

THERMODYNAMICS OF THE QUASI-STATIC GROWTH OF GRIFFITH CRACKS

J. R. RICE

Division of Engineering, Brown University, Providence, RI 02912, U.S.A.

(Received 7 September 1977)

ABSTRACT

RESTRICTIONS on the quasi-static extension, or healing, of Griffith cracks are developed in the framework of irreversible thermodynamics. It is emphasized that thermodynamics requires that $(G - 2\gamma)\dot{l} \geq 0$, where \dot{l} is crack speed, G the Irwin energy release rate, and 2γ the work of reversible separation of the surfaces to be fractured. Implications for 'lattice trapping' models of cracks and for thermally-activated crack motion are discussed, as are the effects on crack growth and healing of a surface-reactive environment, in which case γ must be given a definition appropriate to adsorption-altered surface properties.

1. INTRODUCTION

THE CONDITIONS governing quasi-static extension, or healing, of Griffith cracks are discussed here in the framework of irreversible thermodynamics. By 'Griffith crack' is to be understood a crack which moves in an ideally-elastic lattice without the generation or motion of dislocations, twins, etc. Such conditions are apparently attained in glasses and many ceramics (EVANS, HEUER and PORTER, 1977). They are met much less frequently in metals. But approximate criteria may be given for which an atomistically-sharp crack-tip configuration is stable against blunting by dislocation nucleation from its tip (KELLY, TYSON and COTTRELL, 1967; RICE and THOMSON, 1974), and when these conditions are met up to the point of rupture, it may be assumed that Griffith-like conditions prevail locally near the crack tip. Similar criteria for atomistically brittle response have been given for cracks on interfaces between grains or phases (RICE, 1976; MASON, 1977). When such conditions are met it is possible even in nominally ductile materials that fractures, or segments of macroscopic fractures (e.g. in nucleation of a cavity by cracking of a particle-matrix interface), could take place with Griffith-like conditions in a zone immediately adjacent to the crack tip even though there may be appreciable dislocation motion in surrounding material (THOMSON, 1977).

In the present discussion, Griffith cracks are assumed to be able to undergo quasi-static extension or, in certain conditions, contraction (healing) under sustained, constant loading, at least for a suitably restricted range of loadings. There is ample experimental evidence for this kind of behavior in the presence of *reactive environments* (see the reviews by WACHTMAN (1974) and LAWN and WILSHAW (1975, Ch. 8)), and the thermodynamic formalism to be discussed is wide enough to include such cases. But even for crack growth in high vacuum, some range of quasi-static

crack growth is to be expected because of 'lattice trapping' effects (HSIEH and THOMSON, 1973; LAWN and WILSHAW, 1975, Chs 7, 8). According to this concept, lattice discreteness causes the stress intensity required for rapid (dynamic) growth of a crack to be finitely greater than that required for rapid crack healing. Thus, quasi-static crack motion is possible, e.g. by thermal activation, in the range between these limits, and it is of interest to note that WACHTMAN (1974) cites several observations of quasi-static crack growth in ceramics and glasses under vacuum conditions. Observations of crack healing are also cited.

1.1 Summary of thermodynamic restrictions

Restrictions on crack growth processes will be developed in the next sections based on standard methods of irreversible thermodynamics. The result is a refined interpretation of the GRIFFITH (1920) criterion which, by contrast, focuses on the calculation of an *equilibrium* crack length and cannot deal with kinetic aspects or with the normally, but not always, observed irreversibility of cracks against healing. However, the restrictions on crack growth are first stated directly here, as an introduction, by citing some relevant results from studies by the present writer on the thermodynamics of non-elastic deformation of solids due to structural rearrangements, on the microscale, of constituent elements of material (RICE, 1971, 1975). These rearrangements were characterized by a set of internal variables, each having associated with it a thermodynamic force. While the earlier work focused on rearrangements by dislocation motion, or crystalline slip, the latter work (RICE, 1975) identified forces conjugate to other types of rearrangement, including those by the extension of Griffith cracks.

In particular, the force per unit length of crack front, conjugate to Griffith crack extension over the distance δl normal to the considered portion of crack front, is

$$f = G - 2\gamma. \quad (1)$$

Here, G is Irwin's elastic energy release rate and 2γ is the work of reversibly separating the fracturing surfaces. Such 'forces' are well known from dislocation and crack mechanics, and are defined so that the integral of $f\delta l$ around all segments of crack front is the negative of the free energy difference between equilibrium states before and after the alteration of crack size.

Now, as discussed in the work cited, when actual time-dependent non-elastic processes can be modelled suitably as sequences of constrained equilibrium states (i.e. when instantaneous states during processes are essentially indistinguishable from the equilibrium states which would result if the internal variables were frozen at their instantaneous values), the standard methods of irreversible thermodynamics are applicable. Specifically, where \dot{l} represents the local crack speed $\delta l/\delta t$, the condition for a non-negative entropy production rate was shown to be

$$\int_{\text{c.f.}} f \dot{l} ds \equiv \int_{\text{c.f.}} (G - 2\gamma) \dot{l} ds \geq 0. \quad (2)$$

Here, 'c.f.' denotes 'crack front' and the integral with respect to arc length s is carried over all extending portions of crack front in the considered body. In the circumstances, it seems appropriate to assume that conditions at a given point along

the crack front are governed by the force at that same point, and hence to rewrite the preceding condition as the requirement that

$$(G - 2\gamma)\dot{l} \geq 0. \quad (3)$$

This last inequality, rather than the more usually cited condition that $G = 2\gamma$ for the onset of crack extension, is, I believe, the preferable form of writing the Griffith condition. In the case of stable crack growth by thermal activation over lattice traps in a high vacuum (to avoid environmental effects), the inequality requires that \dot{l} cannot be of a sign different from that of $G - 2\gamma$. This implies that the reversible work of separation *must* lie in the range between the critical G levels for rapid growth (say, G^+) and rapid healing (G^-):

