Energy Absorbed by the Plastic Wake of an Advancing Crack: A Model for Stable Tearing Versus Fast Fracture

Last Update: 25 March 2008

Warning: The author's views on this topic may be contentious.

1. Does Stable Tearing Really Increase the Toughness? (No)

The methodologies used for the engineering application of fracture mechanics can give a false impression regarding the nature of stable tearing. The assessment procedures permit the use of an effective toughness which increases with increasing tear length, according to some tearing modulus or J_R curve. This encourages the belief that the material toughness really does increase in response to tearing. But why should the material immediately ahead of the crack tip (the process zone) care whether the material immediately behind the crack tip has recently torn, rather than having been cracked from the start? Why should the behaviour of the process zone material (specifically, whether it tears or not), depend upon the history of the material on the other side of the crack tip? Why does its toughness increase?

Whether or not the process zone material tears must be determined by the conditions within the process zone, together with the integrated history of the conditions within the same material element. Anything else is simply unscientific. In practice the "conditions" are highly inhomogeneous and anisotropic on the size scale of interest. In truth, these complications will generally play a crucial part in the fracture process. However, let us maintain the simplifying fiction that the behaviour can be understood in terms of continuum stresses and strains. A material can tear and yet still be capable of sustaining a higher load than when the tearing first started. This experimental fact implies that the "conditions" immediately ahead of the new crack tip after tearing are less severe than was the case in the material which tore. If the J integral is indeed a reasonable measure of the proximity to tearing, then this implies that J evaluated on a path lying entirely within the process zone will have reduced after tearing. The reason is that slow tearing of the crack actually reduces the severity of the strain singularity (see, for example, the Alan Zehnder's "Notes on Fracture Mechanics", January 2008, Cornell University).

Such behaviour can be observed in finite element models. It is well known that very different values for J will result from elastic-plastic analyses which unzip a crack gradually compared with analyses which introduce a crack "suddenly", fully formed. The latter will generally concur with expectations based on analytic methods, such as the reference stress approximation. The former can produce values for J which are much smaller, especially when evaluated on contours very close to the crack tip. The contention is that the phenomenon of reduced J near the tip of an extending crack provides the true mechanism which leads to stable tearing. From this perspective, it is not the toughness of the material which increases, but rather the near-tip J which reduces as a consequence of the ameliorated strain singularity.

Note that the phenomenon of reducing J with tearing can only occur in incremental (irreversible) plasticity. It would not happen for a non-linear elastic material, even if its monotonic stress-strain curve were identical. Thus stable tearing is intimately linked with plasticity, i.e. with non-conservative energy-absorbing behaviour.

If toughness does not really increase during stable tearing, does this invalidate the usual assessment procedures? No, the procedures are probably a reasonable engineering guide. It is just the interpretation of the word "toughness" which is at issue. If this is intended to be a local (microscopic) material property, then it does not increase (or, at least, not as usually claimed). However, this local toughness does not correspond to the quantity which is actually measured for tough, ductile materials. J-tests are usually interpreted using a correlation with the work done, U, measured as the area under the load-displacement curve. Thus J is estimated as (roughly speaking[1]) $J = \eta \frac{U}{A}$, where A is the area of the uncracked ligament and η is a dimensionless constant which is roughly 1 for membrane loading and roughly 2 for bending loading.

Such J-U correlations will give an accurate value for J providing the loading is monotonically increasing, and providing that crack extension has not occurred. However, these correlations cannot represent a local J evaluated within the process zone after tearing. The energy U inevitably increases during tearing. This would be the case even if the load were reducing. So the continued use of such J-U correlations necessarily leads to an upward sloping J_R curve. This is the case irrespective of what the stress and strain fields in the process zone are doing. What is actually being measured by this empirical J is the continued ability of *the specimen as a whole* to absorb more energy. The initiation J can be claimed to be a local material property. But once tearing starts, the value of J_R is properly to be regarded as a global parameter which measures the global structural response. This implies that J_R may depend upon the geometry of the specimen and the nature of the loading, and not just upon the material.

