JOURNAL OF MATERIALS SCIENCE 13 (1978) 128–142

Brittle fracture in a ductile material with application to hydrogen embrittlement

R. THOMSON
National Bureau of Standards, Washington D.C., USA

A physical model of fracture in materials is developed which features a brittle crack imbedded in a plastically deformed medium. This model is presented as an alternative to fully ductile failure by hole growth, and general criteria for the two alternatives are discussed. One of these criteria for the existence of an atomically sharp crack is that the dislocation content near the crack tip be limited by the inhomogeneous character of dislocation slip in the crystal. With the dislocation distribution characteristic of Mode III fracture, we derive expressions for the fracture toughness as a function of material parameters. We have extended the theory to the case of hydrogen embrittlement in steels and compare our theoretical predictions with experimental work by others.

1. Introduction

Fracture in materials can occur by either of two physically distinct mechanisms. In the first mode, new surface is formed, and the crack advances by wholly plastic processes, as manifested most commonly by hole growth [1]. In the second mode, fracture occurs by the continuous opening of an atomically sharp crack in cleavage [2]. This mode is variously called brittle fracture or lattice decohesion, but we shall often refer to it also as classic fracture. Although hole growth fracture is a common form of failure in practical materials, brittle fracture hardly ever occurs in the pure form of simple cleavage, and consequently is usually thought to be of little more than academic interest, at least in metals. Indeed, the prevailing opinion is that fracture in metals is a ductile hole growth phenomenon with crack initiation occurring at the point of maximum triaxial stress ahead of a well rounded main crack. However, it is the purpose of this paper to show that brittle fracture is still a valid and useful concept in ductile materials, and that a modified form of classic fracture can describe the process of hydrogen embrittlement in high-strength steels, as well as other fracture phenomena in materials with finite ductility.

We will show that in ductile materials the plastic zone acts to *shield* a brittle crack, embedded within the plastic zone, from the full force of the external stress field. Thus the apparent macroscopic stress intensity factor is characteristic of a tough material. However, the material toughness is at the same time controlled by the properties of the cohesive zone of a brittle crack core. The brittle crack at the core is also consistent with the presence of a well developed plastic zone, provided two essential conditions are met; (1) the brittle crack is stable against spontaneously generated blunting and dislocation formation at the crack tip, and (2) the total shear strain at the crack tip is less than one, so that the dislocation density in the vicinity of the crack tip is significantly smaller than one dislocation per atomic area. We discuss these two conditions briefly in turn.

The first condition was the subject of an earlier paper by Rice and Thomson [3], of which this paper is a natural follow-on. In that paper, we showed that the necessary and sufficient condition for an atomically sharp crack to be stable in a crystal against spontaneous blunting was that the activation energy for the formation of dislocation loops out of the tip of the crack be greater than zero. It was shown in that paper that in the f c c metals, dislocation formation is spontaneous, but that in most other classes of materials, the energy for dislocation formation is greater than zero and often very high. Iron is a borderline case, where

stability is not clear. In computer studies in iron, Kaninen and Gehlen [4] have confirmed that iron may be unstable to dislocation generation at the equilibrium stress. We discuss the effects of hydrogen on fracture in steel later, and we believe that, in the presence of hydrogen, the surface energy in iron is sufficiently lowered that a sharp crack should have no difficulty in remaining stable against spontaneous blunting.

The second condition is also important, since if larger numbers of dislocation can be formed in the vicinity of the crack tip by plastic processes, they will round out the atomically sharp configuration there. Indeed, in the mathematical solution for a crack in continuum plasticity, an initially sharp crack develops a substantial crack opening displacement many atom spacings wide due to the large strain developed at the tip. However, many real materials cannot develop such large amounts of strain homogeneously on an atomic level, because the dislocation sources of strain are distributed very inhomogeneously within the lattice on a scale usually determined by the microstructure of the material. Thus, although on a macroscopic scale where the crystalline microstructure is averaged out, a crack should have an opening displacement qualitatively like the continuum treatment. The crack tip itself need not be blunted (see Fig. 1). At this point, we assume that the crack configuration is like that given in Fig. 1, and return at the end of the paper to a discussion of the necessary criteria for this. Since strain (actually, its gradient) is closely related to the dislocation content, if the strain near the tip is limited, so is the dislocation content. Thus our basic assumption is equivalent to assuming that the dislocation content near the crack tip is limited by strain hardening, and that it cannot build up to the large values implicit in the continuum plastic solution. In dislocation terms, our assumption is related to the tendency of heavily deformed materials to undergo dynamic recovery with the formation of dislocation cells [5]. The interior of the cells are found to be relatively distortion free, while the cells are bounded by incipient sub-grain boundaries containing large densities of dislocations.

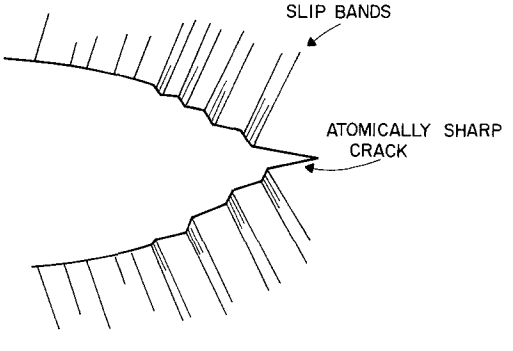

Figure 1 Schematic drawing of a sharp crack in the field of an inhomogeneous dislocation and slip line distribution. The gross features of the crack are broadened by the plastic deformation, but the tip of the crack remains atomically sharp.

As a result of the interplay of these two conditions on the stability of a brittle crack, we view fracture as a competition between the two modes. A material will be subject to classic brittle type fracture (though modified by the considerations of this paper) if a brittle crack is stable against the two types of blunting outlined. Otherwise, the material will fail by a ductile hole growth mechanism.

Our model yields a thermodynamic criterion for the onset of crack growth, even though we shall often couch our work in terms of forces at the crack tip instead of free energies. The kinetics of crack growth will involve much more complicated phenomena such as creep in the plastic zone as discussed by Hart [6], or chemical reactions at the crack tip [2], etc. Our criterion thus forms a lower limit for the onset of fracture, since onset cannot be observed until some finite limiting value of crack velocity has been achieved which is experimentally observable.

Various authors, most notably in recent years Oriani and co-workers [7], have discussed hydrogen embrittlement in terms of lattice decohesion, but without a treatment of the plastic zone. Since even so brittle a phenomenon as hydrogen embrittlement involves large amounts of plastic deformation, most authors have preferred to look for its cause in terms of ways in which the presence of hydrogen can modify the hole initiation and crack process [8]. In our work, the plastic zone forms a crucial and self-consistent part of the discussion, but it is assumed to enclose a core crack which is atomically sharp.