$$G^- \leq 2\gamma \leq G^+. \quad (4)$$

It is of interest that a discrete lattice model of crack growth developed by ESTERLING (1976) leads to a result at variance with (4), in that G^- is found to exceed 2γ , e.g. by as much as 50 per cent. Indeed, Esterling reaches rather strong conclusions on the inadequacy of γ for characterizing the fracture resistance, but a later analysis by FULLER and THOMSON (1977) showed that this violation of (4) is not a true consequence of the Esterling lattice model, and resulted instead from certain approximations that were introduced to deal with non-linearities of the force law. The exact lattice analysis is consistent with (4), as required for consistency with thermodynamic principles. It may further be noted from (3) that kinetic rate laws for quasi-static crack growth, in which \dot{l} is taken as some function of G , must give $\dot{l} = 0$ when $G = 2\gamma$ if they are to be consistent with thermodynamics. In this regard it is interesting to note that FULLER and THOMSON (1977), in their study of non-linear lattice models of fracture, seem to regard it as an open question as to whether what they call the 'thermodynamic' surface energy (γ , or half the work of reversible separation) could differ in general from what they call the 'Griffith' surface energy (half the value of G which makes the thermally-activated crack growth function, $\dot{l} = \dot{l}(G)$, vanish), and, while rightly critical of Esterling's conclusions, they seem nevertheless to regard it as possibly fortuitous that the two surface energies are found to agree exactly for the particular lattice models that they studied. The present results show, however, that by the principles of thermodynamics the two surface energies that they define must be identical in all cases of crack motion under what is understood here as Griffith-like conditions.

It will further be shown that for crack growth in a reactive environment which can adsorb on the newly-created fracture surfaces, inequality (3) remains valid so long as γ is interpreted as a surface energy as altered appropriately to account for adsorption. The implied restriction on kinetic laws has obvious implications for environmentally influenced crack growth. These will be discussed, as will be a new viewpoint that emerges on the normally observed irreversibility of cracks against healing.

2. CRACKED SOLID AS A THERMODYNAMIC SYSTEM

For simplicity, consider a cracked solid which is constrained to deform in plane strain (Fig. 1) in a vacuum or completely inert environment. Here, P represents the

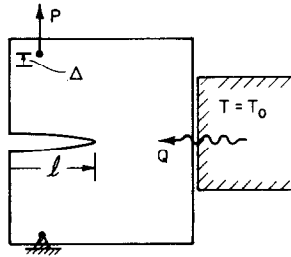


FIG. 1. Cracked solid in contact with heat reservoir.

forces per unit thickness (into the plane of the diagram) exerted on the body, Δ is the work-conjugate displacement (i.e. $\int P d\Delta$ is the work of the applied forces) and l is crack length. The surface of the solid is maintained at temperature T_0 , shown schematically by contact with a heat reservoir in the figure, and Q is the heat adsorbed per unit thickness. It is assumed that all crack motion to be considered is sufficiently slow that the body, viewed macroscopically, acquires negligible kinetic energy and has a temperature field which differs negligibly from T_0 (typical experimental studies of 'slow' crack growth involve crack speeds ranging from 10^{-10} to 10^{-3} m s^{-1} ; see, for example, WACHTMAN (1974)). In this sense, the body, during crack motion, can be regarded as traversing a 'sequence of constrained equilibrium states' corresponding to the sequence of instantaneous crack lengths, and in such a case there would seem to be no disagreements among different schools as to the attachment of meaning to thermodynamic functions or to the proper form of generalization of the laws of equilibrium thermodynamics to irreversible processes. Obviously, thermodynamic restrictions on faster crack motion processes, e.g. with significant temperature non-uniformities and/or with inertial effects, are also of interest. But such cases require for their discussion the determination, and local crack tip interpretation, of non-equilibrium continuum fields. As such, they would seem not to allow the definitiveness of conclusions, and certainly not the simplicity of analysis, of the present, albeit limited, case.

The first law of thermodynamics requires, for the system considered, that

$$\dot{Q} + P\dot{\Delta} = \dot{U}, \quad (5)$$

where $U = U(\Delta, l, T)$ is the internal energy (per unit thickness) of the solid. Note that there is no work term on the left side that is conjugate to l and, in this sense, l can be regarded as an 'internal state variable'. The second law of thermodynamics asserts that an entropy state function $S = S(\Delta, l, T)$ exists for the body and that, in the given circumstances of heat delivery at a uniform surface temperature,

$$\dot{S} \geq \dot{Q}/T, \quad (6)$$

with equality characterizing 'reversible' alterations of state. This last expression may be rewritten in the form

$$\dot{S} = \Lambda + \dot{Q}T, \quad \Lambda \geq 0, \quad (7)$$

where Λ is called the entropy production rate. Now, when \dot{Q} is eliminated between (5)

and (7), we have

$$T\Lambda = P\dot{\Delta} + T\dot{S} - \dot{U} \equiv P\dot{\Delta} - \dot{\Phi} \geq 0, \quad (8)$$

where

$$\Phi = U - TS = \Phi(\Delta, l, T)$$

is the Helmholtz free energy (per unit thickness) and the presumed constancy of temperature is recalled. The last expression may finally be written as

$$T\Lambda = [P - \partial\Phi(\Delta, l, T)/\partial\Delta]\dot{\Delta} - [\partial\Phi(\Delta, l, T)/\partial l]\dot{l} \geq 0. \quad (9)$$

We recall that the difference in the Helmholtz function, between two states at the same temperature, can be calculated as the work of reversibly and isothermally transforming the system from one state to the other. This work will not, of course, generally coincide with the work of actual forces acting during an actual, generally irreversible, transition from the one state to the other. Accordingly, let us set $\Phi = 0$ in the uncracked, unloaded state ($l = P = 0$) at temperature T_0 so that $\Phi(\Delta, l, T_0)$ is the reversible work of isothermally creating the state of Fig. 1. One may create this state by the following two-step sequence of operations: (i) Separate reversibly the two layers of atoms forming the crack walls by pulling against cohesive forces until the layers are effectively out of range of one another. By *definition* of the quantity 2γ as the reversible work of (isothermal) separation per unit area, the contribution to the free energy Φ is $2\gamma l$. (ii) Deform elastically each element of the body lying outside the crack so that such elements have the same strain-state as that actually induced in the body when the crack length is l and the imposed displacement is Δ . Following GRIFFITH (1920), this work can be equated to the isothermal "elastic strain energy" $W(\Delta, l)$ of the body (per unit thickness) as would be computed from the ordinary continuum elastic solution for a cracked solid at temperature T_0 , without regard for the effects of any cohesive molecular forces acting near the crack tip. We note from a well-known property of the elastic strain energy that

$$Pd\Delta = [dW]_{l \text{ fixed}}, \quad \text{or} \quad P = \partial W(\Delta, l)/\partial \Delta, \quad (10)$$

and also that the Irwin energy release rate G is *defined* by

$$G = -\partial W(\Delta, l)/\partial l. \quad (11)$$

Now, from the above considerations,

$$\Phi = 2\gamma l + W(\Delta, l), \quad (12)$$

and the derivatives which enter the expressions (9) for the entropy production rate are

$$\partial\Phi/\partial\Delta = \partial W/\partial\Delta = P, \quad \partial\Phi/\partial l = 2\gamma - \partial W/\partial l = 2\gamma - G.$$

