

Fig. 18.12 Surface flaw shape parameter. (From Ref. 22. Adapted by permission of Prentice-Hall, Inc., Englewood Cliffs, New Jersey.)

To approximate the effects of strain hardening, a flow stress σ_o , taken to be an average of the yield and ultimate strengths, is often used when computing the plastic collapse stress. The plastic collapse stress σ_c is that applied stress which produces σ_o across the remaining uncracked ligament, and is the maximum applied stress that a perfectly plastic material can sustain. This stress may be determined using a limit load analysis. In general, the plastic collapse stress is a function of geometry, type of loading, type of support (boundary conditions), and through-thickness constraint (plane stress or plane strain).^{6,25} For a single through-thickness crack of length a in a strip with width b loaded in tension (see Fig. 18.9), if end rotations are restrained, the plastic collapse stress under plane stress conditions may be approximated by²⁵

$$\sigma_c = \sigma_o(1 - a/b) \tag{18.39}$$

18.5 FATIGUE AND STRESS CONCENTRATION

Static or quasistatic loading is rarely observed in modern engineering practice, making it essential for the designer to address himself or herself to the implications of repeated loads, fluctuating loads, and rapidly applied loads. By far, the majority of engineering design projects involve machine parts



Fig. 18.13 Failure assessment diagram.

subjected to fluctuating or cyclic loads. Such loading induces fluctuating or cyclic stresses that often result in failure by fatigue.

Fatigue failure investigations over the years have led to the observation that the fatigue process actually embraces two domains of cyclic stressing or straining that are significantly different in character, and in each of which failure is probably produced by different physical mechanisms. One domain of cyclic loading is that for which significant plastic strain occurs during each cycle. This domain is associated with high loads and short lives, or low numbers of cycles to produce fatigue failure, and is commonly referred to as low-cycle fatigue. The other domain of cyclic loading is that for which the strain cycles are largely confined to the elastic range. This domain is associated with lower loads and long lives, or high numbers of cycles to produce fatigue failure, and is commonly referred to as high-cycle fatigue. Low-cycle fatigue is typically associated with cycle lives from 1 up to about 10^4 or 10^5 cycles. Fatigue may be characterized as a progressive failure phenomenon that proceeds by the *initiation* and *propagation* of cracks to an unstable size. Although there is not complete agreement on the microscopic details of the initiation and propagation of the cracks, processes of reversed slip and dislocation interaction appear to produce fatigue nuclei from which cracks may grow. Finally, the crack length reaches a critical dimension and one additional cycle then causes complete failure. The final failure region will typically show evidence of plastic deformation produced just prior to final separation. For ductile materials the final fracture area often appears as a shear lip produced by crack propagation along the planes of maximum shear.

Although designers find these basic observations of great interest, they must be even more interested in the macroscopic phenomenological aspects of fatigue failure and in avoiding fatigue failure during the design life. Some of the macroscopic effects and basic data requiring consideration in designing under fatigue loading include:

- 1. The effects of a simple, completely reversed alternating stress on the strength and properties of engineering materials.
- 2. The effects of a steady stress with superposed alternating component, that is, the effects of cyclic stresses with a nonzero mean.
- 3. The effects of alternating stresses in a multiaxial state of stress.
- 4. The effects of stress gradients and residual stresses, such as imposed by shot peening or cold rolling, for example.
- 5. The effects of stress raisers, such as notches, fillets, holes, threads, riveted joints, and welds.
- 6. The effects of surface finish, including the effects of machining, cladding, electroplating, and coating.
- 7. The effects of temperature on fatigue behavior of engineering materials.
- 8. The effects of size of the structural element.
- 9. The effects of accumulating cycles at various stress levels and the permanence of the effect.
- 10. The extent of the variation in fatigue properties to be expected for a given material.
- 11. The effects of humidity, corrosive media, and other environmental factors.
- 12. The effects of interaction between fatigue and other modes of failure, such as creep, corrosion, and fretting.

18.5.1 Fatigue Loading and Laboratory Testing

Faced with the design of a fatigue-sensitive element in a machine or structure, a designer is very interested in the fatigue response of engineering materials to various loadings that might occur throughout the design life of the machine under consideration. That is, the designer is interested in the effects of various *loading spectra* and associated *stress spectra*, which will in general be a function of the design configuration and the operational use of the machine.

Perhaps the simplest fatigue stress spectrum to which an element may be subjected is a zeromean sinusoidal stress-time pattern of constant amplitude and fixed frequency, applied for a specified number of cycles. Such a stress-time pattern, often referred to as a completely reversed cyclic stress, is illustrated in Fig. 18.14*a*. Utilizing the sketch of Fig. 18.14, we can conveniently define several useful terms and symbols; these include:

$$\sigma_{\rm max}$$
 = maximum stress in the cycle

$$\sigma_m$$
 = mean stress = $(\sigma_{max} + \sigma_{min})/2$

 σ_{\min} = minimum stress in the cycle

$$\sigma_a$$
 = alternating stress amplitude = $(\sigma_{\text{max}} - \sigma_{\text{min}})/2$

$$\Delta \sigma$$
 = range of stress = $\sigma_{\rm max} - \sigma_{\rm min}$

$$R = \text{stress ratio} = \sigma_{\min} / \sigma_{\max}$$

A = amplitude ratio =
$$\sigma_a/\sigma_m = (1 - R)/(1 + R)$$



Fig. 18.14 Several constant-amplitude stress-time patterns of interest: (a) completely reversed, R = -1; (b) nonzero mean stress; (c) released tension, R = 0.

Any two of the quantities just defined, except the combinations σ_a and $\Delta\sigma$ or the combination A and R, are sufficient to describe completely the stress-time pattern above.

More complicated stress-time patterns are produced when the mean stress, or stress amplitude, or both mean and stress amplitude change during the operational cycle, as illustrated in Fig. 18.15. It may be noted that this stress-time spectrum is beginning to approach a degree of realism. Finally, in Fig. 18.16 a sketch of a realistic stress spectrum is given. This type of quasirandom stress-time pattern might be encountered in an airframe structural member during a typical mission including refueling, taxi, takeoff, gusts, maneuvers, and landing. The obtaining of useful, realistic data is a challenging task in itself. Instrumentation of existing machines, such as operational aircraft, provide some useful information to the designer if his or her mission is similar to the one performed by the instrumented machine. Recorded data from accelerometers, strain gauges, and other transducers may in any event provide a basis from which a statistical representation can be developed and extrapolated to future needs if the fatigue processes are understood.

Basic data for evaluating the response of materials, parts, or structures are obtained from carefully controlled laboratory tests. Various types of testing machines and systems commonly used include:

- 1. Rotating-bending machines:
 - a. Constant bending moment type
 - b. Cantilever bending type
- 2. Reciprocating-bending machines.



Fig. 18.15 Stress-time pattern in which both mean and amplitude change to produce a more complicated stress spectrum.

- 3. Axial direct-stress machines:
 - a. Brute-force type
 - b. Resonant type
- 4. Vibrating shaker machines:
 - a. Mechanical type
 - b. Electromagnetic type
- 5. Repeated torsion machines.
- 6. Multiaxial stress machines.



Fig. 18.16 A quasirandom stress-time pattern that might be typical of an operational aircraft during any given mission.

- 7. Computer-controlled closed-loop machines.
- 8. Component testing machines for special applications.
- 9. Full-scale or prototype fatigue testing systems.

Computer-controlled fatigue testing machines are widely used in all modern fatigue testing laboratories. Usually such machines take the form of precisely controlled hydraulic systems with feedback to electronic controlling devices capable of producing and controlling virtually any strain-time, load-time, or displacement-time pattern desired. A schematic diagram of such a system is shown in Fig. 18.17.

Special testing machines for component testing and full-scale prototype testing systems are not found in the general fatigue testing laboratory. These systems are built up especially to suit a particular need, for example, to perform a full-scale fatigue test of a commercial jet aircraft.

It may be observed that fatigue testing machines range from very simple to very complex.

The very complex testing systems, used, for example, to test a full-scale prototype, produce very specialized data applicable only to the particular prototype and test conditions used; thus, for the particular prototype and test conditions the results are very accurate, but extrapolation to other test



Fig. 18.17 Schematic diagram of a computer-controlled closed-loop fatigue testing machine.

18.5 FATIGUE AND STRESS CONCENTRATION

conditions and other pieces of hardware is difficult, if not impossible. On the other hand, simple smooth-specimen laboratory fatigue data are very general and can be utilized in designing virtually any piece of hardware made of the specimen material. However, to use such data in practice requires a quantitative knowledge of many pertinent differences between the laboratory and the application, including the effects of nonzero mean stress, varying stress amplitude, environment, size, temperature, surface finish, residual stress pattern, and others. Fatigue testing is performed at the extremely simple level of smooth specimen testing, the extremely complex level of full-scale prototype testing, and everywhere in the spectrum between. Valid arguments can be made for testing at all levels.

18.5.2 The S-N-P Curves-A Basic Design Tool

Basic fatigue data in the high-cycle life range can be conveniently displayed on a plot of cyclic stress level versus the logarithm of life, or alternatively, on a log-log plot of stress versus life. These plots, called *S*-*N* curves, constitute design information of fundamental importance for machine parts subjected to repeated loading. Because of the scatter of fatigue life data at any given stress level, it must be recognized that there is not only one *S*-*N* curves for a given material, but a family of *S*-*N* curves with probability of failure as the parameter. These curves are called the *S*-*N*-*P* curves, or curves of constant probability of failure on a stress-versus-life plot. A representative family of *S*-*N* curves is illustrated in Fig. 18.18. It should also be noted that references to the "*S*-*N* curve" in the literature generally refer to the mean curve unless otherwise specified. Details regarding fatigue testing and the experimental generation of *S*-*N*-*P* curves may be found in Ref. 1.