2. The plastic shielding factor

The basic assumptions of our work are illustrated in Fig. 2, where we sketch the stress in a material as a function of distance from the crack tip. Region I is the region of elasticity, region II is the region of continuum plasticity, and region III is

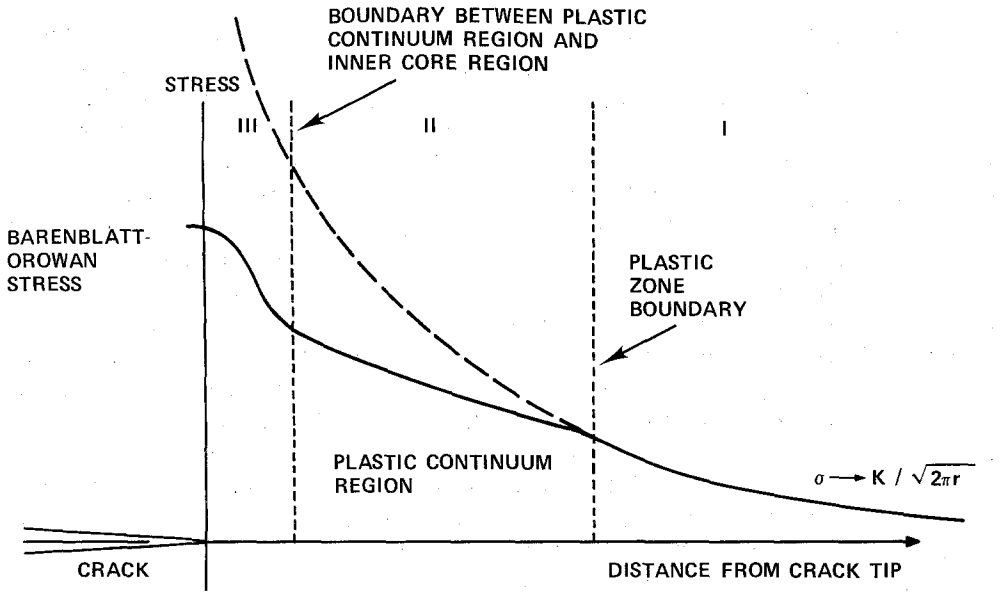

Figure 2 Sketch of the stress in a medium containing a core brittle crack shielded by a plastic zone. The continuum solution cannot be projected into the core crack region because of the inhomogeneous distribution of dislocation sources at an atomic level.

the region dominated by the elastic field of the core crack.

2.1. The plastic zone

We are concerned in this paper with the Mode I or opening crack. Although considerable theoretical work has been expended on the Mode I opening mode plastic solutions, the analysis is difficult, and recourse must be made early in the analysis to numerical solutions. Here, we are interested primarily in the physical nature of the fracture problem, and prefer to work with a simplified model, because it allows us to develop a relatively analytic approach. We shall therefore use the anti-plane strain solution of Rice [9] for the Mode III crack, because it permits us to write analytic expressions, and trust that the predictions it makes can be applied directly to the Mode I crack. For this purpose, we assume a strain hardening material of form

$$\sigma = \sigma_0(\epsilon/\epsilon_0)^n, \epsilon > \epsilon_0 . \qquad (1)$$

n is the work hardening index, typically in the range 0.1 to 0.3. σ and σ_0 are the shear stress measured at some point in the plastic zone and critical shear stress, respectively. ϵ and ϵ_0 are similar quantities for the engineering shear strains and n is the strain hardening coefficient. (The engineering strain is defined by $\epsilon = \partial u_z/\partial\lambda$ where λ is the principal shear direction, and u_z is the anti-plane displacement.

We shall work from the solutions presented by Rice [9]. According to Rice, the strain is a constant on a circle whose radius is R and whose centre is on the cleavage plane at a distance X in front of the crack tip, where

$$R(\epsilon) = \frac{K^2}{2\pi\sigma_0^2}\left(\frac{\epsilon_0}{\epsilon}\right)^{1+n}$$

$$x(\epsilon) = \frac{1-n}{1+n} R(\epsilon) \qquad (2)$$

The solution is depicted in Fig. 3, which shows the direction of the principal strain, ϵ. In Equation 2, K is the apparent stress intensity factor as measured in the elastic region, and defined by the asymptotic behavior of the stress as a function of distance, ρ, from the apparent position of the crack, given by

$$K = \sigma\sqrt{2\pi\rho} \qquad (3)$$

We shall write an approximation of Equation 2 which will be necessary in order to arrive at expressions we can handle analytically, namely that

$$x(\epsilon) = R(\epsilon) \qquad (4)$$

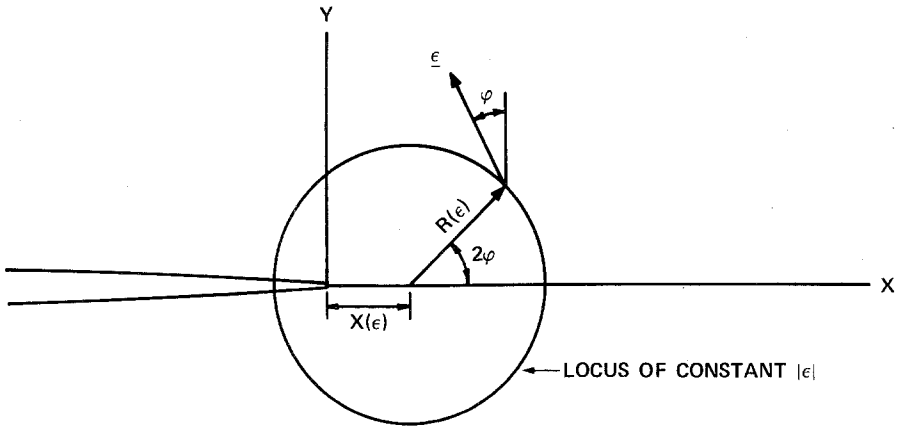

Figure 3 Strain in the plastic zone, from Rice, the solution for the strain in the plastic zone is a circle of radius $R(\epsilon)$ and centred on the fracture plane a distance $X(\epsilon)$ ahead of the crack tip. The angle of the principal stress is the angle ϕ measured from the vertical as shown.

In this case, the circles of constant strain all touch the crack tip, as shown in Fig. 4. When the equation of these circles is expressed in cylindrical co-ordinates, then we can write the strain in the entire plastic zone as

$$\frac{r}{2\cos\theta} = R,$$

$$\epsilon_{3\theta} = \epsilon = \epsilon_0\left(\frac{K^2\cos\theta}{\pi\sigma_0^2 r}\right)^{1/1+n}, \qquad (5)$$

where r and θ are cylindrical co-ordinates with the crack tip as origin, and the cleavage plane is the X-axis.

2.2. The dislocation field

We need to convert the strain field, [4], into a dislocation density field. This dislocation field is often termed the "geometrical dislocation density", and the actual dislocation density may in fact be larger because of the presence of random dislocations whose strain fields cancel at large distances. The geometric dislocation density is determined from the average continuum plastic strain field by the theory of dislocation continua, as outlined, for example, in the review article by DeWit [10]. The geometric component of the screw dislocation density, N, is related to the *elastic portion* of the strain field by

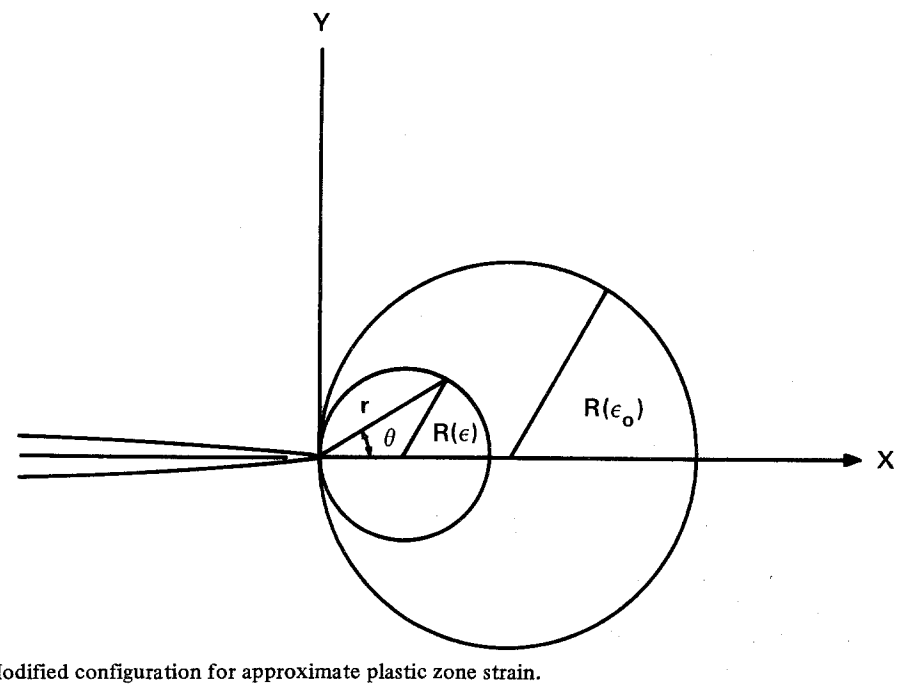

Figure 4 Modified configuration for approximate plastic zone strain.

