Methodologies for Predicting Fatigue Life

R.K. Holman and P.K. Liaw

This article reviews the basic techniques employed in fatigue life prediction. The stress-life, local-strain, and fracture-mechanics methods as applied to life prediction under constant amplitude loading and variable amplitude loading are discussed. Life prediction methodology under variable amplitude loading is also discussed, with particular emphasis on the linear-damage accumulation approach, or Miner's rule. Finally, a discussion of various cycle-counting techniques for variable amplitude loading is given.

INTRODUCTION

The problem of fatigue is one of the most important considerations in design. Among the factors that can affect the fatigue life of a component are mean stresses, geometry/size effects and surface conditions, residual stresses, load interaction effects (sequence effects), environments, and others.

The vast majority of fatigue data available is for constant amplitude (CA) loading. While this type of loading is not generally representative of real-world conditions, the relative ease of testing and analysis is attractive. The stress-life method was developed first and is by far the easiest to apply. However, its limitations make it less widely applicable than the other methods. The local-strain technique is slightly more complex, but it is applicable to a much wider range of conditions and can account for several of the factors that stress-life cannot. The fracture-mechanics approach is relatively new compared to the others, but its application has become widespread. While the stress-life and local-strain approaches determine the total life (including both crack initiation and crack propagation), fracture mechanics gives estimates of crack propagation life, making it a damage-tolerant method (i.e., a cracked member's remaining life can be evaluated).

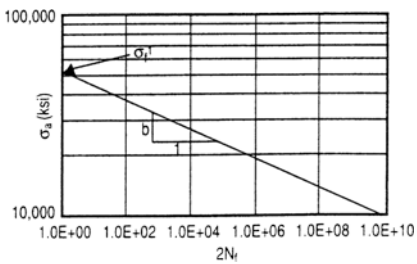

Figure 1. A stress-life curve illustrating the determination of relevant constants.

Fatigue life analyses for variable amplitude (VA) loading estimate the life under conditions more closely approximating those experienced by a component under actual service conditions. A number of factors affect the life under VA loading which are not accounted for in CA-based analyses.

CONSTANT AMPLITUDE LOADING

Stress-Life Approach

Introduced by A. Wöhler in 1860, life estimates implementing the stress-life method include both the cycles to crack initiation and the cycles to propagate the crack to component failure.[1] Stress life is a simple technique wherein data from CA (generally fully reversed) stress-controlled tests are plotted as the nominal stress amplitude σ_a (S_a) versus the number of cycles to cause failure N_f at that stress level. The method assumes the resulting S-N curve, when plotted on a log-log scale, will result in a straight line that can be represented by[1-4]

$$\sigma_a = \frac{\Delta\sigma}{2} = \sigma_f'(2N_f)^b \qquad (1)$$

where σ_f' is the fatigue-strength coefficient (which is generally approximately equal to the true fracture strength for metals), and b is the fatigue-strength exponent. Values of these constants can be found for many materials in the literature or can be determined from a plot of experimental S-N data (Figure 1). This curve should intercept the y-axis at the ultimate tensile strength of the material.

The idea of an endurance limit came from plots of stress amplitude versus fatigue cycle life. Some materials (e.g., many steels) demonstrate a cycle behavior wherein, below a certain stress amplitude, failure will not occur for any number of cycles, giving infinite life. This results in a plateau in the stress-life curve, and this limiting stress amplitude is known as the endurance limit. Many materials, however, exhibit no such limit (e.g., aluminum alloys). In these cases, the endurance limit is generally defined as the stress amplitude to cause failure at 10^7 cycles.

A number of factors can affect the S-N curve, and ideally it would be desirable to have data generated under conditions approximating those of the actual component in question. Among these factors are the material and its processing, mean stress, residual stress, loading frequency, geometry, surface finish, temperature, and environment.[1,3] In the presence of a mean stress (σ_m), an approximation can be made by converting the stress amplitude to an equivalent completely reversed stress amplitude that would result in the same cycles to failure. Three commonly used relations for making this conversion in stress-life calculations are (Figure 2):[5]

the Goodman relation

$$\sigma_a = \sigma_e\left[1-\left(\frac{\sigma_m}{\sigma_{UTS}}\right)\right] \qquad (2)$$

the Gerber relation

$$\sigma_a = \sigma_e\left[1-\left(\frac{\sigma_m}{\sigma_{UTS}}\right)^2\right] \qquad (3)$$

and the Soderberg relation

$$\sigma_a = \sigma_e\left[1-\left(\frac{\sigma_m}{\sigma_{YS}}\right)\right] \qquad (4)$$

where σ_e is the endurance limit of the material, σ_{UTS} is the ultimate tensile strength of the material, and σ_{YS} is the yield strength of the material. The Soderberg line is generally quite conservative. The Goodman line is fairly accurate for brittle materials and conservative for ductile materials; and the Gerber parabola generally describes the behavior of ductile materials well.[1]

The effect of notches or stress raisers can also be included in an approximate matter. A theoretical elastic stress concentration factor (K_t) at the notch can be determined for the sample geometry and loading condition, but the so-determined

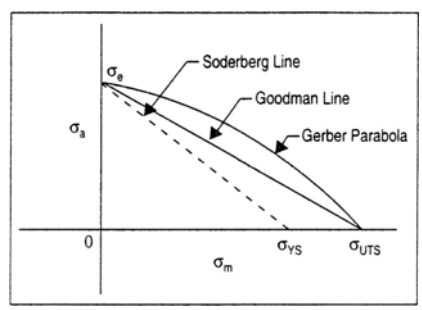

Figure 2. A constant life diagram.[5]

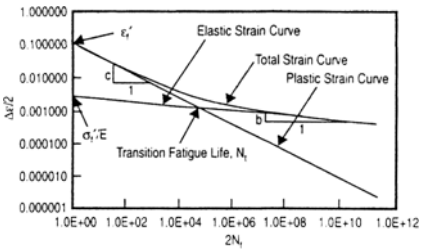

Figure 3. Typical strain-life curves and the relevant constants for steel.

