The Integrity and Durability of Structures and Machines

John Knott
The University of Birmingham

Abstract

The paper's title relates to the difference in emphasis placed on ensuring the *integrity* of a structure, contrasted with that relating to the *durability* of a machine. The former is usually treated as a "one-off" assessment: the latter is more involved with calculating the lifetime and specifying appropriate inspection periods. Integrity assessments do, of course, take account of changes with time and this is illustrated in the paper by considering the temporal variation of the assessment point on a Failure Assessment Diagram (FAD) and the factors affecting the movement in its position: both crack growth and changes in material properties. Probabilistic effects are also treated. Durability issues are addressed by detailed consideration of the "lifing" of a turbine disc in a gas turbine, demonstrating the importance of initial defect size and its control. These principles are generally applicable to a wide range of other machine components.

Introduction

For a variety of engineering enterprises, a full assessment of 'structural integrity' should incorporate not only the "one-off" appraisal of the margin between safe operation and failure at a given moment in time, but also a prediction of future behaviour. There are three areas in which such predictions are important: i). in specifying appropriate inspection periods for existing plant, ii). in making decisions on possible plant life extension (PLEX), iii). in setting the design life for new plant. Since *failure is defined as loss-of-function*, the *function* of the enterprise should always be considered, to ensure that the relevant failure modes are being considered: not only failure by fast fracture or by tensile plastic collapse, but other modes, such as the buckling of slender cross-sections subjected to compression. Note, for example, that the first stage rocket motor casing for the *Saturn* rocket could not, un-pressurised, support the weight of the rocket without buckling, and a gantry was needed to support the rocket before take-off.

The title of the paper emphasises *Structures* and *Machines*. There are two main types of engineering structure. One, such as a bridge crossing a river, is *functionally static*. It deflects in response to traffic loads, winds or waves, it expands or contracts as the temperature fluctuates, but its prime function is to span the river and bear traffic. The second, such as a ship's hull or an aircraft body, is transported from place to place, but has to maintain *rigidity* with respect to the payload and external forces: in this sense, it is still *functionally 'static'*. *Machines* are *functionally dynamic*: assemblages of gears, crankshafts, pistons and other components, which move relative to each other, often at high speed; producing power-drives for transportation, turbine systems for propulsion and power generation, machine-tools for manufacturing industry. There is not always a sharp distinction between structures and machines. As part of its (structural) function, London's *Tower Bridge* has a roadway that is raised and lowered

to allow shipping to ply up and down the *Thames* river. An aeroplane wing incorporates “flaps” which make essential movements during take-off and landing.

Three other specialised functions may be noted. In *Fairground Rides*, the function is to provide (safely) the thrills of perceived danger, subjecting passengers to much higher accelerations and g-forces than those experienced in normal life. In *Musical Instruments*, the function is to amplify an initial vibration, quite contrary to normal engineering practice. Note that some of the highest strength steel wires in use (UTS > 2500MPa) are to be found in instruments such as pianos and electric guitars. The third specialised function could be termed an *Architectural Statement*. Here, the function is to create something quite different from what has been done before: whether a Building (such as the *Guggenheim Museum* in Bilbao) a Sculpture (the *Statue of Liberty*) or one of the Millennial ‘Mirabilia’ (the *London Eye*). The likelihood of failure of each of these specialised functions can be assessed using the methodologies used for *Structures* and *Machines*, although the input stress level parameters will not necessarily conform to those embodied in conventional engineering codes.

Below the creep range, there is a difference in emphasis when considering the behaviour of *Structures*, subjected functionally to quasi-static loading, and that of *Machines*, subjected functionally to significant dynamic load spectra. Assessments for *structures* tend to emphasise fast fracture or plastic collapse, e.g. in the failure assessment diagram (FAD) of the R6 procedure, Milne et al. [1]. The critical material properties involved are the fracture toughness or the flow strength. Sub-critical crack growth is taken account of in calculations of defect size throughout the design life, but tends to be regarded as a subsidiary input to the FAD. In contrast, the emphasis placed on the through-life performance of *machines* relates primarily to sub-critical crack growth: predominantly fatigue, perhaps enhanced by environmental interactions. Here, the material property of interest is the *crack growth-rate* and, in highly stressed components, the life is dominated by control of the initial defect size: even quite large variations in fracture toughness have rather little effect on the number of cycles, or time, to failure. In terms of the plastic collapse failure condition, the *flow strength* also has rather little effect on life: in contrast, it has a major effect on crack growth, because the *applied stress* scales with flow strength, and the crack growth-rate depends on stress to a power. Recognising the occurrence of intermediate situations, the present paper distinguishes between the *Integrity* of *structures* (based on a final failure criterion) and the *Durability* of *machines* (based on calculations of the number of cycles or time to failure).

It would be nice to illustrate the contrast between these two extremes by the failures of two famous bells: the *Liberty Bell*, in the USA, and the *Tsar Kolokol*, in Russia. The former failed on 8 July 1835 as a result of repeated ringing, since 1753, causing the growth by fatigue of an initial casting defect. This is clearly a *durability* issue. A reputed cause of failure of the 216 tonnes *Tsar Kolokol* (cast in 1733-5) was the result of a single application of an unexpectedly large overload when it hit the ground, having fallen from the top of the bell-tower when the support system collapsed under its weight. Prediction of this event would have required an *integrity* assessment. Unfortunately for the illustration, there is evidence that the bell cracked during the casting operation. Whatever the cause, in both cases, loss of function was the lack of ability to continue to be rung with appropriate resonance. These are unusual events: large bells in English churches and cathedrals have been rung for well over 600 years

without failure and the 46.5 tonnes Great Bell in the Great Bell Temple in Beijing dates from 1403 and is still in functionally 'sound' condition. Recent assessments of the failure mechanisms of the steel wires in instruments such as electric guitars indicate that these are due, in play, to fatigue cracks growing through the wire to a length such that fast fracture ensues, Olver et al. [2]. In contrast, the over-tautening of the metal 'E' string on a violin or mandolin during tuning can lead to a simple tensile instability (in a material with essentially zero work-hardening capacity). For a 0.2 mm diameter wire, even 100% ductility gives a plastic displacement of only 0.2 mm (necking at 45^o): over a 500 mm length of wire, this is 0.04% plastic strain. For a wire with UTS >2500 MPa, (and similar yield strength) with a modulus of 210 GPa, the elastic strain at failure is of order 1.2%. There is correspondingly a large amount of elastic stored energy, which is released as kinetic energy imparted to the two halves of the broken string as they fly apart, giving the semblance of a very 'brittle' nature to a failure that is microscopically 100% ductile (in practice some 60% reduction–in-area). Such issues may appear to be only of academic interest in defining the terms 'brittle' and 'ductile', but similar events at large scale mean that great care has to be taken when testing long cables to failure in the test-house, or in assessing the potential consequences of 'whiplash' failures of cables and wire ropes in structural engineering applications.

A further distinction between *Structures* and *Machines* relates to the *sizes of defects* involved in the assessments, and to the methods by which these sizes can be measured and controlled. In general, the stresses applied to structures such as bridges, pressure vessels, pipelines and engineering plant are rather modest, seldom more than 400 MPa, and the number of full stress-range, 'on-off' fatigue cycles, is usually not high. (Over the last thirty years in Europe, the increased axle weights of lorries have increased the fatigue loadings on roadway bridges). Under normal service operation, critical defect sizes are likely to be tens of mm in size. There are several techniques for detecting surface flaws (dye penetrant, magnetic particle, eddy current) and sub-surface crack-like flaws are explored, using, preferably, ultrasonics (although modern image processing techniques are helping to improve the quality of information from X- or γ-radiography). A 2MHz frequency ultrasonic probe gives a wavelength of ~3mm in steel, and this may be regarded as the limit of spatial resolution in general structural engineering practice. In contrast, a nickel-base alloy in a gas turbine (a machine) may be subjected to a stress range of 1000 MPa on every "take-off/landing" cycle. To ensure 10,000 such cycles without failure, the initial defect size should be no greater than 0.15 mm (see later). This is well below the resolution limit for ultrasonics, and the only method of ensuring durability is to exert "process control": careful control of material processing and component manufacture, backed up by exhaustive testing of test-pieces and component features, followed-up by full-size rig tests. For both structures and machines, an extra degree of control/assurance is achieved by a (20% overload) 'proof test' or an over-speed. This not only assures that failure will not occur, for as-manufactured material, at a load somewhat higher than the design load, but also gives some added protection by blunting defects and creating a level of compressive stress ahead of the defect tips.

The Failure Assessment Diagram (FAD) and Temporal Variations

In considering the different aspects of *integrity* and *durability*, the first step is to examine two different diagrams. One is the Failure Assessment Diagram (FAD) – see

fig.1, ref [1] and Knott and Withey [3]. The other is a schematic graph of crack length *vs.* time or number of cycles – fig. 2. The FAD has as ordinate, the K_r ratio: the ratio of the applied stress intensity factor, K_{app}, to the material's resistance to fast fracture: the fracture toughness K_{1c}, or a value derived from a critical value of the J integral, K_{Jc}, or K_{mat}. The abscissa is the L_r ratio, which is the ratio of applied load to plastic collapse load, defined as $L_r = 1$ at general yield, based on the yield strength, but allowed to extend to $L_r > 1$ to allow for work-hardening. (Older versions of the FAD had as the abscissa a term S_r, representing the collapse load based on a flow strength taken as the linear average of the yield strength and the UTS). Different cut-offs are employed to take account of the different work-hardening characteristics of different steels. An assessment point is located on the diagram, using the appropriate values of K_r and L_r, e.g. point P in fig. 1. If P lies within the failure locus, the structure is deemed to be *safe*: integrity has not been breached. A point lying outside the failure locus would normally only be considered in an assessment following a structural failure. It is of some importance, in any periodic review of integrity/safety, to form a view on the safety margin or 'reserve factor' associated with the assessment. This is often represented as the ratio OQ/OP, where the point Q is the intercept with the *failure locus* of the extrapolation of the line OP from the origin, O, of the FAD, through the assessment point, P, see fig.1. As described below, this is at best, a very rough approximation and more detailed consideration needs to be given to the issue.