$$N = \frac{1}{b}\left(\frac{\partial \epsilon_{32}^{el}}{\partial x} - \frac{\partial \epsilon_{31}^{el}}{\partial y}\right) \quad (6)$$

b is the Burgers vector of the dislocations, ϵ_{31}^{el} is the elastic part of the total strain, given by

$$\epsilon_{31}^{el} = \sigma_{31}/2\mu \text{ etc.} \quad (7)$$

The dislocation density in the plastic zone is then given by

$$N = \frac{\epsilon_0}{b}\left(\frac{K^2 \cos\theta}{\pi\sigma_0^2}\right)^{n/1+n} \frac{\cos 2\theta}{r^{1+n/1+n}} + \frac{3}{2b}\frac{n\epsilon_0}{1+n}\left(\frac{K^2}{\pi\sigma_0^2}\right)^{n/1+n} \frac{\sin\theta \sin 2\theta}{\cos^{1/1+n}\theta} \frac{1}{r^{n/1+n}} \quad (8)$$

Near the crack tip, the first term is dominant, and we approximate the screw dislocation density near the tip as

$$N = -\frac{\epsilon_0}{b}\left(\frac{K^2 \cos\theta}{\pi\sigma_0^2}\right)^{n/1+n} \frac{\cos 2\theta}{r^{1+2n/1+n}} \quad (9)$$

2.3. The core crack

For the reasons explained in the Introduction, the dislocation field cannot be extended to atomic dimensions, and we arbitrarily assume a cut-off distance, R_c, below which local order in the crystal is assumed to be maintained. In this region, the crack is rigorously atomically brittle in configuration. If the solution is to be self-consistent at the cut-off radius, R_c, then the stress generated at R_c by the core crack must be the same as the stress generated at R_c by the continued external stress and dislocation field. If K_c is the stress intensity of the core crack, then

$$\sigma(R_c) = \frac{K_c}{\sqrt{(2\pi R_c)}} \quad (10)$$

and from Equations 1 and 2

$$\frac{\sigma(R_c)}{\sigma_0} = \left(\frac{K^2}{2\pi\sigma_0^2 R_c}\right)^{n/1+n} \quad (11)$$

Since we shall be dealing with a thermodynamic criterion for fracture, K_c is related to the true surface energy of the material, γ, by the Griffith relation

$$K_c^2 = 2\gamma Y \quad (12)$$

Y is Young's modulus. In Equation 12, we have neglected ν^2 (ν is Poisson's ratio) in relation to unity, and have introduced a Mode I fracture, whereas the previous equations relate to Mode III. This is an unavoidable inconsistency in our work since we wish to derive analytical results for "real" fractures (i.e. Mode I), but cannot solve for the dislocation field of a Mode I fracture without resorting to numerical methods.

By combining Equations 10 and 11, we can eliminate $\sigma(R_c)$, and write a simple relation between K and R_c

$$K^2 = 2\pi\sigma_0^2\left(\frac{\gamma Y}{\pi\sigma_0^2}\right)^{(1+n)/2n} \frac{1}{R_c^{(1-n)/2n}} \quad (13)$$

Associated with the cut-off radius is a maximum density of dislocations, which is obtained from Equation 9. In these terms, Equation 13 becomes

$$K^2 = \left\{\left(\frac{\eta}{b\epsilon_0^3\mu^2}\right)^{1-n}\left(\frac{\gamma Y}{\lambda}\right)^{1+2n}\right\}^{1/3n}$$

$$\lambda = 2^{(1-3n)/(1+n)}\pi^{n(n-1)/(1+n)(1+2n)}$$

$$\eta = Nb^2 \quad (14)$$

η is a dimensionless parameter, giving the maximum dislocation density in atomic units. σ_0 has been converted to ϵ_0 by means of $\sigma_0 = \mu\epsilon_0$, where μ is the shear modulus. In deriving Equation 14, we have also averaged out the angular dependence explicit in Equation 9.

Equations 13 and 14 are the central results of the paper, and are interchangeable, depending upon whether the important physical parameter is the cut-off radius or the maximum dislocation density. We shall work with the latter.

The equation for K has the correct qualitative dependence on the physical parameters, Y, ϵ_0, γ and η. It varies directly with γ, consistent with the empirical result that when γ is larger, K must increase. However, we note that the dependence on γ also depends upon the strain hardening property of the material, and that K^2 is not proportional to the first power of γ as in the elastic case. Likewise, increasing the yield strength decreases the toughness, as is generally observed. Again, we find that the details of this dependence vary with the strain hardening coefficient. Although η is a special feature of our theory not studied experimentally, clearly as η decreases, the material should approach its perfectly brittle condition, as Equation 14 predicts.

Useful physical insight into K is obtained by noting its variation as the strain hardening coefficient, n approaches its limits. As $n \to 1$, in Equation 14, K takes on its brittle fracture form, Equation 12. In the perfect plasticity limit

TABLE I Some values of K as predicted in Equation 14 (MPa m$^{1/2}$)

n/η	1	10^{-2}	10^{-3}	10^{-4}
0.05	1.65×10^{16}	7.66×10^{-9}	5.22×10^{-6}	3560
0.1		4.56×10^{-4}	1.44×10^{-3}	
0.15	6.42×10^{-4}	829	94.2	10.7
0.16			67.0	
0.2		112	24.1	5.19
0.21		64.8		
0.25	337		10.6	3.36
0.3		15.1	6.17	2.52
1.0	0.932	0.932	0.932	0.932

($n = 0$), the stress in region II of Fig. 2 is a constant, independent of the external stress, and $K \to \infty$. In this case, R takes on a value which is also independent of the external stress, and the core crack is completely screened in an unphysical way. As n decreases from unity, however, K builds up to very large values compared to the brittle limit because of the screening of the core crack by the plastic region. A multiplication factor which is a measure of the effective surface energy of the material is given from Equation 14 by

$$\frac{K^2}{K_c^2} \cong \left\{ \frac{\eta}{\epsilon_0^3 \lambda} \frac{\gamma}{\mu b} \right\}^{(1-n)/3n} \tag{15}$$

A number of aspects of the current theory are more qualitative than we would like, such as the use of Mode III screw dislocation field for the plastic zone. However, one fine point can be rather easily treated, namely, to show that the region inside the plastic zone within which no dislocations are allowed does indeed display the $K_c/\sqrt{(2\pi r)}$ stress dependence which is characteristic of the core crack. This proof is given in the Appendix.

Table I lists some values for K as predicted by Equation 14 for various values of η and n for a representative steel. The ASI 4340 steel on which Oriani and Josephic performed their experiments has a K value in air of about 62 MPa m$^{1/2}$, although at these high values of K, fracture was at least partially by hole growth.

