F. Guiu and R. N. Stevens**

Thermodynamic Considerations of Fatigue Crack Nucleation and Propagation

REFERENCE Guiu, F. and Stevens, R. N., **Thermodynamic Considerations of Fatigue Crack Nucleation and Propagation**, *The Behaviour of Short Fatigue Cracks*, EGF Pub. 1 (Edited by K. J. Miller and E. R. de los Rios) 1986, Mechanical Engineering Publications, London, pp. 407–421.

ABSTRACT Attention is drawn to two paradoxes in the problem of nucleation and propagation of fatigue cracks. One is the fact that an external stress system alone cannot provide the driving force for the nucleation of a crack on a well polished specimen surface since there is a considerable energy barrier to the growth of a crack of virtually zero length. This results in the inability of fracture mechanics to describe such events. Fatigue cracks can only nucleate at the expense of strain energy stored in the material during stress–strain cycling and they are, therefore, formed at regions where, due to material inhomogeneities, large internal strain fields are set up. The existence of non-propagating fatigue cracks is the consequence of this thermodynamic requirement and an excellent example of this is provided by the nucleation of cracks at persistent slip bands.

Another paradox appears when fracture mechanics concepts are applied to the problem of fatigue crack growth, and it refers to the observation that fatigue cracks grow under values of K_{max} lower than the critical value K_{IC} measured in monotonic stressing. This 'thermodynamic impossibility' is difficult to understand by invoking plasticity effects at the crack tip and by modifications to LEFM because, in the context of this theory, localized plastic deformation would be expected to make propagation more difficult, not easier.

It is suggested that the paradox can be resolved by noting that the instability criterion of 'maximum free energy change' for crack propagation is a *necessary* condition only and not a *sufficient* one, since the plastically relaxed stress levels at the crack tip must reach values capable of producing atomic bond rupture. With this consideration in mind the low values of K_I measured in the growth of fatigue cracks can be explained by the existence of processes which can produce atomic bond rupture under stresses lower than those needed to break bonds in tension.

Introduction

The mechanism of nucleation and growth of small cracks on a smooth surface is a central issue in the problem of fatigue failure. Many cracks may be formed on the surface of a material during its fatigue life, but most of these stop growing and only a few become the propagating cracks which will ultimately cause fatigue fracture (**1**)(**2**). It is, therefore, of great practical importance to be able to predict the growth behaviour of short cracks, but efforts towards this goal have been fraught with difficulties because it is claimed that linear elastic fracture mechanics (LEFM) methods are not applicable to this problem (**3**). The conventional way of representing data on fatigue crack growth is illustrated schematically in Fig. 1 where the crack length increment per cycle, da/dN, is plotted against the alternating stress intensity factor, ΔK $(= K_{max} - K_{min})$ in double logarithmic coordinates. The linear region in this plot corresponds to

* Department of Materials, Queen Mary College, Mile End Road, London E1 4NS.

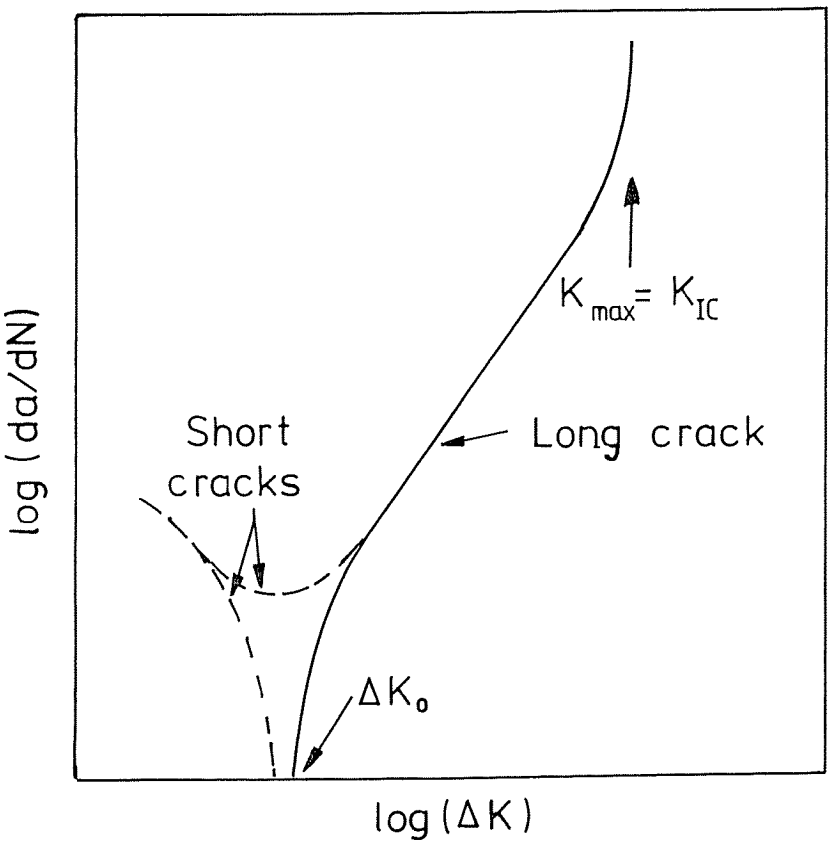

Fig 1 Double logarithmic plot of crack growth rate, da/dN against ΔK showing the normal behaviour of 'long' cracks (full line) and the anomalous behaviour of 'short' cracks (broken lines)