With time in service, the location of the assessment point on the FAD can change as a result of a number of factors. One factor is sub-critical crack growth by fatigue or environmentally assisted cracking (EAC), as represented by fig.2. Other factors involve changes in material parameters: softening or hardening, which affects the flow strength and hence L_r; embrittlement due to segregation of trace impurity elements to grain boundaries or to neutron irradiation, which affects K_{1c} and hence K_r. A schematic diagram showing the change of assessment point with time is given in fig.3. This indicates a general movement of the assessment point towards the failure locus as a function of a temporal parameter (time, neutron dose, number of fatigue cycles), at various stages: at start of life, at a periodic safety review (PSR), at end of design life, and in assessment of the possibility of plant life extension (PLEX). Each of these periods of time is associated with a characteristic FAD, essentially a 2-D section of the 3-D total diagram, but *the rate of approach* of the assessment point *towards the failure locus* is controlled by sub-critical crack growth rates and/or by the rates of change in material properties. By considering two FADs at successive temporal intervals, it is possible to trace the movement of the assessment point P-P' as determined by each of the factors relating to crack growth and to changes in material properties. This refines the concept of safety margin or reserve factor, replacing the ratio OQ/OP in fig.1.

The first factor is that of crack growth. In a simple test-piece of width, W, containing a crack of length, a, the stress intensity, K, is given in detail by the compliance function for the test-piece [3], but, for illustrative purposes, may be taken as dependent on $a^{1/2}$: i.e. for constant K_{1c}, K_r depends on $a^{1/2}$. In contrast, the collapse load depends on $(W - a)$ for uniform tensile loading i.e. L_{GY} decreases in a linear manner with increase in a, so L_r also increases with a in a linear manner. In bending the collapse load depends on $(W - a)^2$, so L_r depends on a^2. In each case, a given increase in a causes the K_r component of the vector PP' to be significantly less than the L_r component. A 10% increase in a causes K_r to increase by 5%, L_r in tension to

increase by 10% and L_r in bending to increase by 20%. The consequence is that the assessment point P does not approach the failure locus along the extension of OP (fig.1), but at a decreased angle to it: see fig.4. This situation is, however, drastically changed when the crack is initially small and located at the root of a notch or other stress concentrator. The analysis of this situation shows that, up to about 20% of the root radius of the notch, the value of K at the crack tip increases very rapidly, such that the value of K_r also increases rapidly. The situation is shown in figs 4 and 5 (taken from [3]). The initial rise in K_r is steep, followed by a change to a shallow slope as the crack becomes long. In each of these situations, the *rates of approach* of point P towards the failure locus, by whatever route is followed, will be governed by the appropriate *crack growth rates* for whichever sub-critical crack growth mechanism is experienced.

Effects of *softening* and *hardening* as a result of service exposure simply cause the assessment point P to move along a horizontal line on the FAD: towards $L_r = 1$ if the material is softening: away from $L_r = 1$ if the material is hardening, see fig.6. *Softening* processes in steel components operating at moderately high temperatures, e.g. in boilers or steam pipes in powers stations or in chemical reactor vessels, operating in the range 500-600^{o}C, can be attributed to the further (long-term) tempering of a QT steel, which was originally tempered in the range, say 630-670^{o}C. The rate of such softening, i.e. the *rate of approach* of point P towards the failure locus, is controlled by the *diffusion* in the iron matrix of *alloying elements*, such as Cr or Mo, which are added to form carbides, relatively stable at the operating temperatures, but eventually coarsening. Data for such diffusion rates are available and long-term forward prediction can be effected using activation energies for diffusion. Developments such as 9Cr1Mo steel, to replace 2.25Cr1Mo, have been made to retard carbide coarsening, and hence softening, rates. Note that, if the room temperature yield strength of a material includes a component due to cold work introduced during processing, this contribution may be removed during long-term exposure at temperature, through recovery and annealing. In steel, this depends on the self-diffusion rates of iron, which become significant above ~500^{o}C.

Hardening as a result of service exposure is less frequently encountered, because many materials are specified in an initially hard state, but two forms of hardening can be recognised: *strain-ageing* (in low-strength steels and Mg-containing aluminium alloys) and *neutron-irradiation hardening* (in materials exposed to a neutron fluence in a reactor). Strain-ageing in steel is caused by the diffusion of C or N atoms to pin dislocations introduced by prior cold work and can, in the extreme, increase the subsequent flow stress by up to 200 MPa. Diffusion rates of C, and particularly N, are significant at room temperature, but spectacular effects on the stress-strain-curve are observed for tests carried out between ~150-200^{o}C (the Portevin–Le Chatelier effect). Note here, however, that continuing plastic strain is being applied to a test-piece (creating new dislocations with every strain increment). In general engineering practice, the hardening induced by the first strain-ageing sequence makes it more difficult for subsequent loads to produce plastic flow, so that the contribution of strain-ageing is often limited to a single increment. Similar effects may be obtained in Mg-containing aluminium alloys, in which Mg atoms diffuse to pin dislocations.

Hardening in pressure-vessel steels as a result of exposure to *neutron irradiation* arises in practice from two effects, Knott and English [4]. One is classified as *matrix*

damage. A neutron produced by the fission reaction (theoretically of any energy >40keV, but more usually thought of as a 'fast' neutron, having an energy >0.1MeV, up to ~2 MeV) interacts with an iron atom to produce a primary knock-on atom (PKA), having substantial kinetic energy. The PKA travels through the matrix, losing energy by creating a 'displacement cascade' of vacancies(V) and (self-)interstitials(I), some of which are present as V-I Frenkel pairs, with others as isolated point defects or in the form of point-defect clusters. The clusters, in particular, provide strong barriers to dislocation movement and hence harden the material. In modern practice, the most appropriate measure of neutron dose, D, is in terms of displacements per atom, dpa. A small further hardening component is produced by the γ-ray flux created in fission processes. Interacting predominantly through the Compton effect, the hardening is similar to that of low-energy 'thermal' neutrons.

The second form of *hardening* is particularly associated with sumerged-arc welded (SAW) joints and is *due to the presence of Cu* in the weld (added to improve electrical conductivity in the SAW process). Cu in Fe is a recognised precipitation hardening system, but the normal ageing temperatures are usually >450^{o}C, a temperature considerably higher than those experienced by pressure vessel steels in Magnox or light water reactors (US and USSR), which normally run at <350^{o}C. It appears, however, that the extra point defects created by neutron irradiation allow Cu atoms to become mobile at the lower temperatures, and permit precipitation hardening to occur. In terms of the *rates* of hardening, and hence the *rate* of point P *moving away* from the failure locus, the appropriate temporal measure is neutron dose, D, in dpa, which can be related to time *via* number of hours at power. The matrix hardening is found to vary with the square root of D [4]. Initially, the rate of Cu precipitation would be expected to vary also with square root of D, reflecting the role of irradiation-induced vacancies, but the actual hardening must reflect the mechanism of Cu precipitate hardening, when the particles are very small (possibly, in the form of Cu zones constrained by the iron matrix to be in bcc configuration). One such mechanism is 'modulus hardening', but, a mechanism has recently been proposed which relates to the shearing of zones and transformation of Cu to the fcc form [5]. With further irradiation, it would be expected that the system would overage, but, in safety analyses, after an initial increase, the contribution from Cu is maintained at its peak level. This is conservative with respect to the prediction of transition temperature shifts [4], but, strictly, could be regarded as non-conservative with respect to the point P and the L_r failure criterion. Since the precipitation has been promoted by point defects, there is still a hardening component due to matrix damage, so that, even if the copper is over-ageing, P' never moves to the right of P, the assessment point for unirradiated material.

In contrast to the horizontal shifts to point P occasioned by changes in flow strength, the occurrence of *embrittlement* without change in yield strength leads to a vertical upward shift (see fig.6). A simple example of this is given by classical 'reversible' temper embrittlement, which is attributed to the segregation to prior austenite grain boundaries of trace impurity elements, such as P, Sn or Sb, on slow cooling to room temperature after tempering. Detailed investigation, using isothermal experiments, indicates that embrittlement in forging steels conforms to a flat-topped 'C'-shaped curve, between temperatures of ~580^{o}C and ~420^{o}C. The temperature corresponding to the minimum embrittlement time is ~560^{o}C. It is recognised that similar segregation can occur on cooling thick sections of QT structural steels, such as

2.25Cr1Mo or A533B/A508, after stress relief heat-treatment following welding: under some circumstances, segregation during stress-relief can lead to stress-relief (reheat) cracking. In terms of the temporal variation of the FAD, however, such *embrittlement* has occurred *before* the structure has *entered service*, and should be incorporated in the start-of-life FAD, although similar effects may contribute to Type 4 cracking in service. Classical temper embrittlement in originally unembrittled steel may be observed for chemical reactor vessels operating in (or for long periods, somewhat below) the generally recognised critical temperature range. It is also possible for segregation to occur in neutron irradiated pressurised water reactor steels (particularly in former Eastern Bloc VVER reactors, Davies [6]): as for Cu, the excess vacancy concentration presumably assists segregation at the lower temperatures. The transition temperature shifts in such reactor pressure vessels are large and correlate with the initial P (and Cu) concentration(s) in the steel. In these cases, for both chemical and nuclear reactors, the *rate of approach* of point P to the failure locus is a function of the *rate of embrittlement*. Recent work on both 2.25Cr1Mo and A533B has demonstrated a linear decrease, both of local critical fracture stress ahead of a notch, and of K_{1c}, with the P level on grain boundaries, beyond a threshold level associated with ~20% intergranular (i.g.) fracture, Ding et al [7,8]. The *amount of segregation* is *proportional to the square root of time* (with the free energies for segregation containing a P-Mo interaction term). The conclusion with respect to K_r is that initially the point P does not move: although segregation is occurring it is insufficient to produce the area fraction of ~20% i.g. fracture to reduce K_{1c}. *K_r then increases with the square root of time*. It could eventually stop, when all the available impurity element is fully segregated, or may even decrease, if de-segregation can occur at very long times.

