VOL. 82, NO. 5 JOURNAL OF GEOPHYSICAL RESEARCH FEBRUARY 10, 1977

The Inception of Faulting in a Rock Mass With a Weakened Zone

JOHN W. RUDNICKI[1]

Division of Engineering, Brown University, Providence, Rhode Island 02912

This paper investigates models for the inception of earth faulting based on the deformation of a rock mass containing an embedded weakened zone. Constitutive laws appropriate to dilatant, frictional, inelastic behavior are used to characterize the weakened zone material. Two distinct types of instability, corresponding to possible models of seismic mechanisms, are identified. These are 'localization' instabilities, at which essentially homogeneous deformation gives way to localized shearing, and 'runaway' instabilities, at which no further quasi-static deformation is possible and inertial effects dominate. Conditions derived for the onset of these instabilities demonstrate that the amount of postpeak deformation in the weakened zone prior to instability is strongly dependent on the deviatoric state of stress induced within the weakened zone and on the detailed nature of the inhomogeneities. In particular, instability is predicted much nearer to peak load for very narrow weakened zones and for states of deviatoric pure shear than for states of axisymmetric compression. Hence, the premonitory events predicted by 'dry crack' precursor models, which associate crack closure with the postpeak regime, would be dramatically different for these two cases. More generally, systematic differences may be observed between strike slip and thrust type faults. A discussion of the qualitative effects of coupled stress–pore fluid diffusion on instability suggests a new interpretation of the dilatancy-diffusion model and indicates that premonitory events predicted by this model may also depend on the amount of postpeak deformation prior to instability.

INTRODUCTION

It has been proposed that dilatancy and the nonuniformity of material properties in a seismically active area are sufficient to produce observed seismic precursors without fluid diffusion [*Myachkin et al.*, 1972, 1974; *Stuart*, 1974; *Brady*, 1974*a*, *b*, 1975*a*]. Despite disagreement over some aspects of the mechanism involved [*Brady*, 1975*b*; *Stuart*, 1975] and a lack of discriminating evidence, this hypothesis is an alternative to that of dilatancy-diffusion [*Nur*, 1972; *Scholz et al.*, 1973] as an explanation of precursory phenomena. However, although these hypotheses and the differences between them [*Myachkin et al.*, 1975] have been discussed in qualitative terms, neither hypothesis has been examined in terms of a quantitative model, incorporating constitutive properties which are consistent with experiments, to determine if they can, in fact, explain the observations in accordance with the laws of mechanics. A principal difficulty is an incomplete understanding of brittle rock failure in general and of mechanisms for the inception of earth faulting in particular. On the other hand, most laboratory investigations of rock failure, despite the evident importance in faulting of heterogeneity of material properties, have been confined necessarily to samples which are relatively homogeneous [e.g., *Brace*, 1964; *Mogi*, 1966; *Wawersik and Fairhurst*, 1970] or which have a flaw which completely transverses the specimen [e.g., *Swanson and Brown*, 1971; *Hobbs*, 1970].

This paper examines some models for the inception of earth faulting based on the deformation of a rock mass in which the heterogeneity of material properties is idealized, for mathematical convenience, as a single ellipsoidal zone. This zone exhibits dilatant, inelastic behavior due to frictional sliding on fissures and microcracking, and it is 'weakened' in the sense that the threshold stress marking the onset of inelasticity is sufficiently less than that in the surrounding, more competent material that the surrounding material remains essentially elastic. Deformation results from the action of remotely applied strains (Figure 1); these may be regarded as boundary conditions imposed by large-scale tectonic processes arising from plate motions or by motion on another portion of the fault system.

Two distinct types of instability in the deformation of the rock mass are identified as possible models for seismic mechanisms. The conditions which are derived for the onset of these instabilities demonstrate that the processes prior to rupture are strongly dependent not only on the weakened zone shape but also on the prevailing stress state. In particular, for a given weakened zone shape, much greater deformation in the postpeak regime of the weakened zone stress-strain curve is required for instability under stress states of axisymmetric compression than under those of deviatoric pure shear. According to dry crack models, which associate the return of anomalous wave speed ratios to normal values with postpeak deformation in the fault zone, these results indicate that different precursor events will be observed for faulting associated with different stress states. More generally, different premonitory events may be expected for strike slip and thrust type faults. A qualitative discussion of coupled deformation–pore fluid diffusion effects indicates that the results are relevant, as well, to considerations of the dilatancy-diffusion hypothesis, although a more thorough analysis is needed.

First, a constitutive law, proposed by *Rudnicki and Rice* [1975] (hereafter referred to as RR) to describe the time-independent behavior of brittle rock, is introduced to characterize the weakened zone response. The parameters are given a concrete interpretation in terms of simple deformation states and experimental observations. Next, the conditions for instability are derived in terms of the weakened zone shape, the stress state, and parameters of the weakened zone constitutive law. These are discussed in terms of mechanisms for premonitory events.

CONSTITUTIVE LAW IN THE WEAKENED ZONE

Weakened Zone Behavior

The weakened zone material is considered to exhibit stress-strain behavior which is representative of rock in a temperature and pressure regime appropriate to brittle deformation.

[1] Now at the Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91109.

Paper number 6B0745.

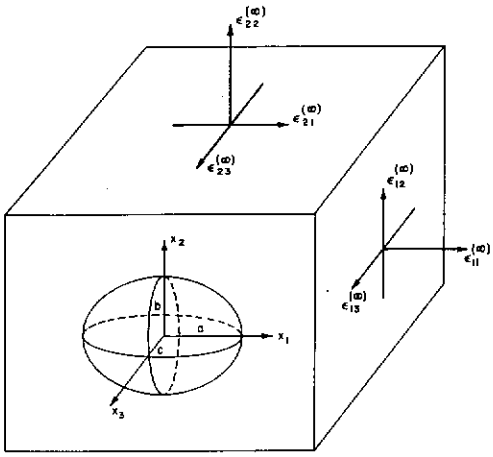

Fig. 1. An ellipsoidal weakened zone with semiaxes $a \geq b \geq c$, embedded in a rock mass subjected to remotely applied boundary strains due to large-scale tectonic motions.

Inelastic behavior results primarily from microcracking and frictional sliding on fissure surfaces. Uplift from sliding at asperities and local tensile cracking at the tips of fissures or other discontinuities give rise to macroscopic inelastic volume change or dilatancy. Typical stress-strain curves are depicted schematically in Figure 2 [e.g., *Brace*, 1964; *Jaeger and Cook*, 1969; *Wawersik and Fairhurst*, 1970]; the onset of inelastic behavior (marked by the onset of dilatancy) occurs at σ_A, and the peak stress is σ_B. Although the precise condition of the material in weakened or fault zones is not well-known, it may be more appropriately regarded as rubblelike or broken, rather than intact, rock. However, laboratory experiments comparing the behavior of intact and prefractured specimens [*Hobbs*, 1970; *Swanson and Brown*, 1971] indicate that the qualitative description given above applies in this case as well.

The weakness of the material in this zone is reflected in the lower values for the stresses σ_A and σ_B as compared with the corresponding stresses in the more competent surrounding material. This interpretation is consistent with experiments comparing the strength of 'virgin' samples with that of rocks weakened by cyclic fatigue or by fracturing. For example, *Haimson* [1974] in a study of four rocks under cyclic loading (Tennessee marble, Indiana limestone, Berea sandstone, and Westerly granite) found that for the number of fatigue cycles tested, the uniaxial peak stress was reduced to 60–80% of the monotonic value. The difference between the peak stresses of intact and prefractured specimens was found by *Swanson and Brown* [1971] to be of the order of 3.5 kbar for Westerly granite and 100 bars for Cedar City tonalite, decreasing to zero at high confining pressure, for confining pressures of 1–7 kbar, and it was found by *Hobbs* [1970] to be 400–600 bars for Ormonde siltstone, Bilsthorpe silty mudstone, and Hucknall shale (for Bilsthorpe mudstone the difference was somewhat less) for confining pressures between 34 and 140 bars. For the rocks tested by Hobbs the difference between the stress marking the onset of dilatancy (σ_A) in the broken samples and that in the intact samples was slightly less than the difference in σ_B. *Scholz and Kranz* [1974] and *Haimson* [1974] also noted a decrease in σ_A with continued cycling. However, *Zoback and Byerlee* [1975] attributed this effect in the former to the fact that the loading was in uniaxial stress, as it was not observed in their tests of cycling under triaxial conditions.

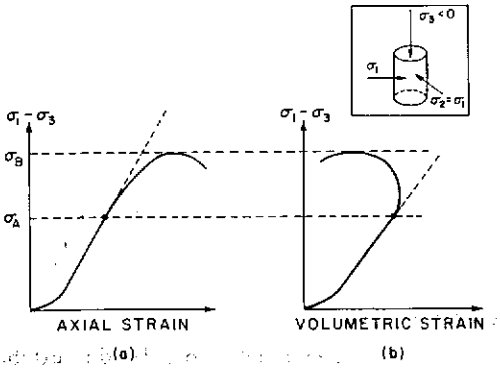

Fig. 2. Representative stress-strain curves for brittle rock in axisymmetric compression.