The mean S-N curves sketched in Fig. 18.19 distinguish two types of material response to cyclic loading commonly observed. The ferrous alloys and titanium exhibit a steep branch in the relatively short life range, leveling off to approach a stress asymptote at longer lives. This stress asymptote is called the *fatigue limit* (formerly called endurance limit) and is the stress level below which an infinite number of cycles can be sustained without failure. The nonferrous alloys do not exhibit an asymptote, and the curve of stress versus life continues to drop off indefinitely. For such alloys there is no fatigue limit, and failure as a result of cyclic load is only a matter of applying enough cycles. All materials, however, exhibit a relatively flat curve in the long-life range.

To characterize the failure response of nonferrous materials, and of ferrous alloys in the finitelife range, the term *fatigue strength at a specified life*, S_N , is used. The term fatigue strength identifies the stress level at which failure will occur at the specified life. The specification of *fatigue strength* without specifying the corresponding life is meaningless. The specification of a *fatigue limit* always implies infinite life.



Fig. 18.18 Family of S–N–P curves, or R–S–N curves, for 7075-T6 aluminum alloy. Note: P = probability of failure; R = reliability = 1 - P. (Adapted from Ref. 31, p. 117; with permission from John Wiley & Sons, Inc.)



Fig. 18.19 Two types of material response to cyclic loading.

18.5.3 Factors That Affect S-N-P Curves

There are many factors that may influence the fatigue failure response of machine parts or laboratory specimens, including material composition, grain size and grain direction, heat treatment, welding, geometrical discontinuities, size effects, surface conditions, residual surface stresses, operating temperature, corrosion, fretting, operating speed, configuration of the stress–time pattern, nonzero mean stress, and prior fatigue damage. Typical examples of how some of these factors may influence fatigue response are shown in Figs. 18.20 through 18.35. It is usually necessary to search the literature and existing data bases to find the information required for a specific application and it may be necessary to undertake experimental testing programs to produce data where they are unavailable.

18.5.4 Nonzero Mean and Multiaxial Fatigue Stresses

Most basic fatigue data collected in the laboratory are for completely reversed alternating stresses, that is, zero mean cyclic stresses. Most service applications involve nonzero mean cyclic stresses. It is therefore very important to a designer to know the influence of mean stress on fatigue behavior so that he or she can utilize basic completely reversed laboratory data in designing machine parts subjected to nonzero mean cyclic stresses.

If a designer is fortunate enough to find test data for his or her proposed material under the mean stress conditions and design life of interest, the designer should, of course, use these data. Such data are typically presented on so-called *master diagrams* or *constant life diagrams* for the material. A master diagram for a 4340 steel alloy is shown in Fig. 18.36. An alternative means of presenting this type of fatigue data is illustrated in Fig. 18.37.

If data are not available to the designer, he or she may estimate the influence of nonzero mean stress by any one of several empirical relationships that relate failure at a given life under nonzero mean conditions to failure at the same life under zero mean cyclic stresses. Historically, the plot of alternating stress amplitude σ_a versus mean stress σ_m has been the object of numerous empirical curve-fitting attempts. The more successful attempts have resulted in four different relationships, namely:

- 1. Goodman's linear relationship.
- 2. Gerber's parabolic relationship.
- 3. Soderberg's linear relationship.
- 4. The elliptic relationship.



Fig. 18.20 Effect of material composition on the S–N curve. Note that ferrous and titanium alloys exhibit a well-defined fatigue limit, whereas other alloy compositions do not. (Data from Refs. 26 and 27.)



Fig. 18.21 Effect of grain size on the S–*N* curve for 18S aluminum alloy. Average diameter ratio of coarse to fine grains is approximately 27 to 1. Nominal composition: 4.0% copper, 2.0% nickel, 0.6% magnesium. Note that at a life of 10⁸ cycles of the mean fatigue strength of the coarse-grained material is about 3000 psi lower than for fine-grained material. (Data from Ref. 28; adapted from *Fatigue and Fracture of Metals*, by W. M. Murray, by permission of the MIT Press, Cambridge, Massachusetts, copyright, 1952.)



Fig. 18.22 Effect on the *S*–*N* curve of grain flow direction relative to longitudinal loading direction for specimens machined from crankshaft forgings. Nominal composition: 0.41% carbon, 0.47% manganese, 0.01% silicon, 0.04% phosphorous, 1.8% nickel. $S_{u} = 139,000$ psi, $S_{yp} = 115,000$ psi, *e* (2.0 in.) = 20%. (Data from Ref. 29.)



Fig. 18.23 Effects of heat treatment on the S–N curve of SAE 4130 steel, using 0.19-in.diameter rotating bending specimens cut from $\frac{3}{6}$ -in. plate, 1625°F, oil quenched, followed by three different tempers. Temper No. 1: $S_u = 129,000$ psi; $S_{yp} = 118,000$ psi. Temper No. 2: $S_u = 150,000$ psi; $S_{yp} = 143,000$ psi. Temper No. 3: $S_u = 206,000$ psi; $S_{yp} = 194,000$ psi.



Fig. 18.24 Effects of welding detail on the S–N curve of structural steel, with yield strength in the range 30,000–52,000 psi. Tests were released tension ($\sigma_{min} = 0$). (Data from Ref. 30.)

A modified form of the Goodman relationship is recommended for general use under conditions of high-cycle fatigue. For tensile mean stress ($\sigma_m > 0$), this relationship may be written

$$\frac{\sigma_a}{\sigma_N} + \frac{\sigma_m}{\sigma_u} = 1 \tag{18.40}$$

where σ_u is the material ultimate strength and σ_N is the zero mean stress fatigue strength for a given number of cycles N. For a given alternating stress, compressive mean stresses ($\sigma_m < 0$) have been empirically observed to exert no influence on fatigue life. Thus, for $\sigma_m < 0$, the fatigue response is identical to that for $\sigma_m = 0$ with $\sigma_a = \sigma_N$.

The modified Goodman relationship is illustrated in Fig. 18.38. This curve is a failure locus for the case of *uniaxial* fatigue stressing. Any cyclic loading that produces an alternating stress and mean stress that exceeds the bounds of the locus will cause failure in fewer than N cycles. Any alternating stress—mean stress combination that lies within the locus will result in more than N cycles without failure. Combinations that just touch the locus produce failure in exactly N cycles. The modified Goodman relationship shown in Fig. 18.38 considers fatigue failure exclusively. The reader is cautioned to insure that the maximum and minimum stresses produced by the cyclic loading do not exceed the material yield strength σ_{yp} such that failure by yielding would be predicted to occur.

For a given design life N, Eq. (18.40) may be used to estimate whether fatigue failure will occur under any nonzero mean stress condition if the ultimate strength σ_u and the completely reversed ($\sigma_m = 0$) fatigue strength σ_N for the material are known. These material properties are usually available.

If the machine part under consideration is subjected not only to nonzero mean stress, but also to a multiaxial state of stress, then multiaxial fatigue must be considered. Historically, the majority of fatigue-related research has been focused on uniaxial loading conditions, and consequently multiaxial fatigue is not as well characterized.



Fig. 18.25 Effects of geometrical discontinuities on the *S*-*N* curve of SAE 4130 steel sheet, normalized, tested in completely reversed axial fatigue test. Specimen dimensions (t = thickness, w = width, r = notch radius): Unnotched: t = 0.075 in., w = 1.5 in. Hole: t = 0.075 in., w = 4.5 in., r = 1.5 in. Fillet: t = 0.075 in., w_{net} = 1.5 in., w_{gross} = 2.25 in., r = 0.0195 in. Edge notch: t = 0.075 in., w_{net} = 1.5 in., w_{gross} = 2.25 in., r = 0.0195 in. Edge notch: t = 0.075 in., w_{net} = 1.5 in., w_{gross} = 2.25 in., r = 0.0195 in.



Fig. 18.26 Size effects on the S-N curve of SAE 1020 steel specimens cut from a 3¹/₂-in.diameter hot-rolled bar, testing in rotating bending. (Data from Ref. 32.)



Fig. 18.27 Effect of surface finish on the *S*-*N* curve of 0.33% carbon steel specimens, testing in a rotating cantilever beam machine: (*a*) high polish, longitudinal direction; (*b*) FF emery finish; (*c*) No. 1 emery finish; (*d*) coarse emery finish; (*e*) smooth file; (*f*) as-turned; (*g*) bastard file; (*h*) coarse file. (Data from Ref. 33.)



Fig. 18.28 Effect of shot peening on the S-N curves for welded and unwelded steel plate. (Data from Ref. 30.)



Fig. 18.29 Effect of operating temperature on the S–N curve of a 12% chromium steel alloy. Alloy composition = 0.10% C, 0.45% Mn, 0.21% Ni, 12.3% Cr, and 0.38% Mo. (Data from Ref. 34.)



Fig. 18.30 Effects of corrosion on the S–N curves of various aircraft materials tested in push–pull loading in seawater or sea spray. (Data from Ref. 35.)



