Fatigue and Fracture Mechanics in Aerospace Structures

George A. Kardomateas[1] and Philippe H. Geubelle[2]

[1] *School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, USA*

[2] *Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

1 HISTORICAL EVOLUTION OF FATIGUE

Although the occurrence of some forms of fatigue was probably not uncommon during the bronze and iron ages, it was not until the middle of the nineteenth century that fatigue was recognized as a problem that had to be addressed by engineers. The primary sources of concern were incidents in which fatigue failures were occurring in railway axles. The increased use of rail transportation led, particularly in Germany and England, to a general consensus that experimental studies, which could provide data for design limitations, should be initiated.

Early experimental investigators of fatigue behavior very logically attempted to duplicate as nearly as practically possible the conditions that were developed in the service incidents, which led to failures. Wohler in Germany, for example, initially conducted fatigue tests on full-size railway axles. Subsequently, he conducted tests on small specimens, and designed machines that produced cyclic bending, reversed bending, uniaxial loading, and torsion. A concise review of Wohler's contributions appeared in an article in the 23 August 1867 issue of the British weekly journal, *Engineering*. Descriptions of other early work are given in the texts by Timoshenko (1953) and Moore and Kommers (1927). A more recent summary that also contains descriptions of both features of modern testing systems and commonly used test specimens is contained in a text by Fuchs and Stephens (1980).

Fatigue tests are conducted on full-scale components such as an aircraft wing attached to a fuselage in order to duplicate the complex diffusion of stress into critical areas. They are sometimes conducted on sub-assemblies and components that are parts of a total structural system. Finally, they are often conducted on small specimens that are designed to provide fatigue data, which reveal the effects of stress state, surface preparation, environment, and loading history on a material of interest.

Encyclopedia of Aerospace Engineering.
Edited by Richard Blockley and Wei Shyy
 ISBN: 978-0-470-68665-2

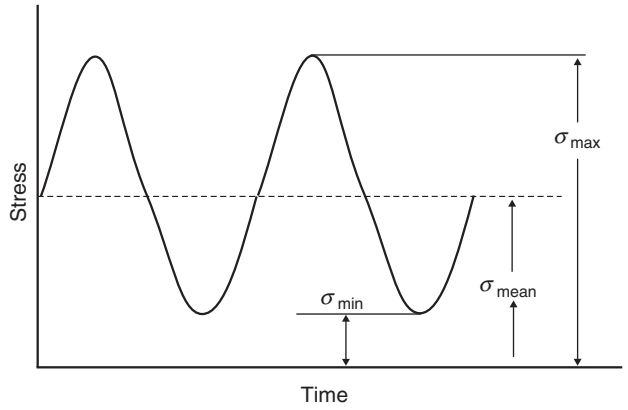

Figure 1. Cyclic variation of stress.

Prior to about 1960 the primary goal of fatigue testing was to obtain stress versus cycles to failure data on specimens that initially were nominally free of cracks. There has been a widespread increase in the use of precracked specimens since then, however, and the data obtained have been used in conjunction with the application of fracture mechanics to form the basis for the development of a new philosophy of design. The goal of this philosophy is to determine the load carrying "tolerance" of a component in which a crack is present (damage tolerance).

2 ENGINEERING CHARACTERIZATIONS OF SAFE-LIFE

For applications in which the amplitude of the cyclic loading is more or less uniformly repeatable and the desired lifetimes involve millions of cycles with moderate levels of loading, design can be based on data obtained from tests in which the loading is of the type depicted in Figure 1. This is the stress-life approach. Another topic concerns high loading levels and smaller cycles to failure and it is described as a strain-life approach.

Figure 1 shows a typical cyclic stress history. Two loading parameters that are used are, by definition, the stress range

$$\Delta\sigma = \sigma_{\max} - \sigma_{\min} \tag{1}$$

and the stress ratio

$$R = \frac{\sigma_{\min}}{\sigma_{\max}} \tag{2}$$

Since $\sigma_{\min}$ can be chosen to be compressive, it follows that R can be negative.

3 THE STRESS-LIFE DIAGRAM (S–N CURVE)

To obtain a stress-life diagram, uniform, constant amplitude tension–tension (positive stress ratio R) and tension–compression (negative R) stress-life tests are conducted. The test traditionally used to obtain basic stress-life data (also called S–N curve) is the rotating bend test for which the stress ratio $R = -1$. Figure 2 shows the schematic of a four-point rotating bend test. The applied load is static but the specimen is in rotation. Because of the rotation, the entire surface material is tested under maximum stress. A cylindrical, hourglass gage section that has a highly polished surface is used. The high rotational speeds that are employed make it possible to accumulate large number of cycles in a reasonably short period of time. The stress values used from reversed bending tests are the maximum bending stresses computed from elementary beam theory.

Because of data scatter, multiple tests are conducted at each stress level, and the test data are summarized on stress–log cycles to failure plots, of the type shown in

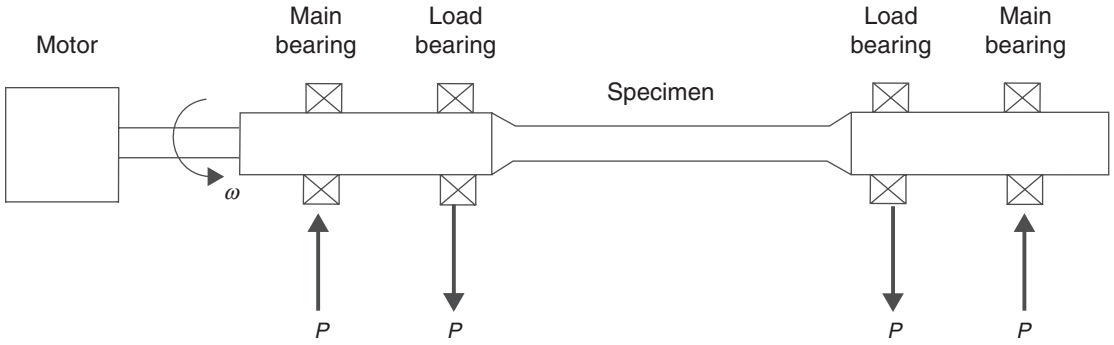

Figure 2. Schematic of a four-point rotating bend test. The main features are: (a) the motor, which gives the rotation; (b) the two main bearings (supports); and (c) the two load bearings (where the static load is applied).