stress usually leads to overly conservative life estimates; the actual stress concentration effect is somewhat less than the theoretical value. Instead, the effect of a notch in stress-life calculations is generally approximated by an experimentally determined correction known as the fatigue notch factor K_f given by[1,4]

$$K_f = \frac{\text{unnotched bar endurance limit}}{\text{notched bar endurance limit}} \quad (5)$$

This is sometimes related to the theoretical stress concentration factor by the notch sensitivity index, q[1,4]

$$q = \frac{K_f - 1}{K_t - 1} \quad (6)$$

A q of zero means the notch has no effect, while a q of one results from $K_f = K_t$. The correction factor K_f is generally applied[1] by either adjusting the entire curve (all $\Delta\sigma$) by K_f (possibly overly conservative) or adjusting the endurance limit by K_f. The fatigue-notch factor is necessary because of the presence of plasticity at the notch tip. However, because the stress-life analysis is based on an elastic state of stress, the occurrence of extensive plastic deformation ahead of the notch may preclude the use of the stress-life method. This also implies that in low-cycle situations (where the plastic strain is dominant), the stress-life approach may not be advisable.

Local-Strain Approach

The local-strain approach, while somewhat more complex than the stress-life approach, is much more versatile and can often give a more accurate life prediction in all but the most simple loading conditions. Like the stress-life method, the local-strain method gives a prediction of the total life (i.e., the life for crack initiation and propagation together). To apply the local-strain method, one must know the cyclic stress-strain relationship [$\varepsilon = f(\sigma)$], the strain-life (ε-N) curve, and (as the name implies) the local state of strain and stress at a notch or stress raiser. Because the life prediction is based on these local values, life is determined where failure would be expected to occur (i.e., at a notch/stress raiser). The (local) strain-life curve is described by[3,4]

$$\varepsilon_a = \frac{\sigma_f'}{E}(2N_f)^b + \varepsilon_f'(2N_f)^c \quad (7)$$

where ε_a is the total strain amplitude ($\varepsilon_{plastic} + \varepsilon_{elastic}$), E is the elastic modulus, ε_f' is the fatigue-ductility coefficient, and c is the fatigue-ductility exponent. These values are tabulated for a number of materials in the literature; they can also be determined experimentally from a stress-life plot or elastic strain-life plot, and a plastic strain-life plot, as shown in Figures 1 and 3.[4]

Equation 7 contains a term for the plastic strain contribution, $\varepsilon_f'(2N_f)^c$, and the elastic contribution,

$$\frac{\sigma_f'}{E}(2N_f)^b$$

Note that the elastic portion

$$\varepsilon_a = \frac{\sigma_f'}{E}(2N_f)^b$$

is equivalent to the stress-life equation $\sigma_a = \sigma_f'(2N_f)^b$ (since $\sigma_a = \varepsilon_a E$). For short fatigue life, the plastic term is dominant, while at a long fatigue life, the elastic portion is the dominant term (Figure 3). The result is that the local-strain method is useful for both low-cycle and high-cycle fatigue problems.[1,3] The boundary between these two regions is the intersection of the plastic and elastic curves, known as the transition fatigue life, and is given by Equation 8.[3]

$$N_t = \frac{1}{2}\left(\frac{\sigma_f'}{\varepsilon_f' E}\right)^{\frac{1}{c-b}} \quad (8)$$

As stated, the local stress-strain state at the notch/stress raiser must be known. These values can be approximated fairly easily through the use of Neuber's rule and the cyclic stress-strain curve. When a notched member is loaded such that the resulting strains are elastic, the local stress and strain are each raised by the same ratio (i.e., K_t, the stress concentration factor, or K_f). However, once plastic strain begins, the local strain rises above that given by K_t, and the local stress falls below it (Figure 4). Thus, there is a local stress concentration factor $K_\sigma < K_t$ and a local strain concentration factor $K_\varepsilon > K_t$. According to Neuber's rule, the geometric mean of these two values remains constant at K_t[6]

$$K_t = \sqrt{(K_\sigma K_\varepsilon)} \quad (9)$$

where

$$K_\sigma = \frac{\sigma}{s} = \frac{\Delta\sigma}{\Delta S} \quad (10)$$

$$K_\varepsilon = \frac{\varepsilon}{e} = \frac{\Delta\varepsilon}{\Delta e} \quad (11)$$

Since K_f is less than K_t, substituting K_t into Equation 9 will result in less conservative values for local stress and strain. Making this substitution, as well as substituting Equations 10 and 11 into Equation 6, yields[4,7]

$$K_t = \sqrt{\frac{\Delta\sigma}{\Delta S}\frac{\Delta\varepsilon}{\Delta e}} \quad (12)$$

Assuming fully plastic yielding does not occur, Δe may be replaced with its elastic value $\Delta S/E$, which after some manipulation gives[1,3,4]

$$\Delta\sigma\Delta\varepsilon = \frac{(K_f \Delta S)^2}{E} \quad (13)$$

In this equation, K_f and E are constants for a given material and geometry. ΔS is known for the loading conditions, so that the right side of Equation 13 is a constant $\Delta\sigma\Delta\varepsilon = C$. This is the equation of a hyperbola; to obtain its solution, it may be plotted on the cyclic stress-strain curve. The intersection of the curve with the hysteresis loop of the load cycle gives the local values of stress and strain in the notch. Life can then be determined from the strain-life relations.

The local-strain approach should also account for the effects of a non-zero mean stress. Commonly, an equivalent completely reversed stress amplitude σ_r that would give the same number of cycles to failure as the mean stress loading is estimated. Two popular methods of calculating this equivalent stress amplitude are the Morrow parameter and the Smith-Watson-Topper parameter.