Fig. 18.31 Effect of fretting on the S-N curve of a forged 0.24% steel. (Data from Ref. 36; reprinted with permission from McGraw-Hill Book Company.)



Fig. 18.32 Effect of operating speed on the fatigue strength at 10⁸ cycles for several different ferrous alloys. (Data from Ref. 37, p. 381.)



Fig. 18.33 Effect of ultimate strength on the S-N curve for transverse butt welds in two steels. (Data from Ref. 38.)

If the applied multiaxial stresses are in-phase and the principal axes do not rotate during the cyclic loading, one commonly used means of estimating fatigue failure is to compute an effective or equivalent alternating and mean stress. These effective stresses are then subsequently treated as uniaxial stresses and used in conjunction with the modified Goodman diagram, as described in the preceding paragraphs. Effective stresses are often derived using combined stress theories of failure for static loading. For example, using the distortion energy theory, Eq. (18.3) is used to define the effective alternating and mean stresses (σ_a)_e and (σ_m)_e as follows

$$(\sigma_a)_e = \frac{1}{\sqrt{2}} \sqrt{(\sigma_{1a} - \sigma_{2a})^2 + (\sigma_{2a} - \sigma_{3a})^2 + (\sigma_{3a} - \sigma_{1a})^2}$$
(18.41)

$$(\sigma_m)_e = \frac{1}{\sqrt{2}}\sqrt{(\sigma_{1m} - \sigma_{2m})^2 + (\sigma_{2m} - \sigma_{3m})^2 + (\sigma_{3m} - \sigma_{1m})^2}$$
(18.42)

where σ_1 , σ_2 , and σ_3 represent principal stresses and the subscripts *a* and *m* refer to alternating and mean components.

Equations (18.40), (18.41), and (18.42) provide a means of predicting fatigue failure under conditions of cyclic multiaxial states of stress assuming the stress amplitudes remain constant throughout the life of the part. Numerous other stress-based multiaxial fatigue methodologies have been proposed, some of which employ different means of quantifying the effect of mean stresses. No consensus has yet been reached as to which multiaxial fatigue methodology best predicts fatigue failure, and predictions made using equivalent stress methodologies should be considered approximate. For more detailed discussions of multiaxial fatigue, the reader is referred to Refs. 1, 42–45.

18.5.5 Spectrum Loading and Cumulative Damage

In virtually every engineering application where fatigue is an important failure mode, the alternating stress amplitude may be expected to vary or change in some way during the service life. Such variations and changes in load amplitude, often referred to as *spectrum loading*, make the direct use of standard S-N curves inapplicable because these curves are developed and presented for constant stress amplitude operation. Therefore, it becomes important to a designer to have available a theory or hypothesis, verified by experimental observations, that will permit good design estimates to be made for operation under conditions of spectrum loading using the standard constant amplitude S-N curves.



Fig. 18.34 Various ways of presenting the influence of nonzero mean stress on the fatigue behavior of 2014-T6 aluminum alloy. (Adapted from Ref. 39.)



Fig. 18.35 Illustration of the influence of accumulated fatigue damage on subsequent fatigue behavior of carbon steel. Note: Life of virgin material at $\sigma_a = 45,000$ psi is approximately 30,000 cycles. (Data from Ref. 40; reprinted with permission from John Wiley & Sons, Inc.)



Fig. 18.36 Master diagram for 4340 steel. (From Ref. 41, p. 317.)



Fig. 18.37 Best-fit S–N curves for notched 4130 alloy steel sheet, $K_t = 4.0$. (From Ref. 11.)



Fig. 18.38 Modified Goodman relationship.

The basic postulate adopted by all fatigue investigators working with spectrum loading is that operation at any given cyclic stress amplitude will produce *fatigue damage*, the seriousness of which will be related to the number of cycles of operation at that stress amplitude and also related to the total number of cycles that would be required to produce failure of an undamaged specimen at that stress amplitude. It is further postulated that the damage incurred is permanent and operation at several different stress amplitudes in sequence will result in an accumulation of total damage equal to the sum of the damage increments accrued at each individual stress level. When the total accumulated damage reaches a critical value, fatigue failure occurs. Although the concept is simple in principle, much difficulty is encountered in practice because the proper assessment of the amount of damage incurred by operation at any given stress level S_i for a specified number of cycles n_i is not straightforward. Many different cumulative damage theories have been proposed for the purposes of assessing fatigue damage caused by operation at any given stress level and the addition of damage increments to properly predict failure under conditions of spectrum loading. The first cumulative damage theory was proposed by Palmgren in 1924 and later developed by Miner in 1945. This linear theory, which is still widely used, is referred to as the Palmgren-Miner hypothesis or the linear damage rule. The theory may be described using the S-N plot shown in Fig. 18.39.

By definition of the S-N curve, operation at a constant stress amplitude S_1 will produce complete damage, or failure, in N_1 cycles. Operation at stress amplitude S_1 for a number of cycles n_1 smaller



Fig. 18.39 Illustration of spectrum loading where n_i cycles of operation are accrued at each of the different corresponding stress levels S_i , and the N_i are cycles to failure at each S_i .

than N_1 will produce a smaller fraction of damage, say D_1 . D_1 is usually termed the damage fraction. Operation over a spectrum of different stress levels results in a damage fraction D_i for each of the different stress levels S_i in the spectrum. When these damage fractions sum to unity, failure is predicted; that is,

Failure is predicted to occur if:

$$D_1 + D_2 + \dots + D_{i-1} + D_i \ge 1 \tag{18.43}$$

The Palmgren–Miner hypothesis asserts that the damage fraction at any stress level S_i is linearly proportional to the ratio of number of cycles of operation to the total number of cycles that would produce failure at that stress level; that is

$$D_i = \frac{n_i}{N_i} \tag{18.44}$$

By the Palmgren-Miner hypothesis, then, utilizing (18.44), we may write (18.43) as *Failure is predicted to occur if:*

$$\frac{n_1}{N_1} + \frac{n_2}{N_2} + \dots + \frac{n_{i-1}}{N_{i-1}} + \frac{n_i}{N_i} \ge 1$$
(18.45)

or

Failure is predicted to occur if:

$$\sum_{j=1}^{i} \frac{n_j}{N_j} \ge 1 \tag{18.46}$$

This is a complete statement of the Palmgren–Miner hypothesis or the linear damage rule. It has one important virtue, namely, *simplicity*; and for this reason it is widely used. It must be recognized, however, that in its simplicity certain significant influences are unaccounted for, and failure prediction errors may therefore be expected. Perhaps the most significant shortcomings of the linear theory are that no influence of the order of application of various stress levels is recognized, and damage is assumed to accumulate at the same rate at a given stress level without regard to past history. Experimental data indicate that the order in which various stress levels are applied does have a significant influence and also that damage rate at a given stress level is a function of prior cyclic stress history. Experimental values for the Miner's sum at the time of failure often range from about ¹/₄ to about 4, depending on the type of decreasing or increasing cyclic stress amplitudes used. If the various cyclic stress amplitudes are mixed in the sequence in a quasi-random way, the experimental Miner's sum more nearly approaches unity at the time of failure, with values of Miner's sums corresponding to failure in the range of about 0.6 to 1.6. Since many service applications involve quasi-random fluctuating stresses, the use of the Palmgren–Miner linear damage rule is often satisfactory for failure protection.

18.5.6 Stress Concentration

Failures in machines and structures almost always initiate at sites of local stress concentration caused by geometrical or microstructural discontinuities. These stress concentrations, or stress raisers, often lead to local stresses many times higher than the nominal net section stress that would be calculated without considering stress concentration effects. An intuitive appreciation of the stress concentration associated with a geometrical discontinuity may be developed by thinking in terms of "force flow" through a member as it is subjected to external loads. The sketches of Fig. 18.40 illustrate the concept. The rectangular flat plate of width w and thickness t is fixed at the lower edge and subjected to a total force F uniformly distributed along the upper edge. The dashed lines each represent a fixed quantum of force, and the local spacing between lines is therefore an indication of the local force intensity, or stress. In Fig. 18.40a the lines are uniformly spaced throughout the plate, and the stress σ is uniform and calculable as

$$\sigma = \frac{F}{wt} \tag{18.47}$$

In the sketch of Fig. 18.40b a flat rectangular plate of the same thickness has been subjected to the same total force F, but the plate has been made wider and notched to provide the same net section width w at the site of the notch. The lines of force flow may be visualized in very much the same way that streamlines would be visualized in the steady flow of a fluid through a channel with



Fig. 18.40 Intuitive concept of stress concentration: (a) without stress concentration; (b) with stress concentration.

the same shape as the plate cross section. No force can be supported across the notch, and therefore the lines of force flow must pass around the root of the notch. In so doing, force flow lines crowd together locally near the root of the notch, producing a higher force intensity, or stress, at the notch root. Thus, the local stress is raised or concentrated near the notch root, and even though the net section nominal stress is still properly calculated by (18.47), the actual local stress at the root of the notch may be many times higher than the calculated nominal stress. Many common examples of stress concentration may be cited, some of which are illustrated in Fig. 18.41. Discontinuities at the roots of gear teeth, at the corners of keyways in shafting, at the roots of screw threads, at the fillets of shaft shoulders, around rivet holes and bolt holes, and in the neighborhood of welded joints all constitute stress raisers that usually must be considered by a designer. The seriousness of the stress



Fig. 18.41 Some common examples of stress concentration: (a) gear teeth; (b) shaft keyway; (c) bolt threads; (d) shaft-shoulder; (e) riveted or bolted joint; (f) welded joint.

concentration depends on the type of loading, the type of material, and the size and shape of the discontinuity.