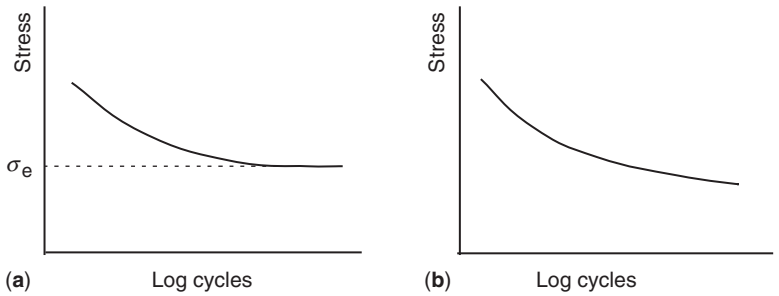

Figure 3. (a) S–N curve with a fatigue (or endurance) limit, σ_e; and (b) S–N curve with no fatigue limit.

Figure 3a and b. These curves are drawn through median points at each stress level. The curves shown in Figure 3a and b represent the two distinct behaviors that are encountered. The curve in Figure 3a decreases asymptotically to a horizontal, constant stress value. Stress amplitudes below this value do not result in failure. Microcracks may develop below this stress level, but growth is arrested. The lower bound stress, σ_e is called the *endurance or fatigue limit*, and these values are often reported in material property tabulations. This limiting stress behavior is most commonly observed for low-strength steels. Most nonferrous metals, however, exhibit curves of the type shown in Figure 3b, that is they do not exhibit a lower bound. As a result, the endurance or fatigue limits that are reported for them are actually the stress level corresponding to a cyclic life of the order of 10^8 cycles.

4 MEAN STRESS EFFECTS: THE GOODMAN RELATIONSHIP

The stress histories for many components do not have a mean stress of zero as the reversed loading, $R = -1$ case. Mean values can be either tensile or compressive. The nonzero mean can be due to either externally applied loading or residual surface layer stresses. The parameters used to describe mean stress effects are defined in terms of the maximum and minimum cyclic stresses. These are the mean stress,

$$\sigma_m = \frac{1}{2}(\sigma_{max} + \sigma_{min}) \tag{3}$$

and the alternating stress,

$$\sigma_a = \frac{1}{2}(\sigma_{max} - \sigma_{min}) \tag{4}$$

These two quantities can be expressed by dimensionless ratios (σ_m/σ_f) where σ_f is the fatigue strength for reversed loading and as (σ_m/σ_y) where σ_y is the yield strength or as (σ_m/σ_u) where σ_u is the ultimate strength. Reference to the values in the denominators provides a basis for developing empirical relations for describing mean stress effects.

A common approach for developing empirical equations is to represent a dependent variable in terms of a power series in an independent variable. To correlate this type of approach with equations that have been widely used we select the ratio (σ_a/σ_f) as the dependent variable and either (σ_m/σ_y) or (σ_m/σ_u) as the independent variable. Thus, we can write, for example,

$$\frac{\sigma_a}{\sigma_f} = C_1 + C_2\left(\frac{\sigma_m}{\sigma_u}\right) + C_3\left(\frac{\sigma_m}{\sigma_u}\right)^2 + \cdots \tag{5}$$

If three terms were to be used, three conditions for values of the ratio (σ_m/σ_u) would be required to evaluate the coefficients. Only two conditions are usually used, however. One procedure, which has been used, is to observe that for $\sigma_m = 0$ then $\sigma_m = \sigma_f$ for failure, hence $C_1 = 1$. Then, keeping only two terms in the series, and observing that for $\sigma_a = 0$, then $\sigma_m = \sigma_u$ for failure, we obtain $C_2 = C_1$. Thus, equation (5) becomes:

$$\frac{\sigma_a}{\sigma_f} + \frac{\sigma_m}{\sigma_u} = 1 \tag{6}$$

This equation, proposed by Goodman (1899), is known as the Goodman equation. This linear relation is shown in Figure 4.

Since test data tend to lie above the straight line, it follows that a nonlinear curve lying above might be preferable. Therefore, proceeding as before and keeping only the C_1 and C_3 terms in (5) results in the following equation, credited to Gerber (1874):

$$\frac{\sigma_a}{\sigma_f} + \left(\frac{\sigma_m}{\sigma_u}\right)^2 = 1 \tag{7}$$

This equation is also shown in Figure 4.

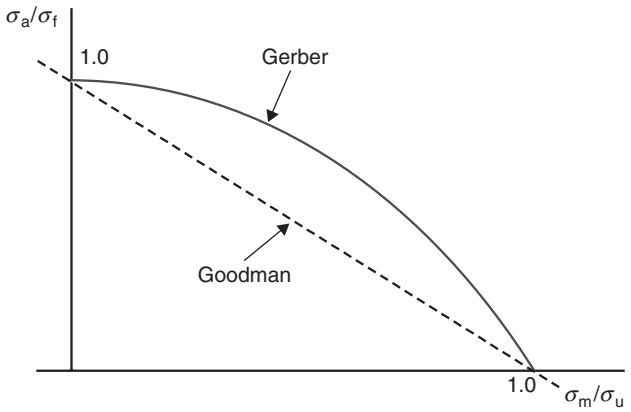

Figure 4. The Goodman diagram (dashed line) and the Gerber diagram (solid line).

5 VARIABLE AMPLITUDE LOADING: THE PALMGREN–MINER RULE

Aerospace structures are subjected to varying amplitudes of loading with time rather than repetitive cyclic loading of constant amplitude. In fact, the loading spectra not only are complex, but differ for different types of aircraft, for example fighter aircraft and transport aircraft.

