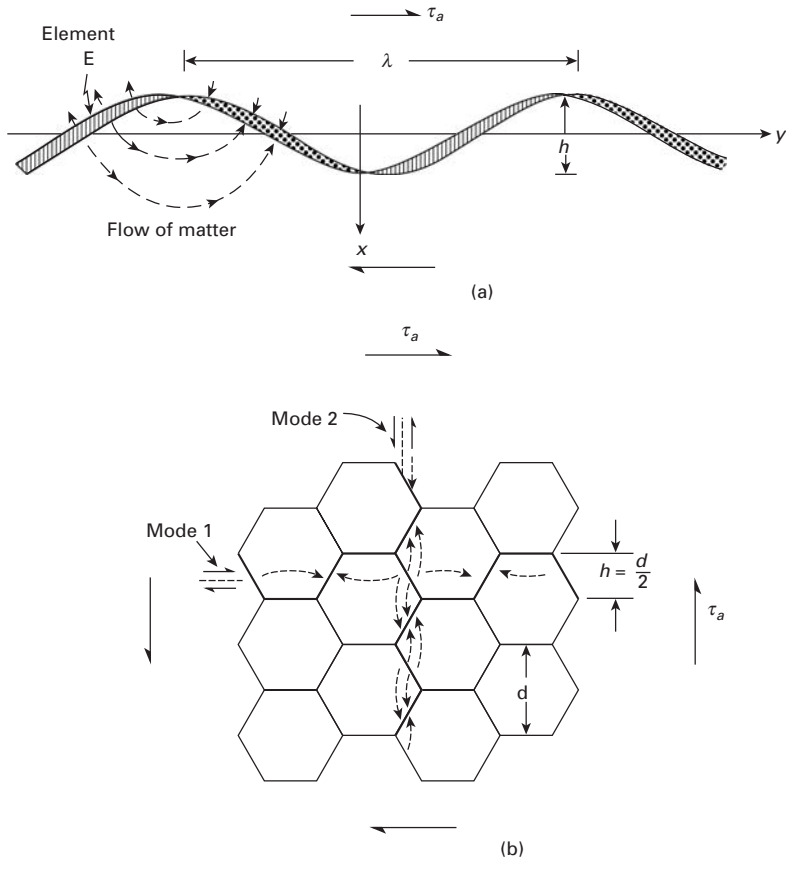


Fig. 13.18 (a) Steady-state grain-boundary sliding with diffusional accommodations. (b) Same process as in (a), in an idealized polycrystal; the dashed lines show the flow of vacancies. (Reprinted with permission from R. Raj and M. F. Ashby, *Met. Trans.*, 2A (1971) 1113.)

13.7 Grain-Boundary Sliding

Grain-boundary sliding usually does not play an important role during primary or secondary creep. However, in tertiary creep it does contribute to the initiation and propagation of intercrystalline cracks. Another deformation process to which it contributes significantly is superplasticity; it is thought that most of the deformation in superplastic forming takes place by grain-boundary sliding.

The grain-boundary sliding rate is controlled by the accommodating processes where the sliding surface deviates from a perfect plane. One can readily see that we cannot have a perfect plane defined by the boundaries between different grains; we cannot look separately at the sliding between two grains having a common interface. The requirements of strain compatibility are such that we have to model the interface as sinusoidal, as is depicted in Figure 13.18. The applied stress τ_a can produce sliding only if it is coupled with diffusional flow that transports material (or vacancies) over a maximum distance of λ, the wavelength of the irregularities. Figure 13.18(b) shows the same effect in a polycrystalline aggregate. The individual grain boundaries

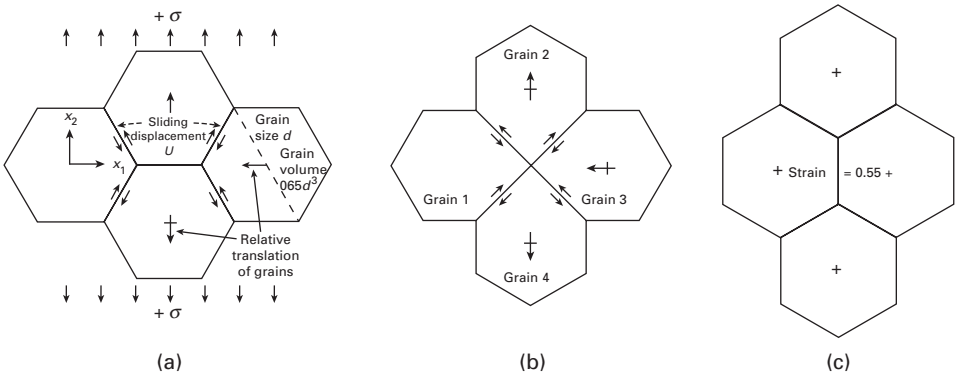


Fig. 13.19 Grain-boundary sliding assisted by diffusion in Ashby–Verrall's model. (Reprinted with permission from M. F. Ashby and R. A. Verrall, *Acta Met.*, 21 (1973) 149.)

are translated by a combination of sliding and diffusional flow under the influence of the applied stress.

The manner in which the individual grains move and change their relative positions by sliding and diffusional accommodation is shown in Figure 13.19. The sliding of grains under the influence of σ, coupled with minor changes in shape, makes possible the sequence (a)–(b)–(c), which results in a strain of 0.55; the unique feature of this mechanism is that the sequence is accomplished with relatively little strain *within* the grains.

13.8 Deformation-Mechanism (Weertman–Ashby) Maps

Deformation-mechanism maps, so named after two people who first introduced them, are a graphical description of creep, representing the ranges in which the various deformation modes are rate-controlling steps in the stress-versus-temperature space. Weertman–Ashby plots assume, for simplicity, that there are some independent and distinguishable ways by which a polycrystal can be deformed, but still retain its crystallinity:

1. Above the theoretical shear strength, plastic flow of the material can take place without dislocations, by simple glide of one atomic plane over another.
2. Movement of dislocations by glide.
3. Dislocation creep, including glide and climb, both being controlled by diffusion.
4. Nabarro–Herring creep.
5. Coble creep.

The theories developed for these different modes of deformation without loss of crystallinity propose constitutive equations that are used in the establishment of the ranges involved. Figure 13.20 shows a typical map for silver. The theoretical shear stress is approximately equal to $G/20$ and is practically independent of temperature. A small dependence on temperature is exhibited by G and is built into the ordinate of

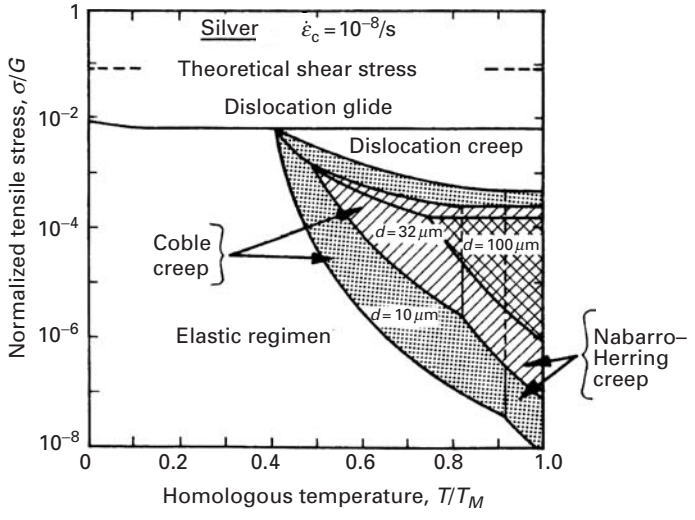


Fig. 13.20 Weertman–Ashby map for pure silver, established for a critical strain rate of 10^{-8} s^{-1}; it may be seen how the deformation-mechanism fields are affected by the grain size. (Adapted with permission from M. F. Ashby, *Acta Met.*, 20 (1972) 887.)

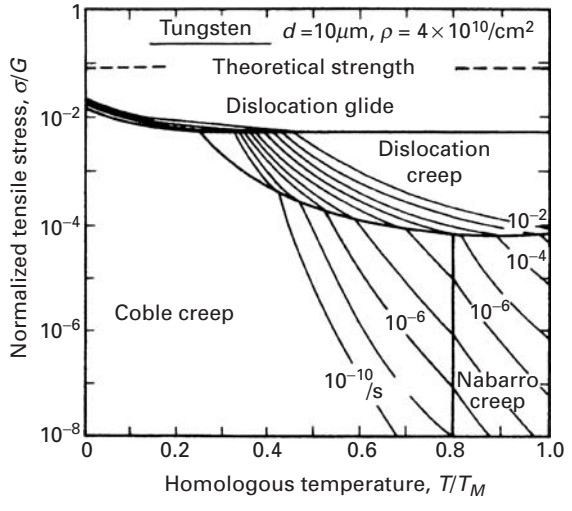


Fig. 13.21 Weertman–Ashby map for tungsten, showing constant strain-rate contours. (Reprinted with permission from M. F. Ashby, *Acta Met.*, 20 (1972) 887.)

the figure. For values of σ/G between 10^{-1} and 10^{-2}, slip by dislocation movement is the controlling mode at all temperatures. It can be seen that the grain size affects the extent of the fields. Three grain sizes are represented: 10, 32, and 100 μm. The fields also depend on the strain rate. The map shown in Figure 13.21 was made for a strain rate of 10^{-8} s^{-1}. The Coble and Nabarro–Herring mechanisms, especially, are affected by the grain size, because of their nature.

Deformation-mechanism maps have technological applications. Consider, for example, a turbine blade operated in a temperature and stress range that is known. The specific stress–temperature profile can be plotted on a deformation-mechanism map in the form of a line. Different parts of the blade undergo different deformation modes. These modes, the rate of creep of each portion, and the respective constitutive equation can be read from the map. Multiaxial stress states can be resolved by calculating the maximum shear stress

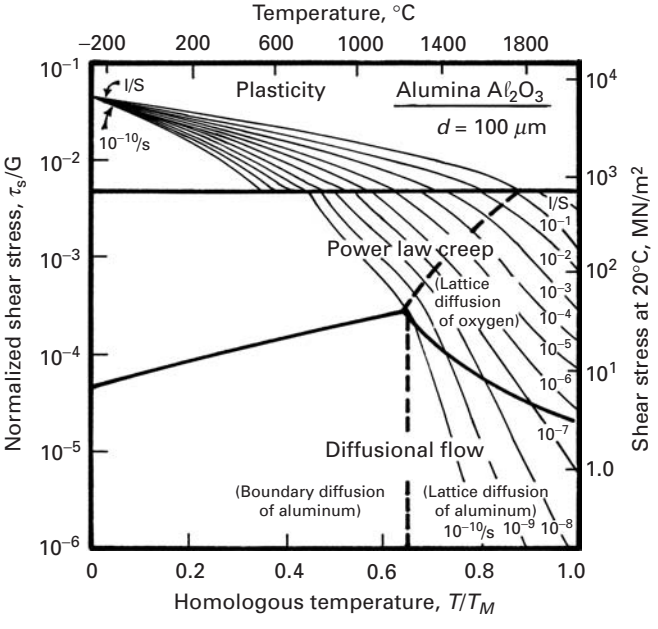


Fig. 13.22 Weertman–Ashby map for Al_2O_3 with a grain size of 100 μm. (Adapted from H. J. Frost and M. F. Ashby, *Deformation-Mechanism Maps*, (New York: Pergamon Press, 1982), p. 100.)

or the effective stress. A strengthening mechanism is helpful only if it retards the creep rate in the correct portion of the map. For instance, dispersion-hardening is effective in controlling dislocation glide and climb, but cannot effectively stop Nabarro–Herring or Coble creep.

From the deformation-mechanism map, we can, in addition to determining the dominant mechanism for a certain combination of stress and temperature, find the strain rate (creep rate) that will result. For this, we have to apply the appropriate constitutive equations and plot the constant strain-rate contours. This is shown in Figure 13.21 for tungsten. The lines allow ready identification of the creep rate. The region in Figure 13.20 consisting of the elastic regimen is occupied by Coble creep in Figure 13.21. The reason for this is that Figure 13.20 applies to one constant strain rate (10^{-8} s^{-1}), whereas Figure 13.21 is built for a whole range of strain rates. Hence, at a strain rate of 10^{-8} s^{-1}, the metal might respond elastically, whereas at a strain rate orders of magnitude lower, Coble creep becomes significant.

Similar maps can be built for ceramics, and a representative map is shown in Figure 13.22. The different domains, as well as the curves for constant strain rates (from 10^{-10}/s to 1/s), are illustrated in the plot. Note that the different diffusing ions (Al^{3+} and O^{2-}) have to be considered.

13.9 Creep-Induced Fracture

Figure 13.23 shows fracture by nucleation and growth of voids at the grain boundaries. Whereas at low and moderate temperatures metals usually fail by transgranular void or crack formation, at high

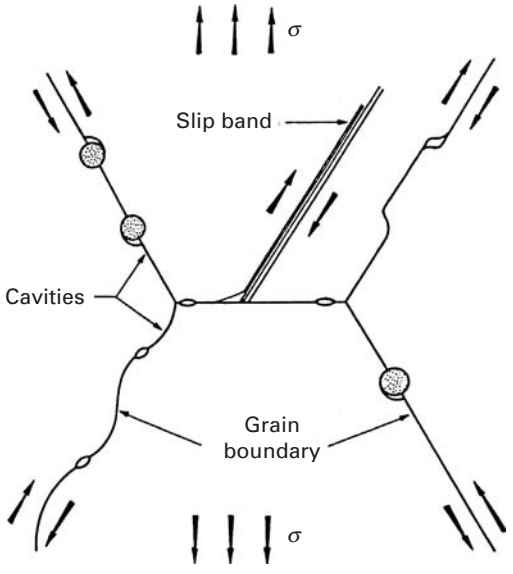


Fig. 13.23 Mechanisms of intergranular nucleation. (From W. D. Nix and J. C. Gibeling, in *Flow and Fracture at Elevated Temperatures*, ed, R. Raj (Metals Park, OH: ASM, 1985).)

temperatures, and especially after creep and superplastic forming, intergranular cavities start forming; they subsequently grow and coalesce, leading to a fracture morphology which is readily distinguishable: observation by scanning electron microscope reveals entire grains on the fracture surface. There are several reasons why cavities (voids and cracks) form at the grain boundaries.

- Impurity segregation at grain boundaries leads to a weakening of these regions. Since the boundary has a different composition from the matrix, under these conditions, we may have:

$$(T_m)_{gb} < (T_m)_\ell \,,$$

  where T_m is the melting point and gb and ℓ refer to the grain boundary and lattice, respectively. This leads to a selective weakening of the boundaries, at high temperatures.
- Interaction of grain-boundary sliding with geometrical (compatibility) constraints. We saw in Section 13.7 that grain-boundary sliding needs to be coupled with plastic deformation of cavities at grain boundaries due to interaction with grain-boundary inclusion. This is due to a step in the grain boundary, and to waviness of the boundary.
- Slip band–boundary interaction. At higher stresses dislocation glide is important in creep, and the intersection of slip bands with boundaries leads to stress concentrations, which may result in opening up of wedges.
- Enhanced grain-boundary diffusion. The combined effect of stress concentrations and higher grain-boundary diffusion coefficients (can be several orders of magnitude higher than lattice diffusion) leads to void growth at the grain boundaries, which is much faster than in the grain interiors.

As a result of the nucleation and growth of cavities, the creep rate is accelerated even under constant stress testing, and the creep curve starts deviating from stage II. (See Figure 13.1; this marks the onset of stage III, or tertiary creep.)

Creep failure is governed by maximum creep strain in many cases. This is the basis for the Monkman–Grant equation (Equation 13.9, Section 13.2):

$$\dot{\varepsilon}_s t_r = k,$$

where $\dot{\varepsilon}_s$ is the creep rate in the steady state, t_r is the time-to-rupture, and k is a constant. The above expression is obtained from:

$$\int_0^{t_r} \dot{\varepsilon}_s dt = \dot{\varepsilon}_s t_r = \varepsilon_r.$$

We assume that ε_r, the ruptuse strain, is a constant for a material. For nickel, the rupture strain, is equal to 4×10^{-2}.

Raj and Ashby[15] developed a model to predict the time-to-rupture time, t_r, as a function of a number of material parameters. Their expression has the form:

$$t_r = \frac{3\pi^{1/2}}{32} \frac{kT}{\Omega \delta D_{gb}} \frac{\lambda^3}{\sigma} \frac{f_v(\alpha)}{f_b(\alpha)} \int_{A_{\min}}^{A_{\max}} \frac{dA}{f(A)},$$

where the symbols have the following meanings: A = areal fraction of grain boundaries occupied by cavities; δ = grain-boundary thickness; Ω = atomic volume; D_{gb} = grain-boundary diffusion coefficient; λ = spacing between cavities; T = temperature in K; k = Boltzmann's constant.

The function $f(A)$ in the integral is equal to:

$$f(A) = \frac{\left[1 - \left(\frac{A_i}{A}\right)^{1/2}\right](1 - A)}{A^{1/2}\left[\frac{1}{2}\ln\left(\frac{1}{A}\right) - \frac{3}{4} + A\left(1 - \frac{A}{4}\right)\right]}.$$

Rupture occurs when A reaches a maximum value $A_{\max}$. The functions $f_v(\alpha)$, $f_B(\alpha)$ represent the dihedral angle (voids at grain boundaries are lens-shaped and not spherical) and the geometry of voids, respectively. When A reaches a value equal to $A_{\max}$, the specimen fails.

It may be seen that the Raj–Ashby equation is consistent with the Monkman–Grant equation when the stress exponent in creep is equal to one. This occurs for Nabarro–Herring, Coble, and Harper–Dorn creep (Equations 13.12, 13.13, and 13.14). For instance, substituting the Coble creep equation into the Raj–Ashby equation, we obtain:

$$\dot{\varepsilon}_s t_r = \frac{3\pi^{1/2}}{32} \frac{A_c}{\Omega} \frac{b^3 \lambda^3}{d^3} \frac{f_v(\alpha)}{f_b(\alpha)} \int_{A_{\min}}^{A_{\max}} \frac{dA}{f(A)}.$$

The right-hand side of this equation consists of a number of material parameters. Thus, the product $\dot{\varepsilon} t_r$ should be a constant, which should, in theory, be equal to the rupture strain, ε_r.

[15] R. Raj and M. F. Ashby, *Acta Mater.*, 23 (1975) 653.

Example 13.4

From Figure E13.4 for Zircaloy under two conditions determine the Monkman–Grant rupture strain at three different strain rates. The failure strain is 0.1. Is the Monkman–Grant equation obeyed?

Solution: We take the three strain rates:

$$\dot{\varepsilon}_1 = 5x10^{-5},$$
$$\dot{\varepsilon}_2 = 5x10^{-4},$$
$$\dot{\varepsilon}_3 = 5x10^{-3}.$$

The following rupture times are obtained from the plot:

$$t_{r1} = 2,000s,$$
$$t_{r2} = 200s,$$
$$t_{r3} = 20s.$$

The products are:

$$\dot{\varepsilon}_1 t_{r1} = \dot{\varepsilon}_2 t_{r2} = \dot{\varepsilon}_3 t_{r3} = 10^{-1}$$

This is equal to the rupture strain. The Monkman-Grant equation is obeyed.

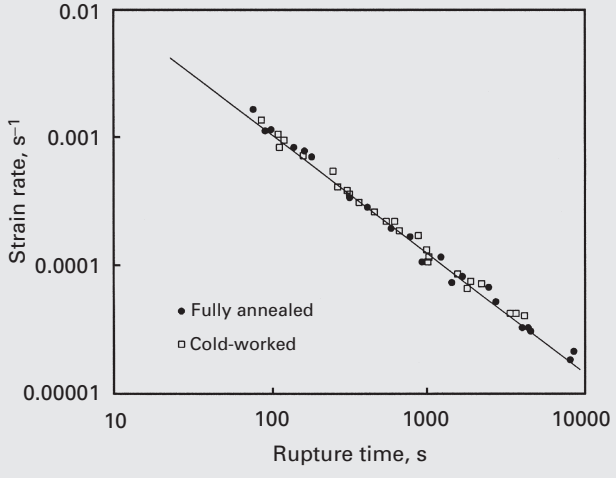
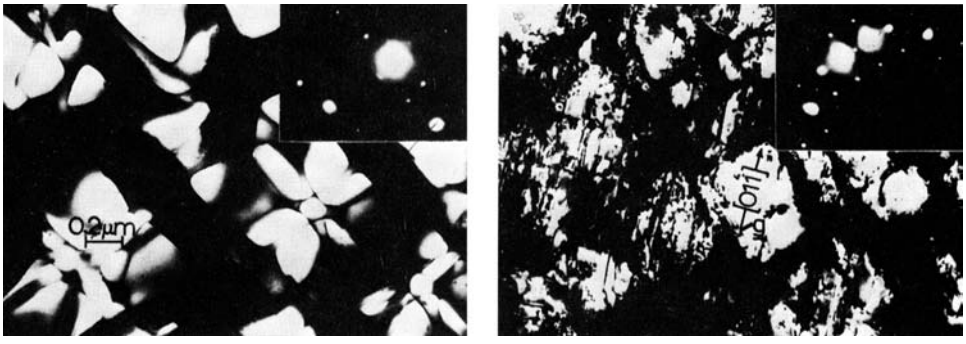


Fig. E13.4 Monkman–Grant plot for Zircaloy. (Courtesy of M. E. Kassner and T. Haynes.)

13.10 Heat-Resistant Materials

High-temperature materials can be classified into two groups: metals and ceramics. High-temperature alloys are, in their turn, classified into superalloys and refractory alloys. The latter are alloys of elements with high melting points, such as tantalum, molybdenum, and tungsten. The superalloys are usually alloys developed for elevated



(a) (b)

Fig. 13.24 Transmission electron micrograph of Mar M-200; notice the cuboidal γ' precipitates. (Courtesy of L. E. Murr.)

temperature service, usually based on group VIIIa elements, where relatively severe mechanical stressing is encountered and where high surface stability is frequently required.

The development of superalloys was initiated in the 1930s, and their first use was in turbo superchargers of reciprocating aircraft engines. The introduction of the turbine in the 1940s was a strong motivator for subsequent developments. Superalloys encompass the nickel, iron, cobalt, and iron–nickel systems. The majority of authors do not include chromium-based alloys in this group. The maximum service temperature (temperature capability) has increased continuously in the past; it can be around 1,200 °C. The life of turbines has increased from 5,000 to over 20,000 hours. The combined effects of high stresses, high temperatures, and long lifttimes have required improvements in the following properties:

- *Short-term mechanical properties*: yield stress, ductility.
- *Long-term mechanical properties*: low- and high-cycle fatigue, creep, creep–fatigue.
- *Hot corrosion resistance*: the principal deterioration processes are oxidation, chlorination, sulfidation, and carburization.

Nickel-based superalloys are the most important group; most commercial nickel-based alloys have more than ten constituent elements and over ten trace elements. These can be divided into the following categories, depending on the function and position of the element in the periodic chart:

- Elements that form substitutional solid solutions in the austenitic matrix: cobalt, iron, chromium, vanadium, molybdenum, tungsten.
- Elements that form precipitates: aluminum, titanium, niobium, tantalum. Figure 13.24 shows the cuboidally shaped γ' precipitates Ni_3Al, Ni_3Ti, and $Ni_3(Al, Ti)$ that are aligned along specific planes of the austenitic matrix.
- Carbide-forming elements: chromium, molybdenum, tungsten, vanadium, niobium, tantalum, titanium.
- Elements that segregate along the grain boundaries: magnesium, boron, carbon, zirconium.

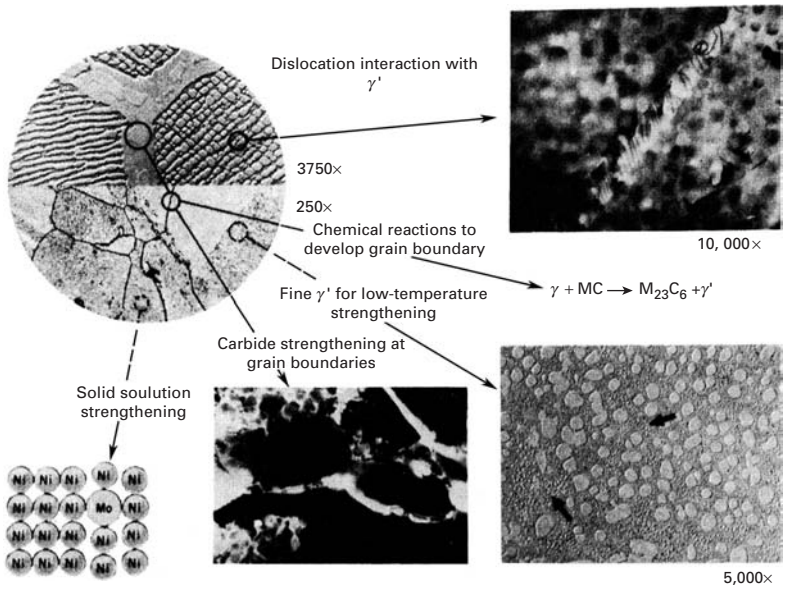


Fig. 13.25 Major microstructural strengthening mechanisms in nickel-based superalloys. (Reprinted with permission from C. T. Sims and W. C. Hagel, eds., *The Superalloys* (New York: Wiley, 1972), p. 33.)

- Elements forming protective and adherent oxides: chromium, aluminum.
- Rare-earth elements.

The microstructure of superalloys reflects the concern of using all possible strengthening mechanisms to retard creep. Figure 13.25 is a composite of these features. One has to retard the movement of dislocations. This is achieved by substitutional solid solution atoms and by a great volume percentage of the $Ni_3(Ti,Al)$ phase γ'. The grain boundaries are strengthened by precipitation of $M_{23}C_6$ carbides on them. Secondary γ', very fine, is precipitated in the space between neighboring primary γ', which is larger. One also wants to carefully avoid the topologically close-packed (TCP) phases R and σ, that occur accidentally and after long exposure to high temperatures, embrittling the alloy.

An interesting phenomenon occurs when γ' strengthened superalloys are stressed at high temperatures. This phenomenon is called "rafting" and consists of the preferential growth of the gamma prime precipitates in a certain orientation. Figure 13.26(a) shows the initial configuration of the precipitates, aligned with the {100} axes of the cubic matrix. In gamma prime strengthened superalloys, the precipitates are actually the major phase (~70%), with the matrix the minor phase. After the MarM-200 alloy shown in Figure 13.26(a) was creep deformed at 1253 K for 28 hours along the [010] direction (stress = 350 MPa), the precipitates coarsened along that loading direction; this is shown in Figure 13.26(b). They are seen to connect with each other and form continuous sheets.

Figure 13.27 shows the stress-rupture properties of a number of nickel-based superalloys. The stress required for rupture in 1,000 hours is plotted against the temperature. The load-bearing ability in

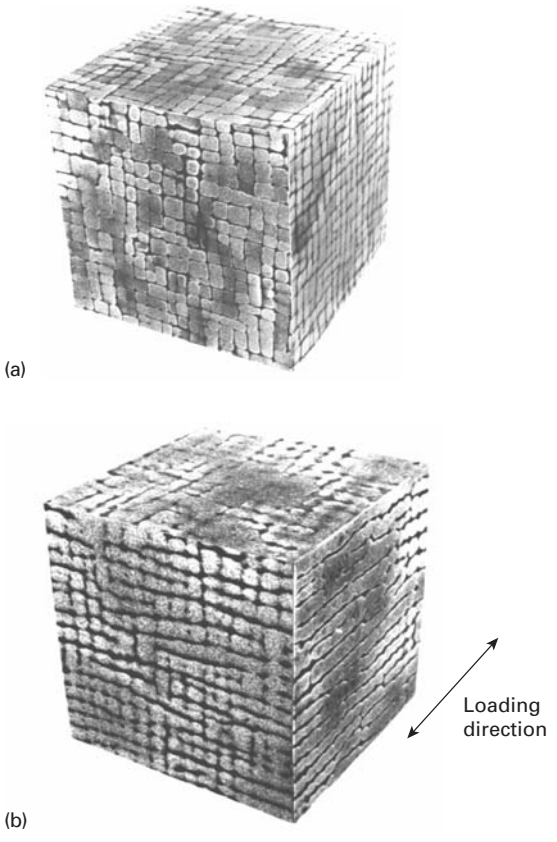


Fig. 13.26 Rafting in MAR M-200 monocrystalline superalloy; (a) original configuration of gamma prime precipitates aligned with three orthogonal cube axes; (b) creep deformed at 1253 K for 28 hours along the [010] direction, leading to coarsening of precipitates along loading direction. (From U. Glatzel, "Microstructure and Internal Strains of Undeformed and Creep Deformed Samples of a Nickel-Based Superalloy," Habilitation Dissertation, Technische Universität, Berlin, 1994.)

the upper range of the use of these superalloys is only a fraction of the one at lower temperatures. The range 800 to 1,000 °C is a very critical one.

Figure 13.28 shows a cross-section of a gas turbine. The temperature of combustion gases is on the order of 1500 °C. The rotational speed of the turbine can be as high as 30,000 rpm, creating centrifugal stresses of the order of 200 MPa. Polycrystalline superalloys have been succeeded by directionally solified superalloys having most grain boundaries aligned with the blade direction. This minimizes creep by grain-boundary sliding. A further development is the single crystal blade, which has no grain boundary. Figure 13.29 shows such a blade. Notice (arrow) that it has holes. Indeed, internal cooling channels are introduced into modern blades to allow air to circulate inside and create, upon exiting the blade, a protective layer at a lower temperature than the exhaust gases. Turbines with ceramic coatings called thermal barrier coatings (TBC) are used; they shield the metallic turbine from the excessively high temperatures. This poses an enormous challenge, because the metallic and ceramic components have different coefficients of thermal expansion, creating stresses that crack the coating. This problem is alleviated by the introduction of an intermediate layer. Figure 13.29(b) shows the significant increase (>100 °C) that can be accomplished using TBCs.

