# Diffusive void bifurcation in stressed solid

Z. Suo and W. Wang
*Mechanical and Environmental Engineering Department, Materials Department, University of California, Santa Barbara, California 93106*

(Received 4 January 1994; accepted for publication 1 June 1994)

Interconnects are susceptible to solid diffusion under residual stress, electric current, and elevated temperature. As atoms diffuse, voids nucleate, drift, and enlarge. At some point, the voids of rounded shape can collapse to narrow slits and sever the interconnects. The fatal slits are often found to be transgranular, i.e., each slit cuts across a single grain. They have raised many concerns, but the underlying mechanism has remained unclear. It is proposed that a void changes shape due to surface diffusion under the combined action of surface energy, elastic energy, and electric current. The void will be rounded if surface energy prevails, but will collapse to a slit if the elastic energy or the electric current prevails. A cylindrical void in an infinite crystal under biaxial stresses, but under no electric current, is analyzed. Four things are done, as follows: (1) A suitable thermodynamic potential is minimized and maximized to select, among a family of ellipses, equilibrium void shapes. The bifurcation diagram consists of a subcritical pitchfork and two Griffith cracks. (2) A void under biased stresses is analyzed to illustrate the effect of imperfections. (3) Exact initial bifurcation modes are determined. The critical loads for the successive modes are closely separated, indicating that the shape evolution will be sensitive to initial imperfections. (4) A variational principle for shape evolution under stress, current and surface energy is identified. Stress-induced evolution time is estimated by using this variational principle.

## I. INTRODUCTION

Making reliable on-chip interconnects has been a persistent challenge as integrated circuits evolve. Present-day interconnects are made of aluminum alloys, and are less than a few microns in width. Submicron lines have already been used in some products. Copper interconnects are being explored. In this article, data for pure aluminum are quoted to illustrate various points, but the physical processes apply to any metals. The interconnects operate under severe conditions: high stress, intense current, and temperature exceeding one-third of the melting point (933 K for aluminum). Diffusion-mediated degradation is ubiquitous as the brute forces act in the small dimensions.[1,2]

The stress results from the mismatch in thermal-expansion coefficients of the metal lines and the surrounding insulators. Pure aluminum in bulk has low-yield strength, below 100 MPa at room temperature, and usually is not under high stress; yet, high stress prevails in fine lines constrained by stiff insulators. The stress is raised in two ways. First, the thermal-expansion mismatch results in a triaxial stress that cannot be fully relieved by plastic flow; for typical thermal history, tensile stress around 400 MPa is found by finite element calculations and x-ray measurements.[3–5] Second, in the small dimensions dislocations are severely bent and can only move under high stress; even without triaxial constraint the stress in thin films may exceed 200 MPa.[6]

In addition to the stress, the interconnects carry intense electric current, sometimes exceeding $10^{10}$ A/m$^2$. Both stress and current cause atoms to diffuse, known respectively as stressmigration and electromigration. Evidence has recently accumulated that narrow, transgranular slits can form and sever the lines.[7–9] The sequence of the events has also been revealed: A round void nucleates first, enlarging and drifting, and then collapses to a narrow slit.[10] Since forming a slit transfers much less mass than growing a rounded void across the linewidth, a slit can significantly reduce the interconnect lifetime. Consequently, the transgranular slits have raised much concerns, even though they may not be as prevalent as rounded voids, as judged from their less frequent appearance in the published micrographs.

In a previous article, we have shown how electric current alone can cause the shape instability;[11] yet, most slits form under both electric current and thermal stress. In this article we focus on the role played by the stress under no electric current. Instability under combined stress and current will be analyzed elsewhere using some of the methods developed here. The interconnects operate in such a temperature range that, within the time of interest, ample atoms diffuse on the void surface but negligible atoms diffuse in the lattice. The void is assumed to reside inside a perfect grain so that grain boundaries are inaccessible for diffusion. Creep is assumed to be slow compared to surface diffusion and therefore neglected. Also neglected is instantaneous dislocation glide, which seems to be a reasonable first approximation, given the high stress in the interconnects. As such, surface diffusion is the only dissipative process included in this analysis.

As diffusion varies the void shape, the solid varies energy by either varying the elastic field or creating the surface. The instability of the void shape is an outcome of the competition between the variation in the elastic and the surface energy. Figure 1 illustrates a small cylindrical void in an elastic solid under biaxial stresses. The two-dimensional problem conveys the essence of the competition; the three-dimensional version will be treated elsewhere. Focus on the problem of perfect symmetry: a circular void in an infinite isotropic solid under biaxial stresses $\sigma_1=\sigma_2=\sigma$. The perfect circle is obviously an equilibrium shape: Nothing is unbal-

 0021-8979/94/76(6)/3410/12/$6.00 

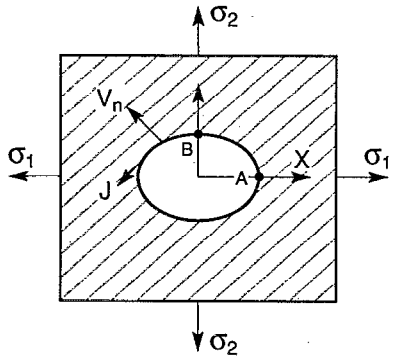



FIG. 1. The cross section of a cylindrical void in an elastic solid under biaxial stresses.

anced to drive surface diffusion; however, this equilibrium becomes unstable if the stress is high, as discussed below. Imagine a void perturbed from the circular shape, say, an ellipse in Fig. 1. Now both stress and surface energy drive the atoms to diffuse on the void surface, but in the opposite directions. Let $K$ be the curvature of, and $w$ the elastic energy density on, the void surface. Because $K_A > K_B$, the surface energy strives to move atoms from $B$ to $A$ and restore the circular symmetry. Because $w_A > w_B$, the elastic energy strives to move atoms from $A$ to $B$ and amplify the asymmetry. The void collapses if the elastic energy prevails over the surface energy.

This picture forms the basis of a dimensional analysis. Let $a_0$ be the initial radius of the void, $\sigma$ the thermal stress, $\gamma$ the surface energy, and $E$ Young's modulus. The relative importance of the elastic energy and the surface energy is described by a dimensionless number

$$\Lambda = \frac{\sigma^2 a_0}{\gamma E}. \tag{1.1}$$

When $\Lambda$ is small, surface energy dominates, and the void will be rounded. When $\Lambda$ is large, strain energy dominates, and the void will collapse into a narrow slit. The circular void collapses when $\Lambda$ exceeds a critical value $\Lambda_c$. The analysis in this article shows that $\Lambda_c = 3/8$. For aluminum with $E = 7\times10^{10}$ N/m$^2$ and $\gamma = 1$ N/m, under stress $\sigma = 4\times10^8$ N/m$^2$, the critical radius is $a_0 = 164$ nm. Any larger void will collapse under this stress level. The mechanism works under both tensile and compressive stress.

The same phenomenon is anticipated for other material systems. A technically important example is residual gas pores inside single-crystal oxide fibers, subjected to both high temperatures and mechanical loads. The mechanism can limit the lifetime of the composite materials based on these fibers (private communication with Evans). In this article, however, we set aside these potential applications and concentrate on the general formulation of the problem, and on the implications for the interconnects.

Surface diffusion mediated instability in elastic solids has been studied by several investigators.[12–15] The lead phenomenon, which has engaged the previous studies, is that an initially flat surface may undulate due to surface diffusion, driven by elastic energy against surface energy. Recent kinetic simulation has shown that cracklike slits may result from such undulation.[13,15] A dimensionless group similar to $\Lambda$ has appeared in these studies, with $a_0$ replaced by the wavelength. In writing this article we have been inspired by these studies, and by the recent synthesis of the spatiotemporal complexity on the basis of nonequilibrium thermodynamics and dynamical systems.[16]

In this article the void evolution is viewed as an irreversible process, and formulated in a sufficiently general way that other mechanisms of energy variation or entropy production can be readily incorporated. A variational principle is identified which governs the evolution under combined action of surface energy, elastic energy, and electron wind force. Questions typical for any evolutionary process also have direct bearing on the voids in the interconnects: (1) Under what conditions does a circular void become unstable? (2) What is the destination of the evolution, a slit of vanishing thickness, or something still quite rounded? (3) Given nominally the same experimental conditions, why are the slits not always observed? (4) How fast does the shape change? (5) What is the role of stress bias or other imperfections? Energetics and kinetics are considered separately in two sections; together they illuminate the phenomenon.