Accordingly, the entropy production rate is

$$\Lambda = -[\partial\Phi(\Delta, l, T)/\partial l]\dot{l}/T = (G - 2\gamma)\dot{l}/T, \quad (13)$$

and the requirement of a non-negative entropy production rate is, since $T > 0$,

$$(G - 2\gamma)\dot{l} \geq 0,$$

as stated earlier in (3).

Only minor changes are necessary to deal with a three-dimensional, rather than plane strain, configuration. Thus, for a crack of area A , not necessarily straight-fronted,

$$\Phi = 2\gamma A + W(\Delta, \text{crack position}), \quad (14)$$

where now Φ and W have units of *energy* rather than energy per unit thickness. Analogously to (11), G is defined at points along the crack front (c.f.) by the requirement that

$$[\delta W]_{\Delta \text{ fixed}} = - \int_{\text{c.f.}} G \delta l ds, \quad (15)$$

to first order in the δ -quantities for *arbitrary* distributions $\delta l = \delta l(s)$ of infinitesimal advance δl of the crack normal to its front. The entropy production rate during a crack growth process may then be shown, by similar steps to those above, to be

$$\Lambda = \frac{1}{T} \int_{\text{c.f.}} (G - 2\gamma) \dot{l} ds, \quad (16)$$

where \dot{l} denotes the local speed of the crack normal to its front. Hence $\Lambda \geq 0$ implies the inequality (2) and, given the relation of G to the local field of stress prevailing near the crack tip, it seems reasonable to interpret this as a requirement that inequality (3) apply pointwise along the crack front.

2.1 Comparison with cohesive zone and lattice models

Figure 2 illustrates an Elliot–Barenblatt ‘cohesive zone’ model of the crack-tip region. The effect of molecular forces is represented by cohesive stresses σ which restrain the crack walls from relative displacement δ ; σ is regarded as some function of δ and, because of the definition of 2γ as the reversible work of separation, 2γ is equal to the area under the σ vs δ relation. The material outside the cohesive zone is modelled as an elastic continuum but now, in contrast to the classical elastic crack solution, it sustains no singularity of crack opening stress. It is of interest to note that the external loading which just equilibrates the crack against growth or healing (the distinction between the two loads vanishing in this case) has been shown to be

$$G = \int_0^{\delta_c} \sigma(\delta) d\delta \equiv 2\gamma, \quad (17)$$

where G is the energy release rate based on the classical elastic crack solution and where it is assumed that the cohesive zone size is negligible by comparison to overall

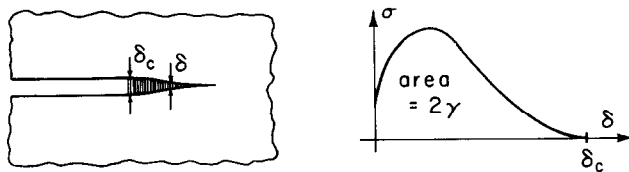


FIG. 2. Cohesive zone crack model.

crack size; δ_c is defined in Fig. 2. The result was proven by WILLIS (1967) based on a linear elastic model of the surroundings and by RICE (1968) for general non-linear elastic models, using path-independent integrals taken around the cohesive zone in a manner later extended by ESHELBY (1970) to the full inclusion of geometric non-linearities.

Of course, the GRIFFITH (1920) equilibrium crack condition is $G = 2\gamma$, so what is being said is that cohesive zone models of the type just discussed are in exact agreement with the Griffith model. However, experimental results can diverge from the predictions of these models (although not from (3)), in that loads in excess of the Griffith level can be sustained even when there is no discernible evidence of plastic flow processes in the crack wake. Of course, cracks are known also to be essentially irreversible against healing in normal circumstances, but this matter is best deferred to the inclusion of surface reactions with the surrounding environment in Section 3.

The divergence between representative experimental results and the Griffith/cohesive zone models may to some extent be attributed to lattice effects. As has already been mentioned, the studies of HSIEH and THOMSON (1973) and ESTERLING (1975) support the concept of 'lattice trapping' of cracks, in that G^+ is finitely greater than G^- . But, analogously to the continuum model of Fig. 2, these lattice models are formulated in a manner that tacitly assumes ideally elastic (in fact, linear) response of all lattice bonds except those bridging the prospective crack plane. This precludes the possibility of, for example, the nucleation of incipient shear dislocations above and below the crack in circumstances for which the dislocations are not driven out beyond a few lattice spacings (at which point they would become unstable and move to much greater distances (RICE and THOMSON, 1974)), but serve nevertheless to greatly diminish the concentration of stress ahead of the crack. Such incipient dislocations would, presumably, be withdrawn from the material by the large surface 'image' forces that dominate after the crack progresses, so that they leave no permanent record. But they could, perhaps, greatly increase G^+ over 2γ , perhaps so much so that thermally-activated crack motion proceeds at an undetectably slow pace unless $G \gg 2\gamma$. This discussion is, of course, *highly speculative* although some support seems to be provided by the limited results reported by GAHLEN, HAHN and KANNINEN (1973) based on a non-linear bcc lattice model with two-body potentials intended to simulate the properties of Fe.

In any event, such lattice considerations serve in principle to illuminate the structure of kinetic laws relating \dot{l} to G (and T) in particular materials, but do not contribute to the basic thermodynamic restriction that $(G - 2\gamma)\dot{l} \geq 0$, which applies independently of the atomistic details of Griffith cracking.