Constitutive Law

The constitutive law which will be used to characterize the response of the weakened zone material is a simple generalization, in a manner analogous to the Prandtl-Reuss equations often used in metal plasticity, of the elementary forms of stress-strain laws used in soil and rock mechanics. This is one of the elastic-plastic laws proposed by RR to describe time-independent brittle rock behavior, and it incorporates pressure dependence of inelastic deformation and inelastic volume change. 'Plastic' is used here in the generic sense of referring to any inelastic deformation, in particular, inelasticity which in the present case arises from frictional sliding and microcracking.

Although this constitutive law has been treated extensively by RR, it will be presented again in some detail in order to make clear its structure and to emphasize the relation with experimental data. For an increment of elastic unloading from some general stress state the deviatoric and volumetric strain increments are given by

$$2\,d\epsilon_{ij}' = d\sigma_{ij}'/G \qquad d\epsilon_{kk} = d\sigma_{kk}/3K \tag{1}$$

respectively, where a prime on a tensor denotes the deviatoric part and G and K are the appropriate incremental shear and bulk moduli (Figure 3). These moduli, in general, will decrease with continued microcracking during the deformation. For example, in the axisymmetric compression test the data of *Lee et al.* [1972] on Hawkesbury sandstone show that Young's modulus ($E = 2G(1 + \nu)$) for unloading at a point slightly past peak stress has decreased to 0.85 of its value in the regime of linear behavior, and the data of *Wawersik and Fairhurst* [1970] on Tennessee marble suggest that in the postpeak regime of general failure, E decreases to fractions ranging from 0.9 to 0.5 of the initial value. Although some clustering of cracks in these samples occurred in the postpeak regime, they were still relatively homogeneous, and no macroscopic fault had formed. A slight decrease in the unloading moduli may also be observed in cyclic fatigue tests [e.g., *Zoback and Byerlee*, 1975; *Haimson*, 1974]. In general, continued loading results in inelastic deformation, and the inelastic, or plastic, portions of the strain increments must be added to those of (1). These are defined as the strain remaining after an infinitesimal cycle of loading followed by elastic unloading which restores the stress state, and they are written

$$2\,d^P\epsilon_{ij}' = \frac{\sigma_{ij}'}{h\bar{\tau}}\left[d\bar{\tau} + \mu\,\frac{d\sigma_{kk}}{3}\right]$$
$$d^P\epsilon_{kk} = \beta(2\,d^P\epsilon_{ij}'\,d^P\epsilon_{ij}')^{1/2} \tag{2}$$

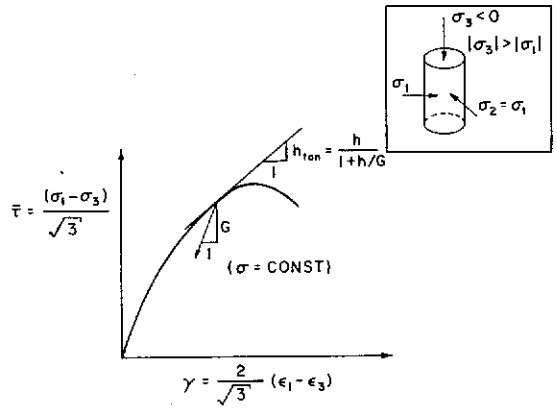

Fig. 3. Constitutive law (equation (4)) specialized to axisymmetric compression. The geometric interpretation of the hardening modulus h and the shear modulus for unloading G is illustrated.

where $\bar{\tau} = (\sigma_{ij}'\sigma_{ij}'/2)^{1/2}$ and β, μ, and h are the dilatancy factor, internal friction coefficient, and plastic hardening modulus, respectively.

Illustration for Axisymmetric Compression

If $\sigma = -\sigma_{kk}/3$ (positive in compression), $\epsilon = \epsilon_{kk}$, and $\gamma = (\epsilon_1 - \epsilon_3)/3^{1/2}$, (2) reduces to

$$d^P\gamma = (d\bar{\tau} - \mu\, d\sigma)/h$$
$$d^P\epsilon = \beta\, d^P\gamma \tag{3}$$

where $\bar{\tau} = (\sigma_1 - \sigma_3)/3^{1/2}$. An experiment may now be imagined in which an axially symmetric sample (Figure 3) is subjected to a constant hydrostatic pressure, say, $\sigma = \sigma^*$, and $\bar{\tau}$ is increased to the value at which the onset of inelastic behavior occurs, say, $\bar{\tau} = \bar{\tau}^*$. Repeating this procedure for different values of σ traces an 'initial yield surface' (Figure 4*a*). The coefficient of internal friction μ_0 defines the slope of this yield curve, and it represents the pressure dependence of the onset of inelastic behavior. Using data of *Brace et al.* [1966] on Westerly granite and aplite and of *Bieniawski* [1967] on norite and quartzite, RR found values of μ_0 ranging from 0.4 to 0.9. Data from *Hobbs* [1970] on both broken and intact samples generally yielded values which were also in this range, although μ_0 for some intact samples of the silty mudstone and the siltstone was as high as 1.15. Data from *Schock et al.* [1973] on Climax Stock granodiorite gave values of μ_0 from 0.3 to 0.7.

Note that this interpretation of μ differs from that based on a curve of failure stress or of peak stress at different confining pressures. Such curves predict somewhat higher values of μ: 0.9–1.3 for the data of *Brace et al.* [1966] and *Bieniawski* [1967] and 0.7–1.1 for the data of *Schock et al.* [1973]. An example of the difference between the failure curve and the curve denoting the onset of inelasticity is shown in Figure 4*b*, in which the data from Figure 2 of *Schock et al.* [1973] have been replotted. The failure surface would coincide with a yield surface only if the mechanism of failure was precisely the same as that for inelastic deformation. Figure 4*b* also illustrates that μ, for low confining pressures, tends toward the higher values of the range quoted.

In general, the 'current yield surface' (Figure 4*a*) separates the regions of elastic unloading from those of continued inelastic deformation. Deformation increments making $d\bar{\tau} < \mu\, d\sigma$ correspond to elastic unloading. Those tending to make $d\bar{\tau} > \mu\, d\sigma$ correspond to continued inelastic loading, and the plastic shear and volumetric strain increments are given by (3). As shown in Figure 3, the hardening modulus h is related to the slope of the $\bar{\tau}$ versus γ curve by $h_{tan} = h/(1 + h/G)$. Consistent with the behavior of frictional materials, $d^P\gamma$ depends not only on the shear stress but also on the hydrostatic component of stress. Despite some indications to the contrary from microscopic examination of stress cracks by *Tapponnier and Brace* [1976] it seems most likely that dilatancy arises from uplift in sliding at asperities, local tensile fissuring at crack tips, and the propping open of microcracks. Thus it is plausible that at any given deformed state, $d^P\epsilon$ bears a fixed relation to $d^P\gamma$ expressed by the dilatancy factor β. In Figure 4*a* the inelastic strain increment, depicted as a vector, would be normal to the yield surface if $\beta = \mu$. However, this 'normality' assumption of classical plasticity, appropriate when slip depends only on the thermodynamical conjugate force [*Rice*, 1974], would be too restrictive for frictional materials.

From the data of *Brace et al.* [1966, Figure 6], RR inferred typical values of β as 0.2–0.4 and noted that β diminishes slightly with confining pressure. Using a stiff-testing machine, *Crouch* [1970] measured volumetric strains for a norite and a sandstone strained far into the postpeak regime. Values of β inferred from this work are approximately 1.1 for the norite and 0.8 for the sandstone. It is unclear to what degree these large values of β may be attributed to deformation into the postpeak regime instead of the character of the rock type. The data of Brace et al. indicate that β may increase by a factor of 2 from the inception of inelastic behavior to conditions near peak stress, and while such an increase is not evidenced in the

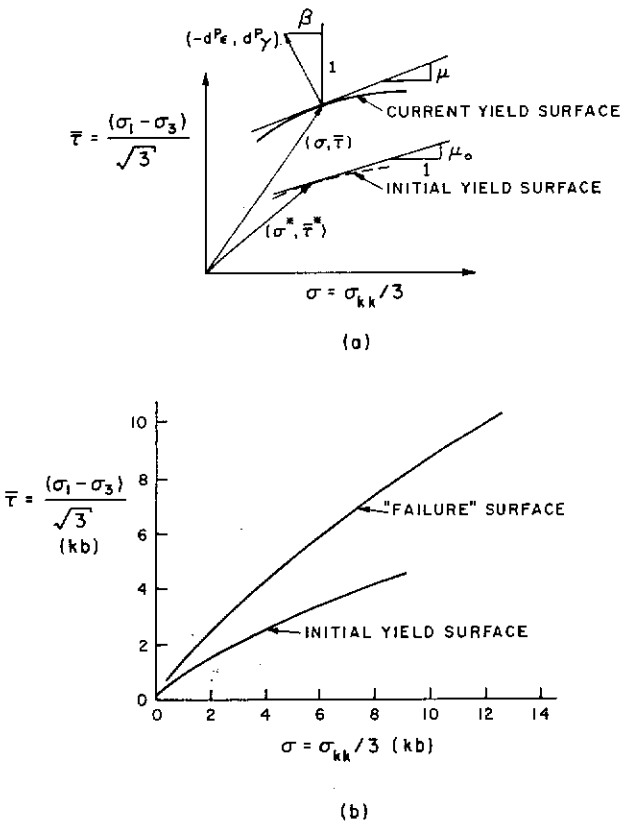

Fig. 4. (*a*) Initial and current yield surfaces in stress space. The geometric interpretation of the friction coefficient μ and the dilatancy factor β is illustrated. (*b*) Initial yield surface and 'failure' surface for Climax Stock granodiorite, replotted from Figure 2 of *Schock et al.* [1973].

data from Crouch, further increases in β may occur in the postpeak regime.