The Morrow parameter for an equivalent completely reversed stress amplitude is given by[2,3,8]

$$\sigma_r = \frac{\sigma_a}{1 - \frac{\sigma_m}{\sigma_f'}} \quad (14)$$

From the analysis given in Reference 3, this yields the following modified strain-life curve

$$\varepsilon_a = \frac{\sigma_f'}{E}\left(1 - \frac{\sigma_m}{\sigma_f'}\right)(2N_f)^b + \varepsilon_f'\left[\left(1 - \frac{\sigma_m}{\sigma_f'}\right)^{c/b}(2N_f)^c\right] \quad (15)$$

Thus, with a given mean stress and load cycle, life can be estimated either numerically or graphically. The equivalent fully reversed stress amplitude, as expressed by the Smith-Watson-Topper parameter, is given by[2,3]

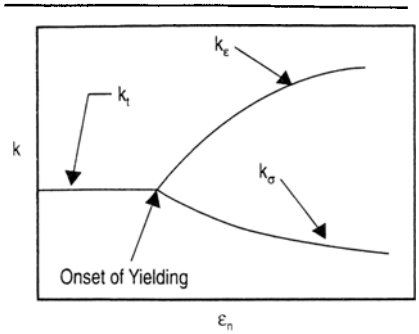

Figure 4. A plot of stress concentration factor versus notch strain demonstrating Neuber's rule.[3]

$$\sigma_f = \left[(\sigma_{max})\varepsilon_a E\right]^{\frac{1}{2}} \quad (16)$$

Again, from an analysis given in References 3 and 9

$$\sigma_{max}\varepsilon_a = \frac{(\sigma_f')^2}{E}(2N_f)^{2b} + \sigma_f'\varepsilon_f'(2N_f)^{b+c} \quad (17)$$

A plot of $\sigma_{max}\varepsilon_a$ versus $2N_f$ can then be used to estimate life; for a given loading condition (i.e., σ_m, σ_a, and ε_a), the fatigue life can be read directly from the plot.

It is apparent that the local-strain method can be applied over a greater range of conditions than the stress-life approach. In particular, if considerable plastic strains occur (e.g., low-cycle fatigue) or if structural constraints are present (e.g., a stress concentration), the strain-based method is preferable.[1] The local-strain approach also has advantages when considering VA loading.

Fracture-Mechanics Method

Stress-life and local-strain approaches measure the total life (initiation plus propagation life). The application of fracture mechanics to life prediction is solely a crack-propagation life method. Life calculations are carried out from a specific initial crack size (which may be a preexisting crack, a flaw, an intrinsic defect characteristic of the material, etc.)[4] to a final crack size at failure, determined from the material's inherent toughness; this idea of toughness, or resistance to crack growth, is one of the basic concepts of fracture mechanics.

Life predictions using fracture mechanics are built from two main relations. The first describes the stress field around the advancing crack tip, and the second characterizes the steady-state growth rate of the advancing crack. The stress field around the crack tip is given by the stress intensity factor, K. It is important to note that this factor describes the elastic stress field around the crack tip; if considerable yielding occurs, another expression for the stress field, J, should be used.[10] The basic relation for the stress intensity factor is given by

$$K = FS\sqrt{\pi a} \quad (18)$$

where F is a factor related to the geometry of the crack and component, and S is the engineering stress. Equation 18 may also be written as

$$K = FS\sqrt{\frac{\pi a}{Q}} \quad (19)$$

where Q is a crack-shape factor.

Thus, K is only a function of load and geometry. Its limiting value K_c (i.e., the value that will cause catastrophic failure of the component) is referred to as the fracture toughness of the material. Interestingly, this value is a material property, independent of loading condition. The crack length at failure (when the right side of Equation 18 equals K_c) can thus be determined from Equation 18 and used for the final crack length in a life-prediction analysis.

$$a_c = \frac{1}{\pi}\left(\frac{K_{IC}}{FS_{max}}\right) \quad (20)$$

K_{IC} refers to mode one (tensile) failure. The form of Equation 18 useful in fatigue-life calculations is

$$\Delta K = F\Delta S\sqrt{\pi a} \quad (21)$$

where ΔK is the stress-intensity-factor range produced by cycling over the stress range ΔS.

The steady-state growth rate of the advancing crack is characterized by the empirical equation proposed by Paris[11]

$$\frac{da}{dN} = C(\Delta K)^m \quad (22)$$

where da/dN is the crack growth rate, and C and m are constants.

A logarithmic plot of da/dN vs. ΔK (Figure 5) produces a straight line with a slope m and intercept C.[12,13] As seen from the plot, at low growth rates/stress-intensity-factor range, the linear relationship breaks down, and the crack-growth rate goes to zero at a threshold value ΔK_{th} below which crack growth does not occur (for large cracks only; this does not hold for very small cracks, as described in Reference 14). The location of this threshold value varies with R-ratio, the ratio of the minimum stress in the cycle to the maximum stress.[3] Also, at very high stress-intensity-factor ranges, the curve deviates upward as unsteady crack growth occurs.

Equation 22 is the most basic form of the growth-rate equation. A number of modifications are often made to account for various factors that may affect the growth-rate behavior. Chief among these modifications are those made to account for the effects of R-ratio. Increasing the R-ratio has the effect of increasing the crack growth rate, particularly at low growth rates (hence, the effect on ΔK_{th}).[15-23]

As mentioned, R-ratio is the ratio of the minimum stress (or stress intensity) to the maximum stress (or stress intensity)

$$R = \frac{S_{min}}{S_{max}} = \frac{K_{min}}{K_{max}} \quad (23)$$

Two popular methods of estimating the effect of R-ratio on growth rates are the Forman equation and Walker equation. The Forman equation is[15,24]

$$\frac{da}{dN} = \frac{C_2\Delta K^n}{(1-R)K_c - \Delta K} \quad (24)$$

where C_2 and n are constants. From $\Delta K = K_{max} - K_{min}$ and $K_{min} = RK_{max}$

$$\Delta K = K_{max}(1-R) \quad (25)$$

Substituting Equation 25 into Equation 24 gives[25]

$$\frac{da}{dN} = \frac{C_2\Delta K^n}{(1-R)(K_c - K_{max})} \quad (26)$$

It is apparent from Equation 26 that at high K_{max} (approaching K_c), the growth rate increases rapidly toward infinity; this is in agreement with the behavior observed by experiment (Figure 5). This makes it applicable to both intermediate- and high-growth rates.[3] To evaluate the constants C_2 and n,

$$\frac{da}{dN}(1-R)(K_c - K_{max})$$

is plotted versus ΔK on a log-log scale; the slope of the resulting line is n, and the intercept is C_2.