## II. ENERGETICS

The suitable thermodynamic potential, consisting of both elastic and surface energy, is a functional of the void shape. Approximate equilibrium void shapes are selected, among a family of ellipses, by minimizing and maximizing the potential. The bifurcation diagram is a composite of a subcritical pitchfork and the Griffith cracks. A void under biased stresses is analyzed to illustrate the effect of imperfections.

### A. Why does a circular void collapse?

Figure 1 illustrates the cross section of a cylindrical void in a solid, subjected to biaxial stresses on the external boundary, but not on the void surface. The cross-sectional shape of the void is arbitrary. The work done by the load either varies energy in the solid, or produces entropy in the diffusion process. The first law of thermodynamics requires that

$$(\text{energy rate}) + (\text{dissipation rate}) = (\text{work rate}). \tag{2.1}$$

The solid varies energy either in the body or on the surface. Denote $w$ as the strain energy per volume and $\gamma$ the surface energy per area. They are taken to be independent from each other for practical purposes. That is, $\gamma$ is independent of the applied stress, and the strain field in the body is determined by the elasticity theory neglecting the effect of surface energy. The total elastic energy and surface energy are

$$U_e = \int_{\text{body}} w\, dA, \quad U_s = \int_{\text{surface}} \gamma\, dL. \tag{2.2}$$

For the two-dimensional problem, they are energy per length, integrated over the cross-sectional area of the solid $A$, and the arc length of the void $L$, respectively. Under the fixed mechanical load, the suitable potential is

$$\Phi = U_e + U_s - (\text{load} \times \text{displacement}).$$

Furthermore,

$$U_e = (\text{load} \times \text{displacement})/2$$

for linear elastic solids. Thus, the thermodynamic potential for the linear elastic solid under constant load is

$$\Phi = -U_e + U_s. \tag{2.3}$$

The potential is a functional of void shape. For a given void shape, $U_e$ is determined by the elasticity problem and $U_s$ is integrated over the perimeter of the void. The same potential has also appeared in the linear fracture mechanics as a functional of crack size and, in three dimensions, crack shape.

Equation (2.1) becomes

$$\frac{d\Phi}{dt} + (\text{dissipation rate}) = 0. \tag{2.4}$$

The second law of thermodynamics requires that the dissipation be positive as atoms diffuse, and vanish as the void reaches the equilibrium shape. That is, atoms diffuse to reduce the potential of the system. Of all void shapes, the equilibrium shape minimizes $\Phi$. Because atoms diffuse only on the surface, the void conserves the cross-sectional area as the shape changes. Other kinetic details are unnecessary for equilibrium considerations and are left to the following section.

In Sec. I the shape instability is analyzed by the local states on the surface. Alternatively, it can be analyzed by the global energy. Compare $\Phi$ for the circular void and a void with reduced symmetry, say an ellipse having the same area as the circle. Here and later we use $\Delta$ to signify the difference of a quantity for an ellipse and a circle; for example, $\Delta\Phi = \Phi(\text{ellipse}) - \Phi(\text{circle})$. The ellipse has longer perimeter than the circle, so that $\Delta U_s > 0$. The body with the elliptic hole is more compliant to the external load than the body with the circular hole, so that $\Delta U_e > 0$. Consequently, both the surface and the elastic energy increase when the circle breaks the symmetry; $\Delta\Phi < 0$ if $\Delta U_e > \Delta U_s$. The circular void is unstable when either the elastic energy is large, or the surface energy is small. These considerations also identify $\Lambda$ in Eq. (1.1).

### B. To relax or to collapse

The fate of the voids needs be clarified. Will a noncircular void relax to a circular void, or will it collapse to a narrow slit? In the following, the potential energy is calculated for ellipses having the constant area, and the ellipse that minimizes it is taken to be in equilibrium. The procedure is that of Rayleigh–Ritz: The potential is a functional of the void shape, but only a restricted family of shapes is searched to minimize it. The procedure usually yields approximate equilibrium shapes and the accuracy improves as more families are searched. The family of ellipses with the constant area is parameterized by only one number; yet, it will be shown that this family contains two exact asymptotes: the initial bifurcation from the circle and the slit of vanishing thickness. Thus, it is not unreasonable to expect that the ellipses well approximate the equilibrium shapes between the two asymptotes.

Let $a_0$ be the radius of the initial circular void. The ellipses have the same area as the circle $\pi a_0^2$. Their shapes are described by

$$X = a_0 \left( \frac{1+m}{1-m} \right)^{1/2} \cos\theta, \quad Y = a_0 \left( \frac{1-m}{1+m} \right)^{1/2} \sin\theta. \tag{2.5}$$

The circle corresponds to $m = 0$, the $X$-direction slit to $m \to +1$, and the $Y$-direction slit to $m \to -1$. The ellipses have perimeter

$$P = \frac{a_0}{\sqrt{1-m^2}} \int_0^{2\pi} (1 + m^2 - 2m \cos 2\theta)^{1/2} \, d\theta. \tag{2.6}$$

The integral is evaluated numerically.

The elastic solution for the elliptic voids exists in the literature, from which $U_e$ is calculated (Appendix A). For the body with an elliptic hole and the body with a circular hole, the elastic energy differs by

$$\Delta U_e = 4\pi \frac{\sigma^2 a_0^2}{E} \frac{m^2}{1-m^2}. \tag{2.7}$$

Thus, $U_e$ increases as the ellipse becomes more elongated, the area and load being constant. This has been stated previously on an intuitive basis that a body with an elliptic hole is more compliant than a body with a circular hole. Combining Eqs. (2.6) and (2.7), the difference in the potential is

$$\frac{\Delta\Phi}{2\pi a_0 \gamma} = -2\Lambda \frac{m^2}{1-m^2} + \left( \frac{P}{2\pi a_0} - 1 \right). \tag{2.8}$$

Now $\Phi$ is a function of the shape parameter $m$ for a given control parameter $\Lambda$.

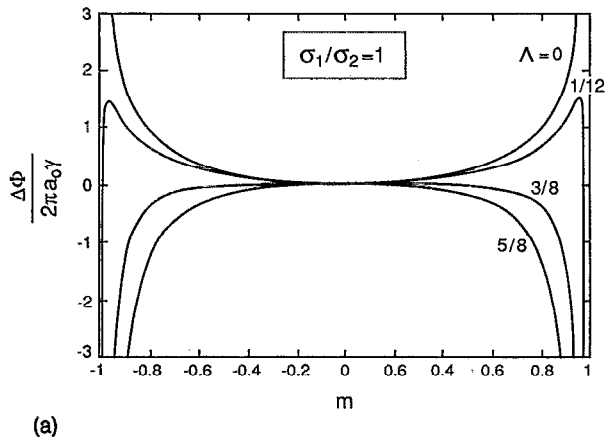
Figure 2(a) displays the function $\Phi(m)$ at several constant levels of $\Lambda$. Each minimum and maximum represents a stable and unstable equilibrium state, respectively. Three types of behaviors emerge depending on the value of $\Lambda$, i.e., the relative importance of elastic and surface energy, as follows.

(1) When $\Lambda = 0$, the stress vanishes; $\Phi$ reaches a minimum at $m = 0$, and maxima at $m = \pm 1$. The circular void is stable and the two slits are unstable: Any ellipse will relax to the circle.

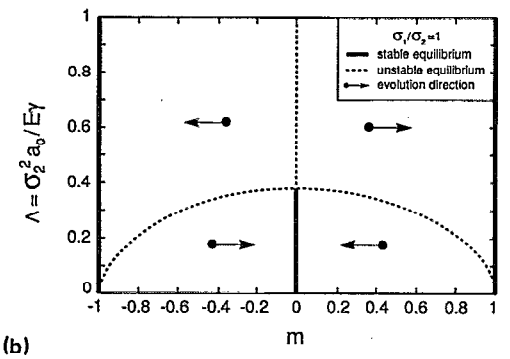
(2) When $\Lambda \in (0, 3/8)$, the stress is finite but surface energy still dominates; $\Phi$ reaches a local minimum at $m = 0$, two maxima at some $\pm m_c$, and two minima at $m = \pm 1$. The maxima act as energy barriers: an ellipse of $|m| < m_c$ will relax to the circle, but an ellipse of $|m| > m_c$ will collapse to the slits.

