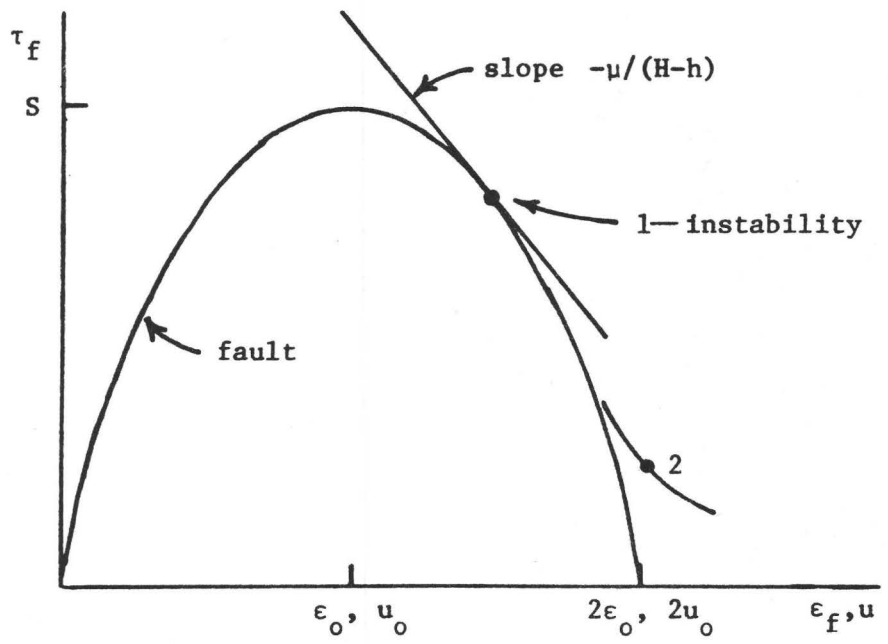
1. Derivative of fault slip (u,w') or strain (ε) with respect to remotely applied displacement (U, W'), strain (ε_∞), or stress (τ_∞).

2. Rice (1975) examined stabilization of dilatant material in homogeneous simple shear deformation

3. Stabilizing mechanisms reviewed and developed in Rice (1977), Rice and Rudnicki (1978) and references cited therein.

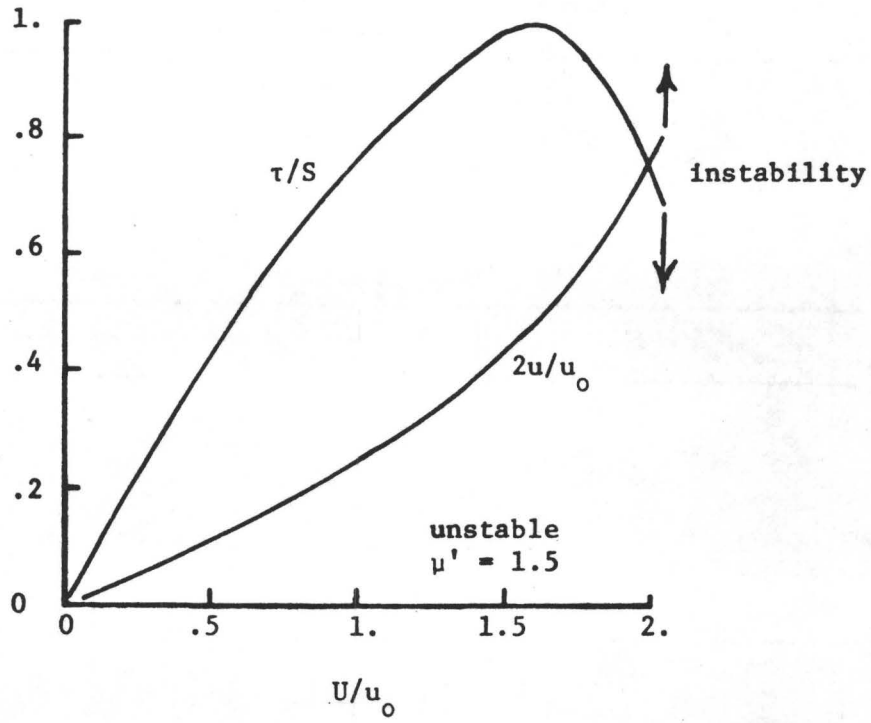


(a) Geometry

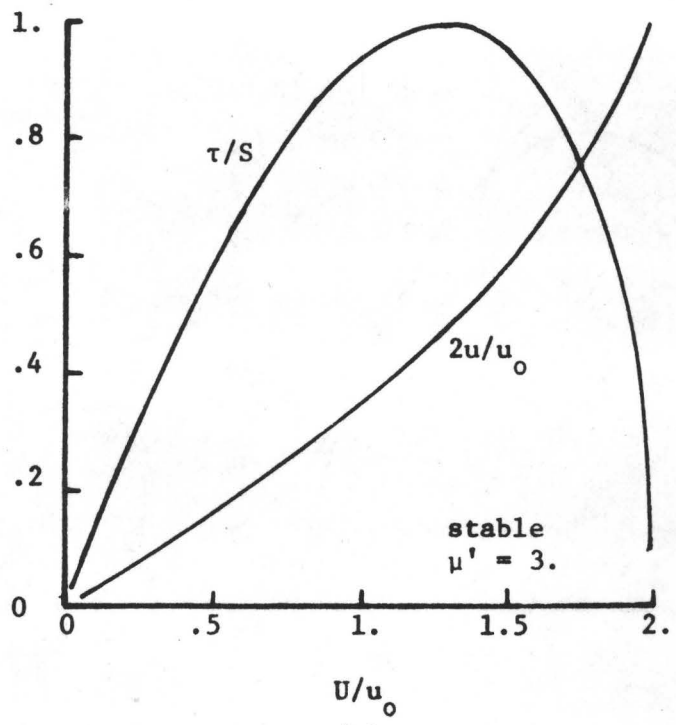


(b) Fault zone constitutive law with strain softening

Fig. 1 The simplest instability model.

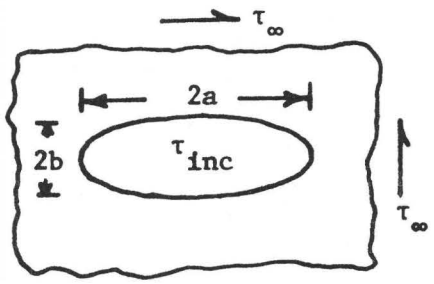


(a)

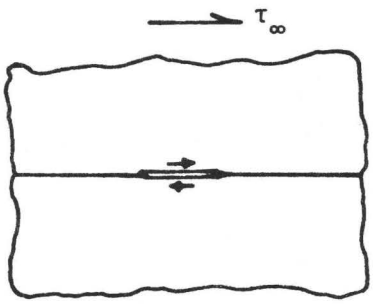


(b)

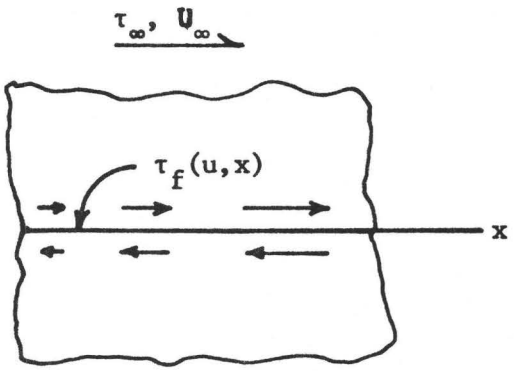
Fig. 2 Fault slip u and stress τ as a function of increasing remote displacement U for unstable μ' = 1 and stable μ' = 3 cases.



(a) Ellipsoidal inclusion

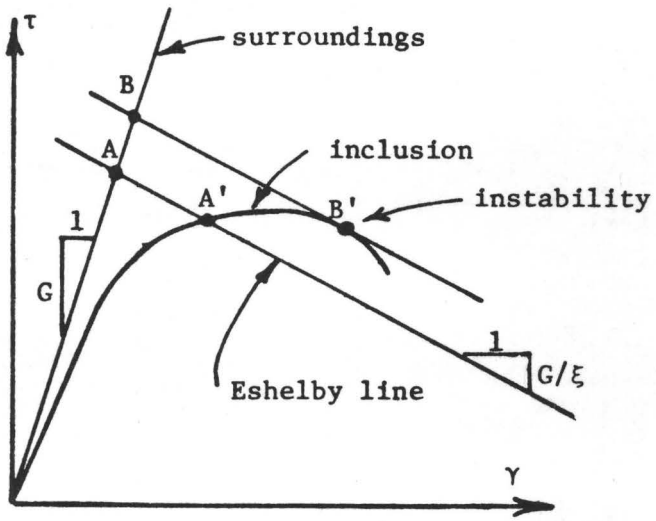


(b) Finite fault

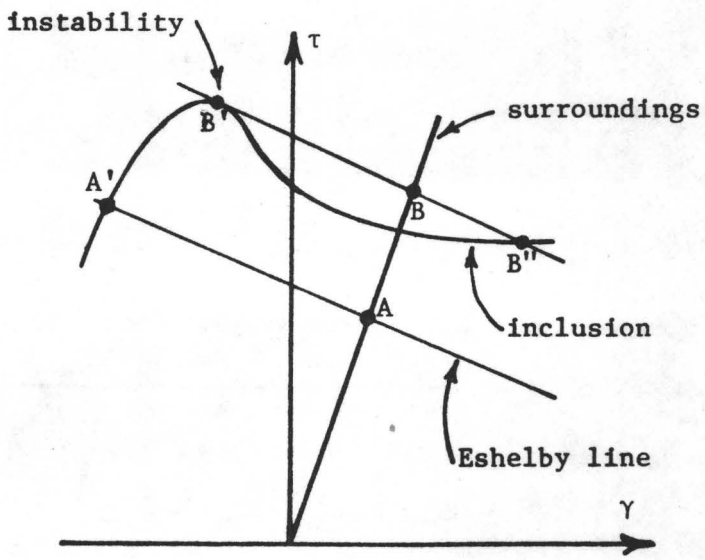


(c) Throughgoing fault

Fig. 3 Instability models

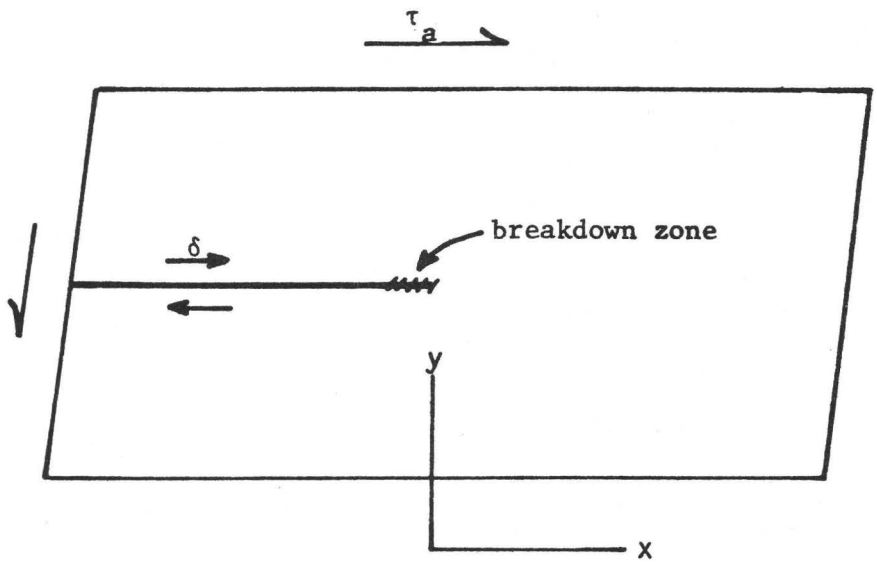


(a) Weak inclusion

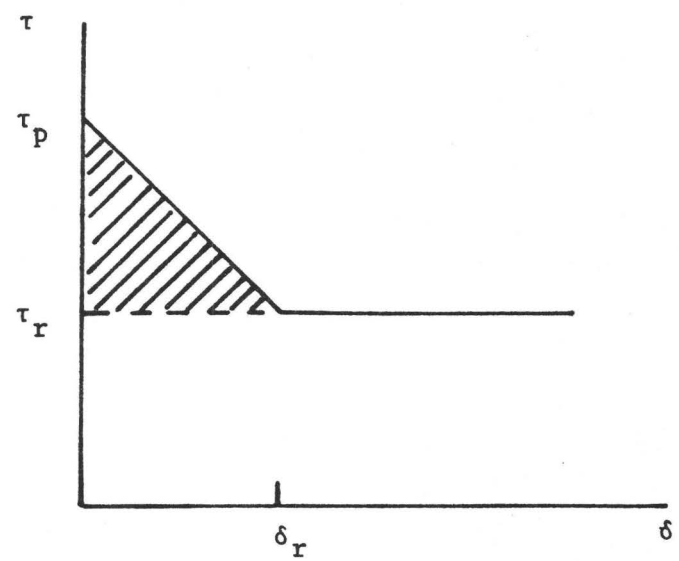


(b) Strong inclusion

Fig. 4 Deformation and instability of inclusion model. Modified from Rice (1977).

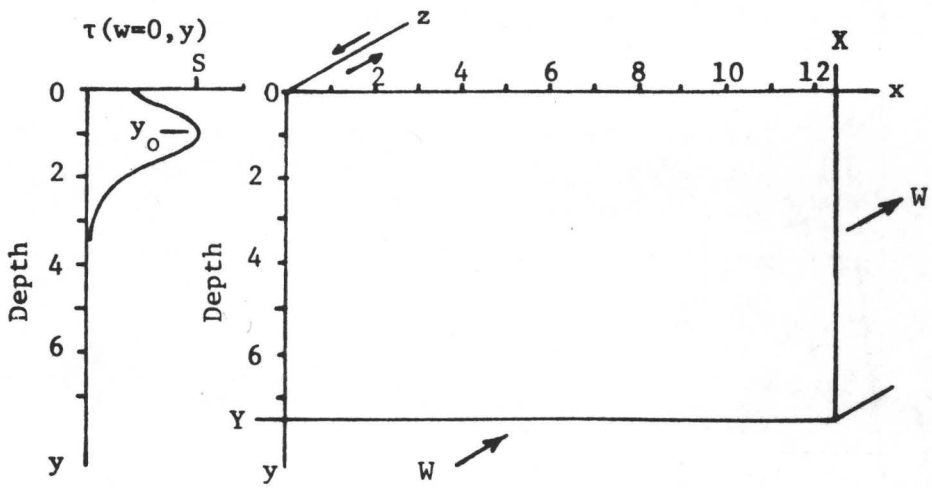


(a) Geometry

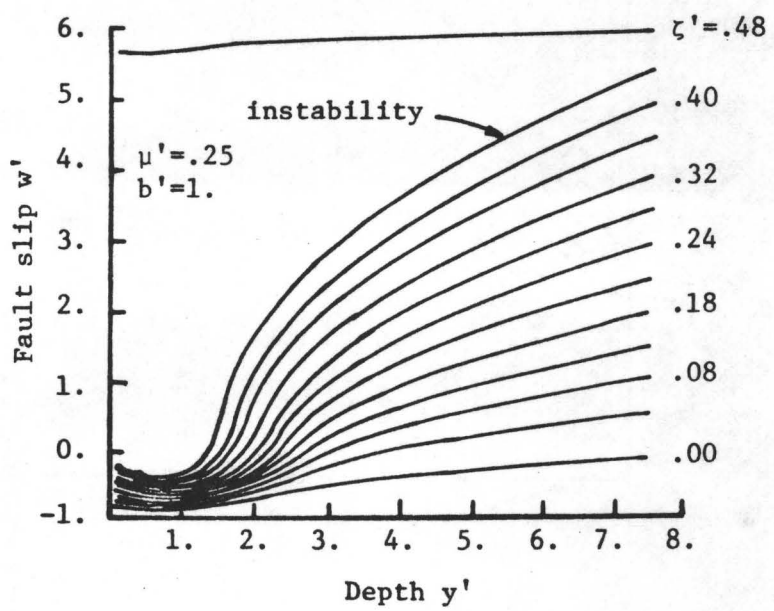


(b) Stress-slip law for fault

Fig. 5 Sharp tip fault model.

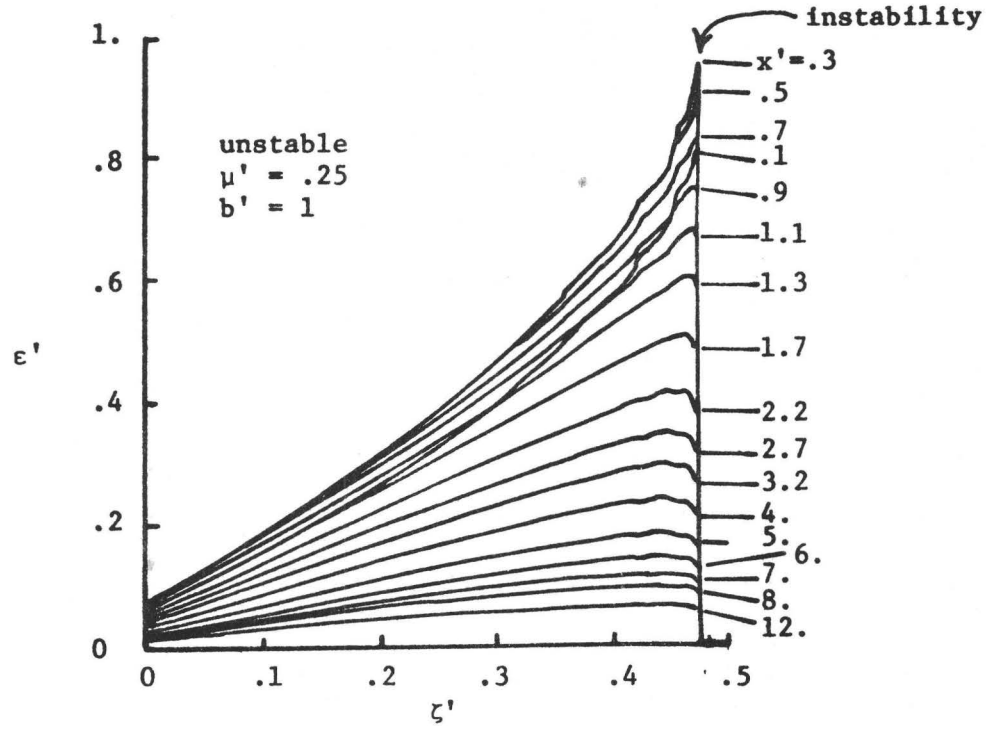


(a) Geometry and peak stress variation

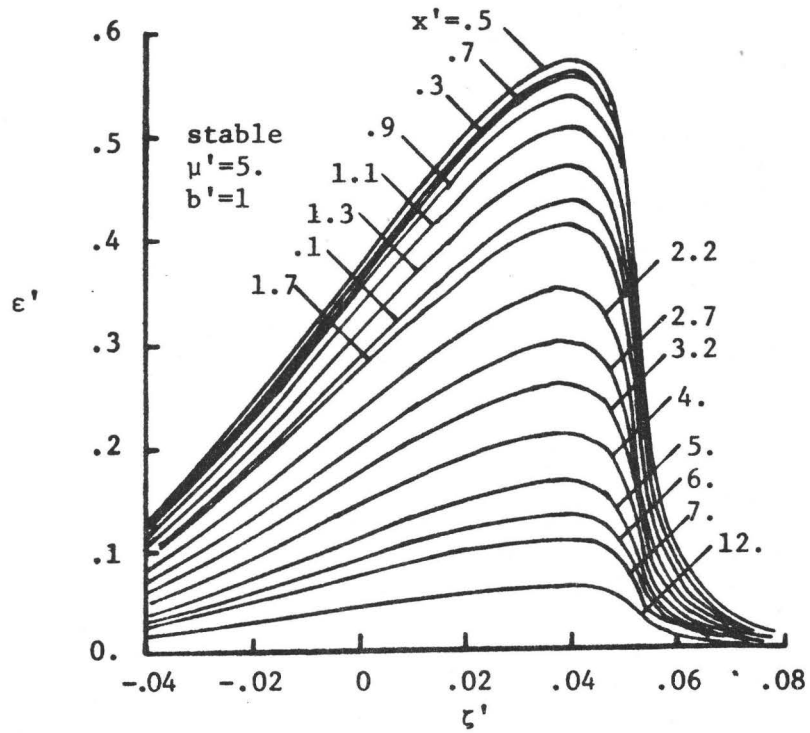


(b) Fault slip vs. depth at successive times ζ'

Fig. 6 Two-dimensional strike-slip fault.