Validity is a key requirement in fracture mechanics. In practical terms, validity is defined as the applicability of laboratory measures of toughness to engineering structures. The approach to justifying the validity of measurements of initiation toughness is to demonstrate that the crack tip fields remain of HRR form at the initiation load. This provides a rational basis for the claim that the same crack tip fields would occur for the same J in a large engineering structure. In practice, the validity requirements generally amount to ensuring that specimens are sufficiently large compared with J_i / σ_y. Similarly size requirements are then carried over into the tearing regime, with limits on the valid extent of tearing being set by maintaining a suitably large value for J_R / σ_y (and possibly other considerations). However, at this point the reasoning becomes questionable. There is now no basis for the implicit claim that the same crack tip fields would occur for the same J_R (or the same tear length) in a large engineering structure. This theoretical underpinning has been challenged by the fact that J_R is empirically a global measure of toughness, and not related directly to the crack tip field after tearing. The validity of torn toughnesses is therefore less secure that that of initiation toughness as regards its theoretical basis. However, the likelihood is that the usual procedures do give a sound indication of the tearing behaviour of large structures. This has been validated experimentally by many large scale test programmes.

2. Why Is Stable Tearing Stable?

Irrespective of the details of what is happening within the process zone, there is a simple requirement which must be fulfilled if fast fracture is to occur: the energy released by an increment of crack advance must equal or exceed the energy absorbed by the structure. Conversely, tearing must be stable if the energy absorbed by the

[1] There are refinements in practice involving the elastic part and/or the uncracked part, which I would not wish to bore you with even if I were competent to do so.

structure exceeds the energy released in advancing the crack. In the case of yielding materials, the minimum energy absorbed by the structure is that required to push the plastic zone to the new crack tip position, leaving a plastic wake behind. Consequently, quantifying the energy requirements of the plastic wake provides a lower bound to the energy absorption of the structure. (Other mechanisms may also absorb energy, such as the process of tearing itself). Hence, simple models based only on yield properties of the material can lead to sufficient conditions for crack stability. A simple model along these lines is presented below.

The 'plastic wake' is illustrated, roughly, pictorially below,

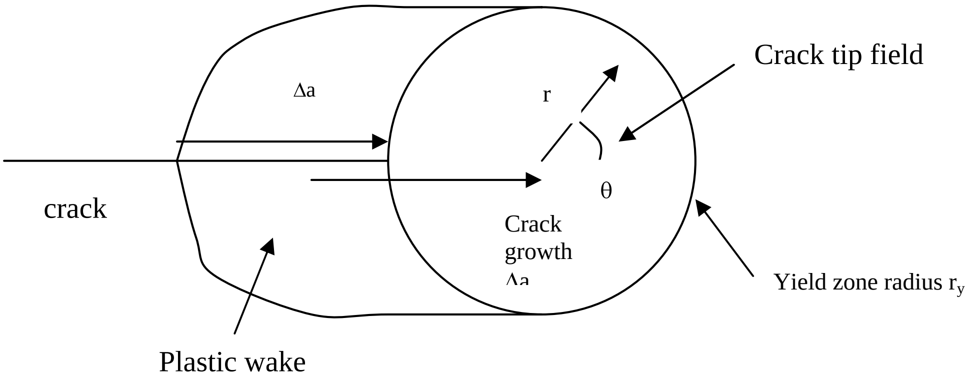


Figure 1

As the crack advances, the irreversibility of plastic deformation leaves large plastic strains parallel to the crack line, behind the 'active' crack tip field of the advancing crack.