(3) When $\Lambda \in (3/8, \infty)$, the stress dominates; $\Phi$ reaches the maximum at $m = 0$, and minima at $m = \pm 1$. The circle is unstable but the slits are stable: any elliptic void will collapse to the slits.

The above information is projected onto the $(\Lambda, m)$ plane, Fig. 2(b). The heavy solid and dotted lines correspond to the stable and unstable equilibrium states, respectively. The two slits $m = \pm 1$ are stable for any $\Lambda > 0$, but unstable




(a)




(b)

FIG. 2. Biaxial stress state $\sigma_1=\sigma_2=\sigma$: (a) The potential as a function of the void shape $m$ at several levels of $\Lambda$; (b) the bifurcation diagram is a combination of a subcritical pitchfork and two Griffith cracks.

for $\Lambda=0$. The circle $m=0$ is metastable when $\Lambda<3/8$, but unstable when $\Lambda>3/8$. The dotted curve is the unstable equilibrium states, referred to as $m_c$ in the preceding paragraph. These lines divide the $(\Lambda,m)$ plane into four regions. A point in each region corresponds to an ellipse under a constant level of $\Lambda$, evolving toward a stable equilibrium state, either the circle or the slits. The evolution direction in each region is indicated by an arrow. An ellipse below the dotted curve relaxes to the circle, and an ellipse above the dotted curve collapses to a slit. An initially circular void will collapse if $\Lambda$ exceeds the critical value $\Lambda_c=3/8$. This value has been used in the Introduction to calculate the critical void radius under a given stress.

### C. Pitchfork and crack

The bifurcation diagram, Fig. 2(b), is better appreciated as follows. First focus on how the perfect circle breaks the symmetry, i.e., the subcritical pitchfork bifurcation at $\Lambda_c$. The shape parameter $m$ measures the order in the critical point theory.[16] Expand Eq. (2.8) in powers of $m$,

$$\frac{\Delta\Phi}{2\pi a_0\gamma}=\left(-2\Lambda+\frac{3}{4}\right)m^2+\left(-2\Lambda+\frac{33}{64}\right)m^4+\cdots. \tag{2.9}$$

Only the two leading terms are retained for small $m$. When $\Lambda>3/8$, the coefficient is positive for $m^2$, so that $m=0$ maximizes $\Phi$. When $\Lambda<3/8$, the coefficient is positive for $m^2$ but negative for $m^4$, so that $m=0$ minimizes $\Phi$. Consequently, $\Lambda=3/8$ is the critical point above which the circle is unstable. Equilibrium requires that $d\Phi/dm=0$, i.e.,

$$2(-2\Lambda+\tfrac{3}{4})m+4(-2\Lambda+\tfrac{33}{64})m^3=0. \tag{2.10}$$

When $\Lambda<3/8$, $\Phi$ reaches the two maxima at

$$m^2=\tfrac{64}{15}(\tfrac{3}{8}-\Lambda),\quad \Lambda\to\tfrac{3}{8}. \tag{2.11}$$

The analysis determines the critical point, $\Lambda_c=3/8$, and the asymptotic behavior of the dotted curve in Fig. 2(b) as $m\to0$; yet, the analysis is not rigorous in that the equilibrium shapes are only searched among the ellipse family. This concern is removed in Sec. III B where the ellipse is shown to be an exact initial bifurcation mode.

The other limiting case, the slit of vanishing thickness, reproduces the Griffith theory of cracks.[17] Keep only the unbounded terms in Eq. (2.8) as $m\to\pm1$,

$$\frac{\Delta\Phi}{2\pi a_0\gamma}=-\frac{2\Lambda}{1-m^2}+\frac{4}{\pi\sqrt{1-m^2}}. \tag{2.12}$$

For a given $\Lambda$, the potential attains the maxima when

$$\sqrt{1-m^2}=\pi\Lambda,\quad \Lambda\to0. \tag{2.13}$$

This is the asymptote in Fig. 2(b) as $m\to\pm1$. From Eq. (2.5) the crack length is $a=a_0\sqrt{2/(1-m)}$ as $m\to1$, so that Eq. (2.13) becomes

$$\frac{\sigma^2 a}{\gamma E}=\frac{2}{\pi}, \tag{2.14}$$

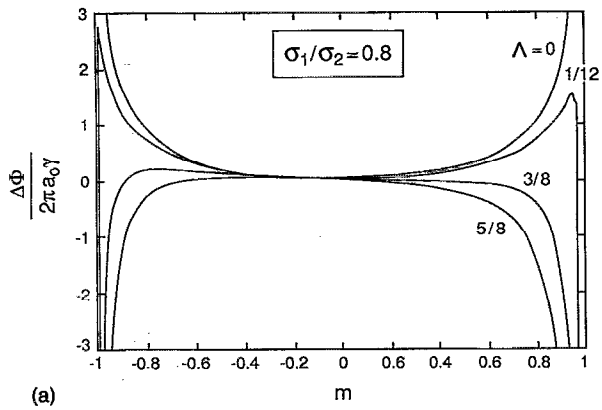
which is just the Griffith condition for crack growth.

The connection with the Griffith cracks is not fortuitous. Both phenomena are based on the competition between the elastic and the surface energy, i.e., on the potential (2.3), although the kinetic processes are different: A void changes the shape by surface diffusion, and a crack extends by atomic decohesion. The energetics of the two phenomena coincide in the limit when the void approaches the crack. The conclusion should apply to other loading configurations if a void is sufficiently elongated to be approximated by a crack. Let $\mathcal{G}$ be the elastic energy release rate for the crack. As atoms diffuse on the surface, the elongated void will become shorter if $\mathcal{G}<2\gamma$, but longer if $\mathcal{G}>2\gamma$. The connection is useful because $\mathcal{G}$ has been solved in fracture mechanics for many configurations.

### D. Stress bias and other imperfections

A void in an interconnect deviates from the perfect symmetry in many ways. Surface energy is anisotropic in crystals; for example, the {111} planes in aluminum have the lowest surface energy, and are the preferred void surfaces. The interconnect is finite and encapsulated by insulators; the elastic modulus misfit causes asymmetry. The stresses in two directions are not exactly the same. Given these imperfections, the circular symmetry breaks even at vanishing stress. What use, then, is the perfect problem?

The significance of the perfect problem is understood as follows. If an imperfection is not too large in magnitude, it only changes the potential $\Phi$ slightly. Changing with it will



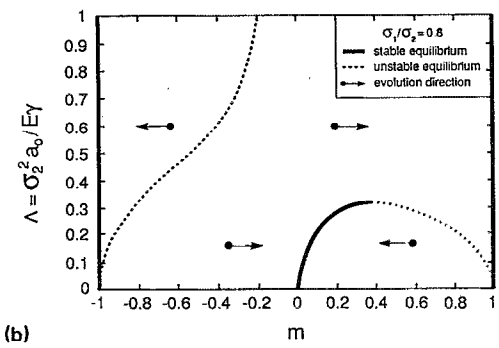
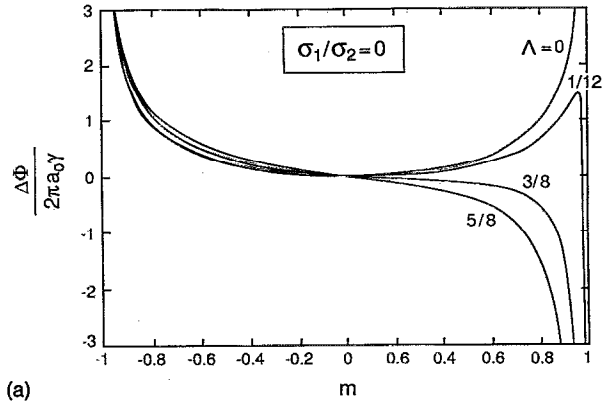




FIG. 3. Biased biaxial stress state $\sigma_1/\sigma_2=0.8$; (a) the potential as a function of the void shape $m$ at several levels of $\Lambda$; (b) stability conditions projected on the $(m,\Lambda)$ plane.