Even in the absence of intergranular segregation, neutron irradiation embrittles steels. The effects have traditionally been measured in terms of (upwards) shifts in notched impact transition temperature, ΔTT, which translate to shifts in curves of fracture toughness *vs.* temperature. The Master Curve [9], for example, gives:

$$K_o = 31 + 77\exp[0.019(T - T_o)] \qquad 1).$$

in which K_o is the value of K_{Jc} in $MPam^{1/2}$ corresponding to a 63.2% probability, p, of cleavage failure, T is temperature, and T_o is the transition temperature where the median toughness (p = 50%) corresponding to a 25mm thick 'high constraint' specimen is 100 $MPam^{1/2}$ and K_o = 108 $MPam^{1/2}$. An upward shift in T_o implies a decrease in K_o at a given temperature, T, but the precise magnitude of this decrease depends first on the equivalence between ΔTT and ΔT_o; secondly, on whether the form of the K_{Jc}:T curve is the same for irradiated material as for unirradiated. Given these factors, there is experimental support for the dependence of ΔTT on the square root of dose, D, in dpa [4], so that there is a basis for calculating the rate of decrease of K_{1c}, and hence increase of K_r with dose as the temporal variable, convertible to time using the time at power. It might be noted that the effect of cold pre-strain is to reduce the value of fracture toughness at a given low test temperature, where the fracture mode is transgranular cleavage, and that such material may be used as a surrogate for irradiated material, e.g. in exploring effects of testpiece size on fracture, Balart and Knott [10]. For both pre-strain and irradiation, the higher yield strength of the material allows the critical fracture stress ahead of a crack to be achieved with a smaller plastic zone than is the case for unembrittled material.

The overall effects of irradiation on the movement of the assessment point, P, on the FAD with time (dose) are shown in fig. 7. Irradiation simultaneously decreases L_r, because the steel is hardened, and increases K_r, because K_{1c} is decreased. The result is that the vector PP' is now oriented in a leftwards, upwards direction in contrast to effects of crack growth, as in figs. 4 and 5, where the direction is always to the right. Although trends in hardening and K_{1c} are both related to the square root of dose/time, this does not imply a slope of -1 for the vector PP', because the L_r component relates to the percentage increase in hardness and the K_r component relates to the percentage decrease in fracture toughness. These percentages will vary, depending on the original, unirradiated values for the flow strength and the fracture toughness. The differences in direction of the vector PP' for crack growth (figs.4,5) for hardening and embrittlement (fig.6), and for neutron embrittlement (fig.7) emphasise the over-generalised nature of the 'safety margin' ratio OQ/OP in fig. 1 and the need for detailed analysis, when addressing the issues presented schematically in fig. 3.

The FAD and Probabilistic Analyses

Many structural integrity assessments are now set within a framework of probabilistic analysis and risk assessment, with due respect to the consequences of failure and societal concern. In particular, the UK's Health and Safety Executive (HSE) has issued three relevant reports: i). *The Tolerability of Risk from Nuclear Power Stations* (1992), ii). *Reducing Risks, Protecting People ("R2P2")* (2001), iii). *The health and safety risks and regulatory strategy related to energy developments* (28 June 2006).A starting-point is the FN-Curve: the frequency at which an event might kill N people plotted *vs* N. Appropriate curves are drawn in fig.8. Although this diagram is not explicitly given in *R2P2,* para. 136 of that report states "Where societal concerns arise because of the risk of multiple fatalities occurring in one event from a single major industrial activity, HSE proposes ..... that the *risk* of an accident causing the *death of fifty people* or more in a single event should be *less than one in five thousand per annum*". In terms of exposure to ionising radiations, the 2006 report (para 302) states that the maximum allowable *annual* whole body dose for workers at a nuclear plant should be less than 20 milliSieverts (mSv) and for members of the public, less than 1 mSv (values at Sizewell B are ~0.5 mSv p.a. and ~0.005 mSv p.a. respectively; natural radiation is ~2 mSv p.a.). The 1 mSv regulatory figure would be regarded as an "Incredibility of Failure" IOF event, corresponding to a (notional) probability of less than 10^{-7}. Reasoning suggests that the probabilities of release of hot, inflammable emissions from chemical plant (such as those giving rise to the *Flixborough* disaster in the UK in 1973) should be set at similar levels.

For the nuclear case, such release is likely to occur only if there is a major brittle failure of the pressure vessel, with release of energy, sufficient for fractured pieces of the vessel to impact on, and breach, the biological shield. Releases due to fractures of pipe-work are likely to be contained within the shield: an issue for plant operators, but not for the general public. The achievement of the required low probability of release rests on a number of factors: the frequency of occurrence of a fault condition (e.g. a pressurised thermal shock) that is able to induce a violent, brittle fracture, and the probability of the material failing under the stresses induced by the fault condition. In terms of resistance to brittle fracture, the safety case assessment would, for conservatism, use the FAD appropriate to the end of (design) life, because this

combines the maximum amount of sub-critical crack growth, together with the maximum degradation in the material's fracture toughness. In general, the direction of the end-of-life vector PP' could lie anywhere between a shallow slope pointing upwards to the right (figs.4,5) and a rather steep slope pointing upwards to the left (fig.7): conceivably, by chance, it could lie along PQ. A simple diagram to illustrate the calculation of probabilities is given in fig.9 which shows the overlap of two distributions: one, the critical defect size under the stress system produced by the faulted condition; the other, the probability of a defect of given size being present in the structure. Often, this second distribution cannot be established as an independent factor and has to be replaced by the capability of the NDT system to detect defects. To support the defect size/NDT capability argument, attempts are made to use expert judgment to set limits on the maximum sizes of defect that need be considered. The failure probability is given by the overlap of the two distributions: unfortunately, in a region where information is least reliable. In the present paper, attention is paid to the distributions of fracture toughness and, to a lesser extent, yield strength, as inputs to the FAD.

A schematic FAD with probability limits is given in fig. 10. It may be noted that the variation in yield strength is much smaller than that in fracture toughness. Values for the standard deviations on yield strength given in Marshall [11] are, for Japanese and French A533B, ~3.5 % of the mean, and for French A508 Class 3 and for US A533B/A508 Class 2, ~6% of the mean. At the 0.01% level (equivalent to ~3.6 s.d.s), the yield strength is likely to be, at most, only some 22% below the mean. In contrast, values of K_{1c} in commercial pressure vessel steels exhibit a wide distribution, and the low probability values, so important to the prediction of overall failure probability, have to be extrapolated from a relatively limited database. One method is to fit available data to a functional form and to extrapolate this to the required low probability level. Early approaches [11] used Gaussian distributions, but the more recent Master Curve [9] approach employs a 3-parameter Weibull form. For Weibull moduli greater than 4, it is difficult to distinguish between the Gaussian and Weibull forms. Free fitting, by Wallin [12] of the Weibull distribution to the Euro data-base to determine the K_{min} threshold value in some cases gave *negative* values.In the Master Curve methodology [9], a (temperature independent) K_{min} value of 20 $MPam^{1/2}$ is assumed. If this is to be taken as the lower bound in safety cases, it provides an ultra-conservative outcome. It may also be noted that the 3-parameter Weibull distribution is such that there is an interaction between K_{min} and the Weibull modulus: the common value of 4 assumed in the methodology is not tenable. Knott [13] has attributed these contradictions to *spatial heterogeneity* in the material and Zhang and Knott [14] showed how this, if unrecognised, can lead to negative values when extrapolations are made: see fig. 11. Following the argument that distributions for "quasi-homogeneous" should be Gaussian (single valued functions plus normally distributed random experimental errors) it is recommended that simple extrapolations should be made only if the s.d. is less than 5-6 $MPam^{1/2}$. In all other cases, detailed metallurgical investigation and re-assessment is necessary.

Durability Issues in Machines: the Turbine Disc

The issues relating to *durability*-dominated behaviour in machines are conveniently encapsulated in a discussion of the lifing of a high-pressure turbine disc, in a gas-turbine used for aero-engines or electrical power generation. Typically the high-

pressure (HP) aero-engine disc is 0.5-1m in diameter, 0.1–0.2m thick and bears a set of blades of aerofoil cross-section around its circumference, usually attached by "fir-tree root" fittings (which introduce stress-concentrations). The first stage HP turbine disc is subjected to the highest temperatures but all discs in a turbine experience high stress because design minimises the overall mass of the turbine. Discs are classified as safety-critical components. A disc failure, if uncontained, could not only destroy an engine, but could also cause heavy fragments to impact at speed onto the body of the aircraft.

The main threat to integrity is perceived as the initiation and propagation of fatigue cracks under the combination of major and minor cycles involved in the service duty. Various lifing philosophies may be employed. The first is that of "life-to-first-crack", where "first-crack" is defined as a surface-breaking, semi-circular crack of radius 0.375mm (surface length 0.75mm). Under current U.K./European Standards for both military and civil aircraft, service lives can be "declared" (approved) only from the results of "spin-rig" tests carried out on full-size discs at representative stresses and temperatures. Such tests are expensive (over £100,000 per test) and the consequence is that only a few full-size tests are carried out. More information can be obtained from smaller-scale testpieces (which will be also employed in a "screening" sense to evaluate new alloys or process routes) but, in the UK, approval rests on full-scale testing. The U.S (FAA) regulations permit the declaration of service lives solely on the basis of specimen tests.

It is assumed that fatigue lives are distributed in a log-normal manner, with a scatter of not more than 6 in life between –3s.d. and +3s.d. The -3s.d. limit is used for approval purposes. This also must be at the 95% confidence level, which implies quite large safety factors if the number of full-size disc samples is small. For a mean life of 10,000 (equivalent) major cycles, the safety factor is 4 if one disc only is tested, 3.05 if five are tested, 2.86 if ten are tested, Harrison and Weaver [15]. If a new design enters service, the leading disc is removed after 50% of its declared "safe life" and is subjected to further spinning to "dysfunction" (onset of rapid crack growth), whilst the remainder are allowed to proceed to 75% of declared "safe life". Once the result from the ex-service disc is known, it is added to those from the original disc tests, the 95% confidence level is re-calculated, and the final "declared safe life" for the remaining discs in service is established. This procedure is conservative and leads to "very remote" possibilities of disc failure. It may be unduly conservative, leading to further potential "life" being wasted.