3. Dislocation drag forces on the crack

In deriving the equilibrium external stress on a crack, we have calculated the static force on a crack tip from its plastic zone, but have not considered the energy required for the crack to regenerate its plastic zone as it moves, leaving behind a wake of plastic deformation. The energy to form this wake will constitute a drag force on the crack as it moves. We shall estimate this drag by calculating the energy required if the crack leaves all of its plastic deformation as a wake. On the basis of elementary geometry, the force due to the wake is

$$\frac{dE}{dx} = -2E_0 \int_R^{R'} N(y)dy. \tag{16}$$

E_0 is the line energy of a dislocation, E is the total stored energy in the wake. $N(r)$ is the density of "geometric" dislocations in the plastic zone, x is the direction of motion of the crack, and y is the direction vertical to the crack plane, Fig. 5. R is the radius of the depleted zone, below which the dislocation density is flattened, and R' is the radius of the plastic zone. The factor of 2 in Equation 16 derives from the fact that the plastic zone extends above as well as below the crack surface.

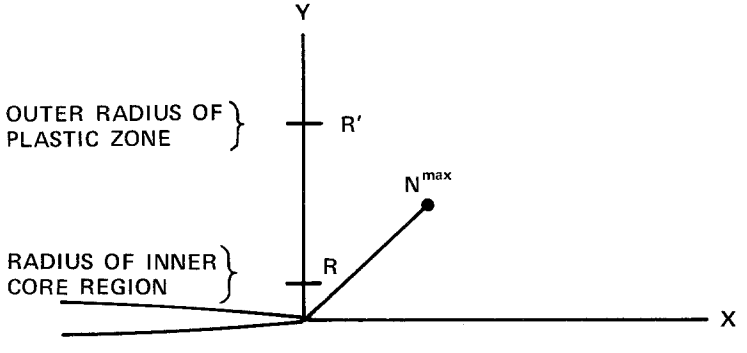

Figure 5 Configuration for calculation of plastic zone drag. N^{max} is the value at a given value of y where $N(r)$ reaches a maximum value.

We take the dislocation density from Equation 9, and slightly over estimate the results by setting $\cos\theta = 1$. Then

$$\frac{dE}{dx} = -2E_0N_0 \int_R^{R'} \frac{dy}{y^{(1+2n)/(1+n)}}$$

$$N_0 = \frac{\epsilon_0}{b}\left(\frac{K^2}{\pi\sigma_0^2}\right)^{n/(1+n)} \quad (17)$$

Taking the upper limit to be infinity, which we can do without significant error, taking the value of R from Equation 2, and K from Equation 14, we have

$$\frac{Y}{K^2}\frac{dE}{dx} = \frac{2^{(1-4n^2+3n^3)/2n(1-n^2)}}{\pi^{(1-2n)/3(1+n)}} \left\{ \epsilon_0{}^{3(1-n)}\left(\frac{\mu b}{\gamma}\right)^{1+n} \frac{1}{\eta^{1-2n}}\right\}^{1/3n} \quad (18)$$

In Equation 18 we have written the ratio of the dislocation drag forces to the crack extension force derived in Equation 14. We have grouped dimensionless quantities together for greater ease of interpretation. Again, we have some difficulty in choosing which elastic constant to use. Using μ gives a lower limit, Y gives an upper limit.

The quantity $\mu b/\gamma$ is a number of order 10 to 100 which we have displayed for a number of materials in Table II. Since even a very strong alloy will have $\epsilon_0 < 10^{-2}$, and since any reasonably ductile material will have a value of $\eta > 10^{-4}$, it is seen that Equation 18 will usually (but not always) be a number less than 1, and the dislocation drag term is expected to be small in relation to the equilibrium crack extension force. The main uncertainty in this conclusion is the fact that the "geometrical" density considered in Equation 16 is likely to be only a fraction of the total dislocation density which includes the random dislocation component. The conclusion still seems to be safe in many cases, however, even if the geometrical component is only 10% of the total. Equation 18 does not include the kinetic effects of dislocation drag, consistent with our attempt to make a thermodynamic estimate of the critical K. Kinetic drag force terms have been estimated by Burns [12] in a single crystal hexagonal metal, where glide and cleavage planes are identical. They have also been studied for more general creep case by Hart [6]. For slow crack growth, these creep effects will be important. We return to a discussion of thermodynamic threshold versus kinetics in a later section.

TABLE II Values of $\mu b/\gamma$ for Selected Materials*

Materials	$\mu b/\gamma$	Yb/γ
Pb	5.765	15.08
Au	4.796	13.96
Al	8.515	23.89
Ni	10.798	29.40
Na	3.752	6.380
Fe	8.726	31.60
W	25.79	66.06
LiF	26.10	67.79
NaCl	26.00	57.38
MgO	28.68	72.48
Al_2O_3	18.45	47.97
Be	34.21	88.95
Zn	32.90	85.54

*Because of the uncertainty in the appropriate elastic constant to be used in our work, we list two values, one for the shear modulus, and one for Young's Modulus. Values of γ are taken from Rice and Thomson [3].

4. Effect of external atmosphere

4.1. Surface energy effects

In Equation 14, we related the stress on the core crack to the surface energy of the opening crack by the original argument of Griffith as modified by Barenblatt. In an external atmosphere (we shall assume a diatomic gas) the surface energy is lowered in a way which was applied to cracks by Petch [13]. We shall extend his treatment in order to apply it more fully. The thermodynamic criterion for equilibrium of a core crack of unit length is given by

$$\delta F_c = \delta F_{el} + \delta F_s + \delta F_g = 0 \quad (19)$$

F_c is the total free energy of the region in the immediate region of the core crack, and not including the plastic zone. (We deal in this section only with the response of the core crack to the shielded stress of its immediate surroundings, and of the surface created by the opening of the sharp crack. Thus, free energy changes do not include the plastic zone.) δF_c is the change in the free energy when the crack is advanced by one atomic distance all along its length. We shall write $\delta F_{el} = Gb_0$, where G is the crack extension force of the core crack, and b_0 is the lattice constant. δF_s is the surface contribution to the free energy from the new surface, and δF_g is the free energy change in the external gas. By standard thermodynamic arguments [14], if the crack moves forward by

one atomic spacing, the surface and gaseous free energies change by the amounts

$$\delta F_g = -\tfrac{1}{2}\mu_g \Gamma(2b_0)$$
$$\delta F_s = 2\gamma b_0 + 2b_0 \Gamma \mu_s \qquad (20)$$

It is assumed the crack is a line one meter in length, and advances one atomic spacing all along its length. Γ is the number of gas atoms on the surface per m^2, assumed to be chemisorbed as an atomic species. μ_g is the chemical potential of molecular gas, and μ_s is the chemical potential of atoms on the surface. In equilibrium for a diatomic gas, $\mu_s = \frac{1}{2}\mu_g$. Thus at threshold,

$$G = 2\gamma \qquad (21)$$

where γ is the true thermodynamic surface tension of the covered surface.

When the gas is present, the Gibbs adsorption equation can be integrated to give

$$\gamma = \gamma_0 - \frac{kT}{2}\int_0 \frac{\Gamma(T)}{P}\,dP \qquad (22)$$

To find $\Gamma(T)$, we shall adopt the Langmuir Equation, although where an empirical isotherm is available, it could also be used. To obtain the Langmuir isotherm, we note that the partition sum of the adsorbed gas on the surface when it is immobile on that surface is

$$Z_s = Z_{conf} Z_g \qquad (23)$$

Z_g is the partition sum of the single atom on the surface relative to that atom bound in a gaseous molecule, and Z_{conf} is the configurational part of Z. We find that

$$\ln Z_{conf} = N_0 \ln N_0 - (N_0 - \Gamma)\ln(N_0 - \Gamma) - \Gamma \ln \Gamma + \Gamma E_b/kT + \ln \zeta_g$$
$$Z_g = e^{E_b/kT}\zeta_g \qquad (24)$$

where N_0 is the total number of surface sites. E_b is the binding energy of a hydrogen atom to the surface referred to its state in a free molecule. Aside from the Boltzmann factor, ζ_g is the partition sum of a gas atom at the surface site. If the atom is tightly bound and in its lowest vibration state, $\zeta_g = 1$. An upper bound corresponds to the adsorbed atom being only weakly bound in the two dimensions of the surface, when $\zeta_g = 2\pi a^2\, mkT/h^2$. Since the free energy of the surface per unit of surface, f_s, is given by

$$f_s = -kT \ln Z_s \qquad (25)$$

and the chemical potential of gas atoms on the surface is

$$\mu_s = \frac{\partial f_s}{\partial \Gamma} \qquad (26)$$

and $\mu_g = 2\mu_s$, we finally have

$$-2E_b + \mu_g = kT \ln\left(\frac{\theta^2}{(1-\theta)^2}\right) - kT \ln \zeta_g^2$$
$$\theta = \Gamma b_0^2 \qquad (27)$$