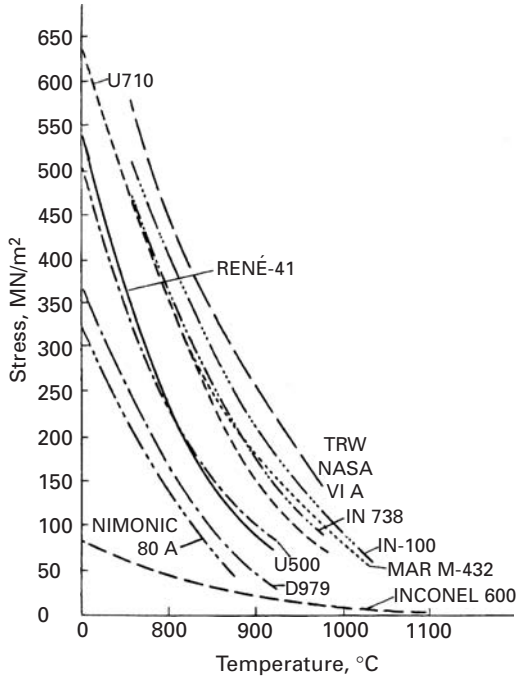


Fig. 13.27 Stress versus temperatures curves for rupture in 1,000 hours for selected nickel-based superalloys. (Reprinted with permission from C. T. Sims and W. C. Hagel, eds., *The Superalloys* (New York: Wiley, 1972), p. vii.)

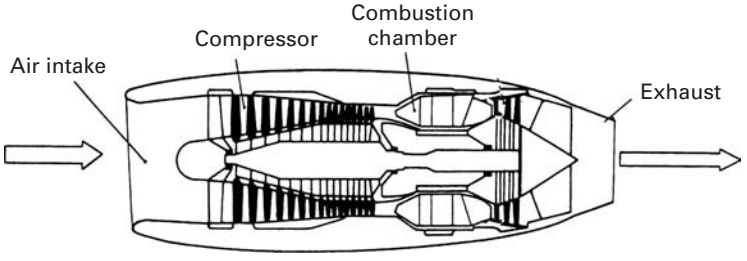


Fig. 13.28 Cross-section of a gas turbine showing different parts. The temperature of gases in combustion chamber reaches 1500 °C.

The use of single-crystal turbine blades represents a significant technological development: Grain-boundary sliding is eliminated by this technique, and an increase in temperature capability of approximately 50 °C over that of polycrystalline superalloys is achieved. The method of producing single-crystal turbine blades involves investment casting in a controlled thermal environment. An added advantage is that the composition of the alloy is simpler than in polycrystalline superalloys.

High-performance ceramics are prime candidates for structural components in advanced automotive gas turbine engines. Operating conditions for such components involve high mechanical and thermal stresses at elevated temperatures; hence, the creep resistance of these materials is of great importance. Excessive creep deformation can affect the dimensional stability of the component, ultimately leading to a loss of function. Generally, there is a high resistance to slip and diffusional mechanisms in high-performance ceramics such as SiC and Si_3N_4. Damage mechanisms such as cavitation, solution of silicon nitride into the glassy phase at the grain

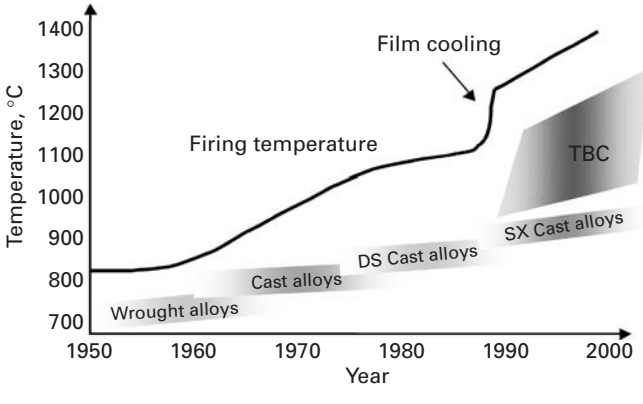


Fig. 13.29 (a) Single crystal turbine blade developed for stationary turbine. (Courtesy of U. Glatzel.) (b) Evolution of maximum temperature in gas turbines; notice the significant improvement made possible by the introduction of thermal barrier coatings (TBCs). (Courtesy of V. Thien, Siemens.)

boundary, and grain-boundary sliding are associated with creep in these ceramics. As mentioned earlier in this chapter, different creep mechanisms give different creep activation energies. Cavitation can result in a reduction in the strength of the material and lead to time-dependent failure (i.e., creep rupture). Viscous flow in any glassy grain-boundary phase can lead to excessive creep deformation. Silicon nitride is a good example of a high-performance ceramic material to use for illustrating some of the unusual problems that must be faced

before their full potential can be realized. Silicon nitride has excellent short-term strength and fracture toughness. Generally, hot, isostatically pressed silicon nitride shows superior properties to silicon nitride processed by pressureless sintering or uniaxial hot pressing. The amount and nature of any densification aids can significantly affect creep behavior. Yttria is a common densification aid used in silicon nitride. Another important variable is the testing method (compression, bending, or tensile testing) used for creep. Although tensile testing of ceramics is not very common, it has been used to study the high-performance ceramics.

Studies aimed at evaluating the long-term mechanical performance of silicon nitride (trade designation NT 154) have shown that cavitation along two-grain junctions controls both creep deformation and creep rupture strength.[16] Silicon nitride is available with different purity levels. Ferber et al. observed that the creep and creep rupture behavior of silicon nitride (NT 164) were significantly improved, compared to that of the commercial material (NT 154), if one could ensure the absence of cavitation along two-grain junctions.[17] These authors attributed the growth of cavities to the following processes, which occur in a sequence: (1) solution of silicon nitride into the intergranular phase at the cavity boundary; (2) transport of the dissolved species along the grain boundary; and (3) precipitation of the species at low-stress sites remote from the cavity. Yet another factor in the high-temperature behavior of nonoxide ceramics such as SiC and Si_3N_4 is their oxidation resistance in air. Oxidation of silicon nitride, rather than creep, was observed to initiate a stress-oxidation damage zone in the material. Finally, we reiterate the importance of the testing method. Wiederhorn et al. found asymmetric behavior of Si_3N_4 in creep:[18] A linear response was obtained in compression, but a power-law response held in tension, with the creep exponent n in the range $2 < n < 5$. These researchers observed minimal cavitation in compression; in tension, however, cavities formed at multigrain junctions, and the creep strain was proportional to the volume fraction of cavities. Thus, cavitation is responsible for creep strain in tension, but not in compression. This discussion should bring home to the reader some important differences between creep mechanisms in metals and in nitrogen ceramics. In metals, a lot of creep strain can occur, but not much of it is due to cavitation. Also, the tertiary creep of metals is absent in silicon nitride.

Ceramics and ceramic composites possess a higher temperature capability than metals. Whereas ceramics tend to be brittle, the addition of reinforcing fibers adds toughness to them. Temperatures

[16] D. C. Cranmer, B. J. Hockey, S. M. Wiederhorn, and R. Yeckley, *Ceram. Eng. Sci. Proc.*, 12 (1991) 1862.

[17] M. K. Ferber and M. G. Jenkins, *J. Am. Ceram. Soc.*, 75 (1992) 2453; and M. K. Ferber, M. G. Jenkins, and T. A. Nolan, *J. Am. Ceram. Soc.*, 77 (1994) 657.

[18] S. M. Wiederhorn, B. J. Hockey, W. E. Luecke, R. Krause, and J. French, unpublished results.

approaching 2,000 °C can be reached with acceptable creep rates. As mentioned earlier, chemical degradation becomes very important at these temperatures, especially for nonoxide ceramics. (Oxides are, obviously, immune to oxidation.) A high-temperature material of some promise is $MoSi_2$, especially as a matrix for high-temperature structural composites (see Chapter 12).

13.11 Creep in Polymers

As mentioned in the preceding sections, creep is a thermally activated process and thus becomes important at high temperatures. The term *high temperature* is a relative one; it is more convenient to use the term *homologous temperature*, $T_H = T/T_m$, where T is the temperature of interest and T_m is the melting of the material (both in kelvins). Typically, creep becomes a significant deformation mode for metals at a homologous temperature greater than 0.4 and for ceramics at a homologous temperature greater than 0.5. (In the case of amorphous polymers, one uses the glass transition temperature T_g rather than the melting point T_m.) At low temperatures, most metals and ceramics show time-independent deformation. In general, polymers show a much larger dependence on time and temperature than metals and ceramics do; that is, polymers show creep effects at much lower stresses and temperatures. This stems from their weak van der Waals interchain forces. In polymers, time-dependent deformation becomes important even at room temperature. Two terms are used to describe the time-dependent behavior of polymers: *creep* and *stress relaxation*. In creep, one applies a constant stress, and the strain response is measured as a function of time. In stress relaxation, one applies a constant strain, and the response is measured in terms of a decrease in stress as a function of time. We have discussed some aspects of these two phenomena in Chapter 2 in connection with viscoelasticity.

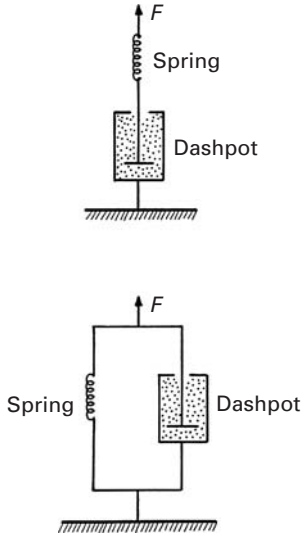


Fig. 13.30 Spring–dashpot analogs (a) in series and (b) in parallel.

For a glassy, viscoelastic polymer subjected to a constant stress σ_0, there is an initial elastic strain recovery, followed by a slow, time-dependent recovery. This viscoelastic response may be modeled as a spring and a dashpot in series (also called the Maxwell model), as shown in Figure 13.30. An application of stress to this system results in a strain ε in the system. This strain is the sum of two contributions, and we can write

$$\varepsilon = \varepsilon_1 + \varepsilon_2, \tag{13.20}$$

where ε_1 is the strain in the spring and ε_2 is the strain in the dashpot. The stresses in the spring and the dashpot are identical, because the two are in series, i.e.,

$$\sigma = \sigma_1 = \sigma_2. \tag{13.21}$$

Then, we can write the following relationships for the elastic (Young's modulus, E) and the viscous case (viscosity, η),

$$\frac{d\sigma}{dt} = E\frac{d\varepsilon_1}{dt}, \quad \sigma = \eta\frac{d\varepsilon_2}{dt}. \tag{13.22}$$

From Equations 13.20 and 13.21, we get

$$\frac{d\varepsilon}{dt} = \frac{d\varepsilon_1}{dt} + \frac{d\varepsilon_2}{dt} = \frac{1}{E}\frac{d\sigma}{dt} + \frac{\sigma}{\eta}. \tag{13.23}$$

Note that the series, or Maxwell, model does not correctly predict the behavior of a viscoelastic material under constant stress or creep conditions (i.e., $\sigma = \sigma_0$), because, in this case,

$$\frac{d\varepsilon}{dt} = \frac{\sigma_0}{\eta}. \tag{13.24}$$

That is, the Maxwell model for creep or constant-stress conditions predicts that the strain increases linearly with time. (See Figure 13.31(a).) Most polymers, however, show de/dt increasing with time. The Maxwell model is more realistic in the case of a stress relaxation test, during which we impose a constant strain $\varepsilon = \varepsilon_0$ and $d\varepsilon/dt = 0$. Under these conditions, Equation 13.23 can be written as

$$0 = \frac{1}{E}\frac{d\sigma}{dt} + \frac{\sigma}{\eta},$$

or

$$\frac{d\sigma}{\sigma} = -\left(\frac{E}{\eta}\right)dt.$$

We can integrate this expression to get

$$\sigma = \sigma_0 \exp\left(-\frac{E\,t}{\eta}\right). \tag{13.25}$$

The quantity η/E is referred to as the *relaxation time* τ, and we can rewrite Equation 13.25 as

$$\sigma = \sigma_0 \exp(-t/\tau). \tag{13.26}$$

Equation 13.26 says that the stress decays exponentially with time, as shown in Figure 13.31(b) (Maxwell). This is quite reasonable for many polymers; however, the process of stress relaxation does not go on indefinitely in real materials.

In another model, called the Voigt model, the spring and the dashpot are arranged in parallel (Figure 13.30(b)). This means that the strains in the two components are identical, i.e.,

$$\varepsilon = \varepsilon_1 + \varepsilon_2, \tag{13.27}$$

and the stresses in the two components add, to give the stress on the system, i.e.,

$$\sigma = \sigma_1 + \sigma_2. \tag{13.28}$$

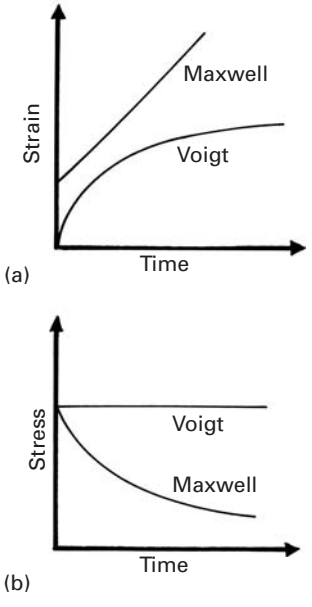


Fig. 13.31 (a) Strain–time and (b) stress–time predictions for Maxwell and Voigt models.

It may be shown from Equations 13.27 and 13.28 that

$$\frac{d\varepsilon}{dt} = \frac{\sigma_0}{\eta} - E\frac{\varepsilon}{\eta}. \tag{13.29}$$

Let us now examine the predictions of the Voigt model for creep, or constant stress loading, and for stress relaxation. For the constant-stress situation, $\sigma = \sigma_0$, Equation 13.29 becomes

$$\frac{d\varepsilon}{dt} + E\frac{\varepsilon}{\eta} = \frac{\sigma_0}{\eta}.$$

This differential equation has the solution

$$\varepsilon = \left(\frac{\sigma_0}{E}\right)\left[1 - \exp\left(\frac{-E\,t}{\eta}\right)\right]. \tag{13.30}$$

Remembering that the quantity η/E is the relaxation time τ, we find that the variation in strain with time at a constant stress (creep) is given by

$$\varepsilon = \left(\frac{\sigma_0}{E}\right)\left[1 - \exp\left(\frac{-t}{\tau}\right)\right]. \tag{13.31}$$

This relationship is shown in Figure 13.31(a); the prediction of the Voigt model is quite realistic, because $\varepsilon \to \sigma_0/E$ as $t \to \infty$.

For the stress relaxation case, we have an imposed constant strain $\varepsilon = \varepsilon_0$, and therefore, $d\varepsilon/dt = 0$. The Voigt model predicts that

$$\frac{\sigma}{\eta} = \frac{E\varepsilon_0}{\eta},$$

or

$$\sigma = E\varepsilon_0.$$

This linear elastic response, however, shown in Figure 13.31(b), does not conform to reality.

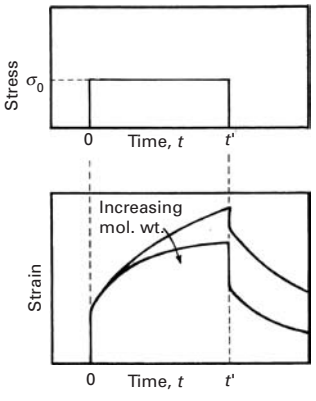


Fig. 13.32 Strain response as a function of time for a glassy, viscoelastic polymer subjected to a constant stress σ_0. Increasing the molecular weight or degree of cross-linking tends to promote secondary bonding between chains and thus make the polymer more creep resistant.

The molecular weight of a polymer can affect its creep behavior. The strain response of a polymer as a function of time, $\varepsilon(t)$, is shown in Figure 13.32. Also shown is the effect of molecular weight. The effect of increasing the degree of cross-linking is in the same direction as that of increasing the molecular weight. Both tend to promote secondary bonding between chains and thus make the polymer more creep resistant. Compared to glassy polymers, semicrystalline polymers tend to be more creep resistant. Polymers containing aromatic rings in the chain are even more creep resistant. Both increased crystallinity and the incorporation of rigid rings add to the thermal stability, and thus to the creep resistance, of a polymer.

In a constant-stress test of the kind just described, a parameter of interest is the *creep compliance J*. This is the ratio of strain to stress. Since the strain will be a function of time, the compliance will also be a function of time. Thus,

$$J(t) = \varepsilon(t)/\sigma_0. \tag{13.32}$$

From Equation 13.32 and 13.31, we can write the creep compliance as

$$J(t) = \varepsilon(t)/\sigma_0 = (1/E)[1 - \exp(-t/\tau)]. \tag{13.33}$$

If one plots a series of creep compliances as a function of time, both on logarithmic scales, over a range of temperature, one gets the curve shown in Figure 13.33(a). It turns out that such individual plots can be superposed by horizontal shifting (along the log-time axis) by an amount log a_t, to obtain a master curve shown in Figure 13.33(b). In Figure 13.33(a), we use arrows to indicate the horizontal shift of data to obtain a master curve corresponding to a reference temperature of the polymer. This figure shows that, when creep compliance is a measured at a series of temperatures, with the glass transition temperature T_g as the reference temperature, then curves above T_g are shifted to the right, while curves below T_g are shifted to the left.

As discussed earlier, a thermally activated process shows a dependence of its on temperature that can be described by an Arrhenius-type expression. When viscous flow occurs in a polymer, the network structure breaks and re-forms locally. The thermal energy for such viscous flow is available above the glass transition temperature T_g. Below T_g, the thermal energy is not high enough for the breaking and remaking of the bonds, and the material does not flow easily. In the viscoelastic range, time and temperature have similar effects on polymers. There are two easy ways of studying such behavior. In the first of these, we can impose a constant deformation on the polymer and follow the resultant stress. This will give us a stress relaxation modulus as a function of time. The other technique involves the application of a constant stress and measuring the deformation as a function of time. This will give us a curve of compliance vs. time. A very useful principle called time–temperature superposition allows us to take the data at one temperature and superimpose them on data taken at another temperature by a shift along the log-time axis. This principle is of great practical use, in as much as obtaining data over a full range of creep compliance or stress relaxation behavior can involve years. The principle allows one to shift data taken over shorter time spans, but at different temperatures, to obtain a master curve that covers longer time spans. Williams, Landel, and Ferry found that the logarithm of a_T (the time-shift factor) follows a simple expression, viz.,

$$\log a_T = -C_1(T - T_s)/(C_2 + T - T_s),$$

where C_1 and C_2 are constants and T_s is a reference temperature for a given polymer.[19] If we take the reference temperature to be the glass transition temperature T_g, then $C_1 = 17.5$ and $C_2 = 52$ K. If the reference temperature T_s is taken to be about 50 °C above T_g, then $C_1 = 20.4$ and $C_2 = 101.6$ K.

The amount of shift can be calculated by the Williams–Landel–Ferry expression. The master curve for creep, obtained by superposing horizontally shifted curves, is shown in Figure 13.33(b). Another way of treating this problem is shown in Figure 13.33(c), where we plot

[19] M. L. Williams, R. F. Landel, and J. D. Ferry, *J. Amer. Chem. Soc.*, 77 (1955) 3701.

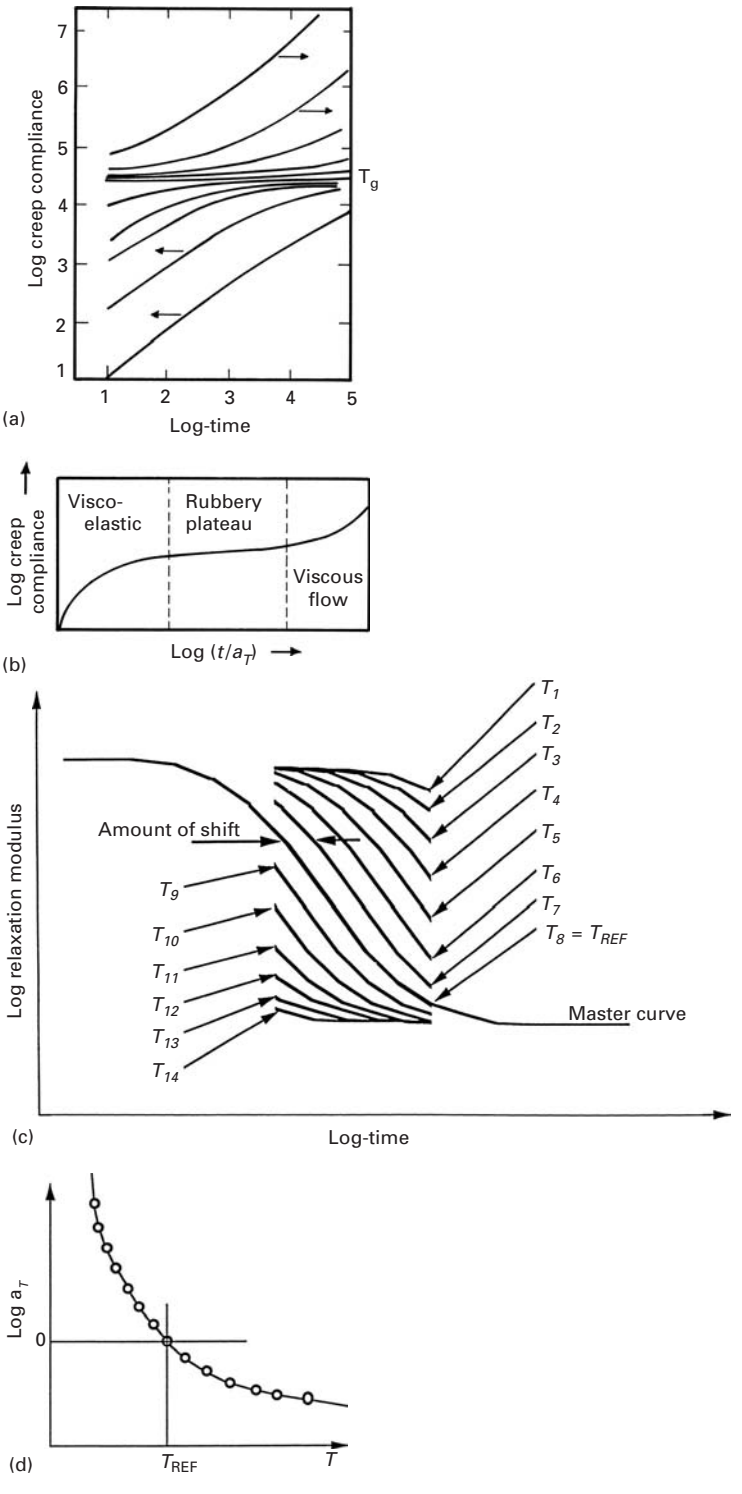


Fig. 13.33 (a) A series of creep compliances vs. time, both on logarithmic scales, over a range of temperature. (b) The individual plots in (a) can be superposed by horizontal shifting (along the log-time axis) by an amount log a_T, to obtain a master curve corresponding to a reference temperature T_g of the polymer. (c) Shift along the log-time scale to produce a master curve: modulus vs. time. (Courtesy of W. Knauss.) (d) "Experimentally" determined shift factor.

the stress relaxation modulus as a function of time, both on logarithmic scales. The "experimentally" determined time-shift factor, as a function of temperature, is shown in Figure 13.33(d).

Now recall the model in which spring and dash-pot are arranged in parallel, i.e., the Voigt model. The model is used to explain the stress relaxation behavior of a polymer. We impose a constant strain ε_0 and follow the drop in stress $\sigma(t)$ as a function of time. (See Figure 13.34.) Instead of a compliance term, we now have a stress relaxation modulus, given by

$$E(t) = \sigma(t)/\varepsilon_0.$$

In the case of stress relaxation also, one can obtain a master curve, as shown schematically in Figure 13.35. Also shown in the figure is the effect of cross-linking and molecular weight. Stress relaxation in polymers is of great practical significance when the polymers are used in applications involving gaskets and seals. At times, this effect can be exploited beneficially: for example, in a situation where residual stresses are not desirable, we can incorporate a polymer to undergo easy stress relaxation in response to residual stresses.

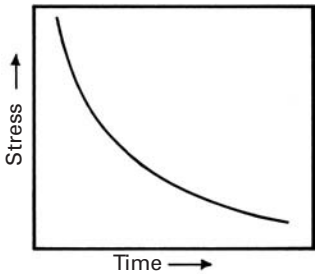


Fig. 13.34 A constant imposed strain ε_0 results in a drop in stress $\sigma(t)$ as a function of time.

Example 13.5

Data on stress relaxation modulus vs. time for polyisobutylene (also known as chewing gum) are shown in Figure E13.5(a). The data span a range of 10^{-2} to 10^{2} hours in time. Obtain the curve of the time-shift factor for a reference temperature of 298 K. Obtain a master curve for polyisobutylene based on time–temperature superposition of data.

Solution: By using the Williams–Landel–Ferry expression

$$\log a_T = -17.5(T - T_{\text{ref}})/[52 + (T - T_{\text{ref}})]$$

for a temperature range of -80 °C to 80 °C, we obtain the curve of the time-shift factor vs. temperature shown in Figure E13.5(b). Using this time-shift curve, we superimpose the individual stress relaxation modulus curves given in the statement of the problem (Figure E13.5(a)), to obtain the master curve shown in Figure E13.5(c). Note that the time scale ranges from 10^{-14} to 10^{2} hours.

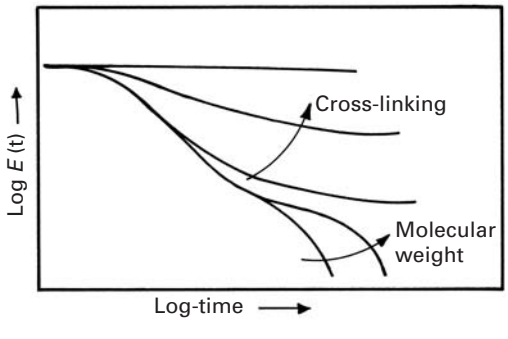


Fig. 13.35 A master curve obtained in the case of stress relaxation, showing the variation in the reduced modulus as a function of time. Also shown is the effect of cross-linking and molecular weight.

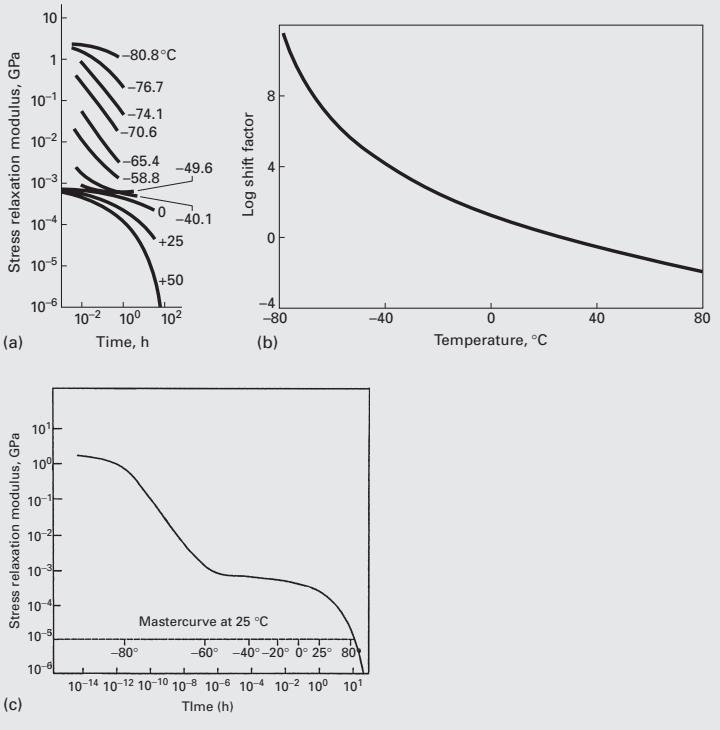


Fig. E13.5 (a) Stress relaxation modulus; (b) time–shift factor; (c) master curve. (From E. Catsiff and A. V. Tobolsky, *J. Polymer Sci.*, 19 (1956) 111.)

Example 13.6

The creep strain rate of a polymer is given by the expression

$$\dot{\varepsilon} = 4.5 \times 10^{11} \exp(-100 \text{ kJ}/RT),$$

where T is the temperature in kelvins and R is the universal gas constant. How much time will it take for a rod of this polymer to extend from 10 mm to 15 mm at 100 °C?

Solution:

$$\dot{\varepsilon} = \Delta\varepsilon/\Delta t = 4.5 \times 10^{11} \exp[-100,000/(8.3 \times 373)]$$
$$= 4.2 \times 10^{-3} \text{ s}^{-1}$$

Hence,

$$\Delta t = 0.5/(4.2 \times 10^{-3}) = 119 \text{ s}.$$

Example 13.7

The activation energy for stress relaxation in a polymer is 50 kJ/mol. The relaxation time at 25 °C is 90 days. What is the relaxation time at 125 °C?

Solution:

$$
\begin{aligned}
1/\tau &= \frac{E}{\eta} = A\exp(-Q/RT),\\
(1/\tau_{25})/(1/\tau_{125}) &= \exp(-Q/R298)/\exp(-Q/R398),\\
\tau_{125} &= \tau_{25}\exp[(Q/R)(1/398 - 1/298)],\\
\tau_{125} &= 90\exp[((50\times10^{3})/8.314)(1/398 - 1/298)],\\
&= 90\times6.4\times10^{-3},\\
&= 0.57 \text{ day}.
\end{aligned}
$$

Example 13.8

A nylon cord, used to tie a sack, has an initial stress of 5 MPa. If the relaxation time for this cord is 180 days, in how many days will the stress reduce to 1 MPa?