A second philosophy is that of "two-thirds dysfunction" in which the 0.75mm "first crack" criterion is replaced by a set fraction (here two-thirds) of the total *life* to failure or to the onset of rapid crack growth (dysfunction). For the lower strength disc alloys, this criterion may not differ widely from that of the 0.75mm "first crack", but in high strength alloys, the crack length at "two-thirds dysfunction" can be less than 0.75mm. The context is that disc alloys should not be regarded as "defect free", but as containing an initial distribution of defects to which linear elastic fracture mechanics (LEFM) analyses can be applied to calculate fatigue-crack growth-rates.

This is particularly the case for the newer, higher-strength discs for which the processing has changed from the conventional cast-and-wrought route (which exhibited unacceptable "banding" of alloy-rich and alloy-lean regions in the final

forging) to the HIPping (hot isostatic pressing) and forging of gas-atomised power. For such material, small particles of (alumino-silicate) refractory lining can occasionally be scoured from the crucible or tun-dish and may pass through the powder-sieve to enter the disc. The size of such a particle could potentially lie in the range 0.1-0.4mm; and is clearly recognisable as a "defect" which serves as a fatigue-crack initiator. The following section describes the application of the LEFM treatment of fatigue to assess the significance of such defects with respect to fatigue life. Although the example is specific, the analysis is generic with respect to other types of defect of similar size.

Long-crack growth-rate data for typical nickel-based disc alloys heat-treated to yield strengths in the range 820-980MPa conform to a Paris-law relationship of the form:

$$da/dN = 4x10^{-12}\Delta K^{3.3} \qquad 2).$$

in which da/dN is the crack growth increment per cycle in m/cycle and $\Delta K = (K_{max}-K_{min})$ is the instantaneous stress-intensity factor range. Additionally, fatigue-crack growth-rates initiated from corner cracks of 0.25mm radius agree well with the long-crack data. Equation 2). is a good starting point for estimating the lifetimes of discs containing defects of the size of refractory inclusions. The calculations refer to a lifetime of 10,000 (equivalent) major cycles, taking the stress-range to be equal to the 0.2% proof strength, assuming zero-tension loading and with no allowance made for the fatigue crack threshold. Note that the design engineer bases stress levels on the strength of the material: the higher the strength, the higher the stress level. The fracture toughness is taken as $100MPam^{1/2}$. The value of ΔK is equated to that for a semi-circular edge crack of radius a:

$$\Delta K = 0.67\Delta\sigma\,(\pi a)^{1/2} \qquad 3).$$

Table 1 gives results, based on fatigue crack growth integrations for the initial "critical" size of defect, a_i ("critical" in the sense that it leads to failure in 10,000 cycles) and the size to which a defect grows after 3300 cycles and 6700 cycles ("one-third dysfunction" and "two-thirds dysfunction" respectively). The "first-crack" surface length of 0.75mm, $2c$, corresponds to a value a = 0.375mm for a semi-circular crack.

Table I: Crack sizes as functions of Stress Range and Number of Cycles

Stress Range = 0.2%Proof Strength. MPa	a_i mm	a(3300) mm	a(6700) mm	a(10,000) mm
800	0.30	0.50	1.20	11.1
850	0.23	0.40	0.97	9.8
950	0.13	0.23	0.58	7.9
1000	0.10	0.18	0.46	7.1
1200	0.04	0.08	0.20	4.9

The strength level 800-850MPa corresponds to that of *Waspaloy*, 950-1000MPa equates to that of *Astroloy API*, and 1200MPa to that of *René 95*. It is instructive to compare the figures in Table 1 with the "life-to-first-crack" and "two-thirds dysfunction". At the *Waspaloy* strength level, "life-to-first-crack" (a = 0.375mm) is clearly highly conservative compared with "two-thirds dysfunction". It is achieved, in round figures, between 1500 cycles (800MPa) and 3000 cycles (850MPa), compared with 6700 cycles. At the *Astroloy API* level, there is a degree of conservatism, but this is not excessive: 5400 cycles (950MPa) or 6300 cycles (1000MPa). On the other hand, at the 1200MPa level, the "life-to-first-crack" is some 8300 cycles, well beyond "two-thirds dysfunction" and with only 1700 cycles left before final failure. The figures obviously have to be modified to take account of the safety factors associated with a 95% confidence level but indicate the way in which strength level (stress range) affects the assessment strategy: see also fig. 2. The reason that *Waspaloy* has a reputation as a "fatigue-tolerant" alloy has as much to do with low applied stress levels as with inherent resistance to fatigue-crack propagation.

These calculations have been made simply for the major "take-off/landing" cycles, when the engine experiences a stress-range from zero to full design stress. In addition to these, the engine in flight will undergo further cyclic loading of "significant" range, in that the threshold is exceeded, although the life at any such cyclic load range would be greater than 10^4 cycles. Such cyclic loads are incorporated into the lifing calculations using Miner's Law. Traditionally, this is based on the S-N curve and is put in terms of the number of cycles N_j at a given stress range $\Delta\sigma_j$ and the number of cycles to failure at that stress range, $N_{f(j)}$. Then, for all j,

$$\Sigma\,(N_j/N_{f(j)}) = 1 \qquad 4).$$

Although this is an empirical approach, it can, with a few simplifying assumptions, be derived from successive integrations of the fatigue crack growth-rate, assuming no "overload" accelerations or retardations, Knott [16]. In this way, "major-cycle" duty cycles can be factored-in to give the "equivalent" 10,000 life cycles. A logical implication is that duty cycles (of both stress and temperature) for every aircraft mission should be recorded to enable continuous modifications to be made to lifing assessments and maintenance/overhaul schedules.

Using equation 3). the initial values of ΔK for *Waspaloy* (800-850MPa) are 15.3-16.4MPam$^{1/2}$, for *Astroloy* (950-1000MPa) 11.9-12.8MPam$^{1/2}$ and for *René 95* (1200MPa) 9MPam$^{1/2}$. Values for long-crack threshold values at low stress-ratio in nickel-base superalloys range from 5.8 to 16.3MPam$^{1/2}$, Tsubota et al [17], but these values are to a large extent a function of the "wind-down" procedure employed, which leads to closure forces in the plastic wake behind the crack tip. Highly crystallographic slip can give rise to pronounced surface roughness, which produces closure at high K_{min} values. Studies of short crack growth in Astroloy API, Kendall et al. [18], suggest a threshold value of approximately 3MPam$^{1/2}$. The calculations in the previous section are therefore somewhat conservative, but the general nature of the conclusions still holds.

It is of interest to examine the extent to which the threshold values may be used to estimate the tolerable levels of vibration stresses: "minor cycles". At 9,000 r.p.m (150Hz) the number of such cycles accumulated over a 2h flight is approximately 10^6

greater than the number of flight cycles, so that non-exceedance of threshold, if this could be achieved, would be a reassuring safeguard. Considering *Astroloy* at a yield strength of 950-1000MPa, the threshold value of 3 $MPam^{0.5}$ may be combined with initial defect sizes of 0.13mm and 0.1mm, using equation 3). to deduce that the maximum tolerable vibration stress level should be of order 220-250MPa. Given the possibility of a defect in the region around a fir-tree root, having a significant stress-concentration (say, 4 - 5), this might be reduced to a value of order 50-60MPa. The initial defect sizes of 0.13mm and 0.1mm are however only "critical" in the sense that they lead to failure in 10,000 cycles. Given rigorous process control, the significance of minor cycles rests on the detectability of growing cracks, since a 10^6 magnification factor implies $3x10^9$ cycles if the inspection interval is set at one-third the life to dysfunction. From Table I, for *Astroloy*, an initial defect could have grown to 0.18-0.23 mm in 3300 cycles. A detection capability with resolution no better than 0.36-0.46mm reduces the above stress limits by 1.4, i.e. to 160-180MPa globally, or to 35-42MPa near stress concentrators. The value of *a*for "life-to-first-crack" is 0.375mm, $2a = 2c$ for the surface length of a semi-circular crack = 0.75mm.

In practice, it is often found that the LEFM calculations are unduly conservative with respect to the behaviour of practical systems. This can be attributed to effects of compressive residual stress, which act at two levels. First, microscopic, "tessellated" stresses can be generated around the refractory inclusions. During the initial solution heat-treatment of the alloy, at temperatures of order 1100^o-1200^oC, inclusion and matrix are in stress-free equilibrium. On cooling to a lower temperature, differential thermal contraction between the inclusion (low coefficient of thermal expansion, CTE) and the matrix (higher CTE) sets up a self-equilibrating residual stress distribution (compressive in the inclusion, tensile in the surrounding matrix), the magnitude of which depends on ΔT, the difference between solution temperature and operating temperature. This stress distribution has the effect of "pinning" short cracks, causing them to grow more slowly than predicted by simple analysis. The effect is much larger at room temperature (large ΔT) than at 600^OC (smaller ΔT) – see Woollin and Knott [19].

The second effect is at the meso/macro level. Surface compressive stresses are induced by shot (glass-bead) peening of external surfaces, particularly the bore and bolt holes, and retard the growth of short cracks in near-surface regions. Assessment of the effect of shot-peening on fatigue life is achieved by carrying-out fatigue tests on "big blocks" of material, which have been subjected to peening treatments identical to those experienced by discs. The fatigue surfaces are examined in detail and measurements are made of the striation spacing, da/dN, as a function of crack length a. Knowing the number of cycles to failure the measurements can be back extrapolated to determine an "effective initial flaw size" EIFS, which can then be used in conjunction with LEFM analysis to provide a lifing procedure for service discs. These data represent an important technological acquisition, but are, of course, specific to a particular combination of alloy (yield strength, work-hardening characteristics), defect content and peening procedure. There is advantage to be gained in modelling the effect of compressive stress in a quantitative manner, particularly to explore the sensitivity of performance to details of peening parameters and defect content and character.