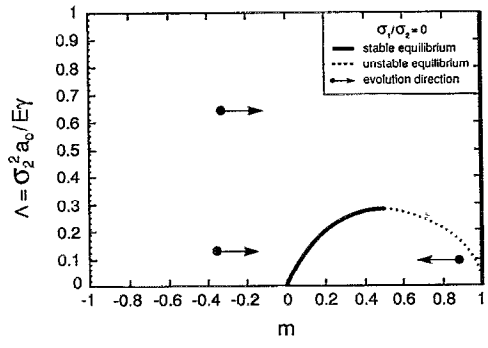




FIG. 4. Uniaxial stress state $\sigma_1=0$, $\sigma_2\neq 0$; (a) the potential as a function of the void shape $m$ at several levels of $\Lambda$; (b) stability conditions projected on the $(m,\Lambda)$ plane.

be the locations of the minima and maxima. The lines in Fig. 2 will bend somewhat, but the essential features should remain unchanged. Even for large imperfections there will still be regions, just as in Fig. 2(b), where a void will relax to a rounded shape (no longer a perfect circle), and other regions where a void will collapse to a slit.

It is obviously impractical, and often unnecessary, to analyze all the imperfections. To illustrate the general idea, we study a void in an isotropic crystal under a biased stress state, i.e., $\sigma_2\neq\sigma_1$ in Fig. 1. The potential is

$$\Delta\Phi=-\frac{2\pi a_0^2}{E}\left(\frac{m}{1-m}\,\sigma_2^2-\frac{m}{1+m}\,\sigma_1^2\right)+\gamma(P-2\pi a_0). \tag{2.15}$$

The first term, the elastic energy, is evaluated in Appendix A. As evident from Eq. (2.15), the tensile and the compressive stresses give the identical response. We will consider the case $\sigma_2>\sigma_1\geqslant 0$, and modify the control parameter as $\Lambda=\sigma_2^2 a_0/E\gamma$.

Figures 3(a) and 3(b) are for $\sigma_1/\sigma_2=0.8$; they are representative for any stress ratios in the interval $0<\sigma_1/\sigma_2<1$. Several asymmetries are noted when comparing Figs. 2(a) and 3(a). For small $\Lambda$, the local minimum no longer occurs at $m=0$, nor do the two maxima at the same value of $|m|$. At a critical value, still denoted as $\Lambda_c$, the minimum and the maximum on the right-hand side annihilate, but the maximum on the left-hand side persists. In Fig. 3(b), the values of $m$ minimizing $\Phi$ are the heavy solid lines, and the values of $m$ maximizing $\Phi$ are the dotted lines. As expected, under the biased stress, the equilibrium shape is noncircular even for a small value of $\Lambda$. The heavy solid curve ends at $\Lambda_c$, and is continued by the dotted curve.

Figures 4(a) and 4(b) are the corresponding diagrams under uniaxial stress state, $\sigma_1=0$ and $\sigma_2\neq 0$, Fig. 1. For a small $\Lambda$, $\Phi$ has a local minimum and a local maximum. For a large value of $\Lambda$, $\Phi$ monotonically decreases as $m$ increases. In Fig. 4(b), the slit $m=-1$ is an unstable equilibrium state, and the slit $m=1$ is a stable equilibrium state. The heavy solid curve ends at $\Lambda_c$, and is continued by the dotted curve. Note that the critical value $\Lambda_c$ is reached at about $m=0.5$, corresponding to an ellipse with axes ratio 3.

An inspection shows that Fig. 3(b) will degenerate to Fig. 2(b) as $\sigma_1/\sigma_2\to 1$, and to Fig. 4(b) as $\sigma_1/\sigma_2\to 0$. They all have the identical Griffith limit as $m\to 1$, as anticipated. Although Figs. 2(b)–4(b) look very different, they mean practically the same thing. The initial void shape is usually not too different from a circle. It will relax to a more or less rounded void if $\Lambda<\Lambda_c$, but collapse to a slit of vanishing thickness if $\Lambda>\Lambda_c$. Plotted in Fig. 5 is $\Lambda_c$ as a function of the stress ratio. The critical number does not vary significantly for the entire range of the stress ratio.

## III. KINETICS

In this section, kinetic concepts for surface diffusion in elastic solids are reviewed.[12–15] We present them in the language of nonequilibrium thermodynamics,[18] so that other

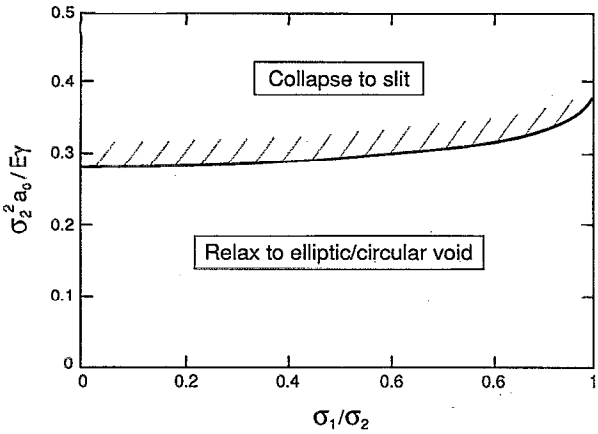



FIG. 5. The critical number $\Lambda_c$ as a function of the stress ratio.

mechanisms of energy variation or entropy production can be added readily. Electromigration is used to illustrate the procedure. Exact initial bifurcation modes of a circular void are determined; the critical loads for successive modes are closely spaced, indicating that complicated void shapes may evolve for slightly different initial imperfections. A variational principle governing shape change is identified and used to estimate the evolutionary rate.

### A. Evolution is an irreversible process

Nonequilibrium thermodynamics has three elements: mass conservation, the two thermodynamic laws, and phenomenological kinetic relations. These are examined in turn for the present problem. Conservation of atoms sets a kinematic constraint: At any point the surface recedes if the flux has a positive divergence, i.e.,

$$V_n=\Omega\,\frac{\partial J}{\partial L}. \tag{3.1}$$

On the left-hand side, $V_n=\mathbf{n}\cdot\partial\mathbf{X}/\partial t$ is the normal velocity of the surface, where $t$ is the time, $\mathbf{X}$ the position vector of a point on the surface, and $\mathbf{n}$ the unit vector normal to the surface pointing into the solid; $V_n>0$ if the surface recedes. On the right-hand side, $J$ is the surface atomic flux, i.e., the number of atoms per time across per length, $L$ is the arc length, and $\Omega$ is the atomic volume.

Express the first law (2.4) explicitly as

$$\frac{d\Phi}{dt}+\int FJ\,dL=0. \tag{3.2}$$

The integral extends over the void perimeter. Everything else having been defined, this equation uniquely defines $F$ as the diffusion driving force on each atom. Furthermore, Eq. (3.2) links the global energy variation with the local kinetic process. The first term is the potential energy increase rate, and the second term the dissipation rate associated with surface diffusion. The second law of thermodynamics requires that the dissipation be positive when atoms diffuse, and vanish when an the void attains equilibrium. Consequently, the potential energy decreases as the void evolves toward the equilibrium shape. Equation (3.2) is valid for any virtual surface velocity $V_n$ and flux $J$ compatible in the sense of Eq. (3.1), even if $F$ and $J$ are not connected by any kinetic relation. This rigorous understanding will lead to an explicit formula for $F$ in Eq. (3.5), and an evolutionary variational principle in Sec. III C.

When a piece of the void surface recedes, the body becomes more compliant to the constant load and the void has longer perimeter, so that both $U_e$ and $U_s$ increase. Formalizing these observations with $\Phi$ defined by Eqs. (2.2) and (2.3), one can show that

$$\frac{d\Phi}{dt}=-\int(w-\gamma K)V_n\,dL. \tag{3.3}$$

The elastic energy density is evaluated on the surface; for a traction-free cylindrical void, $w=\sigma_t^2/2E$, $\sigma_t$ being the hoop stress. The curvature $K$ is taken to be positive for a convex void. Equation (3.3) shows that the potential changes only when the surface moves, as is sensible. The capillary term in Eq. (3.3) is the same as the Laplace–Young relation for soap films. Replacing $V_n$ in Eq. (3.3) by $J$ using Eq. (3.1), and then integrating by parts, one reaches

$$\frac{d\Phi}{dt}=\int J\,\frac{\partial}{\partial L}(\Omega w-\Omega\gamma K)dL. \tag{3.4}$$

Since Eqs. (3.2) and (3.4) are valid for any virtual flux, a comparison of them gives

$$F=\frac{-\partial(\Omega w-\Omega\gamma K)}{\partial L}. \tag{3.5}$$

Given a void shape with the prescribed load, $w$ is determined by the elasticity problem and $K$ by the geometry. Atoms diffuse in the direction of $F$; the void reaches equilibrium when $F$ vanish at every point on the surface. It is sometimes convenient to think in terms of quantity $(\Omega w-\Omega\gamma K)$, the chemical potential. Atoms diffuse toward the position with lower chemical potential. This forms the basis of the instability argument in Sec. I.