the stable growth rate of fatigue cracks as described by the Paris–Erdogan relation (**4**),

$$\frac{da}{dN} = C\Delta K^n \tag{1}$$

where C and n are material constants. The unstable crack propagation regime is approached as K_{max} tends to K_{IC}, the critical value of the stress intensity factor. Below the threshold value, ΔK_0, fatigue cracks grow at undetectably small rates and appear to remain dormant. The anomalous behaviour of 'short' cracks is represented by the broken lines of Fig. 1 and attempts have been made to explain and quantify this anomalous behaviour in terms of local plasticity effects, microstructural and environmental factors, and by using empirical

elastic–plastic constitutive laws (**3**). It seems that most investigations into the problem of nucleation and growth of physically short cracks fail to recognize the real reason for the limitations of conventional LEFM methods and continuum mechanical approaches.

If crack growth is to occur at all, in either static or dynamic loading conditions, there must be a positive driving force, i.e., growth of the crack must reduce the energy of the system. Since crack growth creates new surfaces, thereby increasing the energy of the system, energy must be extracted from the external loading system or from the general strain energy stored in the body if a net energy decrease is to be achieved. The essential idea of the classical Griffith theory (**5**) is that this results in a critical length of crack below which the driving force for growth is negative. It is, therefore, paradoxical that cracks can grow from virtually zero length in fatigue conditions. The use of both fracture mechanics and continuum mechanics formulated in terms of the stress intensity factor, ΔK, and critical strain or displacement criteria tends to obscure this difficulty, but if the physics of fatigue crack growth are to be understood the problem must be addressed.

Another paradox arises when we consider the stable growth of 'long' fatigue cracks within a range of K_{max} values smaller than K_{IC}. This paradox becomes more evident when we consider the modifications to the Griffith theory proposed by Orowan (**6**) and Irwin (**7**). In this modification crack growth is supposed to increase the energy of the system by an amount greater than the energy of the new surfaces because of plastic deformation. A greater critical length of crack is predicted by this modification than by the original theory. But the slow, cyclic growth of fatigue cracks occurs below this modified limiting length (i.e., at values of $K_{max} < K_{IC}$) and the accumulated evidence is that plastic deformation is essential to this growth process in spite of the fact that plastic deformation is supposed to impose an additional energy penalty on crack growth.

This paper aims to consider these two paradoxes and point to the way in which they can be resolved. In the process it is hoped that some progress will be made towards a better understanding of the physics of fatigue fracture. Clearly, a theory implying that crack growth violates the laws of thermodynamics is unlikely to provide much gain in our understanding of the processes involved, and classical fracture mechanics comes into this category if improperly applied to the initiation and growth of fatigue cracks.

The driving force for crack growth

The classical theory of fracture originated by Griffith (**5**) is a thermodynamic theory. Unlike most of the applications of thermodynamics, the equilibrium is unstable rather than stable, and the appropriate thermodynamic potential has a maximum rather than a minimum, the maximum giving the instability criterion.

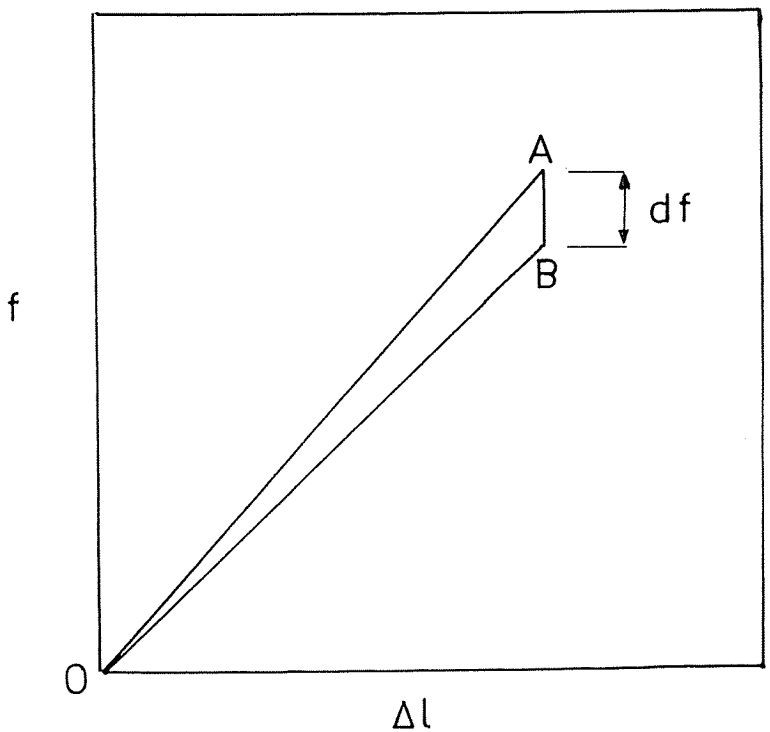

Fig 2 Load elongation curves for a plate with a crack of length a (OA) and for a plate with a crack of length $a + da$ (OB), when no plastic deformation has occurred

Consider a system consisting of a plate of length l and width w under a longitudinal tension and having an edge crack, length a, normal to the tension direction. If crack growth occurs with the system held at fixed length (the fixed grips condition) the Helmholtz energy F of the system is at a maximum value at the critical crack length. If, on the other hand, the tensile force is held constant (the constant load condition) then the thermodynamic potential whose maximum indicates the unstable condition is a modified Gibbs free energy, G' ($= F - fl$, where f is the tensile force). The strain energy of the plate and the surface energy of the crack are both Helmholtz free energies. For simplicity we shall use the fixed grip condition throughout, but, of course, the same results are obtained irrespective of the external constraints.