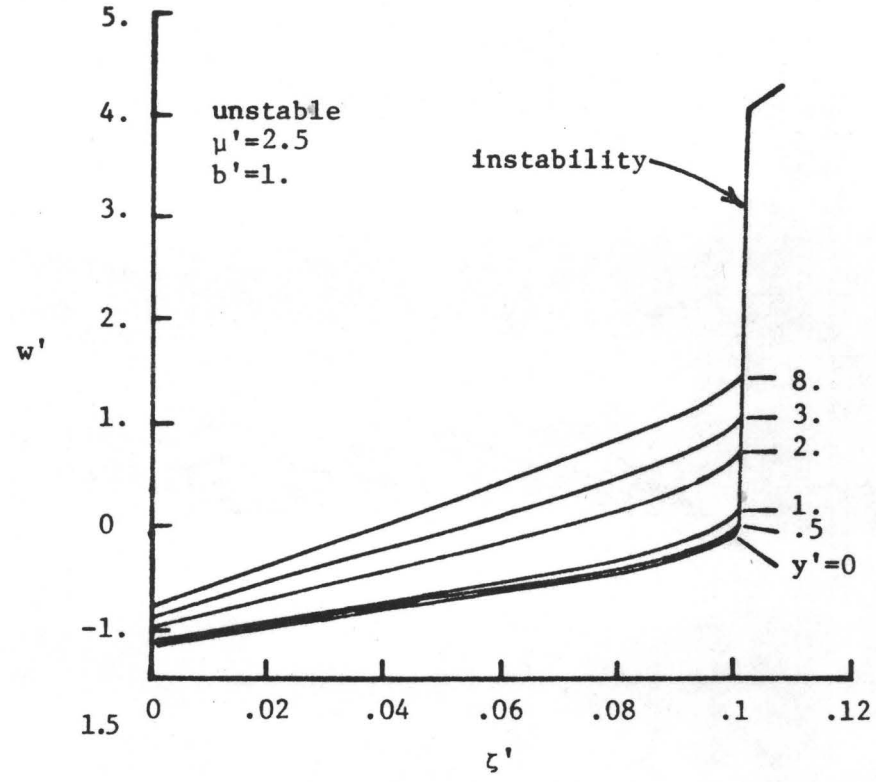


(a)

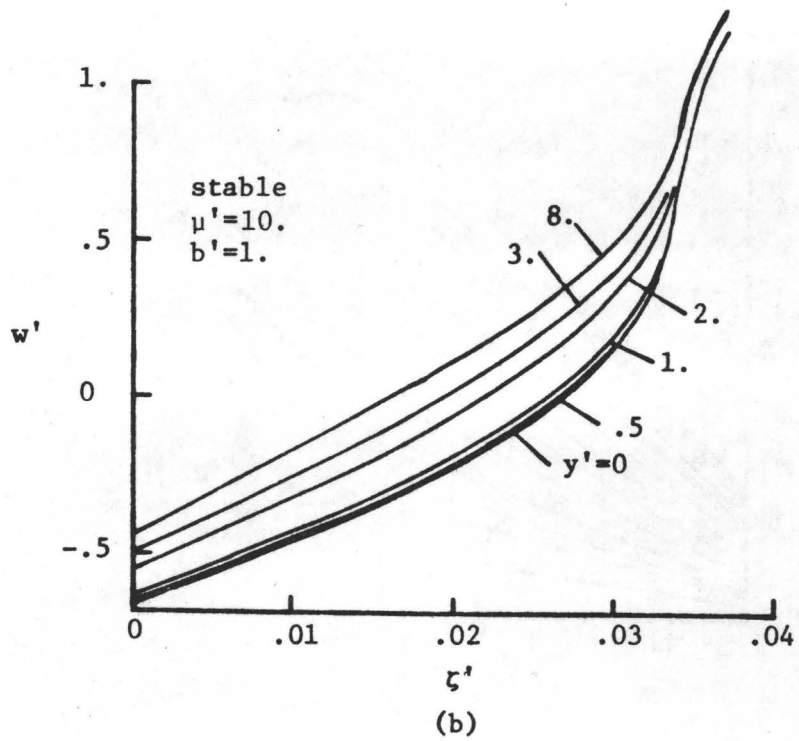


(b)

Fig. 7 Free surface shear strain ε' vs. time ζ' for a two-dimensional, strike-slip fault: finite element solution. x' is distance from fault trace.



(a)



(b)

Fig. 8 Fault displacement w' vs. time ζ' at various depths y' for two-dimensional strike-slip fault: dislocation solution.

MODELS FOR THE DISTRIBUTION OF STRESS AND STRAIN ADJACENT TO THE SAN ANDREAS FAULT

D.L. Turcotte and R.T. Clancy

Department of Geological Sciences
Cornell University
Ithaca, New York 14853

D.A. Spence

Department of Engineering Sciences
University of Oxford
Oxford, England

F.H. Kulhawy

School of Civil and Environmental Engineering
Cornell University
Ithaca, New York 14853

Two studies are reported in this paper. The first concerns the interaction between an elastic lithosphere and a visco-elastic asthenosphere caused by periodic fault displacement. The results of an analytical solution show that the periodic fault motion is damped out by the interaction several lithospheric thicknesses from the fault. The second study reports a finite element investigation of a depth dependent fault rheology. After a great earthquake the stress on the locked portion of the fault is assumed to increase linearly with depth; the stress on the lower plastic portion of the fault is assumed to be independent of time and to decrease linearly with depth. The depth of the locked portion of the fault is decreased with increasing strain accumulation in order to remove the stress singularity at its lower end. The surface distribution of strain and stress is obtained.

INTRODUCTION

There is no question that the San Andreas fault represents a major boundary between the Pacific and North American plates. If the fault accomodates the relative motion between the two plates then it must lie on a small circle about the pole of rotation that defines the relative motion of the plates. Taking the pole of rotation (50.9°N, -66.3°E) given by Minster et al. (1974), good agreement between the circle of rotation and the general trend of the San Andreas fault is obtained. This is evidence that the fault is the primary boundary between the Pacific and North American plates.

In order to understand the cyclical accumulation and release of stress and strain on the San Andreas fault it is necessary to hypothesize appropriate models. A two-dimensional model for the accumulation and release of slip on a strike-slip fault has been proposed by Turcotte and Spence (1974). In this model the lithosphere was treated as rigid elastic plates. The fault was taken to be the boundary between two plates sliding laterally past one another. The upper part of the fault was assumed to be locked (zero displacement) and the lower part of the fault zone and the lower boundary of the plates were assumed to be free sliding (zero shear stress). An analytical solution for the distribution of stress and strain was given. During a great earthquake the accumulated stress and strain was assumed to be relieved.

Implicit in this model is the assumption that the failure stress on the locked brittle zone is large compared with the yield stress on the deeper plastic zone. After a great earthquake the remaining stress on both the upper brittle zone and the lower plastic zone is small compared with the initial stress on the intially locked brittle zone before the earthquake. This assumption has been questioned by Savage (1975) who proposed a transfer of stress between the upper brittle zone and the lower plastic zone.

In order to explain the longitudinal behavior of the San Andreas fault a three-dimensional model for the accumulation and release of stress and strain was proposed by Spence and Turcotte (1976). Two semi-infinite plates acted on by shearing forces were assumed to be sliding freely past each other except for two locked sections. These sections corresponding to the northern and southern locked sections elastic of the fault (Turcotte, 1977) were assumed to be locked to a specified depth. At greater depths the fault was assumed to be free sliding as well as on other sections of the fault. Using the technique of matched asymptotic expansions, analytic expressions were obtained for the three-dimensional distribution of strain and stress. According to this elastic analysis the locked sections impede the lateral sliding between the plates, and the plate interactions extend large distances from the locked sections.

Asthenospheric Damping

It is known from studies of postglacial rebound that the asthenosphere exhibits viscous behavior with a viscosity near 10^{21} poise. An important question is whether interactions between the elastic lithosphere and viscous asthenosphere damps the periodic motions associated with fault slip at plate boundaries.

An approximate equation for stress diffusion has been derived by Elsasser (1969) assuming an elastic plate overlies a viscous layer. This has been applied to periodic displacements at plate boundaries by Bott and Dean (1973) and by Anderson (1975). Melosh (1976) has considered the interaction of an elastic lithosphere with a non-Newtonian asthenosphere. Since the elastic constants of the asthenosphere are nearly equal to the elastic constants of the lithosphere, in most cases, it is not consistent to retain the elastic behavior of the lithosphere while dropping the elastic behavior of the asthenosphere. The problem of the time-dependent response of an elastic lithosphere overlying a viscoelastic asthenosphere has been considered by Nur and Mavko (1974). They

considered a uniform initial displacement on a fault to a prescribed depth and the subsequent relaxation problem.

Spence and Turcotte (1977) have considered the problem of two semi-infinite lithospheric plates of elastic material with a constant thickness h assumed to be in contact along their common edge. This common edge models the fault, which is taken as the strip $x = 0$, $-h < y < 0$ of a plane z = constant. Beneath the semi-infinite plates the asthenosphere is modelled as a semi-infinite, elasto-viscous material with constant properties. The stress-strain relation in the asthenosphere is taken to be the stress-relaxing Maxwell law. The model is illustrated in Figure 1.

The deformation both in the plate and in the asthenosphere must satisfy the equilibrium equation.

$$\rho \frac{\partial^2 w}{\partial t^2} = \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} \tag{1}$$

(ρ = density). In the lithosphere the stresses are elastic, i.e.,

$$\sigma_{ij} = G e_{ij} \tag{2}$$

where G is the shear modulus and e_{ij} are the antiplane strain components

$$e_{xz} = \frac{\partial w_L}{\partial x}, \quad e_{yz} = \frac{\partial w_L}{\partial y} \tag{3}$$

(the subscripts L, A will denote displacements in the lithosphere and in the asthenosphere, respectively). Substitution of (2) and (3) into (1) gives the wave equation

$$\frac{\rho}{G} \frac{\partial^2 w_L}{\partial t} = \Delta w_L \tag{4}$$

where $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2$, which holds within the lithosphere.

In the asthenosphere (2) is replaced by the Maxwell law which is given in differential form by

$$\frac{\partial e_{ij}}{\partial t} = \frac{1}{G}\frac{\partial \sigma_{ij}}{\partial t} + \frac{1}{\eta}\sigma_{ij} \tag{5}$$

where η is the Newtonian viscosity. The viscoelastic relaxation time is then

$$T = \frac{\eta}{G} \tag{6}$$

It is reasonable to assume that the densities and shear moduli of the lithosphere and asthenosphere are the same. Combining (3) and (5), eliminating the resulting stress components by means by (1) and integrating with respect to time (the constant of integration can be taken to be zero for periodic motions) gives

$$\frac{\rho}{G}\left(\frac{\partial^2 w_A}{\partial t^2} + \frac{1}{T}\frac{\partial w_A}{\partial t}\right) = \Delta w_A \tag{7}$$

On the upper surface of the plate the shear traction vanishes:

$$y = 0: \quad \sigma_{yz} = G\frac{\partial w_L}{\partial y} = 0 \tag{8}$$

while on the interface between the lithosphere and asthenosphere the displacement w and shear stress σ_{yz} are continuous, i.e.,

$$\left.\begin{aligned} &y = -h: \\ &\qquad w_L = w_A \\ &\qquad \frac{\partial w_A}{\partial t dy} = \left(\frac{\partial}{\partial t} + \frac{1}{T}\right)\frac{\partial w_L}{\partial y} \end{aligned}\right\} \tag{9}$$

We are primarily interested in processes which occur on a time scale of the order of relaxation time T. Introducing dimensionless scaled time and coordinates by

$$t = T\tilde{t}, \quad x = h\tilde{x}, \quad y = h\tilde{y}, \quad w = h\tilde{w} \tag{10}$$

(4) and (5) become

$$\varepsilon^2 \frac{\partial^2 \tilde{w}_L}{\partial \tilde{t}^2} = \Delta \tilde{w}_L \tag{11}$$

$$\varepsilon^2 \left(\frac{\partial}{\partial \tilde{t}} + 1 \right) \frac{\partial \tilde{w}_A}{\partial t} = \Delta \tilde{w}_A \tag{12}$$

respectively, where $\tilde{\Delta} = \partial^2/\partial \tilde{x}^2 + \partial^2/\partial \tilde{y}^2$ and

$$\varepsilon = \frac{h}{T} \left(\frac{\rho}{G} \right)^{\frac{1}{2}} \tag{13}$$

Thus ε can be interpreted as the ratio of the time taken by a shear wave to cross the plate to the viscoelastic relaxation time.

We assume that the lithosphere has a thickness h = 100 km. The lithosphere-asthenosphere boundary is the transition from rigid, plate-like motion to viscous flow of the mantle rock. Since the viscosity of the mantle is associated with thermally activated creep processes it is expected that the viscosity is exponentially temperature-dependent. The base of the lithosphere is an isotherm defining a rigid behavior; this isotherm is approximately 1000°C. If significant frictional heating occurs on the San Andreas fault the isotherms and, therefore, the lithosphere-asthenosphere boundary, may be elevated beneath the fault. We assume that the density of the mantle rock is $\rho = 3.3 \times 10^3$ kg/m^3 and the shear modulus is $G = 7 \times 10^{10}$ N/m^2. The best data on the viscosity of the asthenosphere comes from studies of postglacial rebound. Cathles (1975) has concluded that beneath the lithosphere there is a low viscosity channel with a viscosity $\eta = 4 \times 10^{19}$ N·s/m^2 with a thickness of 75 km and that the remainder of the upper mantle has a viscosity $\eta = 10^{21}$ N·s/m^2. For our viscosity model we will assume that the asthenosphere has a viscosity $\eta = 4 \times 10^{19}$ N·s/m^2. This corresponds to a viscoelastic relaxation time $T = 5.7 \times 10^8$ s($\sim$ 18 years). Substitution of these values into (13) gives $\varepsilon = 3.8 \times 10^{-8}$, a very small number as expected, and with $\varepsilon = 0$ we have

$$\tilde{\tilde{\Delta w}}_L = 0, \quad \tilde{\tilde{\Delta w}}_A = 0 \tag{14}$$

Laplace's equation must be satisfied in both the lithosphere and asthenosphere. The solutions must satisfy the matching conditions (9). In terms of the variables from (10) these become

$$\tilde{y} = -1: \quad \tilde{w}_L = \tilde{w}_A, \quad \frac{\partial^2 \tilde{w}_A}{\partial \tilde{t} \partial \tilde{y}} = \left(\frac{\partial}{\partial \tilde{t}} + 1 \right) \frac{\partial \tilde{w}_L}{\partial \tilde{y}} \tag{15}$$

The solutions of Laplace's equation in the plate $-1 < y < 0$ and in the asthenosphere $y < -1$ can be expressed in terms of the normal derivatives on the common interface. For the lithosphere we write at $y = -1$:

$$\frac{\partial w_L}{\partial y} = q(x,t) \tag{16}$$

and for the asthenosphere

$$y = -1: \quad \frac{\partial w_A}{\partial y} = g(x,t) \tag{17}$$

Henceforth the tildes on the non-dimensional variables will be omitted. The matching condition (15) is then

$$\frac{\partial q}{\partial t} + q = \frac{\partial g}{\partial t} \tag{18}$$

We shall use the notation

$$w_L(x,0,t) = w(x,t) \tag{19}$$

$$w_L(x,-1,t) = w_A(x,-1,t) = w_I(x,t) \tag{20}$$

for the displacements on the free surface and on the lithosphere-asthenosphere interface.

The solution of Laplace's equation for w_A in the asthenosphere in terms of its normal derivative at $y = -1$ is

$$y \leq -1: \quad w_A(x,y,t) = \frac{1}{\pi} \int_0^\infty g(s,t) \ln \left| \frac{s + z + i}{s - z - i} \right| ds \tag{21}$$

where $z = x + iy$. The antisymmetry condition about $x = 0$ is automatically satisfied. We shall solve the problem in terms of the Fourier sine transform

$$w_I^*(\xi,t) = \int_0^\infty (\sin \xi x) \, w_I(x,t) \, dx \tag{22}$$

Since $\int_0^\infty \sin \xi x \ln \left|\frac{s+x}{s-x}\right| dx = (\pi/\xi) \sin \xi s$, the sine transform of (21) with $z = x - i$ yields

$$w_I^* = g^*/\xi \tag{23}$$

where g^* is the sine transform of g with respect to x.

For the elastic lithosphere, the solution of Laplace's equation for the strip $-1 < y < 0$, $x > 0$ giving $\frac{\partial w}{\partial x}$ in terms of $\frac{\partial w}{\partial y}$ on the boundaries is

$$\left.\begin{array}{l} y = 0: \ \dfrac{\partial w_0}{\partial x} \\ y = -1: \ \dfrac{\partial w_I}{\partial x} \end{array}\right\} = -\int_{-1}^{0} \frac{(\sin \pi y)\left(\frac{\partial f}{\partial y}\right) dy}{\cos \pi y \mp \cosh \pi x} - \int_0^\infty \frac{(\sinh \pi s)\, q(s)\, ds}{\cosh \pi s \pm \cosh \pi x}$$

where $f(y)$ is the prescribed displacement at $x = 0$. The boundary value problem is illustrated in Figure 2. Multiplication of these expressions by $\cos \xi x$ and integration with respect to x from 0 to ∞ gives

$$\xi w_0^* = \xi \int_{-1}^{0} e^{\xi y} f(y)\, dy + e^{-\xi} F(\xi, t) - q^*/(\sinh \xi) \tag{25}$$

$$\xi w_I^* = F(\xi, t) - q^*/(\tanh \xi) \tag{26}$$

where

$$F = \frac{\xi}{\sinh \xi} \int_{-1}^{0} \cosh \xi y\, f(y,t)\, dy \tag{27}$$

Since $w_I^* = g^*$, by (23), (26) provides a relation between g^* and q^*. A second is obtained by taking the transform of (18), with the result

$$\left(\frac{d}{dt} + 1\right) q^* = \frac{dg^*}{dt} \tag{28}$$

Combining (23), (26), and (28) yields a single equation for q^*:

$$\frac{1}{a(\xi)} \frac{dq^*}{dt} + q^* = \frac{dF}{dt} \tag{29}$$

where $a(\xi) = e^{-\xi}\sinh \xi$. The solution of (29) is

$$q^* = a\, F(\xi,t) - a^2 \int_{-\infty}^{t} e^{-a(t-\tau)}\, F(\xi,t)\,d\tau \tag{30}$$

The transformed displacement on the surface is then given by (25),, and has the inverse Fourier transform

$$w_0(x,t) = \frac{2}{\pi}\int_{-t}^{0} \frac{x\, f(y,t)\,dy}{x^2 + y^2} + \int_{-\infty}^{t}\int_{-1}^{0} k(x,y,t-\tau)\, f(y,\tau)\,d\tau dy \tag{29}$$

where

$$k(x,y,t) = \frac{2}{\pi}\int_{0}^{\infty} e^{-2\xi - ta(\xi)} \cosh \xi y \sin \xi x\, d\xi$$

The first term represents the instantaneous elasticitiy of the half space $y < 0$, i.e., of the combined lithosphere and asthenosphere, and the second, the relaxation due to viscoelasticity in the asthenosphere. The kernel (x,y,t) can be expanded in a power series in t by termwise integration. When $f(y,t)$ is a function of t independent of depth on the fault, the expressions simplify further.

For the case of a single step-wise displacement at $t = 0$,

$$f(y,t) = \begin{cases} 0 & t < 0 \\ f(y) & t > 0 \end{cases} \tag{30}$$

(30) gives, for $t > 0$

$$q^*(\xi,t) = \xi e^{-\xi - ta(\xi)} \int_{-1}^{0} f(y) \cosh \xi y\, dy \tag{31}$$

For a displacement which is not a function of depth

$$f(y,t) = \begin{cases} 0 & t < 0 \\ & (-1 < y < 0) \\ 1 & t > 0 \end{cases} \tag{32}$$

and (31) reduces to

$$w_0(x,t) = 1 - \frac{2}{\pi}\int_0^\infty e^{-\xi - ta(\xi)}(\sin \xi x)\,\frac{d\xi}{\xi} \tag{33}$$

The instantaneous displacement is obtained by setting t = 0 in (33) with the result

$$w_0\ (x,o+) = 1 - \frac{2}{\pi}\tan^{-1}x$$

in agreement with the leading term of (29). This is precisely the solution for a unit displacement on a crack of unit depth in an elastic half space y < 0.

To evaluate the integral for general values of x and t, a change in the contour of integration gives the alternative form

$$w_0(x,t) = \frac{2}{\pi}\int_0^\infty e^{-(x\eta + t\sin^2\eta)}\sin\phi(\eta,t)\,\frac{d\eta}{\eta} \tag{34}$$

where

$$\phi = \eta + t\sin\eta\cos\eta$$

This expression is suitable for numerical quadrature after the substitution $\eta = \tan\theta$.

Shear stress accumulates on the fault between great earthquakes as a result of the steady motion of the plates at large distances, and is relieved by displacements during the earthquakes. As a model of this process, we look on the current state as being the result of an infinite sequence of equal step displacements at equal time intervals T_s, the time between great earthquakes. If the mean displacement in each such earthquake is d this corresponds to a plate velocity $V_\infty = d/T_s$. The suggested mechanism is indicated schematically in Figure 3. At t = 0+, after slip on the fault the surface stress on the fault is zero, i.e., $\partial w_0/\partial x = 0$. Measuring displacement from the position at t = 0+, there is a steady increase in w_0 with time at all x positions except on the fault x = 0+, and at large values of x the plates are in uniform translation with velocity V_∞.

When the next great earthquake is imminent, at time $t = T_{s-}$, the shear stress on the fault has reached a maximum causing brittle failure on the fault. Between $t = T_{s-}$ and $t = T_{s+}$ slip occurs on the fault and the instantaneous elastic displacement takes place indicated by arrows, but this dies out at large x so there is no sudden change at sufficient distances from the fault.

The overall motion of the plates at inifinity is caused by the underlying mechanism of continental drfit, which could be represented by tensile or compressive forces in the direction of motion or by shear forces applied to the plate boundaries, but the presence of such forces at large distances does not affect the present discussion. We consider an infinite sequence of step displacements

$$-1 < y < 0: \quad f(y,t) = \begin{cases} d/h & 0 < t < T_s \\ 0 & -T_s < t < 0 \\ -nd/h & -(n+1)T_s < t < -nT_s \end{cases} \tag{35}$$

In summing to infinity, the total displacement is unbounded, however, the derivative of (33), summed over the sequence, gives

$$w_0(x,t) = \frac{2d}{\pi h}\int_0^\infty \frac{a(\xi)e^{-\xi - ta(\xi)}}{1 - e^{-T_s a(\xi)}} (\sin \xi x)\, \frac{d\xi}{\xi} \tag{36}$$

giving the displacement from $t = 0+$ to T_{s-} as

$$w_0(x,t) = \frac{2d}{\pi h}\int_0^\infty e^{-\xi}\left(\frac{1 - e^{-t\,a(\xi)}}{1 - e^{-T_s a(\xi)}}\right) (\sin \xi x)\, \frac{d\xi}{\xi} \tag{37}$$

At $t = T_{s-}$, this gives

$$w_0 = \frac{2d}{\pi h}\int_0^\infty e^{-\xi}(\sin \xi x)\, \frac{d\xi}{\xi} = \frac{d}{h} - \frac{2d}{\pi h}(\tan^{-1}\frac{1}{x}) \tag{38}$$

The second term represents the lagging displacement due to adhesion on the fault, which is made up by the instantaneous elasticity of the half space when slip takes place at $t = T_s$, so that

$$w_0(x,T_{s+}) - w_0(x,T_{s-}) = \frac{2d}{\pi h}(\tan^{-1}\frac{1}{x}) \tag{39}$$

The integral (37) can be converted by a change of contour to the form

$$w_0(x,t) = \frac{td}{T_s h} + \frac{2d}{\pi h}\int_0^\infty e^{-x\eta}\frac{A(\eta,t)}{B(\eta,t)}\frac{d\eta}{\eta} \tag{40}$$

where

$$A(\eta,t) = \sin\phi\, e^{(T_s-t)\sin^2\eta} - \sin\eta\, e^{T_s\sin^2\eta} + \sin(\eta-\psi) - \sin(\phi-\psi)e^{-t\sin^2\eta}$$

$$B(\eta,t) = 2\cosh(T_s\sin^2\eta) - 2\cos\psi$$

$$\phi = \eta + \tfrac{1}{2}\sin 2\eta, \quad \psi = \tfrac{1}{2}T_s\sin 2\eta$$

This form is suitable for numerical evaluation, after the transformation $\eta = \tan\theta$.