3. CRACK GROWTH IN A SURFACE-REACTIVE ENVIRONMENT

Now, with reference to Fig. 3, consider a cracked specimen in a single phase, surface-reactive fluid environment at pressure p . This situation is represented by the rigid outer wall and piston arrangement containing the cracked specimen and its fluid environment, both in contact with the heat reservoir at temperature T_0 . It is assumed that the newly-created surfaces emerging from the crack tip adsorb an 'excess' Γ , in

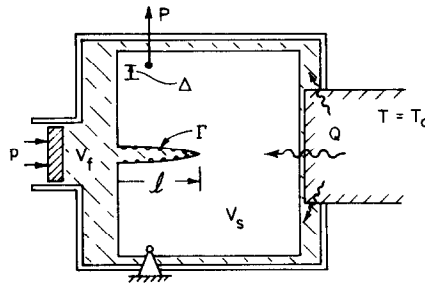


FIG. 3. Cracked solid and surface-adsorbing fluid environment, in contact with heat reservoir. (Γ is excess surface concentration due to adsorption.)

the Gibbs' sense, of mass per unit area from the fluid phase and, for simplicity, it is assumed that this adsorption process has reached completion (i.e. no further increase of the local Γ -value) while points on the newly-created surfaces are still very near to the crack tip. Obviously, the question of whether adsorption takes place *during* the stretching of crack tip bonds, or instead only after these bonds are fully separated, is of great importance to the actual kinetics of the process but not to the form of the thermodynamic restriction that emerges.

If V is the total volume (per unit thickness into the plane of the diagram) enclosed by the rigid outer wall and piston, then the first and second laws of thermodynamics, as in (5) and (7), become

$$\dot{Q} + P\dot{\Delta} - p\dot{V} = \dot{U}, \quad (18)$$

$$\dot{S} = \Lambda + \dot{Q}/T, \quad \Lambda \geq 0, \quad (19)$$

where now U and S are the total energy and entropy of the system, including solid, fluid and surface with adsorbate, Q is the net heat given out by the source and T the essentially uniform temperature of the enclosed system. Hence, by introducing the free energy $\Phi = U - TS$ we have, analogously to (8), with $\dot{T} = 0$,

$$T\Lambda = P\dot{\Delta} - p\dot{V} - \dot{\Phi}. \quad (20)$$

We consider the volume V to be made up of that occupied by the solid, V_s , and that occupied by the fluid, V_f ,

$$V = V_s + V_f, \quad (21)$$

treating the surfaces with adsorbate in the spirit of the Gibbs' concept of a mathematical dividing surface of zero volume. To make the definitions precise, V_s is taken as the volume that would be occupied by a classical elastic continuum loaded with pressure p and force P , and having a crack of length l , and V_f is then *defined* as $V - V_s$. (Note that if the 'actual' fluid mass taken up as adsorbate per unit area is Γ' , and if the adsorbate increases the *nominal* volume of the solid by adding an effective thickness h to its surfaces, then the Gibbs' 'excess' mass per unit area as here defined is $\Gamma = \Gamma' - h\rho_f$, where ρ_f is the mass density of homogeneous fluid at the given pressure and temperature.) The fluid phase is considered to be homogeneous right up to the

mathematical dividing surface, and thus we say that the *mass* of the fluid phase is

$$\rho_f V_f = m_0 - 2\Gamma l. \quad (22)$$

where m_0 is the mass of the fluid phase per unit thickness when there is no crack.

Indeed, since V_f is already defined, this last expression is to be taken as the *definition* of the excess Γ . The free energy Φ is the sum of that of the solid, taken as the ordinary elastic strain energy $W(\Delta, V_s, l)$, plus that of the fluid, regarded as homogeneous and filling the volume V_f with Helmholtz function (or 'strain energy') $w = w_f(p, T)$ per unit mass, plus the *excess* free energy ϕ per unit area due to the presence of the surfaces with adsorbate. Of course, this suffices to define ϕ and, using (22) for the fluid mass, the total Helmholtz free energy is

$$\Phi = W(\Delta, V_s, l) + (m_0 - 2\Gamma l)w_f + 2\phi l, \quad (23)$$

where the dependence of quantities on p and T is not explicitly noted since these remain constant during the crack growth process.

Now, analogously to (10) and (11) one may note that, from the continuum elastic solution,

$$P = \partial W(\Delta, V_s, l)/\partial \Delta, \quad -p = \partial W(\Delta, V_s, l)/\partial V_s, \quad G = -\partial W(\Delta, V_s, l)/\partial l, \quad (24)$$

where the latter defines G . Hence,

$$\dot{\Phi} = P\dot{\Delta} - p\dot{V}_s - [G + 2\Gamma w_f - 2\phi]\dot{l}. \quad (25)$$

Also, by using (21) and (22),

$$p\dot{V} = p\dot{V}_s + p\dot{V}_f = p\dot{V}_s - (2\Gamma p/\rho_f)\dot{l}. \quad (26)$$

With these last two expressions the entropy production rate Λ of (20) is given by

$$\Lambda = \{G - 2[\phi - (w_f + p/\rho_f)\Gamma]\}\dot{l}/T \geq 0. \quad (27)$$

To interpret this expression we note that $(w_f + p/\rho_f)$ is the Gibbs free energy per unit mass of fluid. This quantity is usually referred to as the *chemical potential* μ and it can be regarded as a property of the adsorbate since the adsorbate is presumed to be in equilibrium with the surrounding fluid. Further, one may now *define* a quantity γ so that

$$\gamma = \phi - (w_f + p/\rho_f)\Gamma \equiv \phi - \mu\Gamma, \quad (28)$$

and then the requirement (27) of non-negative entropy production has the same form as earlier, viz.

$$(G - 2\gamma)\dot{l} \geq 0,$$

and this inequality restricts kinetic relations for environmentally influenced Griffith crack growth.

We note that γ as defined in (28) is well known in the thermodynamics of surface adsorption (see, for example, GIBBS (1878)), and is the 'surface energy' (i.e. the work of reversibly creating a unit area of surface) under conditions of chemical composition equilibrium with the adsorbing species in a bulk phase at potential μ . Its role as an effective value for use in Griffith's equation, $G = 2\gamma$, was recognized some time ago by

PETCH (1956) in connection with hydrogen embrittlement. For completeness of the present account, the interpretation of γ and its relation to surface adsorption isotherms (i.e. equations of state $\Gamma = \Gamma(\mu)$, or $\Gamma = \Gamma(p)$, at fixed T) are developed in Section 3.1.