Very close to the new crack tip, the stress and strain fields will vary in some complex manner. As discussed above, sufficiently close to the new crack tip, the fields may be less severe than prior to tearing. However, providing that the plastic zone is large enough, but still small compared with the structure, we can expect a roughly conventional "HRR" field to prevail of much of it. At points where the stress has increased monotonically, i.e. roughly speaking material ahead of the crack tip, the crack tip field will be the same as that for a non-linear elastic material with the same monotonic stress-strain curve. For the usual power-law behaviour,

$$\frac{\varepsilon_{ij}^{P}}{\varepsilon_0} = \frac{3}{2}\left(\frac{\bar{\sigma}}{\sigma_0}\right)^{n-1}\frac{\hat{\sigma}_{ij}}{\sigma_0} \qquad (\hat{\sigma} = \text{deviatoric stress}) \qquad \text{(A)}$$

the crack tip fields sufficiently close to the crack tip will be of HRR form,

$$\sigma_{ij} = \sigma_0\left[\frac{J}{\sigma_0\varepsilon_0 I_n r}\right]^{\frac{1}{n+1}}\tilde{\sigma}_{ij}(\theta, n), \qquad \varepsilon_{ij}^{P} = \varepsilon_0\left[\frac{J}{\sigma_0\varepsilon_0 I_n r}\right]^{\frac{n}{n+1}}\tilde{\varepsilon}_{ij}(\theta, n) \qquad \text{(B)}$$

where $\tilde{\sigma}_{ij}, \tilde{\varepsilon}_{ij}$ are the usual, tabulated, dimensionless HRR functions of angle and hardening index, n.

For points which lie, roughly speaking, behind the advancing crack tip, the fields differ from those above because the plastic strain cannot reduce. In terms of the trajectory on a stress / plastic strain graph the situation is illustrated as follows,

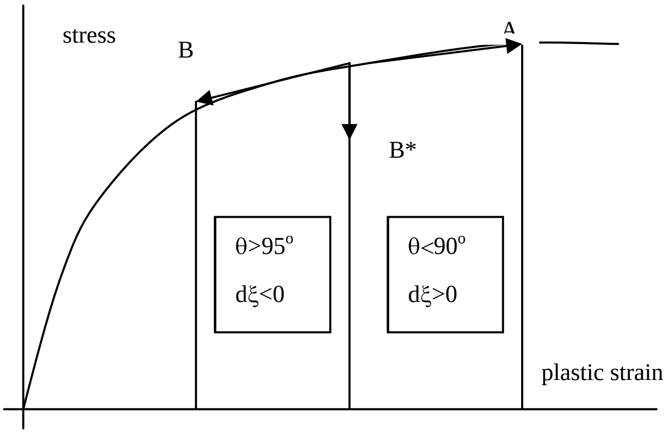
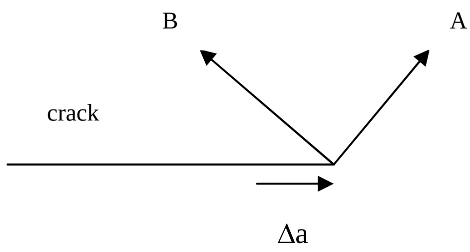


Figure 2



For a point A ahead of the crack, the stress increases as the crack tip advances towards it. The plastic strain, therefore, also increases, as does the "strain energy" density ($\xi = \int \sigma_{ij} d\varepsilon_{ij}$) at A. [NB: For a non-linear elastic material, "strain energy" is the correct term. For a plastic material we note that the irrecoverable "plastic" energy density, $\xi^P = \int \sigma_{ij} d\varepsilon^P_{ij}$, also increases at points ahead of the crack].

Behind the crack tip, however, a non-linear elastic material and a plastic material behave in crucially different ways. A non-linear elastic material will unload back down the original stress-strain curve (see the point B on the graph, above), whereas a plastic material will unload elastically (see the point B^* on the graph, above). At points such as B, the non-linear elastic material will therefore have a decreasing strain energy density, i.e. $d\xi < 0$. In contrast, at points like B a plastic material will have constant "plastic" energy density, i.e. $d\xi^P = 0$. Point B <u>will</u> have a decreasing <u>elastic</u> component of strain energy density, simply because the stress is reducing. However, ignoring this is conservative for our purposes (see discussion below).