Computed values of the incremental displacement $w_0(x,t) - w_0(x,0-)$ as a function of time between offsets for fixed values of x are shown for the three values $T_s = 1, 5, 10$ in Figures 4a, b, c. As $x \to \infty$, (40) shows that

$$hw_0(x,t) \to \frac{td}{T_s} = V_\infty t$$

and the displacement increases linearly with time at the constant plate velocity V_∞. If the offsets occur more frequently, i.e., for small T_s, the approach to a uniform velocity is somewhat more rapid.

It is seen from these results that periodic displacements associated with the fault are damped out in a distance away from the fault which is of the order of the plate thickness. Instantaneous displacements due to faulting are identical to those obtained from a dislocation model.

Finite Element Studies

Although analytical studies can provide important insight into deformations associated with the San Andreas fault, they are limited to relatively simple boundary value problems. In order to study more complex boundary value problems, numerical methods are required. For problems involving elasticity the finite element technique is usually used. The essential feature of finite elements is

the ability to divide a region into separate elements. Material properties and boundary conditions are conveniently prescribed as functions of position by assigning different values to each element.

Variational calculus provides a theoretical basis for the finite element technique. Conventionally, a partial differential equation will govern the unknown distribution in a region. Alternatively, the integral of some function of the unknown (a functional) may be considered to be minimized over the region.

The deformation of an elastic region is governed by the partial differential equilibrium equation.

$$\partial\sigma_{ij}/\partial x_j + \bar{X}_i = 0 \tag{41}$$

where $\bar{X}_i$ are body forces and traction forces which are considered separately as boundary conditions.

The equivalent variational problem is the Principle of Minimum Potential Energy

$$\delta P = \delta U + \delta W = 0 \tag{42}$$

where P is the total potential energy; U, the strain energy, and W, the potential energy of any applied loads. In terms of integration over the region, (42) becomes

$$\delta P = 0 = \overset{<a>}{\iiint_v dU} - \overset{<b>}{\iiint_v \bar{X} \cdot \bar{u} dV} - \overset{<c>}{\iint_s \bar{T} \cdot \bar{u} dS} \tag{43}$$

Term <a> represents the change in strain energy (U) over the volume (V) of the region. Terms <b> and <c> account for work done on the region by body forces ($\bar{X}$) and surface traction forces ($\bar{T}$) respectively; $\bar{u}$ is the unknown displacement field.

For the case of a linear elastic region,

$$dU = \tfrac{1}{2}\, \varepsilon_{ij}\, \sigma_{ji}\, dV = \tfrac{1}{2}\, \varepsilon_{ij}\, C_{jikl}\, \varepsilon_{ji} \tag{44}$$

(43) becomes in matrix notation

$$\tfrac{1}{2}\iiint_V [\varepsilon]^T[C][\varepsilon]dV = \iiint_V [\bar{X}][\bar{u}]^T dV + \iint_S [\bar{T}]dV + \int_S [\bar{T}][\bar{u}]^T dS \qquad (45)$$

Assuming infinitesimal strain,

$$\varepsilon_{ij} = \tfrac{1}{2}\left[\partial u_j/\partial x_i + \partial u_i/\partial x_j\right] \qquad (46)$$

(45) now determines $\bar{u}$ for each element, given the loading forces. The solutions will be described by individual element stiffness matrix equations.

$[K][S] = [F]$

[K] is the stiffness matrix; [S] and [F] are the nodal displacement and force vectors, respectively.

Elements are constructed from nodal points, at which forces and displacements may be either prescribed or determined. It is the number of nodal points within an element which determines the degree of the solution, $\bar{u}$, found for the elements. A linear solution for $\bar{u}$ (i.e., $u = ax + b$) within each element is allowed by two end nodes. Additional, internal nodes permit quadratic and cubic solutions ($u = ax^2 + bx + c$; $u = ax^3 + bx^2 + cx + d$). The advantage of extra, internal nodes is an internally less stiff element. This allows the numerical solution within each element to more closely model the actual response.

Unfortunately, solution time of computers increases drastically with the addition of extra nodes. For this particular study, three-dimensional elements with no internal nodes are employed. Displacements within the elements are linear (i.e., constant shear within each element).

A solution for the combined region of elements is attained by construction of the global stiffness matrix. The stiffness equations for the individual elements are assembled to form the global stiffness matrix. The numerical solution becomes one of a system of simultaneous equations. Displacements at the nodes are solved

for and stresses computed from these displacements. The significant features of this model include:

1) It is an approximate, discretized model of a region.
2) A linear system of simultaneous equations are solved numerically to within computer accuracy.
3) There is the ability to vary material properties from element to element.
4) There is the ability to specify nodal forces and nodal displacements for elements to model boundary conditions.
5) Greater accuracy of the model is attained by a finer element mesh. However, the cost may be prohibitive.

The particular finite element program used in this study is the Structural Analysis Program, Version IV (SAPIV). SAPIV is a very general, structural finite element program developed and refined in later versions at the University of California at Berkeley. The program consists of approximately 15,000 Fortran IV statements.

We initially consider the problem of strain accumulation on the San Andreas fault between great earthquakes. The problem is idealized but not as idealized as that considered by Turcotte and Spence (1975). The finite element mesh employed is given in Figure 5. Ninety-five elements in a slab of one element thickness are constructed from two hundred and thirty nodes. A concentration of elements is placed near the stick-slip region, ac, where stress and strain gradients are largest and of most interest.

The slab represents a vertical cross-section perpendicular to the fault, extending from the fault out into the Pacific plate. The solution is considered asymmetric across the fault plane. The slab is assigned the dimensions of a square cross-section of the lithosphere 50 km by 50 km.

The upper and lower boundaries are assumed to have zero stress (unrestrained). The visco-elastic interaction with the asthenosphere is neglected in this study of fault behavior. Uniform displacements are applied on the edge of the slab away from the fault. As initial conditions the stress conditions applicable just after

a great earthquake (and associated aftershock sequence) are applied on the fault. We assume that the upper part of the fault (to depth $\bar{a} = 0.5$, nondimensionalized with the thickness of the plate) has a stress increasing linearly with depth (constant coefficient of friction on the fault). We also assume that the lower part of the fault has a stress that decreases linearly to the base of the lithosphere. This decrease is associated with the decrease in yield stress with increasing temperature. These initial conditions are the case $\bar{a} = 0.5$ given in Figure 6. The stress at $\bar{a} = 0.5$ is used to define a nondimensional stress.

These initial conditions are obtained as follows (see Figure 5):

1. Nodes on the edge (de) of the slab away from the fault plane are held fixed ($u_z = 0$) using boundary elements.
2. The stress regime of Figure 6 ($\bar{a} = 0.5$) is input on the fault plane (bc) as nodal forces. Forces are initially estimated as σ_{zx}. Area, where σ_{zx} (X) is the piecewise linear function of stress on the fault plane. As there are only a finite number of nodes on the fault plane, a stress average over fault surface areas between the nodes must be calculated to form lumped loads at the nodes.
3. These approximate forces are imput for an initial run. The program calculates resulting stresses in the slab and, in particular, on the fault surface. Differences between calculated stresses and the specified stresses on the fault plane are measured to adjust force input at the fault surface nodes.
4. Once the desired stress profile on the fault is reached, resulting surface displacement and stress profiles (along ce) are noted. Displacements of nodes in the stick-slip region of the fault surface (ac) will be required for the following strain accumulation sequence.

During strain accumulation uniform displacements are applied on the side of the slab away from the fault. Initially, the upper part of the fault $\bar{y} < 0.5$ is assumed to be locked and the lower part of the fault to have the stress initially applied. Without any further modification of the depth of the locked zone a stress

singularity would develop at the lower end of the locked zone, $\bar{y} = 0.5$. In order to eliminate this stress singularity, we allow the lower to migrate upwards. The lower tip of the locked zone is allowed to migrate upwards during each step of strain accumulation until the stress on the free zone ($\bar{y} > \bar{a}$) decreases linearly with depth. This stress accumulation sequence is carried out as follows:

1. Nodes on the stick-slip portion of the fault (ac) are held fixed with the initial displacements. Boundary elements with very high stiffnesses are used for this purpose.
2. The stress profile on the plastic flow region of the fault (ab) is maintained by nodal force loads as before.
3. Edge de is increasingly displaced in the z direction by repetitive use of boundary elements. Displacement on this edge is uniform at all times.
4. Stress on the locked, stick-slip section (ac) increases as displacement at edge de is increased. At some displacement, u_z, on edge de, nodes directly above the crack tip will experience stresses equivalent to plastic yield as determined by the extrapolated plastic flow stress profile. Boundary elements at these nodes are then removed and appropriate nodal forces applied to produce the extrapolated flow stress.
6. Step 4 is repeated until the crack tip a is within two element distances from the surface.

The resulting stress distributions on the fault zone are given in Figure 6. The results are given for equal increments in the depth of the locked section. The corresponding values of the surface ($\bar{y} = 0$) stress are given in Figure 7. The surface displacements for two episodes of strain accumulation are given in Figure 8.

CONCLUSIONS

Both analytical and numerical studies of the stress and strain fields associated with a fault model provide important insight into fault behavior. The analytical study reported here shows the importance of the interaction between the elastic

lithosphere and the viscoelastic asthenosphere in damping the periodic fault motion. The study also shows the critical nature of the interaction of the fault with the asthenosphere. In the model this interaction results in a stress singularity. This singularity may have important implications for fault behavior and must certainly be carefully treated in modelling studies.

In terms of detailed modelling of the San Andreas fault numerical solutions will certaily be required. For problems involving elasticity, plasticity, and fluid flow the finite element technique is favored. The results of a finite element calculation concerning the depth dependent rheology of the fault is reported here. It is concluded that the depth of the locked portion of the fault is likely to decrease as strain accumulates.

Clearly present models are quite idealized. However, even these idealized models point out important aspects of the problem of strain accumulation on major strike-slip faults. Future studies should be expanded to include:

1. More realistic rheologies. The actual rheology of the fault problem is likely to be a complex combination of elastic deformation, plastic yielding, and time-dependent viscous flow.
2. More realistic geometries. Fault curvature may be an essential feature of the behavior of the San Andreas fault. The role of secondary faults may also be essential.

Nevertheless modelling studies must play an important role in understanding the San Andreas fault. Only when a successful model has been produced will we really be able to say that we understand the behavior of the fault.

ACKNOWLEDGEMENTS

This research has been supported in part by the U.S. Geological Survey, Department of the Interior, under USGS Grant No, 14-08-001-G-385.

REFERENCES

Anderson, D.L., 1975, Accelerated plate tectonics, Science 187, 1077-1079.

Bott, M.H.P, and Dean, D.S., 1973, Stress diffusion from plate boundaries, Nature 243, 339-341.

Cathles, L.M., 1975, The Viscosity of the Earth's Mantle, Princeton University Press, Princeton, N.J., 386 pp.

Elsasser, W.M., 1969, Convection and stress propagation in the upper mantle, in The Application of Modern Physics to the Earth and Planetary Interiors, ed. S.K. Runcorn, pp. 223-246, Wiley-Intersciences, London.

Melosh, H.J., 1976, Nonlinear stress propagation in the Earth's upper mantle, J. Geophys. Res. 81, 5621-5632.

Minster, J.B., Jordan, T.H., Molnar, P., and Haines, E., 1974, Numerical modelling of Instantaneous plate tectonics, Geophys. J. Roy. Astro. Soc. 36, 541-576.

Nur, A., andMavko, G., 1974, Postseismic viscoelastic rebound, Science 183, 204-206.

Savage, J.C., 1975, Comment on "An analysis of strain accumulation on a strike-slipe fault" by D.L. Turcotte and D.A. Spence, J. Geophys. Res. 80, 4111-4114.

Spence, D.A., and Turcotte , D.L., 1976, An elastostatic model of stress accumulation on the San Andreas fault, Proc. Roy. Soc. London A349, 319-341.

Spence, D.A., and Turcotte, D.L., 1977, Viscoelastic damping of cyclic displacements of the San Andreas fault, submitted for publication, Proc. Roy. Soc. London.

Turcotte, D.L., 1977, Stress accumulation and release on the San Andreas fault, Pure. Ap. Geophys. 115, 413-427.

Turcotte, D.L., and Spence. D.A., 1974, An analysis of strain accumulation on a strike-slip fault, J. Geophys. Res. 79, 4407-4412.

FIGURE CAPTIONS

Figure 1. Illustration of the model. Two elastic plates with a common edge in the plane x = 0 overlie a viscoelastic half space.

Figure 2. The two dimensional boundary value problem.

Figure 3. Illustration of the surface deformation when the surface plates have constant velocities at large distance from the fault and are subjected to step displacements at the fault. The shaded regions represent surface displacement due to elastic deformation of the plates and asthenosphere.

Figure 4. Solutions for an infinite series of step displacements. The surface displacement is given as a function of the time between steps at various distances from the fault. Three ratios of the time interval between the steps to the viscoelastic relaxation time are considered. (a) $T_s = 1$, (b) $T_s = 5$, (c) $T_s = 10$.

Figure 5. The finite element mesh used in this study.

Figure 6. Shear stress profiles on fault plane.

Figure 7. Surface displacement profiles. Nondimensional displacement, $\bar{u}(u_z G/\sigma_o L)$ is plotted versus nondimensional distance from fault, $\bar{X}(\bar{X}/L)$. Profiles are presented for various crack tip depths, $\bar{a}$ (a/L). The difference between the solid and dashed sequence characterizes displacements during earthquakes, $\Delta\bar{u}$ ($\Delta uG/\sigma_o L$).

Figure 8. Surface stress profiles. Nondimensional shear stress, $\bar{\sigma}(\sigma_{zy}/\sigma_o)$ is plotted versus nondimensional distance from fault, $\bar{X}(X/L)$. Profiles are presented for various crack tip depths, $\bar{a}(a/L)$.

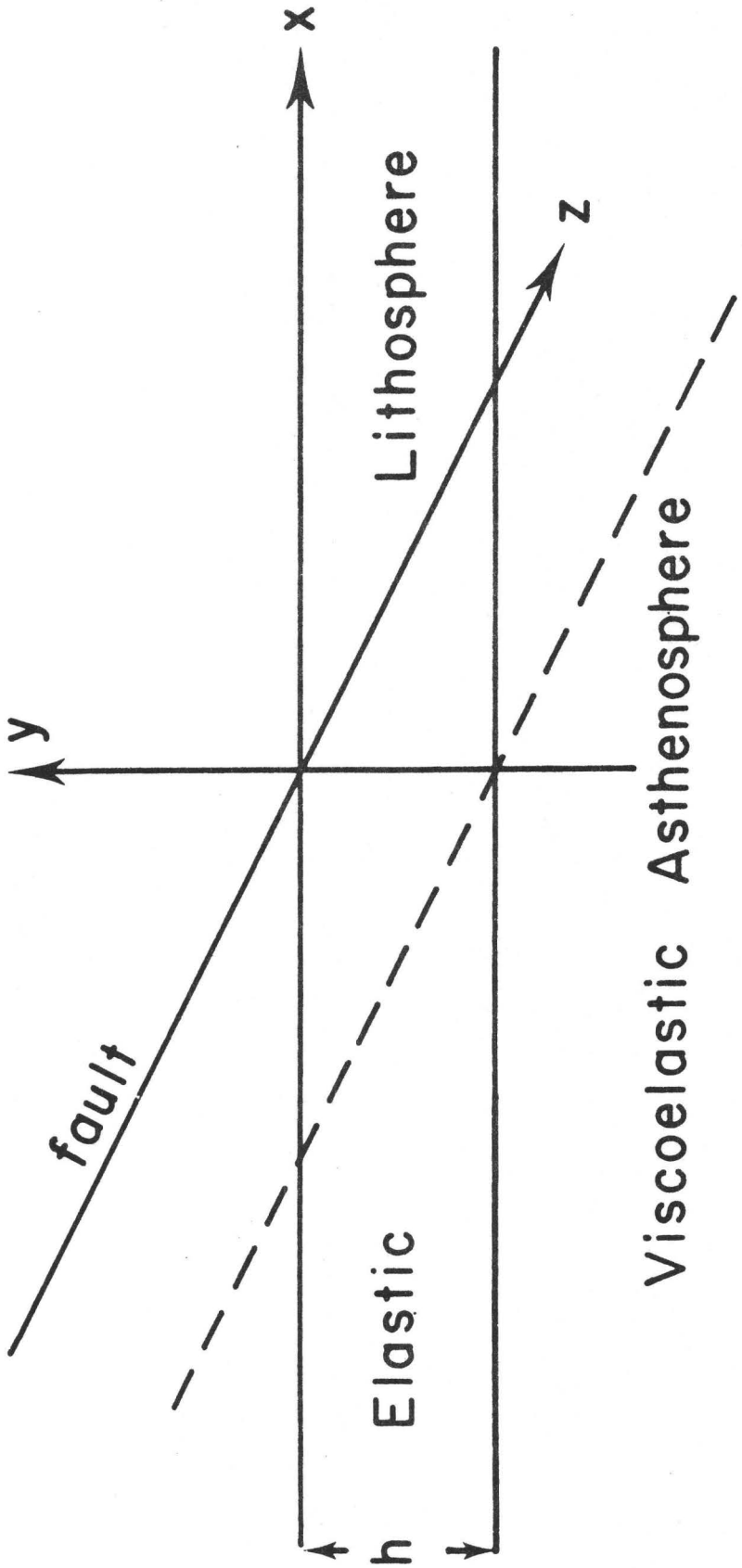

x
y
z
Lithosphere
fault
h
Elastic
Viscoelastic
Asthenosphere

Figure 1

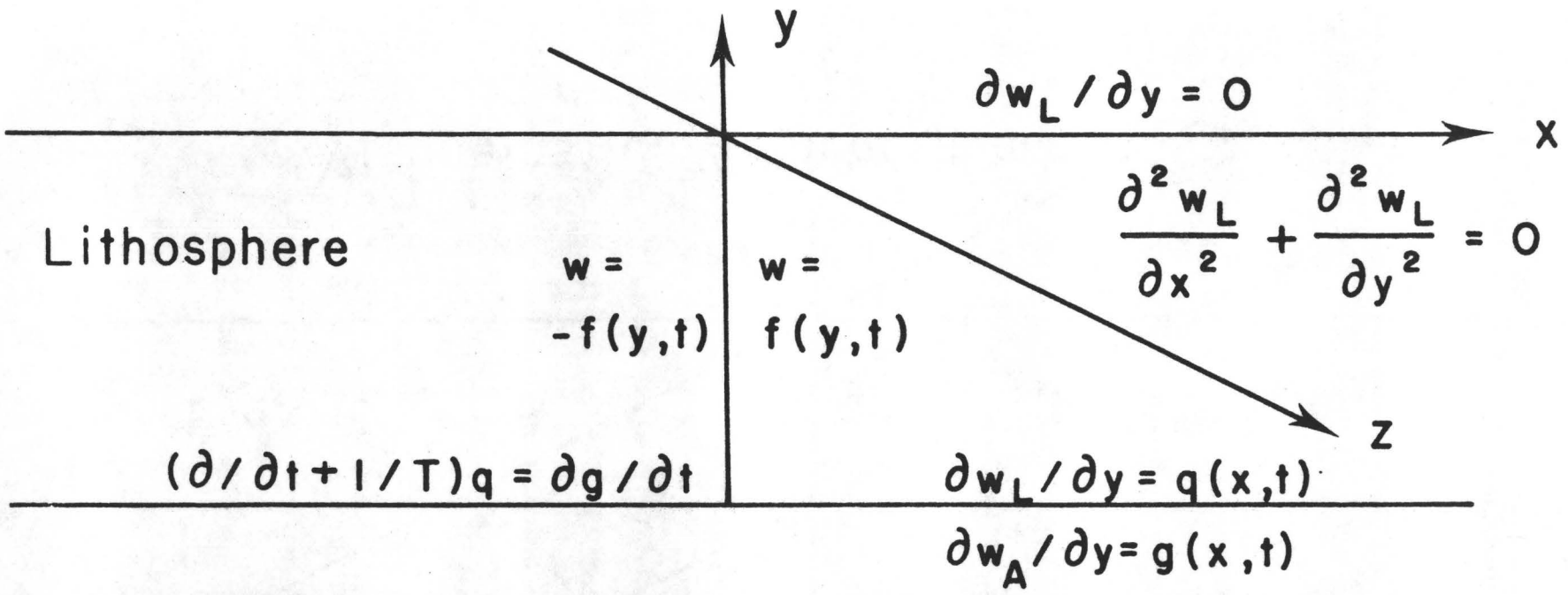

y
x
z
$\partial w_L / \partial y = 0$
$\frac{\partial^2 w_L}{\partial x^2} + \frac{\partial^2 w_L}{\partial y^2} = 0$
Lithosphere
w =
$-f(y,t)$
w =
$f(y,t)$
$(\partial / \partial t + 1/T)q = \partial g / \partial t$
$\partial w_L / \partial y = q(x,t)$
$\partial w_A / \partial y = g(x,t)$
Asthenosphere
$\frac{\partial^2 w_A}{\partial x^2} + \frac{\partial^2 w_A}{\partial y^2} = 0$