The above considerations are applicable for any kinetic relations. In the remainder of the section, a linear kinetic relation is assumed,

$$J=MF. \tag{3.6}$$

The phenomenological constant obeys the Einstein relation $M=D_s\delta_s/\Omega kT$, with $D_s$ the surface diffusivity, $\delta_s$ the effective thickness of the surface atomic layers participating in the diffusion process, $k$ Boltzmann's constant, and $T$ the absolute temperature. The crystal is assumed to be isotropic so that $M$ is constant along the void surface.

The following normalizations prevail in the remainder of this section:

$$(\mathbf{x},l,1/\kappa)=(\mathbf{X},L,1/K)/a_0,\quad \Sigma=\sigma_t/\sigma,\quad \tau=t\Omega^2M\gamma/a_0^4. \tag{3.7}$$

A combination of Eqs. (3.1), (3.5), and (3.6) gives

$$\mathbf{n}\cdot\dot{\mathbf{x}}=-\frac{\partial^2}{\partial l^2}\left(\frac{\Lambda}{2}\Sigma^2-\kappa\right). \tag{3.8}$$

The superimposed dot signifies the derivative with respect to the dimensionless time $\tau$. The right-hand side is completely determined for a given void shape and load; the dimensionless hoop stress $\Sigma$ is determined by the elasticity theory. Consequently, Eq. (3.8) governs the evolution, a moving boundary problem with only one parameter $\Lambda$. Although every individual physical origin is well understood, the evolutionary process can be remarkably complex. In what follows we concentrate on the aspects that have direct bearing on the voids in interconnects.

### B. Initial bifurcation modes

The Rayleigh–Ritz procedure in Sec. II B does not warrant that an ellipse is indeed a bifurcation mode. Here we solve, by a semi-inverse approach, the exact initial bifurcation modes of the circular void under biaxial stress $\sigma_1=\sigma_2=\sigma$, Fig. 1. The initial bifurcation modes are first guessed, and then confirmed, to be hypocycloids.

A hypocycloid is the trajectory of a point fixed on a circular disk which rolls, without slipping, upon the interior of another fixed circle. It is described by

$$x=\cos\theta+m\cos n\theta,\quad y=\sin\theta-m\sin n\theta,$$
$$\theta\in(0,2\pi). \tag{3.9}$$

Here $n$ is a positive integer; $m(\tau)\ll 1$ measures the small perturbation from the circle. They look like curved polygons; for $n=1$ an ellipse, for $n=2$ a triangle, etc. To the first power in $m$, the hypocycloids defined by Eq. (3.9) always have the same area $\pi$.

The elastic solution of a hypocycloids hole exists in the literature, as outlined in Appendix B. To the first power in $m$, the dimensionless hoop stress is

$$\Sigma=2+4mn\cos(n+1)\theta. \tag{3.10}$$

After some manipulations, the first-order perturbation of the dimensionless curvature is found to be

$$\kappa=1+m(n^2+2n)\cos(n+1)\theta, \tag{3.11}$$

and that of the surface velocity

$$\mathbf{n}\cdot\dot{\mathbf{x}}=\dot{m}\cos(n+1)\theta. \tag{3.12}$$

Substituting the above into the evolution equation (3.8), and only retaining the terms of first power in $m$, one obtains that

$$\dot{m}=mn(n+1)^2(8\Lambda-n-2). \tag{3.13}$$

In the above $\cos(n+1)\theta$ has been cancelled from the both sides. That this evolution equation is independent of the position on the surface $\theta$ is significant: With a small perturbation from the circle, the $n$th hypocycloid will grow or shrink only as the $n$th hypocycloid. This confirms that hypocycloids are indeed the initial bifurcation modes. The perturbation grows if $\dot{m}>0$, and shrinks if $\dot{m}<0$. The critical load is reached when $\dot{m}=0$ in Eq. (3.13). Thus

$$\Lambda_n=\frac{n+2}{8}. \tag{3.14}$$

For $n=1$, the bifurcation mode is an ellipse, and the critical load is $\Lambda_1=3/8$.

Observe that the critical loads for successive bifurcation modes are not far apart, being spaced by 1/8. If the control parameter $\Lambda$ is large (large void or high stress), the long-term shape will be selected according to the initial imperfection of the hole. The spatiotemporal complexity can be resolved by tracing the evolution governed by Eq. (3.8) for various initial imperfections. The calculation will not be pursued here, but a variational principle identified in the following subsection will ease the labor.

### C. Variational principles for evolution

Numerical computation is inescapable to trace the evolution in general. Given a void shape, the elasticity problem must be solved first, usually numerically, and the void shape is then updated according to Eq. (3.8). The whole process is repeated for many time steps. Equation (3.8) has been exclusively used in the previous studies, but it may not be efficient because it involves high-order differentiations of the boundary. Instead, variational principles may be used.

Variational principles have been developed for problems such as grain boundary cavitation and powder compaction.[19,20] We find a variational principle governing the shape change in the present problem. Recall that the first law of thermodynamics (3.2) is valid for any compatible virtual velocity and flux. Write this explicitly as

$$\frac{d}{dt}(\delta\Phi)+\int F\delta J\,dL=0. \tag{3.15}$$

The variation in $\Phi$ is due to the shape change. Replacing $F$ in Eq. (3.15) by using the kinetic relation $F=J/M$, one obtains

$$\delta\left(\frac{d\Phi}{dt}+\int\frac{1}{2M}J^2\,dL\right)=0. \tag{3.16}$$

This suggests the following variational principle.

Of all virtual velocities and fluxes compatible in the sense of Eq. (3.1), the actual velocity and flux minimize functional,

$$\Pi=\frac{d\Phi}{dt}+\int\frac{1}{2M}J^2\,dL. \tag{3.17}$$

The integral extends over the perimeter of the void. For a given shape, the first term can only vary with the surface velocity, which in turn relates to the flux by Eq. (3.1). Consequently, $\Pi$ can be viewed as a functional of either $V_n$ or $J$, both being functions of the arc length $L$. Compared with the grain-boundary cavitation problem,[19] a new term $d\Phi/dt$ appears to account for energy variation. An example in Sec. III D will illustrate how this term works.

Using Eq. (3.3), one can also write the functional in alternative forms, e.g.,

$$\Pi=\int\left(-(w-\gamma K)V_n+\frac{1}{2M}J^2\right)dL, \tag{3.18}$$

which may be convenient in some circumstances. Rigorously, Eq. (3.17) only says that $\Pi$ is stationary at the actual

velocity and flux. A proof of minimum on the basis of Eq. (3.18) follows. Let $J$ be the actual flux that satisfies Eqs. (3.1), (3.5), and (3.6). Because atoms only diffuse on the void surface, the function $J(L)$ is periodic with the perimeter of the void. Let $Q(L)$ be an arbitrary function having the same period, $(J+Q)$ be a virtual flux, and the associated virtual velocity be obtained from Eq. (3.1). The virtual flux need not satisfy the kinetic relation (3.6). We now compute the difference $\Pi(J+Q)-\Pi(J)$ using Eq. (3.18). Replacing both the virtual and the actual velocity with their associated fluxes by Eq. (3.1), the difference in $\Pi$ becomes

$$\int \left( -(w-\gamma K)\Omega \frac{\partial Q}{\partial L} + \frac{1}{2M}[(J+Q)^2 - J^2] \right) dL. \tag{3.19}$$

Integrating the first term by parts, one reaches

$$\int \left( \frac{\partial}{\partial L}(\Omega w - \Omega \gamma K) + \frac{J}{M} \right) Q\, dL + \int \frac{Q^2}{2M}\, dL. \tag{3.20}$$

The first integral vanishes because $J$ satisfies Eqs. (3.5) and (3.6). The second integral is always nonnegative because $M>0$, which in turn is required by the second law of thermodynamics. Thus, $\Pi(J+Q)-\Pi(J)\geq 0$ for any virtual flux. This proves the variational principle.