It is also rather convenient to ignore the elastic energy because we do not then require to know the stress at point B^*. One may surmise that, since point B is "unloading elastically" the stress decrements may be estimated from the LEFM field. However, this cannot be exact since, at around the $\theta = 90^o$ position, there would then be a discontinuity and/or a failure of equilibrium where the unloading field is required to match the pure HRR field ahead of the crack. Fortunately we do not need to address this issue.

In the above discussion we have, strictly, used the descriptions "ahead" and "behind" the crack tip as short-hand for those regions in which the stress is increasing or decreasing respectively. The identification of these regions with $\theta < 90^o$ and $\theta > 90^o$ respectively will be shown, below, to be a good approximation, at least in plane strain, but not exact.

The wake energy is defined as the difference in the energy required to grow the crack (from length a_1 to length a_2) caused by the irreversibility of plasticity, i.e.,

$$\text{W.E.} = \int dV\{ \int_{\text{elastic-plastic}} \sigma_{ij} d\varepsilon_{ij} - \int_{\text{non-linear elastic}} \sigma_{ij} \, d\varepsilon_{ij} \} \tag{C}$$

In the first integral we must interpret $d\varepsilon$ as being the sum of elastic and plastic strains. This integral must be carried out respecting the actual stress/strain trajectory at the point in question, consistent with the moving crack tip. In particular, for points which reach a maximum stress and then the stress decreases, the elastic strain trajectory will also decrease, but the plastic strain will remain constant (i.e. $d\varepsilon_{ij}^{P} = 0$ at such points after the maximum stress is reached). Consequently, the plastic part of the first integral is truncated at the maximum stress. This is not the case for the second integral, which can be evaluated simply at the final crack length (independent of stress-strain trajectory). This integral therefore continues to have contributions from the non-linear ("pseudo-plastic") strains when the latter are decreasing. This part of the second integral is obviously negative, corresponding to strain energy being released behind the crack tip (point B in the above graph). Hence the wake energy is positive.

Consider an increment of wake energy due to an increment da in the crack length. Each position will be either a loading or an unloading region. Since the two integrals in (C) differ only in respect of the unloading regions, and because the first integral has no plastic contribution in these regions, the change in wake energy can be identified with (the volume integral of) the decrease in the non-linear elastic strain energy density over the unloading regions. This is correct as long as the linear-elastic contributions to the two integrals is similar, so that it cancels between the two. (This would be exact if points B and B^* in the above graph were at the same stress). We argue below that ignoring the elastic energies is conservative.

From the above discussion it is clear that the irrecoverable plastic wake energy, for a crack which has grown from length a_1 to length a_2, may be written simply as,

$$\text{W.E.} = \int_{a_1}^{a_2} \left[\int_{d\xi<0} \left| \frac{d\xi}{da} \right| dV \right] da \tag{D}$$

where ξ is the non-linear-elastic strain energy density. The volume integral (dV) is carried out over the region for which $d\xi < 0$. Equ.(D) is derived by identifying the increase in the wake energy density, due to crack advance da, with the decrease in ξ (hence the modulus sign). It is zero in regions where ξ increases. Note that, even in regions where ξ is decreasing, it is only the change in ξ which contributes to the wake energy. The current value of ξ at the point in question contributes to the "active" crack tip field, rather than the wake field. Hence, Equ.(D) is consistent with the definition of the wake energy given by Equ.(C).

The plastic wake is a region of residual stress. Close to the crack faces, the normal stress, and the tangential shears, will be small. However, in general, some stresses will persist within the wake, together with their associated elastic strain energy. This energy is recoverable in principle, being elastic, and hence has been excluded from our calculation of the irrecoverable (plastic) wake energy. However, this elastic wake energy is not recovered in practice, and hence ignoring it is conservative [i.e. the formation of the wake actually requires more energy than we have included in Equ.(D)].