Figure 2

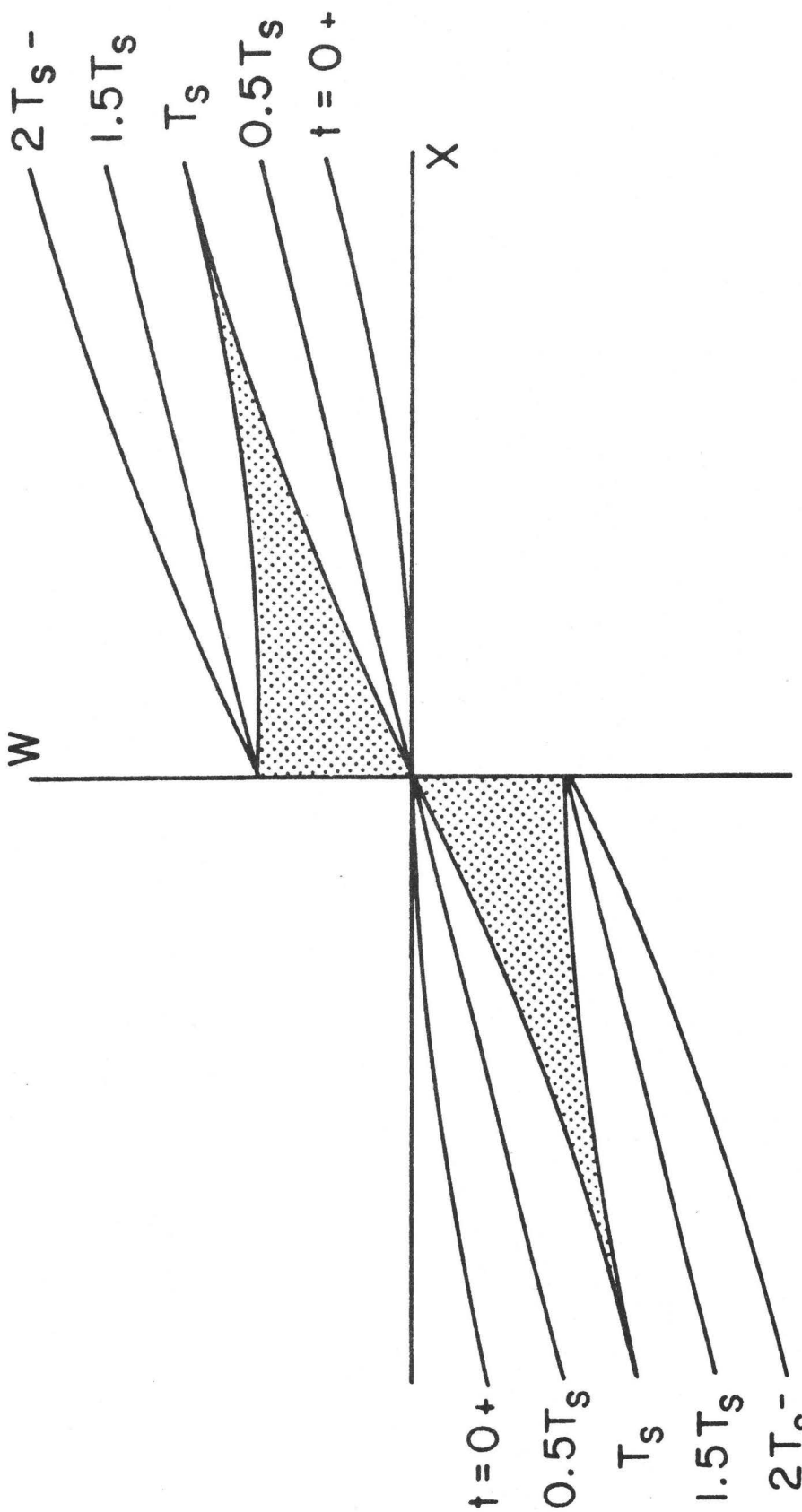


Figure 3

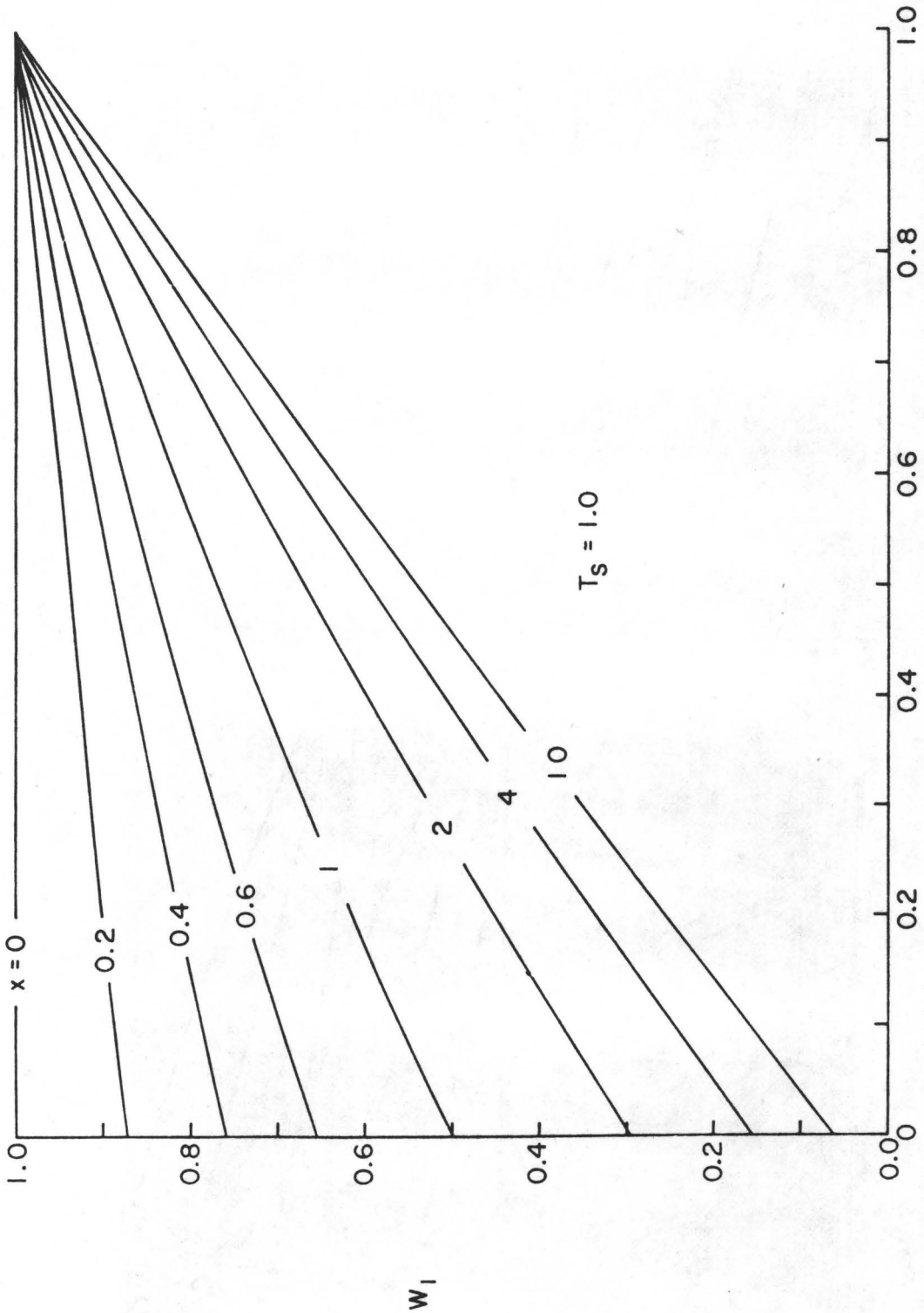


Figure 4a

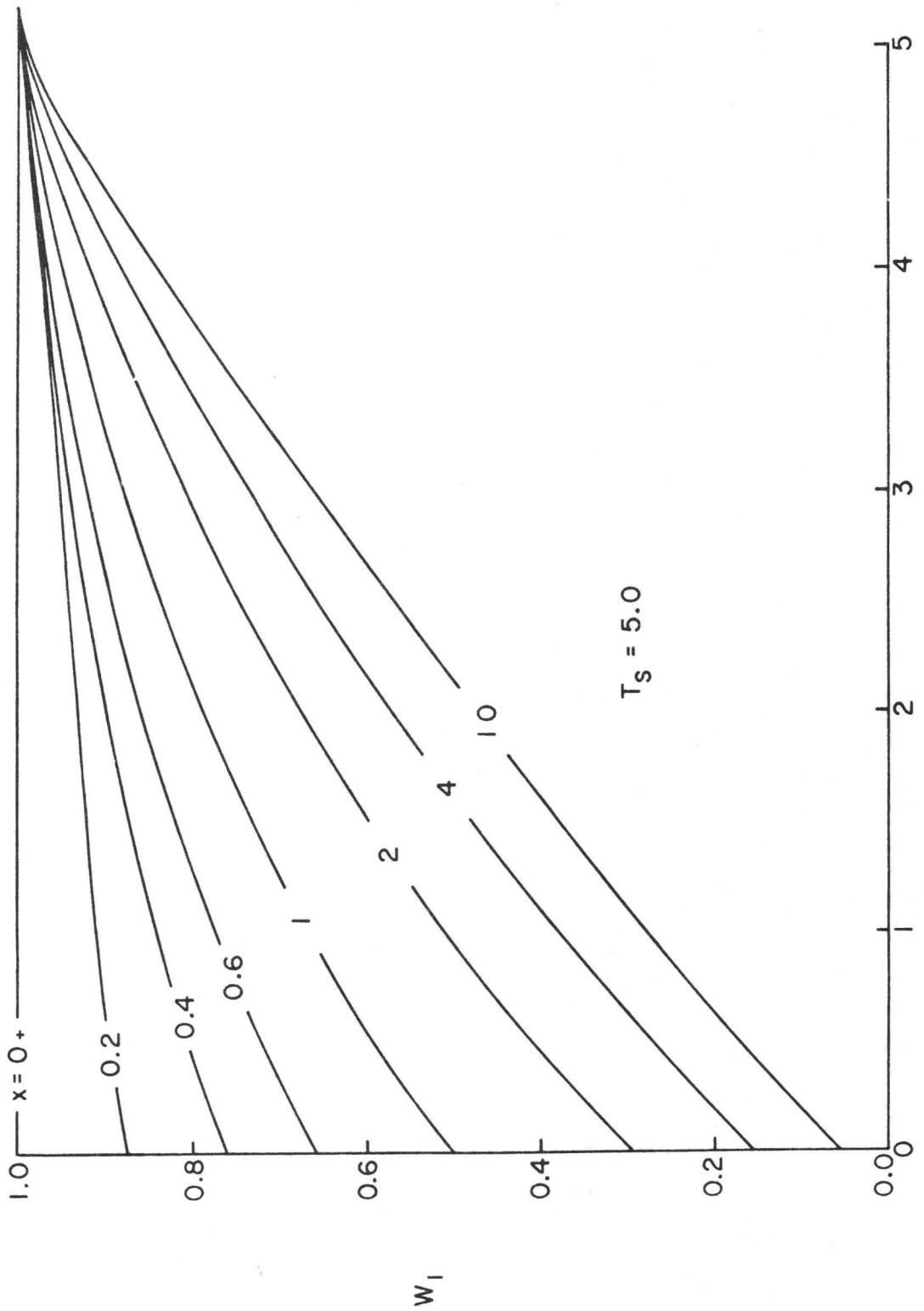


Figure 4b

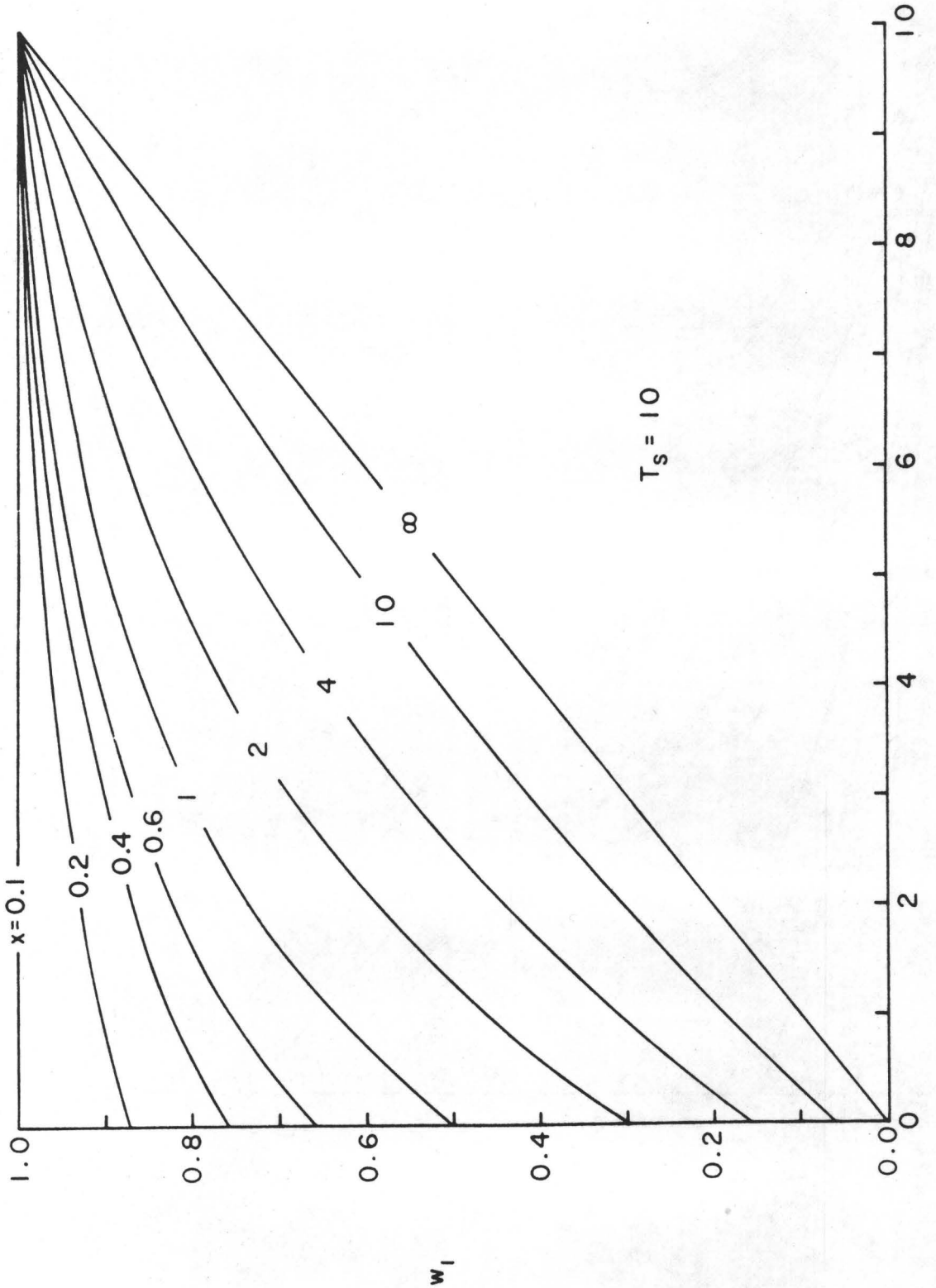


Figure 4c

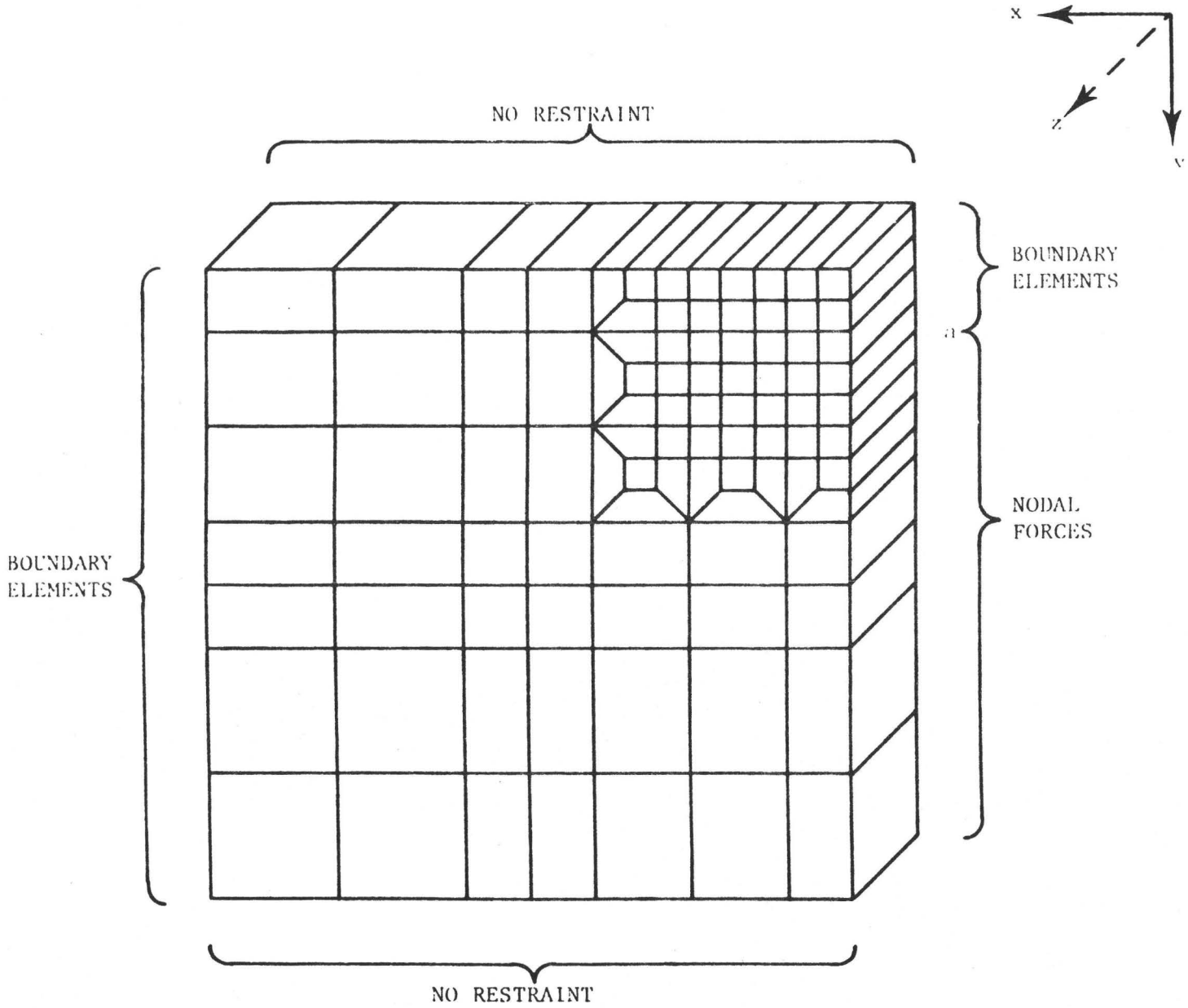

X
Z
Y
NO RESTRAINT
BOUNDARY ELEMENTS
BOUNDARY ELEMENTS
NODAL FORCES
BOUNDARY ELEMENTS
NO RESTRAINT

Figure 5

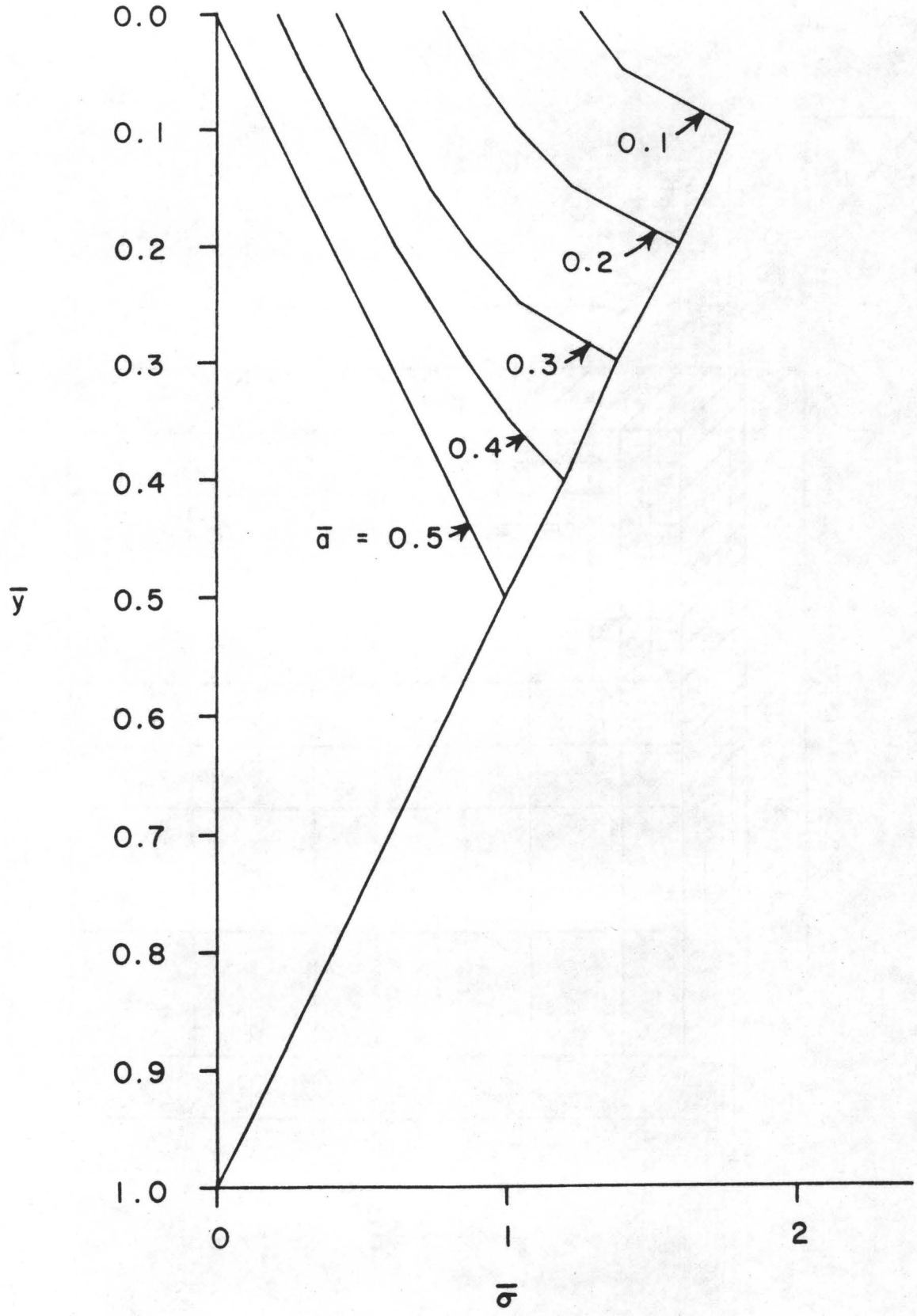


Figure 6

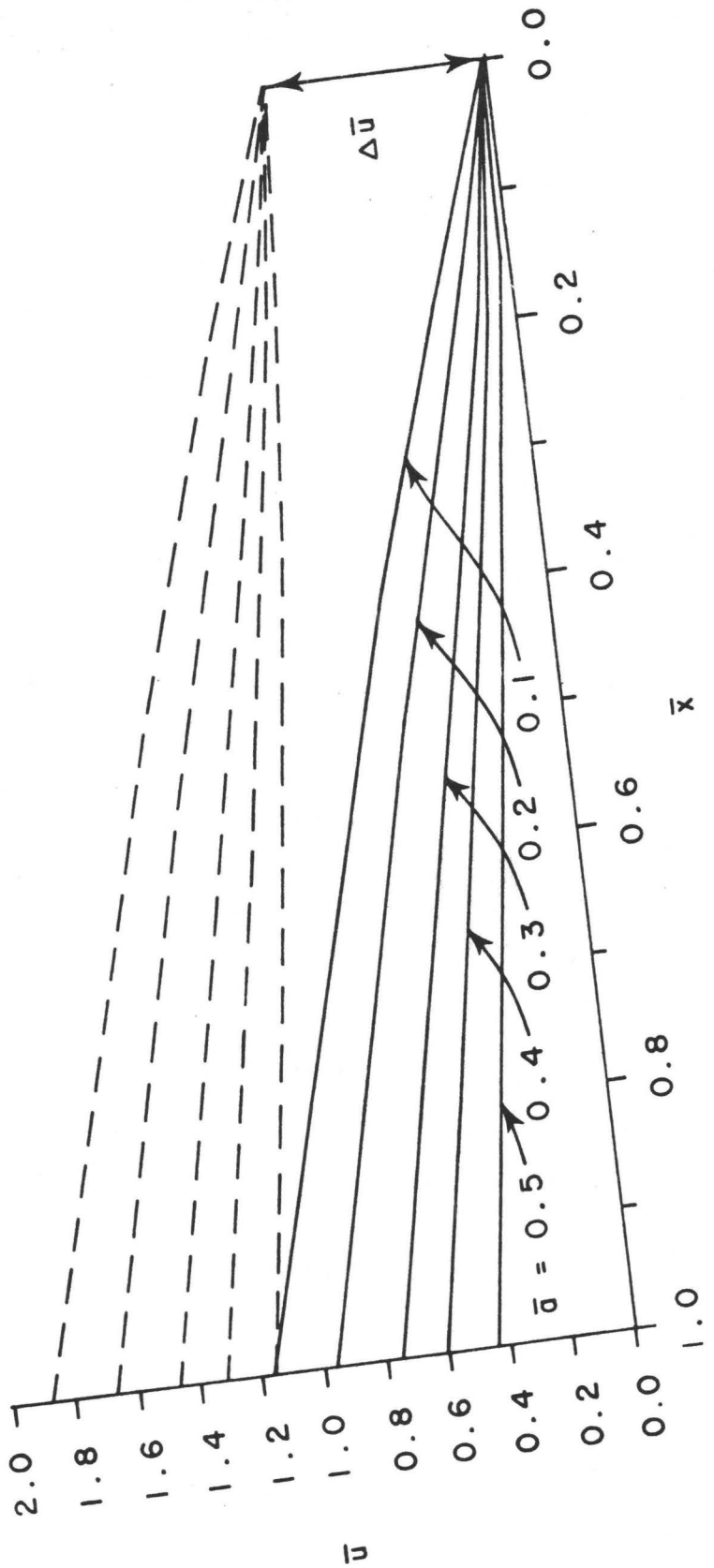

$\Delta\bar{u}$
0.0
0.2
0.4
0.6
0.8
1.0
$\bar{x}$
0.1
0.2
0.3
0.4
$\bar{d}$ = 0.5
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
$\bar{u}$