With appropriate precautions, the residual stress distribution after peening can be determined using X-ray diffraction and successive (electrochemical) removal of surface layers. This distribution can then be used, via a suitable weight function, to calculate the "negative K" produced by the compressive stress. This subtracts from the "applied K" to produce an "effective K". Since the stress is monotonic in character, the "negative K" subtracts from both K_{max} and K_{min} separately. Values less than zero are ignored in the calculation of ΔK. As a final point, it might be noted that, although shot peening is beneficial with respect to retardation of growth from small surface defects, there is a concern that, as disc temperatures rise above 650^{o}C (at the rim) the dislocation sub-structures associated with the plastic deformation begin to anneal-out, removing the compressive stress, but leaving a surface which has been roughened by the peening process, and which is hence more likely to initiate cracks.

Up to this point, discussion has focused on powder alloys, with identifiable refractory inclusion defects, the sizes of which can, in principle, be determined by the mesh size of the sieve used to grade the alloy powder. Setting-aside the effects of shot-peening, the "damage tolerance" lifing philosophy accepts that, for all discs, a defect, such as a machining mark, is always present from the start of life. The initial defect size that can be used in lifing procedures then rests on the non-destructive inspection (NDI) crack-detection size at the 90% to 95% confidence level. Even for surface defects, this is unlikely to be better than $2c$ = 0.2mm surface length (a = 0.1mm) and will be quite substantially worse for sub-surface defects. For ultrasonics NDT at 2MHz, even 2mm is optimistic for buried defects, but, in the absence of internal porosity or refractory inclusions, defects of concern are likely to be located at the surface. In ultra-clean material, 50MHz ultrasound could detect at the 0.1mm level. Comparison with the values of a for different (equivalent) cycles given in Table 1 shows that, at stress levels up to 1000MPa, there is a reasonable chance of detecting surface defects at 3300 cycles (even at 1000MPa, $2a = 2c$ = 0.36mm) and a very good chance at 6700 cycles (at 1000MPa, $2a$ = 0.92mm). For the 1200MPa level, it is unlikely that the defect of main concern will be detected at the 3300 cycles inspection ($2a$ = 0.16mm). Meticulous inspection is needed at 6700 cycles ($2a$ = 0.4mm) to ensure integrity.

In summary, the following principles may be deduced with respect to turbine disc lifing. Rotating components must exhibit high integrity for the planned lifetime (*durability*). Discs are safety-critical components and initial design must therefore embrace lifing methodologies and any necessary re-iterations before a product comes to market. Manufacturing introduces "fatal flaws". For the design/lifing activity to be credible (other than as a "retro-fit" to a service failure) analysis of the manufacturing parameters and lifing consequences needs to be carried out. The role of NDI needs to be thought through carefully, in terms of both spatial resolution and inspection period. For disc alloys, this period (and associated resolution) is a function of stress range (yield strength). The effect of shot-peening is beneficial, but it would be useful to produce a quantitative model so that it can be "designed-in" to the system rather than used as a qualitative "add-on". As disc temperatures increase, the possibility arises of annealing-out the plastic deformation associated with the residual stress whilst maintaining surface roughness and this needs to be investigated carefully. Most of these principles can be applied directly to other machine components. *Stress levels* and *initial defect sizes* will be critical with respect to *durability* and most initial defects of significance will be smaller than the resolution of the NDI technique employed; certainly for buried defects, if not for surface-breaking defects. A variety

of surface hardening techniques may be employed as an alternative to shot-peening (case-carburising, case-nitriding, induction hardening etc.): each of these will also generate a residual stress distribution, which will assist in retarding initiation and the early stages of crack growth. Quantifying the effects of these surface treatments on durability remains a fascinating and important area. In all examples, it is important to recognise that the design engineer tends to scale applied stresses with yield or flow strength level, with the consequence that a given size of initial defect is subjected to higher stresses in harder material, hence increasing sub-critical crack growth-rates and reducing the life. It is of interest in this respect that when hard (carburised, nitrided etc) surface layers are subjected to stress levels based on the matrix yield strength level, they provide good protection against fatigue.

Conclusions

The paper drawn attention to two extremes of assessment: one extreme related to the *integrity* of *functionally static structures*; the other to the *durability* of *functionally dynamic machines*. In the examples chosen, the *integrity* issues have been associated, not only with crack growth, but also with changes in material parameters. The consequence of this is that *both* the *spatial position* and the *temporal position* of the *assessment point* on the FAD with respect to its *proximity to the failure locus* can *vary* quite dramatically, depending on which factors are dominant. The need to set safety cases within a probabilistic approach emphasises low probability events: both the limits of defect distribution and extrapolations of fracture toughness. Extrapolations need to be supported by good physical/metallurgical understanding. The *durability* issues emphasise sub-critical crack growth-rates and initial defect distributions, where the initial defect size to guarantee life is smaller than the NDI detectability limit. Defect control is exercised by rigorous process control and usually by some surface treatment that produces a hard layer in which compressive residual stresses are generated. For all such components, it must be recognised that applied stresses scale with strength level, so that alloys of higher tensile strength appear to exhibit worse behaviour than expected, simply because the cracks grow more rapidly in response to higher applied stresses.

References

(1) Milne, I. Ainsworth, R.A. Dowling, A.R. and Stewart, A.T. *Assessment of Structures Containing Defects*, British Energy ltd. Report R/H/R6 Rev. 3, 1998.

(2) Olver, A.V. Wilson, D. Shaun, P. and Crofton, J., Eng. Fail. Anal Vol. 14, 2007, pp 1224-1232.

(3) Knott, J.F. and Withey, P.A., *Fracture Mechanics - Worked Examples*, The Institute of Materials, Book No. 550, 1993.

(4) Knott. J.F. and English, C.A. Intl. Jnl. Press. Vess. and Piping, Vol. 76, 1999, pp 891-908

(5) Bacon, D.J. and Harry, T. Acta Mater. Vol.50, 2002, pp 209-222

(6) Davies, L.M. A Comparison of Western and Eastern Nuclear Reactor pressure Vesssel Steels AMES Report No. 10, European Commission report EUR 17327 EN, 1997, pp 1-68

(7) Ding, R.G. Islam, A. Wu, S. and Knott, J.F. Mat. Sci. Tech. Vol 21, 2005, pp. 467-475.

(8) Ding, R.G. Rong, T.S. and Knott, J.F. Mat Sci. Tech. Vol. 21, 2005, pp 1255-1260.

(9) ASTM *Standard Test method for Determination of Reference temperature, T_o. for Ferritic Steels in the Transition Range*. E1921-05, 2005, pp. 1-20.

(10) Balart, M.J. and Knott,J.F. Intl. Jnl. Press. Vess. and Piping, Vol. 83 , 2006, pp. 205-215

(11) Marshall, W. An Assessment of the Integrity of PWR Pressure vessels. UKAEA, March 1982, fig. 3.6

(12) Wallin, K Eng. Fract. Mech. Vol.69, 2002, pp. 451-481.

(13) Knott, J.F. Fat. And Fract. Eng. Matls. And Struct, Vol. 29, 2006, pp 714-724.

(14) Zhang X.Z. and Knott J.F. "The Statistical Modelling of Brittle Fracture in Homogeneous and Heterogeneous Microstructures" Acta. Mater. Vol. 48, 2000, pp 2135-2146.

(15) Harrison G.F. and Weaver M.J. in *Engineeering in Fatigue* ed. J.H.Beynon et al. A.A.Balkema (Rotterdam) 1999, pp. 713-720

(16) Knott J.F. in *Advances in Fracture Research* (Proc. 9th Intl.Conf. on Fracture) ed. B.L.Karihaloo et al. Pergamon, 1997, pp. 1213-1224

(17) Tsubota M. King J.E. and Knott J.F. *Proc. First Parsons International Turbine Conference,* ed. D.M.R.Taplin, Parsons Press (Dublin) and Inst.Mechl Engrs. (London), 1984, pp. 189-196.

(18) Kendall J.M., King J.E., Woollin P. and Knott J.F. *Proc. Conf. PM Aerospace Materials* (Luzern) Metal Powder Report Services Ltd. 1987 pp.7.1-7.12

(19) Woollin P. and Knott J.F. *Proc. 3rd Intl. Conf. on Fatigue* ed. R.O.Ritchie and E.A.Starke Jnr. EMAS 1987, II, pp. 1087-1099

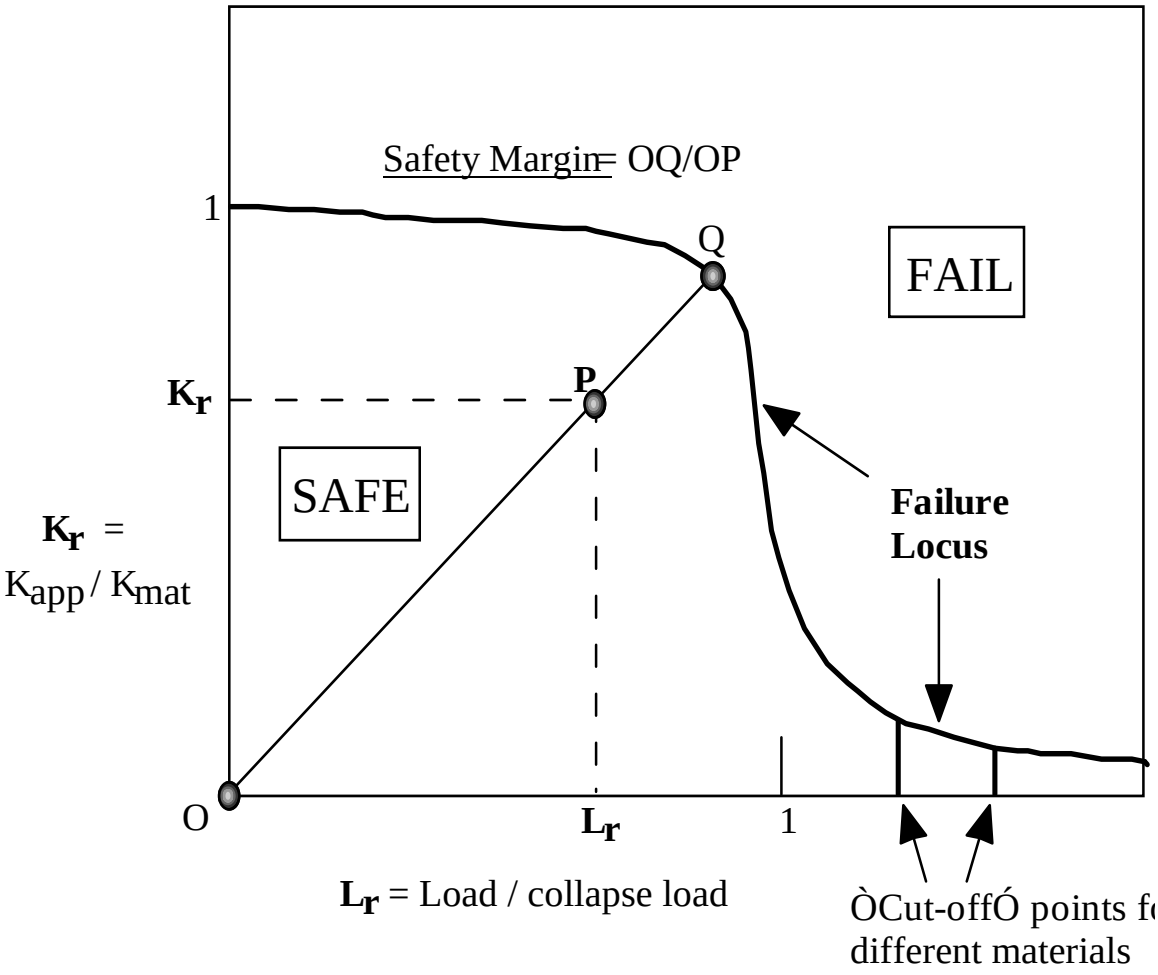

Fig. 1 The Failure Assessment Diagram – see refs. [1] and [3]