The load–elongation curve for the plate with a crack of length a is shown as the line OA in Fig. 2. If the crack increases in length by da then the force at A drops by df and the load–elongation curve is now OB. The area OAB is the decrease in strain energy of the system. The rate of change, $\mathcal{G}$, of elastic energy, F_e, with crack length per unit thickness of plate for a specimen in the fixed grip condition is given by (**8**)

$$\mathcal{G} = -\frac{1}{t}\left(\frac{\partial F_e}{\partial a}\right)_{T,l} = \frac{\alpha\pi(1-\nu^2)\sigma^2 a}{E} \tag{2}$$

where t is the thickness of the plate, T is thermodynamic temperature, ν is Poisson's ratio, σ the stress remote from the crack, α a constant ~1.25 for large l and w, and E is Young's modulus. The quantity $\mathcal{G}$ is called the energy release rate, or the crack extension force. The rate of change in the total Helmholtz free energy of the system per unit thickness of plate with crack length is now found by adding the term due to the increase in surface area of the crack. The result is

$$\frac{1}{t}\left(\frac{\partial F}{\partial a}\right)_{T,l} = -\mathcal{G} + 2\gamma = -\frac{\alpha\pi(1-\nu^2)\sigma^2 a}{E} + 2\gamma \tag{3}$$

where γ is the surface energy. We shall call the negative of the left-hand side of equation (3), i.e., $-(1/t)(\partial F_e/\partial a)_{T,l}$, the driving force for crack growth. Clearly a positive force as defined means that crack growth will reduce the Helmholtz free energy of the system. Setting the right-hand side of equation (3) to zero yields the well known result for the critical crack length, a_0

$$a_0 = \frac{2\gamma E}{\alpha\pi(1-\nu^2)\sigma^2} \tag{4}$$

It is worth observing that the strain energy of the system falls because the average stress in the plate diminishes as the crack grows. The energy motivating crack growth therefore comes from the whole system and not from the region around the crack tip as is sometimes supposed.

Real materials do not appear to obey equation (4). Although the stress for rapid crack propagation is inversely proportional to the square root of the crack length (providing the cracks are 'sharp'), the values of the surface energy, γ, calculated from experiment are 1 to 3 orders of magnitude higher than the thermodynamic surface energy. Orowan (**6**) and Irwin (**7**) proposed that crack growth involved not only the creation of new surfaces but also plastic deformation even in apparently brittle materials and that a plastic work term had to be added to the surface energy. This is a view widely held today and has been emphasized more recently by Weertman (**9**). The surface energy, γ, in equation (4) has therefore to be replaced by a quantity, γ', which is much larger and includes the plastic work done per unit area of crack surface.

If this view is accepted then there arises the problem of explaining how a crack which is sub-critical with respect to equation (4) with γ equal to γ' (the surface energy plus the plastic work) can propagate in fatigue, even if this propagation is slow and progressive rather than catastrophic. The problem is compounded by the fact that it is undoubtedly plastic deformation which allows such cracks to grow in the fatigue situation, whereas according to the Orowan/Irwin modification of the Griffith theory it is plastic deformation which renders them sub-critical.

Since the controlling factor in the propagation of a fatigue crack seems to be

the localized plasticity at the crack tip, critical strain or displacement criteria are used rather than stress intensity criteria in the fracture mechanics approach to fatigue crack growth, and this makes the paradox less evident.

The role of plastic deformation

To resolve the problem raised above we reconsider the effect of plastic deformation on the energetics of crack growth. We can, to a first approximation, employ the methods already used for the classical theory. Although plastic deformation is irreversible we may still be able to evaluate the change in Helmholtz free energy. We assume that the plate is not undergoing general plastic yield and that the plastic deformation occurring is *necessarily* linked with crack growth. We shall also assume, for the purpose of envisaging the energy changes taking place, that crack growth and the accompanying plastic deformation take place sequentially. It is noted that this last assumption is not strictly necessary and does not affect the conclusions.