We also know μ_g for a perfect gas of diatomic molecules. It is given by

$$\mu_g = -kT \ln\left\{\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\left(\frac{kT}{P}\right)\zeta_{Rot}\right\} \qquad (28)$$

where m is the mass of the molecule, h is Planck's constant, ζ_{rot} is the partition sum for the rotator. The Langmuir isotherm then becomes

$$\frac{\theta^2}{(1-\theta)^2} = aP$$
$$a = \left(\frac{2\pi h^2}{mkT}\right)^{3/2}\frac{\zeta_g^2}{kT}\frac{e^{2E_b/kT}}{\zeta_{Rot}} \qquad (29)$$

and the surface tension becomes

$$\gamma = \gamma_0 - \frac{kT}{b_0^2}\ln\left(1 + \sqrt{(aP)}\right) \qquad (30)$$

This equation in its general form has been given by others, but our purpose in laying it out in detail is to note the isotope effect which shows up in the term for the diatomic rotator partition sum, ζ_{rot}. We shall apply it in the following section to the experiments which have been carried out by Oriani and Josephic on hydrogen and deuterium.

Substitution of Equation 29 into Equation 14 gives the dependence on γ to be

$$K = A\left\{\gamma_0 - \gamma_1 \ln(aP)\right\}^{(1+2n)/6n}$$
$$\gamma_1 = kT/2b_0^2 \qquad (31)$$

A is a constant which can be evaluated from Equation 14, and it is assumed that $aP \gg 1$. Another interesting quantity is

$$\frac{dK}{d(\ln P)} = -\frac{1+2n}{6n}K\frac{\gamma_1}{\gamma_0} \qquad (32)$$

$dK/d(\ln P)$ is of interest because it is the slope of the function, K, against pressure when plotted on log paper. In Equation 32 on the right hand side, we have given only the first term in a binomial expansion, valid when $\gamma_0 \gg \gamma_1 \ln(aP)$.

Oriani and Josephic [7] have given some information on this point for the brittle fracture under discussion by performing transient experiments in which no time delay effects have been detected. In this paper, we shall therefore assume that the fracture is brittle, and that the hydrogen is able to attack the opening crack at its tip directly from the gas phase. However, clearly this is a crucial point in our model, and additional experiments should be performed to settle the issue. (We note in passing that in recent work, Rice [16] has combined these two approaches by considering the brittle decohesion of grain boundaries buried in the matrix in a model of brittle fracture which does require bulk hydrogen transport to nucleation site ahead of the main crack.)

4.2. Hydrogen embrittlement

We will take hydrogen embrittlement in high-strength steel as a possible application of our theory. Part of the reason is that not only is this a brittle form of fracture in a material with considerable remaining ductility, but it also involves the chemistry of atmosphere assisted fracture, all of which are included in our work. We emphasize again that we work only with the threshold for fracture, since we have not addressed slow crack growth.

Hydrogen embrittlement covers a variety of phenomena. In the higher strength materials, or where the internal hydrogen concentration is not too high, the fracture is primarily by means of a brittle form of intergranular fracture with the fracture following the prior austenite boundaries of the steel. (See for example, the comments of Oriani and Josephic.) In other cases, the fracture clearly involves hole growth. Our discussion is aimed only at the former case. In this section, we shall deal successively with the main features of this brittle form of hydrogen embrittlement.

4.3. Intergranular fracture

Fractographs show very clear and sharp grain boundary facets on prior austenite boundaries consistent with our model of an underlying brittle fracture [7]. Apparently the prior austenite boundaries are weakened by the presence of large precipitates on the boundaries. The fact that the fracture is brittle is attested by the sharp facets typical of the fractographs. Perhaps these boundaries still retain some misfit or segregated chemical species even in the fully transformed state which lowers their γ relative to other paths in the crystal.

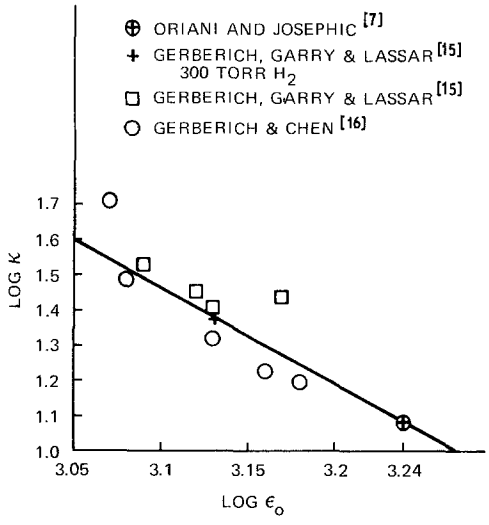

Figure 6 Yield stress data of Gerberich and Chen and others plotted to determine value of n in Equation 14.

4.4. Yield stress dependence

Increasing the yield stress increases the hydrogen susceptibility. This effect appears explicitly in Equation 14. Its precise power dependence involves the strain-hardening coefficient, n. The data available are sketchy. For our purpose, we combine in Fig. 6, the data of Oriani and Josephic [7] and Gerberich *et al.* [15] on external hydrogen effects with the studies by Gerberich and coworkers for internal cathodically charged hydrogen [15, 17]. From Fig. 6, we find the value $n = 0.16$ for the strain-hardening coefficient.

4.5. Grain boundary dependence

Gerberich *et al.* [15] have also studied the effect of grain size on the hydrogen threshold, but we believe this effect is primarily due to the concomitant variation of the yield stress, and is apparently consistent with their findings. Unfortunately, the grain-size dependence is badly confused by the whole question of what the traps for hydrogen in iron are, and how they might change with plastic flow. In addition, Gerberich *et al.* [15] agree that the grain-size dependence may have to do with the distribution of other impurities at the grain boundary, and the additional weakening effect which that might bring. Yet another complicating effect is that the fracture occurs on prior austenite boundaries, so that the two types of grain bound-

aries must be separated. We believe that hydrogen embrittlement studies carried out under conditions where the grain boundary energies are systemmatically varied independently, would be an interesting experiment, as suggested by Gerberich *et al.*

4.6. Acoustic emission

Gerberich *et al.* [15] report acoustic emission from hydrogen embrittlement, and in our view, this is due to the inherently uneven rate of cracking when the fracture plane must follow the grain boundaries.

4.7. Internal versus external hydrogen

Gerberich *et al.* [15] reports that there is a rough equivalence between hydrogen embrittlement induced by cathodically induced hydrogen in which the hydrogen is distributed in the interior of the material, and embrittlement caused by external hydrogen atmospheres. In our view, this result is not unexpected, because the grain boundaries of the prior austenite which form the fracture surface of the cathodically charged material should provide ample hydrogen to modify the final fracture surface energy. In the case of cathodically charged hydrogen, however, the general problem of the role of the prior austenite boundaries as traps for hydrogen in iron makes a quantitative thermodynamic treatment of the grain boundary energy as a function of dissolved hydrogen more difficult. Johnson [18] has emphasized the problem of trap distributions in modifying the effective supersaturation of H in the matrix.