3. Evaluating the Wake Energy

In what follows we shall assume that the HRR stress and strain fields, Equ.(B), prevail everywhere within the domain of integration of Equ.(D). This is necessarily an approximation, as discussed above. But note that the domain of integration is over the unloading region, not the process zone. Hence, the 'abnormal' fields in the process zone will be unimportant. Hence, for many geometries and loadings, the near-tip fields can be approximated as of HRR form even beyond general yield. At larger distances from the crack tip the approximation will become poor. However, we will see that the integrand falls off $\propto 1/r$ so the wake energy is less sensitive to the fields in these more distant regions. Also, the integrals in Equ.(D) can be confined to those regions with non-zero plastic strains (noting that this includes all regions which have ever yielded, even if the current Mises stress is less than the yield stress – i.e. in the plastic wake!).

For a power-law hardening material [Equ.(A)] the strain energy density is simply,

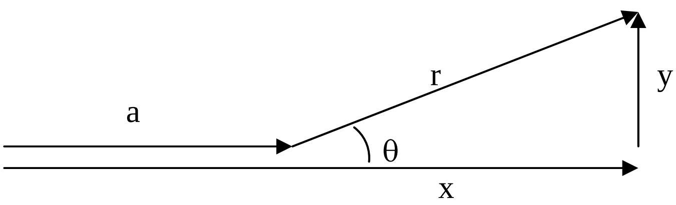
$$\xi = \frac{n}{n+1}\bar{\sigma}\bar{\varepsilon} = \frac{n}{n+1}\sigma_0\varepsilon_0\left(\frac{\bar{\sigma}}{\sigma_0}\right)^{n+1} \qquad \text{(E)}$$

where $\bar{\varepsilon}$ is the Mises equivalent strain. Where the fields are of HRR form, Equ.(B), the energy density becomes,

$$\xi = \frac{n}{n+1}\cdot\frac{J}{I_n r}\cdot\tilde{\bar{\sigma}}^{n+1}(\theta, n) \qquad \text{(F)}$$

where $\tilde{\bar{\sigma}}$ is the dimensionless HRR function for the Mises stress. Note that the dimensionless Mises strain function obeys $\tilde{\bar{\varepsilon}} = \tilde{\bar{\sigma}}^n$ so that the two forms of Equ.(E) are consistent.

To evaluate the integrand of Equ.(D) we use the coordinate system,



Hence, y = r sin(θ), x –a = r cos(θ). The integrand of (D) is thus,

$$\frac{d\xi}{da} = \frac{\partial\xi}{\partial a} + \left[\frac{\partial\xi}{\partial r}\cdot\frac{dr}{da} + \frac{\partial\xi}{\partial\theta}\cdot\frac{d\theta}{da}\right] \qquad \text{(G)}$$

The first term in (G) is due to the crack length dependence of J. The term in square brackets in (G) is due to the material point in question, i.e. that at a fixed value of (x, y), becoming nearer to, or further away from, the crack tip due to the crack advance.

Since we have $\frac{dr}{da} = -\cos(\theta)$ and $\frac{d\theta}{da} = \frac{\sin(\theta)}{r}$ it is readily shown by substitution of (F) that the term in square brackets is simply,

$$\left[\frac{\partial\xi}{\partial r}\cdot\frac{dr}{da} + \frac{\partial\xi}{\partial\theta}\cdot\frac{d\theta}{da}\right] = \frac{n}{n+1}\cdot\frac{J}{I_n r^2}\cdot\frac{\partial}{\partial\theta}\left(\tilde{\tilde{\sigma}}^{n+1}\sin(\theta)\right) \qquad \text{(H)}$$

Hence, using (D), the wake energy increase per unit crack advance, and per unit length of crack front, is given by,

$$\frac{d(W.E.)}{da} = \frac{n}{(n+1)I_n}\int dr \int d\theta \left|\left\{\tilde{\tilde{\sigma}}^{n+1}\frac{\partial J}{\partial a} + \frac{J}{r}\cdot\frac{\partial}{\partial\theta}\left(\tilde{\tilde{\sigma}}^{n+1}\sin(\theta)\right)\right\}\right| \qquad \text{(I)}$$

(Noting that, per unit length of crack front, we have dV = dA = rdrdθ). Equ.(I) is integrated over the region in which the term inside the modulus sign is negative.