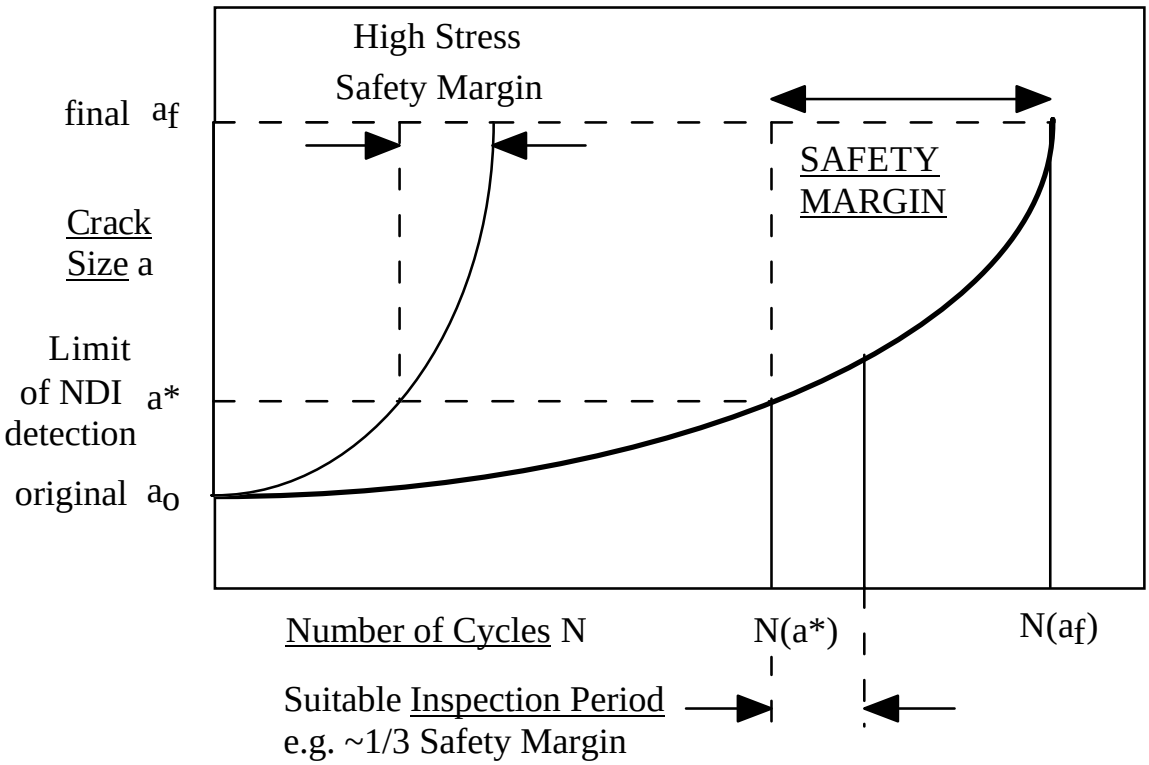

Fig. 2 Schematic diagram showing increase of fatigue-crack length with number of cycles and indicating safety margins between first detection of a crack and failure for different stress levels

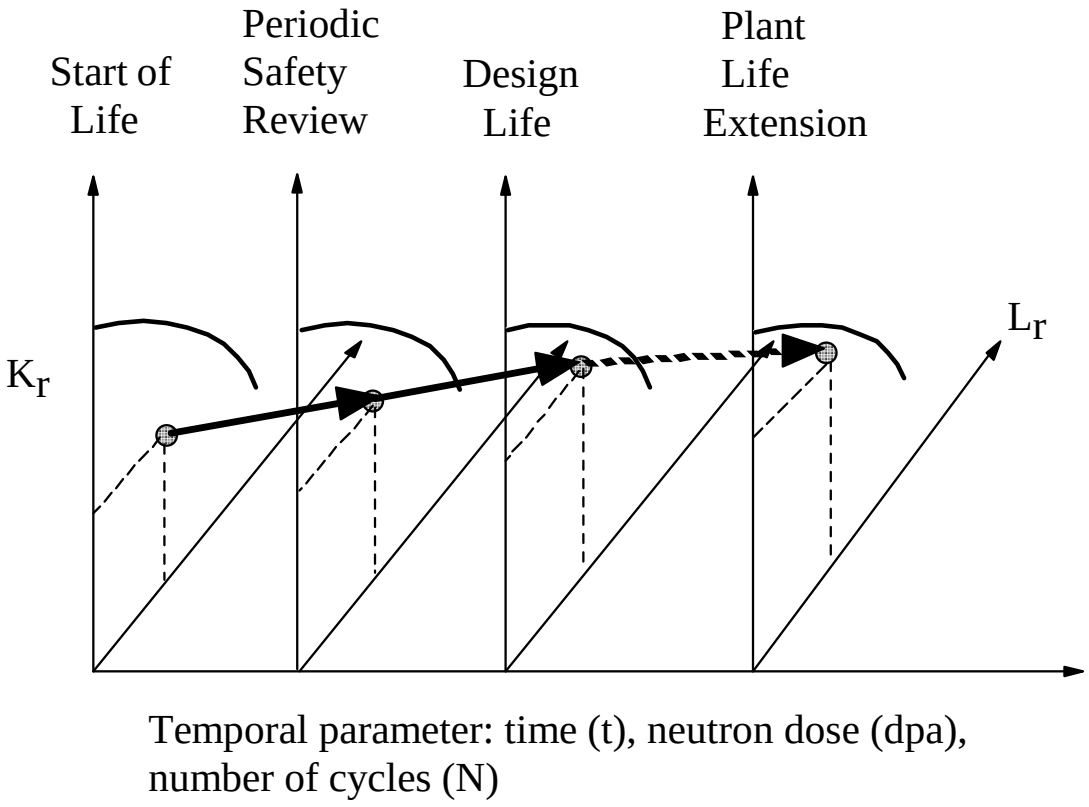

Fig. 3 The temporal variation of the Failure Assessment Diagram

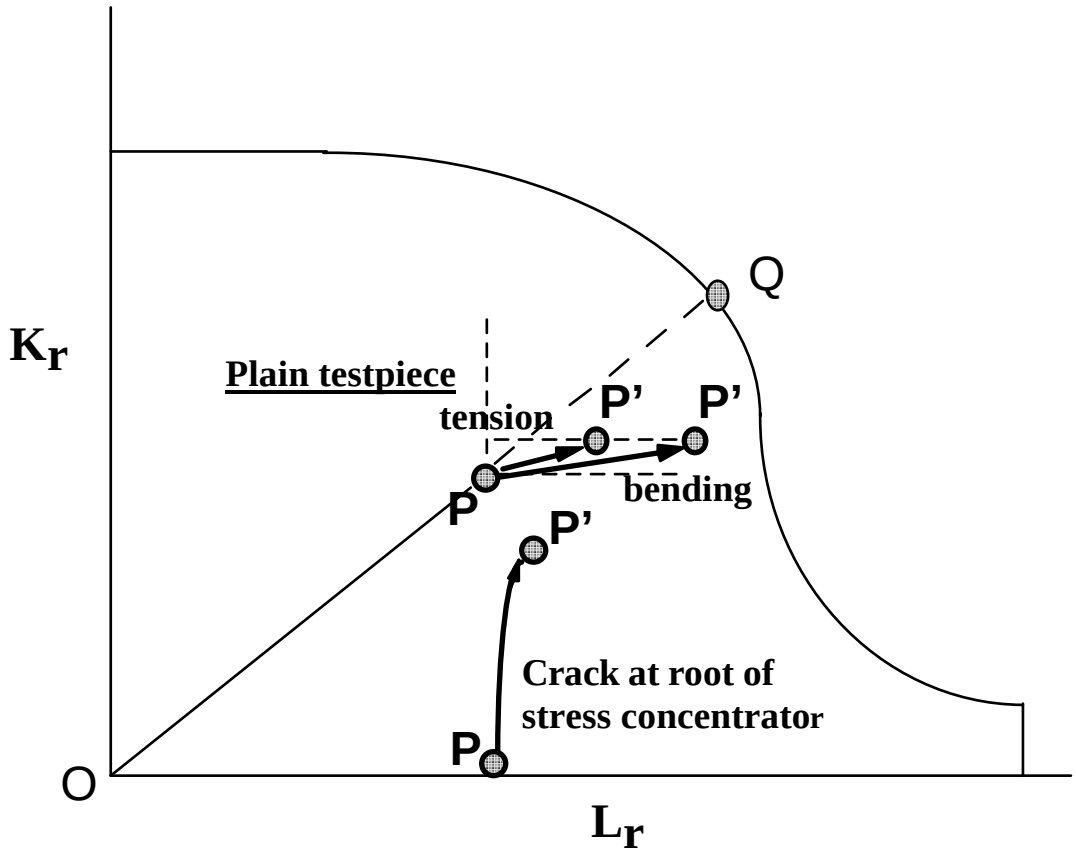

Fig. 4 The Effect of Crack Length on the Position of the Assessment Point on the FAD