Proposals by Palmgren (1924) and Miner (1945) attempt to account for variable amplitude loading effects by the use of the assumption that the fatigue damage incurred during each cycle is independent of the prior loading history. The rule based on this assumption is described as the linear damage rule. If the amplitude of loading is σ_i for n_i cycles and the fatigue life at σ_i is N_i, the fraction of life used would then be n_i/N_i. Therefore, for a total of k cycles in a loading spectrum consisting of n_1 cycles at stress level σ_1, n_2 cycles at stress level σ_2, and so on, fatigue failure occurs when

$$\sum_{i=1}^{N} \frac{n_i}{N_i} = 1 \tag{8}$$

Although the linear damage rule is intended to provide a basis for predicting fatigue life under variable amplitude loading, it has deficiencies that should be recognized. Experiments have shown that the sum of equation (8) can, depending upon the order in which load levels are applied, be either greater or less than unity. If, for example, a block of high-level loading is followed by a block of low-level loading, the sum for failure in equation (8) can be less than unity. For notched specimens the reverse, however, is found to occur. If a block order is changed to a low–high sequence, the sum for failure is greater than unity. It has been argued that if the loading spectrum is a mixture of high–low and low–high sequences, the deviations from unity may be canceled. Alternative methods for assessing the accumulation of damage under variable amplitude loading have been proposed and some of these are discussed by Collins (1981). The introduction of material constants and the complexity of applying them to complex spectra, however, often detract from their usefulness.

There are approaches to include other effects in the basic S–N approach, namely *multi-axial stress states* (e.g., combined torsion and bending), effects of *stress concentration locations* such as holes, notches, and fillets and effects of the *environment* (extremes in temperature and active or corrosive atmospheres); these are discussed in detail in Carlson and Kardomateas (1996).

6 FATIGUE DESIGN APPROACHES: SAFE-LIFE, FAIL-SAFE, AND DAMAGE TOLERANCE

During the 1950s, the loss of several Comet aircraft within a short period of time led to an investigation into the structural integrity of Comet's fuselage. A pressurization test on a sample aircraft revealed that the combination of fatigue and stress concentration at the corner of a window was the probable cause of the failures. These accidents as well as others prompted the development of fatigue methods for aircraft design. The first and oldest approach to dealing with fatigue is to require that a structure should be able to survive several times the intended service lifetime, for example the time to failure of an aircraft wing in a laboratory fatigue test may be required to be four times the expected service lifetime. This type of requirement constitutes a design philosophy described as *safe-life*. The safe-life approach determines a replacement time for aircraft components, usually specified as a number of allowable landings or flight hours. Once a component reaches its replacement time, its safe-life is considered to be used up and it is retired, whether or not any fatigue cracks are present. There are, however, two significant problems with this method: (i) the safety of an aircraft is not protected if a manufacturing or accident or maintenance induced defect is introduced and (ii) safe-life safety factors are quite conservative and thus many components can be quite prematurely retired.

The next approach used in aircraft fatigue design is the *fail-safe* approach, which was developed in the 1960s and is based on designing structures with multiple load paths, such that if an individual element fails, the remaining elements can carry the additional load from the failed element until the damage is detected and repaired. Indeed, there are designs in which the failure of a single component results in a total system failure.

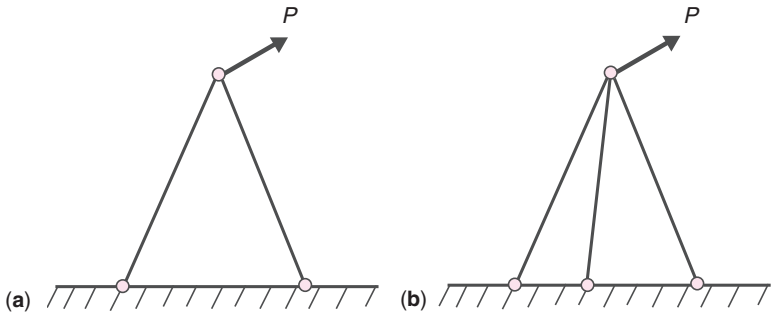

Figure 5. (a) A two-bar truss; and (b) a three-bar truss (fail-safe).

As a simple example, consider a two-bar truss under an external load (Figure 5a). If one bar fails, the remaining bar can no longer serve the intended structural function. By contrast, the failure of one bar in a three-bar truss (Figure 5b) simply results in a redistribution of the bar forces and the truss can still, at least temporarily, serve its intended function of the truss.

The latter example can then be described as being fail-safe. This is obviously a desirable design solution and it is commonly used in aircraft. Ideally, an aircraft designed according to fail-safe principles can sustain damage and remain airworthy until the damage is detected and repaired. This philosophy necessitates periodic inspections to assess the integrity of the load carrying members. However, this philosophy cannot preclude the possibility of simultaneous crack development in multiple load path elements and the inspections in this approach were not based on crack growth principles (fracture mechanics); as a result, the loss of several "fail-safe" aircraft in the mid-1970s emphasized the need to locate cracks and repair them before failure occurred.

The presence of a crack in a component subject to load variations does not necessarily constitute failure. The crack may undergo a time-dependent extension, which is often described as "stable" growth. Eventually, the growing crack may attain what is described as a critical length and then unstable or catastrophic growth can occur. Thus, during a stable growth period, the structural integrity of the system remains intact and the primary concern is the anticipation of when a critical length will be attained. This requires knowledge of the loading history on the cracked component. Prescribed inspection intervals are instituted and fracture mechanics is employed to ensure that a crack would not grow to its critical length within the inspection interval time. Thus, cracks occurring at any time would be caught at the next inspection interval before they have a chance to become critical. Naturally, associated repairs would ensue. Although the task described is complex and a simplified description has been presented, the logic involves evaluating the tolerance of components to the presence of cracks. This is the basis for the evolution of a sophisticated procedure, which is described as *damage tolerance*. In the early 1970s, the US Air Force was the first to adopt the damage tolerance fatigue design approach. With economic and safety advantages over the previous methods, the damage tolerance philosophy was eventually adopted by the commercial aviation.