The energy changes are illustrated in Fig. 3. The crack length increases by da

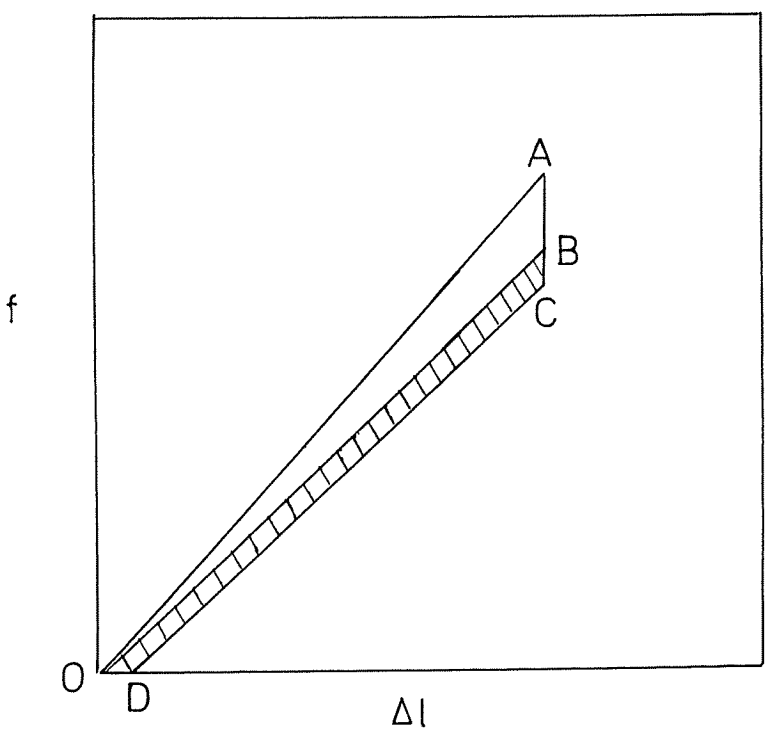

Fig 3 Load elongation curves for a plate with a crack of length a (OA) and for a plate with a crack of length a + da (CO) in which plastic deformation has occurred during crack growth. The shaded area OBCD is the plastic work

under fixed grips. The change in Helmholtz free energy of the system is given by the area OAB as before. Plastic deformation causes the load to fall further from B to C. It is clear that an additional fall in load has to occur since the plastic deformation can only take place by the conversion of elastic strain into plastic strain and the degradation of some of the strain energy into heat. The unloading curve will not now pass through the origin. The additional strain energy lost is equal to the area OBCD.

The implications of this are considerable. It is clear that the sign of the energy change brought about by plastic deformation is the same as the sign of the elastic strain energy change calculated by Griffith. In other words, plastic work, if it has any effect at all on the energetics of crack growth, far from opposing it, will actually provide an increased driving force. We have, however, omitted one small complication. Not all of the energy OBCD is degraded into heat; some of it is stored as the energy of the dislocations created during the plastic deformation which is necessarily associated with crack growth. This gives a term in the energy balance equation having the same sign as the surface energy term. The net effect is to reduce the absolute value of the energy changes associated with plastic deformation somewhat below that represented by the area OBCD.

It will be noted that we have drawn the area OBCD as a rather small fraction of the total OACD. That this is the case can be shown by using an approximate argument based on the Dugdale–Barenblatt (**10**)(**11**) model for the plastic zone length at the crack tip and its effect on the compliance of the system. The plastic zone length, w_p, is given by

$$w_p = \frac{\alpha\pi^2\sigma^2 a}{8\sigma_y^2} \tag{5}$$

where σ_y is the (tensile) yield stress. The effect of the plastic zone is to reduce the compliance of the specimen by an amount equivalent to increasing the crack length by a fraction β ($\sim\frac{1}{2}$) of the plastic zone length, which incidentally illustrates that plastic deformation does decrease the strain energy of the system as argued above. If we put a virtual crack length of $a(1 + \alpha\beta\pi^2\sigma^2/8\sigma_y^2)$ into the Griffith equation for the strain energy of the cracked plate we can find the rate of change of strain energy and separate out the classical term, dF_e (given by multiplying equation (2) throughout by tda), and the term due to plastic deformation. The total change, dF'_e, is given by

$$\mathrm{d}F'_e = -\frac{\alpha\pi(1-\nu^2)\sigma^2 at}{E}\,\mathrm{d}a\left(1 + \frac{\beta\alpha\pi^2\sigma^2}{8\sigma_y^2}\right)^2 \tag{6}$$

Evidently, when multiplied out, the first term on the right-hand side will be the classical strain energy term, dF_p and the remainder will be the changes due to plastic deformation. Hence the strain energy change dF_p due to plastic deformation at the crack tip is given by

$$dF_p = dF_e\left(\frac{\alpha\beta\pi^2\sigma^2}{4\sigma_y^2} + \frac{\alpha^2\beta^2\pi^4\sigma^4}{64\sigma_y^4}\right) \quad (7)$$

For small scale yielding, $\sigma_y > \sigma$, and $dF_p < dF_e$. If $\sigma_y = 3\sigma$ then $dF_p \sim 0.1\ dF_e$. It is, therefore, clear that the plastic work does not oppose crack growth but, since it is small, it will not greatly affect the thermodynamic instability criterion. The term due to the dislocations introduced will be of opposite sign and necessarily be of smaller absolute value than the plastic work term since most of the work done by plastic deformation is degraded into heat. It will, therefore, have even less effect.

If this is correct, it is clear that cracks in real materials are not in unstable equilibrium when they begin to propagate rapidly. They are well past the unstable equilibrium condition and must be in a metastable state when their length lies between a_0 given by equation (3) and the experimentally observed length, a', at which rapid propagation occurs. We conclude that there is no energy balance at the point of rapid propagation; an excess of energy is available for crack growth in this circumstance.

The condition that the driving force (the negative of equation (3), slightly modified to take account of the small effects of plastic deformation) must be equal to, or greater than, zero is certainly a necessary one. Unless the energy of the system can be reduced thereby, there will be no crack growth, slow or rapid. However, it may be that the condition is not sufficient, and as a possible explanation for the non-propagation of thermodynamically unstable cracks it is proposed that the stresses at the crack tip must exceed the theoretical cohesive stress as well as the crack satisfying the energy criterion (**12**). If the theoretical cohesive stress is not exceeded at the crack tip, separation of the surfaces will not occur even if the Helmholtz free energy of the system were to be reduced in the process. The stresses at the crack tip depend on the curvature at this point and it is suggested that the role of plastic deformation is to blunt the crack, thus reducing the local stresses and rendering it metastable. This view explains why crack-tip geometry is so important in practice, whereas it has negligible influence on the magnitude of the crack extension force.