We now wish to illustrate the versatility of the variational principle by considering how to include electromigration. Only an outline of concepts will be given here. A detailed implementation is beyond the scope of this article, and will be reported elsewhere. As an electric current flows in an interconnect, the drifting electrons exert on each atom on the void surface a force $F^* = -q^* E_t$, known as the electron wind force.[1] Here $q^*$ $(>0)$ is a phenomenological constant having the unit of electric charge, and $E_t$ is the electric field tangential to the void surface. The negative sign signifies that the force directs along in the electron flow, which is opposite to the electric field. Including the work done by the electron wind force, the first law (3.2) becomes

$$\frac{d\Phi}{dt} + \int FJ\, dL = \int F^* J\, dL. \tag{3.21}$$

As before, $J$ can be any virtual flux, and $\Phi$ varies with the virtual $V_n$ compatible with $J$. It follows that, of all virtual velocities and fluxes that satisfy Eq. (3.1), the actual velocity and flux minimize the functional

$$\Pi = \frac{d\Phi}{dt} + \int \left( \frac{1}{2M} J^2 - F^* J \right) dL. \tag{3.22}$$

In numerical analysis, a void of an arbitrary shape can be represented by a conformal mapping from a unit circle. The mapping involves infinitely many coefficients, but only a finite number of them is used in numerical analysis. Consequently, the void shape is approximately represented by the chosen coefficients. The electric field on the void surface $E_t$ is readily determined by this mapping. The variational principle (3.22) is then used to compute the time rates of the mapping coefficients. This procedure will trace the evolution of these coefficients, and thereby the void shape.

It also follows from Eq. (3.21) that

$$F = F^* - \frac{\partial(\Omega w - \Omega \gamma K)}{\partial L}. \tag{3.23}$$

That is, the diffusion driving force consists of the electron wind force and the thermodynamic forces resulting from the elastic and the surface energy. Let $\phi$ be the electric potential governed by the Laplace equation. The electric field tangential to the void surface is given by $E_t = -\partial\phi/\partial L$. If $q^*$ is constant along the surface, the diffusion driving force in Eq. (3.23) becomes

$$F = \frac{-\partial(-q^*\phi + \Omega w - \Omega \gamma K)}{\partial L}. \tag{3.24}$$

The quantity in the bracket plays the same role as the chemical potential.

The above variational principles apply to a void subjected to periodic boundary conditions. For other problems such as powder compaction suitable boundary terms must be added. It is straightforward to extend the variational principles for a surface in the three dimensions. There will be two surface flux components, $J_1$ and $J_2$, and Eq. (3.1) will be replaced by the surface divergence. In Eq. (3.17), the line integrals will be replaced by surface integrals, the surface integrals by volume integrals, and $J^2$ by $J_1^2+J_2^2$. The variational principle can be used by dividing the surface into finite elements, so that surface evolution problems can be readily solved in three dimensions.

### D. How fast does the void evolve?

In using the variational principles, the void shape is approximately described by a finite number of parameters, all evolving with time. The more parameters, the better the description. As an example, the evolution rate of a void under biaxial stress $\sigma_1=\sigma_2=\sigma$ will be estimated. The initial imperfections are such that the void evolves in the first mode, from a circle to a slit. As before, we will approximate the shapes between the two limits by a family of ellipses, a choice compromising the accuracy of the prediction and the complexity of the calculation. The family of ellipses is

$$X = a_0 \alpha \cos\theta, \quad Y = a_0 \alpha^{-1} \sin\theta. \tag{3.25}$$

The dimensionless semiaxis $\alpha(\tau)$ is the parameter that describes the void shape; it is restricted to be $\alpha>1$. The ellipses conserve the area, elongating and shrinking in the $X$ and $Y$ directions, respectively, at the same rate.

The variational principle of version (3.17) will be used here. The dimensionless velocity is

$$v_n = \mathbf{n}\cdot\dot{\mathbf{x}} = \left( \frac{\dot\alpha}{\alpha} \right) \left( \frac{d\theta}{dl} \right) \cos 2\theta. \tag{3.26}$$

The dimensionless flux $j$ is so defined that Eq. (3.1) becomes $v_n = dj/dl$. Owing to the symmetry, $j=0$ when $\theta=0$. Integrate Eq. (3.26) and one obtains

$$j = \frac{\dot\alpha}{2\alpha} \sin 2\theta. \tag{3.27}$$

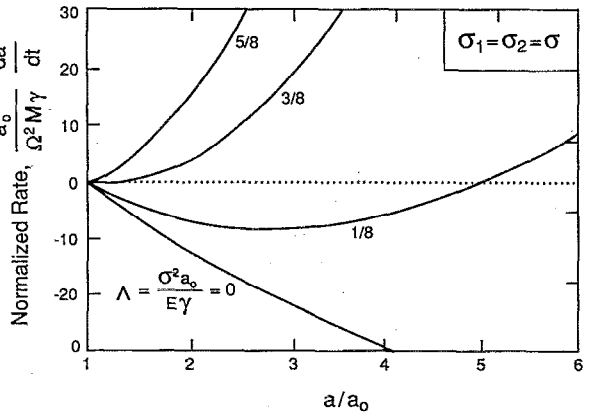



FIG. 6. The evolution rate of an elliptic void.

The potential $\Phi$ for an ellipse has been given by Eq. (2.8); $m$ and $\alpha$ are related by comparing Eqs. (3.25) and (2.5). The normalized functional takes form

$$\frac{\Pi}{(2\pi a_0\gamma)(\Omega^2 M\gamma/a_0^4)}=(-\Lambda I_0+I_1)\dot{\alpha}+\frac{1}{2}I_2\dot{\alpha}^2. \quad (3.28)$$

The coefficients $I$ depend on $\alpha$ only. Straightforward calculations yield

$$I_0=\alpha(1-\alpha^{-4}), \quad (3.29)$$

$$I_1=\frac{2}{\pi}\int_0^{\pi/2}\frac{\sin^2\theta-\alpha^{-4}\cos^2\theta}{\sqrt{\sin^2\theta+\alpha^{-4}\cos^2\theta}}\,d\theta, \quad (3.30)$$

$$I_2=\frac{1}{2\pi\alpha}\int_0^{\pi/2}\sin^2 2\theta\sqrt{\sin^2\theta+\alpha^{-4}\cos^2\theta}\,d\theta. \quad (3.31)$$

The integrals are computed numerically. Minimizing $\Pi$ in Eq. (3.28) by setting $\partial\Pi/\partial\dot{\alpha}=0$, one obtains the approximate evolution rate

$$\dot{\alpha}=(\Lambda I_0-I_1)/I_2. \quad (3.32)$$

Two limiting cases are obtained explicitly. For ellipses close to the circle, $\alpha\to 1$, the initial velocity is

$$\dot{\alpha}=4(\alpha-1)(8\Lambda-3),\alpha\to 1^+. \quad (3.33)$$

The rate $\dot{\alpha}$ is linear with the perturbation $(\alpha-1)$. This result is consistent with Eq. (3.13), noting $\alpha=1+m$ to the first power in $m$. For very elongated ellipse, $\alpha\gg 1$, neglecting the terms with $\alpha^{-4}$ in $I$, one finds that

$$\dot{\alpha}=\frac{15}{2}\alpha\left(\frac{\pi}{2}\Lambda\alpha-1\right), \quad \alpha\gg 1. \quad (3.34)$$

The rate is quadratic with the semiaxis for a very elongated void, and vanishes when the Griffith condition Eq. (2.14) is satisfied.

