Dislocation Screening and the Brittle-to-Ductile Transition: A Kosterlitz-Thouless Type Instability

M. Khantha, D. P. Pope, and V. Vitek
Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104
(Received 22 February 1994)

We propose a new model for the brittle-to-ductile transition based on the Kosterlitz-Thouless concept of dislocation screening. In this model, thermal fluctuations assisted by the applied stress drive the spontaneous generation of dislocations and the instability occurs well below the melting temperature. In the limit of zero stress, our model reduces to the Kosterlitz-Thouless theory of the melting transition, and, in the opposite limit of zero temperature, we obtain the Rice-Thomson result for the brittle-to-ductile transition.

PACS numbers: 62.20.Fe, 46.30.Nz, 61.72.Lk, 62.20.Mk

The brittle-to-ductile transition (BDT) is a classic phenomenon exhibited by almost all materials with the possible exception of fcc metals. The change in the fracture behavior from brittle cleavage to ductile failure occurs usually in a narrow range of temperature accompanied by a dramatic increase in the fracture toughness [1] and a concomitant large increase in the dislocation density [2]. In materials with low initial dislocation density (e.g., Si [2]), the crossover is sharp (<5°) and the transition temperature (T_c) is roughly half of the melting temperature, whereas, in steels with high initial dislocation density, the transition is more gradual (<30°) and occurs below room temperature [1]. T_c is generally a function of the applied rate of stress intensity and its variation can be correlated with the temperature dependence of dislocation mobility in certain materials [2,3]. The massive dislocation generation at T_c is a necessary condition for the BDT. The strain-rate dependence of this behavior is an additional distinct feature associated with the transition.

In the existing models, the BDT is viewed either as a competition between crack propagation and thermally activated generation of a single dislocation at the crack tip [3–6] or as a thermally activated generation of a single dislocation mobility-controlled dynamical process [7]. The former describes the aspect of dislocation generation only while the latter considers the strain-rate dependence of T_c assuming dislocation generation to be possible. The physical understanding of the phenomenon, however, remains incomplete because a thermally activated mechanism of dislocation generation or mobility does not lead to a sudden well-defined transition at a characteristic temperature with a dramatic increase in the fracture toughness without imposing *ad hoc* conditions [6,7].

We propose here a model based on a new and different approach to the spontaneous generation of dislocations leading to the BDT. Unlike the traditional methods, which describe the stress-induced instability at *zero temperature* of a *single* dislocation loop at the crack tip, we consider the thermally induced instability of many small loops in the presence of an applied stress. The creation of many atomic-size loops by thermal activation induced a temperature-dependent *cooperative screening effect* that enhances the subsequent growth of the loops. This effect is physically distinct from the usual dislocation shielding of the crack tip stress [8]. We discuss the mechanism of cooperative dislocation generation in this paper. Elsewhere we derive analytical "similarity" solutions that describe the dislocation dynamics in the vicinity of the crack tip. The characteristic strain-rate dependence of T_c can then be deduced by combining the dynamical solutions with the criterion for instability obtained below.

The concept of dislocation screening was originally introduced by Kosterlitz and Thouless (K-T) [9] in an entirely different context, namely, two-dimensional (2D) phase transitions. In the K-T theory, the unstable generation of dislocations is driven by thermal fluctuations only, without the aid of an applied stress. It occurs close to the melting temperature and gives rise to a dislocation-mediated melting transition [10,11]. In our model, both the external stress and thermal fluctuations assist the growth of dislocation loops. This procedures a K-T type instability (but not a phase transition in the thermodynamic sense) well below the melting temperature at a stress level corresponding to the Griffith threshold needed for brittle crack propagation. This temperature corresponds to the BDT. We recover the familiar Rice-Thomson result [4] for dislocation generation in the limit of zero transition temperature when the applied load is equal to the Griffith threshold. We obtain the condition for the BDT using a simple 2D model. Our approach is applicable to intrinsically brittle materials with low initial dislocation density and the predicted values of T_c are in good agreement with observations.