In summary, the constitutive law for the weakened zone is written

$$2\,d\epsilon_{ij}' = \frac{d\sigma_{ij}'}{G} + \frac{\sigma_{ij}'}{h\bar{\tau}}\left[\frac{\sigma_{kl}'\,d\sigma_{kl}}{2\bar{\tau}} + \mu\,\frac{d\sigma_{kk}}{3}\right]$$
$$d\epsilon_{kk} = \frac{d\sigma_{kk}}{3K} + \frac{\beta}{h}\left[\frac{\sigma_{kl}'\,d\sigma_{kl}}{2\bar{\tau}} + \mu\,\frac{d\sigma_{kk}}{3}\right] \tag{4}$$

RUNAWAY INSTABILITY

The boundaries of the rock mass are taken to be sufficiently far away that the weakened zone is effectively embedded in an infinite matrix and a special property of an ellipsoidal inclusion in an infinite body may be exploited. *Eshelby* [1957] demonstrated that if an ellipsoidal region of an isotropic, homogeneous elastic body undergoes a 'transformation' strain, which in the absence of the constraint of the surrounding material would be the homogeneous strain ϵ_{ij}^T, then the resulting stress and strain in the inclusion are uniform. The strain is given by

$$\epsilon_{ij}^C = S_{ijkl}\epsilon_{kl}^T$$

where the S_{ijkl} are shape factors depending only on the geometry of the ellipsoid and Poisson's ratio. The factors S_{ijkl} are symmetric in ij and kl, but in general, $S_{ijkl} \neq S_{klij}$ (see appendix). Values of the S_{ijkl} for oblate spheroids (semiaxes $a = b > c$) are tabulated in Table 1. If a uniform strain $\epsilon_{ij}^{(\infty)}$ is now applied at infinity, the inclusion strain becomes

$$\epsilon_{ij}^{(i)} = \epsilon_{ij}^{(\infty)} + \epsilon_{ij}^C$$

and the uniform inclusion stress is (since ϵ_{ij}^T occurs without stress)

$$\sigma_{ij}^{(i)} = L_{ijkl}(\epsilon_{kl}^{(i)} - \epsilon_{kl}^T) \tag{5}$$

where $L_{ijkl} = (K - 2G/3)\delta_{ij}\delta_{kl} + G(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})$; G and K are the linear elastic shear and bulk moduli, respectively; and $\delta_{ij} = 1$, if $i = j$, and $\delta_{ij} = 0$ otherwise.

Now take an ellipsoid which has the same size and shape as the untransformed (before application of ϵ_{ij}^T) inclusion of matrix material and apply to it the homogenous strain $\epsilon_{kl}^{(i)}$. If the constitutive law of this ellipsoid is such that the resulting stress is $\sigma_{ij}^{(i)}$, then it may replace the inclusion of matrix material, without violating continuity of displacement or surface traction at the interface [*Eshelby*, 1957]. Alternately, specifying the constitutive law, not necessarily elastic, and finding the appropriate ϵ_{ij}^T in accordance with (5) and the relation

$$\epsilon_{ij}^{(i)} - \epsilon_{ij}^{(\infty)} = S_{ijkl}\epsilon_{kl}^T \tag{6}$$

solve the problem of the ellipsoidal inhomogeneity embedded in a linear elastic matrix deformed by uniform strain at infinity.

Since the Eshelby theory supposes small strain, changes of geometry are neglected as second order, and if the elastic moduli are interpreted as incremental values, (5) and (6) may be taken to relate increments of stress and strain. Combining these equations yields, after some rearrangement,

$$S_{klij}(L^{-1})_{mnij}\,d\sigma_{mn}^{(i)} = S_{klij}\,d\epsilon_{ij}^{(i)} + d\epsilon_{kl}^{(\infty)} - d\epsilon_{kl}^{(i)} \tag{7}$$

Equation (4) may be inverted for the weakened zone stress increments in terms of the strain increments

$$d\sigma_{kl}^{(i)} = L_{klij}\,d\epsilon_{ij}^{(i)} - L_{klij}P_{ij}[G(\sigma_{mn}'^{(i)}/\bar{\tau}) + \mu K\delta_{mn}]\,d\epsilon_{mn}^{(i)} \cdot (h + G + \mu K\beta)^{-1} \tag{8}$$

where $P_{ij} \equiv (\sigma_{ij}'^{(i)}/2\bar{\tau}) + (\beta/3)\delta_{ij}$. Obviously, the $d\sigma_{ij}^{(i)}$ may be eliminated between (7) and (8), leaving an expression relating the weakened zone strain increments and the remote boundary strain increments

$$d\epsilon_{mn}^{(i)}\{\delta_{ml}\delta_{nj} - S_{ijkl}P_{kl}[G(\sigma_{mn}'^{(i)}/\bar{\tau}) + \mu K\delta_{mn}]\} = d\epsilon_{ij}^{(\infty)} \tag{9}$$

Because of the symmetry of $d\epsilon_{kl}^{(i)}$ and S_{ijkl}, (9) is a quasi-linear system of six equations for the six weakened zone strain increments. The condition for runaway instability, that is, the condition for which the ratio of a strain increment in the weakened zone to a remote boundary strain increment may become unbounded is

$$\det|\mathbf{M}| = 0 \tag{10}$$

where $\mathbf{M}$ is the 6×6 matrix whose elements are given by the bracketed term of (9). Otherwise, if $\det|\mathbf{M}| \neq 0$, the coefficient matrix possesses a unique inverse, and a unique solution for the $d\epsilon_{kl}^{(i)}$ in terms of the far-field strain is possible.

Equation (10) may be regarded as a condition on one of the weakened zone constitutive parameters, taken to be the hardening modulus h, in terms of the other constitutive parameters, the geometry of the weakened zone, and the stress state. Because h is a decreasing function of strain, the value for which the instability condition is first met in the progress of deformation is the maximum value of h satisfying (10). Hence the critical value of the hardening modulus at runaway is

$$h_r = -G\left[1 - \sum_{\substack{i,j=1\\ i\neq j}}^{3}(\sigma_{ij}'^{(i)}/\bar{\tau})^2 S_{ijij}\right] - \mu K\beta + \sum_{i,j=1}^{3}[G(\sigma_{ii}'^{(i)}/\bar{\tau}) + \mu K]S_{iijj}[(\sigma_{jj}'^{(i)}/2\bar{\tau}) + \beta/3] \tag{11}$$

where all summations have been written explicitly. Surveying the values of the S_{ijkl} tabulated in Table 1 and observing that the terms $\sigma_{ij}'/\bar{\tau}$ are bounded by unity reveal that h_r will be negative and only in certain circumstances will it approach

TABLE 1. Values of S_{ijkl} for Oblate Spheroids (Semiaxes $a = b > c$) for Various Aspect Ratios c/a

Aspect Ratio c/a	$S_{1111} = S_{3333}$	S_{2222}	$S_{1122} = S_{3322}$	$S_{2211} = S_{2233}$	$S_{1133} = S_{3311}$	$S_{1212} = S_{3232}$	S_{1313}
0.01	0.014	0.994	−0.003	0.241	0.001	0.491	0.007
0.1	0.120	0.938	−0.021	0.177	0.005	0.427	0.057
0.25	0.247	0.843	−0.032	0.106	0.008	0.356	0.119
0.5	0.376	0.702	−0.028	0.045	0.007	0.295	0.184
0.75	0.451	0.589	−0.015	0.015	0.004	0.265	0.223
1.0	0.500	0.500	0.0	0.0	0.0	0.250	0.250

The distinguished axis is the x_2 direction. Calculation for $\nu = 0.2$. $S_{ijkl} = S_{jikl} = S_{ijlk}$; other entries not shown are zero.

zero. Although (11) is a result of general applicability and presents no difficulties of computation in a particular case, the salient features of the runaway instability and the effects of the various parameters are best illustrated by considering two important limiting cases.

Spherical Weakened Zone

For a spherical weakened zone the relations (6) have the particularly simple form

$$\alpha\epsilon_{kk}^{T} = \epsilon_{kk}^{(i)} - \epsilon_{kk}^{(\infty)} \qquad \kappa\epsilon_{ij}'^{T} = \epsilon_{ij}'^{(i)} - \epsilon_{ij}'^{(\infty)}$$

where the prime again denotes the deviatoric part of the tensor, $\alpha = (1 + \nu)/3(1 - \nu)$ and $\kappa = 2(4 - 5\nu)/15(1 - \nu)$, and ν is Poisson's ratio. The expression for the critical hardening modulus simplifies accordingly:

$$h_r = -G(1 - \kappa) - \mu K\beta(1 - \alpha) \tag{12}$$

The ranges of the parameters α and κ are 0.3333–1 and 0.400–0.5333, respectively; typical values for brittle rock are $\nu = 0.2$ and $\alpha = \kappa = \frac{1}{2}$. Thus runaway instability can occur for this weakened zone geometry only after deformation far into the postpeak regime.