Stress raisers may be classified as being either *highly local* or *widely distributed*. Highly local stress raisers are those for which the volume of material containing the concentration of stress is negligibly small compared to the overall volume of the stressed member. Widely distributed stress raisers are those for which the volume of material containing the concentration of stress is a significant portion of the overall volume of the stressed member.

The theoretical elastic stress concentration factor, K_i , is defined to be the ratio of the actual maximum local stress in the region of the discontinuity to the nominal net section stress calculated by simple theory as if the discontinuity exerted no stress concentration effect; that is,

$$K_r = \frac{\text{actual maximum stress}}{\text{nominal stress}}$$
(18.48)

and the magnitude of K_i is found to be a function of geometry and type of loading, but not a function of material. It should be noted that this definition of K_i is valid only for stress levels within the elastic range, and must be suitably modified if stresses are in the plastic range.

The fatigue stress concentration factor, K_f , is defined to be the ratio of the effective fatigue stress at the root of the discontinuity to the nominal fatigue stress calculated as if the notch has no stress

concentrating effect. This definition is made for the high-cycle fatigue range and must be suitably modified for the low-cycle fatigue range. Thus, the fatigue stress concentration factor may be defined as

$$K_f = \frac{\text{effective fatigue stress}}{\text{nominal fatigue stress}}$$
(18.49)

Stress concentration factors are determined in a variety of different ways, including direct measurement of strain, utilization of photoelastic techniques, application of the principles of the theory of elasticity, and finite-element analysis. Numerical values for a wide variety of geometries and types of loading are presented in Ref. 46. Examples of typical charts for selection of stress concentration factor K_t are shown in Fig. 18.42. With a sharply notched specimen or machine part, it is clear that even a moderate load may produce actual stresses at the root of the notch that exceed the yield point of the material locally. The local yielding causes a redistribution of stresses, and the theoretical elastic stress concentration factor K, no longer describes the ratio of actual to nominal stresses accurately, since the actual maximum stress is relatively lower compared to the nominal stress than it would be if the material remained elastic. That is, the stress concentration factor is diminished in magnitude by local plastic flow, whereas the local strain is made larger than would be predicted by elastic theory.

Mathematical solutions of elastic-plastic stress and strain distributions around notches are relatively difficult to obtain, even using numerical solutions and digital computer techniques. One of the more successful approximations for stress concentration due to a circular hole in a very wide plate under tension has been given as47

$$K = 1 + 2\frac{E_s}{E}$$
(18.50)

where E = Young's modulus $\overline{E_s}$ = secant modulus K = stress concentration factor

Figure 18.43 illustrates values of stress concentration factor and strain concentration factor computed for a specific case by (18.50) and compared with values measured using very small strain gages. The agreement between calculated and measured values is good. Unlike the theoretical stress concentration factor K_t , the fatigue stress concentration factor K_t is a function of the material as well as geometry and type of loading. To account for the influence of material characteristics, a notch sensitivity index q has been defined to relate the actual effect of a notch on fatigue strength of a material to the effect that might be predicted solely on the basis of elastic theory. The definition of notch sensitivity index q is given by

$$q = \frac{K_f - 1}{K_f - 1} \tag{18.51}$$

where K_f = fatigue stress concentration factor K_r = theoretical stress concentration factor

q = notch sensitivity index valid for high-cycle fatigue range

The reason for subtracting unity from the numerator and denominator in this definition is to provide a scale for q that ranges from zero for no notch effect to unity for full notch effect. That is, for full notch effect K_f is equal to K_f . The notch sensitivity index has been found to be a function of both material and notch radius. Scatter in the experimental data is a serious problem in evaluating notch sensitivity index, as may be seen in the experimental results shown in Fig. 18.44. The notch sensitivity index for a range of steels and an aluminum alloy are shown in Fig. 18.45 for axial, bending, and torsional loading. These curves provide sufficient accuracy for most design applications and clearly demonstrate that the notch sensitivity index is a function of both the material and the notch root radius. An expression for fatigue stress concentration factor may be written from (18.51) as

$$K_f = q(K_t - 1) + 1 \tag{18.52}$$

where the theoretical elastic stress concentration factor K, may be determined on the basis of geometry and loading from handbook charts such as those depicted in Fig. 18.42. The notch sensitivity index q may be read from charts, such as the one shown in Fig. 18.45.

For uniaxial states of cyclic stress it is sometimes convenient to use K_f as a "strength reduction factor" rather than as a "stress concentration factor." That is, for uniaxial stressing only, a designer may choose to divide the fatigue limit by K_f rather than multiplying the applied nominal cyclic stress



Fig. 18.42 Stress concentration factors for a shaft with a fillet subjected to (a) bending, (b) axial load, or (c) torsion. (From Ref. 46, adapted with permission from John Wiley & Sons, Inc.)

by K_f . Although, conceptually, it is clearly more correct to think of K_f as a stress concentration factor, computationally, it is equivalent, and often simpler, to use K_f as a strength reduction factor. For multiaxial states of stress, however, K_f must be used as a stress concentration factor, since an appropriate value of K_f used as a strength reduction factor would be undefined.

The fatigue stress concentration factor (or strength reduction factor) determined from (18.52) is strictly applicable only in the high-cycle fatigue range, that is, for cycle lives of 10^5-10^6 cycles and larger. For ductile materials and static loads, effects of stress concentration may usually be neglected. Thus, in the intermediate- and low-cycle life range from a quarter cycle (static load) up to about 10^5-10^6 cycles, the stress concentration factor changes from unity to K_f . As shown in Fig. 18.46,



Fig. 18.42 (Continued)

the notched and unnotched *S*–*N* curves converge as they approach the low-cycle end of the range and coincide at the quarter cycle point *A*. Many materials exhibit fatigue stress concentration factors very near unity for lives less than 1000 cycles. Estimates of fatigue stress concentration factor are often made by constructing a straight line on a semilogarithmic *S*–*N* plot from the ultimate strength at a life of 1 cycle to the unnotched fatigue strength divided by K_f at a life of 10⁶ cycles. Such a straight line construction is shown in Fig. 18.46. The ratio of unnotched to notched fatigue strength values read at a specific life then becomes the estimate of fatigue stress concentration factor to be used for that life.

Finally, it should be noted that experimental investigations have indicated that for fatigue of *ductile materials* the fatigue stress concentration factor should be applied *only* to the *alternating component* of stress and not to the steady component of stress that exists in any nonzero mean cyclic stress. For



Fig. 18.43 Effect of plasticity on stress and strain concentration. (After Ref. 41.)



Fig. 18.44 An indication of scatter in the experimental determination of notch sensitivity index for alloys of aluminum and steel. (a) 24ST aluminum under completely reversed axial loading. (b) Steel alloys under alternating bending. (After Ref. 48; reprinted with permission from McGraw-Hill Book Company.)

fatigue loading of *brittle* materials, the stress concentration factor should be applied to the steady component as well.

18.5.7 Low-Cycle Fatigue

Two domains of cyclic loading were identified in the introduction to this section. One domain is that for which the cyclic loads are relatively low, strain cycles are confined largely to the elastic range, and long lives or high numbers of cycles to failure are exhibited. This behavior has traditionally been called high-cycle fatigue. The other domain is that for which the cyclic loads are relatively high, significant amounts of plastic strain are induced during each cycle, and short lives or low numbers of cycles to failure are exhibited if these relatively high loads are repeatedly applied. This type of behavior has been commonly called *low-cycle fatigue* or, more recently, cyclic strain-controlled fatigue.

18.5 FATIGUE AND STRESS CONCENTRATION



Fig. 18.45 Curves of notch sensitivity index versus notch radius for a range of steels and an aluminum alloy subjected to axial, bending, and torsional loading. (After Ref. 49; reprinted with permission from McGraw-Hill Book Company.)

tigue. The transition from low-cycle-fatigue behavior to high-cycle fatigue behavior generally occurs in the range from about 10^4 to 10^5 cycles, and some investigators define the low-cycle-fatigue range to be failure in 50,000 cycles or less.⁵⁰ Although the usual objective of an engineering designer is to provide long life, there are several circumstances in which the low-cycle fatigue or strain-controlled life response is of great importance. For example, in the design of high-performance devices such as missiles and rockets, the total design lifetime may be only a few hundred or a few thousand cycles from launch to delivery, and low-cycle-fatigue analysis and design methods are of direct interest. In



Fig. 18.46 S–N curves for notched and unnotched specimens subjected to completely reversed axial loading. (After Ref. 28, by permission of The MIT Press.)

the design of other high-performance devices, such as aircraft gas-turbine blades and wheels, nuclear pressure vessels and fuel elements, or steam turbine rotors and shells, the occurrence of occasional large mechanical or thermal transients during operation may give rise to significant damage accumulation due to a few hundred or a few thousand of these large cycles over the design lifetime, so that low-cycle-fatigue design methods are of great importance. Even if the loads on a machine or structure are nominally low, the material at the root of any critical notch will experience local plasticity that is cyclically strain controlled because of the constraints imposed by the surrounding bulk of elastic material, and the methods of low-cycle or strain-controlled fatigue will again be important in life prediction of such components. If a typical S-N curve is examined in the low-cycle-fatigue region, it will be found that over the range from a quarter cycle up to around 10^3 cycles the fatigue strength is nearly constant and close to the ultimate strength of the material. That is, the S-N curve remains relatively flat throughout this region in which the material cyclically experiences general yielding and gross plastic deformation. In this region of macroscopic plastic behavior the fatigue life is much more accurately described as a function of the cyclic strain amplitude rather than the cyclic stress amplitude. Stress-strain behavior under these circumstances is characterized by a stress-strain hysteresis loop, as shown in Fig. 18.47, with evidence of a measurable plastic strain in the specimen or machine part. This plastic behavior is typically nonlinear and history dependent, and it has been observed that the stress-strain response of most materials changes significantly with cyclic straining into the plastic range. Some materials exhibit cyclic strain hardening and others exhibit cyclic strain softening, as illustrated in Fig. 18.48. The stress-strain response of most materials changes significantly with applied cyclic strains early in life, but typically the hysteresis loops tend to stabilize so that the stress amplitude remains reasonably constant under strain control over the remaining large portion of the fatigue life. Based on the stable hysteresis loops for a family of different constant strain amplitudes, a curve passed through the tips of these hysteresis loops, as shown in Fig. 18.49, defines a "cyclic" stress-strain curve for the material. In Fig. 18.50 cyclic stress-strain curves are compared with static or monotonic stress-strain curves for several different materials.