As mentioned, the objective of the damage tolerance approach is to detect cracks in principal structural elements (PSEs) before their critical length. A PSE is defined as any aircraft structural component carrying flight, ground, or pressurization loads, whose failure could result in the loss of the aircraft. The goal is to establish inspection intervals for these elements based upon the time it takes to grow a crack from an initial detectable size to the critical crack length.

The first task in the aircraft damage tolerance approach is to define the usage profile. This profile describes the various flight conditions, such as taxi, climb, cruise, descent, and landing impact, and the amount of time spent at each gross weight, speed, and altitude. The usage profile is then used to create a load factor spectrum at the center of gravity of the aircraft. The next task is to identify the PSEs, convert the load factor spectrum into a stress spectrum for each location, and incorporate the effects of the service environment. Using crack growth (da/dN) equations, such as the Paris or Forman equations (described next), the stress spectrum is combined with material properties data and stress intensity factor solutions applicable to each PSE to determine the number of cycles for a crack to reach the critical length starting from the detectable length. This number is usually divided by a factor of two to arrive at the inspection interval. This ensures that, should a PSE develop a crack, it will be inspected at least once before the crack propagates to failure.

Unlike the safe-life approach where components are retired whether or not they are damaged, in the damage tolerance approach components are only replaced if a crack is found during an inspection. It is important to note that any size crack found during an inspection mandates replacement

of the damaged component (even if the crack is at the just detectable size and therefore the component could last until the next inspection). Another advantage of the damage tolerance approach is that crack growth is rather deterministic unlike the large scatter associated with the S–N methodology. Thus, it allows for a reduction of safety factors in design.

7 STRESS-BASED ANALYSIS OF FRACTURE

As indicated earlier, the adoption of fracture mechanics concepts in the fatigue analysis of structural components goes back to the early 1960s with the pioneering work of Paris (1960), who relied on advances made a few years earlier in the description of the stress field in the vicinity of crack tips (Williams, 1957). The presence of singularities in the linearly elastic solution of the stress field in the vicinity of sharp crack tips and notches had been observed based on Inglis' (1913) classical solution of an elliptical hole in a large plate subjected to a remote tensile traction σ_∞ (Figure 6). Inglis showed that the stress concentration $\sigma_{yy}/\sigma_\infty$ at $x = \pm a$, $y = 0$ is given by $1+2a/b$, where a and b, respectively, denote the axes of the ellipse that are perpendicular and parallel to the loading axis. In the limiting case of a very sharp crack (for which $b/a \to 0$),

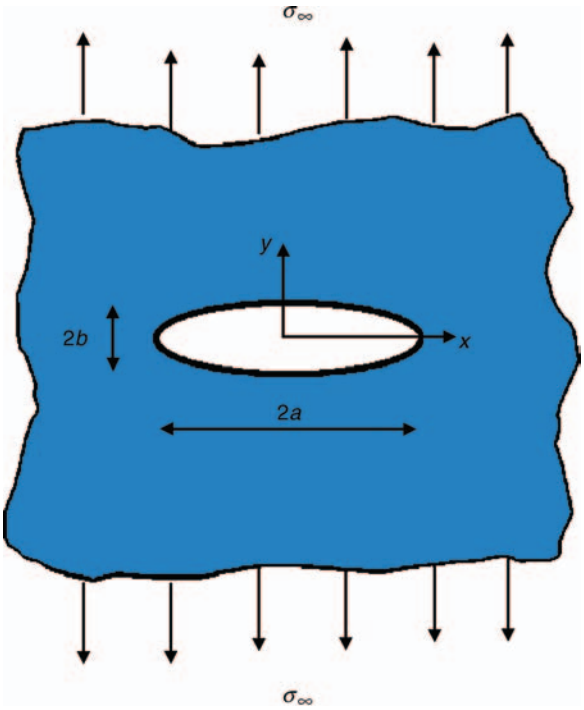

Figure 6. Elliptical hole in a plate subjected to remote tensile loading σ_∞.

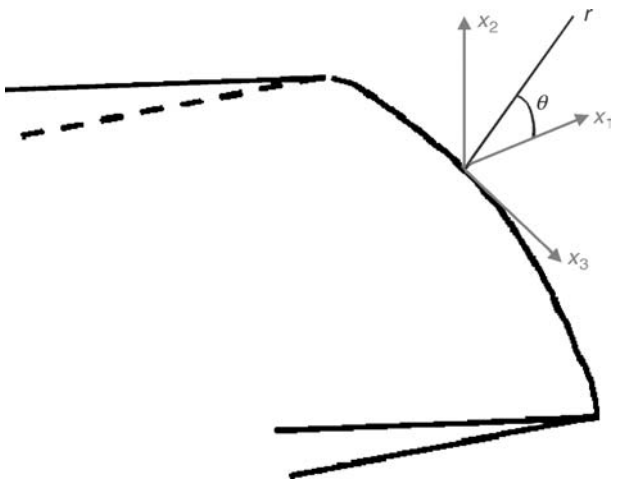

Figure 7. Coordinate system for near-tip solution, with $r = \sqrt{x_1^2 + x_2^2}$ and $\theta = \tan^{-1}(x_2/x_1)$.

the predicted stress concentration tends to infinity suggesting the existence of a stress singularity in the vicinity of the crack tips.