Flat, Cracklike Weakened Zone

This configuration approximates a flat, tabular weakened zone, and it is an important limiting case in that runaway instability will occur earlier in the progress of deformation for this shape than for those of more nearly equal dimensions. One could proceed with the analysis by expanding the shape factors S_{ijkl} for small values of aspect ratio, neglecting terms beyond first order, and continuing as before, but a more direct approach makes use of a result from fracture mechanics.

In light of Eshelby's result that the strain is constant in an ellipsoidal inhomogeneity embedded in a linear elastic solid uniformly strained at infinity, a penny-shaped crack is the limit of a very flat ellipsoidal inhomogeneity having vanishing moduli. Because the ellipsoid strains must be compatible with the matrix strains at the interface, the crack surface displacements determine, apart from an inessential rigid rotation, the homogeneous strains in the ellipsoid. Expressions for the surface displacements of a penny-shaped crack of radius a loaded by surface tractions and stresses (or strains) at infinity are well-known from fracture mechanics solutions [e.g., *Sih and Liebowitz*, 1968]:

$$u_1 = 4(1 - \nu)(\sigma_{21}^{(\infty)} - \sigma_{21}^{(i)})(a^2 - x_1^2 - x_3^2)^{1/2}/\pi G(2 - \nu) \tag{13}$$

$$u_2 = 2(1 - \nu)(\sigma_{22}^{(\infty)} - \sigma_{22}^{(i)})(a^2 - x_1^2 - x_3^2)^{1/2}/\pi G$$

where u_1 and u_2 are the displacements in the x_1 and x_2 directions, respectively; the x_2 direction is normal to the crack surface, and the x_1 axis may be chosen, without loss of generality, so that $\sigma_{23}^{(\infty)} = 0$. These, of course, are the additional displacements due to inserting the crack in the existing elastic field. Because the penny-shaped crack is the limit of a very flat spheroid, the equation of the boundary in the deformed state is

$$(x_1^2 + x_3^2)/a^2 + x_2^2/c^2 = 1$$

and $(a^2 - x_1^2 - x_3^2)^{1/2}$ may be replaced in the expression for the displacements by $\pm x_2(a/c)$ (the positive sign applies for the upper surface of the crack, and the negative sign for the lower). Thus the weakened zone strain increments may be written

$$d\epsilon_{21}^{(i)} = (d\sigma_{21}^{(\infty)} - d\sigma_{21}^{(i)})/2G\eta$$

$$d\epsilon_{22}^{(i)} = (d\sigma_{22}^{(\infty)} - d\sigma_{22}^{(i)})/G\zeta \tag{14}$$

where $\eta = (c/a)\pi(2 - \nu)/4(1 - \nu)$ and $\zeta = (c/a)\pi/2(1 - \nu)$. The remote strain increments are related to $d\sigma_{21}^{(\infty)}$ and $d\sigma_{22}^{(\infty)}$ by

$$d\sigma_{21}^{(\infty)} = 2G\, d\epsilon_{21}^{(\infty)}$$

$$d\sigma_{22}^{(\infty)} = (K - 2G/3)\, d\epsilon_{kk}^{(\infty)} + 2G\, d\epsilon_{22}^{(\infty)}$$

Equations (14) supplant (7) and could have been obtained by asymptotic expansion of the S_{ijkl}. The remainder of the inclusion strains are simply equal to the elastic strains in the matrix to order c/a, where $c/a \ll 1$. Thus the relevant coefficient matrix is again 2×2, and the solution for the critical hardening modulus is

$$h_r = -G[1 - (1 - \eta)(\sigma_{21}^{(i)}/\bar{\tau})^2] - (\mu\beta K - (1 - \zeta G/M)[G(\sigma_{22}'^{(i)}/\bar{\tau}) + \beta K] \cdot [G(\sigma_{22}'^{(i)}/\bar{\tau}) + \mu K]/M\} + O(c/a)^2 \tag{15}$$

where $M = K + 4G/3$.

Although (15) was obtained by considering the limit as c/a goes to zero, the result gives the first term of a perturbation expansion in the aspect ratio. Hence the condition for 'small' aspect ratios is that they be negligible with respect to unity, and (15) may give acceptable results for c/a of the order of a few tenths. Note that β and μ occur symmetrically here, although not in (11). If the compressive stress normal to the plane of the weakened zone ($-\sigma_{22}$) exceeds in magnitude the average in-plane normal stress, then $\sigma_{22}'^{(i)}/\bar{\tau}$ is negative; in this sense, compression normal to the plane of the weakened zone inhibits the onset of instability, although it is clear from (15) that the effect also depends on the relative magnitude of K/G and of $\beta + \mu$.

If the stress induced by large-scale tectonic processes is principally shear and greatly exceeds in magnitude any ambient stress levels, as would be expected near a large transform fault system (e.g., the San Andreas), a further limiting case may be obtained. Thus if one makes the approximation $\sigma_{22}'^{(i)}/\bar{\tau} \approx 0$ and $\sigma_{21}^{(i)}/\bar{\tau} \approx 1$, (15) reduces to

$$h_r/G = -(c/a)\frac{\pi(2 - \nu)}{4(1 - \nu)} - \frac{4}{9}\mu\beta\frac{(1 + \nu)}{(1 - \nu)}\left[1 + (c/a)\frac{\pi(1 + \nu)}{8(1 - \nu)^2}\right] + O(c/a)^2 \tag{16}$$

where the relation $K = 2G(1 + \nu)/3(1 - 2\nu)$ has been used. The critical hardening modulus is negative and comprises a term of the order of the aspect ratio plus a term of the order of $\beta\mu$, both multiplied by the shear modulus for unloading. For the range of β and μ quoted earlier and values of c/a consistent with the approximation the contributions of the two terms are roughly equal. It is apparent from Table 2, in which (16) is evaluated for several values of c/a and $\beta\mu(\nu = 0.2)$, that

TABLE 2. Values of h_r/G for a Cracklike Weakened Zone According to (16), With $\nu = 0.2$

c/a	$\beta\mu = 0.09$	$\beta\mu = 0.18$	$\beta\mu = 0.36$	$\beta\mu = 0.60$
0.01	−0.079	−0.139	−0.261	−0.423
0.05	−0.153	−0.217	−0.346	−0.518
0.10	−0.245	−0.314	−0.452	−0.635

add .4909 $\mu\beta(c/a)$ to each entry

instability is inhibited both by large aspect ratios and by dilatancy and pressure sensitivity. In general, runaway instability for the cracklike weakened zone will occur for values of h equal to small-to-moderate negative fractions of the unloading shear modulus; it may occur just after peak load only for a very narrow zone, if, in addition, the factor for $\beta\mu$ is small.

LOCALIZATION INSTABILITY

In the treatment of the runaway instability the strain in the weakened zone was assumed to remain homogeneous. Another possibility is the development of concentrated, nonuniform deformation in a narrow band within the weakened zone itself. One approach to this problem is that localization may be explained in terms of an instability in the constitutive description of homogeneous deformation. Adopting this approach, *Rice* [1973] outlined a general method for addressing such bifurcations, and RR studied in detail the localization of deformation for constitutive laws intended to model the behavior of brittle rock. It is important to emphasize that this approach addresses the inception of instability in a 'perfect' system: the deformation is idealized as homogeneous up to the point of localization. Although the analysis does not address the growth of the nonuniformity after this point, in many cases the subsequent deformation is so intensely concentrated that localization is synonymous with the onset of rupture. Thus RR identify a limiting material instability in the sense that local nonuniformities, say, in crack content or pore fluid pressure, may cause some localization in the actual system prior to that predicted by the analysis; however, the onset of localization instability does mark the point at which rapid growth of such perturbations will occur. The limiting nature of the localization instability has been discussed in a rigorous fashion by *Rice* [1976].

An alternative and complementary approach to localization is to focus on the growth of initial nonuniformities or favorably oriented microflaws. This approach has been adopted by *Brady* [1974*a*]: conditions are obtained for the growth of a nonuniformity in crack content or 'clusters' of cracks. The condition for failure is then expressed in terms of the stress within these clusters rather than in terms of parameters applying for homogeneous deformation.