The usual method of displaying the results of low-cycle fatigue tests is to plot the logarithm of strain amplitude or strain range versus the logarithm of number of cycles (or reversals) to failure. Sometimes the plastic strain amplitude or strain range is plotted, and sometimes the total strain amplitude or strain range is plotted as the ordinate. Early experimental investigations had indicated that if the plastic strain amplitude were plotted versus cycles to failure on a log-log plot, the data would approximate a straight line with a slope of about -0.5. Subsequent investigations have indicated that the slope ranges from about -0.7 to -0.5. Such plots seem to be remarkably similar for a wide range of materials,⁵¹ as indicated in Fig. 18.51. Experimental evidence accumulated by various investigators in recent years seems to indicate that the cyclic life is better related to total strain than to plastic strain, especially at the longer life end of the low-cyclic range. An example of a plot of strain amplitude versus life is shown in Fig. 18.52, separately showing the plastic strain amplitude and the total strain amplitude for a nickel-steel alloy.



Fig. 18.47 Hysteresis loop associated with cyclic loading that produces low-cycle-fatigue damage.



Fig. 18.48 Illustration of cyclic strain hardening and cyclic strain softening phenomena under strain control. (From Ref. 55, copyright ASTM; reprinted with permission.)

Data such as those shown in Fig. 18.51 led to the proposal of an empirical equation relating plastic strain $\Delta \epsilon_n$ to failure life N_f for completely reversed strain cycling under uniaxial stress conditions in the low-cycle-fatigue regime. The relationship, independently proposed by Manson⁵² and Coffin,53 may be expressed as

$$\frac{\Delta\epsilon_p}{2} = \epsilon_f' (2N_f)^c \tag{18.53}$$

where $\Delta \epsilon_p / 2$ = plastic strain amplitude

- ϵ'_{f} = fatigue ductility coefficient, defined as strain intercept at one load reversal, that is, at $2N_f = 1$ (see Fig. 18.53) $2N_f =$ total reversals to failure c = fatigue ductility exponent defined as slope of plastic strain amplitude versus reversals
- to failure curve on log-log plot (see Fig. 18.53)

Later work by many investigators, capitalizing on the Manson-Coffin equation of (18.53), has indicated that total strain amplitude, the sum of elastic strain amplitude plus plastic strain amplitude, may be better correlated to life. As shown in Fig. 18.53, the total strain amplitude is the sum of elastic plus plastic components. This has been modeled mathematically by Morrow et al.⁵⁴ as

$$\frac{\Delta\epsilon}{2} = \frac{\sigma'_f}{E} (2N_f)^b + \epsilon'_f (2N_f)^c \tag{18.54}$$

where constants b and σ'_t/E are the slope and one-reversal intercept of the plastic curve in Fig. 18.53, and constants c and ϵ'_{i} are the slope and one-reversal intercept of the plastic curve in Fig. 18.53. Typically,⁵⁵ b ranges from about -0.05 to -0.15 and c ranges from about -0.5 to -0.8.



Fig. 18.49 Cyclic stress-strain curve compared to monotonic stress-strain curve for SAE 4340 steel. (From Ref. 55, copyright ASTM; reprinted with permission.)

Although these constants are best evaluated from cyclic testing, they may be approximated from static properties, if fatigue data are unavailable. This may be done by taking σ'_f equal to true fracture strength σ_f , ϵ'_f equal to true fracture ductility ϵ_f , c equal to -0.6 and b equal to $-0.16 \log 2\sigma_f/\sigma_u$. However, actual fatigue data should be used where available.

From the schematic representation of (18.54) in Fig. 18.53, it may be noted that at short lives the plastic strain amplitude component dominates, whereas at longer lives the elastic strain amplitude component dominates. The point at which the elastic and plastic curves intersect has been called the "transition life." A thoughtful consideration of (18.54) and Fig. 18.53 leads to the observation that for design lives less than the transition life, materials with high fatigue ductility (high fracture ductility) are superior, whereas design lives greater than the transition life demand materials with high values of true fracture strength. This contradictory set of requirements, namely, high strength with high ductility, requires careful consideration on the part of the designer to match the appropriate material to the application, with careful attention to the magnitude of operating strain amplitude. This point is emphasized in Fig. 18.54 where three idealized materials are shown, one very high strength material (strong), one very ductile material (ductile), and one whose properties lie between the two extremes (tough). These curves cross at about 10^3 cycles (2 \times 10³ reversals) at a total cyclic strain amplitude of about 0.01. Thus, one would select the "strong" material for design life requirements greater than about 10³ cycles, pick the "ductile" material for design life requirements shorter than about 10³ cycles, and "optimize" with the "tough" material for spectrum loading of a more complicated nature. It is interesting to note that all types of materials seem to have about the same fatigue resistance for total strain amplitude around 0.01, corresponding to a failure life of about 10³ cycles. Presently attainable combinations of true fracture strength and true fracture ductility are illustrated in Fig. 18.55.

The effects of nonzero mean *strain* under low-cycle fatigue conditions have been studied by relatively few investigators, principally for the case of tensile mean strain. The experimental results

18.5 FATIGUE AND STRESS CONCENTRATION



- Incremental step

Material	Condition	0.2 percent yield strength, monotonic σ_{yp} /cyclic σ_{yp} ksi	Strain-bardening exponent n(monotonic)/n'(cyclic)	Cyclic behavior
OFHC copper	Annealed	3/20	0.40/0.15	Hardens
	Partial annealed	37/29	0.13/0.16	Stable
	Cold worked	50/34	0.10/0.12	Softens
2024 aluminum alloy	T4	44/65	0.20/0.11	Hardens
7075 aluminum alloy	T6	68/75	0.11/0.11	Hardens
Man-Ten steel	As-received	55/50	0.15/0.16	Softens and hardens
SAE 4340 steel	Quenched and tempered, 350 BHN	170/110	0.066/0.14	Softens
Ti-8Al-1Mo-1V	Duplex annealed	145/115	0.078/0.14	Softens and hardens
Waspaloy		79/102	0.11/0.17	Hardens
SAE 1045 steel	Quenched and tempered, 595 BHN	270/250	0.071/0.14	Stable
	Quenched and tempered, 500 BHN	245/185	0.047/0.12	Softens
	Quenched and tempered, 450 BHN	220/140	0.041/0.15	Softens
	Quenched and tempered, 390 BHN	185/110	0.044/0.17	Softens
SAE 4142 steel	As-quenched, 670 BHN	235/	0.14/	Hardens
	Quenched and tempered, 560 BHN	245/250	0.092/0.13	Stable
	Quenched and tempered, 475 BHN	250/195	0.048/0.12	Softens
	Quenched and tempered, 450 BHN	230/155	0.040/0.17	Softens
	Quenched and tempered, 380 BHN	200/120	0.051/0.18	Softens

Fig. 18.50 Cyclic stress-strain behavior of several materials. (From Ref. 56, copyright ASTM; reprinted with permission.)

of these few investigations indicate that the effect of a compressive mean strain on low-cycle fatigue life is essentially the same as the effect of a tensile mean strain if their magnitudes are the same. These results also indicate that mean strain effects are of primary importance only in the operating range where the *plastic* strain component dominates, that is, at design lives less than the transition life for the material.

The effects of nonzero mean *stress* are of primary importance only in the operating range where the *elastic* strain component dominates, that is, at design lives greater than the transition life of the material.



Fig. 18.51 Low-cycle-fatigue plot for several different materials. (From Ref. 51, copyright American Society for Metals, 1959; reprinted with permission.)



Fig. 18.52 Strain amplitude versus life for 18% Ni maraging steel, separately showing elastic, plastic, and total components of the strain. (From Ref. 55, copyright ASTM; reprinted with permission.)



Fig. 18.53 Schematic representation of elastic, plastic, and total strain amplitude versus fatigue life. (From Ref. 55, copyright ASTM; reprinted with permission.)



Fig. 18.54 Idealized representation of cyclic strain failure resistance of various types of materials. (After Ref. 55, copyright ASTM; reprinted with permission.)



Fig. 18.55 Monotonic fracture strength-ductility combinations presently attainable for various alloy classes. (From Ref. 55, copyright ASTM; reprinted with permission.)