Solution:

$$
\begin{aligned}
\sigma &= \sigma_0\exp(-t/\tau),\\
1\text{ MPa} &= 5(-t/180),\\
t &= -180\ln 1/5,\\
t &= 290 \text{ days}.
\end{aligned}
$$

13.12 Diffusion-Related Phenomena in Electronic Materials

Interconnect lines in integrated circuits are subject to diffusional processes that can lead to void growth. These voids, inconjunction with stresses, may lead to failure. There are two mechanisms responsible for void growth:

- *Thermal stresses*. Figure 13.36(a) shows an interconnect. The core, which has a rectangular cross section, is the component carrying the electrical current. Electrical currents as high as 10^{10} A/m^2 are carried by the metallic conductor (typically, aluminum). The conductor is covered by a passivation film (e.g. silicon nitride). The difference in thermal expansion coefficient between the two leads them to tensile stresses in the conductor. Assuming a rigid passivation layer, the stress in the conductor can be estimated from:

$$\sigma_t = 3\Delta\alpha\Delta T K, \tag{13.34}$$

  where ΔT is the temperature change follwing passivation, $\Delta\alpha$ is the difference in thermal expansion coefficient between conductor and passivation layer, and K is the bulk modulus of the

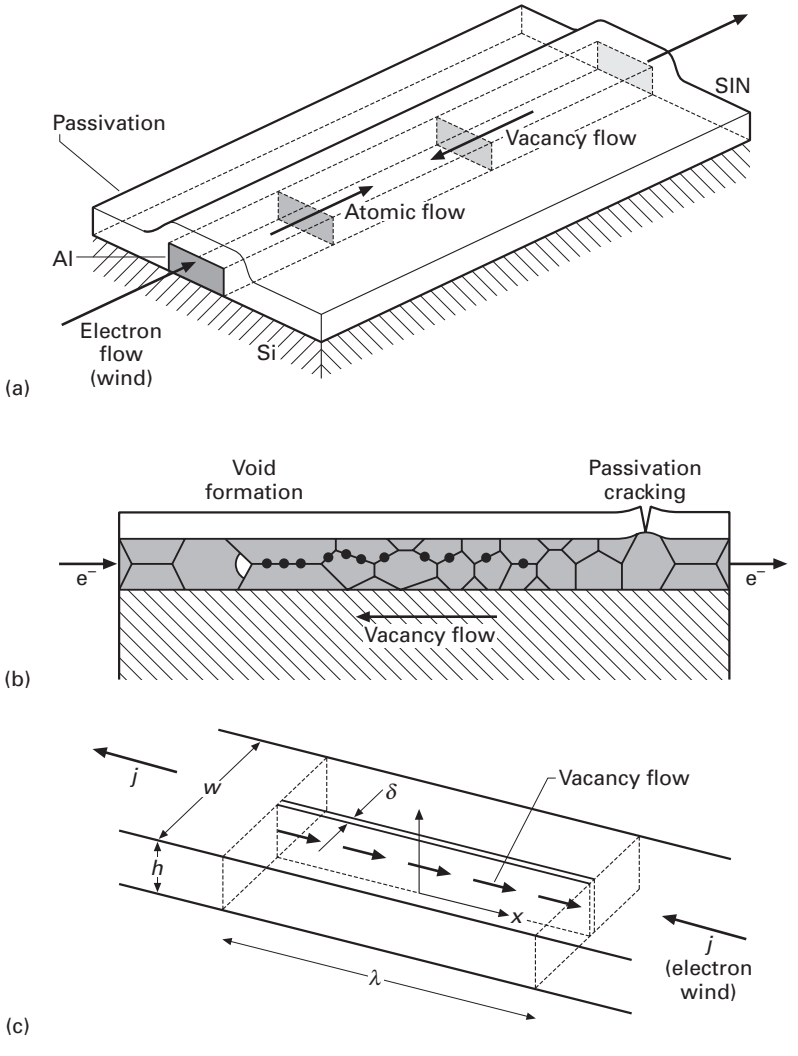


Fig. 13.36 Metal interconnect line covered by passivation layer subjected to electromigration; (a) overall scheme; (b) voids and cracks produced by thermal mismatch and electromigration; (c) basic scheme used in Nix–Arzt equation, which assumes grain-boundary diffusion of vacancies counterbalancing electron wind. (Adapted from W. D. Nix and E. Arzt, *Met. Trans.*, 23A (1992) 2007.)

conductor. Glaxner et al.[20] estimate that these hydrostatic (tensile) stresses can be as high as 500 MPa. Diffusion of vacancies, under these stresses, at a temperature of 200 °C (typical of the conductor) can lead to the formation of voids within the metal or cracks in the passivation layer.

- *Electromigration*. The high current densities in the metal interconnect lines produce an "electron wind." This electron wind causes the migration of matter through, primarily, the diffusion of vacancies. The electron wind will produce at atomic flow in the same direction, with a vacancy flow in the opposite direction. The vacancy flow, which occurs principally along grain boundaries, is interrupted at transverse grain boundaries. The vacancies condense and form voids. This is shown in Figure 13.36(b).

[20] R. J. Glaxner, B. M. Clemens, and W. D. Nix, *J. Mater. Res.*, 12 (1997) 2081.

Figure 13.36(c) shows the flow of electrons and vacancies in a grain with length λ. Diffusion occurs primarily along grain boundaries because the grain-boundary diffusion coefficient, D_{gb}, is orders of magnitude higher than the lattice diffusion coefficients at temperatures below $0.5T_m$, where T_m is the melting point in kelvin. The thickness of the grain boundary is δ. Nix and Arzt[21] developed the following equation for the stress due to electromigration:

$$\sigma_{NA} = \frac{1}{2}\frac{\lambda}{\Omega} eZ^* \rho j, \tag{13.35}$$

where Ω is the atomic volume, λ is the grain-boundary length, δ is the grain-boundary thickness, eZ^* is the effective charge of the metal ions, ρ is the resistance, and j is the current density.

Example 13.9

Calculate the Nix–Arzt stress in an aluminum line if the current density is 10^{10} A/m^2 and the grain size is 20 μm.

Given:

$eZ^* = 3.2 \times 10^{-18}$ coulombs,
$\rho = 28 \times 10^{-9}$ Ωm,
$\delta = 3 \times 10^{-10}$ m,
$\Omega = 16.6$ A^3,
$j = 1010$ A/m^2.

Solution: Inserting these values into Equation 13.33:

$$\sigma_{NA} = 520 \text{ MPa}.$$

This is indeed a high value.

13.13 Superplasticity

Some metallic alloys and ceramics show a peculiar behavior called *superplasticity*. This is the ability to flow, in tension, to very large elongations. Figure 13.37 shows a dramatic illustration of superplastic behavior. The specimen was extended at a temperature of 413 K, and a total strain of 48.5 was reached without failure. The phenomenon of superplasticity was observed for the first time in 1934. A great deal of activity has taken place in the area since then, and superplastic forming has become a successful industrial process.

Superplasticity has been obtained in a number of alloy systems, including titanium alloys (Ti–6% Al–4% V), iron-based alloys, and

[21] W. D. Nix and E. Arzt, *Met. Trans.*, 32A (1992) 2007.

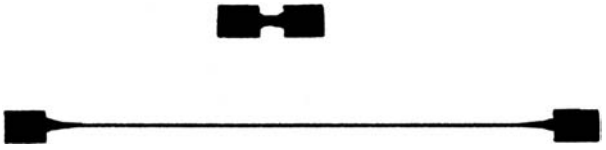


Fig. 13.37 Superplastic tensile deformation in Pb–62% Sn eutectic alloy tested at 415 K and a strain rate of 1.33×10^{-4} s^{-1}; total strain of 48.5. (From M. M. I. Ahmed and T. G. Langdon, *Met. Trans. A*, 8 (1977) 1832.)

aluminum alloys. High-strength nickel-based superalloys have been found to exhibit superplastic behavior, and the process of "gatorizing" (supposedly named after an alligator living in the lake in front of the research institute) has been developed by Pratt and Whitney. The potential of superplastic forming is especially bright for titanium alloys, which are known to be very difficult to form, because of their HCP structure. Superplasticity has also been discovered in ceramics.

The basic reason that some materials can deform superplastically when others cannot is related to how they respond to changes in strain rate. The example of hot glass (above the transition temperature) comes to mind. We are all familiar with Coke bottles stretched to very high strains; these interesting items are sold in curio shops. Glass shows a Newtonian viscous behavior above a certain temperature. Fiberglass is formed in such a manner and can be pulled to extremely fine fibers. In a lamellar flow, Newtonian viscosity is defined by (see Equation 3.29)

$$\tau = \eta \frac{dv}{dy},$$

where dv/dy is the variation in velocity of the fluid with distance y, τ is the shear stress necessary to create the velocity gradient dv/dy, and η is the viscosity. The derivative dv/dy is equivalent to the shear strain rate $\dot{\gamma}$. (See Section 3.6.2.) Thus, we can write

$$\tau = \eta \dot{\gamma}. \tag{13.34}$$

The stress-versus-strain-rate relationship from many materials is not linear, but of the form (see Equation 3.23)

$$\sigma = K \dot{\varepsilon}^m, \tag{13.35}$$

where K and m are constants and m is called the strain-rate sensitivity. In general, m varies between 0.02 and 0.2, for homologous temperatures between 0 and 0.9 (90% of the melting point in K). Hence, one would have, at the most, an increase of 15% in the yield stress by doubling the strain rate. Comparing Equations 13.34 and 13.35, we see that a value of $m = 1$ will give a Newtonian viscous solid. Such a material would not undergo tensile instability and could be stretched indefinitely. Figure 13.38(a) shows schematically how a high value of m will inhibit tensile instability (necking) and, consequently, enhance plasticity in tension. The specimen is being deformed, in tension, at a velocity v. The length increases from L_0 to L_1 and then to L_2. At L_2, necking starts. If the material has a high value of m, this instability

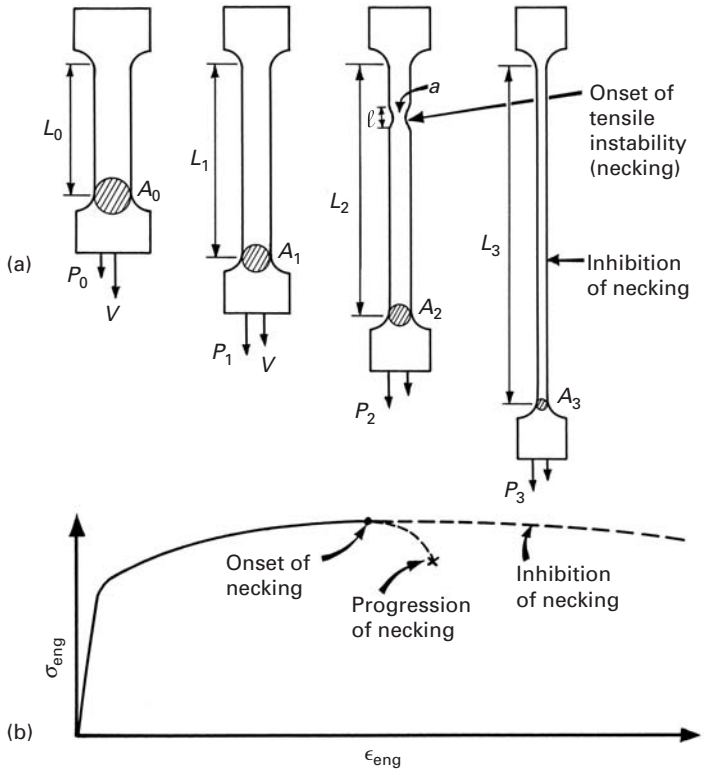


Fig. 13.38 (a) Schematic representation of plastic deformation in tension with formation and inhibition of necking. (b) Engineering-stress–engineering-strain curves.

will be inhibited because of the localized strengthening effect. When the length is L_2, the strain rate over the specimen is

$$\dot{\varepsilon}_2 = \frac{v}{L_2}.$$

In the incipient neck region, which acts as a "minispecimen" embedded in the large specimen, one has

$$\dot{\varepsilon}'_2 = \frac{v}{\ell}.$$

Since $\ell < L_2$, one has $\dot{\varepsilon}'_2 > \dot{\varepsilon}_2$.

The strain rate sensitivity can be obtained from Equation 13.35 by applying that equation to two strain rates and eliminating K. When we do this, we obtain

$$m = \frac{\ln(\sigma'_2/\sigma_2)}{\ln(\dot{\varepsilon}'_2/\dot{\varepsilon}'_2)}.$$

For a high value of m, the strength σ'_2 in the neck region is much higher than σ_2, and further progression of plastic deformation at that region is halted. Mathematically,

$$\sigma'_2 > \frac{P_2}{a}.$$

When m is low, σ'_2 is not sufficiently high, and we have

$$\sigma'_2 < \frac{P_2}{a}.$$

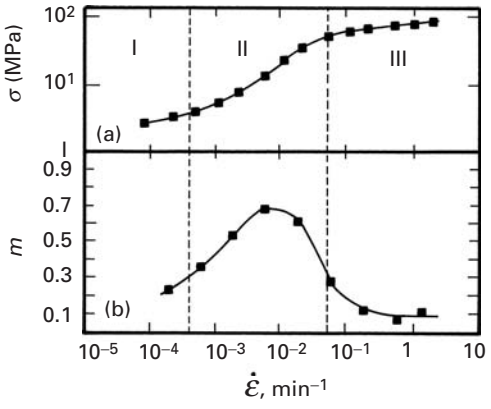


Fig. 13.39 Strain-rate dependence of (a) stress and (b) strain-rate sensitivity for Mg–Al eutectic alloy tested at 350 °C (grain size 10 μm). (After D. Lee, *Acta. Met.*, 17 (1969) 1057.)

a is the cross-sectional area in the neck region. The deformation continues to concentrate itself at the neck, with the attendant reduction in area caused by the constancy of volume. This leads to failure. Figure 13.38(b) shows the two alternative paths beyond the maximum in the engineering-stress–engineering-strain curve. Thus, one concludes that superplasticity is the result of the inhibition of necking as a result of a high value of m.

Under certain conditions of temperature and strain rate, some metals and ceramics exhibit an enhancement of m. Curve (a) in Figure 13.39 shows the stress as a function of strain rate for an Mg–Al eutectic alloy tested at 350 °C. One can see that in region II, the stress rises more rapidly with strain rate. Curve (b) in the same figure shows the strain rate sensitivity m as a function of strain rate. The maximum, $m = 0.6$, occurs for a strain rate of 10^{-2} s^{-1}. $m = \partial \ln \sigma / \partial \ln \dot{\varepsilon}$ is the slope of curve (a) in the figure. Figure 13.40 shows the variation in $\partial L / L_0$ (the tensile fracture strain) with $\dot{\varepsilon}$ for a Zr–22% Al alloy. The maxima (at the three temperatures) in $\Delta L/L_0$ correspond roughly to the center of region II. Further proof of the effect of m on the extent of superplastic flow is provided by Figure 13.41, which contains data from several studies. Data for alloys of Fe, Mg, Pu, Pb,–Sr, Ti, Zn, Zr are plotted, and the correlation is excellent. As m approaches unity, $\Delta L/L_0$ reaches extraordinarily high values.

The microstructural requirement for a high value of m is a small grain size. The testing temperatures are usually above $0.4T_m$, where T_m is the absolute melting point, and the strain rates in which superplasticity is observed are usually intermediate (10^{-4} s$^{-1} < \dot{\varepsilon} < 10$ s^{-1}). Superplasticity is usually enhanced by thermal cycling, i.e., straining the material sequentially at two different temperatures. All alloys that show structural superplastic behavior have a very fine grain size (<10 μm). For these small grain sizes, and at the deformation temperatures, most of the plastic deformation takes place by grain-boundary sliding, and not by the conventional dislocation mechanisms in the interior of the grains. Specimens deformed to very large strains routinely exhibit an equiaxial grain structure, in contrast with

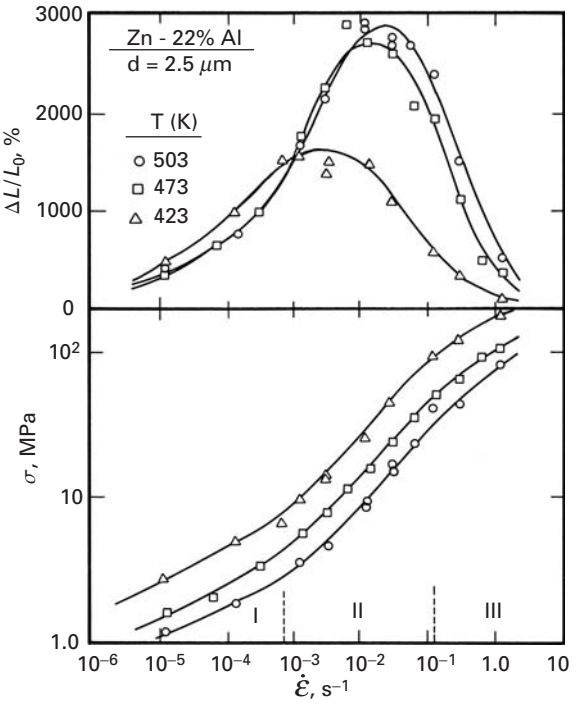


Fig. 13.40 Tensile fracture strain and stress as a function of strain rate for Zr–22% Al alloy with 2.5-μm grain size. (After F. A. Mohamed, M. M. I. Ahmed, and T. G. Langdon, *Met. Trans. A*, 8 (1977) 933.)

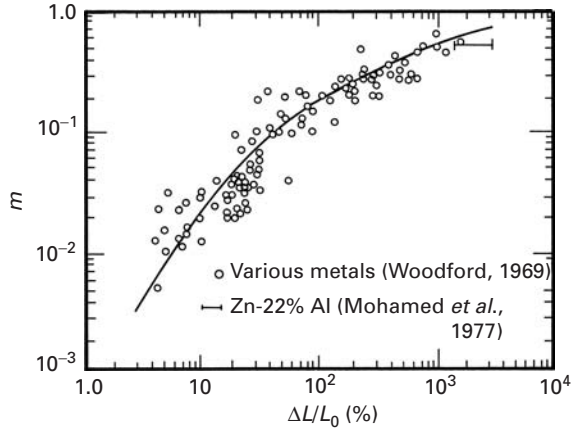


Fig. 13.41 Effect of strain-rate sensitivity *m* on maximum tensile elongation for different alloys (Fe, Mg, Pu, Pb–Sr, Ti, Zn, Zr based). (From D. M. R. Taplin, G. L. Dunlop, and T. G. Langdon, *Ann. Rev. Mater. Sci.*, 9 (1979) 151.)

conventional deformation, in which the strain undergone by the individual grains is equal to the overall strain, and the grains assume an elongated shape. Grain-boundary sliding accounts for 50–70% of the overall strain. Superplastic materials may be likened to sand: the granules retain their shape with plastic deformation. In contrast to sand, however, superplastic materials do not have interstices between the grains. Thus, some plastic accommodation must occur as the grains slide past each other. The contribution of grain-boundary sliding to plastic deformation is more substantial the greater the grain-boundary surface per unit volume. Since the grain-boundary surface is inversely proportional to the grain size, this explains why the

Table 13.2 | Materials Exhibiting Very High Tensile Strains[a]

Material	Maximum strain (%)
Al–33% Cu eutectic	1,500
Al–6% Cu–0.5% Zr	1,200
Al–10.7% Zn–0.9% Mg–0.4% Zr	1,500
Bi–44% Sn eutectic	1,950
Cu–9.5% Al–4% Fe	800
Mg–33% Al eutectic	2,100
Mg–6% Zn–0.6% Zr	1,700
Pb–18% Cd eutectic	1,500
Pb–62% Sn eutectic	4,850
Ti–6% Al–4% V	1,000
Zn–22% Al eutectoid	2,900
Al(6061)–20% SiC (whiskers)	1,400
Partially stabilized zirconia	120
Lithium aluminosilicate	400
Cu–10% Al	5,500
Zirconia	350
Zirconia + SiO_2	1,000

[a]Adapted from D. M. R. Taplin, G. L. Dunlap, and T. G. Langdon, *Ann. Rev. Mater. Sci*., 9 (1979) 15.

contribution of grain-boundary sliding is less important in materials with large grain sizes.

Table 13.2 shows the tensile elongation of a number of superplastic materials. It is interesting to note that superplasticity has been found in composites and ceramics. Sherby and co-workers obtained tensile elongations of 1,300% in an aluminum alloy/SiC whisker reinforced composite. The Sherby team used temperature cycling. These researchers also detected superplasticity in ultrahigh-carbon steel and were able to attribute the splendid properties of the Damascus sword to superplastic forming. Thus, the use of superplasticity is centuries old. For ceramics, superplasticity (in tension) is a technology with great potential. Wakai et al. obtained tensile elongations of 120% in an yttria-stabilized polycrystal containing 90% tetragonal zirconia and 10% cubic zirconia.[22] A grain size of 0.3 μm produced a strain rate sensitivity of $m = 0.5$ at 1450 °C. This elongation was exceeded in work done by Nieh et al., who obtained a value of 350% at 1,550°C.[23] Nanocrystalline ceramics (Section 5.5) are especially attractive in this regard. Superplasticity was also obtained in an aluminosilicate and other ceramic systems.

One of the major problems in superplastic forming is the formation of voids at grain boundaries. Cavitation during superplastic

[22] F. Wakai, S. Sakaguchi, and Y. Matsuno, *Advanced Ceramic Materials*, 1(3) (1986) 259.
[23] T. G. Nieh, C. M. McNally, and J. Wadsworth, *Scripta Met*., 22 (1988) 1297.

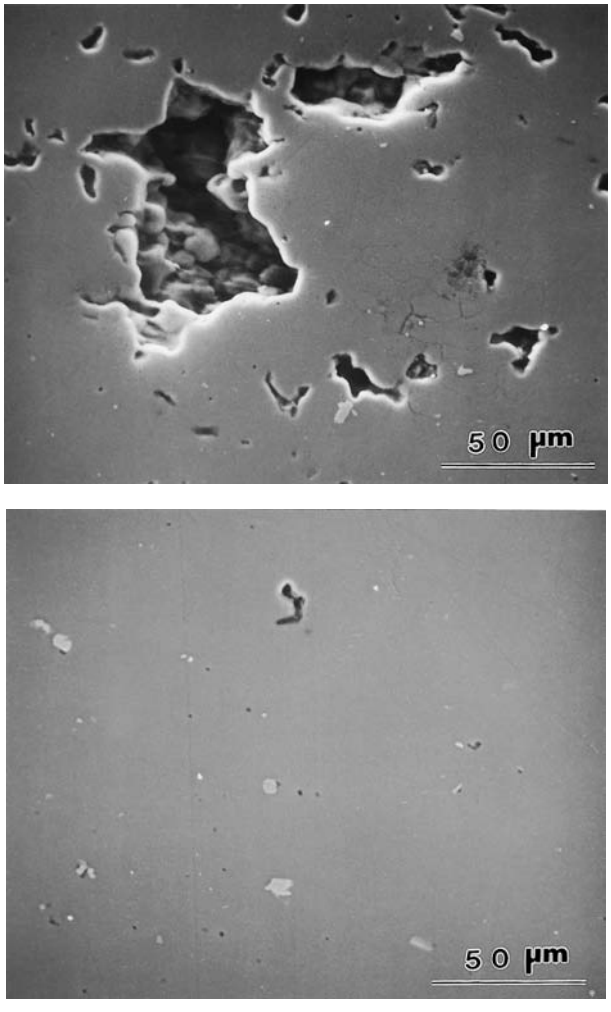
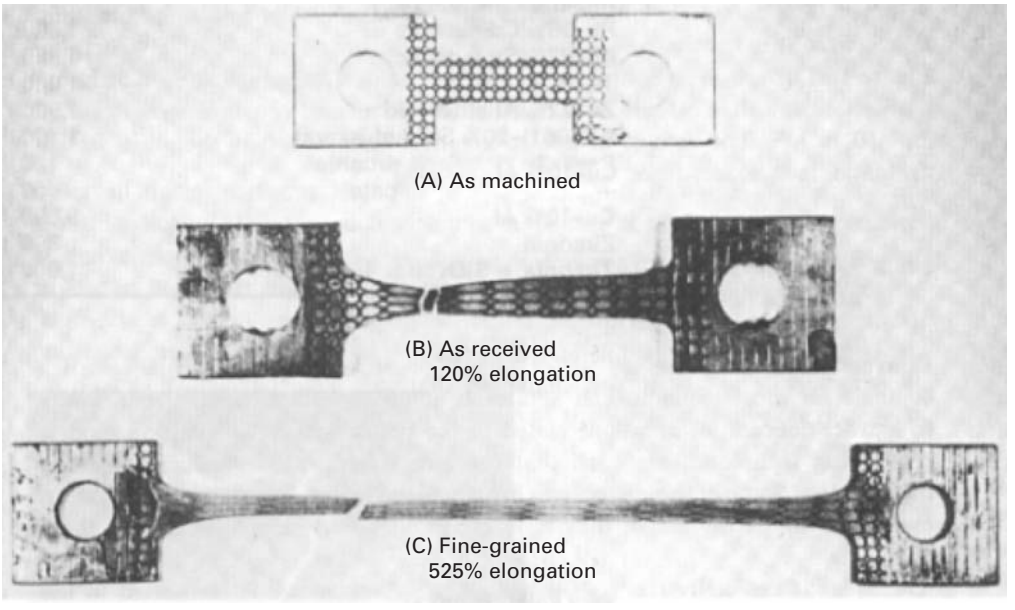


Fig. 13.42 Cavitation in superplastically formed 7475-T6 aluminum alloy ($\varepsilon = 3.5$) at 475 °C and 5×10^{-4} s^{-1}. (a) Atmospheric pressure. (b) Hydrostatic pressure $P = 4$ MPa. (Courtesy of A. K. Mukherjee.)

forming results in a deterioration of the mechanical properties of parts formed by superplasticity. Cavities form because of incompatible deformation of adjacent grains and weaken the material. These voids can be reduced or eliminated by superimposing a hydrostatic stress upon the applied tensile stress. This is illustrated in Figure 13.42. The aluminum alloy shown exhibited considerable cavitation at a plastic strain of 350%, or 3.5. The application of a superimposed hydrostatic pressure of 4 MPa, through a gaseous medium, decreases the cavitation substantially. Otherwise, the cavitation would lead to premature failure. Figure 13.43(a) shows the effect of grain size on the elongation of an 7475 Al alloy. Figure 13.43(b) shows a number of specimens superplastically deformed up to failure. The initial specimen is at the bottom, and the effect of increasing the superimposed pressure is shown from bottom to top. The pressures (and elongations to failure) are, respectively, 330% (atmosphere), 720% (1.4 MPa); 830% (2.8 MPa), and 1330% (and no failure at 5.6 MPa).

(a)



(b)

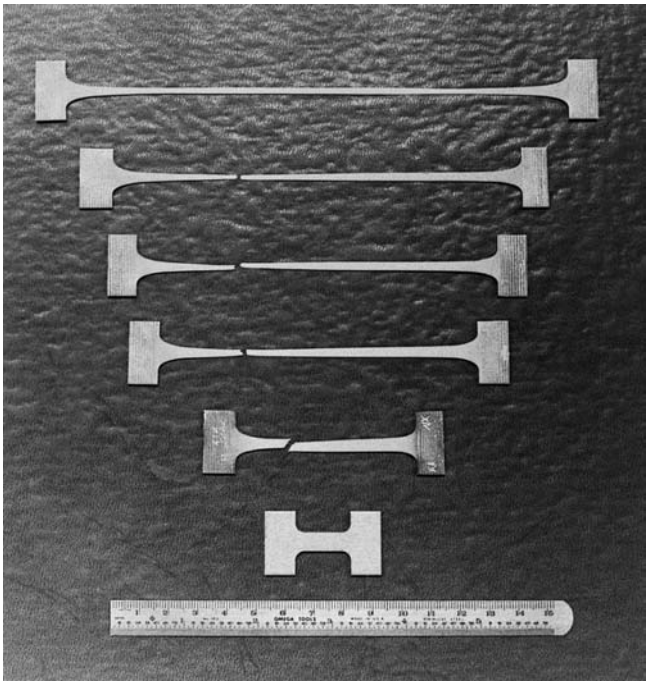


Fig. 13.43 (a) Effect of grain size on elongation: (A) Initial configuration. (B) Large grains. (C) Fine grains (10 μm) (Reprinted with permission from N. E. Paton, C. H. Hamilton, J. Wert, and M. Mahoney, *J. Metal, 34* (1981) No. 8, 21.) (b) Failure strains increase with superimposed hydrostatic pressure (from 0 to 5.6 MPa). (Courtesy of A. K. Mukherjee.)

Suggested Reading

W. R. Cannon and T. G. Langdon. "Creep of Ceramics." *J. Mater. Sci.*, 18 (1983) 1 (Part 1); 23 (1988) 1 (Part 2).

A. H. Chokshi, A. K. Mukherjee, and T. G. Langdon, "Superplasticity in Advanced Materials," *Matls. Sci and Eng. R: Reports*, 10 (1993) 237–274.

A. H. Chokshi and T. G. Langdon. "Characteristics of Creep Deformation in Ceramics." *Matls. Sci. Techn.*, 7 (1991) 577.

H. J. Frost and M. F. Ashby. *Deformation-Mechanism Maps*. Oxford: Pergamon Press, 1982.

F. Garofalo. *Fundamentals of Creep and Creep Rupture in Metals*. New York, NY: Macmillan, 1965.

J. Gittus. *Creep, Viscoelasticity and Creep Fracture in Solids*. New York, NY: Halsted Press (Wiley), 1975. *J. Eng. Mater. Technol.*, 101 (1979) 317.