The results of RR will be used to investigate whether localization instability in the weakened zone may precede runaway instability. One of the constitutive laws investigated by RR (their equation (10)) is the counterpart of (4) written in a form which is invariant to rigid spins and valid for arbitrary deformation magnitudes. However, under the assumptions of the present analysis their results may be applied directly. The critical value of the hardening modulus at localization has the simple form (RR, equation (20))

$$\frac{h_{cr}}{G} = \frac{(1+\nu)}{9(1-\nu)}(\beta-\mu)^2 - \frac{1+\nu}{2}\left(N + \frac{\beta+\mu}{3}\right)^2 + O(\bar{\tau}/G) \tag{17}$$

where N is the intermediate principal deviatoric stress divided by $\bar{\tau}$. N has the values $-1/3^{1/2}$, 0, and $1/3^{1/2}$ for axisymmetric extension, deviatoric pure shear, and axisymmetric compression, respectively. Since the magnitude of $\bar{\tau}$ is that of a typical stress component, terms of order $\bar{\tau}/G$ are needed only when the critical hardening modulus itself is very small. Equation (17) is plotted schematically as a function of N in Figure 5 (from Figure 6 of RR), and Table 3 (from Table 1 of RR) tabulates h_{cr} for various values of N, β, and μ.

From (17) the maximum value of h_{cr} occurs for $N = -(\beta + \mu)/3$ and is given by

$$h_{max} = G\frac{(1+\nu)}{9(1-\nu)}(\beta-\mu)^2 \tag{18}$$

This equation demonstrates that for brittle rock, for which typically $\beta \neq \mu$, localization instability can occur under rising load. More generally, for states of stress near deviatoric pure shear ($N = 0$), the predicted values of h_{cr} are near zero, and, depending on the values of β and μ, may be slightly positive or negative. However, for axisymmetric compression a hardening modulus equal to a large negative fraction of G is predicted for localization.

The stress state in the weakened zone, characterized by N, depends not only on the applied, far-field stresses but also on the geometry of the weakened zone. For a spherical weakened zone it is apparent from the form of the S_{ijkl} that the stress state will preserve the character of the far-field stress state. However, for the cracklike weakened zone an examination of the asymptotic form of the shape factor or of the penny-shaped crack solution reveals that the deformation is constrained essentially to combinations of simple shear and uniaxial compression. For such a mode of deformation the conditions for localization are met at $h = 0$, and consequently, localization is predicted near peak stress for very narrow weakened zones.

Effects of Yield Surface Vertices and Anisotropy

The formulation of the constitutive law, (4), assumed that the subsequent yield surfaces in stress space retain the shape of the initial yield surface or 'harden isotropically.' Using a model of brittle rock as sliding on a collection of randomly oriented fissures, RR have shown that subsequent yield surfaces tend to form a pointed vertex at the current stress point and that the isotropic hardening idealization will be inadequate in describing the stiffness of response to abrupt changes in the pattern of straining. Although a modification of the isotropic hardening constitutive law, proposed by RR to approximate the response at a vertex, considerably complicates the localization calculation, its essential effect is to reduce the magnitude of h_{cr}/G as shown schematically in Figure 5. Localization for states of stress near deviatoric pure shear occurs closer to peak stress, but the character of the

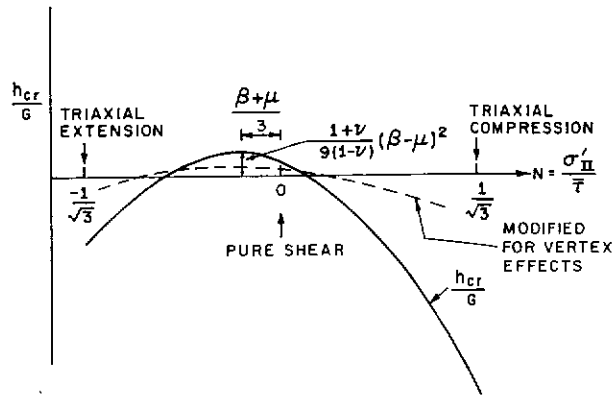

Fig. 5. From Figure 6 of *Rudnicki and Rice* [1975]. Schematic shows the variation with stress state of the critical hardening modulus. Solid line indicates predictions based on the constitutive law (equation (4)). Dashed line shows predictions modified for vertex effects. Drawing is not to scale.

predictions based on the constitutive law (4) is preserved. However, for axisymmetric compression the predicted values of h_{cr} are substantially less negative than those predicted by (17). Nevertheless, negative values of magnitude ranging up to roughly 0.1 G are still predicted for localization.

Another feature which is not included in the constitutive law (4) is the anisotropy which develops, even in samples which are isotropic when unstressed, owing to the preferential growth of microcracks in a direction perpendicular to the greatest principal stress (positive in tension) [*Wawersik and Brace*, 1971; *Tapponnier and Brace*, 1976]. Although the localization analysis is again complicated by the inclusion of anisotropy, preliminary studies for transverse isotropy, intended to model the developed anisotropy in an axisymmetric specimen, indicate that the effect is similar to that of vertex formation: the magnitude of h_{cr}/G is reduced, and the reduction is much greater for axisymmetric deformation states. This anisotropy may be responsible for the fact that the values measured by *Mogi* [1967] for the angle between the normal to the plane of localization and the direction of the least principal stress were consistently greater than those predicted by RR [*Cleary and Rudnicki*, 1976].

Because the localization instability is sensitive to details of the constitutive modeling (RR), particularly for axisymmetric stress states, the effects of yield surface vertices and anisotropy are most important here. However, for the runaway instability these effects are not pronounced and may be taken into account approximately simply by regarding the moduli which enter (11) as average values.

Localization in Axisymmetric Compression Experiments

It is a frequent observation that the failure of brittle laboratory specimens in axisymmetric compression involves localization, at least in a gross sense: relatively random microcracking progresses to a clustering of microcracks from which develops a more or less distinct failure plane. Because the analysis of RR idealizes the deformation as homogeneous until the point of localization, at which the deformation is intensely concentrated, there is inevitable arbitrariness about what constitutes localization. However, the limiting nature of the instability indicates that the sensible interpretation for the onset of localization instability is the appearance of a macroscopic fault in the sample.

Although many studies have shown that only slight changes occur in the rock structure until very near the peak stress [e.g., *Brace*, 1964; *Scholz*, 1968; *Friedman et al.*, 1970], precise investigations of when faulting does occur are complicated by instabilities induced by the testing machine. If the postpeak stress-strain curve descends more rapidly than the unloading characteristic for the testing machine (Figure 6), the sample will 'run away,' owing to the stored elastic energy in the machine, until the point of localization instability is reached. Hence apparent faulting will be identified at the point of runaway, even though the actual material or localization instability may have occurred much later.

Nevertheless, experiments using stiff-testing machines with rapid unloading capacity [*Wawersik and Fairhurst*, 1970; *Crouch*, 1970; *Rummel and Fairhurst*, 1970; *Wawersik and Brace*, 1971] indicate that faulting does not occur until after peak stress and, for at least some rock types, not until very much past peak stress. For Westerly granite this is evidenced most dramatically in a reexamination of the samples of *Wawersik and Brace* [1971] by *Tapponnier and Brace* [1976]; comparison of the comments of the latter authors in their Table 1 with the schematic stress-strain diagram of their Figure 2 indicates that faulting did not occur until roughly one third of the way down the descending portion of the curve. Figure 4 of *Crouch* [1970] for a sandstone clearly indicates that faulting occurred at the break in the curve at a point where the stress had decreased to less than one half of the peak value. Furthermore, a careful examination of this work on postpeak deformation suggests that microcracking often occurs in isolated events until well after peak stress and that some of the inhomogeneity of microcracking which does exist prior to faulting is attributable to the end constraints. For example, Plate 2*b* of *Wawersik and Brace* [1971] and Figure 7 of *Rummel and Fairhurst* [1970] suggest that deformation sufficiently diffuse to

TABLE 3. Values of h_{cr}/G at Instability for Various Stress States

μ	β	Axially Symmetric Extension ($N = -1/3^{1/2}$)	Maximum Value ($N = -(\mu + \beta)/3$)	Pure Shear ($N = 0$)	Axially Symmetric Compression ($N = 1/3^{1/2}$)
0	0	−0.200 (50.8)	0 (45.0)	0 (45.0)	−0.200 (39.2)
0.3	0	−0.122 (54.9)	0.015 (49.3)	0.009 (48.4)	−0.260 (43.2)
0.3	0.15	−0.106 (57.0)	0.004 (51.6)	−0.010 (50.2)	−0.314 (45.2)
0.3	0.3	−0.085 (59.2)	0 (53.9)	−0.024 (51.9)	−0.362 (47.2)
0.6	0	−0.025 (59.2)	0.060 (53.9)	0.036 (51.9)	−0.303 (47.2)
0.6	0.15	−0.031 (61.6)	0.034 (56.3)	−0.004 (53.7)	−0.377 (49.2)
0.6	0.3	−0.031 (64.0)	0.015 (58.9)	−0.039 (55.6)	−0.423 (90.0)*
0.6	0.45	−0.027 (66.6)	0.004 (61.7)	−0.070 (57.4)	−0.385 (90.0)*
0.6	0.6	−0.019 (69.5)	0 (64.9)	−0.096 (59.3)	−0.342 (90.0)*
0.9	0	0.089 (64.0)	0.135 (58.9)	0.081 (55.6)	−0.303 (90.0)*
0.9	0.15	0.063 (66.6)	0.107 (61.7)	0.020 (57.4)	−0.295 (90.0)*
0.9	0.3	0.041 (69.5)	0.060 (64.9)	0.036 (59.3)	−0.282 (90.0)*
0.9	0.45	0.024 (72.7)	0.034 (68.6)	0.088 (61.3)	−0.264 (90.0)*
0.9	0.6	0.011 (76.6)	0.015 (73.2)	0.135 (63.4)	−0.242 (90.0)*

From *Rudnicki and Rice* [1975, Table 1]; recalculated for $\nu = 0.2$. θ_0 in degrees is given in parentheses. Calculations are based on the constitutive law (4). $N = \sigma_2'/\bar{\tau}$, where $\sigma_1' \geq \sigma_2' \geq \sigma_3'$ are the principal deviatoric stresses.