Just as in the case of high-cycle fatigue, the assessment of low-cycle-fatigue damage under conditions where the cyclic strain amplitude ranges over a spectrum of values requires the use of a cumulative damage theory. Many investigators have studied cumulative damage effects in low-cycle fatigue with the general conclusion that a linear damage rule of the Palmgren–Miner type yields acceptable results if local stress–strain behavior can be accurately determined as a function of applied loads and if cycle counting is properly conducted.

If, for example, the fatigue-modified Neuber rule described in Section 18.5.8 is utilized to determine local stress and strain amplitude history and if the rain-flow cycle counting method, also described in Section 18.5.8, is properly applied to the local strain-time spectrum, results of using the Palmgren-Miner linear damage rule of (18.46) have been found to give acceptable life predictions. The range of values for $\Sigma(n/N)$ corresponding to failure under a variety of spectrum loading conditions has been reported to be from about 0.6 to about 1.6 for a variety of materials. This is a narrower range for low-cycle fatigue than has been observed for high-cycle fatigue. It should be noted, however, that for multiaxial states of stress, the linear damage rule may be much less reliable.

A detailed treatment of the effects of multiaxial states of stress on low-cycle-fatigue behavior is beyond the scope of this discussion, but proposed techniques have been presented for estimating low-cycle-fatigue life under these conditions. (See Refs. 57, 58, and 44.) The proposed technique involves the definition of an *equivalent stress* and an *equivalent total strain range*, both calculable from the multiaxial states of stress and strain. Life estimates based on the equivalent total strain range under conditions of a multiaxial state of stress can then be made from uniaxial state-of-stress low-cycle-fatigue data expressed as total strain range versus cycles to failure.

Although strains are often produced mechanically, it is perhaps even more usual to find the cyclic strains produced by a cyclic thermal field. If a machine part undergoes cyclic temperature changes and if the natural thermal expansions and contractions are either wholly or partially constrained, cyclic strains and stresses result. These cyclic strains produce fatigue failure just as if the strains were produced by external mechanical loading. Thus, all the results described earlier for low-cycle fatigue (and high-cycle fatigue if strains are all elastic) are at least qualitatively applicable to thermal fatigue as well. It must be recognized, however, that thermal fatigue problems involve not only all the complexities of mechanical loading but, in addition, all the temperature-induced problems as well.

While thermal low-cycle-fatigue phenomena and mechanical low-cycle-fatigue phenomena are very similar, and are mathematically expressible by the same types of equations, the use of mechanical low-cycle-fatigue results to predict thermal low-cycle-fatigue performance must be undertaken with care.

Some of the differences between thermally and mechanically induced low-cycle fatigue that give rise to apparent discrepancies in low-cycle-fatigue data for these two cases are (see pp. 270–272 of Ref. 58):

1. In thermal fatigue the plastic strain tends to become concentrated in the hottest regions of the body, since the yield point is locally reduced at these hottest regions.

18.5 FATIGUE AND STRESS CONCENTRATION

- 2. In thermal fatigue there is often a localized region of strain developed by virtue of plastic flow during the compressive branch of the strain cycle to produce a bulging at the hottest region. This is followed by a necking tendency adjacent to the bulge caused by plastic flow during the tensile branch of the strain cycle upon cooling.
- 3. Cyclic variations in temperature may, in and of themselves, have important effects on the material properties and ability to resist low-cycle-fatigue failure.
- 4. There may be interaction effects caused by superposition of simultaneous variations in temperatures and strains.
- 5. Rates at which the strain cycling is induced may have an important effect, since the testing speeds in thermal fatigue tests are often greatly different from the rates used in mechanical low-cycle-fatigue tests.

For these reasons caution is necessary in the prediction of thermal low-cycle-fatigue behavior from mechanical low-cycle-fatigue results or vice versa.

18.5.8 Three-Phase Approach for Fatigue Life Prediction

In recent years it has been recognized that the fatigue failure process involves three phases. A crack initiation phase occurs first, followed by a crack propagation phase; finally, when the crack reaches a critical size, the final phase of unstable rapid crack growth to fracture completes the failure process. The modeling of each of these phases has been under intense scrutiny. For the crack initiation phase, the most promising approach seems to be the "local stress–strain method."

The basic premise of the local stress-strain approach is that the local fatigue response of the material at the critical point, that is, the site of crack initiation, is analogous to the fatigue response of a small, smooth specimen subjected to the same cyclic strains and stresses.^{44,45,54,61} This concept is illustrated schematically in Fig. 18.56 for a simple notched plate under cyclic loading. The cyclic stress-strain response of the critical material may be determined from the characterizing smooth specimen through appropriate laboratory testing. To properly perform such laboratory tests, the local or experimental means. Thus, valid stress analysis procedures, finite-element modeling, or experimental strain measurements are necessary, and the ability to properly account for plastic behavior must be included. In performing smooth specimen tests of this type it must be recognized that the phenomena of cyclic hardening, cyclic softening, and cycle-dependent stress relaxation, as well as sequential loading effects and residual stress effects, may be experienced by the specimen as it accumulates fatigue damage presumed to be the same as at the critical point in the structural member being simulated. Some data have been accumulated to support the validity of this postulate.⁵⁹



FAILURE CONSIDERATIONS

Computer simulation of the smooth specimen simulation has also been shown to be feasible.^{44,60,61} To successfully utilize this simulation technique, it is necessary to have access to both monotonic and cyclic material properties, since the stress–strain response of most materials changes significantly with cyclic straining into the plastic range as already noted in Fig. 18.50. If cyclic materials response data are not available from the literature or an accessible data bank, it is necessary to perform enough smooth specimen testing to characterize the cyclic stress–strain response and fracture resistance of the material. With cyclic materials properties available, the computer simulation model for prediction of crack initiation must contain the following abilities:

- 1. To compute local stresses and strains, including means and ranges, from the applied loads and geometry of the structure.
- 2. To count cycles and associate mean and range values of stress and strain with each cycle.
- 3. To convert nonzero mean cycles to equivalent completely reversed cycles.
- 4. To compute fatigue damage in each cycle from stress and/or strain amplitudes and cyclic materials properties.
- 5. To compute damage cycle by cycle and sum the damage to give desired prediction of crack initiation.

To compute local stresses and strains from the external loading and geometry, Neuber's rule is often used in conjunction with the cyclic stress-strain properties and fatigue stress concentration factor. Neuber's rule may be modified for application to fatigue loading^{44,45,61,62} by utilizing the fatigue stress concentration factor K_f together with nominal stress range ΔS , local stress range $\Delta \sigma$, and local strain range $\Delta \epsilon$ to develop the expression

$$\frac{(K_f \Delta S)^2}{E} = \Delta \sigma \,\Delta \epsilon \tag{18.55}$$

All the terms on the left-hand side of (18.55) are known from the geometry and from loading and material properties of the structure. To resolve the right-hand term, an empirical expression for the cyclic stress-strain curve that is satisfactory for most engineering metals may be obtained by separating the cyclic strain amplitude $\Delta \epsilon/2$ into elastic and plastic components to yield

$$\frac{\Delta\epsilon}{2} = \frac{\Delta\epsilon_e}{2} + \frac{\Delta\epsilon_p}{2} = \frac{\Delta\sigma}{2E} + \left[\frac{\Delta\sigma}{2k'}\right]^{1/n'}$$
(18.56)

where k' and n' are the cyclic strength coefficient and cyclic strain-hardening exponent, respectively, determined from the intercept and slope of a log-log plot of cyclic stress amplitude versus cyclic plastic strain amplitude. Some values for n' are shown in Fig. 18.50. The use of (18.56), then, provides a means for computing local stresses and strains, when used in conjunction with (18.55). The need to include cyclic hardening and softening in the prediction model is dependent on the accuracy of other parts of the model, and in some cases these transient phenomena may be regarded as second-order effects. Cycle-dependent stress relaxation may be important in cases in which occasional large overload cycles produce large residual stresses in the structure that relax with additional cycles of local plastic strain.

To properly interpret complex load, stress, or strain versus time histories requires that an appropriate cycle counting method be used. The rain-flow cycle counting method, illustrated in Fig. 18.57, is probably more widely used than any other method. The strain-time history is plotted so that the time axis is vertically downward, and the lines connecting the strain peaks are imagined to be a series of roofs. Several rules are imposed on rain dripping down these roofs so that cycles and half-cycles are defined. Rain flow begins successively at the inside of each strain peak. The rain flow initiating at each peak is allowed to drip down and continue except that, if it initiates at a minimum, it must stop when it comes opposite a minimum more negative than the minimum from which it initiated. For example, in Fig. 18.57 begin at peak 1 and stop opposite peak 9, peak 9 being more negative than peak 1. A half-cycle is thus counted between peaks 1 and 8. Similarly, if the rain flow initiates at a maximum, it must stop when it comes opposite a maximum more positive than the maximum from which it initiated. For example, in Fig. 18.57 begin at peak 3 and 8. Similarly, if the rain flow initiates at a maximum, it must stop when it comes opposite a maximum more positive than the maximum from which it initiated. For example, in Fig. 18.57, the half-cycle beginning at peak 3 ends beneath peak 2, hus counting a half-cycle between peaks 2 and 3. A rain flow must also stop if it meets the rain from a roof above. For example, in Fig. 18.57, the half-cycle beginning at peak 3 ends beneath peak 2. Note that every part of the strain-time history is counted once and only once.