Using an asymptotic method, the nature of that singularity was analyzed in 1958 by Williams, who showed that the stress field σ_{ij} in the vicinity of any sharp crack in a linearly elastic material takes the form

$$\sigma_{ij}(r,\theta) = \frac{K_{\mathrm{I}}}{\sqrt{2\pi r}} g_{ij}^{\mathrm{I}}(\theta) + \frac{K_{\mathrm{II}}}{\sqrt{2\pi r}} g_{ij}^{\mathrm{II}}(\theta) + \frac{K_{\mathrm{III}}}{\sqrt{2\pi r}} g_{ij}^{\mathrm{III}}(\theta) + O(r^0) \tag{9}$$

where the Cartesian (x_1, x_2, x_3) and cylindrical (r, θ, x_3) coordinate systems attached to the crack front are defined in Figure 7. The Roman subscripts and superscripts I, II, and III refer to the three basic failure modes (referred to as the I, II, and III *fracture modes*), which are shown schematically in Figure 8 and, respectively, denote the in-plane opening, in-plane shear, and out-of-plane failure modes. In the linearly elastic case, the stress field in the vicinity of the crack front can be written as a combination of these three failure modes. The angular functions g_{ij}^k entering (9) are trigonometric functions of the angular coordinate θ. For example, for the opening fracture mode ($k = \mathrm{I}$),

$$g_{11}^{\mathrm{I}}(\theta) = \cos\frac{\theta}{2}\left(1 - \sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right)$$
$$g_{22}^{\mathrm{I}}(\theta) = \cos\frac{\theta}{2}\left(1 + \sin\frac{\theta}{2}\sin\frac{3\theta}{2}\right) \tag{10}$$

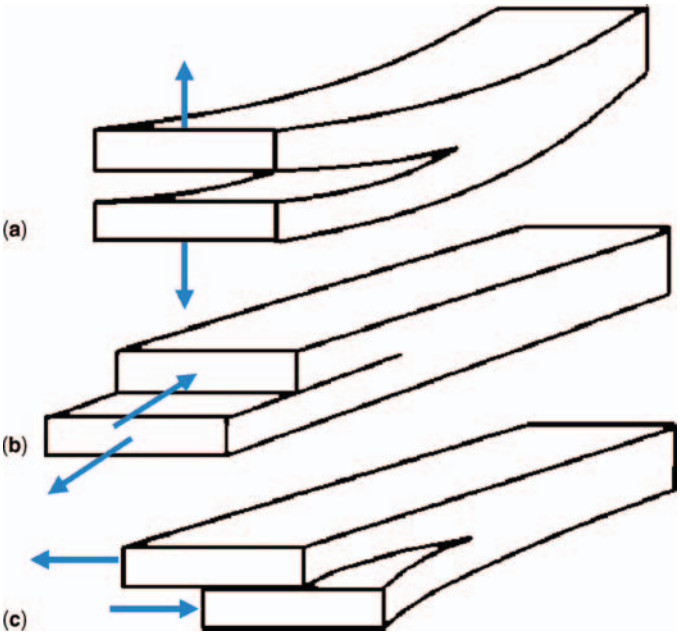

Figure 8. Fracture modes: (a) mode I or opening mode; (b) mode II or in-plane shear mode; and (c) mode III or out-of-plane shear mode.

$$g_{12}^{\mathrm{I}}(\theta) = \cos\frac{\theta}{2}\sin\frac{\theta}{2}\cos\frac{3\theta}{2} \quad g_{13}^{\mathrm{I}}(\theta) = g_{23}^{\mathrm{I}}(\theta) = 0 \qquad (11)$$

The angular functions associated with the other modes, and the expression of the near-tip strain and displacement fields, can be found in all reference books on fracture mechanics (see, e.g., Kanninen, 1985; Suresh, 1991; Anderson, 1995). Due to the canonical nature of the near-tip solution, the three scalar parameters K_{I}, K_{II}, and K_{III} in equation (9) define the stress, strain, and displacement fields in the vicinity of the crack front. These parameters, referred to as the *stress intensity factors*, incorporate the effects of the loading and problem geometry.

Due to the importance of the stress intensity factors in the solution of fracture problems, a wide range of analytical, numerical, and experimental methods have been proposed to solve for K_{I}, K_{II}, and K_{III}, and compendia of stress intensity factors have been created for a large number of geometries and loading conditions (see, e.g., Sih, 1973; Rooke, 1986; Wu and Carlsson, 1991).

Furthermore, since the stress intensity factors uniquely define the near-tip stress field, failure criteria based on critical values of these parameters have been proposed to predict the onset of crack propagation in a material, in the form

$$K_{\mathrm{I}}\,(\text{loading, geometry}) = K_{\mathrm{Ic}} \qquad (12)$$

where K_{Ic} is a material parameter (with units of $\mathrm{Pa\,m^{\frac{1}{2}}}$) referred to as the *fracture toughness*. Its value ranges from about 0.5 $\mathrm{MPa\,m^{\frac{1}{2}}}$ for brittle polymers to more than 300 $\mathrm{MPa\,m^{\frac{1}{2}}}$ for very ductile metals.

It should be noted that, in the case of time-dependent loading of a stationary (or slowly growing) crack, such as in cyclic and impact loading situations, the near-tip stress fields maintain the asymptotic form described by equation (9), except that the stress intensity factors are then time dependent. The "strength" of the stress singularity, that is the value of the exponent entering the radial variation of the singular term in the near-tip stress field expansion ($-1/2$ in (9)), changes in the case of notches and corners. The nature of the stress singularity also changes for rapidly propagating cracks (i.e., for crack speeds representing a substantial fraction of the wave propagation speeds in the material) and for interfacial cracks, but the existence of near-tip singular stress field defined by stress intensity factors remains.

Finally, the mode mixity present in the vicinity of the crack front, that is the relative importance of the mode I, II, and III stress intensity factors, has been shown to play a key role in the propagation behavior of cracks. For an isotropic material under in-plane (mode I/II) loading, cracks have been shown to propagate primarily in the symmetric (mode I) case and various criteria for crack path predictions have been proposed, including those based on symmetry ($K_{\mathrm{II}} = 0$) and maximum hoop stress (Anderson, 1995).

8 ENERGY-BASED ANALYSIS OF FRACTURE

Another way to approach fracture problems, which circumvents the apparent inconsistency of the existence of a singular solution in the vicinity of the crack tip, is based on the concept of *energy release rate* G. This concept, which was adopted by Griffith (1921) in his pioneering study of the fracture of brittle materials, quantifies the decrease in total potential energy Π associated with an infinitesimal propagation of the crack, that is

$$G = -\frac{\partial \Pi}{\partial A} \qquad (13)$$

where A denotes the crack area. For a plate of width b, $\mathrm{d}A = b\ \mathrm{d}a$, where a denotes the crack length. The energy release rate, which has units of $\mathrm{J/m^2}$, thus represents the energy available for propagation of the crack. The energy-based criterion for crack initiation thus takes the form

$$G = G_{\mathrm{c}} \qquad (14)$$

where G_c denotes the fracture toughness of the material, that is the energy required to create fracture surfaces.