Figure 7

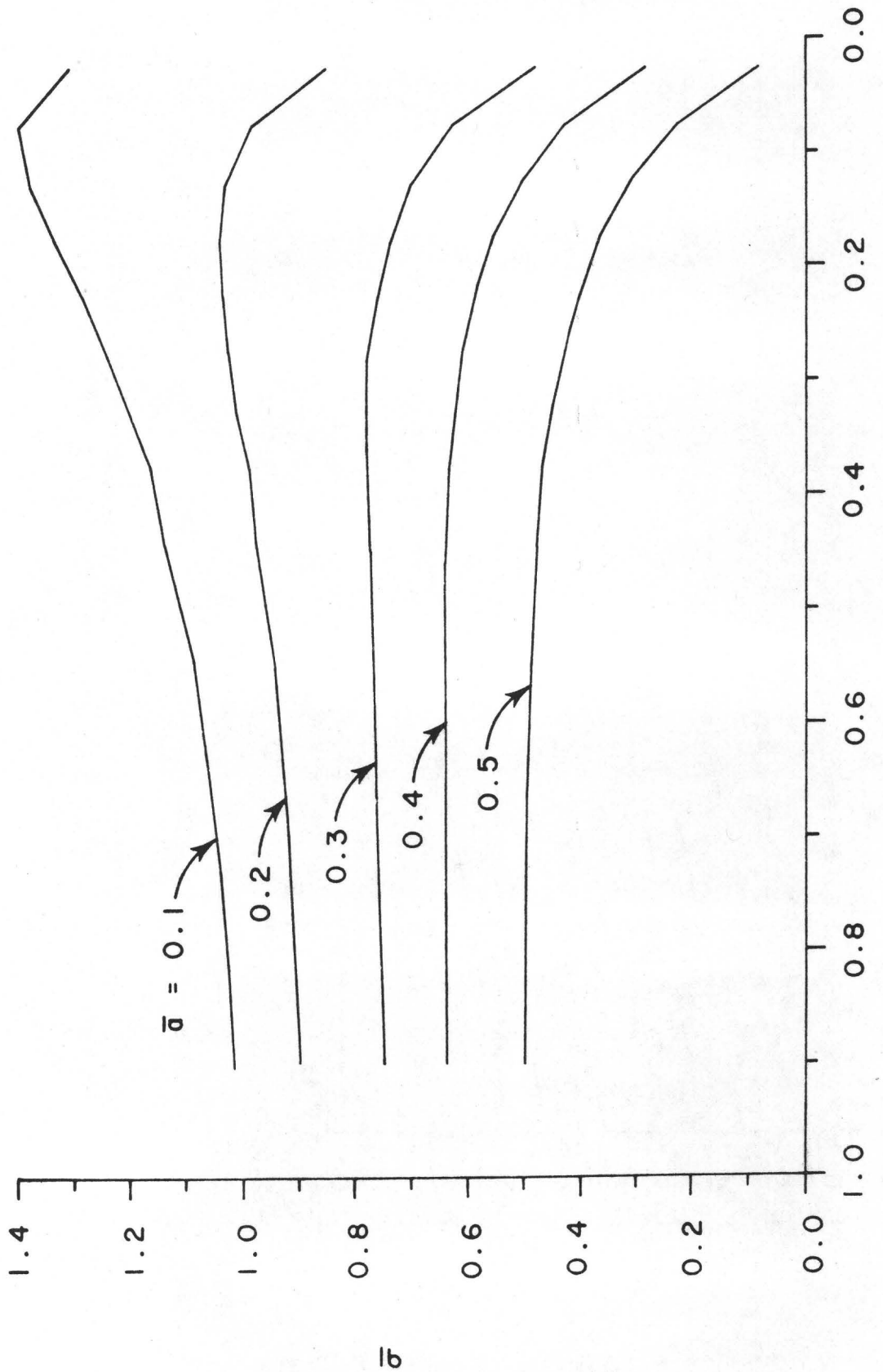

1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
1.0
0.8
0.6
0.4
0.2
0.0
$\bar{q}$
$\bar{x}$
$\bar{a}$ = 0.1
0.2
0.3
0.4
0.5

Figure 8

Inherent Instability of Quasistatic Creep Slippage on a Fault

J. Weertman
Department of Materials Science and Engineering
and Department of Geological Sciences,
Northwestern University
Evanston, Illinois 60201

ABSTRACT

The inherent instability of quasistatic creep slippage is studied for two examples. One example is that of a symmetric transform fault at which a single spreading center terminates. The other example is the transform fault between two spreading centers. It is assumed that the friction stress of a fault decreases with increasing slip velocity.

INTRODUCTION

In earlier papers (Weertman, 1967, 1971a) an analysis was developed with which it could be shown for one special example that an inherent instability exists during creep slippage on a fault. In this paper this intrinsic instability is studied further.

It is assumed in our analysis that the friction stress on a fault decreases in value as a result of creep slippage. It appears rather certain that the friction stress at a given place on a fault at a given time cannot have a unique value if earthquakes are to occur on it. If the friction stress has a unique value that cannot be altered by slip the energy that is available to produce motion across a fault must all be dissipated in the immediate vicinity of the fault in frictional losses. No (elastic) energy can be released into the formation of seismic waves. Figure 1 illustrates in a schematic way why this claim is true for the two dimensional

case.

Figure 1a shows schematically (as a solid line) the friction stress σ_f as a function of distance x along a fault. (The frictional stress σ_f may be a function of position because of variations of rock type along a fault or because, at a given instant in time, the pressure of fluid within the fault is a function of position).

Figure 1a shows schematically (as a dashed line) the internal stress σ_i that exists on a fault because of prior non-uniform slip across the fault. Let $D_i(x)$ represent all the prior displacement that has occurred at position x across the fault. Let the derivative of $D_i(x)$ be represented by $B_i(x)$ where

$$B_i(x) = - \partial D_i(x)/\partial x \tag{1}$$

The internal stress $\sigma_i(x)$ that is produced by prior non-uniform slip is given by

$$\sigma_i(x) = \frac{\mu}{2\pi\alpha} \int_{-\infty}^{\infty} \frac{B_i(x')}{x-x'} \, dx' \tag{2}$$

where μ is the shear modulus and under plane strain conditions the constant $\alpha = 1 - \nu$ where ν is Poisson's ratio if the displacement $D_i(x)$ is in the x direction and $\alpha = 1$ if the displacement $D_i(x)$ is in the z direction where the z direction is perpendicular to the x direction and is in the plane of the fault. (Under plane stress conditions $\alpha = 1/(1+\nu)$.) As is well known the displacement derivative $B_i(x)$ can be considered to be equal to the density of a collection of infinitesimal dislocations. Figure 1b shows schematically these infinitesimal dislocations that prior non-uniform slip has left on the fault. Let $S_i(x)$ represent the additional, "applied" stress that exists on the fault at a given instant in time that is produced by all

processes that cause the earth to be deformed and which drive the motion of the plates. Figure 1c shows a plot of the stress difference $\sigma_f - (\sigma_i + S_i)$ where σ_f and σ_i are the stresses shown in Fig. 1a. In the regions on the fault where the term $\sigma_f - (\sigma_i + S_i) = 0$ only an infinitesimal increase in the "applied" stress on the fault is required to obtain slippage.

The processes that cause the earth to be deformed will produce an additional stress S to act across the fault as time goes by. This additional stress will produce further creep slippage on the fault starting in the regions where $\sigma_f - (\sigma_i + S_i) = 0$. The new slippage in general is non-uniform and an additional internal stress $\sigma(x)$ will be produced on the fault and an additional infinitesimal dislocation distribution B(x) will be added to the fault. (The stress $\sigma(x)$ and the dislocation distribution B(x) are related to each other through equation (2) if in that equation the stress $\sigma_i(x)$ is replaced with the stress $\sigma(x)$ and the term $B_i(x)$ is replaced with the term B(x).) Eventually, everywhere on the fault the sum of the stresses $S_i + S + \sigma_i + \sigma$ will equal the frictional stress σ_f. Thus

$$S_i + S + \sigma_i + \sigma = \sigma_f(x) \qquad (3)$$

Once equation (3) is satisfied creep slippage will occur everywhere along the fault. If the fault can be considered to be essentially infinitely long both the value of the stress S and the new dislocation stress σ will no longer change with time. Subsequent creep displacement will be constant everywhere across the fault. Only if the frictional stress σ_f can fall in value would it be possible for the rapid slippage that can generate seismic waves take place on the fault.

It should be pointed out that there is a formal solution for rapid propagation across a fault in an infinite medium when the fault is in the condition for which equation (3) is satisfied. Suppose, as is shown in Fig. 2, a localized dislocation distribution propagates down a fault at a uniform velocity $V = c_r$ where c_r is the Rayleigh wave velocity if the dislocations are of edge character and $V = c$ where c is the shear wave velocity if the dislocations are of screw character. These moving dislocations produce no stress across their fault plane (Weertman and Weertman, in press). Because they are in uniform motion their total elastic and kinetic energy is a constant and no additional work is required to keep them in motion. Thus once created a localized group of infinitesimal dislocations can move rapidly down the fault, at least until anelastic damping losses within the rock masses on either side of a fault stop them. However, the formation of a localized group of fast moving dislocations requires energy. They cannot be formed unless stored energy is released somewhere along the fault when the frictional stress σ_f drops suddenly in value. In a finite medium, say of dimension L, the passage of a group of infinitesimal dislocations of total Burgers vector b*, (where $b^* = \int B(x)dx$) also will cause the stress S to drop by an amount approximately equal to $\mu b^*/L$ behind the dislocations.

It is to be concluded that if the frictional stress σ_f at a given place on a fault has a unique value and if the stress applied to the rock some distance from the fault is increased monotonically that the patches of localized creep slippage on the fault will become larger until all these patches connect up with each other. Thereafter uniform slip takes place everywhere across the fault. The applied stress can no longer increase in value. This is true

even if the frictional stress varies in value in the horizontal direction as well as in the vertical direction. Slow changes in the value of σ_f that are produced by slow water flow around a fault will not change this conclusion. However, if the frictional stress can plunge downwards in value such a fault is "primed" for catastrophic fast slippage to take place on it.

But if the value of frictional stress σ_f can drop suddenly in value instabilities on the fault will occur even before it is "primed" everywhere along its length. One example is given in Weertman (1971a). Consider now other examples.

Spreading Center that Terminates at a Symmetric Transform Fault

Consider a hypothetical two dimensional spreading center that is shown in Fig. 3 that terminates at a symmetric transform fault. Because of the continuous emplacement of dikes at the spreading center a uniform velocity D_0 is forced upon the material on either side of the spreading center. The spreading center is considered to terminate at the origin ($x = 0$) and it continues to infinity in the upper half plane of Fig. 3. A symmetric transform fault exists on either side of the spreading center termination point. The emplaced dikes are imagined to give rise to normal stresses on either side of the $x = 0$ plane that are just sufficiently large enough to force material at $x = 0$ in the upper half plane to move with a constant velocity in the x direction with a velocity $\dot{D}_0$ to the right and a velocity $\dot{D}_0$ to the left as indicated in the figure. No shear stress is applied at large distances that would produce a shear stress across the transform fault.

(It should be emphasized that dike formation requires an overall tensile stress to act parallel to the x-axis of Fig. 3 (and Figs. 10 and 12). If a compressive stress were to act the magma of a dike, before it solidified, would be forced out of the crack that is a dike onto the earth's surface (Weertman, 1971b). The magma, while it solidifies, can "push" rock on either side of a spreading apart only if a tensile stress is already present. Otherwise the rock on either side of the magma will "squeeze" the magma out and no spreading will occur. In Fig. 3 (and Figs. 10 and 12) it is imagined that at any one time almost all the material in the spreading center is solidified rock and that dikes that contain molten magma at a given instant in time are only of limited number and have limited lengths. Hence the spreading center can be considered, for all practical purposes, to be solid everywhere.)

Assume that the friction stress obeys a law in which the friction stress decreases continuously with increasing displacement velocity $\dot{D} = \partial D/\partial t$. Such a friction law is shown schematically in Fig. 4. Next let the continuously decreasing friction curve of Fig. 4 be approximated by a step function curve shown in Fig. 5 in which the friction stress has a constant value equal to σ_f up to a critical displacement velocity $\dot{D}_c$ and at larger displacement velocities the the friction stress drops to the value $\sigma_f = \Delta\sigma$.

Now consider the slippage on the symmetric transform fault from the moment the spreading center is first created. Consider first

the case in which the spreading center forces a displacement velocity at the origin (x = 0) that is equal to $\dot{D}_0$ where $\dot{D}_0$ is smaller than the critical velocity $\dot{D}_c$ of the friction law that is given in Fig. 5.

As slippage propagates from the spreading center the stress $\sigma(x)$ that acts on the fault is equal to the frictional stress σ_f where slip has occurred. (The slippage that is being considered is a slow creep displacement across the transform fault.) At a given instant in time let slip occur out to a distance equal to b on either side of the spreading center. Thus $\sigma(x) = \sigma_f$ out to x = b on the right hand side of the center and $\sigma(x) = -\sigma_f$ on the left hand side of the center out to x = -b, as shown in Fig. 6. (The negative value for $\sigma(x)$ occurs for the left hand side because the slip displacement on the left hand side is in the opposite sense of the slip displacement on the right hand side.)

If slippage takes place on only a finite segment of a fault and if the stress on that segment is known then the dislocation distribution function B(x) is calculated very easily from the following equation of elasticity theory

$$B(x) = -\frac{2\alpha}{\pi\mu}(b^2 - x^2)^{\frac{1}{2}} \int_{-b}^{b} \frac{\sigma(x')dx'}{(x-x')(b^2-x'^2)^{\frac{1}{2}}} \qquad (4)$$

where x = 0 is taken to be the center of the slipped region which extends from x = -b to x = b. (In the regions for which $|x| \geq b$ the density B(x) = 0.) The following subsidiary equation must also be satisfied

$$\int_{-b}^{b} \frac{\sigma(x)dx}{(b^2-x^2)^{\frac{1}{2}}} = 0 \tag{5}$$

If $\sigma(x)$ is an antisymmetric function of x equation (5) always is satisfied.

If the stress function $\sigma(x)$ that is shown in Fig. 6 is substituted into equation (4) the following equation is found for B(x)

$$B(x) = \frac{2\alpha\sigma_f}{\pi\mu} \log \left| \frac{b + (b^2-x^2)^{\frac{1}{2}}}{b - (b^2-x^2)^{\frac{1}{2}}} \right| \tag{6}$$

In Fig. 7 is shown a plot of the dislocation density B(x) that is given by equation (6). The slip displacement D(x), which is equal to $\int_x^b B(x)dx$, is found by integrating equation (6) and it is equal to

$$D(x) = \frac{2\alpha\sigma_f}{\pi\mu} \left\{ \pi b - 2b\sin^{-1}(x/b) - x\log\left| \frac{b + (b^2-x^2)^{\frac{1}{2}}}{b - (b^2-x^2)^{\frac{1}{2}}} \right| \right\} \tag{7}$$

The displacement velocity $\dot{D}(x)$ is found by taking the time derivative of equation (7). The value of $\dot{D}(x)$ is

$$\dot{D}(x) = \pm \frac{2\alpha\dot{b}\sigma_f}{\pi\mu} \left\{ \pi - 2\sin^{-1}|x/b| \right\} \tag{8}$$

where the + sign is used for $x > 0$ and the minus sign for $x < 0$. The term $\dot{b} = db/dt$ is the velocity with which the outer edges of the slipped zone propagate away from the spreading center. At x = 0 the displacement velocity $\dot{D}(0) = \pm \dot{D}_0$. Hence $\dot{b}$ is given by

$$\dot{b} = \mu\dot{D}_0/2\alpha\sigma_f \tag{9}$$

and the term $\dot{b}$ has a constant value. It should be noted that the magnitude of $\dot{D}(x)$ decreases monotonically from the value $\dot{D}_0$ at $x = 0$ to the value 0 at $x = b$. Hence if $\dot{D}_0$ is smaller than $\dot{D}_c$ then $\dot{D}(x)$ is everywhere smaller than $\dot{D}_c$ in the slipped zone. Thus equations (6), (7), and (8) represent a satisfactory solution for the slippage on a symmetric transform fault at a newly created spreading center. (The solution, however, may be one for a fault in "unstable equilibrium".)

Now suppose the spreading velocity $\dot{D}_0$ is larger than $\dot{D}_c$. Near the spreading center the stress $\sigma(x)$ on the fault need only equal the term $\sigma_f - \Delta\sigma$ in order to have slippage. Far from the fault the slippage velocity should decrease and there the stress $\sigma(x)$ must equal the term σ_f. Hence the stress function $\sigma(x)$ must be that shown in Fig. 8 when creep slippage is "well behaved". The stress $\sigma(x) = \sigma_f - \Delta\sigma$ for $0 < x < a$, where a is a constant with a value $a < b$, and $\sigma(x) = \sigma_f$ for $a < x < b$. Also $\sigma(x) = -\sigma(-x)$.

If the stress function $\sigma(x)$ that is shown in Fig. 8 is substituted into equation (4) the following equation for B(x) is found

$$B(x) = \frac{2\alpha(\sigma_f-\Delta\sigma)}{\pi\mu} \log \left|\frac{b + (b^2-x^2)^{\frac{1}{2}}}{b - (b^2-x^2)^{\frac{1}{2}}}\right| + \frac{2\alpha\Delta\sigma}{\pi\mu}\log\left|\frac{(b^2-a^2)^{\frac{1}{2}}+(b^2-x^2)^{\frac{1}{2}}}{(b^2-a^2)^{\frac{1}{2}} - (b^2-x^2)^{\frac{1}{2}}}\right| \quad (10)$$

Figure 9 shows a plot of B(x) given by equation (10) for the case of $\Delta\sigma/\sigma_f = 0.25$ and $a/b = 0.7$. If equation (10) is integrated in order to obtain D(x) and then the time derivative of D(x) is taken the following equation is found for $\dot{D}(x)$

$$\dot{D}(x) = \pm \frac{2\alpha}{\pi\mu} [\{\pi - 2\sin^{-1}|x/b|\}\{(\sigma_f-\Delta\sigma)\dot{b} + \Delta\sigma\frac{(b\dot{b}-a\dot{a})}{(b^2-a^2)^{\frac{1}{2}}}\} +$$

$$\Delta\sigma\dot{a}\log\left|\frac{|x|(b^2-a^2)^{\frac{1}{2}} + a(b^2-x^2)^{\frac{1}{2}}}{|x|(b^2-a^2)^{\frac{1}{2}} - a(b^2-x^2)^{\frac{1}{2}}}\right| \quad] \tag{11}$$

where again the + sign is used for $x > 0$ and the - sign for $x < 0$. The term $\dot{a}$ is the time derivative of the term a.