*For these cases, h_{cr} is given with $N_{\min} = \sigma_3'/\bar{\tau}$ substituted for N.

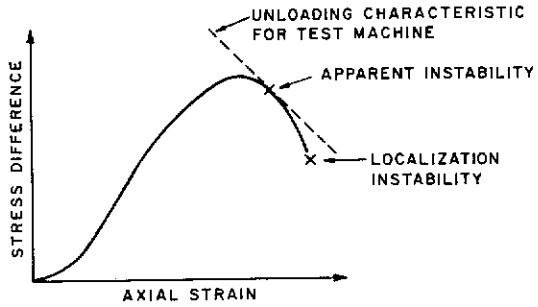

Fig. 6. Schematic showing the runaway instability caused when the material stress-strain curve descends more rapidly than the unloading characteristic for the testing machine.

be considered homogeneous for the purpose of this analysis may persist well into the postpeak regime, although clustering of cracks inevitably occurred as the point of faulting was approached. Although Figure 7*b* of Rummel and Fairhurst indicates a large crack emanating from the lower left boundary of the sample, there is only isolated cracking in the interior, and it is likely that the larger fracture was due to end effects rather than intrinsic material instability. Indeed, the analysis of RR suggests that localization instabilities can occur earlier under plane strain than axisymmetric deformation fields, and the constraint at the ends of a specimen forces a strain field that is not only much larger in magnitude by comparison with the average deformation but also much closer to a plane strain–like state.

Ideally, one would like to study fault formation by monitoring the distribution of strain in the sample, but there are obvious experimental difficulties with such a procedure. However, one promising method is the acoustic emission technique used by *Lockner and Byerlee* [1975] to study a granite and a sandstone under 1 kbar of confining pressure. Although they reported a clustering of acoustic emission events prior to peak stress for one sandstone sample, another sample, deformed at a much slower rate, showed localization only in the postpeak regime. For the granite there was no evidence of localization prior to peak stress.

The complexity of brittle rock failure is well documented, and it has been demonstrated by RR that predictions of localization are very sensitive to the precise structure of the constitutive law. Nevertheless, although some localization in laboratory samples may begin near peak stress, the predicted values of h_{cr} are consistent with the observations cited here for which faulting, that is, localization in the sense of limiting material instability, was not prematurely induced by the loading machine. Furthermore, although the discussion here has been limited to axisymmetric compression experiments, *Cleary and Rudnicki* [1976] have pointed out that the variation with N of predicted values of h_{cr} (Figure 5) is consistent with that of the slopes at failure observed by *Mogi* [1971]. These observations support the assertion that localization can be modeled as an instability in the constitutive description of homogeneous deformation, and it is recommended that this point of view merits greater attention in the study of brittle rock failure. At the same time, it must be recognized that details of the constitutive law, on which the predictions of localization depend, are the least well-known aspect of the theory.

Comparison of the Instabilities

Conditions have been derived for the onset of the runaway instability and the localization instability, and these have been examined for the two limiting cases of a spherical weakened zone and a flat cracklike weakened zone. For a spherical weakened zone the onset of the runaway instability (equation (12)), even under conditions most favorable to instability (e.g., $\mu = \beta = 0$ and κ equals its maximum value (0.533)), is predicted for values of the hardening modulus which are negative and of magnitude equal to approximately half the unloading modulus. On the other hand, the onset of the localization instability for this geometry depends primarily on the nature of the far-field stress state through the parameter N (equation (17)). Predictions for the onset of localization, even under the most severe circumstances, that is, predictions for axisymmetric compression based on the smooth yield surface idealization (Table 3), indicate that localization will precede runaway ($h_{cr} > h_r$) for the spherical weakened zone. Although for very narrow weakened zones, runaway instability is predicted for $h_r \sim -(\mu\beta + c/a)G$ (equation (16)), it was argued that the localization instability, owing to the constraint of the weakened zone deformation state, will occur near $h_{cr} = 0$. Therefore localization will occur, in general, prior to runaway, although the onset of the instabilities may be competitive for very narrow weakened zone shapes.

The condition for the runaway instability was that a remote boundary strain increment could induce an unbounded strain increment in the weakened zone. This condition indicates the absence of any static solution for the next increment of deformation. The onset of this instability obviously entails a sudden loss of load-carrying capacity for the rock mass and corresponds to the onset of a seismic instability.

Localization of deformation has been treated with the use of the results of RR. For stress states near that characterized by $N = -(\beta + \mu)/3$, localization may occur under rising load. Such deformation may be associated with creep events. However, the concentration of straining in the localized zone will likely drive the stress here past peak, and it is doubtful that conditions of rising load will prevail very long in the localized zone. More typically, localization will occur under falling load, and in this case the process may accelerate and increasingly concentrate the deformation. For example, if the weakened zone is roughly spherical, localization results in a planar region of concentrated deformation that constitutes a further weakened zone. At instability for the spherical zone the critical value of h for the very narrow zone, not only for localization but also for runaway, has already been attained. Hence further, rapid concentration of deformation and transition to a dynamic or seismic instability are expected. Although the analysis specifically addresses only the inception of instability, the results suggest the possibility of acceleration of localization as a seismic mechanism.

Coupled Deformation–Fluid Diffusion Effects

The preceding analysis of conditions for instability has been carried out as if the rock mass were completely dry. If the rock mass is fluid infiltrated and if the time scale of stress alteration induces significant local pore pressure changes, the analysis can be considerably complicated by the coupling of the deformation to the diffusion of pore fluid. Because of the inhomogeneity of material properties, which has been idealized as an ellipsoidal weakened zone, the pore pressure changes induced by rapid stress alteration will be nonuniform at the outset. *Rice and Cleary* [1976] have identified two ways in which coupled stress-diffusion can contribute to the stabilization of a shear fault, and these may be used to assess qualitatively the effects on the instabilities.

The first mechanism concerns the alteration of the effective values of elastic moduli according to whether the material responds in a 'drained' (pore fluid pressure is a constant) or 'undrained' (fluid mass content is a constant in each elemental volume) fashion. In particular, *Rice and Cleary* [1976] have shown that the undrained Poisson's ratio ν_u is in the range $\nu \leq \nu_u \leq 0.5$, where ν is the drained value and the upper limit is attained if the solid and fluid constituents are separately incompressible. The shear modulus, of course, remains the same; the undrained bulk modulus is accordingly elevated in comparison to the drained value. Hence the elastic response is stiffer for changes in load which are rapid in comparison to the time scale of fluid diffusion.

Alteration of the apparent value of the modulus governing inelastic deformation will also result from stress-diffusion effects and dilatancy. If new pore space is created at a rate exceeding the influx of pore fluid, pore suctions will be induced. This will cause an increase in the 'effective' compressive stress $(\sigma - p)$, even when the total stress is constant. Further inelastic deformation due to frictional sliding and microcracking will be inhibited, and the rock is said to be 'dilatantly hardened.' In terms of the constitutive law (4) the apparent value of the hardening modulus is increased over the value for drained behavior, and it may, in fact, be positive while the corresponding value for the underlying drained deformation is negative. Dilatant hardening has been observed in laboratory specimens by *Brace and Martin* [1968] and has been studied by *Rice* [1975] for the special deformation state of simple shear. *Rice* [1975] showed that the amount of dilatant strengthening which can be effected is limited by diffusive instability: when the condition for localization instability, (17), phrased in terms of the hardening modulus for drained behavior, is met, local nonuniformities in the amount of shear tend to grow rather than decay.

As is shown by the analyses of *Rice* [1975] and *Rice and Cleary* [1976], the essential effect of coupled deformation–pore fluid diffusion is to introduce a time dependence into the material response. Thus it is plausible that the onset of the runaway instability will be predicted approximately by (11) if the weakened zone pore fluid pressure is included as an additional hydrostatic stress component and if the moduli entering (11) are the effective values. The way in which the action of pore fluid diffusion may stabilize runaway instability is evident: the condition for instability may be met in terms of the drained moduli, but its onset will be delayed until the moduli relax from their undrained values. Similarly, the progress of diffusive instability, which *Rice* [1975] has shown to set in when the localization condition is met in terms of the drained moduli, will be governed by the rate of pore fluid diffusion. Thus although advocates of dry crack or diffusionless models for earthquake precursors maintain that pore fluid diffusion is not required to explain the observed premonitory effects, if pore fluid is present, diffusion will set a time scale of the instability. Nevertheless, it has not yet been established that this effect is of sufficient magnitude and spatial extent to account for precursory events, as is proposed by dilatancy-diffusion proponents.