Equation (3.32) is plotted in Fig. 6. Denote $a$ as the semiaxis of an ellipse and $\alpha=a/a_0>1$. The ellipse will elongate further if the velocity is positive, but relax toward a circle if the velocity is negative. Three behaviors emerge for different values of $\Lambda$. For $\Lambda=0$, the rate is negative for any

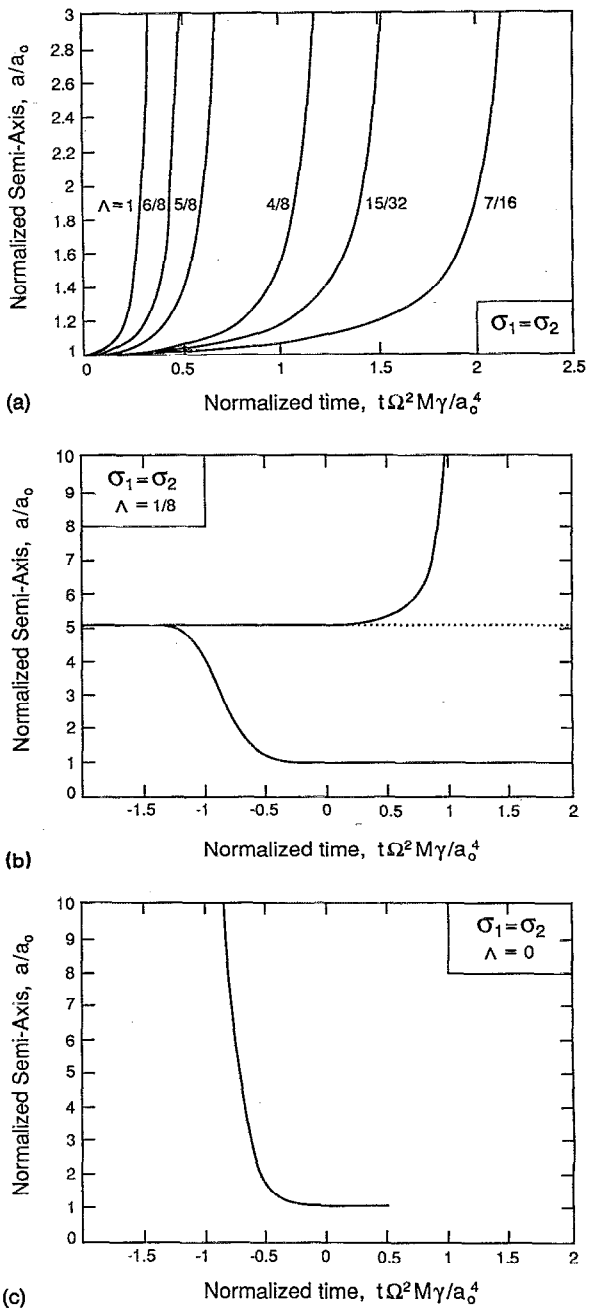



FIG. 7. The time for one ellipse to evolve to another.

ellipse, which will relax to the circle under the action of surface energy alone. For $\Lambda\in(0,3/8)$, the rate is negative for ellipse not too far from the circle, but positive for very elongated ellipses. For $\Lambda\in(3/8,\infty)$, the rate is positive for any ellipse, which will collapse to a slit.

The evolution time from one ellipse to another is obtained by integrating Eq. (3.32). The three types of behaviors are plotted in Figs. 7(a), 7(b), and 7(c), respectively. Since the evolution is an irreversible process, the time always increases in these figures. When $\Lambda\in(3/8\infty)$, any ellipse will become more elongated, Fig. 7(a). The curves are plotted by arbitrarily assigning the initial value $\alpha=1.01$ at $\tau=0$. One can also read from the diagram the time needed for, say, an ellipse with $\alpha=1.2$ to evolve to an ellipse with $\alpha=2.0$ under

load level $\Lambda=5/8$. When $\Lambda\in(0,3/8)$, there is an unstable equilibrium shape for each value of $\Lambda$, which is marked as the dashed line in Fig. 7(b) for $\Lambda=1/8$. More elongated ellipses will collapse to slits, but less elongated ellipses will relax to the circle. When $\Lambda=0$, any ellipse will relax to the circle, Fig. 7(c). This diagram gives the relaxation time of the ellipses after the load is removed. Conversely, the surface kinetic constant $M$ can be deduced if the relaxation time is measured experimentally.

Approximate, but explicit, formulas can be obtained from Eqs. (3.33) and (3.34). For example, if both initial and final axes, $a_i$ and $a_f$, are not too far from the circular radius $a_0$, the time duration is integrated from Eq. (3.33),

$$t_f-t_i=\frac{kTa_0^4}{32\Omega\gamma D_s\delta_s}\left(\frac{\sigma^2 a_0}{E\gamma}-\frac{3}{8}\right)^{-1}\ln\left(\frac{a_f-a_0}{a_i-a_0}\right). \tag{3.35}$$

The formula can also be used to estimate the order of magnitude even when $a_f$ is not very close to $a_0$.

If the void shape is described by many parameters, Eq. (3.28) will consist of a linear and a bilinear form of their rates. Setting the partial derivatives of $\Pi$ with respect to the rates to be zero, one obtains simultaneous linear algebraic equations for the rates. They are solved by Gaussian elimination to yield coupled initial value problems, which are then integrated by any standard procedure. Of course, as the void shape evolves, the elasticity problem must be independently solved at each step. The implementation as described is now in progress.

## IV. CONCLUDING REMARKS

A void in an elastic solid collapses into a transgranular slit when the dimensionless group $\sigma^2 a_0/E\gamma$ exceeds a critical value. The mechanism works under both tensile and compressive stress. It is important to examine this mechanism among the related ones. The transgranular slits will dictate the interconnect lifetime only if grain-boundary cavitation does not prevail. The latter has been thoroughly studied;[21] the stress to initiate grain-boundary cavitation is $\sigma\approx 2\gamma/a_0$, which is substantially lower than the stress to initiate a transgranular slit. Several considerations might explain why the transgranular slits occur in interconnects. First, a void drifts in a line under electric current, so that even a void nucleated at a grain boundary or a triple junction may sever the line inside a grain.[10] Second, for a line having the bamboolike grain structure, encapsulated by dielectrics, atomic sinks are partially eliminated which, in turn, limits the growth rate of a rounded void.[22] Third, electric current is expected to contribute in causing the shape instability.[11] Fourth electromigration will redistribute stress in the interconnects.[23]

As discussed in Sec. I, surface diffusion is the only dissipative process included in the present analysis. Among other dissipative processes, plastic creep gives rise to the greatest uncertainty. If operating at sufficiently high rate, creep will relax the thermal stress in the interconnect, and reduce the stress concentration at the tip of an elongated ellipse. Consequently, plastic creep tends to prevent a void from collapsing, or to blunt the tip of an existing elongated void. Creep can be incorporated into the analysis, which adds numerical complexities, but the major uncertainty arises from the lack of precise knowledge of the creep law in submicron dimensions. Some basic development is needed before such numerical analysis is worthwhile.

Although this work has been motivated by the interconnects, the phenomenon is anticipated for other material systems where stress is high and grain boundaries are inaccessible for diffusion. For example, single-crystal oxide fibers under mechanical load may suffer delayed fracture due to this mechanism. Unlike the existing stress corrosion cracking mechanisms, this one does not invoke environmental effects. It will cause a preexisting void inside a well-coated material to collapse; general fracture follows when the void is sufficiently elongated. It is hoped that experiments with better controlled systems will soon succeed in sorting out these matters.

## ACKNOWLEDGMENTS

We wish to thank Professor E. Arzt of the Max-Planck-Institut at Stuttgart, and Dr. J. E. Sanchez of Advanced Micro Devices for their generosity in showing us unpublished micrographs. Since the preprint of this article was circulated, several colleagues have kindly informed us that void instability in an elastic solid by mass transport had been studied in other technical problems (see Refs. 25 and 26). The work was supported by the Defense Advanced Research Projects Agency through the University Research Initiative under the Office of Naval Research Contract No. N-0014-92-J-1808, by the National Science Foundation through Grant No. MSS-9202165, and through a Young Investigator Award MSS-9258115 for Z.S.

## APPENDIX A: ELASTIC ENERGY CALCULATION

An infinite body containing an elliptic hole subjected to remote stresses stores infinite amount strain energy; yet, one can compute the energy difference between a body containing an elliptic hole and a body containing a circular hole subjected to the same remote stress state. Denote $\Delta U_e=U_e(\text{ellipse})-U_e(\text{circle})$. The ellipse and the circle have the same area $\pi a_0^2$. The hole can be regarded as an elastic medium with vanishing stiffness, so it can have strains. Of course, the stresses inside the hole vanish. For an infinite body containing an elliptic hole under remote stress, it is well known that the strain inside the hole is uniform. Denote the applied remote stresses as $\sigma_{ij}^{\infty}$, and the strains inside the hole as $\epsilon_{ij}^0$. We show that

$$\Delta U_e=\tfrac{1}{2}\pi a_0^2\sigma_{ij}^{\infty}\Delta\epsilon_{ij}^0. \tag{A1}$$

Here $\Delta\epsilon_{ij}^0$ are the differences in the strains inside the elliptic hole and those inside the circular hole.