It perhaps should be mentioned that all this has no great consequence for fracture mechanics, which can be legitimately regarded as a phenomenological theory based on the well-founded observation that crack propagation takes place at a critical value of the elastic plus mechanical energy release rate for a given crack geometry. What is affected is our understanding of the physics of fracture and this is important when considering the growth of small cracks.

Slow propagation of metastable cracks in fatigue

If this interpretation is accepted, one of the paradoxes of fatigue crack growth can be easily resolved. There is plenty of energy available to drive crack growth, by *any suitable mechanism*, for crack lengths between the Griffith

critical length a_0 and the length, a' corresponding to the value of K_{IC} (see Fig. 1). Hence, in fatigue conditions, stable crack growth is possible within a range of stress intensity factors smaller than K_{IC}, and the plastic deformation which is intimately associated with the crack tip provides the mechanism necessary for it.

All the existing models of fatigue crack growth, which are well supported by experimental evidence, envisage that the crack grows by a 'shear decohesion' mode rather than by 'tensile decohesion' (**13**)(**16**). This can be achieved by the generation, or annihilation, of dislocations at the crack tip (i.e., plastic deformation), but this shear mode of crack growth will also need to satisfy both an energy criterion and a stress criterion. In the shear mode, however, the stress criterion is easier to satisfy because the theoretical shear stress for creation of dislocations is lower than the theoretical cohesive stress by a factor of 2 to 30, depending on the material (**17**), and crack growth by a shear mode can proceed under a lower value of stress than for tensile cracking (i.e., at a value of K_{max} less than K_{IC}). In addition, the alternating stresses and the plastic deformation associated with them can re-sharpen the crack, keeping the crack tip stresses sufficiently near the theoretical shear stress to allow an increment of growth each cycle. A limited amount of propagation per cycle is a natural consequence of these models since the dislocations produced to allow shear crack growth will themselves reduce the shear stresses at the tip and work hardening will limit the numbers created and the distances they move. Re-sharpening then takes place in the compressive part of the stress cycle.

Nucleation and growth of sub-critical cracks in fatigue

We now turn to the problem of short fatigue cracks, or cracks which are so short that they are sub-critical with respect to the classical Griffith equation. It is a fact that the growth of such cracks from virtually zero length occurs in fatigue (**1**)(**2**) and it is certain that this cannot violate energy conservation. The energy deficit must be made up by processes reducing the energy and which are necessarily coupled to the crack growth process.

In fatigue deformation, and in particular during the stage which precedes the nucleation of a crack, the applied stress does work on the fatigued sample each cycle and a great deal of energy is continually extracted from the mechanical system. Most of this energy is irreversibly degraded into heat, as is evidenced by the stress–strain hysteresis curves observed in fatigue. The main dissipative mechanism is plastic deformation and it is a characteristic of plastic deformation that not all the work done is converted into heat. A small fraction of the plastic work is stored in the material as the energy of the strain fields of the dislocations created. These internal strain fields depend not only on the number of dislocations but also on their spatial arrangements, which can be highly non-uniform and can lead to large energy densities in local regions. In low amplitude fatigue conditions these regions of high energy density tend to

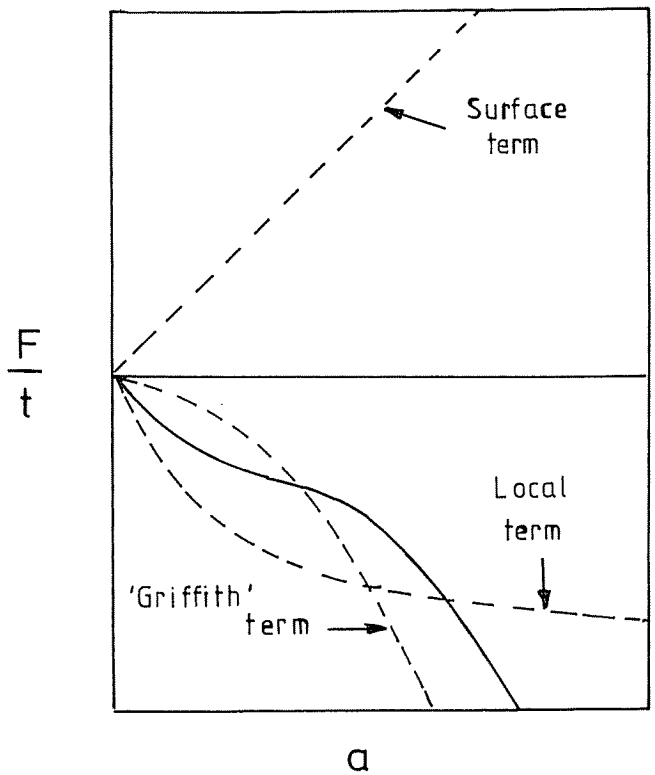

Fig 4 Helmholtz free energy per unit thickness of plate as a function of crack length, a. The various terms contributing to the free energy are shown as broken lines. The total free energy is shown as the full line

develop over many cycles in places where, due to flaws, defects, or inhomogeneities, the plastic strain is concentrated. It is then possible for a crack to nucleate and grow in these regions if the elastic energy stored in them decreases sufficiently when the crack grows.