The concept of screening is closely related to the fact that dislocation loops with radii slightly bigger than the (dislocation) core radius (r_0) can actually be formed by thermal activation at finite temperature in all solids [9]. The probability of existence of atomic-size loops or dipoles in 2D (which represent metastable "topological" defects) at a given temperature is determined only by

0031-9007/94/73(5)/684(4)$06.00

the energy of the loops. Hence, loops and dipoles of such sizes as can be formed by thermal fluctuations will always exist in a solid at a given temperature. The loops grow on average with increasing temperature and/ or stress until an unstable configuration is reached beyond which they can expand freely and give rise to ductile behavior. The equivalent description in 2D, which we shall henceforth adopt, is the dissociation of dipoles at the instability, causing a large increase in the density of "free" dislocations.

The creation of the dipoles induces in the medium a certain amount of plastic strain in addition to the elastic strain due to the applied loads. If the response of the medium to the applied stress is measured in terms of the total strain (as opposed to the purely elastic strain), then the increase due to the plastic component gives rise to an "effective shear modulus" (or susceptibility). This property is a good probe to distinguish between the elastic and plastic response since the effective modulus vanishes (or the susceptibility diverges) when plastic flow begins. The effective property differs very little from the normal definition well below T_c if one considers just a few existing dipoles. However, a cumulative effect is produced as the temperature is increased, because the presence of a few dipoles helps to nucleate more dipoles. This cooperative effect occurs due to the reduction in the self-energy of dipoles owing to the decrease in the effective modulus, caused by the plastic strain induced by dipoles of smaller separations.

The concept of the effective modulus can also be understood in terms of the total energy of a stressed elastic solid containing a small concentration of bound dislocation dipoles at finite temperature. The energy of the system can be expressed as $U^0 + U^d + U^{\text{int}}$ [12], where U^0 is the elastic energy of the body without the dislocations, U^d is the self-energy of the dipoles (calculated using the elastic strain fields of the interacting pair of dislocations forming the dipole), and U^{int} represents the "interaction" energy between the dislocation displacement fields and the applied tractions. This last term is nonzero when the dipoles are formed after the external loads have already been applied and gives rise to the well-known Peach-Koehler force. (In writing the total energy, we have ignored higher order contributions such as the dipole-dipole interactions.) In the K-T theory, the U^{int} term is rewritten in terms of U^0 and U^d by defining renormalized or effective moduli [11]. This identification is useful because instead of treating the effect of all fluctuations explicitly, it is sufficient to consider a test dipole of separation r and treat the effect of all dipoles with separation less than r (which are more numerous) by defining a scale-dependent susceptibility. The latter gives rise to a "screened" interaction between the dislocation pair with opposite Burgers vector forming the test dipole, effectively lowering its self-energy [13]. The divergence of the susceptibility at a certain temperature and stress signals the spontaneous dissociation of dipole fluctuations resulting in plastic flow.

We have so far discussed the homogeneous generation of dislocations leading to the BDT without taking into account the role of the crack. It is well known that the emission of dislocations should necessarily occur at the crack tip in order to blunt the crack [4]. The region of high stresses near the crack tip makes the latter a naturally favorable site for the spontaneous generation of dislocations.

Consider a dislocation-free elastic solid containing a sharp crack with traction-free surfaces loaded in mode I. Let σ be the shear stress acting on a slip plane that is assumed to be coincident with the crack plane for convenience. At finite temperature, there is a small concentration of bound dislocation dipoles generated by thermal activation. In the presence of an applied stress each dipole is in a metastable equilibrium [14]. Hence, a nonzero density of free dislocations exists in the thermodynamic limit even at low temperatures. We assume that this density is small and identify the BDT as caused mainly by the cooperative unstable dissociation of dipoles that leads to a rapid increase in the density of free dislocations. The main approximation in the K-T theory and in our model is the requirement of low initial dipole density which is the case when the core energy E_c is large. It is then sufficient to retain in the partition function only terms of the order of $\exp(-2E_c/k_BT)$. This approximation is particularly appropriate for highly brittle materials in which dislocations are known to possess narrow cores with large core energies.