The equivalence between stress- and energy-based approaches to fracture in the linear elastic case can be demonstrated by computing the change in strain energy (or crack closure work) associated with an infinitesimal advance of the crack tip, leading to the following relation between the energy release rate G and the stress intensity factors K_k ($k = $ I, II, and III):

$$G = \frac{K_{\mathrm{I}}^2 + K_{\mathrm{II}}^2}{\bar{E}} + \frac{1+\nu}{E} K_{\mathrm{III}}^2 \tag{15}$$

where $\bar{E} = E$ for plane stress and $\bar{E} = E/(1-\nu^2)$ for plane strain, and ν is the Poisson's ratio.

A key contribution to the energy approach of fracture is the J-integral introduced by Rice (1968) and defined by

$$J = \int_{\Gamma} \left(W n_1 - \boldsymbol{T} \cdot \frac{\partial \boldsymbol{u}}{\partial x_1} \right) \mathrm{d}\Gamma \tag{16}$$

where Γ denotes any contour that encircles the crack tip, W is the strain energy density that relates stresses and strains through $\sigma_{ij} = \partial W / \partial \varepsilon_{ij}$, n is the unit outward normal to the contour Γ, $\mathbf{u}$ is the displacement vector, and $\mathbf{T}$ is the traction vector. For linearly (and nonlinearly) elastic materials, the J-integral is path independent, that is it does not depend on the contour Γ used to define it. This key property allows to relate near-tip parameters (such as the stress intensity factors) to the far-field loading quantities. Furthermore, Rice showed that, for a (non)linearly elastic solid, J physically corresponds to the rate of change of potential energy with respect to crack advance, that is to the energy that flows into the crack tip region. For a linearly elastic material, it thus reduces to the strain energy release rate G.

To conclude this section on the energy approach of fracture, let us mention that energy-based criteria, such as the maximum energy release rate criterion, have also been proposed to predict the crack path, and have yielded predictions very similar to those associated with stress-based criteria.

9 NONLINEAR FRACTURE MECHANICS

The singular nature of the stress and strain "K-fields" in the vicinity of the crack tip is in contradiction with the linear kinematic assumption underlying the theory of elasticity used to derive these near-tip fields. Addressing this apparent inconsistency and/or providing a more realistic description of the nonlinear processes taking place in the immediate vicinity of the crack front are the focus of nonlinear fracture mechanics. It should be pointed out, however, that despite the apparent inconsistency of the linearly elastic solution, the stress intensity factors introduced play a key role in many fracture problems since the size of the nonlinear region (where large deformation and/or material nonlinearity effects must be accounted for) is often very small. In these cases, referred to as *small-scale yielding*, the K-field solution described by (9) remains dominant outside that small nonlinear zone and therefore continues to characterize the near-tip conditions.

Over the years, various estimates have been proposed for the size of the nonlinear zone. For the monotonic loading case, an early estimate of the plastic zone size r_p was obtained by Irwin (1960) by comparing the stress field ahead of the crack ($\theta = 0$) to the yield stress σ_y of the material, leading in the mode I case to

$$r_p = \alpha \left(\frac{K_{\mathrm{I}}}{\sigma_y} \right)^2 \tag{17}$$

with $\alpha = 1/3\pi$ for plane strain and $\alpha = 1/\pi$ for plane stress. Extending this approach to the "complete" asymptotic solution (9), estimates of the size and shape of the plastic zone can be obtained, as shown in Figure 9 in the plane stress and plane strain cases. Other more complex estimates of the plastic zone size that account for the redistribution of stresses in the nonlinear zone have also been introduced.

Beyond estimating the size and shape of the nonlinear zone, the focus of the nonlinear analysis of fracture

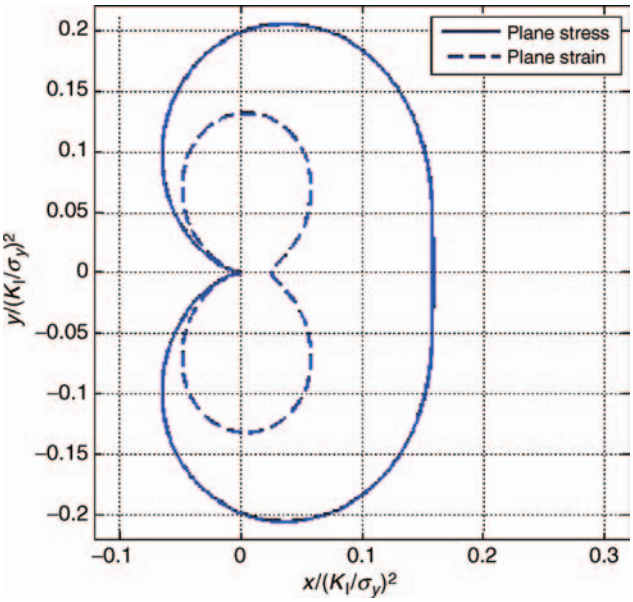

Figure 9. Plane stress and plane strain estimates of the shape and size of the plastic zone ($\nu = 0.3$). The crack is located along the negative x-axis, with the crack tip at the origin.