The energetic conditions for this process are schematically illustrated in Fig. 4, which shows clearly how the strain released from the local regions of internal strain provides the driving force for the nucleation of the crack and part of that required for its initial growth.

The various contributions to the free energy of a stressed plate system when a surface crack nucleates and grows are represented in Fig. 4. The surface energy term is positive and proportional to a; the free energy term arising from the external loading, or from strain energy in regions remote from the crack (the classical Griffith term), is negative and proportional to a^2. The negative of the slope of the curve representing this term is, of course, the crack extension

force, $\mathcal{G}$. The term F_L/t, representing the local strain energy released is also negative but the absolute value of its slope decreases with a as the crack grows and the local strain energy is consumed. It is convenient to call the negative of the slope of this term the *local* crack extension force, $\mathcal{G}_L$. The sum of all these terms gives the total Helmholtz free energy per unit thickness of plate as a function of crack length, a, and is shown as a full line in Fig. 4.

In order for a crack to nucleate from zero length the resultant curve in Fig. 4 must have a negative slope at the origin. Since $\mathcal{G}$ is zero at $a = 0$, the condition that the nucleation of a crack reduces the energy of the system is

$$\mathcal{G}_L(a = 0) > 2\gamma \tag{8}$$

and this is a necessary condition. Provided that it is satisfied, then the three typical situations illustrated in the upper diagrams in Fig. 5 arise when the three free energy terms are added up. If the local crack extension force is very high then the total free energy curve may have no maximum, and there will be no energy barrier for the nucleation and propagation of the crack to any length. This situation is shown in Fig. 5(a).

If $\mathcal{G}_L$ decreases more rapidly than shown in Fig. 5(a) then the total free energy curve has a minimum and a maximum, as in Fig. 5(c). In this case a crack which has nucleated can grow to a stable size, a_0, corresponding to the minimum of the free energy curve, where it will remain a non-propagating crack. There is no driving force for any further crack growth.

The Helmholtz free energy curve in Fig. 5(b) corresponds to the critical situation where $\mathcal{G}_L$ remains just large enough to give a positive driving force for any length of crack.

It is to be noted that since the size of the region of high strain energy density cannot be very large, the local crack extension force, $\mathcal{G}_L$, is expected to drop to zero quite rapidly as crack length increases. Hence, providing that a crack nucleated and propagated by a local region of high strain energy does not get trapped in a thermodynamic equilibrium state, as is the case Fig. 5(c), it will sooner or later begin to show normal 'long' crack behaviour.

In order to show that this energetic argument can qualitatively reproduce well the observed behaviour of short fatigue cracks, the conventional logarithmic plots of growth rate, da/dN, against ΔK are also shown schematically in the lower diagrams in Fig. 5. These have been drawn on the assumption that the crack growth rate is some increasing function of the driving force, i.e., the negative of the slope of the full curves in the upper part of Fig. 5. It can be seen that there are short cracks with anomolously high growth rates relative to the 'nominal' values of ΔK calculated by classical LEFM and that these growth rates tend to *decrease* with increasing ΔK. This is clearly a direct consequence of the fact that the actual driving force is higher than that apparent because of the high value of $\mathcal{G}_L$. The decreasing growth rate with nominal ΔK is the result of the shape of the curve of the local strain energy released by crack growth. Thus cracks apparently nucleate with great ease on the surface of a material