The Walker equation attempts to account for R-ratio effects through the use of a modified ΔK in Equation 22.[25]

$$\overline{\Delta K} = \frac{\Delta K}{(1-R)^{1-\gamma}} \quad (27)$$

where $\overline{\Delta K}$ is an equivalent stress-intensity-factor range at R = 0, and γ is a material constant.

Together, Equations 22 and 27 yield[15]

$$\frac{da}{dN} = \frac{C_1}{(1-R)^{m(1-\gamma)}}\Delta K^m \quad (28)$$

where C_1 is a Walker equation constant. The exponent, m, is thus independent of R-ratio according to this equation.

A log-log plot of Equation 28 gives a straight line with intercept C dependent on R-ratio[3]

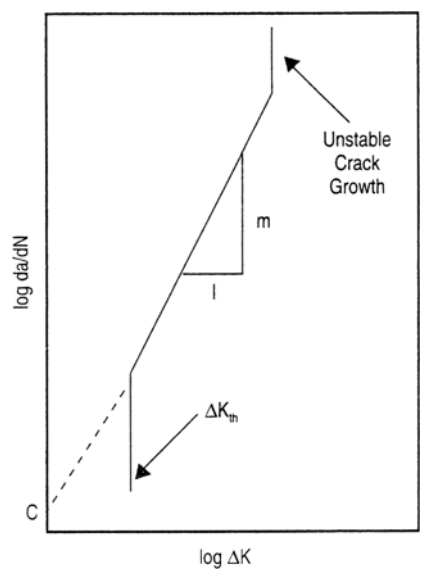

Figure 5. A crack growth rate curve da/dN versus ΔK with constants.

CYCLE COUNTING METHODS

Due to the difficulty in defining cycles in complex variable load histories, methods of cycle counting are necessary. A number of these techniques are described here, with information available in the ASTM standard on cycle counting, E 1049.[34] While rainflow counting may be slightly more complex than simple peak or level-crossing methods, the more realistic load history obtained is generally worth the effort. Not only can such factors as mean stress and sequence effects be included, but the cycles produced are related to the actual stress-strain response of the component.

Level-Crossing Counting

A count is recorded whenever an increasing (positively sloped) portion of the load history crosses a certain level above a reference level. Likewise, a count is recorded whenever a decreasing (negatively sloped) portion of the load history crosses a certain level below the reference level (Figure A). Crossings of the reference level itself are counted only on increasing portions of the curve. Once all counts and their load levels have been recorded, the most damaging cycles are constructed from the available counts. First, the largest cycle (with all the necessary level crossings) is constructed. From the remaining counts, the second largest cycle is constructed. This is continued until all counts have been used.[34]

Peak Counting

Peak counting records relative maxima and minima in the load history and their load levels. Generally, only maxima (peaks) above a reference level are counted, and only minima (valleys) below the reference level are counted (Figure B and Table 1). Alternatively, all peaks and valleys may be counted. To reduce the number of counts to a more manageable level, small amplitude loadings may be neglected by implementing mean-crossing peak counting, wherein only the largest peak or valley occurring between successive crossings of the mean level is counted. Similar to level-crossing counting, the most damaging cycle count is made from the largest peak and valley, then the second largest, etc., until all of the peaks and valleys have been used (Table 1).[34]

In both level crossing and peak counting, the order in which the resulting cycles are applied could affect the life. Thus, alternate methods of constructing the cycles might be preferred.[34]

Simple-Range Counting

In simple-range counting, range refers to the difference between the load levels of successive reversals (points where the slope of the curve changes sign, or peaks and valleys). Positive ranges occur when a peak follows a valley, and negative ranges occur when a valley follows a peak. To reduce the number of counts, ranges below a certain level are often neglected. Either positive or negative ranges may be counted, in which case each range is counted as one full cycle. On the other hand, both positive and negative ranges may be counted together, in which case each is only counted as a half-cycle.[34]

Range-Pair Counting

In range-pair counting, two subsequent ranges (of opposite sign) are considered together. If the second range is greater than or equal to the first range in size, the first range is counted, and the peak and valley are removed from consideration. If the second range is smaller than the first, then the next range in the load history is examined (and becomes the new second range). This is continued until the end of the load history is reached, at which point counting proceeds backwards from the end to include the remaining points. An advantage of range-pair counting is that the mean stress can be included as a second parameter in the matrix of cycles and ranges.[2,34]

Rainflow Counting

Rainflow counting is generally considered the best procedure currently available for counting cycles.[3,4] In rainflow counting, two consecutive ranges are considered together.

The name stems from the manner in which the cycle counting may be visualized: the load history is turned 90° counterclockwise so that the vertical axis is time, increasing downward (Figure C). Now imagine the peaks and valleys as a series of rooftops. The path rain would follow coming down the rooftops defines the cycles and half-cycles of the load history, but the rain must flow subject to several conditions. For example, rain flow begins at the inside of the peaks or valleys. Flow beginning at a peak/valley must stop when it comes opposite a peak/valley of equal or greater size, respectively (i.e., more positive than the starting peak or more negative than the starting valley), giving a half-cycle. Flow also stops when it meets the flow from a higher roof, again giving a half cycle. The cycle counting is analogous to the actual cyclic stress-strain response of the material, such that a range counted as a full cycle will form a complete stress-strain hysteresis loop, whereas a half-cycle will not.[2,4]

The algorithm for rainflow counting can be described as follows. The important terms are the previous (first) range, current (second) range, and the starting point of the history. If the second range is smaller than the first, the first range is momentarily skipped, the second range becomes the first, and the next peak or valley in the history is read to form the new second range (this is continued until a pair is found where the second range is larger than the first.)