4.8. Hydrogen pressure dependence

The most important data from our point of view are those reported by Oriani and Josephic [7], and Gerberich *et al.* [15] on the variation of the threshold of fracture with the pressure of external hydrogen and deuterium. Both these sets of authors, however, were guided by theoretical models quite different from our own. In Fig. 7, we have plotted their data again. Also, we have displayed some new data in new work to be published by Oriani and Josephic on the same material, but taken over a wider range of hydrogen pressure and which avoids certain systematic experimental errors in their earlier work connected with the threshold K measurement. This most recent and more accurate data plots as a reasonably straight line in the graph Fig. 7 over its entire length. We note that both from Equation 32 and from physical reasoning, for small pressure, P, the curve of K versus log P must become asymptotic

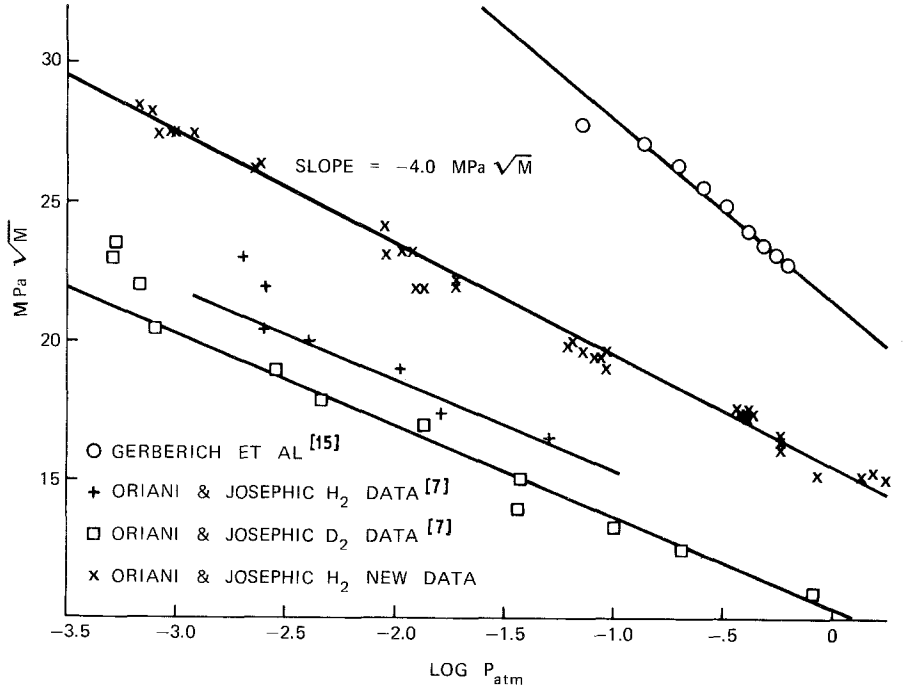

Figure 7 Data of Oriani and Josephic [7] and Gerberich *et al.* [15]. Plotted according to Equations 14 and 32.

TABLE III Calculations of $dK/d(\log P)$ in Equation 32 for various values of n and η (MPa$\sqrt{m}$)*

n/η	10^{-2}	10^{-3}	10^{-4}
0.05	4.69×10^{8}	3.20×10^{5}	218
0.1	1.52×10^{3}	48.2	1.52
0.14		3.54	
0.15	20.0	2.27	0.258
0.16		1.54	
0.18	4.58		
0.2	2.18	0.468	0.101
0.218	2.27		
0.25	0.562	0.177	0.056
0.3	0.225	0.0092	0.037
1.0	7.78×10^{-3}	7.78×10^{-3}	7.78×10^{-3}

*Material parameters are those characteristic of 4340 steel;
$\mu = 8.6 \times 10^{10}$ N m^{-2}, $\gamma = 1.975$ N m^{-2}, $\epsilon_0 = 10^{-2}$, $Y = 2.2 \times 10^{11}$ N m^{-2}, $b = 2.1477$ Å, $T = 294$ K, $b_0 = 2.48$ Å.

to its limiting value in vacuum. Thus the slope must decrease for large negative values of $\log P$, as shown qualitatively by the curve of Gerberich.

In Table I, we have already displayed predicted values of K for a material with parameters corresponding to the steel used by Oriani and Josephic and by Gerberich. Oriani and Josephic report a value for K in air in their 4340 steel of about 62 MPa m$^{1/2}$. If we choose $\eta = 10^{-3}$, which we believe is the most reasonable value, this corresponds to a value of n of about $n = 0.16$ ($\eta = 10^{-3}$) from Table I. We can also estimate the value of n from the slope of the curves in Fig. 7 from Equation 32. This slope is displayed in Table III for the same material parameters as before. From the Table, the best value of n is approximately $n = 0.14$ ($\eta = 10^{-3}$). Thus, from the various lines of argument, a value of n in the vicinity of $n \simeq 0.15$ is a roughly consistent value for this material in our theory with the parameters chosen. For $\eta = 10^{-2}$, a somewhat larger value near $n = 0.2$ is consistent. We are struck by the remarkably consistent predictions for n made from Tables I, III, and from the yield stress dependence of Fig. 6.

4.9. Isotope effect

Oriani and Josephic [7] report that measurable differences can be detected between hydrogen and deuterium. We note in Equation 29 that a depends explicitly on the isotopic mass of the molecular species and in addition, both ζ_g and ζ_{Rot} depend upon the mass implicitly. For $T = 298°$ C, we have

$$\zeta_{Rot} = \begin{cases} 3.04 & H_2 \\ 4.68 & D_2 \end{cases} \qquad (33)$$

There is considerable uncertainty about what value to choose for ζ_g in Equation 24, but it probably is closer to the tightly bound limit than to the loosely bound one. When $\zeta_g = 1$, then the curve of K_{D_2} versus P should be shifted by a factor of approximately 4 relative to that of H_2 on the P axis. If we assume that $\zeta_g/M^{3/2} \sim 1$, then ζ_{Rot} is the dominant term, and the shift is about 1.5. We note that the curve of Oriani and Josephic in Fig. 8 is actually shifted by a factor of about 3 in the predicted direction, which is within the expected range. Oriani in a private communication has pointed out that a similar shift in the solid solubility of H_2 relative to D_2 is predicted by our general treatment, and that none is observed by Heumann and Primas [19]. We are surprised that the mass effects exactly cancel out in the partition sums for gas and solid, and offer two comments: First, perhaps the experimental result should be rechecked; and second, the state of the adsorbed atom on the surface is by no means identical to that of the same atom in the bulk. Thus, cancellation effects in the bulk need not apply quantitatively on the surface also.

5. Slow crack growth and the thermodynamic threshold

Our theory is a thermodynamic theory of the onset or threshold for crack growth. It establishes the necessary energy balance for possible fracture growth, and does not address kinetics. This is clear enough from Section 4, where the surface energy balance is treated in an explicit thermodynamic manner, but not so for the treatments of the plastic zone in Sections 2 and 3. Nevertheless, these treatments, since they are fundamentally continuum-based are of necessity in the same class.

For example, we set the criterion for crack advance to be that the forces opening the atoms in the cohesive region meet the Barenblatt condition, that is, the forces on the atoms at the tip are sufficient to hold the crack open against the cohesive forces acting there between the atoms. The energy condition in the plastic zone itself, discussed in Section 3, is also a thermodynamic argument because we deal merely with the overall energy conservation of the system.

When the basic thermodynamic energy balances are satisfied, there still remains the question of how fast the process occurs. For example, if in the basic molecular processes by which the crack advances, there are activation barriers to be overcome (even though the thermodynamic forces are positive), the rate of crack advance will be governed by rate processes. If no activation barriers exist, then fast crack growth ensues.