is the study of the nonlinear effects on the near-tip fields. One of the most successful studies in that regard led to the Hutchinson–Rice–Rosengren (HRR) asymptotic fields (Hutchinson, 1968; Rice and Rosengren, 1968), which are based on the small strain J_2-deformation theory of plasticity (i.e., monotonic loading). In particular, it relies on the classical Ramberg–Osgood power law relationship between stresses and strains:

$$\frac{\varepsilon_{ij}}{\varepsilon_y} = \frac{3\alpha}{2}\left(\frac{\sigma_e}{\sigma_y}\right)^{n-1}\frac{s_{ij}}{\sigma_y} \tag{18}$$

where E, α, n and σ_y are material parameters, $\varepsilon_y = \sigma_y/E$, s_{ij} denotes the deviatoric stresses, and $\sigma_e = \sqrt{3s_{ij}s_{ij}/2}$ is the effective stress. The resulting asymptotic stress and strain fields take the form

$$\frac{\sigma_{ij}}{\sigma_y} = \left(\frac{J}{\alpha\sigma_y\varepsilon_y I_n r}\right)^{1/(n+1)}\tilde{\sigma}_{ij}(\theta, n) \tag{19}$$

$$\frac{\varepsilon_{ij}}{\alpha\varepsilon_y} = \left(\frac{J}{\alpha\sigma_y\varepsilon_y I_n r}\right)^{1/(n+1)}\tilde{\varepsilon}_{ij}(\theta, n) \tag{20}$$

where J is the value of the J-integral, I_n are explicit functions of the exponent n, while $\tilde{\sigma}_{ij}$ and $\tilde{\varepsilon}_{ij}$ are angular functions. As expected, the near-tip HRR solution equations (19 and (20) recovers the inverse square root singularity of the linearly elastic solution (9) when $n = 1$.

Another approach to incorporate nonlinearity in the near-tip solution is the cohesive model introduced by Dugdale (1960) and Barenblatt (1962), in which all nonlinearities are concentrated to a vanishingly thin region (called the *cohesive zone*) ahead of the crack (Figure 10). In that region, the material failure response is described by a nonlinear relation between the cohesive traction T resisting the crack opening and the displacement jump Δ ahead of the crack tip. The area under the cohesive traction–separation curve corresponds to the energy required for the formation of a crack, that is the fracture toughness G_c. In addition to providing a natural bound for the near-tip stresses (which cannot exceed the maximum value of T) and thereby eliminate issues associated with the existence of a stress singularity, the cohesive modeling of fracture is the theoretical foundation of a numerical scheme (referred to as the cohesive finite element method), which has shown over the past two decades great success in the modeling of a wide variety of quasi-static, dynamic, and fatigue failure problems (Needleman, 1987; Camacho and Ortiz, 1996; Maiti and Geubelle, 2005).

10 FATIGUE CRACK GROWTH (THE PARIS LAW)

Most fatigue problems involve relatively low load levels for which the small scale yielding assumptions (i.e., for which the stress field outside the small nonlinear zone can be described by the asymptotic solution (9)) is applicable. Based on the problem geometry, one can therefore associate the cyclic load applied on the specimen to a range ΔK of the (mode I) stress intensity factor. In his pioneering study, Paris (1960) showed that, when displayed on a log–log plot, the relation between the crack advance per cycle, da/dN, and ΔK looks similar to the curve shown schematically in Figure 11. This curve, usually referred to as the Paris curve, shows three distinct crack propagation regimes. In the steady-state regime (region II), the relation between da/dN and ΔK is a power law:

$$\frac{da}{dN} = C(\Delta K)^m \tag{21}$$

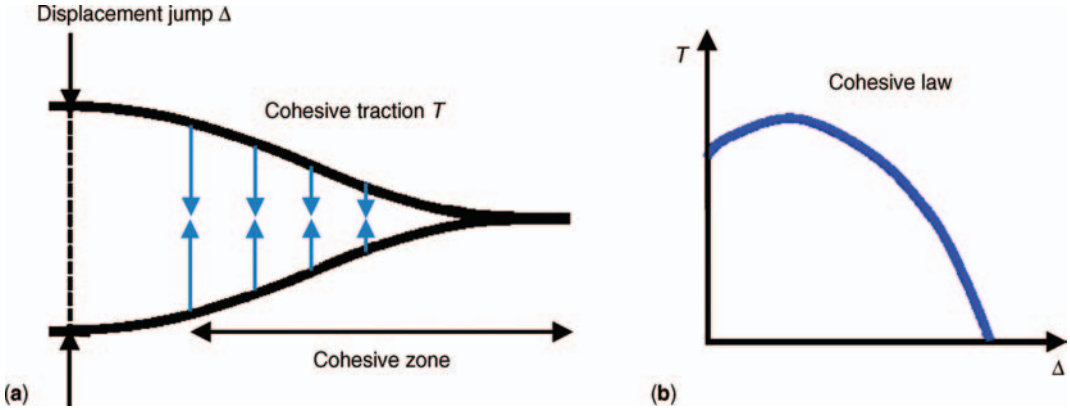

Figure 10. Cohesive zone modeling of a crack: (a) schematic of crack tip region; and (b) cohesive failure law.

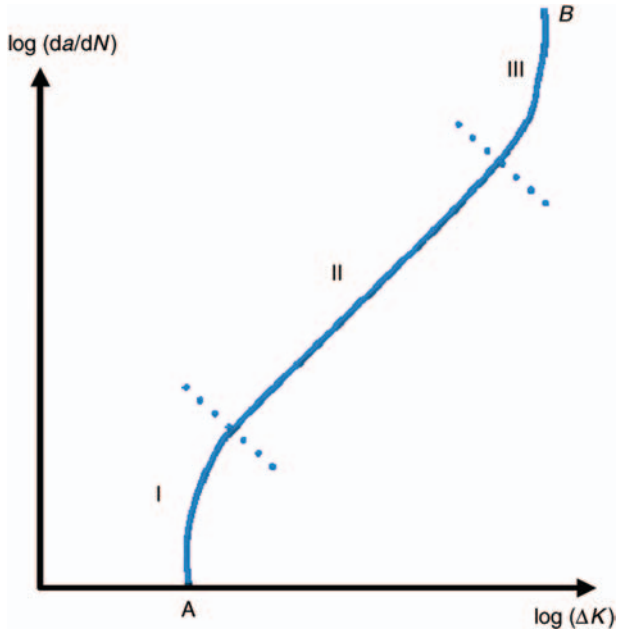

Figure 11. Typical Paris curve between the crack advance per cycle da/dN and the amplitude of the cyclic load (quantified through the associated range of stress intensity factor ΔK), showing three distinct regimes of crack propagation.

where the exponent m (i.e., the slope of the curve on the log–log scale) is of the order of 3 in many metals and is substantially higher (of the order of 10) for polymers.