3.1 Surface energy and surface adsorption

We recall that ϕ has been defined as the *surface excess* of Helmholtz free energy (per unit area) that arises when the system of Fig. 3 is regarded as an ordinary elastic solid surrounded by a fluid phase that is homogeneous up to the mathematical dividing surface with the solid. To interpret ϕ consider the system of Fig. 4,

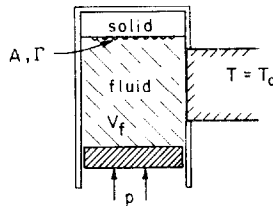


FIG. 4. Adsorption of fluid substance on a solid surface.

maintained at temperature T_0 and consisting of a thin portion of solid having an area A exposed to, and in composition equilibrium with, a fluid under pressure p , now to be regarded as variable.

In the circumstances, one may neglect any variations in volume or energy of the solid (the reader may wish to carry V_s and W through as variables in the development to follow; and the same end result, equation (33), will be found). The entire volume between the exposed solid surface and the piston is V_f , and when we regard this as homogeneous up to the interface, we say that the fluid phase has mass $\rho_f V_f$ and Helmholtz free energy $w_f \rho_f V_f$, where ρ_f and w_f pertain to homogeneous fluid at pressure p and temperature T as before. Then Γ and ϕ are defined as the surface excesses of mass M of the fluid substance and free energy Φ of the system:

$$M = \rho_f V_f + \Gamma A, \quad (29)$$

$$\Phi = \rho_f V_f w_f + \phi A. \quad (30)$$

Now, when we compare two infinitesimally separated equilibrium states of the system, the alteration of free energy equals the work of external forces in the transition between states, and thus

$$\begin{aligned} d\Phi &\equiv d(\rho_f V_f)w_f + (\rho_f V_f)dw_f + A d\phi \\ &= -p dV_f \equiv -(p/\rho_f)d(\rho_f V_f) - p(\rho_f V_f)d(1/\rho_f). \end{aligned} \quad (31)$$

The terms have been rearranged following the identity signs to take advantage of the facts that by a similar consideration applied to a mass of homogeneous fluid,

$$dw_f = -pd(1/\rho_f), \quad (32)$$

and that since M in equation (29) is constant, $d(\rho_f V_f) = -A d\Gamma$. With these, (31) becomes

$$d\phi = (w_f + p/\rho_f) d\Gamma \equiv \mu d\Gamma, \quad (33)$$

which is the Gibbs equation for a surface undergoing isothermal adsorption. Further, a Legendre transformation enables one to re-write the Gibbs equation as

$$d\gamma \equiv d(\phi - \mu\Gamma) = -\Gamma d\mu, \quad (34)$$

which involves the quantity γ of (28) that enters the thermodynamical restriction on crack growth.

In fact, (34) embodies the celebrated ‘Gibbs adsorption theorem’,

$$\gamma = \gamma_0 - \int_{-\infty}^{\mu} \Gamma(\mu') d\mu', \quad (35)$$

where $\Gamma = \Gamma(\mu)$ is the surface adsorption isotherm at the given temperature. The lower limit, $\mu = -\infty$, corresponds to the condition that the piston of Fig. 4 is withdrawn indefinitely, so that the fluid is reduced to a gas of vanishing density and pressure (note from (32) and the definition of μ that $d\mu = dp/\rho_f$, and that $\rho_f \propto p$ for a dilute gas, so that $\mu \rightarrow -\infty$ as $p \rightarrow 0$). This dilute limit corresponds, of course, to vacuum conditions and in that limit, $\gamma \rightarrow \phi \rightarrow \gamma_0$, where now γ_0 denotes the quantity γ introduced in Section 2, and defined so that $2\gamma_0$ is the work of reversible separation in a vacuum (or a completely inert environment). Thus, the Gibbs adsorption theorem shows that if a fluid substance adsorbs ($\Gamma > 0$) as its potential is raised from that of vacuum conditions, then γ is necessarily decreased relative to γ_0 and hence the critical G -value, above which quasi-static crack growth is thermodynamically permissible, is reduced. Indeed, the amount of the reduction in G_{crit} is

$$2(\gamma_0 - \gamma) = 2 \int_{-\infty}^{\mu} \Gamma(\mu') d\mu' = 2 \int_0^p [\Gamma(p')/\rho_f(p')] dp', \quad (36)$$

and is directly calculable from adsorption isotherms. Here, the latter form uses $d\mu = dp/\rho_f$ and envisions that the isotherm has been reported with Γ as a function of pressure in the fluid phase.

To see that γ can indeed be interpreted such that $2\gamma l$ is the reversible work of isothermally-creating crack surfaces of length l on which there is adsorption Γ , it is convenient to consider the following sequence of processes relative to Fig. 3: *First*, with $l = 0$, we withdraw the piston to create a negligibly low pressure. *Second*, the new surfaces are created with reversible work $2\gamma_0 l$ as appropriate to the vacuum-like conditions prevailing. *Third*, the fluid is brought again to its initial pressure under conditions of composition equilibrium with the new surfaces. The net reversible work of the first and third steps is

$$- \int_{p'=p}^{p'=0^+} p' d[m_0/\rho_f(p')] - \int_{p'=0^+}^{p'=p} p' d[m_0/\rho_f(p') - 2l\Gamma(p')/\rho_f(p')],$$

where the terms in brackets are the fluid volumes V_f of (22) as appropriate to each step. These integrals combine after integration by parts to what may be identified

from (36) as $2(\gamma - \gamma_0)l$, plus the work $2lp\Gamma/\rho_f$ that would be done at the piston in separation at constant fluid pressure, and hence the net reversible work of creating the crack with its adsorbate is $2\gamma l$.

3.2 Further comments

For simplicity, the preceding development has been given in terms of a single-phase fluid environment that is capable of adsorbing on the newly-exposed fracture surfaces. More generally, the environments of interest will contain several chemically distinct species (e.g. air consisting of N_2 , O_2 and H_2O vapor; liquid water with dissolved salts; etc.) and each species may have its own surface composition excess, say, Γ_i ($i = 1, 2, \dots$), due to adsorption. Also, a chemical potential μ_i may be associated separately with each species in the surrounding environment and, when the environment is at composition equilibrium with the surface, these potentials may be considered properties of the surface.