APPLICATIONS TO EARTH FAULTING

The analysis has shown that only for weakened zone deformation states near $N = -(\beta + \mu)/3$ (Table 3) can instability occur under rising load ($h > 0$). Although the deformation in very narrow weakened zones will be constrained to such states, falling stress-strain curves are, in general, a necessary feature of modeling processes leading to faulting. Data from laboratory axisymmetric compression tests using stiff-unloading machines support this conclusion, as it is a frequent observation that faulting does not occur until postpeak and that for at least some types of brittle rock, faulting does not occur until relatively late in the region of descending stress-strain curves. Although *Walsh* [1971] suggested, on the basis of a comparison of the relative stiffness of a loading apparatus versus that of the surrounding material for a fault, that descending stress-strain curves could not be realized in situ, the calculation assumed elastic properties and slip on a cracklike surface. This conclusion is consistent with the analysis here for the special case of a cracklike weakened zone but is not applicable for other shapes of weakened zones.

More generally, for most weakened zone shapes the localization instability will precede the runaway instability, although their onset may be competitive for very narrow weakened zones. Thus for a given weakened zone shape the amount of postpeak deformation prior to instability depends on the stress state, at least if the weakened zone is not extremely narrow. For deviatoric states of stress near pure shear ($N = 0$), localization is predicted near peak stress as compared with states of axisymmetric compression ($N = 1/3^{1/2}$). For faulting under different states of stress these results suggest differences in conditions prior to rupture and consequently in precursory events. In particular, it is possible that strike slip and thrust type faults are associated with states of stress that are more nearly deviatoric pure shear and axisymmetric compression, respectively.

The predictions concerning the amount of postpeak deformation prior to instability have a particularly clear consequence for the dry crack models of *Myachkin et al.* [1972, 1974] and *Stuart* [1974] which associate the closing of cracks and consequently the return of anomalous wave speed ratios to normal values with the descending portion of the stress-strain curve. Thus for cases in which the onset of instability has been predicted near peak load the earthquake will occur, according to these models, when the wave speed ratio is still anomalously low.

Although it is implicit in these dry crack models that crack closure is associated with descending stress-strain curves in a weakened zone, the experimental evidence is to the contrary. *Thill* [1972] and *Rummel* [1974] observed continued decreases in the compressional wave speed during deformation in the postpeak regime, and *Crouch* [1970] observed continued dilatancy in this regime. Although there is disagreement about whether the conditions of these experiments adequately reflect the situation in the field, it may be more appropriate to associate the descending portion of the curve not with the closing of cracks but rather with the formation of the anomalous zone, that is, the region exhibiting significant dilatancy. In this case, crack closure outside of the zone of localization may be caused by the acceleration of localization within the anomalous zone. This possibility seems to be similar to that suggested by *Myachkin et al.* [1974], although they were not specific about the mechanism involved and they evidently envisioned crack closure as beginning near peak stress. However, a more detailed analysis which considers the evolution of instability is needed to evaluate this mechanism.

The analysis also suggests an interpretation of the dilatancy-diffusion hypothesis. As was noted by *Rice* [1975], if the rock mass is fluid infiltrated, dilatant hardening (and elastic stiffening) in an initially very narrow weakened zone will cause adjacent regions to sustain greater stress and consequently to undergo inelastic, dilatant deformation. This is one mecha-

nism by which a spatial extent of dilatancy sufficient to cause observable travel time anomalies might be achieved. Because the extent of dilatant hardening is limited by the localization instability, the total dilatant volume achievable by this process will depend on the amount of postpeak deformation prior to instability. Thus according to the analysis here, the extent of the anomalous zone will depend on the deviatoric state of stress in the weakened zone and consequently may be consistently different for strike slip and thrust type faults. *Aggarwal et al.* [1975] reported that for thrust type earthquakes in New York state the characteristic length of the anomalous region was 6–10 times that of the rupture zone, while *Robinson et al.* [1974] suggested that the characteristic lengths of anomalous zones associated with strike slip earthquakes in California are at most 2–3 times the largest dimension of the rupture zone. However, since some of the conclusions of *Robinson et al.* [1974] have been revised but not yet published (anonymous reviewer, private communication, 1976), these observations are at most suggestive. Nevertheless, they do indicate a direction for further field measurements.

After the onset of localization instability the deformation will be concentrated in a narrow rupture zone within the anomalous region. This would explain the apparent discrepancy between the narrowness of zones of fault gouge due to past ruptures and the spatial extent of the inelastic deformation and dilatancy required to cause observable wave travel time anomalies. The progression of instability, however, will be limited by the time scale of pore fluid diffusion as moduli relax from their undrained values. This interpretation of the dilatancy-diffusion hypothesis is certainly an extrapolation of the results of the present model. More detailed modeling which includes coupled deformation-diffusion effects and considers not only the inception but also the evolution of the instability is necessary for a more precise evaluation. Indeed, the questions of whether dilatant hardening and diffusion processes can stabilize a rock mass against a rapid failure, cause dilatancy to achieve a spatial extent sufficient to generate observable anomalies in stress wave travel times, and permit a progression of failure on a time scale comparable to that observed for the premonitory period are central to the dilatancy-diffusion hypotheses and have not as yet been answered.

Pore fluid diffusion has been suggested as one mechanism which may account for the characteristic time dependence of earthquake precursors. Time-dependent crack closure, though it is not explicitly included in the present model, is another possibility. Although there is no apparent time scale associated with this process in the dry crack models of *Myachkin et al.* [1972, 1974] and *Stuart* [1974], *Brady* [1974*a*] has identified a characteristic time for crack closure. Under certain assumptions this time scale is formally the same as that associated with diffusion processes, and preliminary experimental verification of its existence has been obtained by *Brady* [1976].

CONCLUSION

Although the results provide no direct evidence for or against existing earthquake precursor models, they suggest new interpretations for dry crack and dilatancy-diffusion hypotheses in terms of constitutive behavior of brittle rock which is consistent with laboratory experiments. These results are based on a new approach to brittle rock failure as a constitutive instability and the idealization of material inhomogeneities near a fault as a single weakened zone. Although the idealizations considerably simplify the immense complexity of actual tectonic processes, the two distinct instabilities which have been identified represent possible models for the onset of seismic instability. Furthermore, it has been demonstrated that conditions prior to rupture depend strongly on the state of stress and the nature of material inhomogeneities. This suggests that premonitory events will also be sensitive to these features. More specifically, it is anticipated that premonitory events will be different for strike slip and thrust type faults. This prediction may be testable as more detailed field evidence becomes available.

APPENDIX: VALUES OF S_{ijkl} FOR THE GENERAL ELLIPSOID

The factors S_{ijkl} are symmetric in ij and kl, but in general, $S_{ijkl} \neq S_{klij}$. In addition, if the coordinate axes coincide with the principal axes of the ellipsoid, the S_{ijkl} possess the convenient property that the only nonzero entries have the form S_{iiii}, S_{iijj}, or S_{ijij} (no sum here on i or j). Expressions for the S_{ijkl} have been given by *Eshelby* [1957]. The ellipsoid has semiaxes $a > b > c$, with the orientation as in Figure 1.

$$8\pi(1-\nu)S_{1111} = 3a^2I_{aa} + (1-2\nu)I_a$$

$$8\pi(1-\nu)S_{1122} = 3b^2I_{ab} - (1-2\nu)I_a$$

$$8\pi(1-\nu)S_{1212} = (\tfrac{3}{2})(a^2+b^2)I_{ab} + (1-2\nu)(I_a+I_b)/2$$

where

$$I_a = 2\pi abc\int_0^\infty \frac{du}{(a^2+u)\Delta}$$

$$I_{aa} = 2\pi abc\int_0^\infty \frac{du}{(a^2+u)^2\Delta} \qquad \text{(A1)}$$

$$I_{ab} = \tfrac{2}{3}\pi abc\int_0^\infty \frac{du}{(a^2+u)(b^2+u)\Delta} \qquad \text{(A2)}$$

$$\Delta = [(a^2+u)(b^2+u)(c^2+u)]^{1/2}$$

The remaining quantities may be obtained by cyclic permutation. Expressions for I_a, I_b, and I_c may be written alternatively as

$$I_a = \frac{4\pi abc}{(a^2-b^2)(a^2-c^2)^{1/2}}(F-E)$$

$$I_c = \frac{4\pi abc}{(b^2-c^2)(a^2-c^2)^{1/2}}\left\{\frac{b(a^2-c^2)^{1/2}}{ac} - E\right\}$$

$$I_b = 4\pi - I_a - I_c$$

where $F = F(\theta, k)$ and $E = E(\theta, k)$ are elliptic integrals of the first and second kinds with amplitude $\theta = \sin^{-1}(1-c^2/a^2)^{1/2}$ and modulus $k = (a^2-b^2)^{1/2}/(a^2-c^2)^{1/2}$. The remaining quantities may be obtained by the relations

$$I_{ab} = \frac{(I_b - I_a)}{3(a^2-b^2)} \qquad I_{aa} = 4\pi/3a^2 - I_{ab} - I_{ac} \qquad \text{(A3)}$$

In the case of an oblate spheroid ($a = b > c$), (A3) fails, but from (A1) and (A2) it is clear that $I_{ab} = I_{aa}/3$. The other factors are then computed straightforwardly. A similar manipulation is possible for the prolate spheroid ($a > b = c$). Finally, we record that for the oblate spheroid

$$I_a = I_b = \frac{2\pi a^2 c}{(a^2 - c^2)^{3/2}} \left\{ \cos^{-1}(c/a) - \frac{c}{a} \left(1 - \frac{c^2}{a^2} \right)^{1/2} \right\}$$

and for the prolate spheroid

$$I_b = I_c = \frac{2\pi a c^2}{(a^2 - c^2)^{3/2}} \left\{ \frac{a}{c} \left(\frac{a^2}{c^2} - 1 \right)^{1/2} - \cosh^{-1}(a/c) \right\}$$

Acknowledgments. The author is grateful to J. R. Rice for suggesting this research and for his helpful comments during the course of the investigation. Useful discussions with M. P. Cleary and W. D. Stuart are also gratefully acknowledged. This study was supported by the NSF Geophysics Program under grant EAR74-13697 A01 to Brown University.