If the second range is larger than the first, there are two possibilities. If the first range does not include the starting point, it is counted as a full cycle and the relevant values are recorded, and the peak and value are discarded from the history. If the first range does contain the starting point, it is counted as a half cycle and the relevant values are recorded; the first point of the range is removed, and the second point becomes the new starting point of the load history. In either case, after counting the cycle/half-cycle, the next ranges are formed from the three most recent peaks and valleys that were not removed, and the process is repeated. When no more ranges can be counted, those remaining are each counted as half cycles (this is done in Table 2 for the load history of Figure C). The cycles and their ranges (as well as the mean stress values) thus collected can be stored in matrix form and subsequently used for fatigue-life predictions.[34]

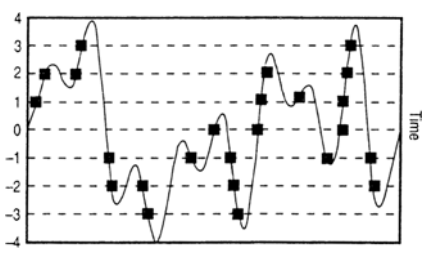

Figure A. A plot illustrating level-crossing counting.

Table 1. Peak Counts from Figure B and Corresponding Constructed Ranges

Load Peaks	Counts
3.8	1
3.7	1
2.5	1
2.3	1
1.5	1
0.5	1
−1.2	1
−1.5	1
−2.5	1
−2.6	1
−3.5	1
−3.9	1
Load Range	
7.7	1
7.2	1
5.1	1
4.8	1
3.0	1
1.7	1

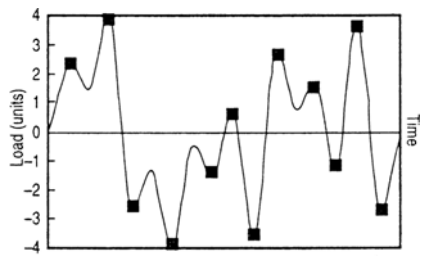

Figure B. A plot demonstrating those peaks that should be counted in peak-cycle counting.

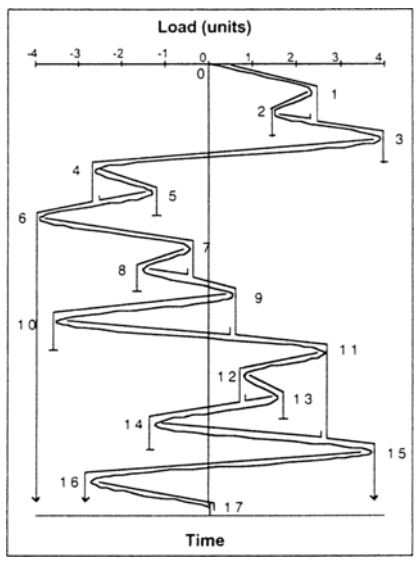

Figure C. An example of rainflow counting for cycle counting.

Table 2. Counted Ranges and the Resulting Number of Cycles from Figure C

Ranges Considered (in order)	Cycles
1–2	1
0–3	0.5
4–5	1
7–8	1
9–10	1
12–13	1
11–14	1
3–6	0.5
6–15	0.5
15–16	0.5
16–17	0.5

$$C = \frac{C_1}{(1-R)^{m(1-\gamma)}} \quad (29)$$

The slope, m, is equivalent to the value for an R-ratio of zero.

A life prediction analysis is made by simply integrating

$$\frac{da}{dN} \text{ [using whatever form of } \frac{da}{dN} = f(\Delta K, R) \text{ chosen]}$$

between the initial and final crack sizes. Again, the initial crack size may be a preexisting crack in a structure (where the remaining useful life is to be estimated) or may be assigned based on defects inherent in the material or structure under consideration.[4] This initial crack size, a_0, is often determined by nondestructive evaluation (NDE) techniques.[26] Thus, the number of cycles to failure is given by

$$N_1 = \int_{a_0}^{a_c} dN = \int_{a_0}^{a_c} \frac{da}{f(\Delta K, R)} \quad (30)$$

In the simplest cases, a closed form may exist for the integration. However, a number of factors can greatly increase the complexity of the integration, such as modifications for R-ratio, making it difficult or impossible to obtain a closed-form solution. Similarly, the geometry factor F (as well as Q) generally changes with the advance of the crack.[15] Thus, numerical-integration techniques are often necessary. Possibly the simplest method is to perform the iteration in a number of small steps, assuming a small crack increment and calculating F, ΔK, and da/dN from this. Then, assuming for small steps that

$$\frac{da}{dN} = \frac{\Delta a}{\Delta N}$$

a ΔN can be calculated for each increment. When the critical crack size is reached, the ΔNs are summed for the life.[15] Similarly, the procedure can be carried out using a suitable integration rule (e.g., Simpson's rule).[3] The integration can also be carried out graphically. Inverting the crack-growth-rate equation yields

$$\frac{dN}{da} = \frac{1}{f(\Delta K, R)} \quad (31)$$

Combining with Equation 30 results in

$$N_1 = \int_{a_0}^{a_c} \left(\frac{dN}{da}\right) da \quad (32)$$

Thus, the area under a plot of dN/da versus a gives the number of cycles to failure.[3] To produce the plot, a series of crack increments are again chosen and used to determine F and ΔK and, thus, da/dN, which is then inverted and plotted versus the crack length.