The net result is that the entropy production inequality for crack growth is

$$\Lambda = (G - 2\gamma)\dot{l}/T \geq 0, \quad \gamma = \phi - \sum_i \mu_i \Gamma_i, \quad (37)$$

where ϕ is again the surface excess of Helmholtz free energy. Also, (33) and (34) become

$$d\phi = \sum_i \mu_i d\Gamma_i, \quad d\gamma = - \sum_i \Gamma_i d\mu_i. \quad (38)$$

It may further be noted that a thermodynamic analysis of the separation process *per se* has recently been developed by the writer (RICE, 1976), with reference to initially coherent interfaces between solid phases on which there is internal adsorption of elements dissolved in solid solution. The formulation distinguishes between adsorption effects at the limiting conditions of slow separation with composition equilibrium in the adsorbed layer (i.e. separation at constant μ), and of more rapid separation with constant adsorbate composition (i.e. separation at constant Γ). Generalizations of the Gibbs adsorption theorem are derived by which the adsorption induced lowerings of the work of separation in these two conditions can be directly evaluated from adsorption isotherms for the coherent solid interface and for the completely separated solid surfaces.

4. THERMODYNAMICALLY ADMISSIBLE KINETIC RELATIONS

In this section, some general features of kinetic relations for quasi-static crack growth are discussed in relation to the thermodynamic framework. Such a kinetic relation can be represented as a plot of crack speed \dot{l} vs G (sometimes, the stress intensity factor, proportional to $G^{1/2}$, is used), and several such relations are illustrated schematically in Fig. 5. First consider crack growth in a vacuum or a completely inert environment. The relevant value of 2γ in equation (3) is $2\gamma_0$, the work of reversible separation in vacuum, and Fig. 5(a) illustrates the salient features of thermodynamically admissible kinetic relations for crack growth. These must satisfy $(G - 2\gamma_0)\dot{l} \geq 0$ and the curve shown is intended to represent the effects of lattice

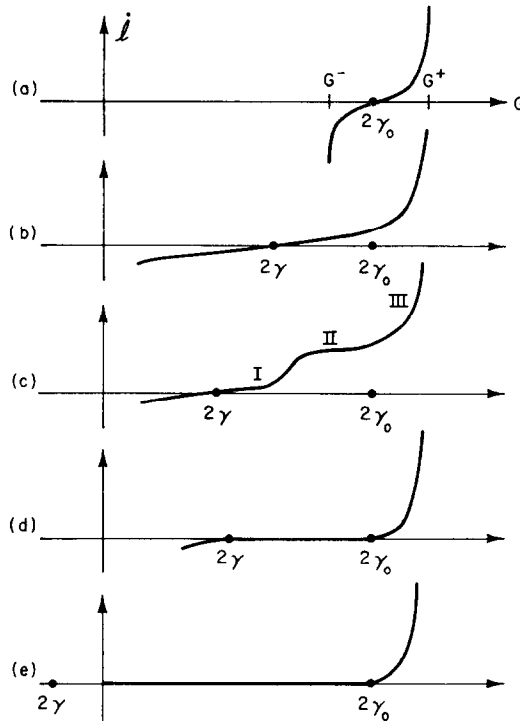


FIG. 5. Thermodynamically admissible kinetic relations: (a) Thermally activated crack growth in vacuum; lattice trapping. (b) Environmentally assisted crack growth. (c) Three typical stages. (d) Environment is surface-active but unable to reach separating crack tip bonds. (e) Strongly surface-active environment makes γ negative but is unable to reach tip; thermodynamics prohibits crack healing when γ is negative.

trapping as discussed in Section 1.1 and at the end of Section 2.1. The crack speed must vanish when $G = 2\gamma_0$, and thermal motions allow the possibility of growth or healing when G differs from $2\gamma_0$. The speed becomes large and positive at G^+ , and large and negative at G^- , where G^+ and G^- are the limits of the trapping range.

WIEDERHORN *et al.* (1974) present data on crack growth in vacuum for six types of glasses and observe that four of these show a range of sub-critical crack growth, with a temperature dependence suggestive of a thermally activated process, whereas two do not. They argue that the presence or absence of sub-critical crack growth under these conditions is explainable in terms of crack tip structure, with a narrow cohesive zone at the tip giving rise to the possibility of thermally activated sub-critical growth and a wide cohesive region favoring abrupt failure.

Figure 5(b) is drawn for the case of crack growth in a surface-reactive environment. Now, thermodynamics requires that $(G - 2\gamma)/l \geq 0$ with γ , defined by (28), being depressed in value relative to γ_0 by an amount dependent on the extent of surface adsorption as in (36). Now, the crack speed must vanish when $G = 2\gamma$, and a range of 'stress corrosion' crack growth, or 'static fatigue', is possible under the environmental influence as illustrated. Since the physical-chemical processes of bond

weakening will, in general, be dependent on the diffusive kinetics of entry of the embrittling species to the crack tip region, the range of very rapid growth will be essentially the same as in vacuum, taking place near the upper limit of the trapping range between $2\gamma_0$ and G^+ .

Indeed, this concept of the influence of a surface reactive environment is broadly consistent with the range of sub-critical crack growth data for ceramics and glasses, mostly under the influence of water vapor, as surveyed by WACHTMAN (1974) and LAWN and WILSHAW (1975). It seems typical from these studies that environmentally influenced kinetic relations can be divided into the three regimes shown in Fig. 5(c). In regime I, the environmental influences are dominant and result in an increasing speed of crack growth with increasing G . But a plateau regime II follows, presumably because of the difficulty of diffusing the environmental species to the crack tip at the speed of crack growth. This is followed by a regime III as shown which is essentially independent of the concentration of the embrittling species and in which the crack growth kinetics become similar to the case without environmental effects. It may be remarked that while in principle the environment-reduced surface energy term 2γ sets the lower limit to the range of sub-critical crack growth, there has not, to the writer's knowledge, been any definitive study reported in which 2γ is calculated independently from the thermodynamics of adsorption (equations (35) and (36)) and thereby shown to coincide with the threshold G -level for crack growth.