REFERENCES

Aggarwal, Y. P., L. R. Sykes, D. W. Simpson, and P. G. Richards, Spatial and temporal variations in t_s/t_p and P wave residuals at Blue Mountain Lake, New York: Application to earthquake prediction, *J. Geophys. Res., 80,* 718–732, 1975.

Bieniawski, Z. T., Mechanism of brittle fracture of rock, 2, Experimental studies, *Int. J. Rock Mech. Mining Sci., 4,* 407–423, 1967.

Brace, W. F., Brittle fracture in rocks, in *State of Stress in the Earth's Crust*, edited by W. R. Judd, pp. 111–174, Elsevier, New York, 1964.

Brace, W. F., and R. J. Martin III, A test of the law of effective stress for crystalline rocks of low porosity, *Int. J. Rock Mech. Mining Sci., 5,* 415–426, 1968.

Brace, W. F., B. W. Paulding, Jr., and C. H. Scholz, Dilatancy in the fracture of crystalline rocks, *J. Geophys. Res., 71,* 3939–3953, 1966.

Brady, B. T., Theory of earthquakes, I, A scale independent theory of failure, *Pure Appl. Geophys., 112,* 701–726, 1974*a*.

Brady, B. T., Seismic precursors before rock failures in mines, *Nature, 252,* 549–552, 1974*b*.

Brady, B. T., Theory of earthquakes, 2, Inclusion theory of crustal earthquakes, *Pure Appl. Geophys., 113,* 149–168, 1975*a*.

Brady, B. T., Comment on 'Diffusionless dilating model for earthquake precursors,' by W. D. Stuart, *Geophys. Res. Lett., 2,* 259–261, 1975*b*.

Brady, B. T., Laboratory investigation of tilt and seismicity anomalies in rock before failure, *Nature, 260,* 108–111, 1976.

Cleary, M. P., and J. W. Rudnicki, The initiation and propagation of dilatant rupture zones in geological materials, in *The Effect of Voids on Material Deformation,* vol. 16, edited by S. C. Cowen, pp. 13–30, American Society of Mechanical Engineers, New York, 1976.

Crouch, S. L., Experimental determinations of volumetric strains in failed rock, *Int. J. Rock Mech. Mining Sci., 7,* 589–603, 1970.

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.

Friedman, M., R. D. Perkins, and S. J. Green, Observation of brittle deformation features at the maximum stress of Westerly granite and Solenhofen limestone, *Int. J. Rock Mech. Mining Sci., 7,* 297–306, 1970.

Haimson, B. C., Mechanical behavior of rock under cyclic loading, in *Proceedings of the Third Congress of the International Society of Rock Mechanics,* vol. 2, part A, pp. 373–378, National Academy of Sciences, Washington, D. C., 1974.

Hobbs, D. W., The behavior of broken rock under triaxial compression, *Int. J. Rock Mech. Mining Sci., 7,* 125–148, 1970.

Jaeger, J. C., and N. G. W. Cook, *Fundamentals of Rock Mechanics,* p. 145, Methuen, London, 1969.

Lee, I. K., O. G. Ingles, and R. C. Neil, Elastic and inelastic deformations of a sandstone and a stabilized soil under precise strain control, *Uniciv Rep. R-88,* Univ. of New South Wales, Kensington, N. S. W., Australia, 1972.

Lockner, D., and J. Byerlee, Acoustic emission and fault formation in rocks, in *Proceedings of the Conference on Acoustic Emissions,* Pennsylvania State University, University Park, in press, 1975.

Mogi, K., Some precise measurements of fracture strength of rocks under uniform compressive stress, *Rock Mech. Eng. Geol., 4,* 41–55, 1966.

Mogi, K., Effect of the intermediate principal stress in rock failure, *J. Geophys. Res., 72,* 5117–5131, 1967.

Mogi, K., Effect of triaxial stress system on the failure of dolomite and limestone, *Tectonophysics, 11,* 111–127, 1971.

Myachkin, V. I., G. A. Sobolev, N. A. Dolbilkina, V. N. Morozow, and V. B. Preobrazensky, The study of variations in geophysical fields near focal zones of Kamchatka, *Tectonophysics; 14,* 287–293, 1972.

Myachkin, V. I., B. V. Kostrov, G. A. Sobolev, and O. G. Shamina, Experimental and theoretical investigations of processes that are possible forerunners of earthquakes, *Izv. Acad. Sci. USSR Phys. Solid Earth,* 676–679, 1974.

Myachkin, V. I., W. F. Brace, G. A. Sobolev, and J. H. Dieterich, Two models for earthquake forerunners, *Pure Appl. Geophys., 113,* 169–181, 1975.

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

Rice, J. R., The initiation and growth of shear bands, in *Plasticity and Soil Mechanics,* edited by A. C. Palmer, pp. 263–274, Engineering Department, Cambridge University, Cambridge, England, 1973.

Rice, J. R., Continuum mechanics and thermodynamics of plasticity in relation to microscale deformation mechanisms, in *Constitutive Relations in Plasticity,* edited by A. S. Argon, pp. 23–79, MIT Press, Cambridge, Mass., 1974.

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, in *Proceedings of the 14th International Congress of Theoretical and Applied Mechanics,* vol. 1, edited by W. T. Koiter, pp. 207–220, North-Holland, Amsterdam, 1976.

Rice, J. R., and M. P. Cleary, Some basic stress diffusion solutions for fluid-saturated elastic porous media with compressible constituents, *Rev. Geophys. Space Phys., 14,* 227–241, 1976.

Robinson, R., R. L. Wesson, and W. L. Ellsworth, Variation of P wave velocity before the Bear Valley, California, earthquake of 24 February 1972, *Science, 184,* 1281–1283, 1974.

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., Changes in the P-wave velocity with increasing inelastic deformation in rock specimens under compression, in *Proceedings of the Third Congress of the International Society of Rock Mechanics,* vol. 2, part A, pp. 517–523, National Academy of Sciences, Washington, D. C., 1974.

Rummel, F., and C. Fairhurst, Determination of the post-failure behavior of brittle rock using a servo-controlled testing machine, *Rock Mech., 2,* 189–204, 1970.

Schock, R. N., H. C. Heard, and D. R. Stephens, Stress-strain behavior of a granodiorite and two graywackes in compression to 20 kilobars, *J. Geophys. Res., 78,* 5922–5941, 1973.

Scholz, C. H., Microfacturing and the inelastic deformation of rock in compression, *J. Geophys. Res., 73,* 1417–1432, 1968.

Scholz, C. H., and R. Kranz, Notes on dilatancy recovery, *J. Geophys. Res., 79,* 2132, 1974.

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

Sih, G. C., and H. Liebowitz, Mathematical theories of brittle fracture, in *Treatise on Fracture,* vol. 2, pp. 67–190, Academic, New York, 1968.

Stuart, W. D., Diffusionless dilatancy model for earthquake precursors, *Geophys. Res. Lett., 1,* 261–264, 1974.

Stuart, W. D., Reply, *Geophys. Res. Lett., 2,* 263–264, 1975.

Swanson, S. R., and W. S. Brown, The mechanical response of prefractured rock in compression, *Rock Mech., 3,* 208–216, 1971.

Tapponnier, P., and W. F. Brace, Development of stress-induced microcracks in Westerly granite, *Int. J. Rock Mech. Mining Sci. Geomechanics Abstr., 13,* 103–112, 1976.

Thill, R. E., Acoustic methods for monitoring failure in rock, in *New Horizons in Rock Mechanics, 14th Symposium on Rock Mechanics,* Pennsylvania State University, University Park, 1972.

Walsh, J. B., Stiffness in faulting and in friction experiments, *J. Geophys. Res., 76,* 8597–8598, 1971.

Wawersik, W. R., and W. F. Brace, Post failure behavior of a granite and a diabase, *Rock Mech., 3,* 61–85, 1971.

Wawersik, W. R., and C. Fairhurst, A study of brittle rock fracture in laboratory compression experiments, *Int. J. Rock Mech. Mining Sci., 7,* 561–575, 1970.

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.

(Received June 4, 1976;
revised October 4, 1976;
accepted October 6, 1976.)