The energy of an "unscreened" test dipole of separation r in an isotropic medium is [10,13]

$$U(r) = 2q^2 \ln\left(\frac{r}{r_0}\right) + 2E_c - \sigma b r \,, \qquad (1)$$

where $q^2 = [\mu_0 B_0/(\mu_0 + B_0)] b^2/2\pi = [\mu_0/(1-v)]b^2/4\pi$, μ_0 and B_0 are two-dimensional shear and bulk moduli in the absence of dislocations, v represents Poisson's ratio, b, r_0, and E_c represent the Burgers vector, core radius, and core energy of the dislocation, respectively. (The two-dimensional elastic constants and stresses are in units of N/m.) We follow the same procedure as in [9,10] and introduce a scale-dependent polarizability $\varepsilon(r)$ to take into account the screening due to all dipoles of separation smaller than r. The force between the dislocations comprising the test dipole is then $2q^2/r\varepsilon(r)$ instead of $2q^2/r$. $\varepsilon(r)$ is related to the susceptibility $\chi(r)$ by $\varepsilon(r) = 1 + 4\pi\chi(r)$. Using a continuum approximation to treat the smaller dipoles, we can write $\chi(r) = \int_{r_0}^{r} \int_0^{2\pi} n(r')\alpha(r')r'\,dr'\,d\theta$. Here, $\alpha(r)$ is the polarizability of a dipole of separation r at temperature T in the limit of zero stress and hence is the same as in the K-T theory; $n(r)$ is the number of pairs of separation r at finite σ and T. Defining $\beta = 1/k_BT$, the expressions for α and n are similar to those in [9]:

$$\alpha(r) = \frac{\beta q^2 r^2}{2},$$
$$n(r) = \frac{1}{r_0^4}\exp[-2\beta E_c - \beta V(r) + \sigma\beta b r], \quad (2)$$
$$V(r) = 2q^2 \int_{r_0}^{r} \frac{dr'}{r'\varepsilon(r')}.$$

Using Eq. (2) in the definition of the susceptibility we obtain a complicated integral equation for $\varepsilon(r)$ which has to be solved self-consistently. The determination of the critical radius (r_c) and temperature (T_c) beyond which $\varepsilon(r)$ increases rapidly is facilitated by transforming the integral equation to a set of coupled nonlinear equations [10]. Let us define

$$l = \ln\left(\frac{r}{r_0}\right), \; h(l) = \frac{\beta q^2}{\pi\varepsilon(r)},$$
$$y(l) = \left(\frac{r}{r_0}\right)^2 \exp\left(\frac{-2\beta E_c - \beta V(r) + \beta\sigma b r}{2}\right). \quad (3)$$

The function $y(l)$ is related to the probability of existence of dislocations at a given T and σ while $h(l)$ represents the screening due to dipoles of separation less than $r_0\exp(l)$. Differentiating the expression for $\varepsilon(r)$, and using Eq. (3), we obtain (to first order in y) a set of "renormalization transformation" equations,

$$\frac{dh^{-1}(l)}{dl} = 4\pi^3 y^2, \; \frac{dy}{dl} = \left(2 - \pi h + \frac{\beta\sigma b r}{2}\right)y, \quad (4)$$

subject to the boundary conditions $h(l=0) = \beta q^2/\pi$ and $y(l=0) = \exp(-\beta E_c + \sigma\beta b r_0/2\pi)$.

In the presence of stress, we obtain a nonautonomous system of coupled nonlinear equations describing the scaling behavior of thermally nucleated dipoles, unlike the autonomous set (the K-T equations) obtained in the absence of stress [10,13]. At low temperatures $h(l)$ is large and, therefore, dy/dl is negative for small values of σ. However, when σ is of the order of the theoretical shear strength $(2q^2/b^2)$, dy/dl changes sign, reflecting the fact that homogeneous nucleation of dislocations is possible at all temperatures when the applied stress is very large. This scenario is not relevant to the BDT as the theoretical shear strength is usually larger than the Griffith threshold so that cleavage occurs before spontaneous emission of dislocations. In order to make the analysis simpler, we approximate the last term in Eq. (4) by $\beta\sigma b r'/2$, where r' is a constant for a given temperature representing the average size of the dipoles. The error introduced is small since both r and r' do not differ much from r_0 at low stresses.