As mentioned, the stress intensity factor, K, is based on an assumption of linear elastic fracture mechanics (LEFM). In high-ductility materials and where the plastically strained zone ahead of the crack tip is large, predictions based on ΔK can result in considerable error. In such cases, elastic-plastic fracture mechanics may be more effective,[4,10,15] wherein the parameter J (or the cyclic J-integral) is used to describe the state of stress/strain at the crack tip. For a discussion of the J-integral as it relates to fatigue crack growth, refer to Reference 27.

The use of fracture mechanics in fatigue life prediction has become widespread. A standard test method exists for determining fatigue-crack-growth rate (ASTM E647),[28] and the fact that a given ΔK (for a given material, stress-ratio, and environment) produces the same growth rate for any combination of stress/crack length makes a single set of material data widely applicable for any type of loading and geometry.[14] A major advantage of fracture mechanics is that it is a damage-tolerant method; the discovery of a crack does not necessarily mean that the component's useful life is exhausted. The crack growth can be monitored, and estimates of remaining life can be made based on the crack length and K_c.[8] Of course, this may necessitate a periodic inspection of cracks in critical applications using NDE methods.[26] This may not be trivial and is often quite costly. But in applications where replacement entails great expense, the cost of periodic inspection may be worthwhile.

VARIABLE AMPLITUDE LOADING

The life prediction procedures discussed so far were based on CA loading. This type of fatigue data is by far the easiest to produce. As such, the vast majority of available fatigue data were generated in CA tests, and many of the relationships used in fatigue were originally developed based on the data from this type of loading. Unfortunately, in real-world applications, CA loading is seldom the case (and is actually the exception). Most fatigue design problems will involve VA loading, which may be random, block-shaped (blocks of CA loading at different amplitudes/mean stresses), etc. Making a VA life-prediction analysis based on an unmodified CA model will often result in considerable error. Thus, a number of methods of life prediction under VA loading have been advanced. Ideally, a VA technique should account for a number of factors that affect life, including sequence effects and overloads, prestrain effects, and mean stress.

It is well established that the order in which the various load levels are applied in VA loading can have a significant effect on the life of the component leading to sequence effects. In general, a high-tensile cycle (overload) followed by lower stress cycles can have the beneficial effect of temporarily retarding crack growth, and, thus, increasing life. This may not seem logical, but the effect is due to the combined (conflicting) influences of a beneficial residual compressive mean stress produced at the crack tip and increased strain damage introduced during the high-stress cycles (with the residual stress effect being dominant).[7]

Compressive overloads have the opposite effect (reducing life), but to a much lesser degree. This leads to the interesting effect that if a high-amplitude reversed load sequence is stopped after a tensile peak and then cycled at a lower stress level, a longer life may be produced than if the same high amplitude load had been stopped after a compressive valley.[7] The magnitude of the effect increases with the ratio of the maximum stress at the high-stress level to the maximum stress at the low-stress level.[3] Also, the effect is the largest when the overloads act mainly in the same direction (i.e., either all tensile or all compressive).[29] Highly irregular loads, however, decrease the effect of overloads. Because the effect is mainly to retard crack growth, neglecting sequence effects will often lead to conservative life predictions unless large compressive overloads are experienced.[3] On the other hand, in unnotched and crack-free specimens, a small number of high overstresses combined with cycling at low stress will actually have the opposite effect, in that the high stresses may initiate cracks early in the life that would have taken much longer to form at the lower stress level. The cracks can, however, then propagate at the lower stress level, leading to a reduction in life.[2] Thus, caution must be exercised if sequence effects can be expected.

The application of a small number of plastic prestrain cycles can affect the fatigue life of the component (usually reducing it). In real-world applications, a certain amount of prestrain is common and, hence, can affect the service life. This effect is generally attributed to crack initiation caused by the prestraining.[2] Prestraining samples prior to testing can help account for this effect.

Mean-stress effects are as important under VA loading as they are under CA loading. However, the mean stress will vary from cycle to cycle under VA. Consequently, a mean stress may be determined for each cycle or average values may be used.

The most common technique for estimating a VA fatigue life is through the use of the linear cumulative damage con-

cept, or Miner's rule. While the method has its shortcomings, attempts to modify it to increase accuracy have often met with only limited success and at the cost of a considerable increase in complexity, as well as often requiring experimental values that are seldom readily available.

Miner's Rule of Linear Damage Accumulation

Miner's rule, or the Palmgren-Miner rule, is a fatigue-life-prediction methodology based on an assumption of linear accumulation of damage in specimens subjected to VA loading. Miner's rule was advanced by Miner[30] in an attempt to apply the CA fatigue data available at the time to more complex VA loading schedules, as might be experienced in actual service.[1,3] According to Miner's rule, at a given stress level, the ratio of the number of loading cycles to the number of cycles at failure corresponds to the fraction of the total life that is exhausted.[4]

Miner based his method on the assumption that the amount of damage accumulated is directly related to the work absorbed by the specimen, and that failure corresponds to a particular amount of work.[30] Let $w_1, w_2, \ldots w_n$ represent the work absorbed at each of n loading levels, and let W represent the total work required to bring about failure. Then, at failure

$$w_1 + w_2 + \ldots + w_n = W \qquad (33)$$

Now, the fraction of the total work absorbed at a particular loading level is equal to the fraction of total life exhausted, such that

$$\frac{w_1}{W} = \frac{n_1}{N_1}; \frac{w_2}{W} = \frac{n_2}{N_2}; \frac{w_n}{W} = \frac{n_n}{N_n} \qquad (34)$$

where $n_1, n_2, \ldots n_n$ are the number of cycles at stress level $s_1, s_2, \ldots s_n$, respectively, and $N_1, N_2, \ldots N_n$ are the number of cycles required to bring about failure at $s_1, s_2, \ldots s_n$, respectively, as given by the S-N curve.