Now note: for a finite value of $\dot{a}$ the logarithmic term in equation (11) approaches infinity in value as $x \to a$. It is thus impossible for $\dot{D}(x) < \dot{D}_c$ for values of x that are slightly larger than a in value. Hence equation (10) can be a solution of the equation (4) only if $\dot{a} = 0$. Thus set $\dot{a} = 0$. From the conditions that $\dot{D}(0) = \pm \dot{D}_0$ and $\dot{D}(a) = \dot{D}_c$ the following equation for the value of a is found.

$$a = b\sin\left[\frac{\pi}{2}\left(1 - \frac{\dot{D}_c}{\dot{D}_0}\right)\right] \tag{12}$$

and the following equation for $\dot{b}$ is obtained

$$\dot{b} = (\mu\dot{D}_0/2\alpha)/(\sigma_f + \Delta\sigma + \Delta\sigma/\cos[(\pi/2)(1 - \dot{D}_c/\dot{D}_0)]) \tag{13}$$

Now if $\dot{a} = 0$ equations (12) and (13) can be satisfied only at <u>one</u> instant of time. At any other period of time, because b increases with time, a different value of a is required to satisfy equation (12). Therefore, in general, no "well behaved" solution is possible. By well behaved is meant a solution for a slow creep slippage on the fault for which an essentially static elastic stress analysis can account for the displacement and the stress that acts across the fault. Hence the slippage on the symmetric transform fault, when $\dot{D}(0) > \dot{D}_c$, must take place sporatically in a fast catastrophic fashion between periods of slow creep or no creep.

The reason that a well behaved, slow creep slip solution cannot be maintained indefinitely is the occurance of the large dislocation density

which is shown in Fig. 9, at the boundaries $x = \pm a$ that separate a high friction stress region from a low friction stress region. If the position of this boundary moves during creep slippage the localized packet of dislocations there also move and in so doing produce a large slip velocity. A large slip velocity reduces the friction stress and the slip process runs away, producing an earthquake. Of course, the infinite value of $B(x)$ at $x = \pm a$ is a consequence of the fact that the stress there is required to jump discontinuously in value. However, if the friction stress decreased rapidly but continuously with increasing sliding velocity in the vicinity of $x = \pm a$ the infinite value of the dislocation density $B(x)$ will be rounded off to a finite value. The density $B(x)$ will still have a much larger value than elsewhere (except at $x = 0$) and consequently a quasi-steady slow creep displacement solution still cannot exist. Only if the friction stress drops gradually in value with increasing displacement velocity would a quasi-steady solution be possible. A quantitative definition of what is a rapid and what is a gradual drop in friction stress can be made by solving the equations of this section with different friction laws in the same manner that an analogous problem of deciding what is a rapid or slow friction stress decrease was treated earlier (Weertman, 1967).

Transform Fault Between Two Spreading Centers

The example of the last section is rather artificial. A newly created spreading center that terminates at a symmetric transform fault if one ever existed, must be a very rarely occuring phenomena. Consider next a phenomenon that does exist in abundance - a transform fault between two spreading centers. Such a fault is shown in Fig. 10.

If the distance between two spreading centers does not change with time and the positions of the two spreading centers themselves remains fixed in space then on average all the slip on the transform fault takes place between the two spreading centers. The average slip on the projections of the fault beyond the spreading centers - the fossil or inactive projections of the fault - must be equal to zero. Now suppose slip at a constant velocity can occur on the fault between the spreading centers. The stress $\sigma(x)$ must equal σ_f everywhere on the transform fault for this slippage to occur if the velocity $2\dot{D}_0$, which is twice the spreading rate of each center (see Fig. 10), is smaller than the term $\dot{D}_c$.

In Fig. 11a is shown a plot of $\sigma(x)$ versus distance on a transform fault. The origin is taken to be the midway point between the two spreading centere. The spreading centers are situated at $x = \pm a$. In the region $-a < x < a$ the stress $\sigma(x) = \sigma_f$. Slip in the region $-a < x < a$ requires limited slip in the regions $|x| > a$. If no limited slip occured in this latter region an infinite stress would build by at $x = \pm a$. Hence, as shown in Fig. 11a, it is reasonable to expect that $\sigma(x) = -\sigma_f$ in the regions $a < |x| < b$ where the value of b is determined through equation (5).

The stress $\sigma(x)$ shown in Fig. 11a requires the stationary distribution of dislocations that is plotted in Fig. 11b. This distribution is found from equations (4) and (5) and is equal to

$$B(x) = \frac{4\alpha\sigma_f}{\pi\mu} \log \left| \frac{x(b^2-a^2)^{\frac{1}{2}} + a(b^2-x^2)^{\frac{1}{2}}}{x(b^2-a^2)^{\frac{1}{2}} - a(b^2-x^2)^{\frac{1}{2}}} \right| \qquad (14)$$

and the distance b is equal to $b = 2^{\frac{1}{2}}a$.

If the net sliding velocity $2\dot{D}_0$ is larger than $\dot{D}_c$ between the two spreading centers the friction stress there is reduced to the value $\sigma_f - \Delta\sigma$. The dislocation density in this situation ($\sigma(x) = \sigma_f - \Delta\sigma$ for $-a < x < a$ and $\sigma(x) = -\sigma_f$ for $a < |x| < b$) is equal to

$$B(x) = \frac{2\alpha(2\sigma_f - \Delta\sigma)}{\pi\mu} \log \left| \frac{x(b^2-a^2)^{\frac{1}{2}} + a(b^2-x^2)^{\frac{1}{2}}}{x(b^2-a^2)^{\frac{1}{2}} - a(b^2-x^2)^{\frac{1}{2}}} \right| \qquad (15)$$

where the distance b is equal to $b = a/\cos\{(\pi/2)(\sigma_f-\Delta\sigma)/(2\sigma_f-\Delta\sigma)\}$

Although the solution that has just been constructed and which is represented by equation (14) (or by equation (15)) and the plots shown in Fig. 11 apparently satisfy the condition of slippage with a constant velocity on the transform fault (that is, in the region $-a < x < a$) there is a serious difficulty associated with this solution. The regions that are the projections of the transform fault (that is, the regions $a < |x| < b$) require a stationary distribution of dislocations to prevent the stress there from building up to an infinite value. But any locked-in dislocations on these fossil projections of the transform fault will be moved at a velocity $\dot{D}_0$ away from the spreading center. But if the dislocations in regions $a < |x| < b$ do move away from the spreading center the stress on both the active transform fault ($-a < x < a$) as well as on the fossil projections of the transform fault is changed. The dislocations on the fossil fault projections could move back towards the spreading center. This motion will produce slip on the projections. However, if they do the direction of the slip is such that in the regions $a < |x| < b$ the stress $\sigma(x)$ must have a reversed value. That is, $\sigma(x)$ must change its value from $\sigma(x) = -\sigma_f$ to $\sigma(x) = \sigma_f$. But the slip on the fossil projections

average over time must equal zero. Hence in the regions near the spreading center (see Fig. 12) on the fossil fault projections periods of slip in one direction (with $\sigma(x) = \sigma_f$) must be followed by periods of slip in the opposite direction (with $\sigma(x) = -\sigma_f$). These alternating slip periods will have the effect of placing locked-in dislocations of alternating sign, such as is shown in Fig. 12, on the fossil transform fault projections. At large distances from the spreading centers the dislocations on the fault projections will be locked-in and the fossil fault itself will be locked. The locked-in dislocations do move away from the spreading centers but this motion does not require slip on the fossil fault.

The slip on the transform fault between the two spreading centers cannot take place at a velocity that is constant with time as well as with position if alternating slip must take place on the fault projections near the spreading centers. If the displacement velocity $2\dot{D}(0)$ is larger than the critical velocity $\dot{D}_c$ it is inevitable that periods of fast, catastrophic slippage, that is, earthquakes, must occur on the transform fault.

Savage (reviewed in Savage, in press) has proposed the existence of "dislocation waves" that might explain the apparent migration of earthquake sequences along an earthquake fault. The somewhat qualitative analysis given above would explain how Savage's dislocation waves and migrating earthquake sequences could get started from a spreading center and move along a transform fault.

CONCLUSIONS

In this paper two more examples have been analysed of an inherent instability of quasi-static creep slip on a fault. In the example of a hypothetical symmetric transform fault this inherent instability can be demonstrated rigorously. In the other, more realistic, example of a transform fault between two spreading centers it is clear that a uniform slip velocity across the fault that is constant both in time and in position along the fault cannot be maintained. It is obvious that slip on such a transform fault also is unstable although no attempt was made in this paper to develop mathematical equations that rigorously would prove this claim.

It is predicted that on the fossil fault projections of a transform fault there should exist locked-in infinitesimal dislocations of alternating sign. The existence of these locked-in dislocations could be proved from first motion studies of small and rare earthquakes that might occur on the fossil fault projections. There should be just as many earthquakes with first motions in one direction as those with first motions in the opposite direction if the picture presented in Fig. 12 is correct.

REFERENCES

Savage, J. C. — in press, Dislocations in Seismology, in Dislocation Theory: A Treatise, F. R. N. Nabarro, editor, North Holland, Amsterdam.

Weertman, J. — 1967, Theory of Infinitesimal Dislocations Distributed on a Plane Applied to Discontinuous Yield Phenomena, Canadian Journal of Physics, Vol. 45, pp. 797-807.

Weertman, J. — 1971a, Allen-Wallace Zone of the San Andreas Fault and Discontinuous Yielding of a Slip Band in a Crystal, Nature Physical Science, Vol. 231, pp. 9-10.

Weertman, J. — 1971b, Theory of Water-filled Crevasses in Glaciers Applied to Vertical Magma Transport Beneath Oceanic Ridges, Journal of Geophysical Research, Vol. 76, pp. 1171-1183.

Weertman, J. and J. R. Weertman — in press, Moving Dislocations in Dislocation Theory: A Treatise, F. R. N. Nabarro, editor, North Holland, Amsterdam.

FIGURE CAPTIONS

Fig. 1. (a) Schematic plot of friction stress σ_f and internal stress σ_i versus distance x along a fault. (b) Infinitesimal dislocations that produce the internal stress σ_i. (c) The stress difference $\sigma_f - (\sigma_i + S_i)$ versus distance x where S_i is the initial value of the "applied" stress. Slippage can take place in the patches where the stress difference is equal to zero.

Fig. 2. Localized distribution of infinitesimal dislocations of density B(x) that moves with velocity V across fault plane.

Fig. 3. Spreading center that terminates at a symmetric transform fault.

Fig. 4. Schematic plot of friction stress versus slippage velocity $\dot{D}$.

Fig. 5. The friction stress plot of Fig. 4 is here approximated with a step function.

Fig. 6. Stress σ on a symmetric transform fault versus x for the case in which $\dot{D}_0 < \dot{D}_c$.

Fig. 7. Dislocation distribution B(x) versus x that corresponds to the stress variation shown in Fig. 6.

Fig. 8. Stress σ versus x for a symmetric transform fault for the case in which $\dot{D}_0 > \dot{D}_c$.

Fig. 9. Dislocation distribution B(x) versus x that corresponds to the stress variation shown in Fig. 8 for the case in which $\Delta\sigma/\sigma_f = 0.25$ and $a/b = 0.7$.

Fig.10. Transform fault between two spreading centers.

Fig.11. (a) Stress variation of σ versus x required to have slippage everywhere on transform fault for the case in which the friction stress σ_f is a constant. (b) Dislocation distribution B(x) versus x that corresponds to the stress variation shown in (a).

Fig.12. Infinitesimal dislocations on transform fault between two spreading centers and locked-in infinitesimal dislocations that exist on inactive "fossil" portions of the fault.