There are two main ways in which kinetics will be manifested; (1) rate processes at the crack tip itself may limit crack growth, (2) rate processes in the plastic zone by which the plastic zone accommodates itself to the core crack can limit the crack growth. *A priori*, there is no way of telling which of these is the most important from our thermodynamic arguments. However, intuitively, in a case like the intergranular hydrogen embrittlement failure, where the driving forces at the crack tip dominate the dislocation wake drag force, we would expect crack tip molecular activation processes to dominate dislocation creep forces. One form of such molecular processes at the tip is the lattice trapping effect discussed in an earlier paper [20]. The additional complexities involving, say, hydrogen gas effects at the crack tip will be physically analogous to, and mathematically similar to, the lattice trapping and have been discussed in a qualitative generic way by Lawn [2].

6. Brittle and ductile fracture

In this paper we have reconsidered some old questions regarding the fundamental basis of brittle versus ductile fracture in materials by developing a model of modified brittle fracture which is consistent with the known properties of ductile materials. Since hole growth is a demonstrated form of fracture in these same materials, the fundamental question then pertains to the criteria governing the appearance of one form of fracture instead of the other, or even more importantly the transition from one form of fracture to the other.

We believe this criterion concerns the ability of the material to create an arbitrary dislocation density up to densities of the order of unity in atomic units, and homogeneous down to atomic dimensions. We have shown that if the crystal is limited in its ability to create such densities at stresses characteristic of the tip stress of a crack, then the fracture process is essentially a brittle phenomenon, even though the macroscopic toughness of the crack imbedded within its plastic zone may be quite high.

We have already commented in the Introduction on the conditions for self-consistency of our model, namely that (1) the crack must be stable against spontaneous dislocation formation, and (2) that sources in the crystal must not be distributed homogeneously on an atomic scale. But condition (2) is always satisfied for any crystalline solid with a nominal degree of order. Dislocation sources must by definition be regions of high stress concentration, such as ledge sites on grain boundaries, incoherent precipitate boundaries, etc., which are not distributed so densely and uniformly over a random prospective fracture surface in the material as to be able to furnish the required homogeneous dislocation densities. We therefore believe that condition (1) is the primary criterion.

The argument of the last paragraph refers to the self-consistency of the classical fracture model, but this same argument also applies to the hole growth model. Suppose a hole is nucleated on an atomic scale in front of a main crack, as postulated in the hole growth model. As in the case of classical fracture, the material around the hole must deform with strains large compared to unity in order for the hole to grow. So long as the radius of curvature of the hole is small compared to the distance between dislocation sources at the available stress, the hole must supply its own dislocations by spontaneous formation at its surface in much the same way as we have argued for the crack. Unfortunately, although this qualitative argument is easy to state, the corresponding quantitative statement is very difficult for the hole case, because of the complex stress state, and the uncertainties regarding the ultimate nature of the original hole nuclei. Hence, although we have given in this paper a self-consistent model for the brittle fracture case, until we can do the same for hole growth, such matters

as the ductile to brittle transition cannot be put on a firm theoretical foundation, and this remains a task for the future.

There is within our model one parameter which particularly invites further study, namely the cut-off dislocation density and the inner core radius, R_c, of Equation 2. Unfortunately, we have found no experimental data which have helped throw much light on this parameter, and we suggest experimental studies regarding dislocation densities and structures near crack tips would probably be the most useful studies in sorting out the fundamental questions we have attempted to pose.

7. Appendix

Equation 14 was derived by assuming that below the dislocation cut-off radius R_c, the stress is given by the solution of a crack in an infinite medium. Here, we shall show that if we take the continuum stress solution given by Rice and subtract from it all the dislocations within the cut-off radius, R_c, that the stress within the depleted zone is that of the core crack near the crack tip. From this result, we shall rederive Equation 14.

Quantitatively,

$$\sigma(x) = \sigma_{ep}(x) - \sigma_{dep}(x) \qquad (A1)$$

σ is the total stress along the x-axis (see Fig. 8), σ_{ep} is the Rice elastic–plastic solution, and σ_{dep} is the stress along the x-axis of the subtracted (or depleted) dislocations within R. (We perform our calculations for simplicity along the x-axis in front of the crack, and do not include the angular dependence.)

We calculate the term, σ_{dep} as the sum of two terms. The first, σ^D, is the direct contribution of the dislocations, and the second is an image term, σ^S.

$$\sigma_{dep}(x) = \sigma^D + \sigma^S \qquad (A2)$$

σ^D is the direct term of the dislocations stress corresponding to a dislocation in an infinite medium. At the tip of the crack, this stress is given by

$$\sigma_{23}^D(x') = \frac{\mu b}{2\pi}\int \frac{N'(r,\theta)}{r'}\cos\theta\, dx \qquad (A3)$$

$N(r,\theta)$ is the deleted dislocation density, given by

$$N' = N(r,\theta) - N_m \qquad (A4)$$

$N(r,\theta)$ is the continuum density given by Equation 8 and N_m is the cut-off maximum density allowed by the material.

The term in σ^S is essential because the deleted dislocations are not in an infinite medium, but one with a cut in it. Hence σ^D is modified by image terms due to the cut. This image term is given by

$$\sigma_{23}^S(x) = \frac{-2|x+a|}{\pi(x^2-2ax)^{1/2}} \int_0^{-a} \frac{\sigma_{23}(x')\sqrt{[-(2ax'+x'^2)]}}{x^2+x'^2+2a(x-x')}\,dx', x>0 \qquad (A5)$$

This expression is obtained in the following standard way. Suppose there is a stress $\sigma_{23}^D(x')$ exerted over the negative x'-axis in an infinite medium. Then we make a cut along the negative x'-axis of length $2a$, beginning at the origin. To satisfy the boundary condition of elasticity, these stresses must be cancelled on the surface by adding a stress $-\sigma_{23}^D(x')$ along the cut surface. Then the stress along the positive x-axis, (ahead of the cut) due to to this stress is given by Equation A2. Equation A5 appears in the elastic theory of fracture, and is discussed in a number of places, e.g. [11]. Our notation differs from these authors, however, because we have placed the origin at the crack tip instead of in the centre of the cut surface. The stress $\sigma_{23}^D(x')$ is the stress at x' which is generated by the dislocations deleted from the depleted zone. Following the argument as outlined by Goodier (see [11], Equations 45 to 55), we have

$$\sigma_{23}^D(x) = \frac{M}{\sqrt{(2\pi x)}} - \sigma_{23}^D(0) + \ldots$$
$$M = -\sqrt{\left(\frac{2}{\pi}\right)}\int_0^{-a} \frac{\sigma_{23}^D(x')}{\sqrt{|x'|}}\,dx' \qquad (A6)$$

In writing A6, we assume that the size of the depleted zone is small compared to the crack length, so that the integral for M extends only over a small portion of the total crack.