B. P. Kashyap, A. Arieli, and A. K. Mukherjee. "Microstructural Aspects of Superplasticity." *J. Mater. Sci.*, 20 (1985) 2661.

M. E. Kassner and M. T. Perez-Prado, *Fundamentals of Creep in Metals and Alloys*. New York, NY: Elsevier, 2004.

F. R. N. Nabarro, and H. L. de Villiers. *The Physics of Creep*. London: Taylor & Francis, 1995.

J. P. Poirier. *Creep of Crystals: High Temperature Deformation Processes in Metals, Ceramics, and Minerals*. Cambridge, U.K.: Cambridge University Press, 1985.

O. D. Sherby and P. M. Burke. "Mechanical Behavior of Crystalline Solids at Elevated Temperature." *Progr. Mater. Sci.*, 13 (1967) 325.

J. Weertman, and J. R. Weertman. "Mechanical Properties, Strongly Temperature Dependent," in *Physical Metallury*, 4th ed., R. W. Cahn and P. Haasen. eds. New York, NY: Elsevier, 1995.

Exercises

13.1 A cylindrical specimen creeps at a constant rate during 10,000 hours when it is subjected to a constant load of 1,000 N. The initial diameter and length of the specimen are 10 and 200 mm, respectively, and the creep rate is 10^{-8} h^{-1}. Find:

(a) The length of the specimen after 10^4, 10^6, and 10^8 hours.
(b) The true and engineering strains after these periods.
(c) The true and engineering stresses after these periods.

13.2 Give three reasons why the extrapolation of creep data obtained over a short period can be dangerous over long periods.

13.3 By means of plots, show how isochronal stress-versus-strain curves can be constructed from creep curves for various stresses at a certain temperature.

13.4 Howson et al.[24] obtained the following stress-rupture results for the superalloy INCONEL MA 754 (a dispersion-strengthened alloy):

[24] T. E. Howson, D. A. Mervyn, and J. K. Tien, *Met. Trans.*, 11A (1980) 1609.

Temperature (°C)	Applied Stress (MPa)	Rupture Life (hours)
760	189.7	–
760	206.9	83.9
760	206.9	111.2
760	224.2	38.6
760	224.2	29.0
760	241.4	6.9
760	258.7	1.8
746	206.9	320.8
774	206.9	65.0
788	206.9	33.2
982	110.4	195.1
982	113.8	136.6
982	113.8	106.9
982	116.5	27.6
982	117.3	106.3
982	120.7	13.0
982	120.7	39.0
996	110.4	52.6
996	110.4	41.3
1010	110.4	20.3
1010	110.4	41.7
1024	110.4	9.4

(a) Verify whether this alloy obeys a Larson–Miller relationship, and find C. Then prepare a master plot, assuming that it does.

(b) Determine the predicted stress-rupture life if the alloy is stressed at 1,000 °C and 50 MPa.

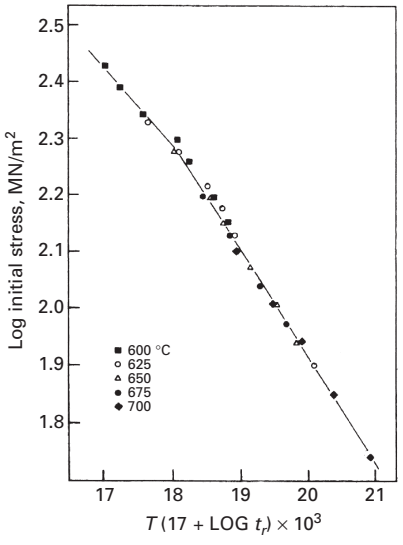


Fig. Ex13.5 Master plot for Larson–Miller parameter for AISI 316 stainless steel. (Courtesy of T. Silveira and S. N. Monteiro.)

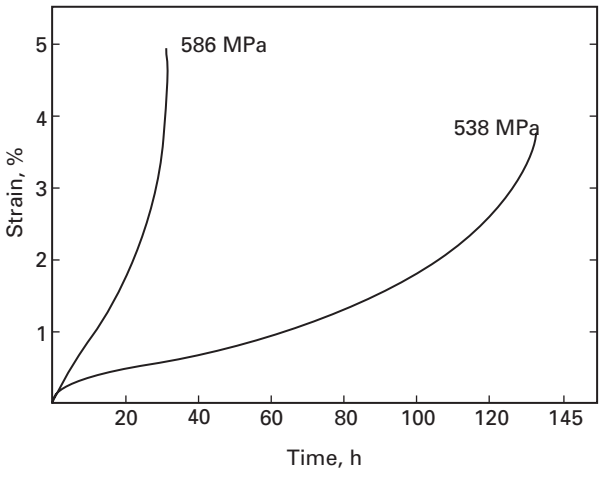


Fig. Ex13.8

13.5 What is the predicted stress-rupture life of AISI 316 steel (18% Cr, 8% Ni, 2–3% Mo) at 800 °C and 160 MPa? (See Figure Ex13.5.)

13.6 Verify whether the data of Exercise 13.5 obey the Manson–Haferd correlation.

13.7 Assuming that pure silver creeps according to the Dorn equation, estimate the rupture time at 400 °C when the silver is subjected to a stress of 50 MPa, knowing that at 300 °C and at the same stress level the rupture time is 2,000 hours.

13.8 Howson et al.[25] studied the creep and stress-rupture response of oxide-dispersion-strengthened (ODS) superalloys produced by mechanical alloying. They determined that the activation energy for creep Q_c was 619 kJ/mol by conducting tests at a constant applied stress of 558.7 MPa at the three temperatures of 746, 760, and 774 °C.

(a) The results shown in Figure Ex13.8 were found for experimental alloy MA 6000 E at 760 °C. Estimate the value of n, and discuss this value in terms of the microstructure exhibited by the alloy (made by means of dispersion-strengthening by inert yttrium oxide dispersoids plus precipitation-strengthening by gamma prime).

(b) By applying Equation 13.4, show how the activation energy can be found. Make the appropriate plot, and find the minimum creep rate at the aforementioned three temperatures. Note that the activation energy is given per mole.

13.9 In Exercise Ex13.9, verify how closely the Monkmon–Grant relationship is obeyed.

13.10 A lead-based alloy (melting point, 327.5 °C) was tested at ambient temperature (23 °C) and three different engineering stress levels: 8.5, 9, and 10 MPa. The curve that was obtained was shown in Figure Ex13.10.

(a) From the temperature aspect, establish whether the room temperature is in the creep domain for Pb.

(b) Obtain the minimum creep rates for the three stress levels.

[25] T. E. Howson, D. A. Mervyn, and J. K. Tien, *Met. Trans.*, 11A (1980) 1609.

Fig. Ex13.10

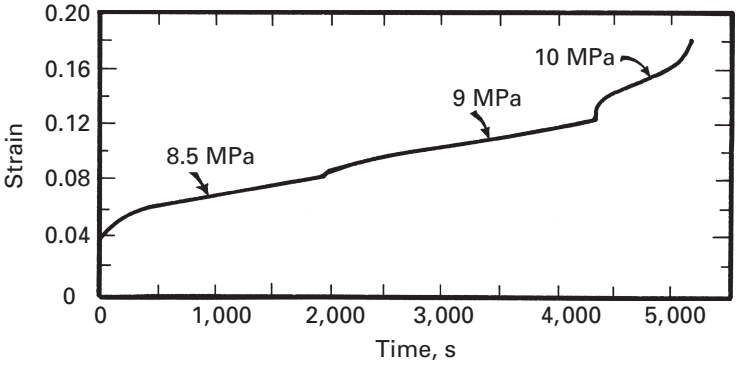
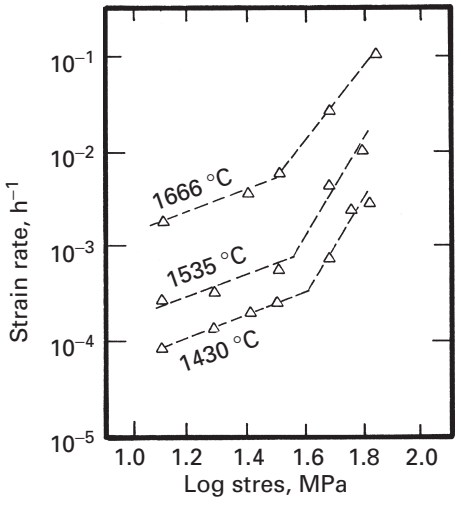


Fig. Ex13.12 Steady-state creep behavior of UO_2 can be divided into two regimens with different stress exponents. The transition stress between the regimens decreases with increasing temperature. (From L. E. Poteat and C. S. Yust, in *Ceramic Microstructures*, R. M. Fulrath and J. A. Pask, eds. (New York: John Wiley & Sons, 1968) 646.)



(**c**) Obtain parameters for the curve at 8.5 MPa, as expressed by Equation 13.1.
(**d**) Obtain the stress exponent in the Mukherjee–Bird–Dorn equation. Based on this value, what mechanism of creep do you expect?

13.11 Using the results from the previous problem, predict the minimum creep rate for the same material if the test would be carried out at 10 °C and stress levels of (a) 10 MPa and (b) 20 MPa. The self-diffusion coefficients for Pb are:

$$D = 5.24 \times 10^{-8} \text{ cm}^2 \cdot \text{h}^{-1} \text{ at } 250^\circ\text{C},$$
$$D = 2.92 \times 10^{-7} \text{ cm}^2 \cdot \text{h}^{-1} \text{ at } 285^\circ\text{C}.$$

13.12 Determine the slopes for the creep behavior of UO_2, shown in Figure Ex13.12. Discuss the deformation mechanisms in the two regions.

13.13 Tungsten is being used at half its melting point ($T_m \approx 3{,}400$ °C) and a stress level of 160 MPa. An engineer suggests increasing the grain size by a factor of 4 as an effective means of reducing the creep rate.

(**a**) Do you agree with the engineer? Why? What if the stress level were equal to 1.6 MPa?
(**b**) What is the predicted increase in length of the specimen after 10,000 hours if the initial length is 10 cm?

(*Hint*: Use a Weertman–Ashby map.)

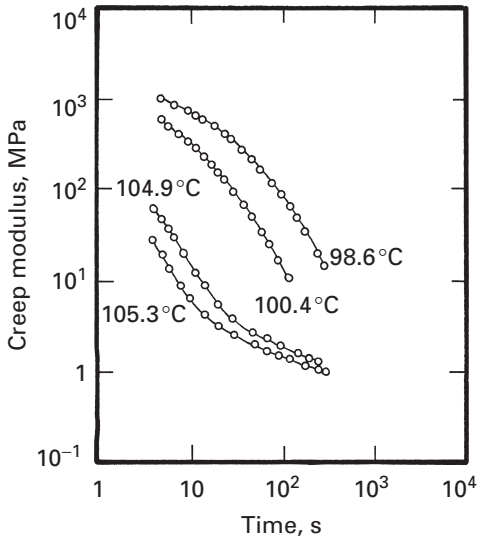


Fig. Ex13.15 (After W. Kurz, J. P. Mercier, and G. Zambelli, *Introduction aux Sciences des Matériaux* (Lausanne, Switzerland: Presses, 1987), p. 287.)

13.14 Stress relaxation in a polymer results from molecular displacements. Thus, one would expect that the effect of temperature on stress relaxation would be similar to that of any other thermally activated process. As described in the text, an Arrhenius-type expression describes the temperature dependence of such phenomena. Thus, the relaxation time is the inverse rate, or

$$\frac{1}{\tau} \propto \exp(-Q/kT),$$

where the symbols have their usual significance. Describe how you would determine, from this expression, the activation energy Q of the molecular process causing the relaxation.

13.15 Curves of creep modulus (the inverse of creep compliance) vs. time for four different temperatures are shown in Figure Ex13.15. Obtain a master curve for the polymer at a reference temperature of 101.6 °C.

13.16 An amorphous polymer has a glass transition temperature of 100 °C. A creep modulus of 1 GPa was measured after 1 hour at 75 °C. Using the Williams–Landel–Ferry expression, determine the time required to reach this modulus at 50 °C.

13.17 The viscosity of an amorphous polymer is 10^5 Pa · s at 190 °C and 2×10^2 Pa · s at 270 °C. At what temperature will the viscosity be 10^9 Pa · s?

13.18 What is the strain undergone by a polymer in tension at 67 °C for one minute if the polymer's strain-rate response is given by $\dot{\varepsilon} = 4.5 \times 10^{28} \exp(-200\ \text{kJ}/RT)$?

13.19 A nylon cord has an initial stress of 2 MPa and is used to tie a sack. If the relaxation time for this cord is 250 days, how many days will it take for the stress to drop to 0.1 MPa?

13.20 How much time will it take for a rod of polymer to extend from 20 mm to 30 mm at 120 °C if it is deformed at a strain rate $\dot{\varepsilon} = 4.5 \times 10^{11} \exp(-100\ \text{kJ}/RT)$?

13.21 Find the initial stress for a nylon cord if the relaxation time for the cord is 100 days and in 50 days the stress is reduced to 1 MPa.

13.22 (a) Determine the strain-rate sensitivity in the superplastic range for the alloys shown in Figures 13.39(a) and 13.40, and explain the values encountered. (b) Why does the maximum in ductility vary with temperature in Figure 13.40?

13.23 Explain why the presence of voids decreases the maximum strain in superplastic deformation.

13.24 A polymer has a viscosity of 10^{12} Pa·s at 150 °C. If this polymer is subjected to a tensile stress of 100 MPa at that temperature, compute the deformation after 10 h. Assume the polymer to behave as a Maxwell solid. Take $E = 5$ GPa, and use the equation $\mathrm{E} = \frac{\sigma}{E} + \frac{1}{3\eta}\sigma t$.

13.25 For a given polymer, the activation energy for stress relaxation was measured to be 10 kJ/mol. If the stress relaxation time for this polymer at room temperature is 3,600 s, what would be the relaxation time at 100 °C?

13.26 In a laboratory experiment for potential creep deformation, it was found that the creep rate ($\dot{\varepsilon}$) of an alloy is 0.5% per hour at 780 °C and 2.8×10^{-2}% per hour at 650 °C.

(a) What is the activatuion energy for creep in the given temperature range?
(b) At a temperature of 550 °C, what is the estimated creep rate?

13.27 A component used in a chemical plant is being used at 600 °C and stress of 25 MPa. The corresponding creep rate is: 3×10^{-12} s $^{-1}$. If the stress is increased to 35 MPa and the temperature to 650 °C, what will be the corresponding creep rate?
Given:

$Q = 150$ kJ/mol,
$n = 4.5$ (stress exponent).

13.28 Honey has a viscosity of 1.5 Pa·s at room temperature. If the activation energy of honey is 20 kJ/mol, what will be its viscosity at 0 °C?

13.29 From the plot in Figure Ex13.29 on p. 711, determine the amount of strain that a titanium component will undergo if subjected to a stress of 30 MPa at a temperature of 700 °C for 1 hour.
Given:

$D_0 = 1$ cm^2 s^{-1},
G (shear modulus) $= 43.8$ GPa,
$Q = 15$ kJ/mol,
$b = 0.3$ nm.

13.30 Give an example of the creep phenomenon in nature (without interference of human paws).

13.31 Using the parameters in Figure 13.6, calculate the rupture time by the Larson–Miller equation: (a) at a temperature of 1089 K and stress of 100 MPa; (b) at a temperature of 1005 K and stress of 200 MPa; (c) at a temperature of 922 K and stress of 300 MPa.

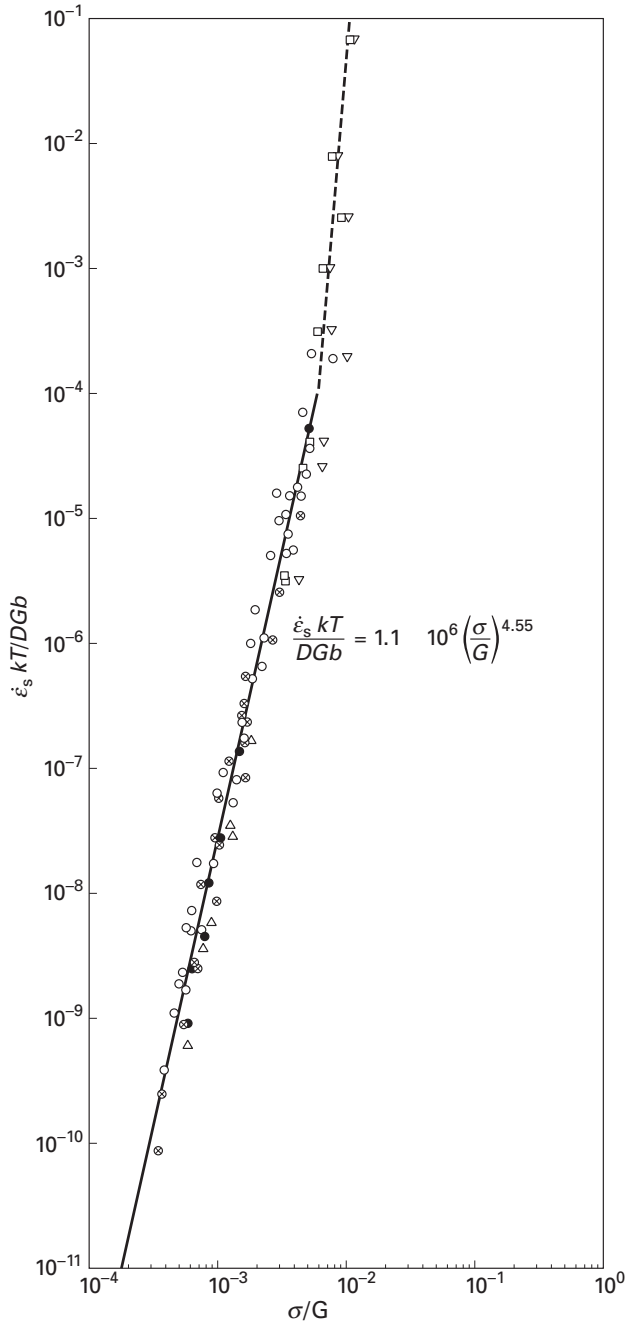


Fig. Ex13.29 Plot of normalized strain rate $\dot{\varepsilon}kT/DGb$ versus normalized stress σ/G for the high-temperature plastic deformation of alpha titanium. (Adapted from H. Conrad, *Prog. Mater. Sci*, 26 (1981) 375.)

13.32 At room temperature, the relaxation time for a polymer is 100 days. What will be its relaxation time if the activation energy is one-fourth of its current value?

13.33 Calculate the creep compliance for stainless steel with a relaxation time of 72 hours.

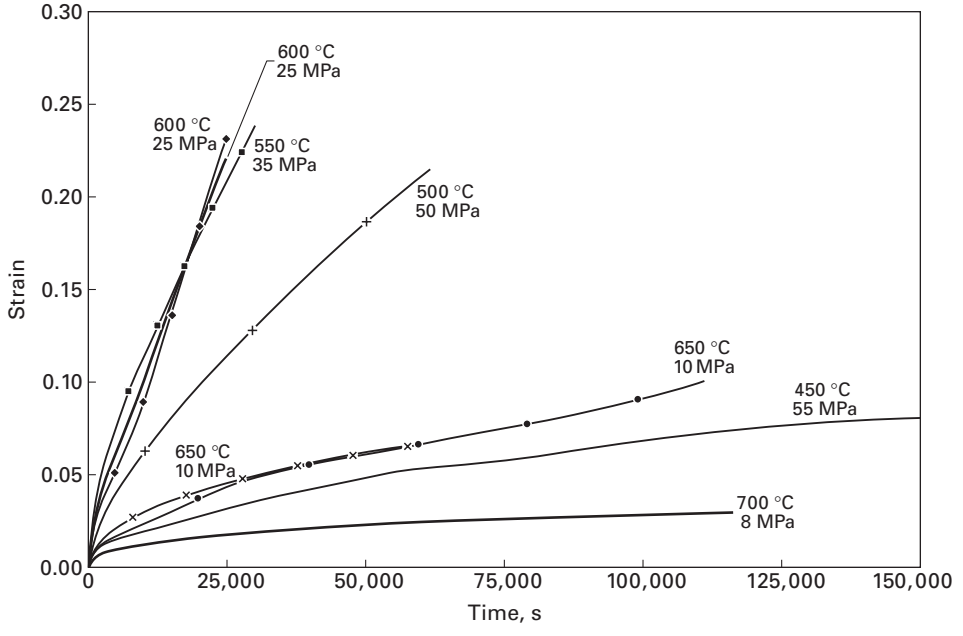


Fig. Ex13.35 Creep curves for Zircaloy at different temperatures and stresses. (Courtesy of M. E. Kassner.)

13.34 Give an example of a superplastic material and find (and describe) an application for it. A web-engine search is recommended.

13.35 **(a)** From the different creep tests carried out on Zircaloy, Figure Ex13.35, determine the stress exponent.
Given:

$G(T) = 36.27$ (GPa) $-0.02T$ (°C),
$D_0 = 5 \times 10^{-4}$ m^2/s,
$Q = 270$ kJ/mol,
$b = 3.23 \times 10^{-10}$ m.

(b) Knowing the rupture time at 600 °C and 25 MPa is 10 hours, calculate the rupture time for the other conditions in Figure Ex13.35 using the Monkman–Grant equation.

Chapter 14

Fatigue

14.1 Introduction

There is some confusion in the literature about the terminology pertaining to fatigue. We define fatigue as a degradation of mechanical properties leading to failure of a material or a component under *cyclic* loading. This definition excludes the so-called phenomenon of static fatigue, which is sometimes used to describe stress corrosion cracking in glasses and ceramics in the presence of moisture. Brittle solids (glasses and crystalline ceramics) undergo subcritical crack growth in an aggressive environment under static loads. Silica-based glasses are especially susceptible to this kind of crack growth in the presence of moisture. If a glassy phase exists at grain boundaries and interfaces, it will be susceptible to such an attack. Thus, static fatigue is more appropriately a stress corrosion phenomenon, rather than a cyclic stress-related phenomenon.

In general, fatigue is a problem that affects any structural component or part that moves. Automobiles on roads, aircraft (principally the wings) in the air, ships on the high sea constantly battered by waves, nuclear reactors and turbines under cyclic temperature conditions (i.e., cyclic thermal stresses), and many other components in motion are examples in which the fatigue behavior of a material assumes a singular importance. It is estimated that 90% of service failures of metallic components that undergo movement of one form or another can be attributed to fatigue. Often, a fatigue fracture surface will show some easily identifiable macroscopic features, such as beach markings. Figure 14.1 shows a schematic of the fracture surface of, say, a steel shaft that failed in fatigue. The main features of this kind of failure are a fatigue crack initiation site, generally at the surface; a fatigue crack propagation region showing beach markings; and a fast-fracture region where the crack length exceeds a critical length. Typically, the failure under cyclic loading occurs at much lower stress levels than the strength under monotonic loading.

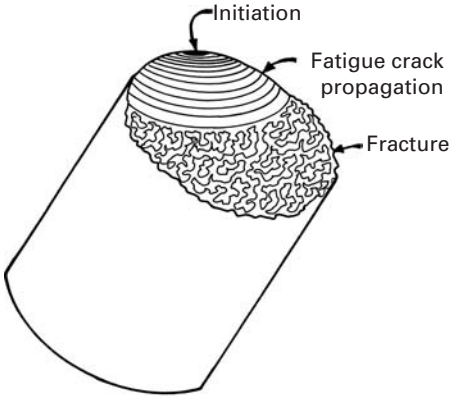


Fig. 14.1 Schematic representation of a fatigue fracture surface in a steel shaft, showing the initiation region (usually at the surface), the propagation of fatigue crack (evidenced by beach markings), and catastrophic rupture when the crack length exceeds a critical value at the applied stress.

One can divide the study of cyclic behavior of materials into the following three classes:

- Stress-life approach
- Strain-life approach
- Fracture mechanics approach.

The stress-life approach is the oldest way of treating the cyclic fatigue data. It is useful when stresses and strains are mostly elastic. The main drawback of this approach is that we are unable to distinguish between the initiation and propagation phases of fatigue life. The strain-life approach is useful when there is a significant amount of plastic strain. The fatigue life is typically quite short under these conditions. In the fracture mechanics approach, we apply the basic ideas of fracture mechanics to cyclic fatigue, i.e., we use the cyclic stress intensity factor as the crack driver. It allows us to estimate the life spent in propagating a crack from an initial size to larger size or to the critical size corresponding to failure.

In this chapter, we present a basic description of the various aspects of fatigue in different materials, followed by a brief examination of the various fatigue-testing techniques.

14.2 Fatigue Parameters and *S–N* (Wöhler) Curves

We first define some important parameters that will be useful in the subsequent discussion of fatigue. These parameters, shown in Figure 14.2, are as follows:

$$\begin{aligned} &\text{cyclic stress range} && \Delta\sigma = \sigma_{\max} - \sigma_{\min'}, \\ &\text{cyclic stress amplitude} && \sigma_a = (\sigma_{\max} - \sigma_{\min})/2, \\ &\text{mean stress} && \sigma_m = (\sigma_{\max} + \sigma_{\min})/2, \\ &\text{stress ratio} && R = \sigma_{\min}\sigma_{\max}, \end{aligned}$$

where $\sigma_{\max}$ and $\sigma_{\min}$ are the maximum and minimum stress levels, respectively.

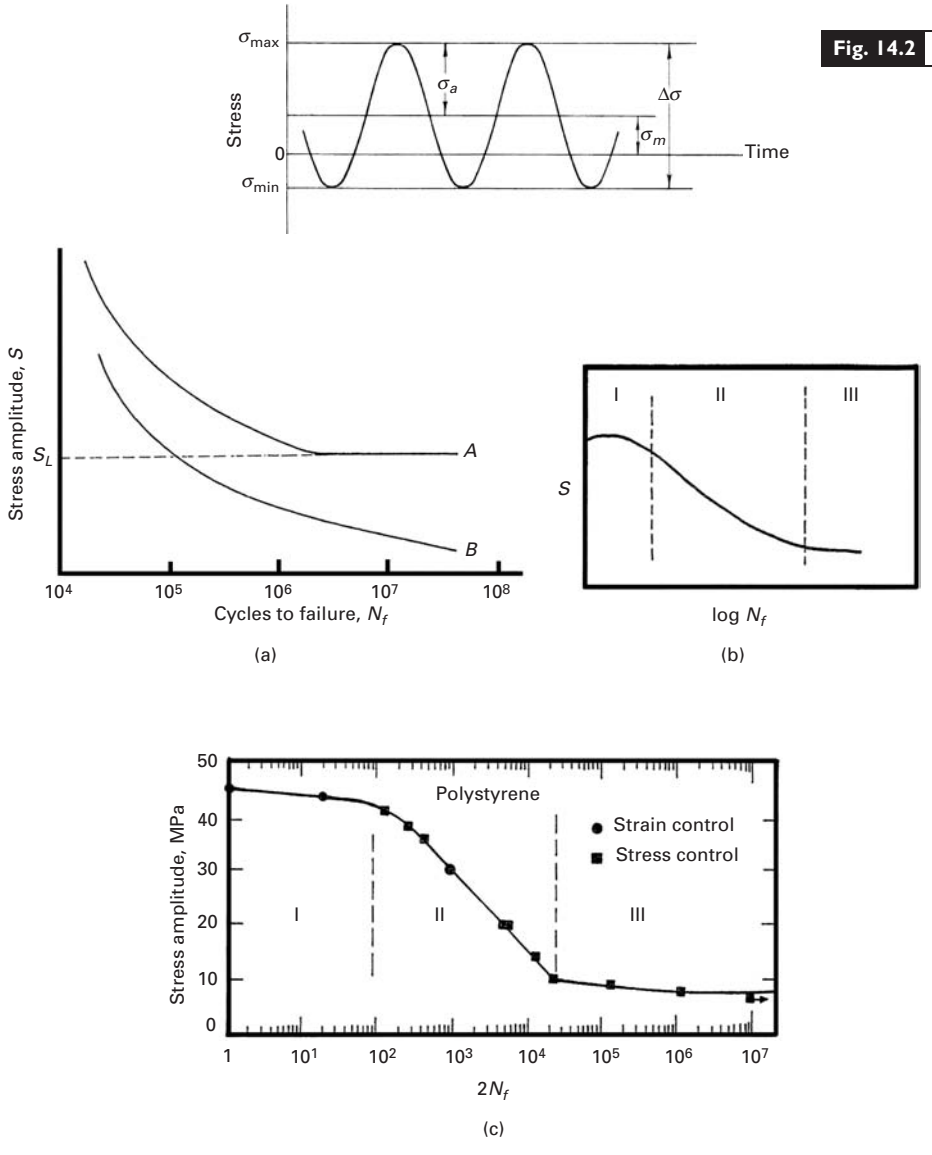


Fig. 14.2 Fatigue parameters.