LAWN and WILSHAW (1975, Ch. 7) develop a microscopic theory of crack growth in a reactive environment, based on thermal activation and a 'kink' mechanism of crack growth (see also ESTERLING (1976)) which can proceed when a molecule from the environment surmounts local energy barriers and attaches itself to a separating crack tip bond. Their model seems most appropriate for regime I and, except for the fact that they neglect 'backward' jumping in their thermal activation calculations, their model would lead to a result consistent with the thermodynamic requirement $(G - 2\gamma)l \geq 0$ with γ calculated as appropriate for dilute adsorption from an ideal gas. POLLET and BURNS (1977) develop a similar thermal activation formalism for crack growth.

4.1 *Thermodynamics and crack healing*

Crack growth is normally regarded as a non-reversible process, and often this non-reversibility is abetted, in materials which do not exhibit Griffith-like conditions of crack growth, by the mechanically ill-fitting fracture surfaces that result when dislocation steps or large plastic tears are left on the crack surfaces. But when Griffith-like conditions prevail, crack healing is not ruled out by thermodynamics or by experiment.

Indeed, the classic work of OBRIEMOW (1930) and OROWAN (1933) on the *partial reversibility* of crack growth in mica is well known, and the extent of healing was found to be strongly sensitive to the presence of the surrounding environment. Further, WIEDERHORN and TOWNSEND (1970) have studied *crack healing* in soda-lime-silica glass cracked in laboratory air. For example, they report an 80 per cent strength recovery for cracks that are transiently opened under shock loading conditions and a 20 per cent recovery for cracks that are held open to the atmosphere for several

minutes. They attribute the time effect to the increased surface adsorption of some environmental species, presuming it to be either O_2 or H_2O vapor.

It has, of course, long been recognized as a difficulty of the Griffith theory that the presumed full reversibility of a crack is in rather poor accord with experience. For example, GRIFFITH'S (1920) critical equilibrium condition is one of unstable equilibrium at which the crack, according to his model, must be regarded as able to extend indefinitely or to completely close and disappear.

The present irreversible thermodynamics formulation leads, I believe, to a far more satisfactory basis for discussion of the problem. First, it is to be observed that the inequality $(G - 2\gamma)\dot{l} \geq 0$ neither prohibits nor requires negative \dot{l} when $G < 2\gamma$; the theory recognizes properly that it is the detailed kinetics on the crack-tip scale, lying outside the scope of macroscopic thermodynamics, which will determine if healing will actually occur in some given case. The schematic crack growth relations of Figs 5(b, c) have been drawn so as to indicate healing, i.e. a negative \dot{l} for $G < 2\gamma$, since this is a thermodynamically permissible response. But whether healing *actually occurs* will depend to a large extent on the ability of the adsorbed films to diffuse back into the environment (their thermodynamically preferred state) as the crack surfaces pinch upon them. The kinetics of this process may be *very slow* and, if the crack is completely unloaded, it seems probable that large islands of adsorbed film would be trapped by the crack surfaces so that healing would be a far from complete process, as observed.

There would seem to be need for a far wider experimental study of crack healing than yet reported. For example, it would seem profitable to study crack behavior at G -levels just slightly below the growth threshold, presumably identifiable as $G = 2\gamma$, so that there is a minimum of film pinching and entrapment to slow the kinetics of desorption. One possible application is to the crustal rocks of the Earth which are known to exhibit macroscopic inelasticity in shear by the slippage on existing microfissures and the opening of tensile fissures, in the prevailing macroscopically compressive stress fields, due to intense local stress concentrations. Such materials are extremely brittle and are known to exhibit stress-corrosion-cracking behavior which is strongly sensitive to the presence of H_2O (e.g. SWOLFS (1972); MARTIN and DURHAM (1975)), typically available in the form of groundwater. This type of local environment-influenced cracking behavior may be of importance in allowing long-term shear strength reductions of natural rocks prior to faulting, and the comparatively modest stress reductions upon faulting (typically 10–100 bar on average within focal regions) may allow the possibility of significant strength recovery by the time-dependent healing of sharp microcracks.

Another important aspect of the crack irreversibility problem emerges when it is recalled that strongly surface-reactive environments may have no detectable influence in causing time-dependent crack growth. Oxygen in dry air is a case in point. The ready oxidation of freshly exposed crack surfaces suggests a strong surface reactivity with O_2 , but it may be presumed that due to the size of the molecule, or to blockage by oxide films already formed, the molecule cannot actually get to the crack tip cohesive zone and cause its *thermodynamically permissible* reduction of the separation energy from $2\gamma_0$ to 2γ .

In such cases, thermodynamics still requires that $(G - 2\gamma)\dot{l} \geq 0$, but it seems

appropriate on physical grounds to rewrite the entropy production rate, when $\dot{l} > 0$, as

$$\Lambda = (G - 2\gamma)\dot{l}/T = (G - 2\gamma_0)\dot{l}/T + 2(\gamma_0 - \gamma)\dot{l}/T \geq 0. \quad (39)$$

Here, the first term represents the entropy production in the crack tip separation process, and this process can now be assumed to proceed as if no environment were present. The second term represents the entropy production of the non-equilibrium adsorption process which occurs as the freshly exposed surface comes into chemical contact with the environment. Since the two processes are essentially decoupled when $\dot{l} > 0$, it seems appropriate to interpret the requirement of non-negative entropy production, in the given circumstances, as the *pair* of inequalities

$$(G - 2\gamma)\dot{l} \geq 0 \quad \text{for all } \dot{l}, \quad (G - 2\gamma_0)\dot{l} \geq 0 \quad \text{for } \dot{l} > 0. \quad (40)$$

As illustrated in Fig. 5(d), the result is that \dot{l} must vanish for all values of G between 2γ and $2\gamma_0$. When $G > 2\gamma_0$, crack growth can occur in a manner similar to growth in a vacuum or inert environment. When $G < 2\gamma$, crack healing is thermodynamically permissible as illustrated in the figure, although such healing will be subject to the kinetic limits discussed earlier for desorption from the pinched or entrapped film.