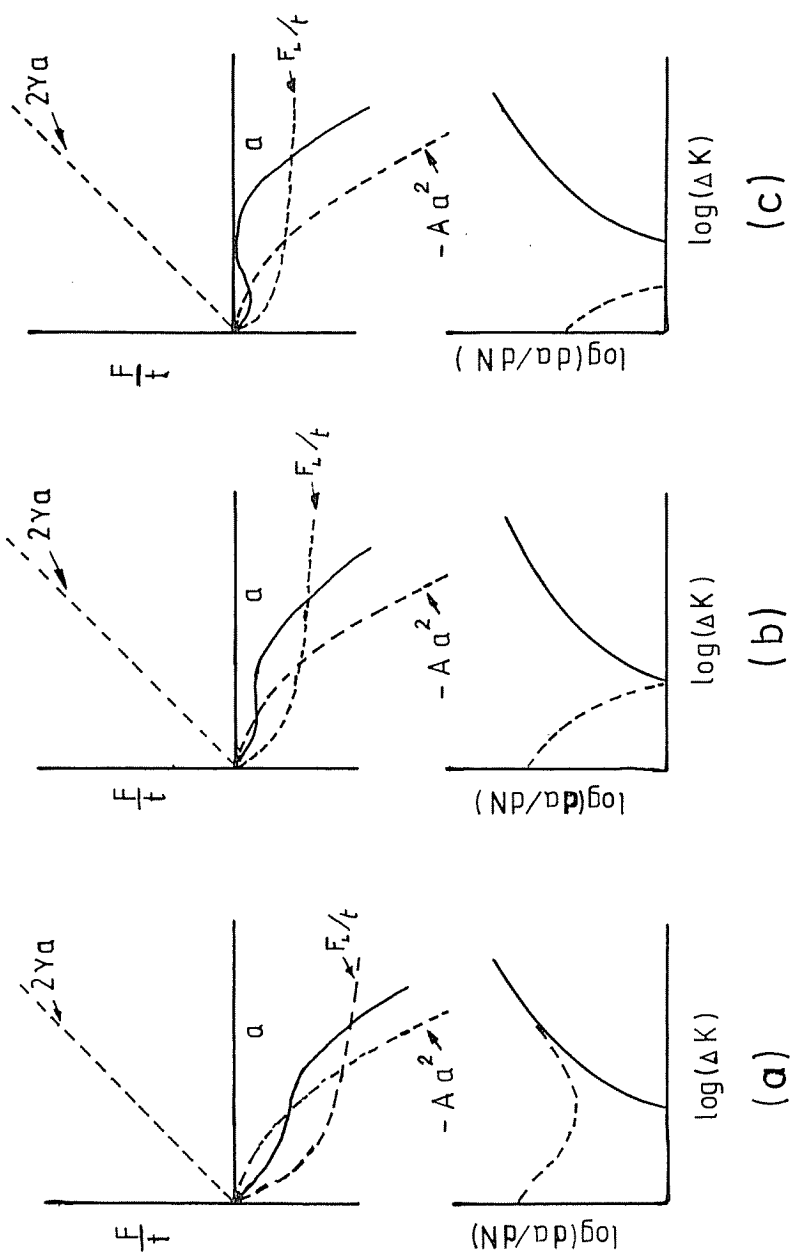

Fig 5 The upper diagrams represent change in Helmholtz free energy with crack length for three different cases: (a) no energy barrier to crack growth, (c) an energy barrier is met after initial crack growth, and (b) the critical transition between (a) and (c). The lower diagrams represent schematically the corresponding conventional crack growth rate curves assuming that the growth rate increases with the driving force

undergoing fatigue. Most grow to a certain length and then stop, having reached a situation in which the driving force is insufficient for further growth (Fig. 5(c)). A few go on to grow to a length at which they become 'long' fatigue cracks (Fig. 5(a)) and one eventually grows to a length at which rapid propagation and failure occurs.

It should be noted that, as pointed out before, a critical stress criterion needs also to be satisfied for a crack to nucleate and grow. It is, therefore, understandable that fatigue cracks are observed to nucleate at regions of high stress concentration which can arise from the geometrical magnification of the remote applied stress as well as from the local internal stress which can reach in some cases very high values. These local sources of internal strain and stress fields, necessary for the nucleation and initial propagation of a fatigue crack can be easily identified in real fatigue situations. Perhaps the best known is the persistent slip band (PSB) which has been subjected to intensive study both theoretical and experimental (**18**)(**19**). Persistent slip bands are the source of large internal stresses, as clearly explained by Brown and Ogin (**19**), who discussed the formation of non-propagating fatigue cracks from PSBs in a way which is a particular example of the general principle outlined above. The PSB stores a high density of strain energy locally and in an appendix by Eshelby to the paper of Brown and Ogin an expression is given for the local crack extension force, $\mathcal{G}_L$, for a crack of length a originating at the intersection of the PSB boundary with the specimen surface and lying along the boundary of the band. This is given by

$$\mathcal{G}_L = \frac{B}{a} \simeq \frac{2\mu h^2 \varepsilon_t^2}{(1-\nu)a} \tag{9}$$

where h is the width of the PSB, μ is the shear modulus, ν Poisson's ratio, and ε_t the shear strain in the band. This implies that the local strain energy released varies with crack length according to

$$(F_L - F_{L,0})/t = -B/\ln a \tag{10}$$

in agreement with the general requirement outlined above and illustrated in Fig. 5. The local crack extension force has a singularity at $a = 0$. This means that the condition in equation (8) would always be satisfied, but the singularity is physically impossible and caution is required in using equation (9) and (10) to discuss the conditions pertaining to the nucleation of a crack.

The expression can be used, however, to illustrate the behaviour of a very short crack with a concrete example. After adding all the free energy terms the necessary condition for the nucleation and propagation of a crack is

$$\frac{1}{t}\left(\frac{\partial F}{\partial a}\right)_{T,l} = 2\gamma - 2Aa - B/a = 0 \tag{11}$$

where $2Aa = \mathcal{G}$ is the classical crack extension force given by equation (2). The quadratic equation (11) has solutions

$$a_0 = \frac{\gamma \pm (\gamma^2 - 2AB)^{1/2}}{2A} \tag{12}$$

corresponding to the unstable and stable equilibrium situations illustrated in Fig. 5. When $\gamma^2 < 2AB$ there are no real roots and there are, therefore, no positions of stable or unstable equilibrium and no energy barrier for the nucleation and growth of a crack. This is the situation illustrated in Fig. 5(a). When $\gamma^2 > 2AB$ then the equation has two roots, one corresponding to the minimum and one to the maximum of the curve in Fig. 5(c), and a short crack can nucleate and grow at first, but then becomes a stable non-propagating crack, remaining stuck at the minimum energy position. The transition between the two cases, represented by Fig. 5(b), occurs when $\gamma^2 = 2AB$. The condition which must be satisfied for a short crack to continue growing and eventually become a long crack is $\gamma^2 \leqslant 2AB$. Using the expressions defined by equations (2) and (9) for the constants A and B this can also be written as $\gamma/\sigma \leqslant 2.6h\varepsilon_t$.

It is noted that if B is zero the solution of equation (11) is the Griffith critical length, $a_0 = \gamma/A$. If the condition $\gamma^2 = 2AB$ is applied then the solution of equation (11) yields $a = \gamma/(2A)$. Thus if the local strain energy release rate is sufficient to allow a crack to grow to a length greater than half the Griffith critical length there will be sufficient energy to allow the crack to become a long crack.