Traditionally, the behavior of a material under fatigue is described by the *S–N* (or $\sigma - N$) curves (Figure 14.3), where S (or σ) is the stress and N is the number of cycles to failure. Such an *S–N* curve is frequently called a *Wöhler curve*, after the German engineer who first observed that kind of fatigue behavior in railroad car wheels in the 1860s. For steels, in general, one observes a fatigue limit or endurance limit (curve *A* in Figure 14.3(a)), which represents a stress level below which the material does not fail and can be cycled indefinitely. Such an endurance limit does not exist for nonferrous metals (curve *B* in the figure). Polymeric materials show essentially similar *S–N* curves. Figure 14.3(b) shows a schematic of an *S–N* curve for a variety of polymers. Polymers that form crazes, such as polymethylmethacrylate (PMMA) and polystyrene (PS), may show a flattened portion in the very beginning, indicated as stage I in the figure. In region II, the stress is not

Fig. 14.3 (a) *S* (stress)–*N* (cycles to failure) curves. (A) Ferrous and (B) nonferrous metals; S_L is the endurance limit. (b) *S–N* curves for polymeric materials. Polymers that form crazes, such as polymethylmethacrylate (PMMA) and polystyrene (PS), may show a flattened portion in the very beginning, indicated as stage I. (c) An example of an actual *S–N* curve showing the three stages in the case of polystyrene.

high enough for crazes to form. Crazed regions are, of course, the sites of microcrack nucleation. They will form in the initial quarter of a tensile cycle in such materials. Recall that the crazes do not form in compression. Such a flat region does not exist for polymers that do not show craze formation, and the *S–N* curve for such polymers will be very similar to that of metals; that is, stage I is simply an extension of stage II. An example of an actual *S–N* curve showing the three stages in the case of polystyrene is presented in Figure 14.3(c). Polystyrene shows extensive crazing at room temperature. Polycarbonate, on the other hand, does not show crazing at room temperature, and its *S–N* curve does not show stage I. No such *S–N* curves are available for ceramics, although, as we shall see later in the chapter, stable subcritical crack propagation under cyclic fatigue can occur in ceramics. Note that the relationship between *S* and *N* is not a single-valued function, but serves to indicate a statistical tendency. Also, the fatigue life determined in terms of *S–N* curves cannot be separated into the initiation and propagation parts of fatigue. Figure 14.4 shows representative *S–N* curves for (a) metals and (b) polymers. It is apparent that the endurance life varies widely. There is a dramatic difference between, for instance, 1.2% C steel and gray cast iron.

14.3 Fatigue Strength or Fatigue Life

Traditionally, fatigue life has been presented in the form of an *S–N* curve (Figure 14.3). With regard to this measure, *fatigue strength* refers to the capacity of a material to resist conditions of cyclic loading. However, in the presence of a measurable plastic deformation, materials respond differently to strain cycling than to stress cycling. Thus, one would expect that the fracture response of a material under cyclic conditions would show a similar difference. In this section, we treat fatigue life in terms of strength or strain versus number of cycles to failure N_f or number of reversals to failure, $2N_f$. It is convenient to consider separately the elastic and plastic components of strain. The elastic component can be readily described by means of a relation between the stress amplitude and the number of reversals (i.e., twice the number of cycles). This is called the Basquin[1] relationship. It may be expressed as:

$$\boxed{\sigma_a = \sigma_f'(2N_f)^b}.$$

Since the deformation is elastic, we can write:

$$\sigma_a = \frac{\Delta\sigma}{2} = \frac{\Delta\varepsilon_e E}{2}.$$

Thus:

$$\Delta\varepsilon_e/2 = \sigma_a/E = (\sigma_f'/E)(2N_f)^b,$$

[1] O. H. Basquin, *Proc. Am. Soc. For Testing and Matls.*, 10 (1910) 625.

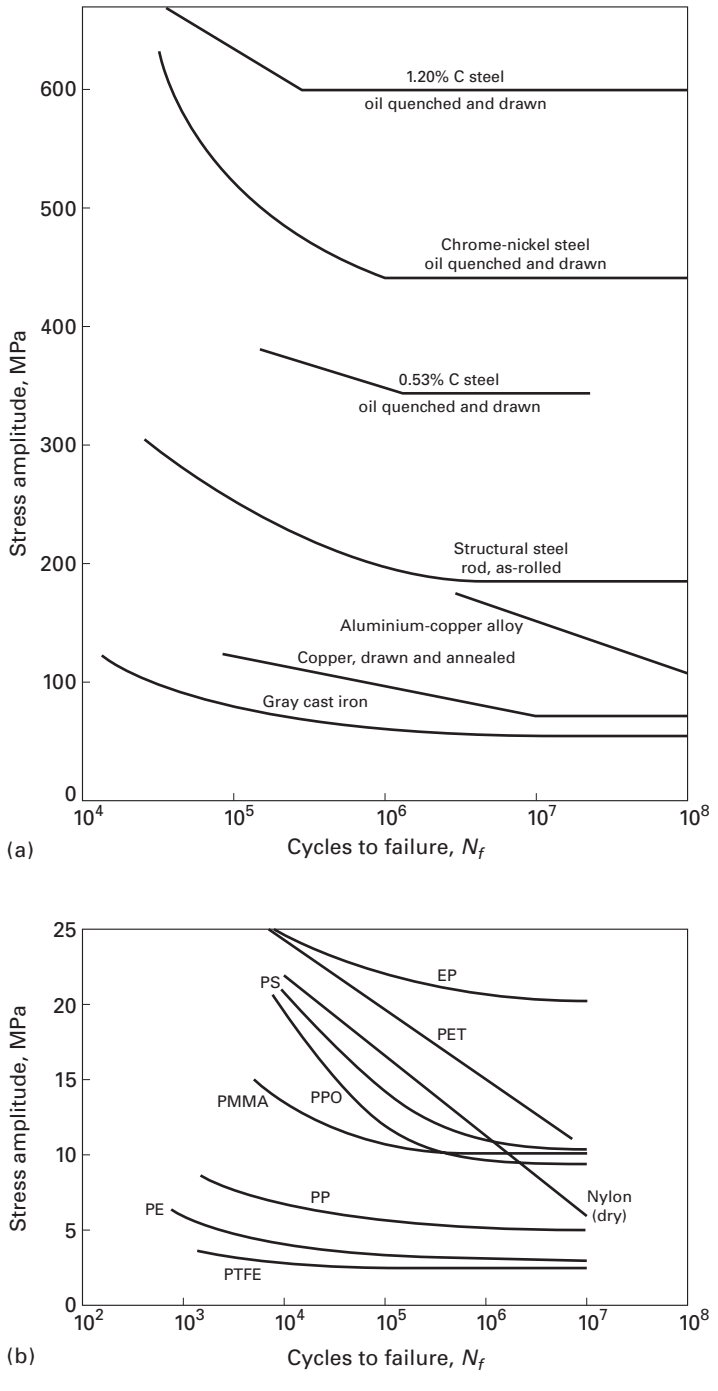


Fig. 14.4 *S–N* curves for typical (a) metals and (b) polymers.

where $\Delta\varepsilon_e/2$ is the elastic strain amplitude, σ_a is the true stress amplitude, σ_f' is the fatigue strength coefficient (equal to the stress intercept at $2N_f = 1$), N_f is the number of cycles to failure, and b is the fatigue strength exponent. This relationship is an empirical representation of the *S–N* curve above the fatigue limit in Figure 14.3. On a log–log plot, it gives a straight line of slope b. One can see that the

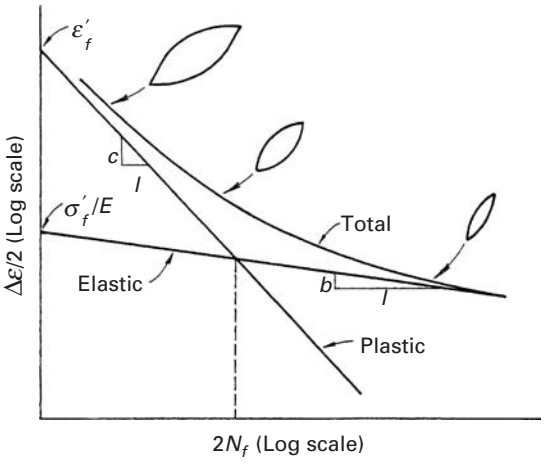


Fig. 14.5 Superposition of elastic and plastic curves gives the fatigue life in terms of total strain. (Adapted with permission from R. W. Landgraf, in *American Society for Testing and Materials, Special Technical Publication* (*ASTM STP*) 467 (Philadelphia: ASTM, 1970), p. 3.)

two expressions are equivalent. In elastic deformation, σ_a/E is equal to the elastic strain amplitude, $\Delta\varepsilon_e/2$.

The plastic strain component is better described by the Manson–Coffin relationship,[2]

$$\Delta\varepsilon_p/2 = \sigma_a/E = \varepsilon'_f(2N_f)^c,$$

where $\Delta\varepsilon_p/2$ is the amplitude of the plastic strain, ε'_f is the ductility coefficient in fatigue (equal to the strain intercept at $2N_f = 1$), $2N_f$ is the number of reversals to failure, and c is the ductility exponent in fatigue. On a log–log plot, the Manson–Coffin relation gives a straight line of slope c. It has been observed that a smaller value of c results in a longer fatigue life. In the regimen of high-strain, low-cycle fatigue, the Manson–Coffin relation assumes great importance. Experimentally, it is frequently more convenient to control the total strain. In many structural components, the material in a critical place (say, at a notch root) may be subjected, essentially, to strain control conditions due to the elastic constraint of the surrounding material. For a material subjected to a total strain range of $\Delta\varepsilon_t$ (elastic and plastic strain), we can determine the fatigue strength by a superposition of the elastic and plastic strain components, i.e.,

$$\Delta\varepsilon_t/2 = \Delta\varepsilon_e/2 + \Delta\varepsilon_p/2 = (\sigma'_f/E)(2N_f)^b + \varepsilon'_f(2N_f)^c.$$

Thus, we expect that the curve of the fatigue life, in terms of total strain, will tend to the plastic curve at large total-strain amplitudes, whereas it will tend to the elastic curve at low total-strain amplitudes, as shown schematically in Figure 14.5. An example of such a behavior from a real material (an 18%-Ni maraging steel) is shown in Figure 14.6. The Manson–Coffin regimen is known as *low-cycle fatigue*, in which there is a discrete amount of plastic deformation in each cycle. The Basquin regimen is called *high-cycle fatigue*. The stresses are primarily

[2] L. F. Coffin, *Trans. ASME*, 76 (1954) 931; S. M. Manson and M. H. Hirschberg, in *Fatigue: An Interdisciplinary Approach* (Syracuse, NY: Syracuse University Press, 1964), p. 133.

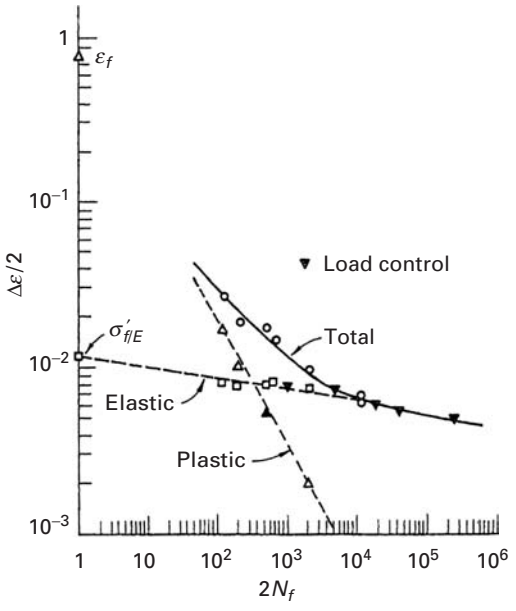


Fig. 14.6 Fatigue life in terms of strain for an 18%-Ni maraging steel. (Adapted with permission from R. W. Landgraf, in *American Society for Testing and Materials, Special Technical Publication ASTM STP* 467, (Philadelphia: ASTM, 1970), p. 3.)

elastic in this domain. The typical range of values for the exponents b (high-cycle fatigue) and c (low-cycle fatigue) are:

$$b : \frac{1}{8} \rightarrow \frac{1}{5} \text{ (Basquin exponent)},$$
$$c : 0.5 \rightarrow 0.6 \text{ (Manson-Coffin exponent)}.$$

14.4 Effect of Mean Stress on Fatigue Life

The mean stress σ_m can have an important effect on the fatigue strength of a material. A simple and crude way to demonstrate the effect of σ_m would be to present *S–N* curves of a given material for different values of σ_m on the same graph. Figure 14.7 shows such curves schematically. Note that, for a given stress amplitude σ_a, as the mean stress increases, the fatigue life decreases.

We can describe the effect of σ_m in a very simple manner. Suppose that the limiting value of any combination of stresses is σ_f, the monotonic true fracture stress. We can think of other arbitrary limits, such as the ultimate tensile stress σ_{UTS} or the yield stress σ_y, but σ_f is the maximum allowable true stress. Figure 14.8(a) shows a schematic plot of alternating stress σ_a (or S) versus σ_m. Note that, for $\sigma_m = 0$, the alternating stress σ_a is a maximum and equal to σ_f. For $\sigma_a = \sigma_f$, the fatigue life is simply one-fourth of a cycle. For an ideal material, one would expect the relationship $\sigma_a + \sigma_m \leq \sigma_f$, the limiting value of any combination of stresses, to be valid. Thus, one can expect a straight line to join points A and B in Figure 14.8(a). Cyclic loading is not possible to the right of line AB. Then, in the presence of a mean stress of, say, $\sigma_f/3$, we will have a maximum allowable stress equal to

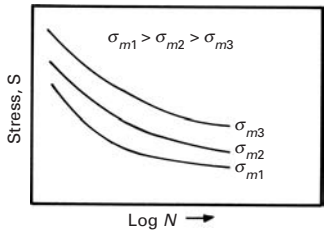


Fig. 14.7 Effect of mean stress on *S–N* curves. The fatigue life decreases as the mean stress increases.

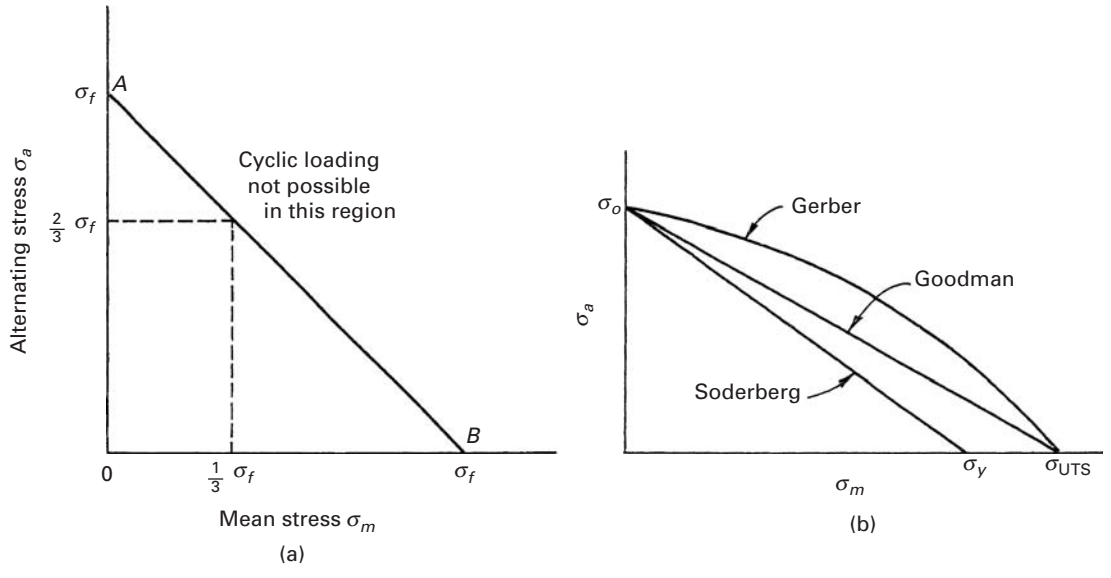


Fig. 14.8 (a) Effect of mean stress on fatigue life. (b) Gerber, Goodman, and Soderberg diagrams, showing mean stress effect on fatigue life.

$2\sigma_f/3$. Note that this description is an oversimplification of the real behavior of the material, in as much as it assumes that the damage produced in each cycle by cyclic plastic strain is independent and noncumulative.

Various empirical expressions have been proposed which take into account the effect of mean stress on fatigue life. Some of these are the following: Goodman's relationship, which assumes a linear effect of mean stress between $\sigma_m = 0$ and σ_{UTS}:

$$\sigma_a = \sigma_0[1 - \sigma_m/\sigma_{\text{UTS}}].$$

Gerber's relationship, which assumes a parabolic effect of mean stress between $\sigma_m = 0$ and σ_{UTS}:

$$\sigma_a = \sigma_0[1 - (\sigma_m/\sigma_{\text{UTS}})^2].$$

Soderberg's relationship, which assumes a linear effect of mean stress between $\sigma_m = 0$ and σ_y:

$$\sigma_a = \sigma_0(1 - \sigma_m/\sigma_y).$$

In all these relationships, σ_m is the mean stress, σ_a is the fatigue strength in terms of stress amplitude when $\sigma_m \neq 0$, σ_0 is the fatigue strength in terms of stress amplitude when $\sigma_m = 0$, σ_{UTS} is the monotonic ultimate tensile strength, and σ_y is the monotonic yield stress. Figure 14.8(b) shows the three relations schematically. Experimentally, it has been observed that the great majority of data falls between the Gerber and Goodman lines. Thus, the Goodman diagram represents a conservative estimate of the mean stress effect. Note that the three expressions involve uniaxial stresses. In most real-life situations, one encounters biaxial or triaxial situations. Hence, one needs to define stresses corresponding to σ_y or σ_{UTS} under multiaxial stress situations. A practical way around this is to use the concept of equivalent

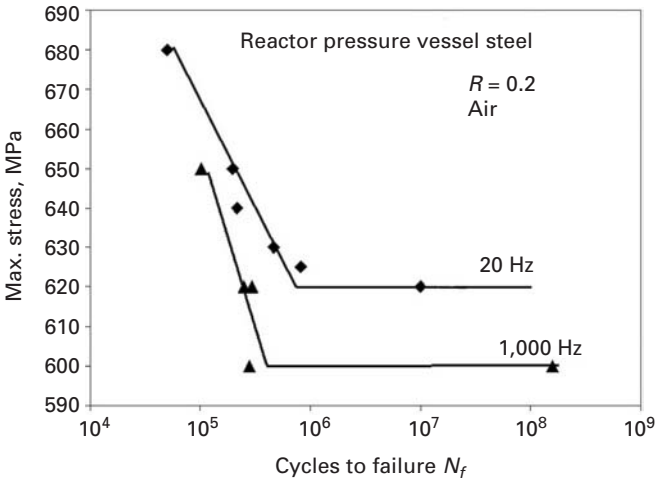


Fig. 14.9 Effect of frequency on the fatigue life of a reactor pressure vessel steel. The fatigue life decreases at 1,000 Hz compared with that at 20 Hz. (Used with permission from P. K. Liaw, B. Yang, H. Tian *et al.*, *ASTM STP* 1417 (West Conshohocken, PA: American Society for Testing and Materials, 2002.))

distortion energy – that is, to compute the distortion energy for uniaxial and multiaxial states and (a) assume that the cycling behavior of a material is equivalent when the material is cycled between two energy distortion values and (b) compute the maximum value of the mean stress by use of von Mises yield criterion. (See Chapter 3.) Then we can find the failure conditions for a given multiaxial stress state.

14.5 Effect of Frequency

Frequency of cycling can have an effect on the fatigue behavior of a material. Figure 14.9 shows the *S–N* curves for a steel used for nuclear reactor pressure vessels at 20 Hz and 1,000 Hz. Note that increasing the test frequency from 20 Hz to 1,000 Hz resulted in lower fatigue life at a given stress level. The decrease of the fatigue life at higher frequencies is attributed to the temperature increase that results in the higher-frequency tests. We should mention that such high-frequency fatigue tests are done in very sophisticated machines. Special servovalves, activated by voice coils, allow frequencies of up to 1,000 Hz to be obtained. Typically, such a machine is housed in a separate, soundproof room with a heat pump that allows cooling, thus preventing overheating of the servovalves.

14.6 Cumulative Damage and Life Exhaustion

The discussion in the preceding sections was restricted to fatigue under simple conditions of constant amplitude, constant frequency, and so on. In real life, the service conditions are rarely so simple. Many components and structures are subject to a range of fluctuating loads, mean stress levels, and variable frequencies. Thus, it is of great importance to be able to predict the life of a component subjected to variable-amplitude conditions, starting from data obtained in simple

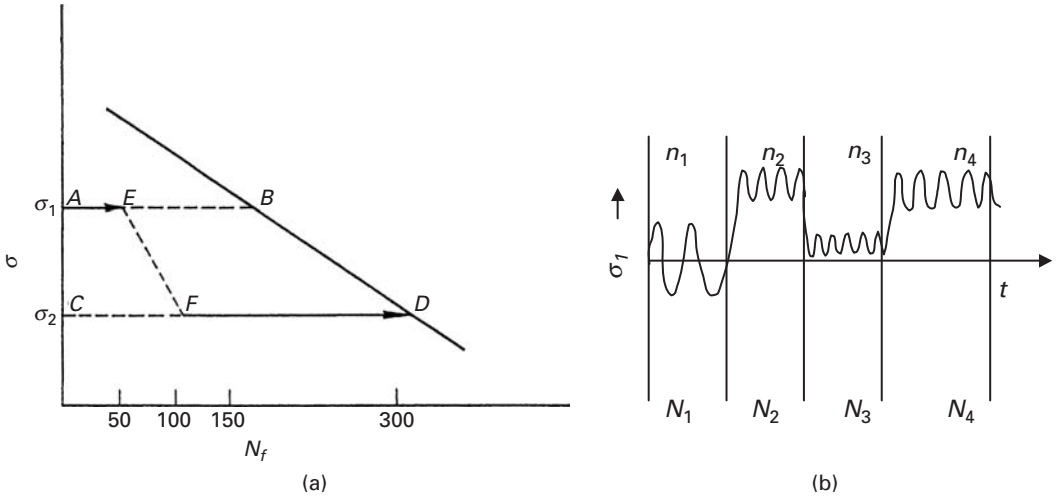


Fig. 14.10 (a) Damage accumulation, in a high-to-low loading sequence. (Adapted with permission from B. I. Sandor, *Fundamentals of Cyclic Stress and Strain* (Madison, WI: University of Wisconsin Press, 1972.)) (b) Sequence of block loadings at four different mean stresses and amplitudes.

constant-amplitude tests. The cumulative-damage theories attempt to do just that.

Basically, these theories consider fatigue to be a process of accumulating damage in a material until a certain maximum tolerable damage is reached. In other words, the phenomenon of fatigue is considered to be an exhaustion process of a material's inherent life (or ductility). A schematic fatigue life diagram, shown in Figure 14.10(a), elucidates the concept. At a constant stress of, say σ_1, the life of the material is 150 cycles, while at σ_2 it is 300 cycles. According to the cumulative-damage theory, in going from A to B or C to D, we gradually exhaust the material's fatigue life. That is, at points A and C, 100% of life at that level is available, while at points B and D, the respective lives are completely exhausted. If fatigue damage does, indeed, accumulate in a linear manner, each cycle contributes the same amount of damage at a given stress level. For example, on cycling the material from A to E, we exhaust one-third of the fatigue life available at σ_1. If we now change the stress level to σ_2, then the percentage of life already exhausted at σ_1 is equivalent to the percentage of life exhausted at σ_2. That is, one-third of fatigue life at σ_2 is equivalent to one-third of fatigue life at σ_1. Thus, in descending from E to F, we get from 50 to 100 cycles, and, as only one-third of fatigue life was exhausted at σ_1, two-thirds of fatigue life – that is, 200 cycles – is still available at σ_2. The same kind of change can be described for a low-to high-stress traversal.

Figure 14.10(b) shows the sequence of cyclic loading periods, each one with a specific mean stress and stress amplitude. The number of cycles in the four blocks are n_1, n_2, n_3, n_4, etc. The total life for each of these blocks is, respectively, N_1, N_2, N_3, N_4, etc. We just add up the fractions. The parallel can be drawn with our health. If a person were to smoke during their entire life, their life expectancy would be, say 50 years. A totally healthy person would live 100 years. If somebody alternates between ten years of smoking and ten years of no smoking, the life expectancy would be shortened to 75.

This linear damage model does not concern itself with the physical picture of the fatigue damage. It does, however, give an empirical way of predicting the fatigue life after a complex loading sequence. The method is generally known as the *Palmgren–Miner rule* or, simply, *linear cumulative-damage theory*.[3] The Palmgren–Miner rule says that the sum of all life fractions is unity; that is,

$$\boxed{\sum_{i=1}^{k} n_i/N_i = 1,} \quad \text{or}$$

$$n_1/N_1 + n_2/N_2 + n_3/N_3 + n_4N_4 + \ldots n_k/N_k = 1, \tag{14.1}$$

where k is the number of stress levels in the block spectrum loading; N_1, N_2, . . ., N_i are the fatigue lives corresponding to stress levels σ_1, σ_2, . . . σ_i, respectively; and, n_1, n_2, . . ., n_i are the number of cycles carried out at the respective stress levels. This rule is obeyed by a series of materials if the underlying assumptions are satisfied. The principal assumption is that the damage accumulation rate at any level does not depend on the prior loading history of the material; in other words, the damage per cycle is the same at the beginning or at the end of fatigue life, at a given stress level. This implies that the magnitude and direction of the change in amplitude (from low to high or high to low) do not have an effect on fatigue life. We also assume that in each block the loading is totally reversible (i.e., $\sigma_m = 0$). The validity of these assumptions is problematic. For example, it is quite likely that, for blocks identical in size and amplitude, a change in load from high to low would be much more dangerous than one from low to high: A crack initiated at high loads can continue to grow at low loads, whereas in the reverse case, at low loads, perhaps the crack would never have formed.

Example 14.1

The *S–N* curve of a material is described by the relationship

$$\log N = 10(1 - S/\sigma_{\max}),$$

where N is the number of cycles to failure, S is the amplitude of the applied cyclic stress, and $\sigma_{\max}$ is the monotonic fracture strength – i.e., $S = \sigma_{\max}$ at $N = 1$. A rotating component made of this material is subjected to 10^4 cycles at $S = 0.5\,\sigma_{\max}$. If the cyclic load is now increased to $S = 0.75\,\sigma_{\max}$, how many more cycles will the material withstand?

Solution:
For $S = 0.5\,\sigma_{\max}$,

$$\log N_1 = 10(1 - 0.5) = 5.$$

[3] A. Palmgren, *Z. Ver. Dtsch. Ing.*, 53 (1924) 339; M. A. Miner, *J. Appl. Mech.*, 12 (1945) 159.

Thus,

$$N_1 = 10^5 \text{ cycles.}$$

For $S = 0.75\ \sigma_{\max}$,

$$\log N_2 = 2.5.$$

So

$$N_2 = 316 \text{ cycles.}$$

Using Palmgren–Miner's rule, we have

$$n_1/N_1 + n_2/N_2 = 10^4/10^5 + n_2/316 = 1$$

or

$$n_2 = 284 \text{ cycles.}$$

Example 14.2

A microalloyed steel was subjected to two fatigue tests at ± 400 MPa and ± 250 MPa. Failure occurred after 2×10^4 and 1.2×10^6 cycles, respectively, at these two stress levels. Making appropriate assumptions, estimate the fatigue life at ± 300 MPa of a part made from this steel that has already undergone 2.5×10^4 cycles at ± 350 MPa.

Solution: We assume an equation of the form (other forms can also be assumed):

$$\begin{aligned}
\Delta\sigma(N_f)^a &= c,\\
800(2 \times 10^4)^a &= 500(1.2 \times 10^6)^a,\\
\frac{800}{500} &= 1.6 = \left(\frac{1.2 \times 10^6}{2 \times 10^4}\right)^a = (60)^a,\\
a &= 0.115,\\
c &= 800(2 \times 10^4)^{0.115}\ \text{MPa}\\
&= 2{,}498\ \text{MPa.}
\end{aligned}$$

At ± 350 MPa,

$$\begin{aligned}
N_{f1} &= \left(\frac{c}{\Delta\sigma}\right)^{1/a} = \left(\frac{2,498}{700}\right)^{1/0.115}\\
&= (3.57)^{8.7} = 6.4 \times 10^4 \text{ cycles.}
\end{aligned}$$

For 2.5×10^4 cycles,

$$\frac{N_1}{N_{f1}} = \frac{2.5 \times 10^4}{6.4 \times 10^4} = 0.39.$$

At ± 300 MPa,

$$\begin{aligned}
N_{f2} &= \left(\frac{c}{\Delta\sigma}\right)^{1/a} = \left(\frac{2,498}{600}\right)^{1/0.115}\\
&= (4.16)^{8.7} = 2.45 \times 10^5 \text{ cycles.}
\end{aligned}$$

From Palmgren–Miner's rule,

$$\frac{N_1}{N_{f1}} + \frac{N_2}{N_{f2}} = 1,$$

$$\frac{N_2}{N_{f2}} = 1 - \frac{N_1}{N_{f1}} = 1 - 0.39 = 0.61.$$

Therefore,

$$\begin{aligned} N_2 &= 0.61 \times N_{f2} \\ &= 0.61 \times 2.45 \times 10^5 \\ &= 1.49 \times 10^5 \text{ cycles.} \end{aligned}$$

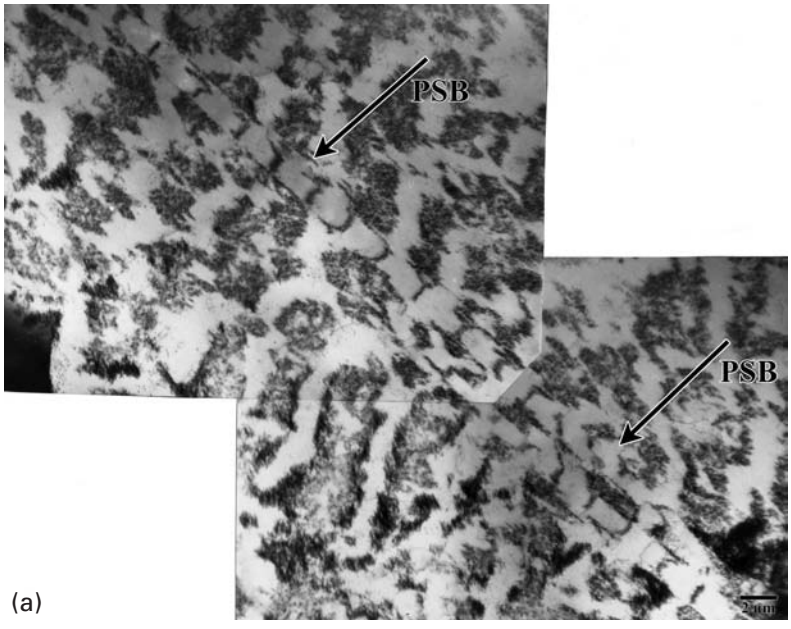
14.7 Mechanisms of Fatigue

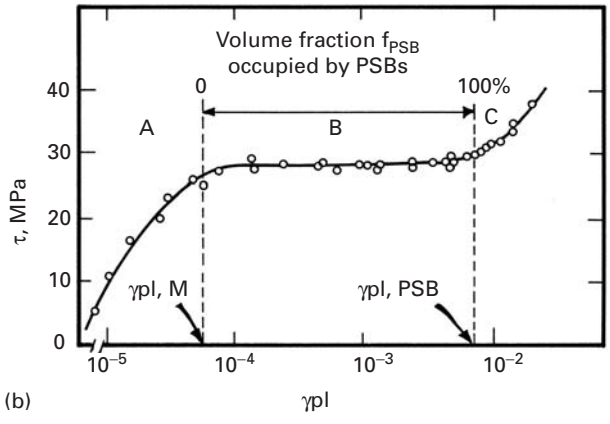
In this section, we describe the physical mechanisms responsible for fatigue mechanisms – mainly, fatigue crack nucleation and propagation. We assume that our starting material does *not* have any preexisting crack or cracklike defects. The fracture mechanics approach, focusing only on the propagation of preexisting cracks, will be examined in Section 14.8.