Let us henceforth consider the stress to be a constant, smaller than the theoretical shear strength. Then there is a critical temperature β_c corresponding to $r' = r_c$ such that $2 + \beta_c\sigma b r_c/2 = \pi h_c$ and $dy/dl = 0$. Above this temperature, both y and h^{-1} increase rapidly due to the spontaneous dissociation of thermally induced dipoles. The rapid increase in the free dislocation density occurs at the temperature

$$T_c = \frac{1}{2k_B}\left(\frac{q^2}{\varepsilon(r_c)} - \frac{\sigma b r_c}{2}\right). \quad (5)$$

When $\sigma = 0$, the result in Eq. (5) is the same as that obtained in the K-T theory [10] where it corresponds to a true fixed point of the renormalization equations. We can also recover the well-known Rice-Thomson result [4] for the stress-induced instability of a single dipole at zero transition temperature by equating $T_c = 0$ and $\varepsilon(r_c) = 1$ since there is no screening effect when only one dipole is present. We obtain $\sigma_c = \mu_0 b/2\pi(1-v)r_c$ which is the critical stress for homogenous nucleation of free dislocations. The emission occurs preferentially at the crack tip due to the stress concentration. If K_I is the elastic stress intensity factor [15] then the stress on a critical dipole in the vicinity of the tip is $K_I/\sqrt{2\pi r_c}$ and spontaneous dissociation occurs when this equals $\mu_0 b/2\pi(1-v)r_c$. We can eliminate K_I by assuming the fracture load to be given by the Griffith threshold $[K_I^2(1-v)/2\mu_0 = 2\gamma]$, where γ is the surface energy of the crack plane, and obtain $r_c/b = \mu_0 b/8\pi(1-v)\gamma$, which is the Rice-Thomson result in two dimensions [4].

We can carry out a linear stability analysis of Eq. (4) around the unstable point in the (y, h^{-1}) plane similar to the K-T analysis [11] and determine the "critical" trajectory that flows into the unstable point. We can then use the boundary conditions to express T_c as the solution of a nonlinear equation:

$$(\beta_c)^{-1} = \frac{\mu_0 b^3[1 - 2\pi\exp(-\beta_c E_c b + \sigma b^2\beta_c r_0/2)]}{8\pi(1-v)(1+\sigma b^2\beta_c r_0/4)}. \quad (6)$$

(In the above expression all quantities are expressed in 3D units for ease of calculation and it is done by transforming the 2D parameters appropriately using the Burgers vector as a suitable length scale.) It is not surprising that the transition temperature is expressed in terms of the elastic constants, the core parameters, and the applied stress because these quantities define the energy of a dislocation in any theory based on linear elastic description which represents a common starting point of all the models proposed so far. However, the existing models [4–6] assume thermally activated generation of dislocations and it is then necessary to introduce additional conditions, such as a specific rate of dislocation generation [6] in order to define a transition temperature. In contrast, the present model predicts a well-defined transition temperature without any additional assumptions and explains how a massive generation of dislocations suddenly becomes possible above a certain temperature for a given stress level in all materials.

We have calculated T_c from Eq. (6) using the same values of the material parameters as in the Rice-Thomson model [4] and using $E_c = (\mu b^2/2\pi)\ln(b/r_0)$ with $v = 0$. For typical values of the surface energy of brittle materials, the local stress at a distance r_0 from the crack tip is of the order of $\mu_0/10$. For Si, our approach predicts $T_c \sim 1000$ K for a partial dislocation and 5000 K for a full dislocation. If the crack tip stress is of the order of $\mu_0/100$, then $T_c \sim 1240$ K for a partial dislocation. For Cr and W, we obtain T_c equal to ~ 1600 K and 3000 K, respectively, using the parameters in [5]. The stress is assumed to be one-tenth of the respective shear modulus for both materials. These results are in qualitative agreement with observations and T_c is typically two-thirds of the melting temperature for very brittle materials.