To prove Eq. (A1), consider a body of external boundary $S_\infty$ containing an arbitrary hole of boundary $S_0$. Denote $n_i$ as the unit normal vector on the surfaces, pointing away from the solid. Let $\sigma_{ij}^{\infty}$ be a stress tensor independent of position. The body is subjected to the traction vector $\sigma_{ij}^{\infty}n_i$ on the

external surface, but free of traction on the void surface. Let $u_j$ be the displacement field in the body. The elastic energy in the body is

$$U_e=\tfrac{1}{2}\int_{S_\infty}\sigma^\infty_{ij}n_iu_j\,dS. \tag{A2}$$

Rewrite the above as

$$U_e=\tfrac{1}{2}\int_{S_\infty+S_0}\sigma^\infty_{ij}n_iu_j\,dS+\tfrac{1}{2}\int_{S_0}\sigma^\infty_{ij}n_iu_j\,dS. \tag{A3}$$

In the second integral, the surface normal is switched to point into the solid.

To understand the first integral in Eq. (A3), one needs an auxiliary body with the same geometry as the original body. The auxiliary body is under the uniform stress $\sigma^\infty_{ij}$ everywhere, with traction $\sigma^\infty_{ij}n_i$ on both internal and external surfaces. Consequently, the first integral in Eq. (A3) is the virtual work done by the traction on the auxiliary body through the displacement of the original body. According to the reciprocal theorem, it is the same, virtually, as the work done by the traction of the original body through the displacement of the auxiliary body. Neither the traction $\sigma^\infty_{ij}n_i$ on the original body nor the displacement of the auxiliary body depends on the hole geometry. Thus, the first integral in Eq. (A3) is independent of the hole geometry. Only the second integral makes the difference when the hole changes shape or size.

To take advantage of the fact that strains are uniform in the elliptic hole, one applies the Gauss theorem to the second integral in Eq. (A3), which changes the displacement to the strain inside the hole, leading to Eq. (A1).

Referring to Fig. 1, for an elliptic hole subjected to remote biaxial stress state, the hoop stresses at points $A$ and $B$, respectively, are[24]

$$\sigma_t(A)=-\sigma_1+\frac{3+m}{1-m}\sigma_2,\quad \sigma_t(B)=-\sigma_2+\frac{3-m}{1+m}\sigma_1. \tag{A4}$$

The corresponding stresses for a circular void are obtained by setting $m=0$. Because the point on the surface is under uniaxial stress state, the hoop strains are given by the stress divided by Young's modulus $E$ of the solid. Compatibility requires that on the surface the hoop strain in the solid is the same as the hoop strain in the "medium" inside the hole. Thus,

$$\Delta\epsilon^0_{xx}=\frac{-4m}{1+m}\frac{\sigma_1}{E},\quad \Delta\epsilon^0_{yy}=\frac{4m}{1-m}\frac{\sigma_2}{E}. \tag{A5}$$

Substituting Eq. (A5) into Eq. (A1), one finds the elastic energy difference between a body with an elliptic hole and a body with a circular hole,

$$\Delta U_e=\frac{2\pi a_0^2}{E}\left(\frac{m}{1-m}\sigma_2^2-\frac{m}{1+m}\sigma_1^2\right). \tag{A6}$$

Plane stress conditions are assumed in the above; under plane strain conditions, replace $E$ by $E/(1-\nu^2)$, $\nu$ being Poisson's ratio.

## APPENDIX B: STRESS AROUND A HYPOCYCLOIDAL HOLE

The hoop stress around a hypocycloidal hole is obtained from Ref. 24 with some manipulations. For two-dimensional elasticity problems, the stress field is solved by two analytic functions $\Omega(z)$ and $\omega(z)$ where $z=x+iy$ and $i=\sqrt{-1}$, namely

$$\sigma_{yy}+\sigma_{xx}=4\ \mathrm{Re}[\Omega'(z)], \tag{B1}$$

$$\frac{\sigma_{yy}-\sigma_{xx}}{2}+i\sigma_{xy}=\bar{z}\Omega''(z)+\omega'(z). \tag{B2}$$

The function

$$z=R(\zeta+m\zeta^{-n}) \tag{B3}$$

conformally maps the exterior of a unit circle on the $\zeta$ plane, $\zeta=\exp(i\theta)$, to the exterior of a hypocycloid on the $z$ plane. Here $n$ is a positive integer and $0\leqslant m<1/n$. The last restriction ensures that the hypocycloid does not have loops and only has cusps if $m=1/n$. A hypocycloid hole under remote biaxial tension, $\sigma_{xx}=\sigma_{yy}=\sigma$ as $|z|\to\infty$, is solved by

$$\Omega(z)=\frac{\sigma R}{2}\left(\zeta-\frac{m}{\zeta^n}\right), \tag{B4}$$

$$\omega(z)=-\frac{\sigma R}{2}\left[\frac{1}{\zeta}-m\zeta^n+\left(\frac{1+m\zeta^{n+1}}{\zeta^{n+1}-mn}\right)\left(\zeta^n+\frac{mn}{\zeta}\right)\right]. \tag{B5}$$

The hole is traction free so that the hoop stress on the surface is giving by the first invariant of the stress tensor, $\sigma_t=\sigma_{xx}+\sigma_{yy}$. From Eqs. (B1), (B3), and (B4) one finds that

$$\sigma_t=4\ \mathrm{Re}\left(\frac{d\Omega/d\zeta}{dz/d\zeta}\right)=\sigma\frac{2-2m^2n^2}{1+m^2n^2-2mn\cos(n+1)\theta}. \tag{B6}$$

In the body of the text, Eq. (3.10) retains the terms up to the first power in $m$.


[1] P. S. Ho and T. Kwok, Rep. Prog. Phys. **52**, 301 (1989).
[2] *Stress-Induced Phenomena in Metallization,* edited by C. Y. Li, P. Totta, and P. Ho (American Institute of Physics, New York, 1992).
[3] R. E. Jones and M. L. Basehore, Appl. Phys. Lett. **50**, 725 (1987).
[4] B. Greenebaum, A. I. Sauter, P. Flinn, and W. D. Nix, Appl. Phys. Lett. **58**, 1845 (1991).
[5] M. A. Korhonen, C. A. Paszkiet, and C.-Y. Li, J. Appl. Phys. **69**, 8083 (1991).
[6] W. D. Nix, Metall. Trans. A **20**, 2217 (1989).
[7] J. E. Sanchez, L. T. McKnelly, and J. W. Morris, J. Appl. Phys. **72**, 3201 (1992).
[8] J. H. Rose, Appl. Phys. Lett. **61**, 2170 (1992).
[9] Y.-C. Joo and C. V. Thompson, Mater. Res. Soc. Symp. Proc. **309**, 351 (1993).
[10] O. Kraft, S. Bader, J. E. Sanchez, and E. Arzt, Mater. Res. Soc. Symp. Proc. **309**, 199 (1993).
[11] Z. Suo, W. Wang, and M. Yang, Appl. Phys. Lett. **64**, 1944 (1994).
[12] R. J. Asaro and W. A. Tiller, Metall. Trans. **3**, 1789 (1972).
[13] C.-H. Chiu and H. Gao, Int. J. Solids Structures **30**, 2983 (1993).
[14] L. B. Freund and F. Jonsdottir, J. Mech. Phys. Solids **41**, 1245 (1993).
[15] W. H. Yang and D. J. Srolovitz, Phys. Rev. Lett. **71**, 1539 (1993).
[16] G. Nicolis and I. Prigogine, *Exploring Complexity* (Freeman, New York, 1989).

[17] A. A. Griffith, Philos. Trans. R. Soc. London Ser. A **211**, 163 (1921).
[18] I. Prigogine, *Introduction to Thermodynamics of Irreversible Processes*, 3rd ed. (Wiley, New York, 1967).
[19] A. Needleman and J. R. Rice, Acta Metall. **28**, 1315 (1980).
[20] R. M. McMeeking and L. T. Kuhn, Acta Metall. **40**, 961 (1992).
[21] T.-J. Chuang, K. I. Kagawa, J. R. Rice, and L. B. Sills, Acta Metall. **27**, 265 (1979).
[22] A. F. Bower and L. B. Freund, J. Appl. Phys. **74**, 3855 (1993).
[23] M. A. Korhonen, P. Borgesen, K. N. Tu, and C.-Y. Li, J. Appl. Phys. **73**, 3790 (1993).
[24] A. H. England, *Complex Variable Methods in Elasticity* (Wiley, London, 1971).
[25] R. N. Stevens and R. Dutton, Mater. Sci. Eng. **8**, 220 (1971).
[26] H. Gao, Q. J. Mech. Appl. Math. **45**, 149 (1992).