If we use values which are typical for copper then we can put $\sigma \simeq 60$ MPa (the saturation stress for Cu single crystals) and $\gamma \simeq 1$ J/m^2. With these values the Griffith crack length, $a_0 \simeq 20$ μm and the short/long transition length is half this, $\simeq 10$ μm. The value of $h\varepsilon_t$ is then 6 nm and if we take a typical observed value for h of 2 μm the strain in the band must be 3×10^{-3} which is certainly a reasonable value (**19**).

The nucleation and initial growth of a fatigue crack at a PSB has been used here as an illustrative example because it is the only case in which an expression for the rate of release of locally stored energy is available. There are, of course, other sources of internal strain which can develop during fatigue and can provide the driving force for crack nucleation. Examples of these are the incompatible deformation of surface grains which promotes the initiation of cracks at grain boundaries (**20**)(**21**), the formation of soft regions and dislocation free channels in a hard matrix with corresponding extrusions at the surface (**22**)(**23**), and the incompatibilities of deformation developed around second phase particles (**23**).

We believe that detailed knowledge of the structure of these regions of high strain energy density and the manner in which the strain energy is released by crack growth is essential if the initiation and growth of short cracks by fatigue is to be properly understood.

References

(1) MILLER, K. J. (1985) Initiation and growth rates of short fatigue cracks, *Fundamentals of deformation and fracture* (Cambridge University Press, Cambridge), pp. 477–500.

(2) BASINSKI, Z. S. and BASINSKI, S. J. (1985) Surface geometry in fatigued copper crystals, *Fundamentals of deformation and fracture* (Cambridge Univerity Press, Cambridge), pp. 583–594.

(3) SURESH, S. and RITCHIE, R. O. (1984) Propagation of short fatigue cracks, *Int. Metals Revs*, **29**, 445–476.

(4) PARIS, P. C. and ERDOGAN, F. (1963) A critical analysis of crack propagation laws, *Trans AIME, J. Basic. Engng*, **85**, 528–534.

(5) GRIFFITH, A. A. (1920) The phenomena of rupture and flow in solids, *Trans R. Soc. (Lond.)*, **A221**, 163–198.

(6) OROWAN, E. (1952) Fundamentals of brittle behaviour in metals, *Fatigue fracture of metals* (John Wiley, New York), p. 154.

(7) IRWIN, G. R. (1948) Fracture dynamics, *Fracture of metals* (American Society of Metals), p. 152.

(8) KNOTT, J. F. (1973) *Fundamentals of fracture mechanics* (Halstead Press (John Wiley), New York).

(9) WEERTMAN, J. (1978) Fracture mechanics: a unified view for Griffith–Irwin–Orowan cracks, *Acta Met.* **26**, 1731–1738.

(10) DUGDALE, D. S. (1960) Yielding of steel sheets containing slots, *J. Mech. Phy. Solids*, **8**, 100–104.

(11) BARENBLATT, G. I. (1962) The mathematical theory of equilibrium cracks. *Advances in Applied Mechanics* (Academic Press, New York), Vol. 7, p. 55.

(12) PETCH, N. J. (1968) Metallographic aspects of fracture, *Fracture* (Academic Press, New York), Vol. 1, p. 351.

(13) LAIRD, C. and SMITH, G. C. (1962) Crack propagation in high stress fatigue, *Phil. Mag.*, **7**, 847–857.

(14) NEUMANN, P. (1974) New experiments concerning the slip processes at propagating fatigue cracks – I, *Acta Met.*, **22**, 1155–1166.

(15) NEUMANN, P. (1974) The geometry of slip processes at a propagating fatigue crack – II, *Acta Met.*, **22**, 1167–1178.

(16) WEERTMAN, J. (1984) Crack growth for the double-slip-plane and the modified DSP crack model, *Acta Met.*, **32**, 575–584.

(17) KELLY, A. (1966) *Strong solids* (Clarendon Press, Oxford).

(18) MUGHRABI, H., WANG, R., DIFFERT, K., and ESSMAN, U. (1983) Fatigue crack initiation by cyclic slip irreversibilities in high-cycle fatigue, *Fatigue mechanisms: advances in quantitative measurement of physical damage, ASTM, STP 811* (American Society for Testing and Materials), p. 5.

(19) BROWN, L. M. and OGIN, S. L. (1985) Role of internal stresses in the nucleation of fatigue cracks, *Fundamentals of deformation and fracture* (Cambridge University Press, Cambridge), pp. 501–528.

(20) MUGHRABI, H. (1975) Wechselverformung und Erdmüdungsbruch von α-Eisenrielkristallen, *Zeitschrift Metalkunde*, **66**, 719–724.

(21) GUIU, F., DULNIAK, R., and EDWARDS, B. C. (1982) On the nucleation of fatigue cracks in pure polycrystalline α-iron. *Fatigue Engng Mater. Structures*, **5**, 311–321.

(22) MUGHRABI, H., ACKERMANN, F., AND HERZ, K. (1979) Persistent slip bands in fatigued face-centered and body-centered cubic metals, *Fatigue mechanisms, ASTM-STP 675* (American Society for Testing and Materials), p. 69.

(23) MUGHRABI, H. (1983) Cyclic deformation and fatigue of multi-phase materials, Proceedings of the 4th RISØ International Symposium on Metallurgy and Materials Science: Deformation of Multi-phase and Particle Containing Materials, Roskilde, Denmark, p. 65.