Finally, it may be remarked that there appears to be no thermodynamic impediment against the occurrence of negative values of γ . As is clear from (36), whenever the surface adsorption is sufficiently strong at the prevailing levels of adsorbate potential (or pressure in the fluid environment) to cause the integrals occurring in (36) to exceed $2\gamma_0$ in value, then γ is negative. Such behavior may be taken as reflecting the fact that many solids are thermodynamically metastable in their environment and would, judged energetically, tend to react to some new form, except for the overriding kinetic barriers to such a process. We recall that G is necessarily non-negative. This is clear in linear elasticity since G is proportional, by Irwin's relation, to the square of the stress intensity factor. More generally, as RICE and DRUCKER (1967) observed, in any loaded elastic system which satisfies the principle of minimum potential energy, advantage may be taken of the fact that the displacement field prevailing before an increment of crack extension is a kinematically admissible field for the state after crack extension, which has the direct consequence that $G \geq 0$. Thus, when $\gamma < 0$, the thermodynamic inequality

$$(G - 2\gamma)\dot{l} \geq 0 \quad \text{implies } \dot{l} \geq 0. \quad (41)$$

Hence in this case, which may be rather common under ordinary environments of crack growth, *thermodynamics prohibits crack healing*.

Indeed, Fig. 5(e) has been drawn on the presumption that a strongly surface reactive environment is present, so that $\gamma < 0$, but that this environment is kinetically limited from access to the crack tip. In such a case, the environment is without effect on the forward process of crack growth, possible when $G \geq 2\gamma_0$, and no crack healing can occur.

4.2 Summary

Irreversible thermodynamics restrictions on the quasi-static growth of Griffith cracks have been derived with due attention to the effect of the surrounding

environment. The formulation would seem to offer advantages over the classical Griffith equilibrium crack model, not least because of the scope illustrated here for inclusion of lattice trapping and stress corrosive effects within the theoretical formalism, and for the insights afforded on the normal irreversibility of cracks against healing. Since the theory is thermodynamic in nature, nothing is given beyond *global restrictions* on the detailed molecular kinetics of crack growth. Obviously, this is an important topic for further study, as are also cases of crack growth in less brittle solids, for which plastic flow is an essential part of the separation process, and of more rapid crack growth processes in which equilibrium elasticity is an inadequate model for the crack-containing solid.

ACKNOWLEDGEMENT

This study was supported by the NSF Materials Research Laboratory at Brown University, and by the University of Michigan through the ARPA Materials Research Council.

REFERENCES

- | | | |
|------------------------------------------------|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| ESHELBY, J. D. | 1970 | <i>Inelastic Behavior of Solids</i> (edited by M. F. Kanninen, W. F. Adler, A. R. Rosenfield and R. I. Jaffee), (Battelle Institute, Materials Science Colloquium. Columbus and Atwood Lake, Ohio. September 15–19, 1969), p. 77. McGraw-Hill, New York. |
| ESTERLING, D. M. | 1976 | <i>J. appl. Phys.</i> 47 , 486. |
| EVANS, A. G., HEUER, A. H. and PORTER, D. L. | 1977 | <i>Fracture 1977</i> (edited by D. M. R. Taplin), (Proceedings of the Fourth International Conference on Fracture. University of Waterloo. 19–24 June 1977), Vol. 1, p. 529. University of Waterloo Press. |
| FULLER, E. R., JR and THOMPSON, R. | 1977 | <i>Ibid.</i> Vol. 3, p. 387. |
| GEHLEN, P. C., HAHN, G. T. and KANNINEN, M. F. | 1973 | <i>Proceedings of the Third International Conference on Fracture</i> (edited by A. Kochendörfer), (Munich. April, 1973), Vol. 2, Paper I–243. Verein Deutscher Eisenhüttenleute, Düsseldorf. |
| GIBBS, J. W. | 1878 | Trans. Connecticut Acad. III, 343. (Also: 1961. <i>The Scientific Papers of J. Willard Gibbs</i> , Vol. One: <i>Thermodynamics</i> , p. 55. Dover, New York.) |
| GRIFFITH, A. A. | 1920 | <i>Phil. Trans. R. Soc.</i> A221 , 163. |
| HSIEH, C. and THOMSON, R. | 1973 | <i>J. appl. Phys.</i> 44 , 2051. |
| KELLY, A., TYSON, W. R. and COTTRELL, A. H. | 1967 | <i>Phil. Mag.</i> 15 , 567. |
| LAWN, B. R. and WILSHAW, T. R. | 1975 | <i>Fracture of Brittle Solids</i> . Cambridge University Press. |
| MARTIN, R. J. and DURHAM, W. B. | 1975 | <i>J. Geophys. Res.</i> 80 , 4837. |

- MASON, D. D. 1977 *Segregation induced embrittlement of grain interfaces*. Sc.M. Thesis, Brown University, Providence.
- OBREIMOW, J. W. 1930 *Proc. R. Soc. A* **127**, 290.
- OROWAN, E. 1933 *Z. Phys.* **82**, 235.
- PETCH, N. J. 1956 *Phil. Mag.*, Ser. 8, **1**, 331.
- POLLET, J.-C. and BURNS, S. J. 1977 *Int. J. Fract.* **13**, 775.
- RICE, J. R. 1968 *Trans. ASME* **90**, Ser. E, *J. appl. Mech.* **35**, 379.
- 1971 *J. Mech. Phys. Solids* **19**, 433.
- 1975 *Constitutive Equations in Plasticity* (edited by A. S. Argon), Ch. 2. Massachusetts Institute of Technology Press.
- 1976 *Effect of Hydrogen on Behavior of Materials* (edited by A. W. Thompson and I. M. Bernstein), (Proceedings of an International Conference. Moran, WY. September 7-11, 1975), p. 455. The Metallurgical Society of AIME, New York.
- RICE, J. R. and DRUCKER, D. C. 1967 *Int. J. Fract. Mech.* **3**, 19.
- RICE, J. R. and THOMSON, R. 1974 *Phil. Mag.* **15**, 567.
- SWOLFS, H. S. 1972 *Underground Waste Management and Environmental Implications* (edited by T. D. Cook), Amer. Assoc. Petrol. Geologists, Memoir 18, p. 224.
- THOMSON, R. 1977 *The Mechanics of Fracture* (edited by F. Erdogan), AMD-Vol. 19, p. 1. American Society of Mechanical Engineers, New York.
- WACHTMAN, J. B., JR 1974 *J. Am. Ceram. Soc.* **57**, 509.
- WIEDERHORN, S. M. and TOWNSEND, P. R. 1970 *Ibid.* **53**, 486.
- WIEDERHORN, S. M., JOHNSON, H., DINESS, A. M. and HEUER, A. H. 1974 *Ibid.* **57**, 336.
- WILLIS, J. R. 1967 *J. Mech. Phys. Solids* **15**, 151.