14.7.1 Fatigue Crack Nucleation

Fatigue cracks nucleate at singularities or discontinuities in most materials. Discontinuities may be on the surface or in the interior of the material. The singularities can be structural (such as inclusions or second-phase particles) or geometrical (such as scratches or steps). The explanation of preferential nucleation of fatigue cracks at surfaces perhaps resides in the fact that plastic deformation is easier there and that slip steps form on the surface. Slip steps alone can be responsible for initiating cracks, or they can interact with existing structural or geometric defects to produce cracks. Surface singularities may be present from the beginning or may develop during cyclic deformation, as, for example, the formation of intrusions and extrusions at what are called the persistent slip bands (PSBs) in metals. These bands were first observed in copper and nickel by Thompson *et al.*[4] They appeared after cyclic deformation and *persisted* even after electropolishing. On retesting, slip bands appeared again in the same places. Later, the dislocation structure in the PSBs was investigated extensively. Figure 14.11(a) shows a TEM micrograph of a polycrystalline copper sample that was cycled to a total strain amplitude of 6.4×10^{-4} for 3×10^5 cycles. Fatigue cycling was carried out in reverse bending at room temperature and at a frequency of 17 Hz. The thin foil was taken 73 μm below the surface. Two parallel PSBs (diagonally across the micrograph) embedded in a veined structure in polycrystalline

[4] N. Thompson, N. J. Wadsworth, and N. Louat, *Phil. Mag.*, 1 (1956) 113.





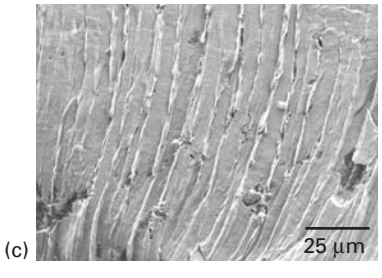


Fig. 14.11 (a) Persistent slip bands in vein structure. Polycrystalline copper fatigued at a total strain amplitude of 6.4×10^{-4} for 3×10^5 cycles. Fatiguing carried out in reverse bending at room temperature and at a frequency of 17 Hz. The thin foil was taken 73 μm below the surface. (Courtesy of J. R. Weertman and H. Shirai.) (b) Cyclic shear stress, τ, vs. plastic cyclic shear strain, γ_{pl}, curve for a single crystal of copper oriented for single slip. (After H. Mughrabi, *Mater. Sci. Eng.*, 33 (1978) 207.) The terms $\gamma_{pl,\,M}$ and $\gamma_{pl,\,PSB}$ refer to cyclic plastic shear strain in the matrix and persistent slip bands, respectively. (c) Intrusions/extrusions in a tin-based solder due to thermal fatigue. (Courtesy of N. Chawla and R. Sidhur)

copper can be seen. The PSBs are clearly distinguished and consist of a series of parallel "hedges" (a ladder). These ladders are channels through which the dislocations move and produce intrusions and extrusions at the surface Figure 14.11(c). Stacking-fault energy and the concomitant ease or difficulty of cross-slip play an important role in

the development of the dislocation structure in the PSBs. Kuhlmann-Wilsdorf and Laird have discussed models for the formation of PSBs in metals.[5] They compared the deformation substructures produced by unidirectional and cyclic (fatigue) deformation and interpreted them in terms of the differences between the two modes of deformation. The principal differences are as follows:

1. Due to the much larger time spans of deformation in fatigue, the dislocation structures formed are much closer to the configurations having minimum energy than the ones generated by monotonic straining. That is, more stable dislocation arrays are observed after fatigue.
2. The oft-repeated to-and-fro motion in fatigue minimizes the buildup of surpluses of local Burgers vectors, which are fairly prevalent after unidirectional (monotonic) strain.
3. Much higher local dislocation densities are found in fatigued specimens.

The characteristic dislocation arrangements observed in FCC metals form in the following manner. In monocrystals, we first have uniform fine slip, followed by the formation of veins consisting of dense bundles of dislocation dipoles and other debris. After this, PSBs are formed. They occur with the onset of saturation and are often associated with slight work-softening. There also seems to be a threshold strain for PSB formation, equal to 8×10^{-5} in the case of copper monocrystals. When subjected to strain-controlled cycling, an initially annealed metal hardens at first and then attains a saturation stress. If we plot this saturation stress against the applied plastic strain, we get another type of cyclic stress-strain curve, an example of which is shown for a single crystal of copper oriented for single slip in Figure 14.11(b). The curve has three stages, one of which is a plateau region, and each stage is characterized by a distinct dislocation structure. At low strains in the plateau region, the structure consisted of a hard matrix containing a loop–patch dislocation structure and a soft PSB with dislocations in a ladderlike arrangement. At large strains in the plateau region (plastic shear strain greater than 2×10^{-3}), most of the matrix phase and a part of the PSB had a well-developed mazelike structure. In the case of a polycrystal, the grains in the softest orientation and with not much constraint from their neighbors deform and harden by the accumulation of dislocations. An example of such a structure in a Cu–Ni polycrystal is shown in Figure 14.12. Dislocation walls form on {100} planes.

The interface between the PSB and the matrix represents a discontinuity in the density and distribution of dislocations. Hence, one would expect PSBs to be the preferential sites for fatigue crack nucleation. The surface of a metal subjected to cyclic stressing will have PSB extrusions and intrusions. Recall that monotonic loading of a

[5] D. Kuhlmann-Wilsdorf and C. Laird, *Mater. Sci. Eng.*, 27 (1977) 137; *Mater. Sci. Eng.*, 37 (1979) 111.

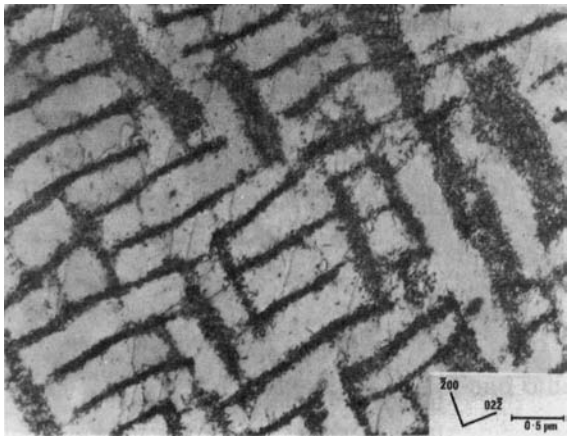


Fig. 14.12 Well-developed maze structure, showing dislocation walls on {100} in Cu–Ni alloy fatigued to saturation. (From P. Charsley, *Mater. Sci. Eng.*, 47 (1981) 181.)

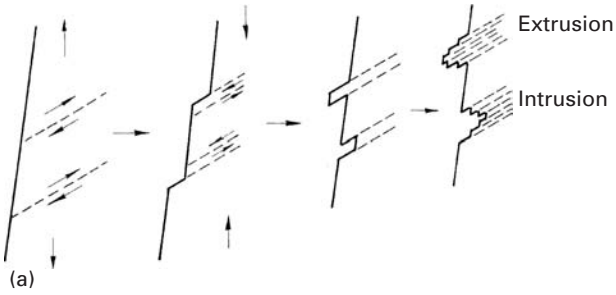


Fig. 14.13 (a) Fatigue crack nucleation at slip bands. (b) SEM of extrusions and intrusions in a copper sheet. (Courtesy of M. Judelwicz and B. Ilschner.)

metal results in the formation of slip steps at the surface. On being subjected to cyclic loading, however, the surface of the metal will have intrusions and extrusions where PSBs emerge. A model for this form of nucleation is shown in Figure 14.13(a). During the loading part of the cycle, slip occurs on a favorably oriented plane, and during the unloading part of the cycle, reverse slip occurs on a parallel plane, because the slip on the original plane is inhibited owing to

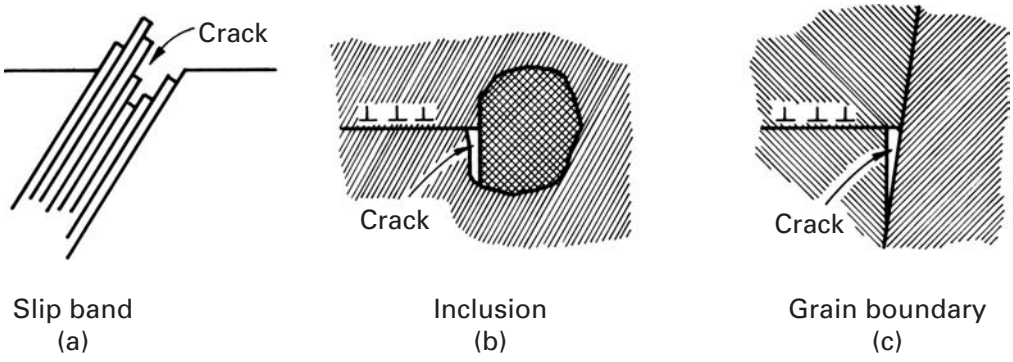


Fig. 14.14 Some mechanisms of fatigue crack nucleation. (After J. C. Grosskreutz, *Tech. Rep. AFML-TR-70–55* (Wright–Patterson AFB, OH: Air Force Materials Laboratory), 1970.)

hardening or, perhaps, the oxidation of the newly formed free surface. The first cyclic slip may create an extrusion or an intrusion at the surface. An intrusion may grow and form a crack by continued plastic deformation during subsequent cycles. An actual example of the formation of intrusions and extrusions in a sample of a copper sheet subjected to 15,000 cycles under a 60-MPa amplitude is shown in an SEM micrograph in Figure 14.13(b). Even during cyclic stressing in the tension–tension mode, this mechanism can function, as the plastic strain occurring at the peak load may lead to residual compressive stresses during the decreasing-load part of the cycle.

Such intrusions/extrusions owing to PSBs can form in ductile metals under the action of thermal fatigue as well. Thermal fatigue involves cyclic loading under thermal stresses. Such thermal stresses occur when a temperature change occurs, which leads to differential expansion in two components that are joined together. Figure 14.11(c) shows an interesting example of this in a tin-based solder in contact with copper; the two being subjected to temperature-induced thermal stress.

Twin boundaries can be important crack nucleation sites in hexagonal close-packed materials such as magnesium, titanium, etc., and their alloys. Inclusions and second-phase particles are commonly the dominant nucleation sites in materials of commercial purity – for example, aluminum, high-strength steels, and many polymers. Grain boundaries can become important nucleation sites at large strain amplitudes and at temperatures greater than about $0.5T_m$, where T_m is the melting point in kelvin, or in the presence of impurities that produce grain-boundary embrittlement (e.g., O_2 in iron). Some of these mechanisms are illustrated schematically in Figure 14.14.

Since most fatigue failures form at the surface of a material, the condition of the surface is very important. Indeed, polishing the surface can significantly increase the fatigue life of the material. A very important technological process to enhance fatigue life is *shot peening*, in which small metallic spheres are accelerated and hit the surface of the part. This bombardment by small particles puts the surface layer of a component in residual compression. This technique is used routinely in industry. Examples include compression coil springs in the automotive industry, wing skins for aircraft and other applications.

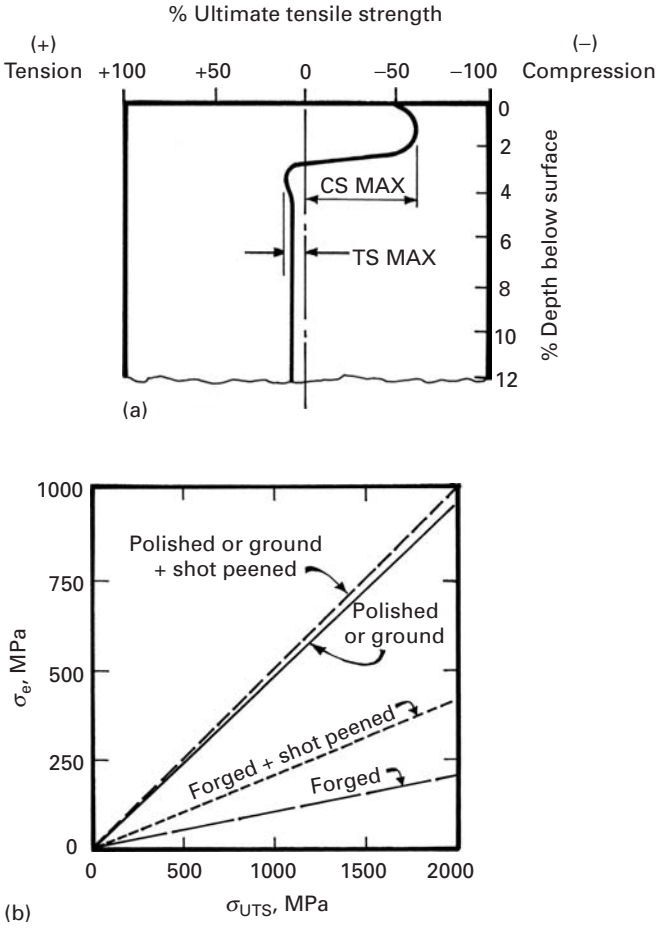


Fig. 14.15 (a) Residual stress profile generated by shot peening of a surface; CS and TS indicate compressive and tensile stress, respectively. (b) Effect of shot peening on fatigue life, σ_e of steels with different treatments as a fucntin of ultimate tensile strength, σ_{UTS}. (After J. Y. Mann, *Fatigue of Materials* (Melbourne, Melbourne University Press, 1967).)

Compressive stresses in surface layers have been introduced by laser. The technique called laser peening, involves a laser system based on neodymium-doped glass. Figure 14.15(a) shows, in a schematic fashion, a surface layer under residual compressive stress due to the cold-working from shot peening. The interior is under a small tensile stress as a result. The effect of shot peening on the endurance limits of steels with different ultimate tensile strengths is shown in Figure 14.15(b). For the as-forged components, the endurance limit is approximately 15% of the ultimate tensile strengths. Shot peening doubles the endurance limit, although it becomes less effective once the part is polished or ground, because the endurance limit is considerably increased by that process.

14.7.2 Fatigue Crack Propagation

At large stress amplitudes, a very large fraction (around 90%) of fatigue life is spent in the growth or propagation of a crack. For a component that contains a notch, this fraction becomes even larger. In as much as in most real structures cracklike imperfections are present, the crack propagation part can be a very important aspect of fatigue.

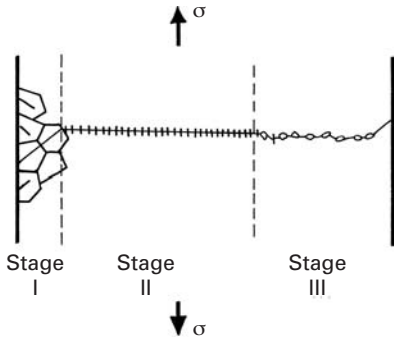
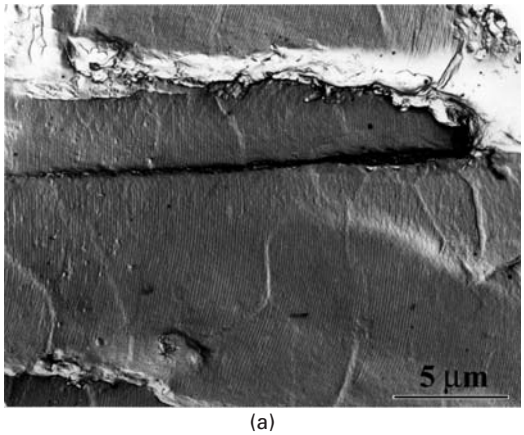
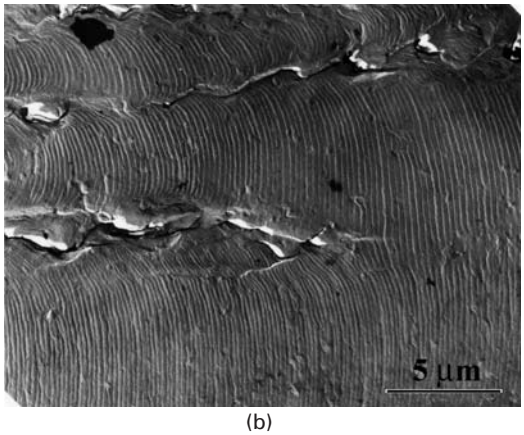


Fig. 14.16 Stages I, II, and III of fatigue crack propagation.

A brief description of crack propagation, in terms of microstructural processes, follows. A few cracks nucleate at the surface and start propagating in a crystallographic shear mode (stage I) on planes oriented at approximately 45° to the stress axis. (See Figure 14.16.) During this stage, cracking occurs along the crystallographic slip planes, and the crack growth is on the order of a few micrometers or less per cycle. Little is known about crack propagation in this stage. Many consider the stage to be an extension of the nucleation process. Once a crack is initiated, say, at a slip band on the surface, it continues along the slip band until it encounters a grain boundary. These crystallographic cracks penetrate a few tenths of a millimeter in this mode. From there on, a dominant crack starts propagating in a direction perpendicular to the stress axis in the tensile mode. This is called stage II, and typically, the fracture surface shows striation markings. The ratio of the extent of stage I to stage II decreases with an increase in stress amplitude. The stress concentration at the tip of the crack causes local plastic deformation in a zone in front of the crack. With crack growth, the plastic zone increases in size until it becomes comparable to the thickness of the specimen. When this occurs, the plane-strain condition at the crack front in stage II does not exist any more, the crack plane undergoes a rotation, and the final part of rupture occurs in plane-stress or shear mode. This corresponds to Stage III, characterized by rapid crack propagation. Microscopic observations of fatigue fracture surfaces frequently show striations in stage II. Propagation occurs in a direction perpendicular to the tensile stress, and in a large number of metals and alloys (principally of Al and Cu), at high amplitudes, the fracture surface shows the characteristic striations. Such striations have been observed in polymers as well. Frequently, each striation is thought to represent one load cycle, and indeed, it has been observed by means of programmed amplitude fatigue tests that in many materials these striations do represent the crack front position in each cycle. An example of the variation in fatigue striations in a 2014-T6 aluminum alloy is presented in Figure 14.17. Figure 14.17(a) shows the striations in the early stage of fatigue life, while Figure 14.17(b) shows the striations in the late stage of cycling. Note the smaller striation spacing in the early stage, indicating a lower crack propagation rate than in the late stage. These micrographs



(a)



(b)

Fig. 14.17 Fatigue striations in 2014-T6 aluminum alloy; two-stage carbon replica viewed in TEM. (a) Early stage. (b) Late stage. (Courtesy of J. Lankford.)

were taken in a TEM from a two-stage carbon replica of the surface of the aluminum alloy. The reason for this is that only TEM could provide the high-enough magnification for viewing the closely spaced striations in the early stage of fatigue. However, the reader is warned that such a correlation is not always available. If it were, one should be able to relate striation spacing to ΔK (see Section 14.8) and obtain a one-to-one correspondence between the macroscopic growth rate and ΔK. One cannot always do this, however, indicating that the crack front may have advanced by a combination of the formation of striations and other fracture mechanisms.

Care should be exercised in the interpretation of fatigue striations. It has been observed in an Fe–Si alloy that, whereas the fatigue striations were 2 μm apart, the actual advance of the crack front per cycle was only 10^{-9} m, or 2,000 times smaller![6] These results show that, under certain conditions, the crack front remains "dormant" for many cycles, while damage accumulates in the material. At a certain point, the crack advances discontinuously. This phenomenon is very common in polymers at low values of ΔK. A craze forms gradually

[6] W. Yu, K. Esablul, and W. W. Gerberich, *Met. Trans.*, 15A (1984) 889.

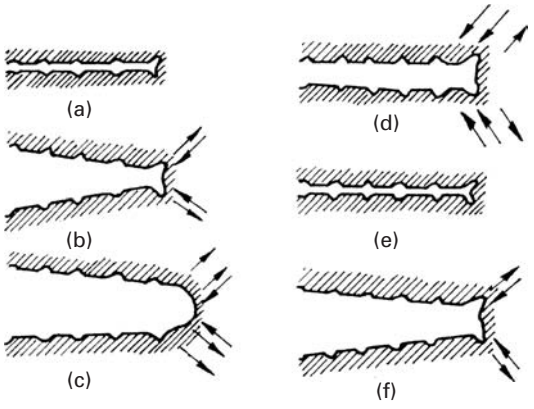


Fig. 14.18 Fatigue crack growth by a plastic blunting mechanism. (a) Zero load. (b) Small tensile load. (c) Maximum tensile load. (d) Small compressive load. (e) Maximum compressive load. (f) Small tensile load. The loading axis is vertical (After C. Laird, in *Fatigue Crack Propagation*, ASTM STP 415 (Philadelphia: ASTM, 1967), p. 131.)

at the tip of the crack during fatigue. When the craze reaches a critical length, the crack advances through it. The process repeats itself periodically.

At higher values of ΔK, striations become less important in the overall crack propagation rate. One model of fatigue crack growth by a striation mechanism is shown in Figure 14.18. This model involves repetitive blunting and sharpening of the crack front. Figure 14.18(a) shows the situation at zero load. During the tensile part of the load cycle, plastic strains at the crack tip cause localized slip on planes of maximum shear. (See Figure 14.18(b).) The situation at maximum tensile load is shown in Figure 14.18(c). The start of the compressive cycle is shown in Figure 14.18(d). The reversal of the loading direction, during compression, causes the crack faces to join (Figure 14.18(e)). However, the new surface created during the tensile part of the cycle is not completely "rehealed," due to slip in the reverse direction. Depending on the material and the environment, a large part of slip during compression occurs on new slip planes, and the crack tip assumes a bent form with "ears," as shown in Figure 14.18(e). At the end of the compression half of the cycle, the crack tip is resharpened, and the propagation sequence of the next cycle is restarted. (See Figure 14.18(f).) This model of plastic blunting and resharpening seems to be valid for any ductile material, including polymers. There is evidence that the crack propagates in a similar manner in stage I, but with only a group of slip planes at 45° operating. However, one must bear in mind that, although the presence of striations confirms a fatigue failure mechanism, an absence of striations does not necessarily preclude fatigue. In fact, a variety of other fracture modes are possible in fatigue. In single-phase materials, transgranular or intergranular fracture modes are possible, while any second phases that are present may lead to dual fracture modes. Figure 14.19 shows the various possible microscopic fracture modes in fatigue.

Fracture surfaces in polymers produced under fatigue conditions also show some characteristic features. Generally, two distinct regions are present: a region of smooth, slow crack growth surrounding the fracture initiation site and a rough region corresponding to rapid

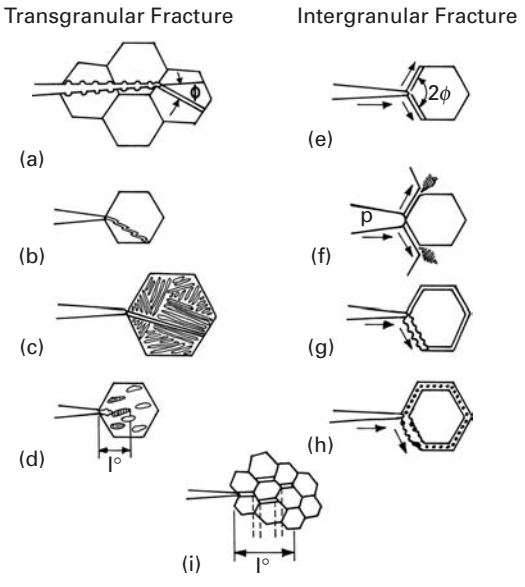


Fig. 14.19 Microscopic fracture modes in fatigue. (a) Ductile striations triggering cleavage. (b) Cyclic cleavage. (c) $\alpha - \beta$ interface fracture. (d) Cleavage in an $\alpha - \beta$ phase field. (e) Forked intergranular cracks in a hard matrix. (f) Forked intergranular cracks in a soft matrix. (g) Ductile intergranular striations. (h) Particle-nucleated ductile intergranular voids. (i) Discontinuous intergranular facets. (Adapted from W. W. Gerberich and N. R. Moody, in *Fatigue Mechanisms*, ASTM STP 675 (Philadelphia: ASTM, 1979) p. 292.)

crack growth. Sometimes, semicircular concentric bands are seen in the smooth region near the starting flaw. These bands are indicative of discontinuous crack growth, and the band width represents the extent of plastic zone or craze that developed ahead of the crack tip. Fracture surfaces produced by cyclic loading in polymers frequently show striations. There is some confusion on the use of the term *striation* in the literature on fatigue in polymers and metals. In metals, the term is used to denote markings on the fracture surface, without regard to any correlation between the striation spacing and the crack growth per cycle. In polymers, the term *striation* is used only when there is a one-to-one correlation between the striation spacing and the crack growth in each cycle. This stems from the fact that there are other types of discontinuous growth which result in fracture surface markings, but without any advance in the fatigue crack. In particular, in polymers there occur discontinuous growth bands (DGBs), which correspond to a burst of fatigue crack growth after some hundreds of fatigue cycles; that is, the crack tip remains stationary for some cycles and then undergoes an advance. DGBs resemble striations, but their spacing is much larger than the crack growth in a cycle. The formation of a DGB is thought to be due to the accumulation of damage ahead of the fatigue crack over many cycles, followed by a sudden jump by the crack. One model explains that a craze forms at the fatigue crack and that the crack initially grows along the craze–matrix interface, as shown in Figure 14.20, and then along the craze filled midrib until the crack is arrested. A repetition of this process results in the appearance of dark bands on the fracture surface. Thus, the DGBs represent the successive positions of crack tips that have been blunted, advanced, and arrested repeatedly.

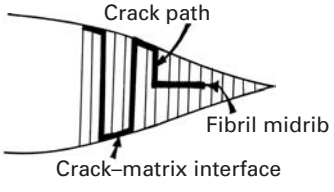


Fig. 14.20 Discontinuous crack growth through a craze at the tip of a fatigue crack. (After L. Konczol, M. G. Schincker and W. Döll, *J. Mater. Sci.*, 19 (1984) 1604.)

14.8 Linear Elastic Fracture Mechanics Applied to Fatigue

The use of large monolithic structures has resulted in widespread application of fracture mechanics. In particular, the phenomenon of fatigue crack propagation can be analyzed in terms of linear elastic fracture mechanics. The basic assumption here is that cracks already exist in a structural component and that they will grow as the component gets used in service. In terms of fatigue crack growth studies, it is also implied that the fatigue life of a component is determined mostly by the crack growth under cyclic loading.