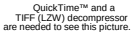
QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

Fig. 5 The FAD for a crack growing from a stress concentrator – see [3]

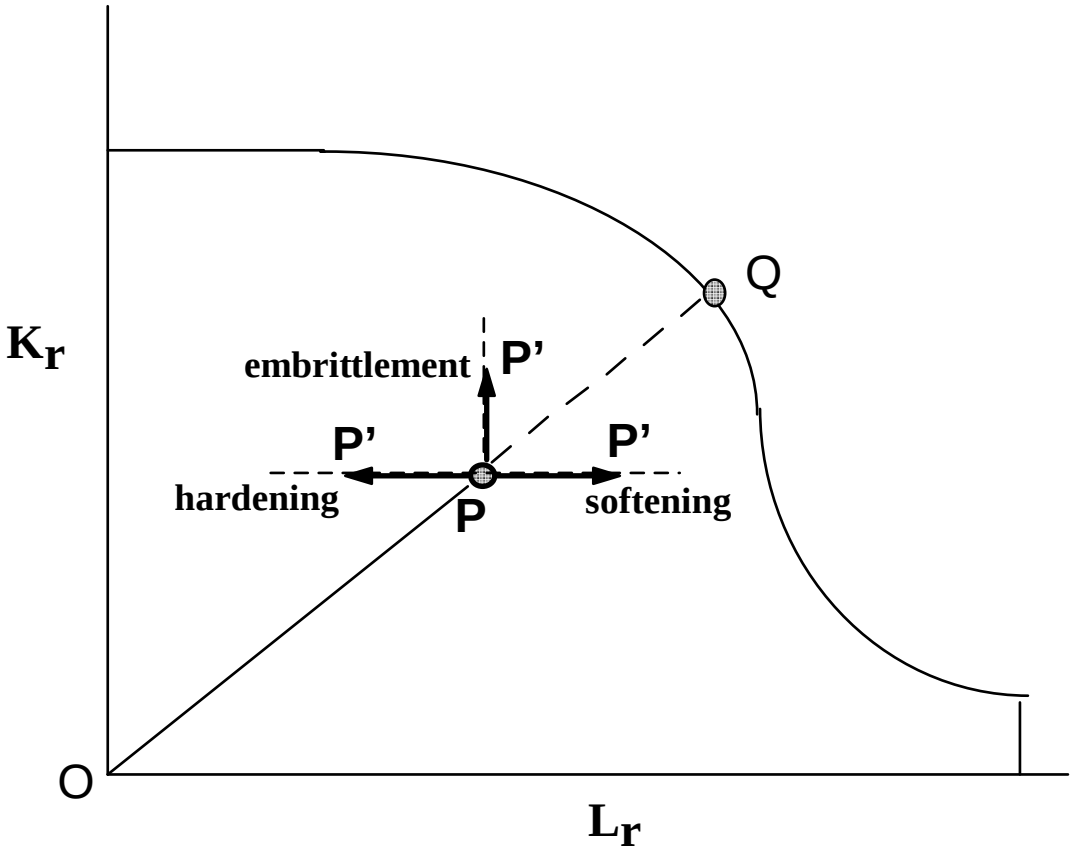

Fig. 6 Effects of Softening, Hardening, Embrittlement on the Position of the Assessment Point

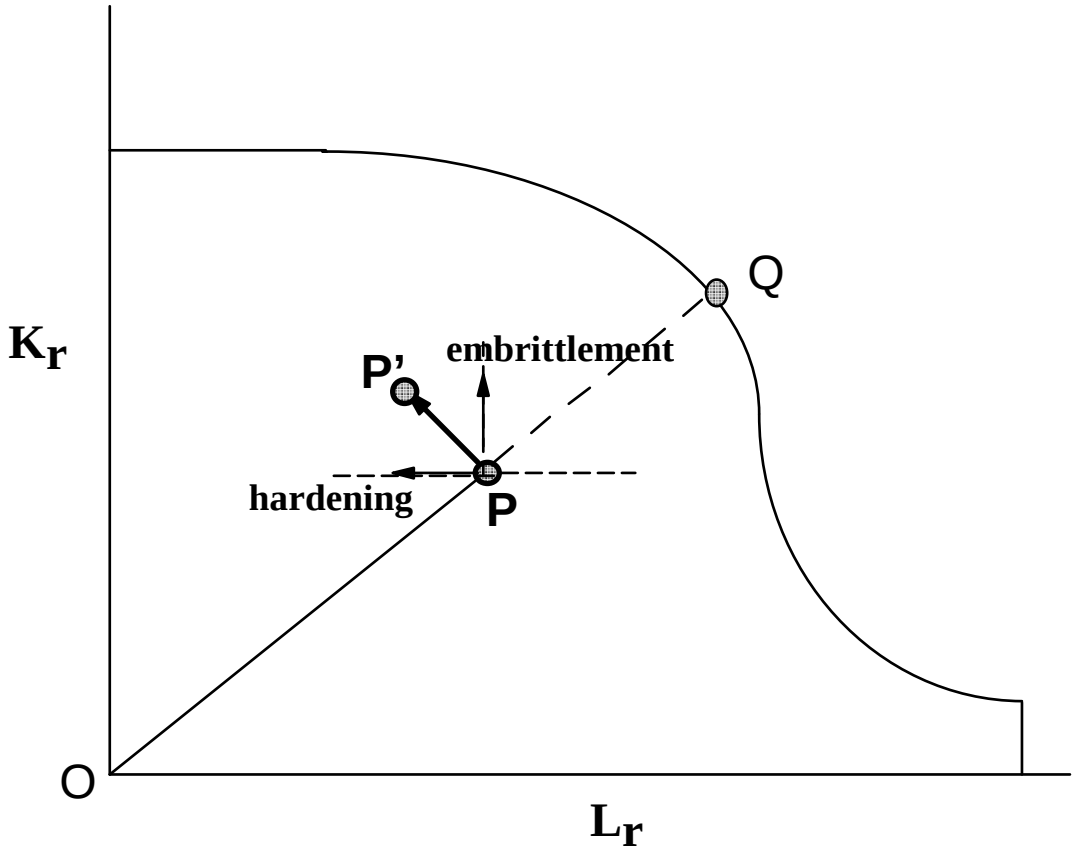

Fig. 7 Effect of Neutron Irradiation on the Position of the Assessment Point

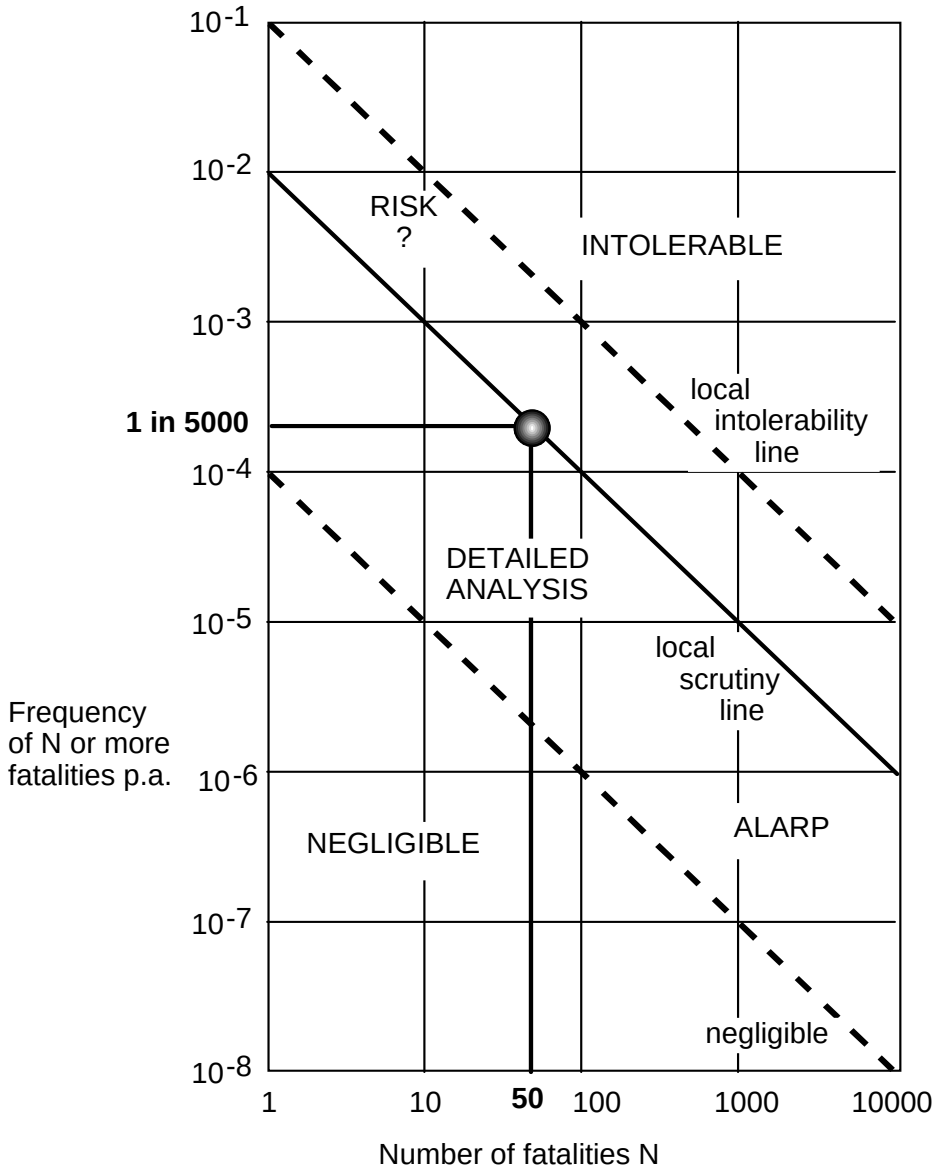

Fig. 8. The "F/N" curve given in "R2P2" 1999, but not in "R2P2" 2001, relating the frequency of fatalities to the number of people involved in the fatality. The <1 in 5000 figure for 50 or more fatalities is referred to in "R2P2" 2001 (see text).