Ultimately we shall have recourse to numerical evaluation of Equ.(I). However it is instructive first to carry out an analytical estimation using some reasonable approximations…

(1) The first term in the integrand of (I) may be expected to be small if the crack growth is small compared with ‘a’, i.e. the variation in J is slight. Alternatively, if the compliance is roughly constant, then $\frac{\partial J}{\partial a} \approx \frac{J}{a}$ and this will be small compared with the second term if the region of integration (i.e. the extent of the plastic zone) is small compared with ‘a’. Consequently, our first order approximation is to neglect the first term in (I).

(2) The functions $\tilde{\tilde{\sigma}}$ are tabulated numerically [Ref.1]. These dimensionless functions, raised to power n+1, are plotted against angle in Figure 1 for three illustrative values of n (3, 5, 7) and for plane strain. They are not very sensitive to ‘n’ and are almost symmetrical about θ = 95°. A reasonable approximation is,

$$\tilde{\tilde{\sigma}}^{n+1} = \cos^3(\theta - 95^o) \qquad \text{(J)}$$

as shown in Figure 1. Figure 2 plots the derivative of $\tilde{\tilde{\sigma}}^{n+1}\sin(\theta)$, that is the angular dependence of the integrand of Equ.(I) with the dJ/da term neglected. This defines the domain of the integration, (I), namely where this function is negative. We see that, to a good approximation, this region is θ > 95°. This justifies our earlier usage of the terms “ahead” or “behind” the crack tip as a reasonable description of the regions in which the non-linear elastic strain energy density is increasing or decreasing respectively.

Adopting approximations (1) and (2), above, the wake energy expression becomes,

$$\frac{d(W.E.)}{da} = \frac{2nJ}{(n+1)I_n}\int_{r_{min}}^{r_{max}} dr \int_{95^o}^{180^o} d\theta \left|\left\{\frac{1}{r}\cdot\frac{\partial}{\partial\theta}\left(\tilde{\tilde{\sigma}}^{n+1}\sin(\theta)\right)\right\}\right| \qquad \text{(K)}$$

The leading factor of 2 comes from the equal contribution of the half-plane below the crack, i.e. for $\theta < -95^o$. Both integrals are trivial, giving,

$$\frac{d(W.E.)}{da} = \frac{2nJ}{(n+1)I_n} . \log\left(\frac{r_{max}}{r_{min}}\right) \qquad \text{(L)}$$

where we have used $\tilde{\tilde{\sigma}}(95^o) = 1$ and $\sin(95^o) \approx 1$. Note that it has not been necessary to use any particular approximation for $\tilde{\tilde{\sigma}}$, such as Equ.(J), since the integral depends only upon the value of the function at its maximum, namely unity.

The dependence of the wake energy on the radial limits of the integration requires some discussion. Note that it becomes divergent if the lower limit is taken as zero, or if the upper limit is taken as infinity. This need not be regarded as alarming. The same thing happens in respect of the strain energy associated with a dislocation, which has the same logarithmic dependence. Nevertheless, it is curious because the total energy associated with the crack tip field in a finite body is finite [an energy density proportional to 1/r gives a total energy proportional to the size of the body]. So how can its derivative with respect to 'a' be divergent? The resolution of the paradox is to consider the change in the total energy. To do this the integral in Equ.(K) is extended over all angles (i.e. from 0 to 180^o), and the modulus sign is removed. This angular integral is clearly zero, and this saves us from any concern about the radial integral being divergent. The physical reason for the integral being zero in this case is that it merely represents an unchanged crack tip field translating along with the crack tip. [Any change in the total energy will be due to the first term in (I), proportional to dJ/da, which we have neglected].