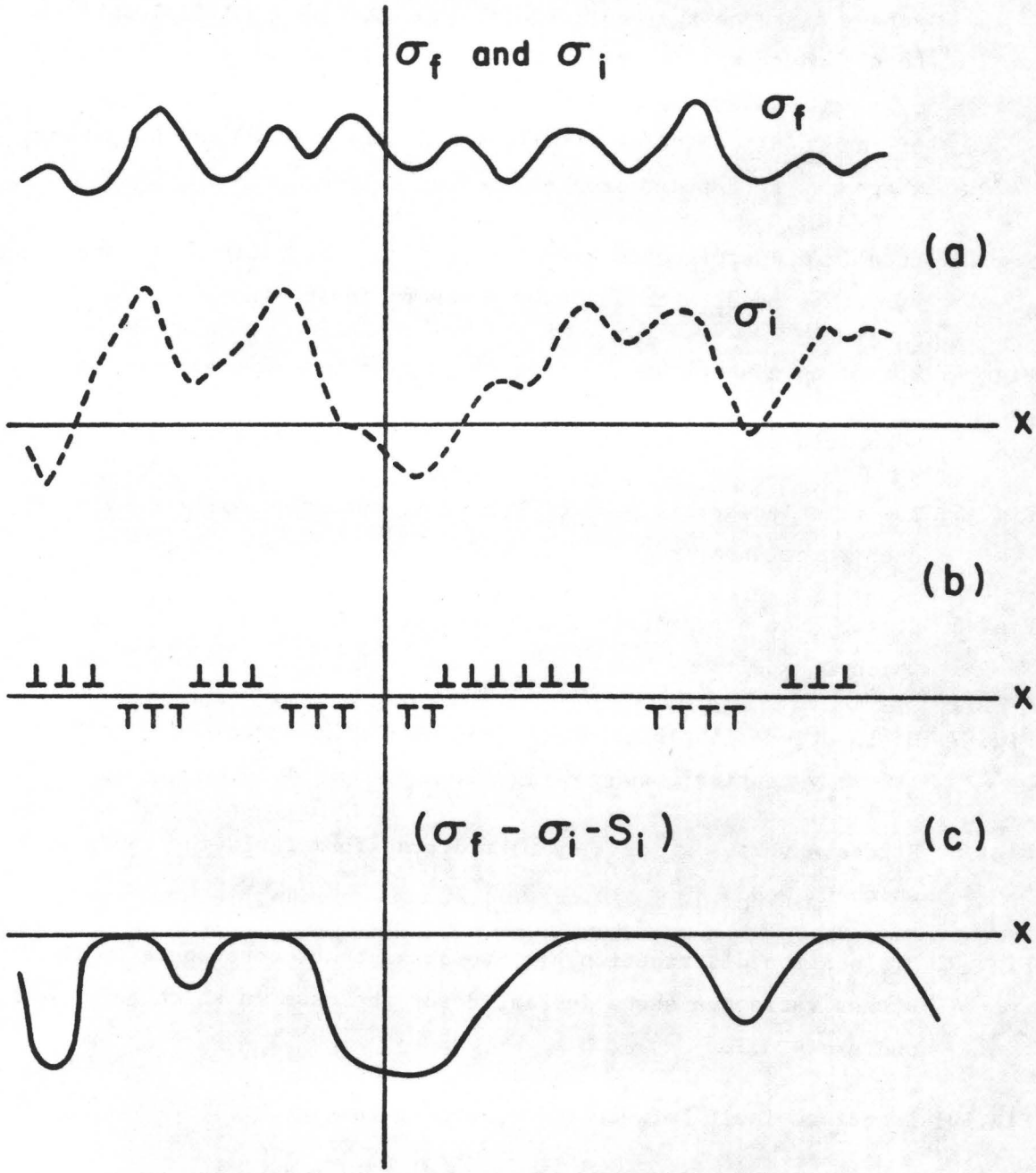


Figure 1

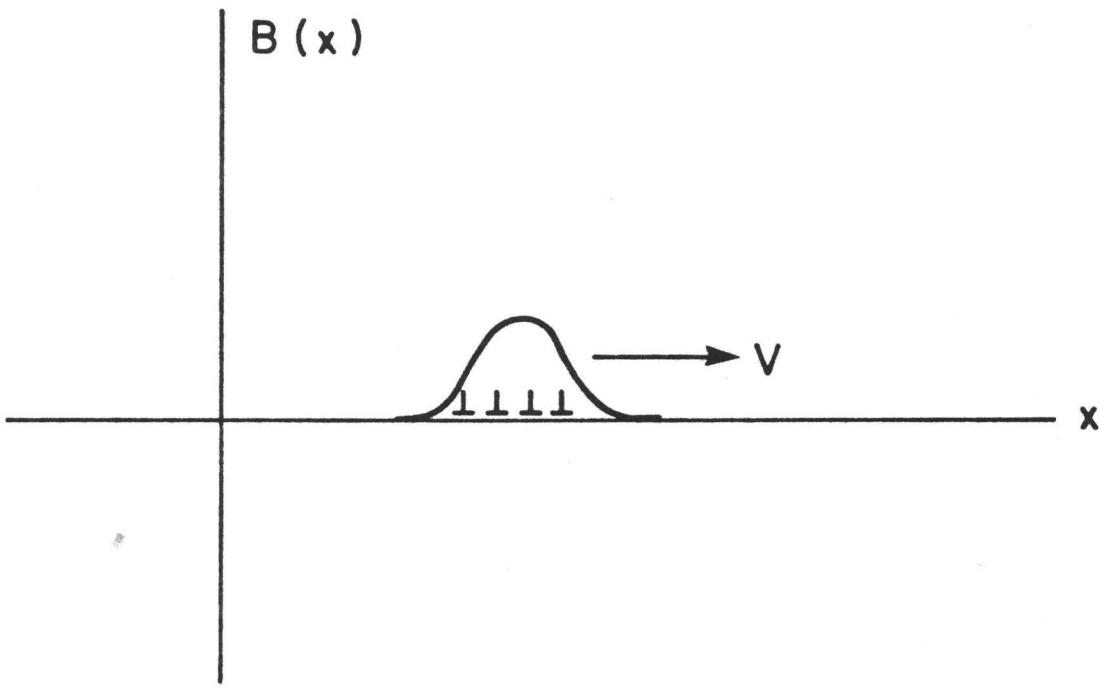


Figure 2

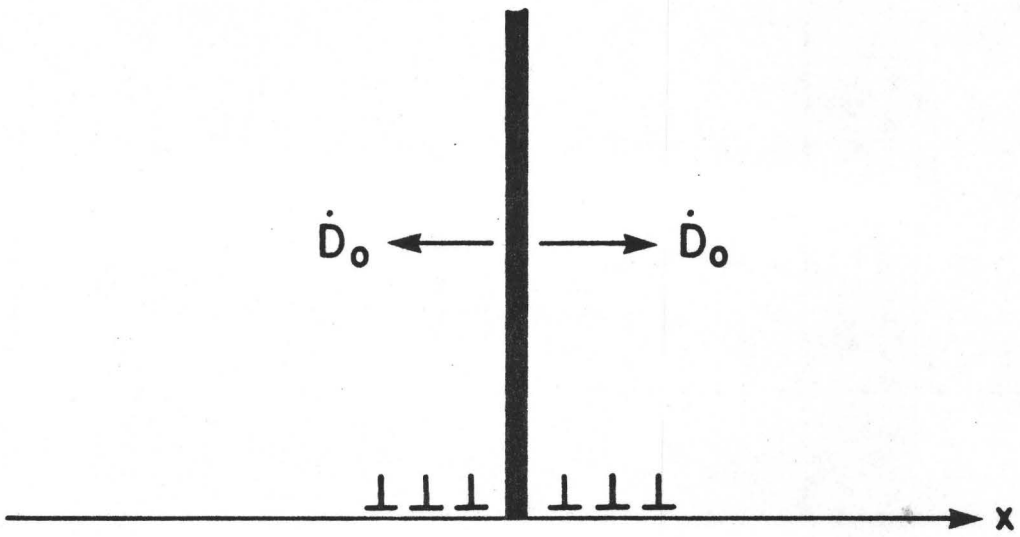


Figure 3

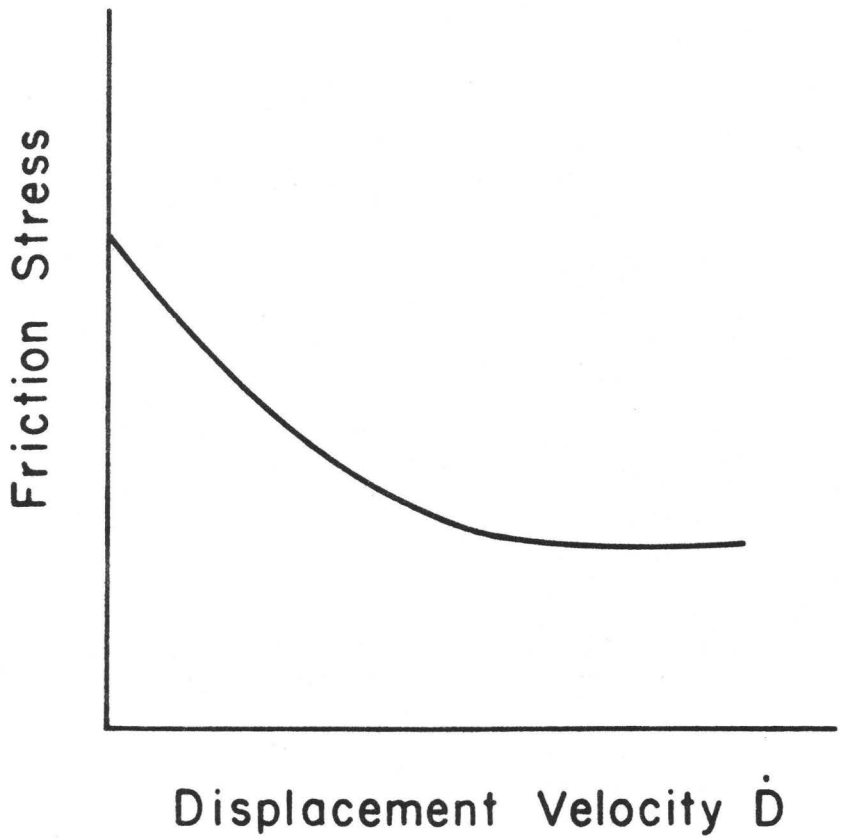


Figure 4

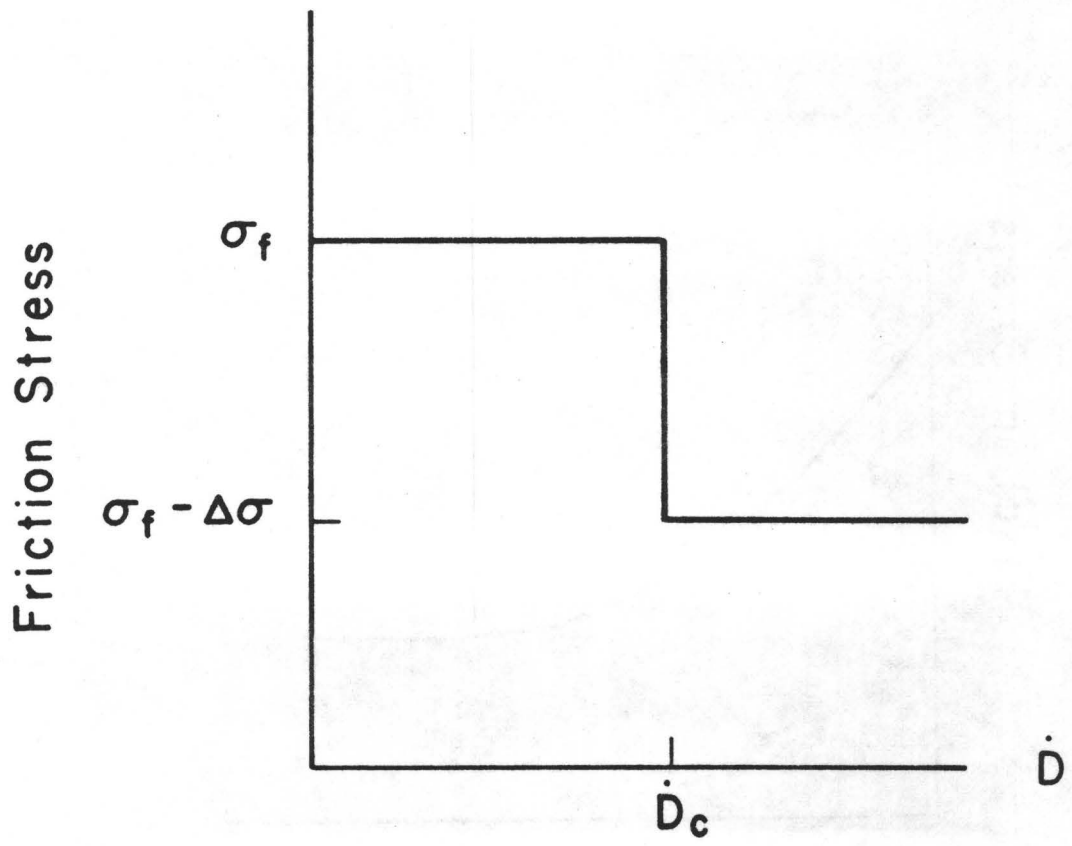


Figure 5

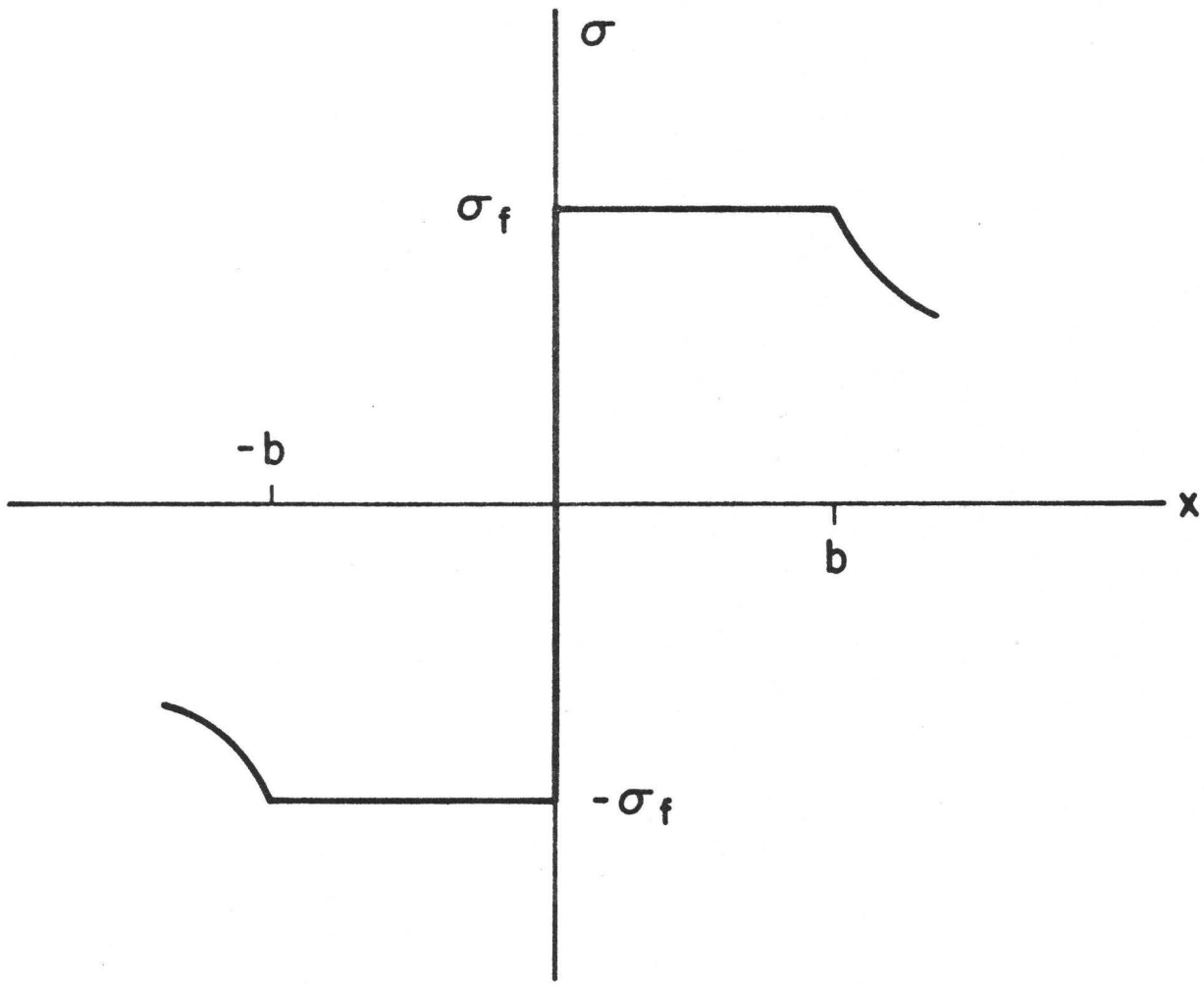


Figure 6

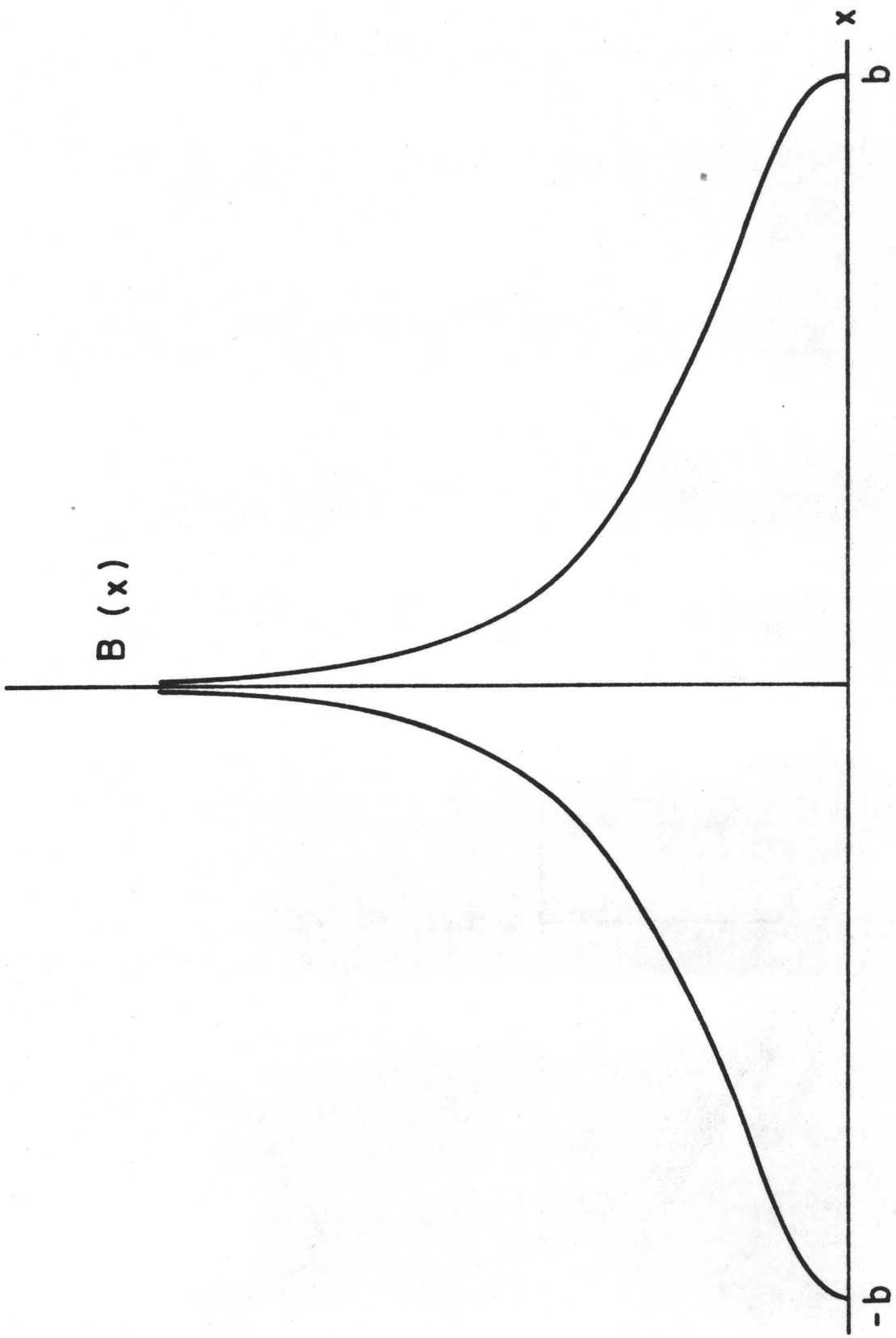


Figure 7

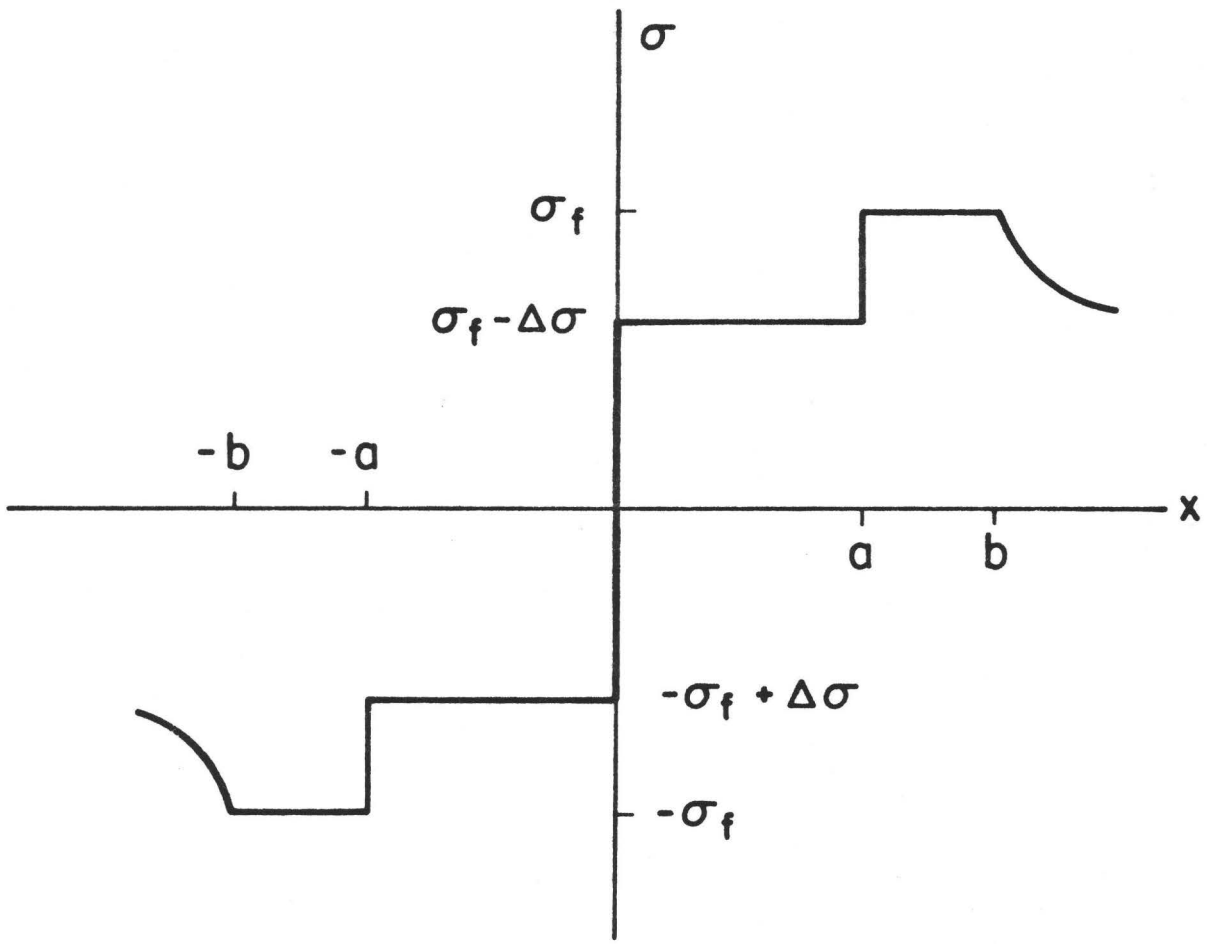


Figure 8

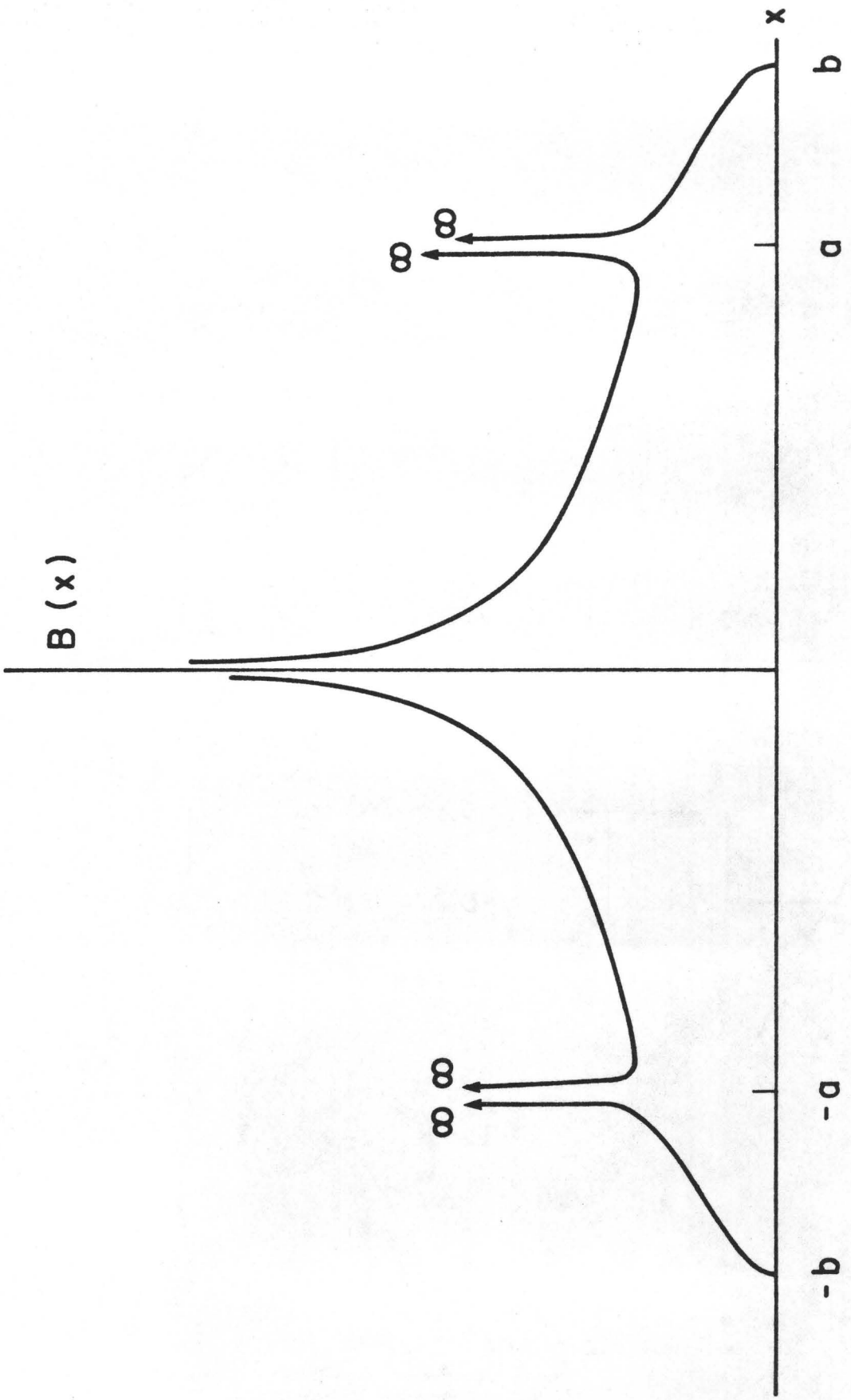


Figure 9

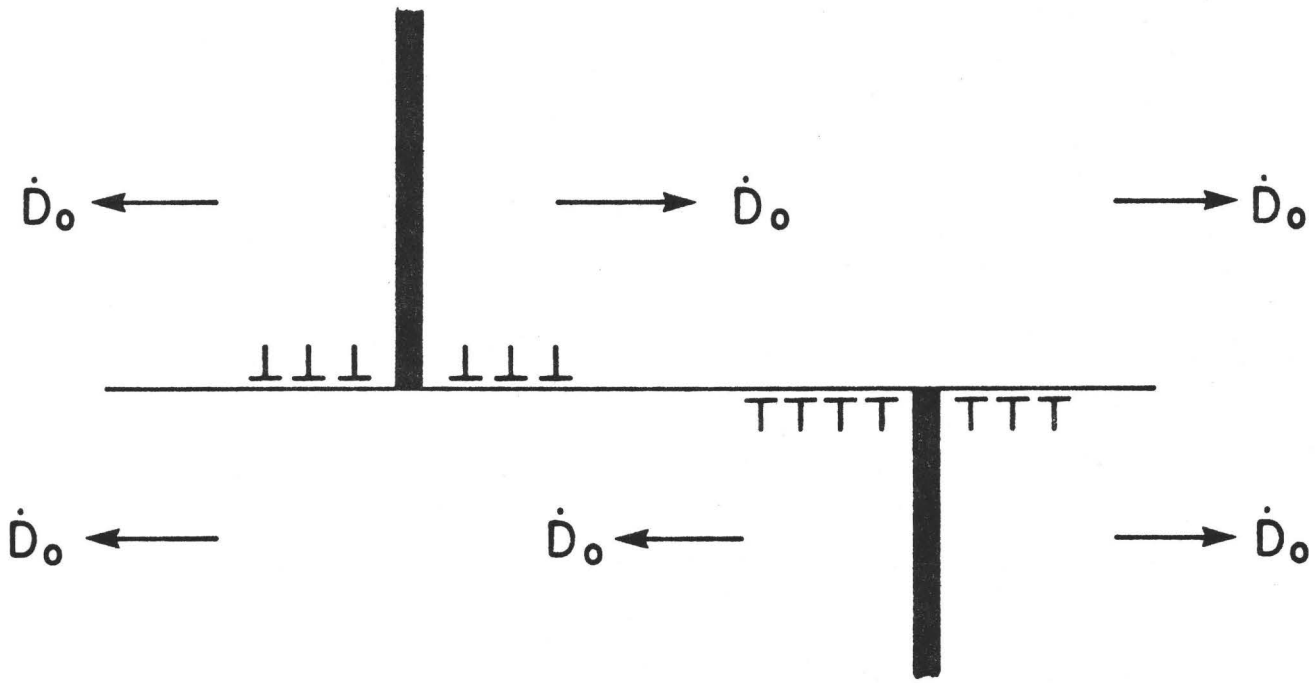


Figure 10

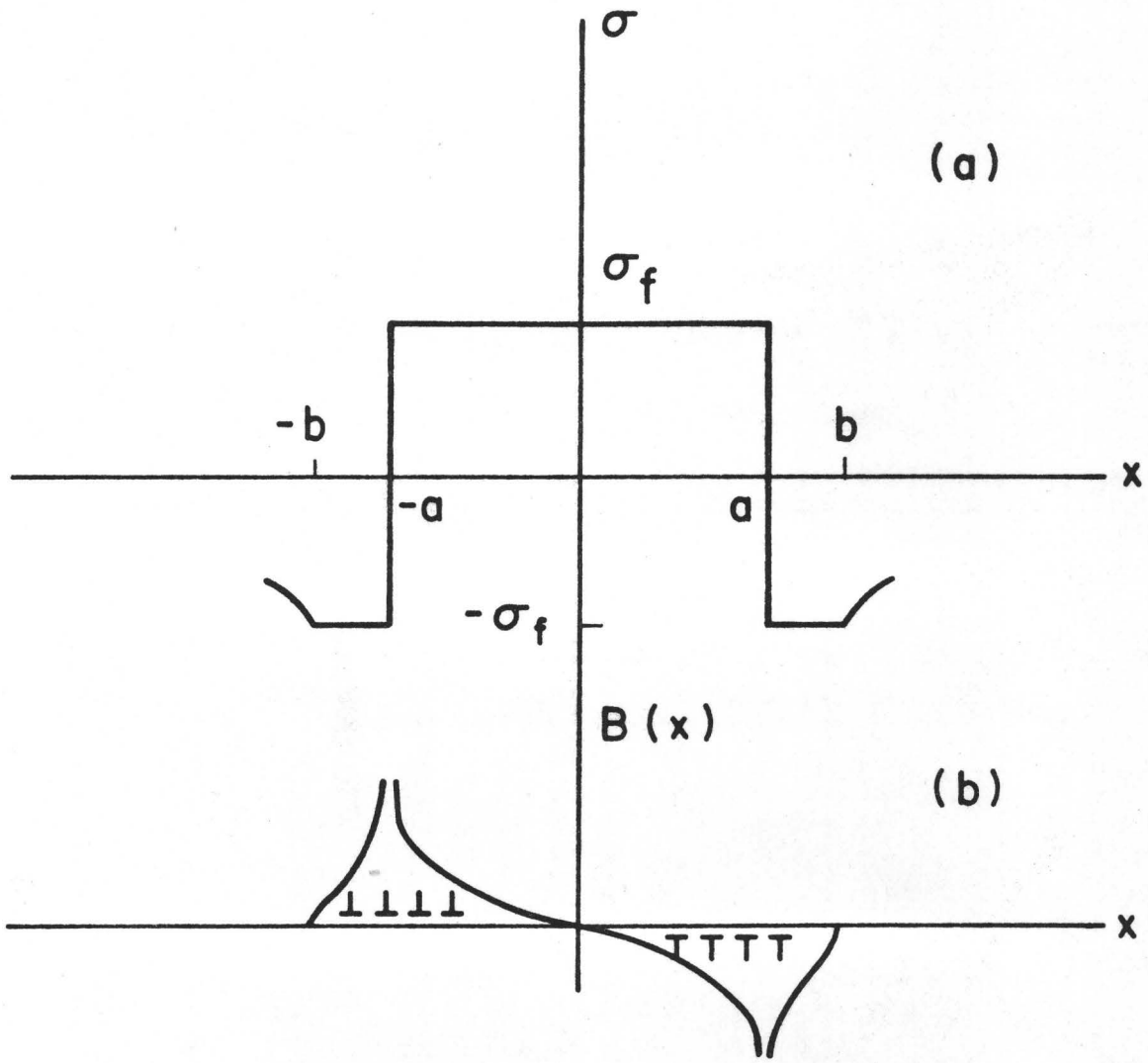


Figure 11

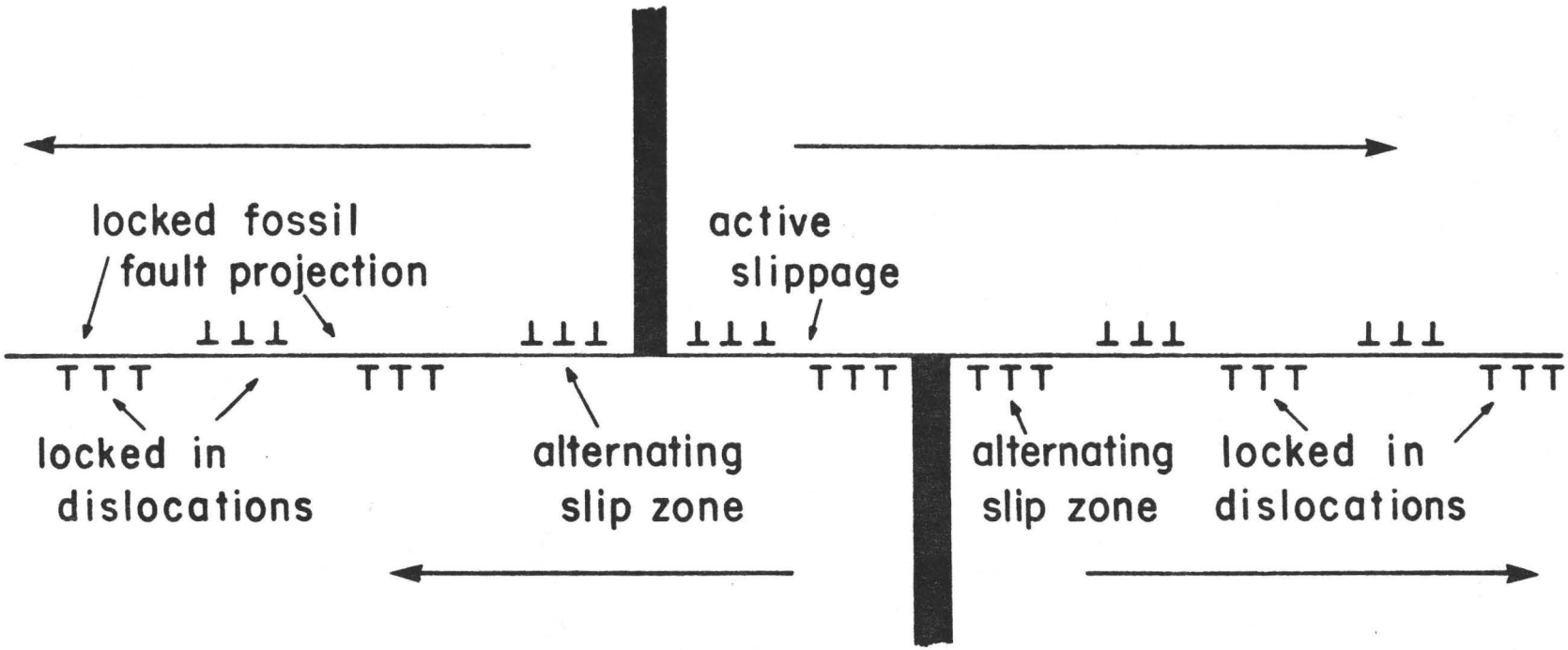

locked fossil
fault projection
active
slippage
locked in
dislocations
alternating
slip zone
alternating
slip zone
locked in
dislocations

Figure 12

Joe Andrews

There's been alot of discussion at this meeting about heterogeneity being basic to fault mechanics and earthquake occurence. For instance, the stress or material properties must be heterogeneous to stop a rupture, and we have earthquakes of all sizes upon a fault. We have seen much evidence for the heterogeneity, in particular, evidence from the South African gold mines. It almost seems as if the process of repeated brittle failure in a natural setting leads to this heterogeneity and irregularity. Whereas if you were talking about a viscous deformation, you would expect a very smooth behavior. I think a major theoretical question is why is brittle failure associated with heterogeneity. A very promising idea is what Jim Rice suggested, that if you have a fresh fracture anywhere on the fault it will be at a different orientation from the major shear direction. We need to have some statistical idea of the heterogeneity on the fault, to understand the strong motion from earthquakes and also just to get at such basic questions as why are there earthquakes of all sizes.

I feel we're still in the basic research stage of the whole problem. For that reason, I tend to focus my thinking on earthquake hazards reduction and not so much on earthquake prediction. Part of the reason for this is I think the definition of prediction is kind of fuzzy. The definition is such that in some sense we are predicting earthquakes right now, everyone of us is predicting an earthquake from some certain section of the fault within certain accuracy and certain magnitudes. We need a socially useful type of definition.

I still think there are still some big uncertainties. For example, if we can't decide whether the stress on the fault is 100 bars or a kilobar, then there are some basic things we need to determine before we can be very sure that the way we spend our money is very efficient.

We need to understand the problem of rupture strength and length of rupture, and what stops a rupture, if we're going to deal with the problem of ruptures growing into regions where there isn't any premonitory evidence nor any way of knowing that the fault is dangerous there. Without this understanding, we won't be able to predict how big an earthquake is going to be. I think we need higher quality standard seismic surveys. We need to use strong motion data for high frequencies, to help understand what the true stress is. We need some good absolute stress measurements. I also think we need to get some temperature measurements on active faults, if possible right after a big earthquake, (or perhaps on faults in the deep African mines) to

get some idea of the energy that's going into heat. I guess, most important, we need higher quality tilt and strain measurements to understand the long term deformation. I have a suspicion that we're going to need better instruments, long base tiltmeters and strain meters. We may not be able to establish the possibilities for earthquake prediction with a lot of low-quality data which yields ambiguous results. It may be expensive, but I think that we can gradually increase our understanding of the tectonics of the fault zone with really precise tilt and strain measurements in conjunction with other types of information. I think we have to approach the earthquake prediction problem, perhaps through the back door, by understanding the basic physics better. In this way we will at least be sure to gain a lot in earthquake hazards reduction, that is, taxpayers will get their moneys worth in hazards reduction, if not in a neat scheme for earthquake prediction in a short time.

James H. Dieterich

Traditionally, three groups, from different disciplines, have worked on the mechanics of earthquakes. Those groups are the theoretical mechanicists, seismologists and laboratory experimentalists. Something this meeting brought home to me is that there is still not enough communication between those groups. The seismologists have questioned the relevance of the theoretical mechanical models to the processes they consider important; namely multiple events, fault heterogeneity and so forth. Some of the models based on those seismological observations have asserted dynamic stresses and complete stress drops on the order of kilobars. In turn, from the point of view of current laboratory studies, I find such seismological models implausible. I think all of our work would benefit from better communication and better attempts to reconcile work in the different groups.

My comments concern the mechanics of dynamic crack propagation, a fairly narrow subject within the scope of this meeting, but a fairly important subject on which the interpretation of seismic data is based. The viewpoint of this subject which I took in my talk is at a very basic research level, a viewpoint which I believe is useful in planning further work in this area. The main objective of the work is to obtain a more accurate picture of the local fracture process on a fault so that seismic data can be better interpreted. In relation to Jim Dieterich's remarks about communication among seismologists, theoreticians, model-makers and so on, I think that it is important to keep in mind that these analytical models provide a very useful language for discussing and interpreting data which may lead to greater insight into the faulting process but, at least at this time, they should not be viewed as providing an absolutely rigid framework by which to draw conclusions.

There are a number of relevant special problems within dynamic fracture which are currently being studied, and I might mention one or two of these problems here. One such problem is concerned with trans-sonic frictional crack propagation on a fault plane, which is a relatively unexplored phenomenon. If we think of fracture as being the transfer of load to prospective fracture sites due to a trans-sonically moving stress concentrator, then this load transfer takes place by means of stress waves travelling at about the same speed that the fracture progresses. This is an unusual situation in fracture mechanics and one worth further study.

Another important problem in this area has to do with fracture criteria. A number of elaborate earthquake models have been proposed and these will undoubtedly have to be analyzed by numerical methods which will require some sort of discretization in the solution procedures. I think it is extremely important to distinguish between the selected fracture criterion and its numerical implementation, i.e., the fracture criterion must not be determined by the nature of the discretization in the numerical procedure. Several other special problems in this area are discussed in the paper.

I think at this point we have many theoretical models and many different laboratory experiments. However, only a very small number of earthquakes have been studied in detail from all aspects, including percursory seismicity, long term observations of tilts and strains, and rigorous analysis of the actual mechanisms of the earthquake and the aftershocks. I think what we need is more rigorous and complete observational studies of perhaps several tens of earthquakes under diverse circumstances. If we had that, it might become more clear which theoretical models and experiments were applicable and which ones were not. Basically, however, I think many more earthquakes must be studied in real detail observationally before we can make much more progress.

Both from the summaries just given and from the conversation generated over the past three days, it seems clear that there is a wide range of approach to and opinions on this subject. There are fundamental aspects of earthquake mechanics that we don't understand, and the order of magnitude of the tectonic stress giving rise to crustal earthquakes is probably at the top of the list. Moreover, there seem to be difficulties in communicating the results of one relevant discipline to another, which is in part due to the often contradictory nature of the separate results. But that's why we have meetings like this, which I have greatly enjoyed.

Hiroo Kanamori

I am more or less of the same opinion as Jim Brune, Jim Savage and Bob Geller. As we have discussed throughout this meeting, fault properties are probably very heterogeneous along the fault zone. It is possible that precursory phenomenon, if any, may be limited to a zone of the initial stress concentration and may not be related to the entire fault length. In view of this, we should be careful in looking at the simple-minded relations on the precursory time versus fault dimension. It is my opinion that, from a practical point of view of earthquake prediction, short-term precursors like foreshocks and land-deformation are very important. In the past, because of the lack of pronounced foreshock activity before major California earthquakes, the usefulness of foreshock activity as a tool for prediction of California earthquakes has been viewed with considerable reservation. In order to study the physical process of pre-earthquake activity, it will be necessary to have broad-band instrumentation near fault zones.

I have one comment and a question: On the first evening of this conference someone asked, "Can we conclude that softening is required for instability?" I think that the answer is "yes" as long as we are considering models with localized zones of slip or inelastic strain imbedded in an otherwise elastic material. Although heterogeneity may be important, I don't think that spatially varying properties alone are enough to cause instability, without local softening or weakening.

My question has to do with our treatment of heterogeneity in earthquake modeling. During this conference we have heard evidence for heterogeneity in the form of wildly varying fault slip and high frequency seismic energy. I'm wondering if these observed heterogeneities can be related in a deterministic way to the distribution or statistics of stress, strength or friction on the fault prior to the earthquake. Or, for all practical purposes, is the earthquake a very random process once rupture has begun? Is there a simple relation between the instantaneous spatial statistics of stress and strength, the ensemble statistics of static parameters contained in the frequency-magnitude relation, and the statistics that determine the seismic spectrum in a given event?

Art McGarr

I want to present several questions which can be answered or which may be capable of being answered observationally. If they are answered conclusively, modelers can save considerable time, money and effort and avoid chasing up false leads. First, the matter of stress levels has been quite central to the discussions of this meeting. There are two essential questions involving the operable stress field. I think we are capable of resolving the controversy about whether the level of stress on the San Andreas fault is less than 100 bars or greater than a kilobar. One procedure would be simply drilling a deep hole and doing a suite of ambient stress measurements at various depths. The capability exists and this hopefully will be done over the next five years. The other question is whether stress can ever be of any use in predicting earthquakes, or even be related in detail to earthquakes. I am very pessimestic about this. If it does show a lot of variability, then it would require incredible luck to make the appropriate observation, plus an incredible amount of money and time. Probably it's simply not practical.

The other questions are some upon which seismologists disagree but for which I think answers are obtainable. These questions are:

- Do b values for foreshocks and aftershocks change or not?

  Some of the people here who are not seismologists may think that b values are lower for foreshocks than for aftershocks, but this question certainly has not been resolved.

- I think we need to get some field evidence on whether there is any observable strain softening.

- As far as I know, no one has ever conclusively shown that there is any field evidence whatsoever for instability preceding failure, and yet a lot of the modeling is done on the basis that this is an observable phenomenon.