The calculation of M is straightforward, but must be approximated in order to obtain analytic expressions (see Fig. 8).

$$M = -\sqrt{\left(\frac{2}{\pi}\right)}\frac{\mu b}{2\pi}\int\frac{dx'}{\sqrt{(x')}}\int\frac{x'+x}{(r'^2)}N(r)\,dr$$
$$= \sqrt{\left(\frac{2}{\pi}\right)}\mu b\int_{crack}\frac{dx'}{\sqrt{(x')}}\int_{\substack{depleted\\ zone}}\frac{x'+x}{r'^2}\epsilon_0\left(\frac{K^2\cos\theta}{\pi\mu^2\epsilon_0^2}\right)^{n/1+n}\frac{\cos 2\theta}{r^{(1+2n)/(1+n)}}\,dr \qquad (A7)$$

In this form, M cannot be integrated except by numerical methods, so we break up the integrations in the following manner

$$M \simeq \sqrt{\left(\frac{2}{\pi}\right)}\sigma_0\left(\frac{K^2}{\pi\sigma_0^2}\right)^{n/1+n}\int\frac{dx'}{\sqrt{(x')}}\left[\int_0^{x'}dr\int_{-1}^{+1}d(\cos\theta)\frac{r}{x'}\frac{\cos 2\theta\cos^{n/1+n}\theta}{r^{(1+2n)/(1+n)}}\right.$$
$$\left.+\int_{x'}^{R_c}dr\int_{-1}^{+1}d(\cos\theta)\cos\theta\frac{\cos 2\theta\cos^{n/1+n}\theta}{r}\right] \quad \text{(A8)}$$

The θ part contributes a numerical factor of the order of unity which we shall set equal to unity. Thus,

$$M \simeq \sqrt{\left(\frac{2}{\pi}\right)}\sigma_0\left(\frac{K^2}{\pi\sigma_0^2}\right)^{n/1+n}\left\{\int_0^{R_c}\frac{dx'}{\sqrt{(x')}}\frac{1}{x'}\int_0^{x'}\frac{dr}{r^{n/1+n}}+\int\frac{dr}{r^{(1+2n)/(1+n)}}\right]$$
$$+\int_{R_c}^{\infty}\frac{dx'}{\sqrt{(x'^3)}}\int_0\frac{dr}{r^{n/1+n}}\Bigg\} \quad \text{(A9)}$$

Finally, we find an approximate expression for M,

$$M \simeq 2\sqrt{\pi}\lambda'\sigma_0\left(\frac{K^2}{\pi\sigma_0^2}\right)^{n/1+n}R_c^{(1/2-n/1+n)}, \text{ where } \lambda' = \frac{5+12n+13n^2}{2\pi\sqrt{2}n(1-n)} \quad \text{(A10)}$$

With these results, we are in position to derive Equation 14 in perhaps a more satisfactory way than in the main text. The stress ahead of the crack now becomes

$$\sigma(x) = -\frac{M}{\sqrt{(2\pi x)}} + \sigma_{23}^{D}(0) - \sigma_{23}^{D}(x) + \sigma_{ep}(x) \quad \text{(A11)}$$

At $x = R_c$, the stress is composed of the $1/\sqrt{x}$ term, as well as the other terms, making the matching procedure used in the text uncertain and qualitative. However, at the crack tip, the dominant terms clearly are the three first terms since the singularity at $r \to 0$ in σ_{ep} is weaker than $1/\sqrt{x}$. We thus write to a good approximation,

$$\sigma(0) = -\frac{M}{\sqrt{(2\pi b)}} = \frac{K_c}{\sqrt{(2\pi b)}} \quad \text{(A12)}$$

Since the stress at the crack tip is given by the cohesive forces of the atoms there, by Equation 12 we can also write

$$-M = 2\gamma Y \quad \text{(A13)}$$

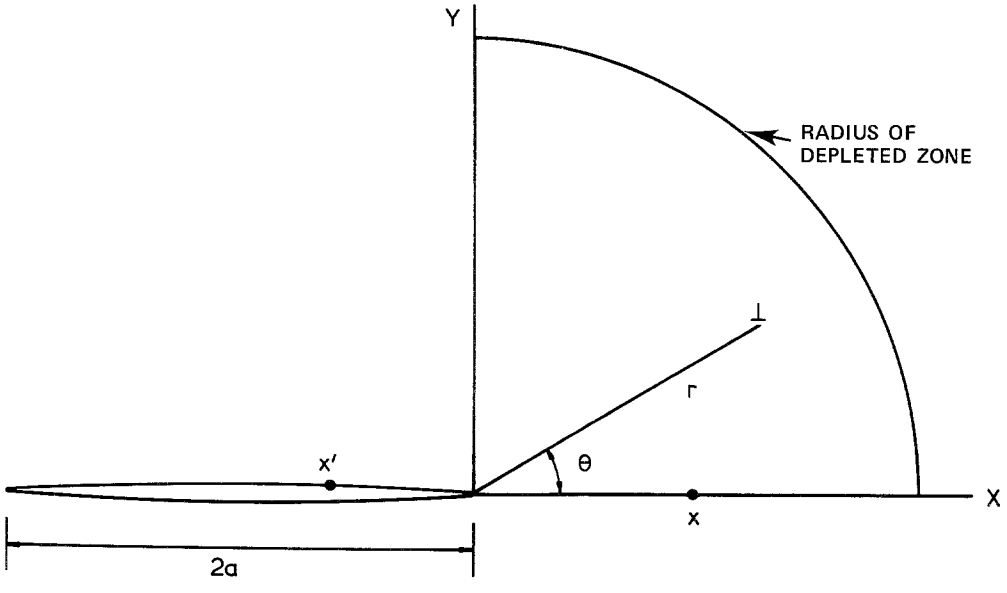

Figure 8 Co-ordinate configuration for stress at crack tip. x' is a point on the crack surface, x is a point ahead of the crack tip at which the overall stress is needed. (r, θ) is the position of a deleted dislocation in the core region.

After substituting Equations 9 and A10 into Equation A13, we again find Equation 14, but with the substitution of the numerical factor λ' for λ. Because of the simple physical ideas behind Equation 14, and because of the inherent approximations in all we have done, we prefer the simpler argument of the main text to the one outlined in this Appendix.

References

1. G. HAHN, M. KANNINEN, and A. ROSENFELD, Ann. Rev. Mater. Sci. **2** (1970) 381.
2. B. LAWN and T. WILLSHIRE, "Fracture of Brittle Solids" (Cambridge U. Press, Cambridge, 1975).
3. J. RICE and R. THOMSON, *Phil. Mag.* **29** (2974) 73.
4. M. KANNINEN and M. GEHLEN, *Int. J. Fract. Mech.* **7** (1971) 471.
5. J. D. EMBURY, "Strengthening by Dislocation Substructures" in "Strengthening Methods in Crystals", edited by A. Kelly and R. Nicholson (Wiley, New York, 1971).
6. E. HART, unpublished internal GE Research Report.
7. R. ORIANI and P. JOSEPHIC, *Acta Met.* **22** (1974) 1065.
8. See papers by W. Gerberich, J. Tien, and others, (Met. Soc. AIME, New York, 1976).
9. J. RICE, *J. Appl. Mech.* **34** (1967) 287.
10. R. DeWIT, "Solid State Physics", Vol. 10, edited by F. Seitz and D. Turnbull, (Academic Press, New York, 1960) p.249.
11. J. GOODIER, in "Fracture", Vol. II, edited by H. Liebowitz (Academic Press, New York, 1968) p.1.
12. S. BURNS, *Acta Met.* **18** (1970) 969.
13. N. PETCH, *Phil. Mag.* **1** (1956) 331.
14. H. ADAMSON, "Physical Chemistry of Surfaces", (Interscience, 1967).
15. W. GERBERICH, J. GARRY, and J. LESSAR, "Effect of Hydrogen on Behavior of Materials" (Met. Soc. AIME, New York, 1976).
16. J. RICE, "Effect of Hydrogen on Behavior of Materials", edited by A. Thompson and I. Bernstein, (Met. Soc. AIME, New York, 1976) p.455.
17. W. GERBERICH and Y. CHEN, *Met. Trans.* **6A** (1975) 271.
18. H. JOHNSON, "Effect of Hydrogen on Behavior of Materials" (Met. Soc. AIME, New York, 1976).
19. H. HEUMANN and PRIMAS, *Z. fur Naturforshung* **21a** (1966) 260.
20. C. HSIEH and R. THOMSON, *J. Appl. Phys.* **44** (1973) 2051.

Received 31 March and accepted 28 April 1977.