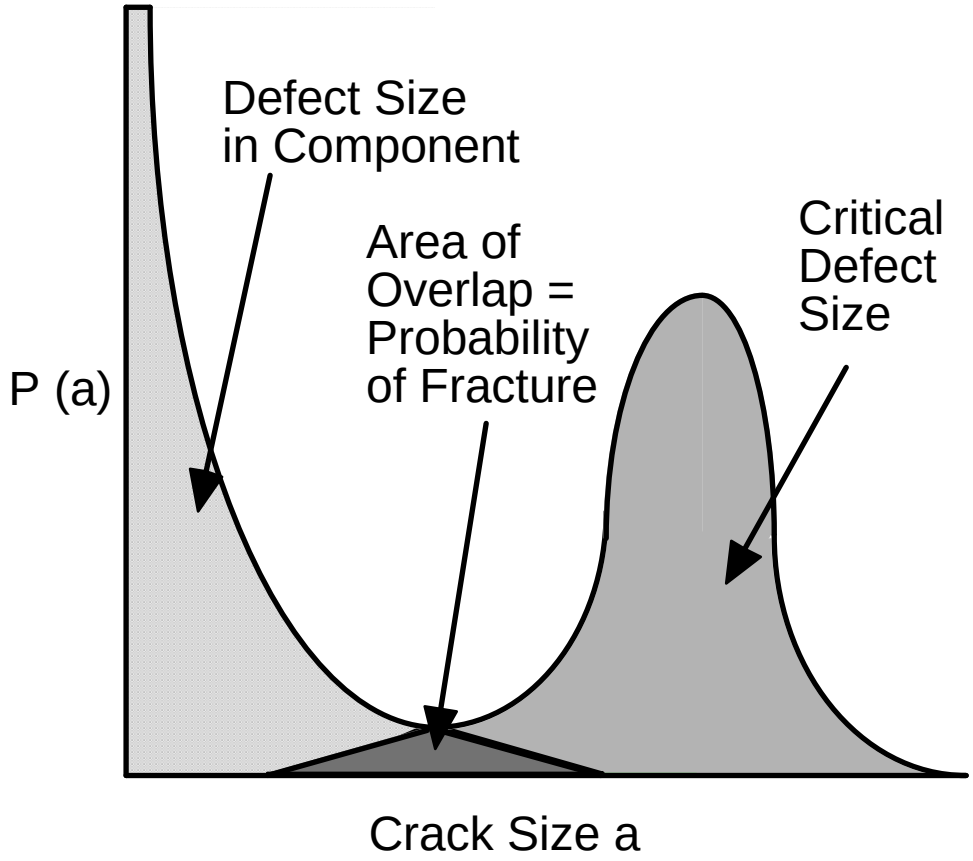

Fig. 9 Schematic diagram indicating the way in which the probabilities of the presence of a defect and the critical size of defect are combined to derive the overall probability of failure

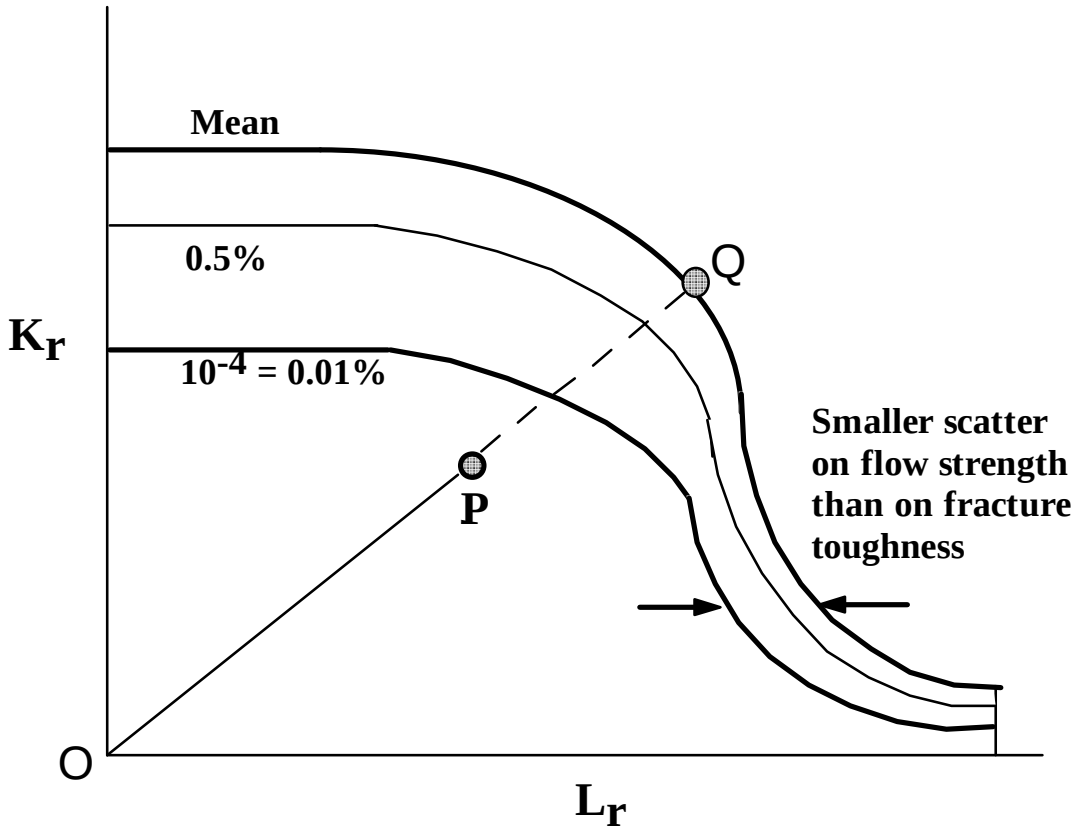

Fig. 10 Probability Limits on the FAD

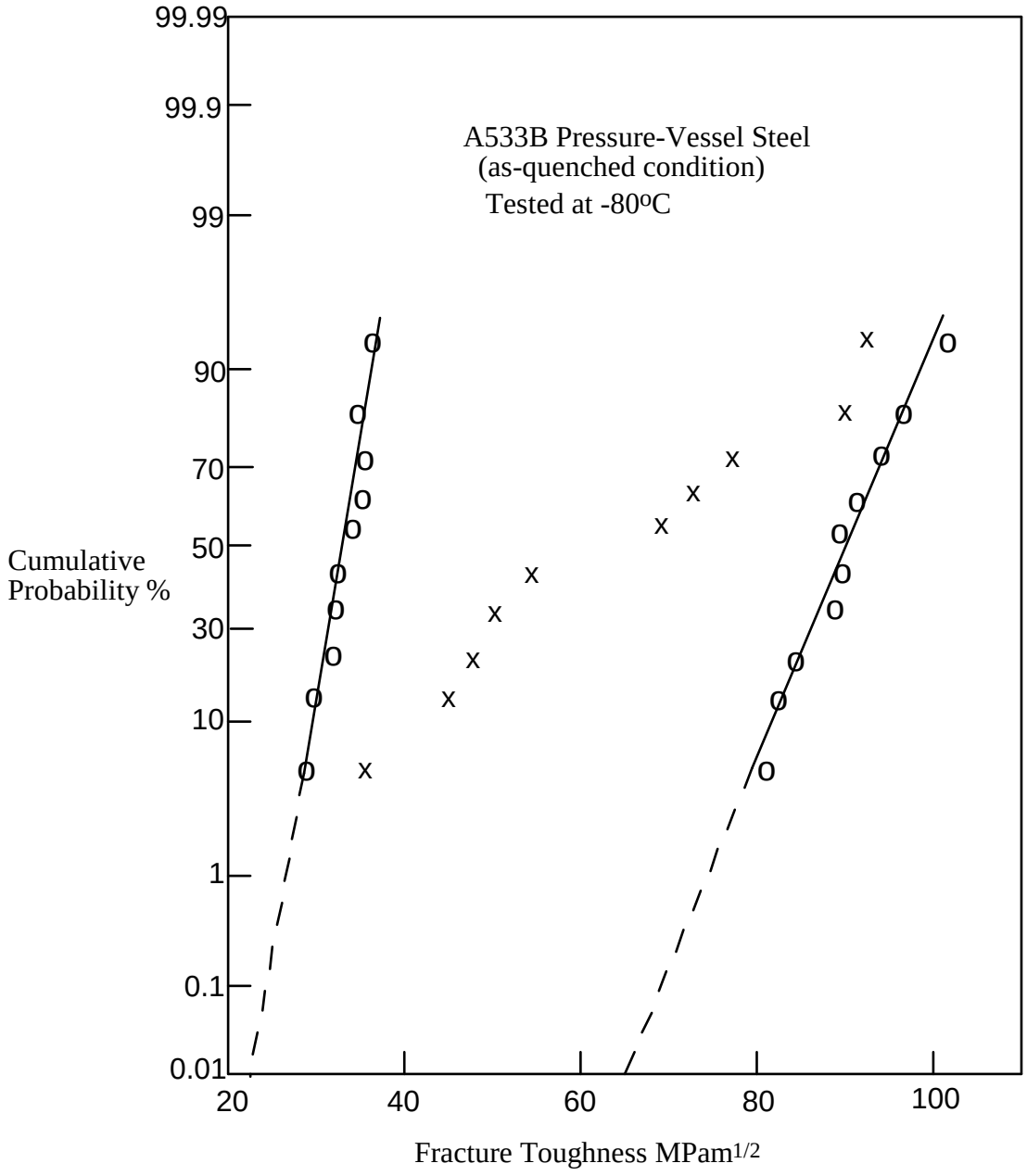

Fig.11. Cumulative failure probabilities, plotted on normal probability paper, for Martensitic, Bainitic and a mixed Martensitic/Bainitic microstructure in A533B Steel. Note that uninformed extrapolation of the points for the mixed microstructure would give a negative fracture toughness at the 0.01% level – see ref. [14]