Now, from Equation 33,

$$\frac{w_1}{W} + \frac{w_2}{W} + \ldots + \frac{w_n}{W} = 1 \qquad (35)$$

Substituting Equation 34 into Equation 35,

$$\frac{n_1}{N_1} + \frac{n_2}{N_2} + \ldots + \frac{n_n}{N_n} = 1 \qquad (36)$$

or,

$$\sum \frac{n_n}{N_n} = 1 \qquad (37)$$

Thus, when the sum in Equations 36 and 37 (i.e., the sum of the fractions of the total life exhausted at each stress level) equals one, failure is predicted.

Besides the assumption that failure will occur after a certain amount of work has been absorbed by the sample, Miner also made several other assumptions:[30] sinusoidal loading, no work hardening, that failure occurs when a crack is visible, and that stresses below that required to cause failure at 10^7 cycles are neglected. To account for R-ratio effects, S-N curves are needed for each R-ratio encountered by the part. Alternatively, a modified Goodman diagram can be used to relate stresses at several R-ratios, but at the expense of accuracy.[3,4,30]

Miner's rule has many advantages, namely its simplicity and ease of applications. Unfortunately, it suffers from a number of drawbacks that make it often inaccurate. First, work since Miner has shown the accumulation of fatigue damage to be a nonlinear process.[4,31] Also, the method fails to account for such factors as load sequence effects (for stress-life based evaluation, as Miner used), mean stress effects, notch effects (e.g., the effect of a single tensile or compressive overload on a notched/cracked specimen), and the effects of prior static loading.[1,4,32]

Miner's experimental results support his hypothesis; in 22 Alclad 24S-T aluminum alloy specimens, the average

$$\sum \frac{n_n}{N_n}$$

was 1.015, with a maximum of 1.45 and a minimum of 0.61.[30] However, Miner's results were based on very simple block-loading schedules, with only two or three stress levels. While Miner claimed the method to be conservative, subsequent experiments have not always borne this out in cases of more complex loading programs or random loading, where results have been shown to be nonconservative by an order of magnitude or more.[4] On the other hand, under special testing conditions, Miner's rule has been shown to be overly conservative, giving

$$\sum \frac{n_n}{N_n}$$

values as high as 300.[4]

Despite its shortcomings, Miner's rule can provide a simple technique for obtaining an approximation of life under conditions of VA loading. When used with an appropriate cycle-counting scheme (e.g., rainflow counting), the method can be used to produce life estimates very quickly for extremely complex random loadings, which can be a considerable task when using other techniques.

Stress-Life and Local-Strain Approaches

The stress-life approach does not account for sequence effects.[3] Thus, if sequence effects can be neglected, the stress-life approach can be used in conjunction with Miner's rule (and if necessary, an appropriate cycle counting scheme) to make fatigue-life estimates under VA loading.[4] However, if sequence effects are expected to be a factor, the use of the stress-life method is generally inadvisable.

The local-strain method is more flexible than the stress-life technique. It can account for both sequence effects and prestrain because of its basis on local-strain conditions and the cyclic stress-strain behavior.[3,4] Using a cycle-counting technique (e.g., rainflow counting), the individual cycles are determined from the load history, and strain range and mean stress are determined for each cycle. The analysis used for CA loading is then used on individual cycles (or blocks of cycles with an equivalent strain amplitude) taking the individual values of stress and strain from the hysteresis loop for the cycle [given by $\varepsilon_a = f(\sigma_a)$] and the strain-life relationships to find the value of N_f for the cycle.[3] This N_f can then be used with Miner's rule, and a life prediction can be made.

Fracture Mechanics

If sequence effects can be neglected, a simple cycle-by-cycle sum of crack increments can be determined. For the conditions of the ΔK and R-ratio of the cycle in question, the change in crack length can be read from a plot of

$$\frac{da}{dN} \text{ vs. } \Delta K$$

($\Delta N = 1$, so $da = \Delta a$).[1,3] Again, rainflow counting can be used to separate the cycles with their ranges and means stresses. When the sum of crack increments equals or exceeds the critical crack length a_c, the cycles are totaled for the fatigue life.

A very simple method for applying CA data to a VA load sequence is to assume that ΔK in the growth rate equation can be replaced by the root-mean-square value of ΔK for the load history, ΔK_{RMS}[33]

$$\Delta K_{RMS} = \left[\frac{\sum_{i=1}^{n} \Delta K_i^2}{N} \right]^{\frac{1}{2}} \qquad (38)$$

This makes for a very simple life calculation, but unfortunately, this method also neglects sequence effects, and in some cases, the life estimates resulting are fairly inaccurate.[1,3,33]

A number of methods that include sequence effects in fracture-mechanics-based life predictions have been conceived. These often involve either crack closure or residual-stress effects. Suresh covers a number of these in Reference 1.

CONCLUSIONS

Although many prediction approaches lack the simplicity necessary for widespread application, or require data that are unavailable and difficult to create, a number are promising. For example, statistical methods for modeling crack growth are now being implemented.[25,26] Also, finite-element modeling of processes such as damage accumulation or crack growth (often based on one of the methods described here) is now being used to estimate life. Similarly, topics such as multiaxial fatigue, thermal fatigue, and crack closure have been omitted here.

The issue of fatigue-crack initiation is currently an important research area in light of the newly developed advanced nondestructive evaluation technologies.[26,37,38] Minute cracks on the order of 100 Å could be detected using atomic force microscopy or scanning tunneling microscopy. It is expected that research advances will be made in detecting and modeling minute cracks, which govern the initial fatigue life.

ACKNOWLEDGEMENTS

The authors are grateful to the National Science Foundation for their financial support. We thank M. Poats, our contract monitor, for her support and encouragement. We also thank Bill Becker of the University of Tennessee for his counsel on the topic of fatigue.