4. First Estimates of Wake Energy Using Equ.(L)

To evaluate (L) we need to choose appropriate size scales to identify with r_{min} and r_{max}. These will necessarily be uncertain. Fortunately the logarithmic dependence means that the resulting wake energy is not very sensitive to their exact magnitude.

For r_{min} the obvious identification is with the grain size. A smaller value for r_{min} would be non-conservative (giving a larger estimate for the wake energy). The COD might be a larger size scale than the grain size, but it is clear that continuum plastic deformation is significant at smaller size scales – but virtue of the mechanisms giving rise to the crack tip blunting. Hence using the COD would seem overly conservative. Hence we shall assume r_{min} = grain size = 50 microns (with a sensitivity study for 200 microns).

We shall identify r_{max} with the 'active' plastic zone size, ($r_y = J/\sigma_0\varepsilon_0 I_n$). Recall that the derivation of Equ.(L) has assumed that the stress and strain fields are of HRR form throughout the domain of integration.This means that extending the integration to r_y necessarily involves a crude approximation. However, unlike the total strain energy, the integrand in the wake energy integral, Equ.(K), falls off as 1/r, so that we may expect the approximation to be reasonable – being less sensitive to the details of the fields further from the crack tip.

We shall illustrate the magnitude of the wake energy for plastic hardening indices (n) of 3, 5 and 7 and for plane strain, for which Ref.1 gives I_n as 5.51, 5.02 and 4.77 respectively. For illustration we shall consider loadings or crack sizes giving rise to

SIFs of 10, 20, 40, 80 or 120 MPa√m, with J given by $(1-\nu^2)K^2/E$ and E = 160 GPa. We shall take yield properties as σ_0 = 200MPa, ε_0 = 0.125% (consistent with the value of E). For n = 5 these data give yield zone radii (r_y) of 0.45mm, 1.81mm, 7.3mm, 29mm and 65mm for the increasing SIF values respectively.

The Table below gives the resulting wake energies per unit crack advance, per unit crack front, normalised by J, i.e., $\frac{1}{J}.\frac{d(W.E.)}{da}$,

n	K = 10	K = 20	K = 40	K = 80	K = 120
3	0.57 (0.20)*	0.95 (0.57)	**1.33** (0.95)	**1.71 (1.33)**	**1.93 (1.55)**
5	0.73 (0.27)	**1.19** (0.73)	**1.65 (1.19)**	**2.11 (1.65)**	**2.38 (1.92)**
7	0.83 (0.32)	**1.33** (0.83)	**1.84 (1.33)**	**2.35 (1.84)**	**2.65 (2.14)**

*for r_{min} = 50μm (200μm)

5. What Does the Normalised Wake Energy Rate $\frac{1}{J}.\frac{d(W.E.)}{da}$ Mean?

When the crack advances by Δa the energy that is made available to drive the crack growth is JΔa. [For strain controlled loading, this energy originates from the strain energy of the body, which therefore decreases. In load control, the energy originates from the external agency that is maintaining the load constant].

Hence, when the wake energy increment exceeds JΔa there is insufficient energy available to allow the crack growth to continue (because plasticity is demanding more energy than is being released by the crack extension). Thus, all values greater than unity in the above Table (shown in bold) must correspond to cracks which are stable under small extensions, i.e. exhibit stable tearing.

The Table implies that unstable behaviour (fast fracture) is only possible, for materials with the yield stress and Young's modulus assumed, for SIFs less than about 40 MPa√m. This appears to align remarkably well with typical lower shelf fracture toughnesses for steels in the cleavage regime (recalling that we have used no fracture test data of any sort, nor even postulated any particular fracture mechanism).

In conclusion, stable tearing is expected when $\frac{1}{J}.\frac{d(W.E.)}{da} > 1$. This is independent of any debate about the validity, or otherwise, of the torn toughness, J_R.