If cycles are to be counted for a "duty cycle" or a "mission profile" that is repeated until failure occurs, one complete strain cycle should be counted between the most positive and most negative peaks in the sequence, and other smaller complete cycles that are interruptions of this largest cycle should also be counted. This will be accomplished by the rain-flow method if the cycle counting is



Fig. 18.57 Example of rain-flow cycle counting method. (After Ref. 63.)

started at either the most positive or most negative peak in the sequence. Using this procedure to identify the maximum and minimum strain (or stress), the range and mean may be tabulated for each cycle in the stress-strain history.

The nonzero mean stress cycles may be converted to equivalent completely reversed cycles by utilizing either the modified Goodman equation or some empirical expression based on specific material data. Using Eq. (18.40), for a tensile mean stress the equivalent completely reversed stress σ_{eqC-R} is

$$\sigma_{eqC-R} = \frac{\sigma_a}{1 - \sigma_m / \sigma_u}, \qquad \sigma_m \ge 0$$
(18.57)

and for a compressive mean stress

$$\sigma_{eqC-R} = \sigma_a, \qquad \sigma_m \le 0 \tag{18.58}$$

To compute the fatigue damage in each cycle associated with the equivalent completely reversed stress and strain range, it is necessary to have available data for strain amplitude versus cycles to failure, N_f (or reversals to failure, $2N_f$), as illustrated in Fig. 18.53. An expression for total strain amplitude has already been given in Eq. (18.54) as a function of total number of cycles to failure.

Damage is summed by utilizing an appropriate cumulative damage theory. Using the procedure described in this section for prediction of crack initiation, the Palmgren-Miner linear damage hypothesis of (18.46) gives results that are as good as any other proposed technique. Thus, when the sum of cycle ratios becomes equal to unity, it is predicted that crack initiation has occurred. The prediction techniques described in this section are practical only with the help of a computer program designed to perform the tedious cycle-by-cycle analyses involved. Another practical difficulty lies in the definition and detection of an "initiated" crack. Nevertheless, the state of the art in prediction of

fatigue crack initiation has improved greatly during the past decade. The local stress-strain approach appears to be emerging as the initiation model of choice in making design predictions of initiation life.

A fatigue crack that has been initiated by cyclic loading, or any other preexisting flaw in the structure or material, may be expected to grow under sustained cyclic loading until it reaches the critical size from which it will propagate rapidly to catastrophic failure in accordance with the principles of fracture mechanics. For many structures or machine elements, the time required for a fatigue-initiated crack or a preexisting flaw to grow to critical size is a significant portion of the useful life.

The fatigue crack growth rate da/dN has been found to often correlate with the crack-tip stress intensity factor range such that

$$\frac{da}{dN} = g(\Delta K) \tag{18.59}$$

where ΔK is the stress intensity factor range, computed using the maximum and minimum applied stresses with $\Delta K = K_{max} - K_{min}$. Most crack growth rate data produced have been characterized in terms of ΔK and plotted either as a log-log or log-linear function of ΔK . For example, Fig. 18.58 illustrates the dependence of fatigue crack growth on stress intensity factor. The crack growth rate, indicated by the slope of the *a* versus *N* curves, increases with both the applied load and crack length. Since the crack-tip stress intensity factor range also increases with applied load and crack length, it is clear that the crack growth rate is related to the applied stress intensity factor range.

To plot the data of Fig. 18.58 in terms of the stress intensity factor range and crack growth rate, the crack growth rate is determined from the slope of the *a* versus *N* curves between successive data points. Corresponding values of ΔK are then computed from the applied load range and mean crack length for each interval. The results of this procedure are shown in Fig. 18.59 for the data presented in Fig. 18.58. It should be noted that all the curves of Fig. 18.58 incorporate themselves into a single curve in Fig. 18.59 through use of the stress intensity factor, and the curve of Fig. 18.59 is therefore applicable to any combination of cyclic stress range and crack length for released loading (R = 0) on specimens of this geometry. Different geometries under different applied stresses will exhibit



Fig. 18.58 Effect of cyclic-load range on crack growth in Ni–Mo–V alloy steel for released tension loading. (From Ref. 73, copyright Society for Experimental Stress Analysis, 1971; reprinted with permission.)



Fig. 18.59 Crack growth rate as a function of stress-intensity range for Ni–Mo–V steel. (From Ref. 73, copyright Society for Experimental Stress Analysis, 1971; reprinted with permission.)

identical crack-tip stress fields if the stress intensity factors are equal. Thus, because the stress intensity factor characterizes the state of stress near the crack tip, the fatigue crack growth rate correlation shown in Fig. 18.59 is applicable to any cyclically loaded component with R = 0 manufactured using the same material. This allows crack growth data generated from simple laboratory specimens to be utilized for approximate crack growth predictions in more complex geometries.

Fatigue crack growth rate data similar to that shown in Fig. 18.59 have been reported for a wide variety of engineering metals. The linear behavior observed using log-log coordinates suggests that Eq. (18.59) may be generalized as follows

$$\frac{da}{dN} = C(\Delta K)^n \tag{18.60}$$

where *n* is the slope of the log da/dN versus log ΔK plot and *C* is the da/dN value found by extending the straight line to a ΔK value of unity. This relationship was first proposed by Paris.⁶⁶ The empirical parameters *C* and *n* are a function of material type, *R* ratio, thickness, temperature, environment, and loading frequency. Standard methods have been established for conducting fatigue

crack growth tests,⁶⁴ and fatigue crack growth rate data may be found in References 10, 11, 13, 14, 15, and 16. Many other fracture mechanics-based empirical correlations other than Eq. (18.60) have been proposed, some of which are discussed by Schijve.⁶⁵ An extensive overview of the fatigue crack propagation problem is provided by Pook.⁷¹

Given an initial crack of length a_i with this crack either initiated by cyclic loading or initially present as a flaw, Eq. (18.60) may be integrated to give the number of cycles N required to propagate a crack to a size a_N such that

$$N = \int_{a_i}^{a_N} \frac{da}{C(\Delta K)^n}$$
(18.61)

Given that ΔK is a function of crack length, numerical integration techniques may be required to compute N. An approximate procedure and several idealized examples are presented by Parker.²³

It must be emphasized that Eqs. (18.60) and (18.61) are applicable only to region II crack growth, as illustrated in Fig. 18.60. Region I of Fig. 18.60 exhibits a threshold ΔK_{th} below which the crack will not propagate. Region III corresponds to the transition into the unstable regime of rapid crack extension. In this region, crack growth rates are large and the number of cycles associated with growth in this region small.

Given an initial crack of length a_i , from Eq. (18.61), the number of cycles required to grow a crack to a critical length a_c such that rapid crack extension would be predicted may be approximated as

$$N_{p} = \int_{a_{i}}^{a_{c}} \frac{da}{C(\Delta K)^{n}}$$
(18.62)

Assuming an initial crack that has been initiated by cyclic loading, the crack propagation life N_p given by Eq. (18.62) may then be added to the crack initiation life N_i computed using the local stress-strain approach previously discussed to obtain an estimate of the total fatigue life N with

$$N = N_i + N_p \tag{18.63}$$

Such estimates are highly sensitive to the length of the initial crack a_i . While the local stress-strain approach may be used to compute the number of cycles N_i required to initiate a crack, the corresponding length of this initiated crack is not given. No consensus has yet been reached regarding the length of this initiated crack. A size of between 2 and 3 mm has been suggested,⁶¹ as cracks of this



stress intensity factor range, log scale

Fig. 18.60 Schematic representation of fatigue crack growth rate.

size normally exist at fracture in the small laboratory specimens used to generate material property data for the local stress–strain approach. An alternative approach would involve the assumption of a preexisting material or manufacturing defect such that $N_i \approx 0$. For example, such an assumption is often made during the analysis of welded joints.⁶⁷ If nondestructive techniques are employed, a reasonable assumption for the size of this initial defect would be the smallest flaw that could avoid detection.

Crack growth rates determined from constant amplitude cyclic loading tests are approximately the same as for random loading tests in which the maximum stress is held constant but mean and range of stress vary randomly. However, in random loading tests where the maximum stress is also allowed to vary, the sequence of loading cycles may have a marked effect on crack growth rate, with the overall crack growth being significantly higher for random loading spectra.

Many investigations have shown a significant delay in crack propagation following intermittent application of high stresses. That is, fatigue damage and crack extension are dependent on preceding cyclic load history. This dependence of crack extension on preceding history and the effects upon future damage increments are referred to as *interaction* effects. Most of the interaction studies conducted have dealt with *retardation* of crack growth as a result of the application of occasional tensile overload cycles. Retardation may be characterized as a period of reduced crack growth rate following the application of a peak load or loads higher and in the same direction as those peaks that follow.

The modeling of interaction effects requires consideration of crack tip plasticity and its subsequent influence. In metals of all types, cracks will remain closed or partially closed for a portion of the applied cyclic load as a consequence of plastically deformed material left in the wake of the growing crack. Under cyclic loading, crack growth will occur during the loading portion of the cycle. Given that a plastic zone exists at the crack tip prior to crack extension, as the material at the crack tip separates, the newly formed crack surfaces will exhibit a layer of plastically deformed material, closing the crack while the applied stress remains tensile. This phenomenon is known as plasticity-induced fatigue crack closure and was first discussed by Elber.⁶⁹ Upon reloading during the following cycle, crack growth will not continue unless the applied load is sufficiently large such that the compressive stresses acting along the crack surfaces are overcome and the crack is fully opened. This load is known as the crack growth rates under both constant amplitude and spectrum loading. Further information regarding crack closure may be found in Refs. 1, 6, 18, 42, 43, 61, 65, and 70.