We can determine K_{Ic} or K_c for a given material in the laboratory and can use the data obtained to determine a failure locus in terms of a critical applied stress and a corresponding critical crack length, or vice versa. (See Figure 14.21.) For example, in Figure 14.21(a), we can observe that, for a given crack length a_1, there is a critical failure stress σ_1 of the material. Conversely, for a given design stress σ_2, there is a critical crack length a_2. In principle, then, the region under the failure locus represents the safe region with respect to a catastrophic failure. Consider, for instance, a component containing a crack length of a_1, at a stress of σ_2, where $\sigma_2 < \sigma_1$. Under these

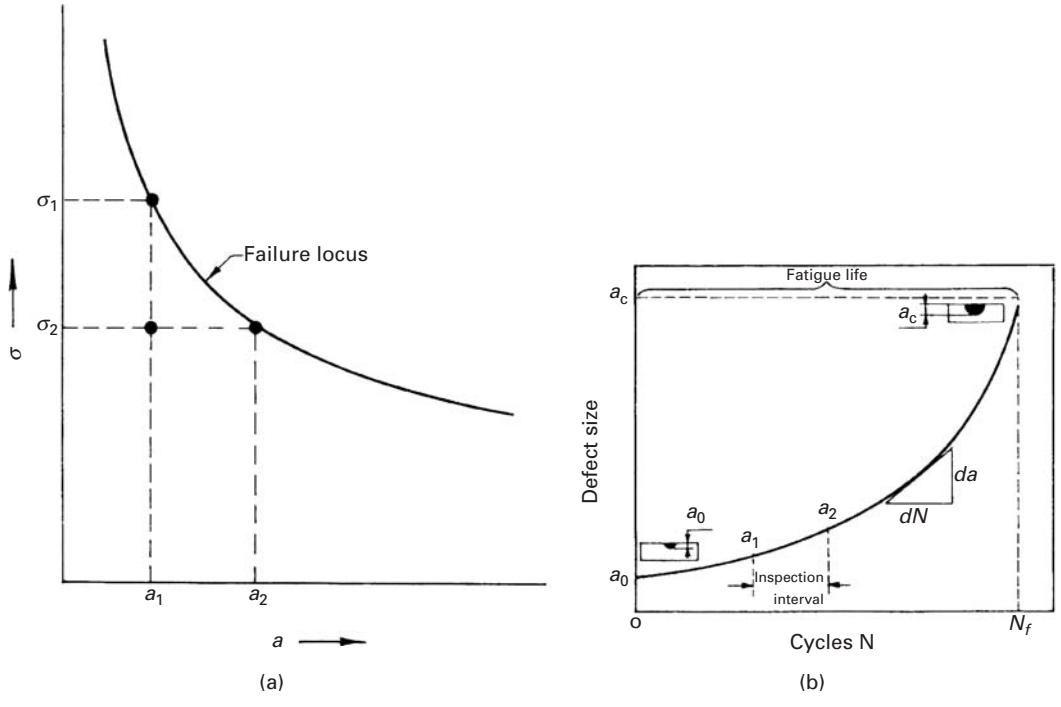


Fig. 14.21 (a) Failure locus. (b) Schematic of crack length a as a function of number of cycles, N.

conditions, the component will be safe because a_1 is smaller than the critical defect size a_2, which corresponds to the applied stress σ_2. This security is based, of course, on the assumption that loading is static and that the crack does not grow in service. But we know very well that cracks in structures do grow during service. An increase in crack length at σ_2, in service, from a_1 to a_2 will eventually lead to structural failure. Thus, although the fracture toughness of a material establishes the failure condition and the residual strength of a structural component, the component's service life or durability is mainly a function of its resistance to subcritical crack growth (i.e., its resistance to crack growth by fatigue, creep, stress corrosion, etc.).

As we pointed out in Chapter 7, linear elastic fracture mechanics accepts the preexistence of cracks in a structural member. The model for the crack tip is the same as that described for nonfatigue regimens. The material containing a crack, under tension, has a small plastic zone at the crack tip, and this plastic zone is surrounded by a rather large elastic region. This being so, we focus our attention on the propagation of cracks under conditions of fatigue. Once again – and it is worth repeating – we do not concern ourselves here with the crack nucleation problem under fatigue. Under cyclic loading, a dominant crack grows, as a function of the number of cycles, from an initial size a_0 to a critical size a_c, corresponding to failure, as shown in Figure 14.21(b). The basic problem is thus reduced to one of characterizing the growth kinetics of the dominant crack in terms of an appropriate driving force. From there, one can estimate the service life and/or schedule inspection intervals under designed loading conditions and service environments. Since crack growth starts from the most highly stressed region at the crack tip, we characterize the driving force in terms of the stress intensity factors at the tip – that is, the range of the stress intensity factor $\Delta K = K_{\max} - K_{\min}$, where $K_{\max}$ and $K_{\min}$ are the maximum and minimum stress intensity factors corresponding to the maximum and minimum loads, respectively. The crack growth rate per cycle, da/dN, can then be expressed as a function of the cyclic stress intensity factor at the crack tip, ΔK. Hence, if a mathematical equation describing the crack growth process and the appropriate boundary conditions is available, we can, in principle, compute the fatigue life (i.e., number of cycles to failure). Paris *et al.* proposed the following *empirical* relationship (known as the Paris–Erdogan relationship) for crack growth under cyclic conditions:[7]

$$\boxed{da/dN = c(\Delta K)^m} \tag{14.2}$$

Here, a is the crack length, N is the number of cycles, ΔK is the cyclic stress intensity factor as defined earlier, and C and m are empirical constants that depend on the material, environment, and test conditions, such as the load ratio R, the test temperature, the waveform,

[7] P. C. Paris, M. P. Gomez, and W. P. Anderson, *The Trend in Engineering*, 13 (1961) 9; P. C. Paris and F. Erdogan, *J. Basic Eng., Trans. ASME*, 85 (1963) 528.

etc. Another empirical relation connecting the parameters C and m is

$$C = A/(\Delta K_0)^m$$

where A and K_0 are some other material constants.

Since many variables affect the crack growth rate in fatigue, we can write, in a very general way,

$$da/dN = F(\Delta K, K_{\max}, R, \text{frequency, temperature, } \ldots). \tag{14.3}$$

Clearly, one cannot obtain such an ideal and detailed characterization. In practice, one collects data under restricted conditions, but consistent with the applications in service. In principle, the rate equation 14.3 can be integrated to determine the service life N_f, or an appropriate inspection interval ΔN, for a structural component. We have

$$N_f = \int_{a_0}^{a_f} \frac{da}{F(\Delta K, \ldots)},$$

or

$$\Delta N = N_2 - N_1 = \int_{a_1}^{a_2} \frac{da}{F(\Delta K, \cdots)}. \tag{14.4}$$

Rewriting, we get

$$N_f = \int_{K_{f\max}}^{K_{f\max}} \frac{dK}{(dK/da)\,F(\Delta K, \cdots)}, \tag{14.5}$$

or

$$\Delta N = N_2 - N_1 = \int_{K_1}^{K_2} \frac{dK}{(dK/da)\,F(\Delta K, \cdots)}. \tag{14.6}$$

If we plot the logarithm of the crack growth rate da/dN against the logarithm of the alternating stress intensity factor $\Delta K = K_{\max} - K_{\min}$ at the crack tip, we get the kind of curve shown in Figure 14.22. The curve has a sigmoidal form with three regions. Region II is the one that shows the Paris–Erdogan type of power-law relation between da/dN and ΔK. The power-law region connects the upper and lower limiting regions. The lower limit on the cyclic stress intensity factor in region I denotes a threshold value below which the crack does not propagate. This limit is called the *threshold cyclic stress intensity factor* ΔK_{th}. The upper limit in region III indicates the conditions of accelerated crack growth rate associated with the start of final rupture.

Many researchers (see, for example, Suresh (1991) in the suggested reading) have discussed the primary operating mechanisms and the important variables in the three stages of fatigue crack propagation:

Stage I: In this stage, the average crack growth per cycle is less than a lattice spacing. Crack propagation mechanisms are characteristic of a discontinuous medium. The microstructure of

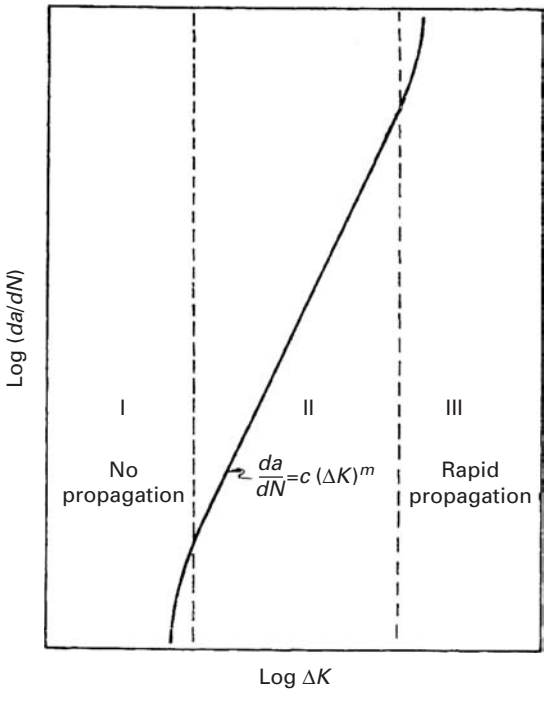


Fig. 14.22 Schematic of crack propagation rate da/dN versus alternating stress intensity factor ΔK.

the material, the stress ratio R, and the environment have a large influence on the crack growth.

Stage II: This is the power-law regimen, where the Paris–Erdogan relationship applies. Crack propagation mechanisms in stage II are characteristic of a continuous medium. The influence of the microstructure, R, the environment, the thickness of the material, etc. on crack growth is small.

Stage III: Crack propagation mechanisms in this stage are similar to those in the static mode (cleavage, intergranular, microvoid coalescence, etc.) In stage III, the microstructure, R, and the thickness of the material have a large influence on crack growth, but the influence of the environment is small.

The Paris–Erdogan power relationship (Equation 14.2) describes the crack propagation rate in stage II for a variety of materials–polymers, metals, and ceramics. It is very useful because of its extreme simplicity. For example, it has been observed experimentally that data points in the form of log (da/dN) versus log ΔK for a given material (with a constant microstructure) from three different samples – an edge crack in a compact-tension sample, a through-the-thickness central crack in a plate, and a plate containing a partial through-the-thickness crack – all fall on the same line. Also, there is experimental evidence that the stress level by itself does not influence the fatigue crack growth rate for levels below the general yielding stress. Thus, we can assume that the parameter ΔK describes uniquely the crack growth rates for many engineering applications. However, rather gross

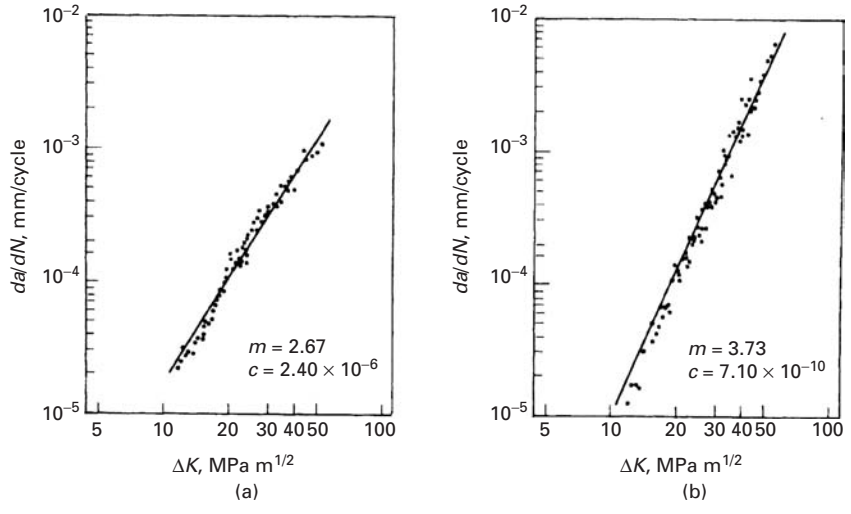


Fig. 14.23 Fatigue crack propagation in an AISI 4140 steel. (a) Longitudinal direction (parallel to rolling direction). (b) Transverse direction (perpendicular to rolling direction). (Reprinted with permission from E. G. T. De Simone, K. K. Chawla, and J. C. Miguez Suárez, *Proc. 4th CBECIMAT* (Florianópolis, Brazil, 1980), p. 345)

microstructural features of a material, such as the directionality imparted by aligned inclusions, can influence fatigue crack growth rates drastically, changing the value of m significantly. Figure 14.23 illustrates the directionality in the fatigue crack propagation rate in an AISI 4140 steel. The exponent m has a higher value in the transverse direction than in the longitudinal (rolling) direction, due to the presence of elongated inclusions.

Example 14.3

Consider long crack propagation under fatigue. Develop an expression for the number of cycles, ΔN, required for a crack to grow from an initial length a_i to a final length a_f, Given that $\Delta K = Y\Delta\sigma\sqrt{\pi a}$ and $da/dN = C\Delta K^m$, where the symbols have their usual significance. Discuss the implications of the expression.

Solution:

$$
\begin{aligned}
da/dN &= C\,\Delta K^{\,m},\\
\Delta N &= \int_{a_i}^{a_f} \frac{da}{C\,\Delta K^{\,m}}\\
&= \int_{a_i}^{a_f} \frac{da}{C\,(Y\,\Delta\sigma\sqrt{\pi})^m a^{m/2}}\\
&= \frac{a_f^{(1-m/2)} - a_i^{(1-m/2)}}{C\,(Y\,\Delta\sigma\sqrt{\pi})^m(1-m/2)}\\
&= \frac{1-(a_i/a_f)^{m/2-1}}{C\,(Y\,\Delta\sigma\sqrt{\pi})^m(m/2-1)}\left(\frac{1}{a_f^{m/2}-1}\right).
\end{aligned}
$$

The implications of the expression are that it is not valid for $m = 2$ and that ΔN is more sensitive to the initial crack length a_i than the final crack length a_f.

Example 14.4

The fatigue crack markings shown in Figure E14.4 were found in a fractured part. Determine the time to rupture of this part if the loading frequency is 10 Hz, the maximum stress applied to the part is 300 MPa, and the minimum stress is zero. The initiation stage of the flaw is 50% of the life of the part.

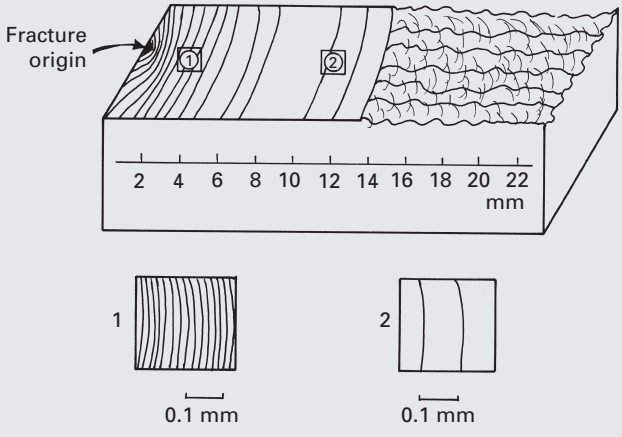


Fig. E14.4

Solution: We have

$$\left(\frac{da}{dN}\right)_1 = 0.016 \text{ mm}, \left(\frac{da}{dN}\right)_2 = 0.1 \text{ mm},$$
$$a_1 = 2 \text{ mm}, a_2 = 10 \text{ mm}.$$

Fracture occurs when $a_f = 14$ mm. Assuming that we have plane strain,

$$\begin{aligned} K_{Ic} &= 1.12\sigma\sqrt{\pi a} \\ &= 1.12 \times 300\sqrt{\pi \times 0.014} \\ &= 70 \text{ MPa m}^{1/2}. \end{aligned}$$

We now find the parameters for the Paris equation:

$$\begin{aligned} \frac{da}{dN} &= C(\Delta K)^m, \\ \Delta K_1 &= 1.12\sigma\sqrt{\pi a_1} = 1.12 \times 300\sqrt{\pi \times 2 \times 10^{-3}} \\ &= 26.6 \text{ MPa m}^{1/2}, \\ \Delta K_2 &= 1.12\sigma\sqrt{\pi a_2} = 1.12 \times 300\sqrt{\pi \times 10 \times 10^{-3}} \\ &= 59.55 \text{ MPa m}^{1/2}, \\ 0.016 \times 10^{-3} &= C(26.6)^m, \\ 0.1 \times 10^{-3} &= C(59.55)^m, \\ m &= 2.27, \\ C &= 0.94 \times 10^{-8}, \end{aligned}$$

$$\frac{da}{dN} = 0.94 \times 10^{-8}(\Delta K)^m = 0.94 \times 10^{-8} \times (1.12 \times 300\sqrt{\pi a})^{2.27},$$

$$\frac{da}{a^{1.135}} = 0.019\, dN.$$

Integrating between the limits, $a_f = 14$ mm and $a_i = 0.02$ mm, we get:

$$\int_{a_i}^{a_f} \frac{da}{a^{1.135}} = \int_{0.02}^{14} \frac{da}{a^{1.135}} = 0.019 \int dN = 0.019N,$$

$$a_0 = 0.02\,\text{mm}$$

$$a_f = 14\,\text{mm}$$

$$0.019N = \frac{-1}{0.135 \times (14 \times 10^{-3})^{0.135}} + \frac{1}{0.135 \times (0.02 \times 10^{-3})^{0.135}}$$

Each cycle corresponds to 0.1 s, so

$$t = 0.1N$$

The total time is equal to the initiation time plus the propagation time:

$$t = 2t = 98.6 \times 2\,\text{s} = 197.2\,\text{s}.$$

Example 14.5

An aluminum alloy has a plane-strain fracture toughness K_{Ic} of 50 MPa $m^{1/2}$. A service engineer has detected a 1-mm-long crack in an automotive component made of this alloy. The component will be subjected to cyclic fatigue with $\Delta\sigma = 100$ MPa with $R = 0$. How many more cycles can this component endure? Take $K = 1.05\,\sigma\sqrt{\pi a}$ and da/dN (mm/cycle) $= 1.5 \times 10^{-24}\ \Delta K^4$ (MPa $m^{1/2})^4$.

Solution: The final crack length a_f can be obtained from K_{Ic} with 100 MPa of applied stress:

$$K_{Ic} = 50 = 1.05 \times 100 \times \sqrt{\pi a_f},$$

$$a_f = \left[\frac{50}{1.05 \times 100\sqrt{\pi}}\right]^2 \text{m} = 0.072\ \text{m} = 72\ \text{mn}.$$

We thus have (being very careful with units!)

$$N_f = \frac{1}{(1.5 \times 10^{-24})(1.05^4)\pi^2(100)^4}\left[\frac{1}{a_i} - \frac{1}{a_f}\right] = \frac{10^{16}}{(1.5)(1.22)(9.87)}\left[1 - \frac{1}{72}\right]10^3 = 10^{16}[0.055][0.986] = 5.5 \times 10^{17}\text{cycles}.$$

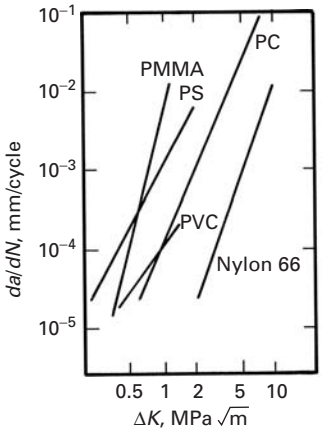


Fig. 14.24 Fatigue crack propagation rates for a number of polymers. (After R. W. Hertzberg, J. A. Manson, and M. Skibo, *Polymer Eng. Sci.*, 15 (1975) 252.)

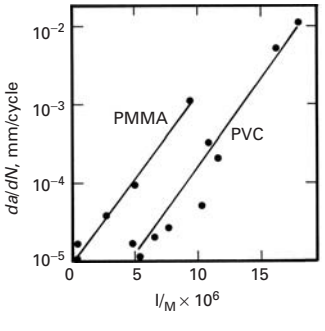


Fig. 14.25 Variation in fatigue crack propagation rates, at fixed values of ΔK (= 0.6 MPa $m^{1/2}$) and test frequency ν (= 10 Hz), as a function of reciprocal of molecular weight for PMMA and PVC. (After S. L. Kim, M. Skibo, J. A. Manson, and R. W. Hertzberg, *Polymer Eng. Sci.*, 17 (1977) 194.)

Example 14.6

When subjected to fatigue under a $\Delta\sigma = 140$ MPa, an alloy showed the following Paris-type fatigue crack propagation relationship:

$$\frac{da}{dN}(\text{m/cycle}) = 0.66 \times 10^{-8}(\Delta K)^{2.25}$$

where ΔK is in MPa $m^{1/2}$. Estimate the number of cycles required for the crack to grow from 2 mm to 8.8 mm.

Solution:

$$\frac{da}{dN} = 0.66 \times 10^{-8}(1.12\Delta\sigma\sqrt{\pi a})^{2.25},$$
$$a^{-1.125}da = 0.66 \times 10^{-8}(1.12)^{2.25} \times (140)^{2.25}(\pi)^{1.125}dN.$$

Integrating, we get

$$\int_{a_0}^{a_c} a^{-1.125}da = 2.0815 \times 10^{-3}\int_0^{N_f} dN,$$
$$N_f \times 2.0815 \times 10^{-3} = -\left[\frac{a^{-0.125}}{0.125}\right]_{0.002}^{0.0088},$$
$$N_f = \frac{2.941}{2.0815 \times 10^{-3}} = 1.4129 \times 10^3 \text{ cycles},$$

or

$$N_f = 1{,}413 \text{ cycles}.$$

A Paris–Erdogan type of relationship can be used to describe the fatigue crack propagation rate da/dN in polymers also. Figure 14.24 shows the fatigue crack rates as a function of ΔK for a number of thermoplastic polymers. Note that semicrystalline nylon 66 is superior in resistance to fatigue crack growth than amorphous, but ductile, polymers such as polyvinyl chloride (PVC) and polycarbonate (PC), which in turn are superior to brittle, amorphous polymers such as PS and PMMA. The latter both show deformation by crazing. Metals such as aluminum alloy and steel (not shown in the figure) would have curves to the right of that of nylon. That is, polymers show a lower resistance to fatigue crack propagation than metals do: Unlike metals, the range of the exponent m for polymeric materials can be quite large, from 4 to 20.

The molecular weight of a polymer is a very important variable for a number of properties, including fatigue crack propagation. In general, as the molecular weight increases, the fatigue strength increases and the fatigue crack propagation rate decreases. Figure 14.25 shows the variation in the fatigue crack propagation rate in PMMA and PVC, at a constant value of ΔK (0.6 MPa $m^{1/2}$) and at 10 Hz, as a function of $1/M$, where M is the molecular weight.

Earlier, it was thought that cyclic fatigue in ceramics did not occur, at least not in an inert atmosphere. This was based on the fact

that, in a ceramic, no dislocation-based plastic deformation occurred at the crack tip. Although dislocation plasticity is generally absent in ceramics, many ceramics show subcritical crack growth, under cyclic loading, at room temperature and at elevated temperatures. The deformation mechanisms under cyclic loading are generally different from those under static loading. While dislocation-based cyclic slip is responsible for fatigue in metals, phenomena such as microcracking, phase transformations, interfacial sliding, and creep can promote an inelastic constitutive response in brittle solids, leading to cycle fatigue. Work done by Suresh and coworkers, as well as others, on fatigue crack growth in a variety of brittle solids in compression, tension, and tension–compression fatigue shows that mechanical fatigue effects – that is, stable crack propagation – under cyclic fatigue conditions – can occur in ceramics at room temperature and in brittle solids as well.[8] Ewart and Suresh were the first ones to show such a cyclic fatigue effect in ceramics by subjecting them to cyclic compression. Researchers have used a variety of loading techniques to obtain fatigue crack growth data in ceramics under cyclic loading, such as four-point flexure, compact tension, and wedge-opening load specimens. Figure 14.26 shows the fatigue crack growth in an alumina sample (grain size $\approx$10 μm) subjected to tension–compression fatigue ($R = -1$) at a frequency of 5 Hz, in terms of da/dN vs. K_{max}, the maximum stress intensity factor. If we take $\Delta K = K_{max}$, it is easy to see that the data correspond to a Paris–Erdogan type of power law, $da/dN = C\Delta K^m$. The figure also shows the data obtained from static loading in terms of crack growth per cycle. We use the relationship $da/dN = (da/dt)/v$, where v is the frequency, and plot this against the maximum stress intensity factor K_{max} at which the static test was performed. The idea of putting the two curves together is to show that the crack growth rate in alumina under cyclic loading is much faster than that under static loading. *S–N* curves obtained in tension–tension cycling of fine-diameter alumina fiber also showed a distinct cyclic loading effect.[9] Work on fatigue growth rate in zirconia partially stabilized with magnesia also confirmed the cyclic fatigue effect.[10] One should note, however, that the values of m for ceramics in the Paris–Erdogan type of power law relationship range from 8 to 42, much higher than the 2–4 range for metals.

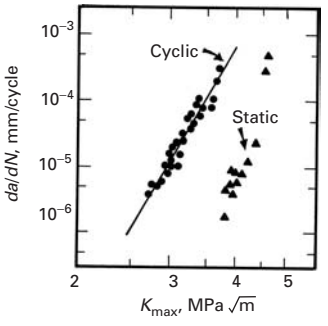


Fig. 14.26 Fatigue crack growth rate da/dN in alumina as a function of the maximum stress intensity factor K_{max} under fully reversed cyclic loads ($v = 5$ Hz). Also indicated are the rates of crack growth per cycle derived from static-load fracture data. (After M. J. Reece, F. Guiu, and M. F. R. Sammur, *J. Amer. Ceram. Soc.*, 72 (1989) 348.)

In many materials, such as intermetallics, composites, and ceramics, there are a number of additional mechanisms operating during cyclic fatigue. These mechanisms are responsible for R curve behavior. These mechanisms are classified as extrinsic and intrinsic. Figure 14.27 shows, in a schematic fashion, a few intrinsic and

[8] L. Ewart and S. Suresh, *J. Mater. Sci.* 22 (1987) 1173; S. Suresh and J. R. Brockenbrough, *Acta Met.*, 36 (1988) 1455; M. J. Reece, F. Guiu, and M. F. R. Sammur, *J. Amer. Ceram. Soc.*, 72 (1989) 348; R. H. Dauskardt, D. B. Marshall, and R. O. Ritchie, *J. Amer. Ceram. Soc.* 73 (1990) 893.

[9] N. Chawla, M. Kerr, and K. K. Chawla, *J. Amer. Ceram. Soc.*, 88 (2005) 101.

[10] R. H. Dauskardt, D. B. Marshall, and R. O. Ritchie, *J. Amer. Ceram. Soc.,* 73 (1990) 893.

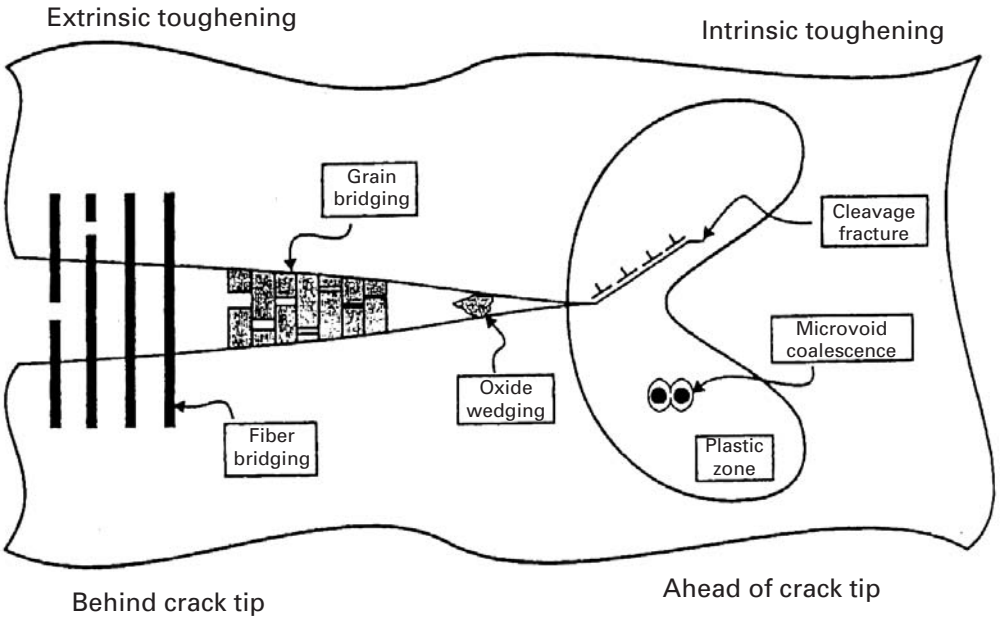


Fig. 14.27 Intrinsic and extrinsic mechanisms of fatigue damage. (After R. O. Ritchie, *Intl. J. Fracture*, 100 (1999) 55.)

extrinsic mechanisms. Intrinsic mechanisms operate ahead of the crack tip. Examples are void formation and coalescence, microcracks, plastic deformation, and phase transformations. Extrinsic mechanisms operate behind the crack tip. Extrinsic toughening mechanisms include fiber bridging (in the case of composites), grain bridging, phase transformations, microcrack toughening, and oxide wedging. Oxides formed on the cracked surface can dislodge themselves and act as microwedges along the crack, impeding it from closing complertely, on unloading. Thus, the stress amplitiude is effectively decreased.

14.8.1 Fatigue of Biomaterials

A common type of fatigue damage in the body is called "stress fracture." These are actually microcracks that form in bones from repeated stressing at excessive loads. Such is the case of long-distance runners whom experience stress fractures in the shins and feet. The body has a marvelous capacity of self-healing and rest usually alleviates the problem and allows the small cracks to close.

In the domain of biomaterials, most implants in the body undergo cyclic loading, which often results in failure. Hip and knee replacement joints, titanium posts in teeth implants, and mechanical heart valve prostheses are prone to fatigue. This problem is aggravated by the physiological environment, and environmental degradation can accelerate fatigue crack growth. For instance, the yield stress of the cobalt (62%)–chrome (28%) alloy used in biological applications is approximately 650 MPa. However, the endurance limit (for 10^7–10^8 cycles) is in the 150–250 MPa range.

Example 14.7

Estimate the number of cycles that a total hip replacement joint will experience in ten years.

Solution: A sedentary person walks approximately one hour per day. This corresponds to 5 km/day. Assuming that each step is equal to 780 mm (the student can verify this!) we have approximately 8,000 cycles/day, or 2.9×10^3 cycles/year. Over a period of ten years, we would have approximately 3×10^4 cycles.