- A very important question asked this morning by Amos and one that is capable of being answered is whether earthquakes ever occur in relatively weak rock. If they don't, then we should be focusing our attention on only the strong rocks. The obvious corrolary to this would be that a zone of fault gouge is not really a very useful concept in trying to model fault movement in seismogenic zones.

- Another subsidiary question along that line is whether or not aftershocks occur in previously unfractured material or whether they result from movement on the fault of the main shock. I don't know whether that is really capable of being answered or not but possibly it is.

- Another question is the importance of aseismic slip. Some modelers seem to think that earthquakes are sort of a pathologic phenomenon and that aseismic slip is relatively important. On the other hand, in seismogenic regions the seismic moments appear to be capable of accounting for virtually all of the relative motion between plates.

- The last item is the importance of pore pressure. If we can do the critical observations which indicate that it is not a very important phenomenon, or doesn't really go on in the seismogenic region, we can perhaps get rid of some unnecessary complexity in our models.

I made a chart here of what I consider the old views of what people had before they came here and their new views as expressed by their presentations just now, and I must say that most of us didn't change our views much. After two and a half days, we still hold roughly the same views we held before we came here, including myself. This is shown by the fact that the most important conclusion I drew from this meeting is that heterogeneities are very important in-situ with two resultant implications. One relative to earthquake prediction is that there is a strongly random component in the failure process, and probably in the preparation for failure there are some elements in the earthquake that are not going to be predictable simply because it is random. What exactly it is going to do, I don't know. Whether this randomness will be expressed in uncertainty in time, place or magnitude, I don't know, but this intrinsic random element in the failure process will cause some difficulties in time prediction.

The other implication is that laboratory experiments will have to be designed to cope with this problem. In the past, laboratory experiments have been performed on samples that were ground to be perfectly flat. Fortunately, they never got to be perfectly flat and that maybe is why we have stick-slip in the laboratory. However, one should think of laboratory experiments designed to include heterogeneities to the correct scaling relationship. This is not easy because of the stiffness, but this type of work should be done.

An important question that I see is that we know heterogeneities are important but are they really essential? Can we have failure without heterogeneities? The way to answer this is by a combination of some kind of dynamic modeling that sees failure occur in a continuous system (hopefully, having

some similarities to observed patterns of seismicity), and we have to to into the field and somehow identify the nature of the heterogeneities present on faults. Are they due to variation in elastic constants, in fault gouge properties, in friction, in pore pressure, etc.?

Jim Rice

I have some comments on what we might look at to understand processes leading to instability, rather than the dynamics of the instability per se. Hence, my comments are directed to the understanding of non-elastic, constitutive response for frictional sliding and general deformation of rocks and to the development of models for the pre-instability period. It seems that the biggest question mark we have is that of the relevance of laboratory studies to the field situation. If we could become convinced that laboratory studies were really relevant, as I expect we probably could, then I think we would conclude that the amount that has been done thus far in laboratory studies to really get at strain softening or strain weakening behavior and time dependent behavior is, in fact, very far short of what's needed. I was very encouraged by the kind of study that Art McGarr reported under the category of relevance of laboratory studies to field situations. I think it would be worthwhile to pursue more examinations of faulting at this scale, as the mining scale is perhaps a good intermediate between the laboratory and earthquake levels. At the same time, I think that measurements of stresses by over coring, hydraulic fracturing and the like are very much needed. Another type of field observation which is very little reported so far relates to the kinds of constitutive parameters that enter porous media theories. In the field, these are almost certainly going to be dominated by large scale joints. It's very difficult to estimate from tests on intact laboratory rocks what parameters control drained versus

undrained elastic constants in the field. The only way I know of getting any evidence of that is to look at things like long time consolidation when you change the magnitude of loading on the surface of rock. I think that's an important area to get at. So far, the only thing I've seen discussed very much in the literature is the field estimation of diffusivity, but this is very far from a complete description of all the constitutive parameters that enter porous media theories. If laboratory studies are relevant, I think we must conclude that we want to know a lot more about time-dependent friction of the type Dieterich discussed, especially in connection with strain softening and instability behavior, and I think it would be worthwhile to seek some laboratory evidence on healing under conditions of temperature matching those at depth. Also, in our constitutive studies of friction, I think it's important to look more at the question of roughness and how much this changes the observations reported so far. I could say other things on where we might go with respect to looking at precursory models of the type Rudnicki and I reported on, but I believe that discussion is contained well enough within our paper.

I would second the specific recommendations made by Rice and I would also second the comment of Jim Dieterich that there have been a number of groups involved here and the communication between them has not been very good. This diversity of viewpoints is indicative of a certain lack of understanding of the fundamental features of processes we're trying to describe. I think that suggests that we need further work on a broad front which includes theoretical models which have explicit assumptions that are carried to logical conclusions. Even if these don't lead to predictions which are directly observable at this time, I think they provide a context for further work, both experimental and observational. This was evident in the discussion of strain softening. At the same time, on the bottom line of those models we have a real uncertainty about what numbers to put into them. Some of the things Rice suggested are areas for further work, both in experimental determination of parameters and in the field determination of transport properties. And finally, the ultimate test is what you can observe in the field. However, I think some of the discussion here has pointed out that interpretations of field observations is influenced by concepts that are carried over from times when we knew less. One of these is the fact that stress drops are ordinarily discussed in terms of a planer surface whereas the actual width of a fault zone may be very important in the way in which we determine stress drops. In short, there's a need for progress on all fronts, theoretical, experimental and observational.

I do believe that we need more observations. The observations we have can be explained with very simple models, for example dislocation models. Indeed, we can explain the observations with several different dislocation models. I'm not sure that we need more elaborate theories; rather, I think we need more interpretation of the actual observations. As an example, the Southern California uplift merits more interpretation because the observations there are quite good.

Bill Stuart

It appears that considerable progress has been made in the formulation of mechanical models for pre-earthquake deformation, earthquake instability, and rupture propagation, although no single model exists for a complete sequence. I believe such a model can be constructed if we obtain more knowledge about shear strain localization, post-peak strain and strain rate constitutive properties and their variation in space, and aging (healing) of disrupted fault materials. Information about these phenomena can be gained from a few laboratory experiments, theoretical analyses, geophysical observations, and study of exhumed faults.

Despite elementary mechanical principles and theoretical results, I notice that archaic concepts and language live on. Three examples are: (1) earthquakes occur on a fault when the local stress reaches the local fault strength, (2) absolute stress values are needed to test instability theories, and (3) laboratory data represent the ultimate truth. In fact earthquake instabilities occur when the net resisting force on a fault declines at a critical rate, only deformation changes are needed to test models, and neither laboratory bulk sample tests nor simulations are perfectly scaled versions of field situations.

COMMENTS BY DONALD L. TURCOTTE

It seems to me that there are two approaches to the problem of earthquake prediction. The first is essentially empirical; to measure all observational quantitites hoping that one or more will provide systematic precursory information on the occurrance of earthquakes. The second is to develop a basic understanding of the fundamental processes associated with the occurrance of earthquakes. This understanding may then provide the basis for the systematic prediction of earthquakes.

Although some might argue that the accumulation and release of stress on the San Andreas fault is so complicated that it is impossible to satisfactorily model the process, I would disagree. Although it may, in fact, be impossible to model every aspect of the fault, I believe it will be possible to model the basic behavior. It is clear from this meeting that much remains to be done before satisfactory fault modelling will be possible. Specialists from many branches of mechanics, materials science, applied mathematics, and computer science will be required. I would hope that the U.S.G.S. program on earthquake prediction and hazard reduction would attract increasing numbers of young scientists in these areas into research on the San Andreas fault and other related problems.

J. Weertman

I have two comments. Slow slippage on a fault has the effect of smoothing away internal stress fluctuations. I'd like to see with simple-minded models ways internal stress fluctuations can be regenerated on a fault during the fast slip process. One way might be what I was talking about at the conference. At the source of the motion it may be impossible to have slow, smooth, continuous slippage. Another obvious way: if the stress drops themselves that occur during fast slippage were functions of position internal stress fluctuations would automatically be regenerated on a fault. The stress drops may be functions of position because of variations of rock types along a fault. Anyway, I'd like to see simple, convincing ways to regenerate the internal stress fluctuations that slow slippage removes from a fault.

My second comment concerns what we heard from Joe Andrews yesterday. He pointed out that there is a difficulty with crack tip stress conditions for propagating slip considered as a propagating crack if it's in a mixed mode (combined mode II and III shear crack). The same problem occurs with a moving dislocation, i.e., there are problems with a fast moving mixed dislocation (partly edge and partly screw in character). This difficulty might mean that for a fault slipping in a mixed mode that the moving boundary between the slipping and the unslipped regions is not a continuous smooth line. The boundary line might have to take on a zigzag form with the zig portions moving in a mode II (edge character) manner and the zag portions in a mode III (screw character) manner. Or the boundary line might move with respect to time in a very irregular fashion. Internal stress irregularities might

be left behind after mixed mode slippage. For any small earthquake where you have a loop of an expanding slipped area most of the slippage at the boundary line is in a mixed mode. Thus mixed mode slippage might be rather important to study.

J. WEERTMAN
MATERIALS SCIENCE DEPT.
NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS 60201