At lower loading amplitudes (region I), the fatigue response tends to a threshold value A below which no crack propagation is observed. In the upper range of loading amplitudes (region III), the crack propagation rate increases rapidly as the load level approaches that associated with crack propagation under monotonic loading (Point B in Figure 11). As described earlier, a detailed knowledge of the location and size of the cracks in the cyclically loaded structure combined with the experimentally extracted parameters describing the fatigue response (19) of the material allows for the prediction of the expected fatigue life of the structure.

11 SUMMARY AND PERSPECTIVE

This section outlined the fundamental knowledge in aerospace fracture and fatigue. Although the basic concepts are universally valid, current applications are mostly on metallic components (aluminum, titanium, steel alloys, etc.). The prediction of fatigue and failure of advanced materials (composites, sandwich structures, ceramics, etc.) requires additional considerations due to the complex nature of their microstructure and it is a timely and ongoing research effort. Another key challenge is the accurate numerical modeling of fracture in complex materials and/or structures under complex cyclic loading conditions including overloads, especially in three dimensions.

Finally, an important topic being currently researched is that of "small fatigue cracks" and "below threshold" fatigue crack growth. Indeed, the fatigue crack growth in the so-called "Paris regime" is well studied and extensively applied in aerospace damage tolerance. In this regime, cracks do not grow below a "threshold" value of ΔK. It has been well established, however, that cracks do indeed grow below the threshold, and a considerable part of the lifetime can be spent in the "below threshold" regime. The growth below threshold is influenced by the alloy microstructure (grain size) and is characterized by increased scatter. There is a need for physically based procedures for a smooth transition from the "Paris regime" to the "below threshold regime," which can be practical and ensure proper damage tolerance procedures for aerospace components.

REFERENCES

Anderson, T.L. (1995) *Fracture Mechanics: Fundamentals and Applications*, CRC Press.

Barenblatt, G.I. (1962) The mathematical theory of equilibrium cracks in brittle fracture. *Adv. Appl. Mech.*, **7**, 55–129.

Camacho, G.T. and Ortiz, M. (1996) Computational modelling of impact damage in brittle materials. *Int. J. Solids and Struct.*, **33**, 2899–2938.

Carlson, R.L. and Kardomateas, G.A. (1996) *An Introduction to Fatigue in Metals and Composites*, Chapman & Hall, London.

Collins, J.A. (1981) *Failure of Materials in Mechanical Design*, Wiley-Interscience, New York.

Dugdale, D.S. (1960) Yielding of steel sheets containing slits. *J. Mech. Phys. Solids*, **8**, 100–108.

Fuchs, H.O. and Stephens, R.I. (1980) *Metal Fatigue in Engineering*, John Wiley & Sons, New York.

Gerber, H. (1874) Bestimmung der zulassigen Spannungen in Eisenkonstructionen. *Z. Bayerischen Architeckten and Ing.-Ver.*, **6**, 101–110.

Goodman, J. (1899) *Mechanics Applied to Engineering*, Longmans-Green, London.

Griffith, A.A. (1921) The phenomena of rupture and flow in solids. *Philos. Trans. Royal Soc., London*, **A221**, 163–197.

Hutchinson, J.W. (1968) Singular behavior at the end of a tensile crack in a hardening material. *J. Mech. Phys. Solids*, **16**, 13–31.

Inglis, C.E. (1913) Stresses in a plate due to the presence of cracks and sharp corners. *Trans. Inst. Naval Arch.*, **55**, 219–241.

Irwin, G.R. (1960) Plastic zone near a crack and fracture toughness. *Proceedings of the Seventh Sagamore Ordnance Materials Conference*, August 16–19 1960, vol. IV, New York. Syracuse University Research Institute, NY, pp. 63–78.

Kanninen, M.F. (1985) *Advanced Fracture Mechanics*, Oxford University Press.

Maiti, S. and Geubelle, P.H. (2005) A cohesive model for fatigue failure of polymers. *Eng. Fract. Mech.*, **72**, 691–708.

Miner, M.A., (1945) Cumulative damage in fatigue. *J. Appl. Mech.*, **67**, 159–164.

Moore, H.F. and Kommers, J.B. (1927) *The Fatigue of Metals*, McGraw-Hill, New York.

Needleman, A. (1987) A continuum model for void nucleation by inclusion debonding. *J. Appl. Mech.*, **54**, 525–531.

Palmgren, A. (1924) Die Lebensdauer von Kugellagern, 2. *Ver. Dtsch. Inge.*, **68**, 339–347.

Paris, P.C. (1960) The Growth of Cracks due to Variations in Loads. PhD thesis. Lehigh University, Bethlehem, PA.

Rice, J.R. (1968) A path independent integral and the approximate analysis of strain concentration by notches and cracks. *J. Appl. Mech.*, **35**, 379–386.

Rice, J.R. and Rosengren, G.F. (1968) Plane strain deformation near a crack tip in a power law hardening material. *J. Mech. Phys. Solids*, **16**, 1–12.

Rooke, D.P. (1986) *Compounding Stress Intensity Factors: Applications to Engineering Structures*, Parthenon Press.

Sih, G.C. (1973) *Handbook of Stress Intensity Factors*, Institute of Fracture and Solid Mechanics, Lehigh University, Bethlehem.

Suresh, S. (1991) *Fatigue of Materials*, *Cambridge Solid State Science Series*, Cambridge University Press.

Timoshenko, S.P. (1953) *History of Strength of Materials*, McGraw-Hill, New York.

Williams, M.L. (1957) On the stress distribution at the base of a stationary crack. *J. Appl. Mech.*, **24**, 109–114.

Wu, X.-R. and Carlsson, A.J. (1991) *Weight Functions and Stress Intensity Factor Solutions*, Pergamon Press.