The above results demonstrate clearly the importance of thermally induced dislocation screening when studying the problem of BDT. It is not surprising that thermal fluctuations play such a crucial role since the phenomenon is strongly temperature dependent. This feature is best seen from experiments in Si single crystals [2]. The transition to ductile behavior is dramatic and occurs in a narrow temperature range of less than 5° accompanied by massive dislocation activity. Such a rapid and sudden dislocation generation is only possible if the generation process exhibits a thermally induced cooperative instability as described above. In materials where the dislocation mobility is a power-law function of the stress, a correlation between the strain-rate dependence of T_c and the temperature dependence of dislocation mobility can be deduced by combining stress-induced dynamical effects in the instability criterion derived above. This is a separate additional feature of the BDT distinct from the sudden massive dislocation generation that is responsible for the transition and is described elsewhere.

In dislocation theory, homogeneous generation of dislocations is usually considered to be impossible, based on the energy needed to nucleate a single dislocation loop. Hence, thermal effects are generally ignored. Both the K-T and the present model show how quite a different conclusion is reached when more realistic assumptions considering many dislocations and possible interactions between them are made. The concepts discussed here provide an alternate mechanism for dislocation generation at finite temperatures. This mechanism is different from the stress-driven Frank-Read sources in which preexisting dislocations generate future dislocations, a mechanism that is known to exist in all solids. This approach may help to resolve, in general, the dilemma in explaining the rapid increase in dislocation density at low values of plastic strain in a material containing an initially low density of Frank-Read sources.

The authors gratefully acknowledge helpful discussions with Professor J. Bassani, Professor D. Luzzi, and Professor G. Tichy. This research was supported by the U.S. Air Force Office of Scientific Research Grant No. F49620-92-J-0019.

[1] J.F. Knott, *Fundamentals of Fracture Mechanics* (Revised) (Butterworths, London, 1979).
[2] C. St. John, Philos. Mag. **32**, 1193 (1975).
[3] H. Vehoff, in *High-Temperature Ordered Intermetallic Alloys V*, edited by I. Baker *et al.*, MRS Symposia Proceedings No. 288 (Materials Research Society, Pittsburgh, 1993), p. 71.
[4] J.R. Rice and R. Thomson, Philos. Mag. **29**, 73 (1974).
[5] J.R. Rice, J. Mech. Phys. Solids **40**, 239 (1992).
[6] J.R. Rice and G.E. Beltz, J. Mech. Phys. Solids **42**, 333 (1994).
[7] P.B. Hirsch, S.G. Roberts, and J. Samuels, Proc. R. Soc. London A **421**, 25 (1989).
[8] R. Thomson, Solid State Physics (Academic, New York, 1986), Vol. 39, p. 1.
[9] J.M. Kosterlitz and D.J. Thouless, J. Phys. **6**, 1181 (1973).
[10] A.P. Young, Phys. Rev. B **19**, 1855 (1979).
[11] D.R. Nelson and B.I. Halperin, Phys. Rev. B **19**, 2457 (1979).
[12] R.W. Lardner, *Mathematical Theory of Dislocations and Fracture* (University of Toronto Press, Toronto, 1974).
[13] A.P. Young, in *Ordering in Strongly Fluctuating Condensed Matter Systems*, NATO Advanced Study Institute, Ser. B, Vol. 50, edited by T. Riste (Plenum, New York, 1980), p. 271.
[14] C.E. Bottani and M. Jacona, J. Phys. Condens. Matter **1**, 8337 (1989).
[15] J.R. Rice, in *Fracture: an Advanced Treatise*, edited by H. Liebowitz (Academic, New York, 1968), Vol. 2, p. 191.