Discussion to this point has been limited to the growth of through-thickness cracks under mode I loading. While mode I loading is often dominant, under the most general circumstances the applied cyclic loads will generate stress intensity factor ranges ΔK_I , ΔK_{II} , and ΔK_{III} at the crack tip, and *mixed mode* fatigue crack growth must be considered. Modeling methodologies for mixed mode fatigue crack growth are discussed in Refs. 19 and 23. In addition, fatigue cracks in machine elements and structures are often not through-thickness cracks but rather surface cracks that extend partially through the thickness. Such surface cracks are often semi-elliptical in shape and the analysis of these cracks is considerably more complicated. Information regarding surface cracks may be found in Refs. 6 and 68.

Recent research has suggested that when fatigue cracks are small, crack growth rates are larger than would be predicted using Eq. (18.60) for a given ΔK .⁷² Small-crack behavior is often important, as a significant portion of the fatigue life may be spent in the small-crack regime. The fatigue crack propagation life N_p will also be influenced by the presence of residual stresses such as might exist as a consequence of welding, heat treatment, carburizing, grinding, or shot-peening. Compressive residual stresses are beneficial, decreasing the rate of fatigue crack growth and increasing propagation life. While approximate methodologies exist for incorporating the effects of residual stress within fatigue crack growth predictions,⁶¹ residual stress distributions are often difficult to characterize.

Reasonable design estimates for fatigue life may be obtained using Eq. (18.63). However, the many uncertainties typically associated with fatigue life predictions emphasize the essential requirement to conduct full-scale fatigue tests to provide acceptable reliability.

18.5.9 Service Spectrum Simulation and Full-Scale Testing

To achieve a reliable fatigue-resistant design configuration, a realistic service loading simulation test or a full-scale fatigue test will generally be required. For service load simulation testing it is essential that both the test specimen and the loading spectrum, including sequence, be representative of actual service conditions. This means that the specimen should be an actual component, a complete structural subassembly, or a complete full-scale machine or structure. It also means that an exact simulation of the service load-time history would be best, if known. Usually, however, an estimated load-time history must be generated on the basis of a statistical representation of similar mission loading spectra obtained from similar instrumented articles already in actual use.

In designing service loading simulation spectra, special attention must be given to the highest load levels to be incorporated because they may exert a major influence on crack propagation and fatigue life. As noted in Section 18.5.8, the life-enhancing retardation phenomenon associated with

large load peaks can be extremely important. It should be recognized that the largest peak loads in service will vary from article to article on a statistical basis, so that some articles will experience the maximum load peak more than once, whereas others will never see this load level. For these reasons it is usual to truncate the design loading spectrum, discarding all load levels that occur fewer than 10 times during the projected design lifetime. Although it may seem intuitively wrong to discard the high load peaks to achieve a more critical simulation test spectrum, it must be realized that the crack retardation effect of the occasional high loads gives a longer test life than if the high loads were omitted. Thus, it is conservative to omit the 10 highest loads from the simulation test spectrum. Likewise, truncation is utilized in analyses for establishing suitable inspection intervals for aircraft so that propagating cracks will have a higher probability of being detected before they approach critical size.

The application of failsafe or limit loads at regular intervals during a service spectrum simulation or full-scale fatigue test *must be avoided*, since they may contribute to crack growth retardation that is not typical of the actual flight spectrum. Limit loading should be applied only at the end of the fatigue test.

Low-amplitude cycles are often omitted from service simulation testing to save time. However, it must be recognized that these cycles may contribute to fatigue crack nucleation through the fretting process, and omitting these low-amplitude cycles, and the fretting they would normally produce, may result in unsafe simulation test predictions of service fatigue life.

Full-scale fatigue testing of an article such as a newly designed aircraft is extremely expensive. Such tests generally would be regarded as accelerated tests, where flight simulation testing over a 6to 12-month testing period is designed to represent 10 or more years of actual service history. Flight simulation testing of an aircraft component on a modern closed-loop fatigue testing machine may be accomplished in one or two weeks or less. The benefits of full-scale testing include (1) discovery of fatigue-critical elements and design deficiencies, (2) determination of times to detectable cracking, (3) obtaining data on crack propagation, (4) determination of remaining safe life with cracks, (5) determination of residual strength, (6) establishment of proper inspection intervals, and (7) development of repair methods. Factors that may influence the full-scale fatigue simulation test results include loading rate, environmental factors, statistical scatter, and loading spectrum deviations. Actual service life is usually shorter than the simulation test life, sometimes by factors of as much as 2 to 4. However, in recent years the agreement between full-scale flight simulation testing and in-service experience has improved significantly. Full-scale flight simulation testing is often continued over the long term so that fatigue failures in the test will lead the fleet experience by enough time to redesign and install whatever modifications are required to prevent catastrophic fleet failures in service before they occur.

18.5.10 Damage Tolerance and Fracture Control

The concept of "damage-tolerant" structure, which has developed primarily within the aerospace industry, is characterized by structural configurations that are designed to minimize the loss of aircraft because of the propagation of undetected flaws, cracks, or other similar damage. There are two major design objectives that must be met to produce a damage-tolerant structure. These objectives are controlled safe flaw growth, or safe life with cracks, and positive damage containment, which implies a safe remaining or residual strength. These objectives should not be considered as separate or distinct requirements, however, because it is only by their judicious combination that effective fracture control can be achieved. Furthermore, it must be emphasized that damage-tolerant design is not a substitute for a careful fatigue analysis and design as discussed earlier, because the achievement of "fatigue quality" through careful stress analysis, geometry selection, detail design, material selection, surface finish, and workmanship is a necessary prerequisite to effective damage-tolerant design and fracture control.

The general goals of damage-tolerant design and fracture control include the selection of fracture resistant materials and manufacturing processes, the design for inspectability, and the use of damage-tolerant structural configurations such as multiple load paths or crack stoppers, in addition to the usual rules of good design practice.

In the application of the fracture control philosophy, the basic assumption is made that flaws do exist even in new structures and that they may go undetected. The first major requirement for damage tolerance, therefore, is that any member in the structure, including each element of a redundant load path group, must have a safe life with assumed cracks present. For any specific application the primary factors influencing the design include the type or class of structure, the quality of the nondestructive inspection, (NDI) techniques used in production assembly, the accessibility of the structure to inspection, the assurance that the member will be inspected on schedule when in service, and the probability that a flaw of subcritical size will go undetected even though periodic in-service inspections are made on schedule.

Most structural arrangements may be classified according to load path as class 1, single load path; class 2, single primary load path with auxiliary crack arrest features; or class 3, multiple or redundant load path. Each of these structural classes is illustrated in Fig. 18.61. Clearly, for the class 1 structure it is essential to satisfy the safe-life-with-cracks requirement, because failure is catastrophic. For class



2 structures, including pressurized cabins and pressure vessels, relatively large amounts of damage may be contained by providing tear straps or stiffeners. There is usually a high probability of damage detection for a class 2 structure because of fuel or pressure leakage, that is, "leak-before-break" design is characteristic of class 2 structures. Class 3 structures are usually designed to provide a specified percentage of the original strength, that is, a specified residual strength, during and subsequent to the failure of one element. This is often called "failsafe" type of structure. However, the preexisting flaw concept requires that all members, including every member of a multiple load path structures because it is appropriate to take a larger risk of operating with cracks if multiple load paths are available.

The development of inspection procedures is an important part of any fracture control program. Appropriate inspection procedures must be established for each structural element, and regions within elements may be classified with respect to required NDI sensitivity. Inspection intervals are established on the basis of crack growth information assuming a specified initial flaw size and a "detectable" flaw size that depends on the NDI procedure. Inspection intervals are established to ensure that an undetected flaw will not grow to critical size before the next inspection, with a comfortable margin of safety. The intervals are usually picked so that two inspections will occur before any crack will reach critical size.

A good fracture-control program should encompass and interact with design, materials selection, fabrication, inspection, and operational phases in the development of any high-performance engineering system.

18.6 CREEP AND STRESS RUPTURE

Creep in its simplest form is the progressive accumulation of plastic strain in a specimen or machine part under stress at elevated temperature over a period of time. Creep failure occurs when the accumulated creep strain results in a deformation of the machine part that exceeds the design limits. *Creep rupture* is an extension of the creep process to the limiting condition where the stressed member actually separates into two parts. *Stress rupture* is a term used interchangeably by many with creep rupture; however, others reserve the term stress rupture for the rupture termination of a creep process in which steady-state creep is never reached, and use the term creep rupture for the rupture termination of a creep process in which a period of steady-state creep has persisted. Figure 18.62 illustrates these differences. The interaction of creep and stress rupture with cyclic stressing and the fatigue process has not yet been clearly understood but is of great importance in many modern high-performance engineering systems.

Creep strains of engineering significance are not usually encountered until the operating temperatures reach a range of approximately 35-70% of the melting point on a scale of absolute temperature. The approximate melting temperature for several substances is shown in Table 18.2.

Not only is excessive deformation due to creep an important consideration, but other consequences of the creep process may also be important. These might include creep rupture, thermal relaxation, dynamic creep under cyclic loads or cyclic temperatures, creep and rupture under multiaxial states of stress, cumulative creep effects, and effects of combined creep and fatigue.