References

1. S. Suresh, *Fatigue of Materials* (New York: Cambridge University Press, 1991).
2. N.E. Dowling, *J. of Materials* (3) (1972), pp. 71–82.
3. N.E. Dowling, *Mechanical Behavior of Materials* (Englewood Cliffs: Prentice Hall, 1993).
4. C.C. Osgood, *Fatigue Design*, 2nd ed. (New York: Pergamon Press, 1982).
5. G.E. Dieter, *Mechanical Metallurgy*, 3rd ed. (New York: McGraw-Hill, 1986).
6. H. Neuber, *J. of App. Mech.*, 28 (1961), pp. 544–550.
7. J.H. Crews, *Effects of Loading Sequence for Notched Specimens Under High-Low Two-Step Fatigue Loading*, NASA TN D-6558 (November 1971).
8. V. Kliman, P. Fuleky, and J. Jelemenská, *Advances in Fatigue Lifetime Predictive Techniques, ASTM STP 1292*, 3rd vol., eds. M.R. Mitchell and R.W. Landgraf (Philadelphia, PA: ASTM, 1996), pp. 305–327.
9. S.M. Tipton and D.A. Newburn, *Advances in Fatigue Lifetime Predictive Techniques, ASTM STP 1122*, eds. M.R. Mitchell and R.W. Landgraf (Philadelphia, PA: ASTM, 1992), pp. 369–382.
10. M. Vormwald, P. Heuler, and T. Seeger, in Ref. 9, pp. 28–43.
11. P.C. Paris and F. Erdogan, *Trans. of the ASME J. of Basic Engrg*, 89 (1967), pp. 528–534.
12. R. Sunder, in Ref. 9, pp. 161–175.
13. A.T. Chang et al., *Advances in Fatigue Lifetime Predictive Techniques, ASTM STP 1292*, 3rd vol., eds. M.R. Mitchell and R.W. Landgraf (Philadelphia, PA: ASTM, 1996), pp. 100–115.
14. J.C. Newman et al., in Ref. 9, pp. 5–27.
15. D. Broek, *The Practical Use of Fracture Mechanics* (Boston: Kluver Academic Publishing, 1989).
16. P.K. Liaw et al., *Acta Met.*, 30 (1982), pp. 2071–2078.
17. P.K. Liaw, T.R. Leax and W.A. Logsdon, *Acta Met.*, 31 (1983), pp. 1581–1587.
18. P.K. Liaw, *Acta Met.*, (1985), p. 33.
19. P.K. Liaw, T.R. Leax and J.K. Donald, *Acta Met.*, 35 (1987), pp. 1489–1502.
20. J. McKittrick, *Met. Trans.*, 12A (1981), p. 1535.
21. S. Suresh and R.O. Ritchie, *Met. Trans.*, 13A (1982), p. 1627.
22. N. Minakawa and A.J. McEvily, *Scripta Metallurgica*, 15 (1981) p. 633.
23. D.L. Davidson, *Fatigue of Engineering Materials and Structures*, 3 (1980), p. 229.
24. C.M. Hudson and P.E. Lewis, *Part-Through Crack Fatigue Life Prediction, ASTM ASP 687*, ed. J. B. Chang (Philadelphia, PA: ASTM, 1979), pp. 113–128.
25. D. Broek, *Elementary Engineering Fracture Mechanics* (Boston, MA: Martinus Nijhoff Publishers, 1982).
26. S. Shanmugham and P.K. Liaw, *ASM Handbook, Fatigue and Fracture*, vol. 19 (Materials Park, OH: ASM, 1996), pp. 210–223.
27. N.E. Dowling and J.A. Begley, *Mechanisms of Crack Growth. ASTM STP 590* (Philadelphia: ASTM, 1976).
28. ASTM Standard E 647, *1995 ASTM Standards*, 03.01.
29. F. Xiangjiong, *Engineering Fracture Mechanics*, 34 (5/6) (1989), pp. 1241–1248.
30. M.A. Miner, *Trans. of the ASME*, 67 (1945).
31. H.A. Lipsitt, D.F. Frank, and G.C. Smith, *Proceedings of the Air Force Conference on Fatigue and Fracture of Aircraft Structures and Materials, AFFDL TR 70-144*, eds. H.A. Wood, et al. (Washington, D.C.: USAF, December 1969).
32. S.S. Manson, J.C. Freche, and C.R. Ensign, *Application of a Double Linear Damage Rule to Cumulative Fatigue*, NASA TN D-3839 (April 1967).
33. H. Alawi, *Trans. of the ASME, Journal of Engineering Materials and Technology*, 111 (4) (October 1989), pp. 338–344.
34. ASTM Standard E 1049, *1995 ASTM Standards*, 03.01.
35. M. Kato et al, *Scripta Metallurgica*, (1984), p. 18.
36. M. Kato and T. Mori, *Mechanics of Materials*, 13 (2) (March 1992), pp. 155–163.
37. S.E. Harvey, P.G. Marsh, and W.W. Gerberich, *Acta Met. et Mat.*, 42 (10) (1994), pp. 3493–3502.
38. M.-R. Lin, M.E. Fine, and T. Mura, *Acta Met.*, 34 (4) (1986).

ABOUT THE AUTHORS

R.K. Holman earned his B.S. in materials science and engineering at the University of Tennessee–Knoxville in 1996. He is entering graduate school this fall at the Massachusetts Institute of Technology. Mr. Holman is a member of TMS.

P.K. Liaw earned his Ph.D. in materials science and engineering at Northwestern University in 1980. He is currently a professor and Ivan Racheff Chair of Excellence in the Department of Materials Science and Engineering at the University of Tennessee. Dr. Liaw is also a member of TMS.

For more information, contact P.K. Liaw, University of Tennessee, Department of Materials Science and Engineering, 427 B. Dougherty Building, Knoxville, Tennessee 37996; (423) 974-6356; fax (423) 974-4115; e-mail pLiaw@utk.edu.