Heart valve prostheses have been manufactured from titanium, cobalt–chromium alloys, and pyrolitic carbon. The problems encountered with metallic heart valves has led to the almost universal adoption of pyrolitic carbon or a pyrolitic carbon/graphite laminate. By 1996, over 600,000 of these valves had been installed. There are two types of commercial pyrolytic heart valves: a tilting-disc and a bileaflet design. Pyrolitic graphite exhibits a high biocompatibility and resistance to blood clotting. This aspect is of great importance, since blood clotting triggered by the immune response can lead to thromboembolism.

The human heart beats 4×10^7 times per year. Both stress/life (*S–N* curves) and damage-tolerant approaches (Paris–Erdogan-type relationship) have been used to estimate the fatigue life of these valves. In the *S–N* analysis, pyrolitic graphite valves are designed for 10^9 cycles (with a failure rate less than one in 100,000). This represents 25 years. However, it is known that there is considerable scatter in the data. The damage-tolerant approach was implemented by Ritchie.[11] Different experiments carried out at $\Delta K = 0.9\ K_{\max}$ yielded the results shown in Figure 14.28. These plots show two striking characteristics:

- The slope of the da/dN versus ΔK plots, m, is extremely high: approximately 100. This behavior is also found in ceramics and signifies the following: once a preexisting cracks starts to grow, it rapidly increases in velocity and leads to failure. This is highly undesirable for applications, since the material is not forgiving.
- The growth behavior shows a large degree of scatter. The threshold stress intensity factor for which growth rates da/dN are on the order of 10^{-11}m/cycle vary from specimen to specimen, and one has:

$$0.7\,\text{MPa}\,\text{m}^{-1/2} \leq \Delta K_{th} \leq 2\,\text{MPa}\,\text{m}^{-1/2}.$$

These results illustrate the importance of a comprehensive fatigue evaluation of biomaterials. Although only 50 pyrolitic graphite heart valve prostheses had failed by 1996, the high value of m shows that a

[11] R. O. Ritchie, *J. Heart Valve Dis.*, 5 (1996) S9.

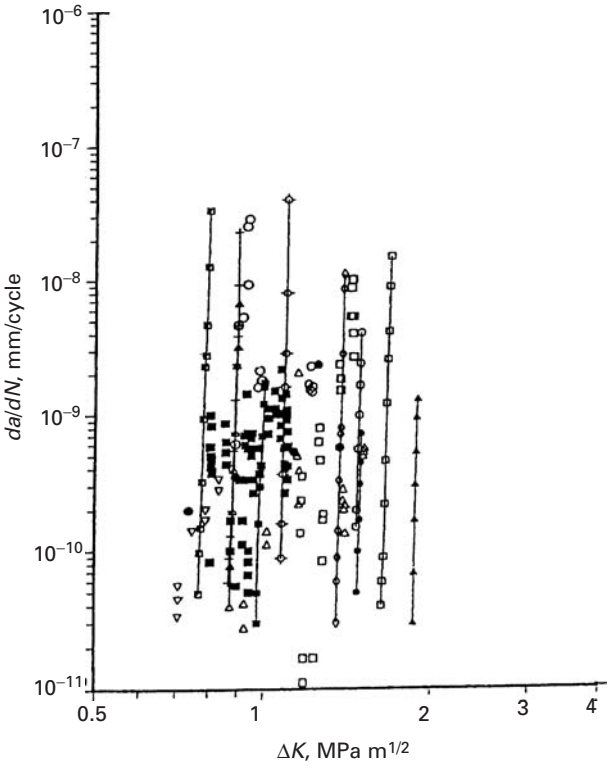


Fig. 14.28 Fatigue crack propagation rates for pyrolitic-carbon coated graphite specimens in a physiological environment; leaflet and compact-tension specimens. (Adapted from R. O. Ritchie, *J. Heart Valve Dis.*, 5 (1996) S9.)

doubling of ΔK (or the applied stress) leads to an increase in da/dN and reduction in life of 20 orders of magnitude. Thus, an accidental "nicking" of the pyrolitic graphite by a sharp scalpel during surgery can dramatically influence the life of the patient. The effect can be calculated through the Paris–Erdogan equation and is left to the curious student.

14.9 Hysteretic Heating in Fatigue

An important aspect of fatigue behavior has to do with hysteretic heating. During each loading cycle, we get a hysteresis loop. The area of the loop represents the energy spent in the cycle. Most of the hysteretic energy is dissipated as heat, and if the material is not a good thermal conductor, as is the case with most polymers, then a temperature rise can occur. Such hysteretic heating and the resultant thermal softening can be important at high strain rates in most insulators. In extreme cases, the temperature rise can be large enough that the material fails by viscous flow or melting. Even room temperature can be quite high for polymers. It is convenient to examine the temperature effects in terms of the homologous temperature T_H. Thermally activated phenomena become operative in most materials at $T_H > 0.4-0.5$. Consider a thermoplastic polymer with a melting point of 300 °C. Taking room temperature to be 300 K, we find that

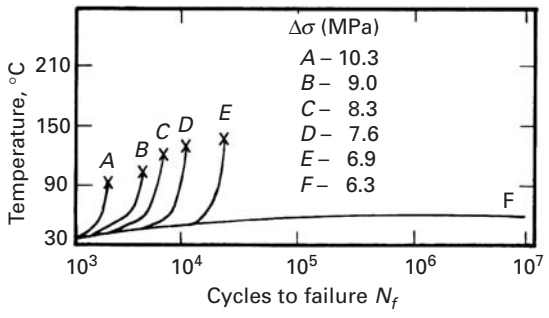


Fig. 14.29 Effect of the applied stress range $\Delta\sigma$ on temperature rise in PTFE subjected to stress-controlled fatigue. The symbol x denotes failure of the specimen. (After M. N. Riddell, G. P. Koo, and J. L. O'Toole, *Polymer Eng. Sci.* 6 (1966) 363.)

T_H for this polymer is 300/573 = 0.52. Thus, even a small increase in temperature above the ambient temperature will put such a polymer in the regimen where significant thermal softening can occur. Thermoplastics commonly have low thermal diffusivity and show nonlinear viscosity. The degree of thermal softening during fatigue will depend on the magnitude and frequency of the applied stress and on the viscoelastic characteristics of the polymer. It is easy to see that an appreciable rise in temperature can result at high stress amplitudes or frequencies. This can lead to a reduced fatigue life at frequencies greater than 10 Hz at room temperature. In fact, the ASTM standard specification D-671–71 calls for the measurement of the temperature at fatigue failure. If the temperature goes above the glass transition temperature of the polymer, thermal softening, fracture, and resolidification of the fractured material can occur.

The hysteretic temperature rise is a function of the dimensions of the specimen also. In a thinner specimen, a greater proportion of the heat generated will be lost to the environment. A thicker specimen, on the other hand, will retain a larger fraction of heat and thus show a lower fatigue endurance limit than a thinner specimen of the same material. Figure 14.29 shows the effect of applied stress on the temperature rise in polytetrafluoroethylene (PTFE) subjected to cycling at 30 Hz, at room temperature under stress control. The endurance limit $\Delta\sigma_L$ for this material under these conditions is 6.5 MPa. For $\Delta\sigma > \Delta\sigma_L$, indicated by the curves marked *A*, *B*, *C*, *D*, and *E* in the figure, a rapid increase in temperature occurred with an increasing number of cycles. For $\Delta\sigma < \Delta\sigma_L$, represented by curve *F* in the figure, the temperature rise was not high enough to cause thermal softening.

Hysteretic heating effects during high-frequency fatigue have also been observed in continuous fiber-reinforced ceramic matrix composites.[12] Unlike the heating effect observed in polymers or polymer matrix composites, the origin of heating in ceramic matrix composites is the frictional sliding between two mating surfaces, such as a fiber–matrix interface or an interlaminar shear.

[12] See, for example, J. W. Holmes and C. Cho, *J. Amer. Ceram. Soc.*, 75 (1992) 929; N. Chawla, Y. K. Tur, J. W. Holmes, J. R. Barber, and A. Szweda, *J. Amer. Ceram. Soc.,* 81 (1998) 1221.

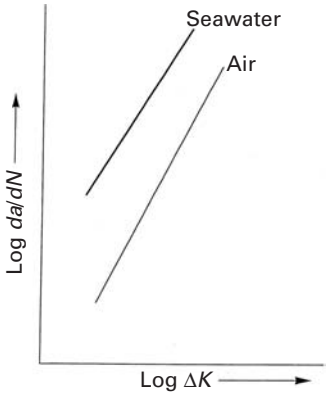


Fig. 14.30 A schematic of fatigue crack propagagtion rate as a function of cyclic stress intensity factor in air and seawater. At any given ΔK, the crack propagation rate is higher in seawater than in air.

14.10 Environmental Effects in Fatigue

Fatigue behavior of materials can be adversely afected by the environment (solid, liquid, or gaseous). We treat the topic of environmental effects on static mechanical properties of materials in Chapter 16. Here we wish to point out that an aggressive environment, under conditions involving cyclic fatigue, can hasten the nucleation of a surface crack and then propagate it to fracture. Such a phenomenon is also referred to as *corrosion fatigue*. Figure 14.30 shows schematically the fatigue crack propagation in a material subjected to cyclic fatigue in air and in seawater.

14.11 Fatigue Crack Closure

Under certain circumstances, surfaces of a fatigue crack can contact each other, and the crack will close even when the far-field stress field is still tensile. The crack does not reopen until a sufficiently high tensile stress is reached in the next loading cycle. This phenomenon, called *crack closure*, was said by a number of researchers to occur as result of crack-tip plasticity. As the applied stress on a material is increased, a plastic zone develops at the crack tip. (See Chapter 7.) As the crack grows, a plastically deformed zone is produced in its wake, while the material surrounding this zone is still elastic. The explanation of this phenomenon was that the plastically deformed zone caused the crack surfaces to close before zero stress was reached. However, for fatigue crack growth to occur, the crack must be fully open. Thus, premature contact between the crack surfaces – i.e., crack closure – results in a lowering of the crack driving force. It follows that one should use an *effective* stress intensity factor range, ΔK_{eff}, rather than ΔK in fatigue crack growth analysis. If the stress at which the crack is just open is σ_{op}, and the corresponding stress intensity factor is K_{op}, then we can define the effective cyclic stress intensity factor as

$$\Delta K_{\text{eff}} = K_{\text{max}} - K_{\text{op}}.$$

Recall that the applied cyclic stress intensity factor is given by $\Delta K = K_{\text{max}} - K_{\text{min}}$, and that $K_{\text{op}} > K_{\text{min}}$. Therefore, we will have

$$\Delta K > K_{\text{eff}}.$$

Elber proposed that ΔK_{eff} explains the R effect on the fatigue crack growth rate.[13] At high values of R, the crack closure effect is small because K_{op} approaches K_{min}, and ΔK_{eff} becomes closer to ΔK. Later, other explanations besides crack-tip plasticity were proposed for the crack closure effect. Among the various phenomena held to be responsible for crack closure are crack surface roughness, asperities in the

[13] W. Elber, *Eng. Fract. Mech.*, 2 (1970) 37.

crack wake from oxides or corrosion products, viscous fluid, and phase transformation ahead of the crack tip. The oxide-induced crack closure is possible in a material that forms an oxide film on the surface easily. When such a material is subjected to cyclic stress near the threshold regimen at low load ratios (R) and in a moist environment, corrosion products (i.e., oxides) of thickness comparable to the crack-tip opening displacements can build up at and near the tip. The oxide film continually breaks and forms behind the tip due to the crack surfaces coming together as a result of plasticity-induced closure and mode-I displacements characteristic of near-threshold crack growth. The crack closes at stress intensities above K_{min}. This mechanism is less likely to operate in a dry, oxygen-free environment and high load ratios. (Plasticity-induced closure is small.) The formation of an oxide film is time dependent and not likely to occur at high frequencies. Roughness-induced crack closure is thought to occur when the fracture surface roughness is comparable in size to the crack-tip opening displacement (CTOD) and significant mode-II deformation occurs. In such a case, cracks can become wedge–closed at contact points above the crack faces. Crack closure causes an increase in stiffness and a decrease in compliance. High values of R result in less crack closure; that is, ΔK_{eff} is closer to ΔK for higher R.

14.12 The Two-Parameter Approach

In a series of papers, Vasudevan *et al*. proposed a new, two-parametric approach to fatigue crack propagation.[14] Among the features of their approach are the following:

- It is not necessary to invoke crack closure to explain fatigue crack propagation.
- Plasticity at the crack tip cannot contribute to crack closure.
- Crack closure induced by oxide, corrosion, or roughness is very local and small.

There are five local parameters: the cyclic stress intensity factor ΔK; the maximum stress intensities K_{max}; the minimum stress intensity K_{min}; the mean cyclic stress intensity factor K_{mean}; and the R ratio. Out of these five, Vasudevan et al. used the applied driving force ΔK and the peak stress intensity K_{max} as the two parameters that are *sufficient and necessary* to analyze fatigue crack propagation data. There are thus two threshold quantities, one in each of the two parameters: the alternating stress intensity factor ΔK^* and K^*_{max}. These two must be satisfied *simultaneously* for crack propagation to occur. Based on data available in the literature on a wide range of alloys, Vasudevan and colleagues constructed a fundamental fatigue threshold curve or a fatigue map, as shown in Figure 14.31. The fundamental curve is

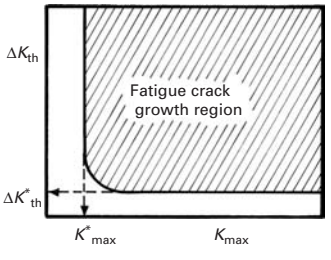


Fig. 14.31 A fatigue threshold curve. (After A. K. Vasudevan, K. Sadananda, and N. Louat, *Mater. Sci. Eng*., A188 (1994) 1.)

[14] See, for example, A. K. Vasudevan, K. Sadananda, and N. Louat, *Mater. Sci. Eng.*, A188 (1994) 1.

independent of testing and geometric parameters; it depends only on the material and environmental parameters. The shape and magnitude of such fundamental curves can vary, depending on the curve of ΔK_{th} vs. the load ratio R. On the basis of these new concepts, Vasudevan *et al*. classified the fatigue crack growth data into five different classes, using the experimental data on ΔK_{th} vs. R. Such a classification could provide a new basis for understanding the synergistic effects of various driving forces (mechanical, chemical, and microstructural) upon fatigue crack growth.

14.13 The Short-Crack Problem in Fatigue

For long cracks, under conditions of applicability of linear elastic fracture mechanics (LEFM), there exists a threshold stress intensity range ΔK_{th}, below which no fatigue crack growth occurs. The value of the cyclic threshold stress intensity depends on a variety of factors: the microstructure of the material; the test environment; the load ratio R $(=\sigma_{min}/\sigma_{max} = K_{min}/K_{max})$; various crack-tip factors such as the amount of overload, cold work, etc.; experimental techniques; and the geometry of the specimen. It has been observed that *short* fatigue cracks, in metals and polymers, can propagate at rates different from those of the corresponding long fatigue cracks under the influence of the same driving force. Generally, for a given ΔK, the growth rates of small cracks are higher than those of long cracks.[15] A *short* crack is a crack that is smaller than the microstructural unit of the materials; for instance, a crack of length equal to grain or precipitate size is a short crack. In practice, one finds that long cracks can be between 1 and 20 mm, while short cracks are smaller than 0.1 mm. The anomalous growth of short cracks is explained in Figure 14.32, a plot of log da/dN vs. log ΔK. Long cracks do not grow below a constant threshold ΔK_{th}. For long cracks, we have

$$K_I = Y\sigma\sqrt{\pi a},$$

and the threshold ΔK_{th} is a constant, as indicated in Fig. 14.32. This is in accord with LEFM, as K_I alone determines the stress state at the crack tip. However, in the short-crack regimen, cracks grow below this threshold value, as indicated by the deviation from the solid line in the figure. Short cracks propagate below the long-crack threshold (ΔK_{th}). The fatigue crack growth rate of short cracks decreases progressively, until a minimum in crack velocity occurs at a crack length on the order of the grain size; that is, $a \approx d$, where d is the grain size. This so-called short-crack anomaly arises when the crack size approaches the dimension of the microstructural feature (e.g., grain size, inclusions, etc.). Under such circumstances, homogeneity is lost, which is implicit in the LEFM treatment of the long crack.

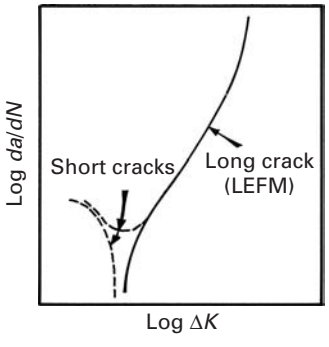


Fig. 14.32 Fatigue crack growth rates for long and short cracks.

[15] See, for example, R. A. Smith and K. J. Miller, *Int. J. Mech. Eng.*, 20 (1978) 201; S. Suresh and R. O. Ritchie, *Intl. Met. Rev.*, 29 (1984) 445.

Sadananda and Vasudevan have extended their two-parametric framework (see Section 14.12) to explain short-crack behavior.[16] According to these authors, a crack grows when both thresholds, ΔK^* and $K_{max}{}^*$, are met simultaneously. A short crack is no exception, and for it to propagate, it must meet these requirements, too. The short crack is different from the long crack in terms of the internal stresses that it encounters: All short cracks grow in internal stress fields that accentuate the applied stress to the level at which the crack propagates. An important conclusion of Sadananda and Vasudevan's work is that a similitude between the long and short crack is maintained!

14.14 Fatigue Testing

Among the reasons for carrying out fatigue testing on a material, we may include the need to develop a better understanding (fundamental or empirical) of the fatigue behavior of the material and the need to obtain more practical information on the fatigue response of a component or structure of the material. The fatigue test samples may thus range from tiny samples tested within, say, the specimen chamber of a scanning electron microscope to complete aircraft wings weighing many tons. It would be futile to try to include everything known about fatigue testing here; instead, we present some of the common techniques and point out some of their salient aspects.

14.14.1 Conventional Fatigue Tests

Conventionally, fatigue testing has been done by cycling a given material through ranges of stress amplitude and recording the number of cycles of failure. The results are reported in the form of *S–N* curves. There are two main types of loading: rotating bending tests and direct stress tests (Figure 14.33). In direct stress machines, the stress distribution over any cross section of the specimen is uniform, and we can easily apply a static mean tensile or compressive load (i.e., the *R* ratio may be varied). However, the more common and popular type of loading has been the rotating bending beam test, described next. Direct loading machines are discussed in Section 14.14.4.

14.14.2 Rotating Bending Machine

Rotating bending tests are perhaps one of the most simple and oldest types of fatigue test. They provide a simple method of determining fatigue properties at zero mean load by applying known bending moments to rotating round specimens. Commercially, many versions are available, the main difference being in the application of the load: at a single point, as in a cantilever loading machine, or by some kind of two- or four-point loading (Figure 14.33). In the latter case, the

[16] K. Sadananda and A. K. Vasudevan, in *Fatigue 96, Berlin, May 6–10, 1996*, (Oxford: Pergamon Press, 1996), p. 375; K. Sadananda and A. K. Vasudevan, in *Twenty-seventh National Symposium on Fracture Mechanics*, ASTM STP-1296, (West Conshohocken, PA: ASTM, 1996).

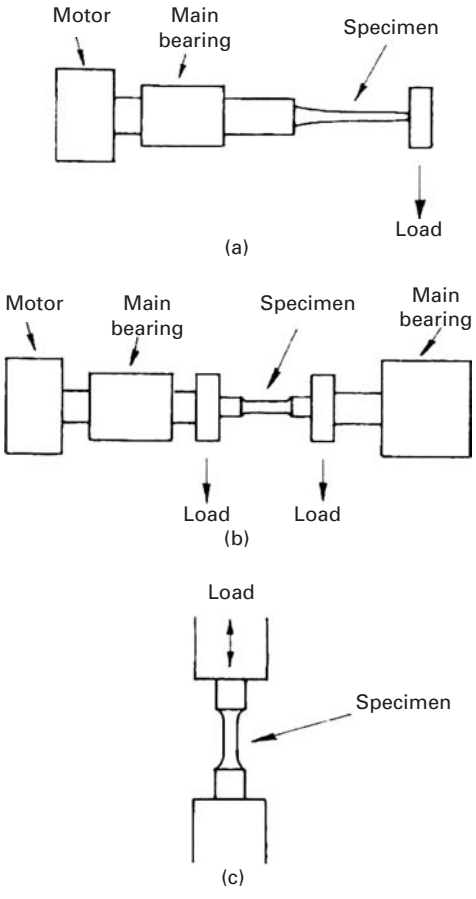


Fig. 14.33 Various loading configurations used in fatigue testing. (a) In cantilever loading, the bending moment increases toward the fixed end. (b) In two-point beam loading, the bending moment is constant. (c) Pulsating tension, or tension–compression, axial loading.

bending moment is constant over the entire test section of the specimen, and thus, we use a specimen of constant diameter. In the cantilever type of loading machine, the specimen either has a narrow waist, so that the maximum bending stress occurs at the smallest diameter, or has a tapered cross section, such that the maximum bending stresses are constant at all cross-sections. The stress at a point on the surface of a rotating bending specimen varies sinusoidally between numerically equal maximum tensile and compressive values in every cycle. Assuming the specimen to be elastic, we have

$$\pm S = \frac{32M}{\pi d^3}, \tag{14.7}$$

where S is the maximum surface stress, M is the bending moment at the cross section under consideration, and d is the diameter of the specimen. In such a test we obtain the number of cycles to failure at a given stress level. The stress level S is continually reduced, and the number of cycles to failure, N_f, increases. A logarithmic scale is used for N, and we obtain an S–N curve. In the case of ferrous materials, we generally attain a fatigue limit or endurance limit S_L (Figure 14.3). Cycling below S_L can be done indefinitely, without resulting in

failure of the material. Such an endurance limit is not encountered in nonferrous metals or polymers. In these cases, one sets an arbitrary number of cycles, say 10^7, and takes the corresponding stress to be the fatigue life of the material.

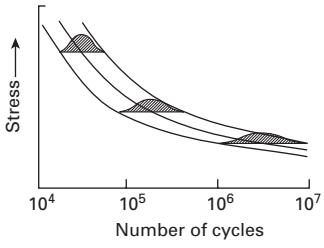


Fig. 14.34 *S–N* curve showing log-normal distribution of lives at various stress levels.

14.14.3 Statistical Analysis of *S–N* Curves

It has been observed that, if a sufficiently large number of identical specimens is fatigue tested at the same stress amplitude, a Gaussian or normal distribution describes the logarithm of the fatigue life distribution. Figure 14.34 shows a schematic *S–N* diagram with a log-normal distribution of lives at various stress levels. There is more of a spread in the lives of a group of specimens tested at a stress level greater than their fatigue limit than in the stress levels necessary to cause failure at a given life. The data from cyclic loading tests (whether rotating bending beam, pulsating tension, or axial tension–compression) must be analyzed statistically. The mean value $\bar{x}$ and the standard deviation σ for a given set of data are given by

$$\overline{x} = \frac{\sum x}{n} \tag{14.8}$$

and

$$\sigma = \left[\frac{\sum (x - \overline{x})^2}{n - 1}\right]^{1/2}, \tag{14.9}$$

where x is the cyclic life of the material at a given stress (the test value) and n is the number of test values (i.e., the number of samples tested to failure at a given stress). With these statistical parameters, one can obtain confidence limits for the probability of survival of the material. The anticipated fatigue life, with a desired level of confidence (C) that at least P of the samples will not fail, may be written as

$$\text{anticipated life } (C, P) = \bar{x} - q\sigma. \tag{14.10}$$

where q is a function of C, P, and the number of test samples used to determine $\bar{x}$ and σ. The selection of a particular confidence limit depends on the importance of the component to the structural integrity of the material. The more important the component, the higher should be the confidence limit and the lower the stress. The q values for a given distribution are available in tabulated form in the literature. Table 14.1 presents the q-values, assuming a normal distribution. With anticipated life (Equation 14.10) and the q tables, we can develop a family of curves showing the probability of survival or failure of a component (Figure 14.35).

14.14.4 Nonconventional Fatigue Testing

In the category of nonconventional fatigue testing, we include practically all modern fatigue testing other than that involving the determination of *S–N* curves. The machines used are direct loading machines. The drive system of the load train receives a time-dependent signal from the controls, converts it into a force, or

Table 14.1 *q*-values for *S–N* Data, Assuming a Normal Distribution*

P(%)	75	90	95	99	99.9	75	90	95	99	99.9
n			$C = 0.50$					$C = 0.75$		
4	0.739	1.419	1.830	2.601	3.464	1.256	2.134	2.680	3.726	4.910
6	0.712	1.360	1.750	2.483	3.304	1.087	1.860	2.336	3.243	4.273
8	0.701	1.337	1.719	2.436	3.239	1.010	1.740	2.190	3.042	4.008
10	0.694	1.324	1.702	2.411	3.205	0.964	1.671	2.103	2.927	3.858
12	0.691	1.316	1.691	2.395	3.183	0.933	1.624	2.048	2.851	3.760
15	0.688	1.308	1.680	2.379	3.163	0.899	1.577	1.991	2.776	3.661
18	0.685	1.303	1.674	2.370	3.150	0.846	1.544	1.951	2.723	3.595
20	0.684	1.301	1.671	2.366	3.143	0.865	1.528	1.933	2.697	3.561
25	0.682	1.297	1.666	2.357	3.132	0.842	1.496	1.895	2.647	3.497
n		$C = 0.90$						$C = 0.95$		
4	1.972	3.187	3.957	5.437	7.128	2.619	4.163	5.145	7.042	9.215
6	1.540	2.494	3.091	4.242	5.556	1.895	3.006	3.707	5.062	6.612
8	1.360	2.219	2.755	3.783	4.955	1.617	2.582	3.188	4.353	5.686
10	1.257	2.065	2.568	3.532	4.629	1.465	2.355	2.911	3.981	5.203
12	1.188	1.966	2.448	3.371	4.420	1.366	2.210	2.736	3.747	4.900
15	1.119	1.866	2.329	3.212	4.215	1.268	2.068	2.566	3.520	4.607
18	1.071	1.800	2.249	3.106	4.078	1.200	1.974	2.453	3.370	4.415
20	1.046	1.765	2.208	3.052	4.009	1.167	1.926	2.396	3.295	4.319
25	0.999	1.702	2.132	2.952	3.882	1.103	1.838	2.292	3.158	4.143

*Reprinted with permission from ASTM STP No. 91 (Philadelphia: ASTM, 1963), p. 67.

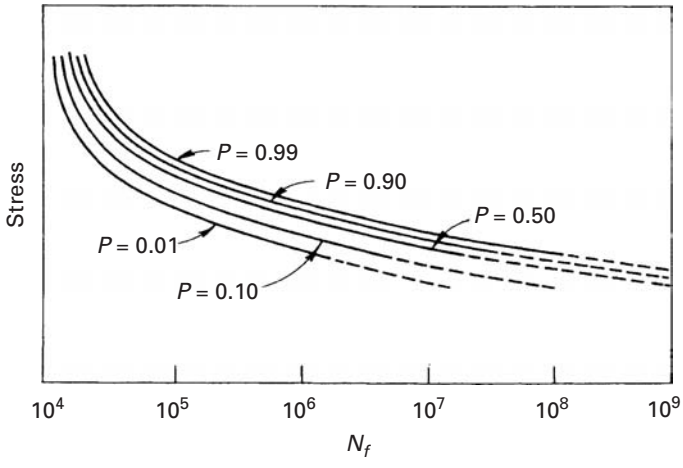


Fig. 14.35 Family of curves showing the probability of survival or failure of a component.

displacement–time excitation, and transfers this excitation to the fatigue specimen. The three common control parameters are force: deflection or displacement, and strain. For most constant-amplitude fatigue tests, a simple harmonic motion is programmed into the drive system. Electronic function generators are commonly used; they

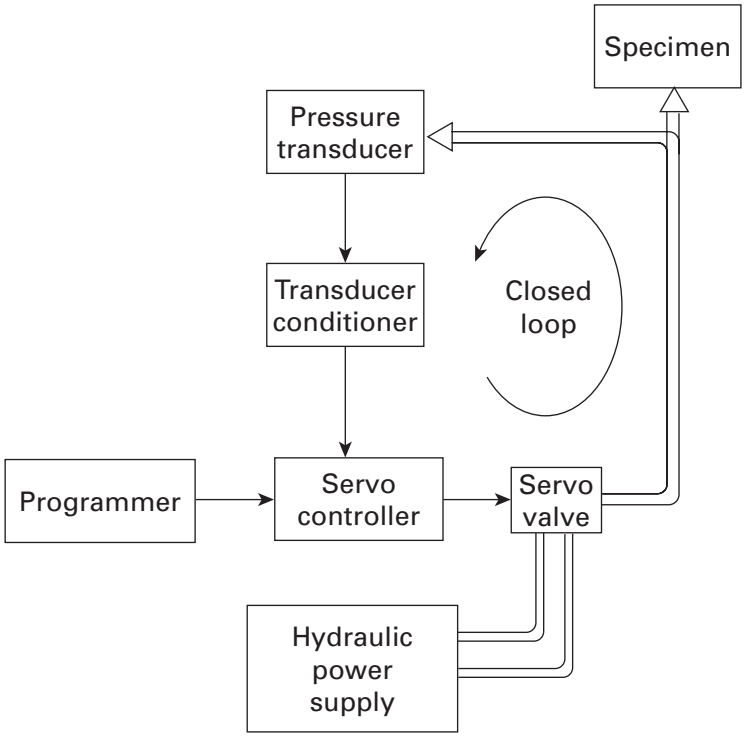


Fig. 14.36 Line diagram of a hydraulically operated closed-loop system.

generate an electrical signal that varies with time in the way the fatigue control parameter is desired to vary with time. A variety of signals can be programmed – for example, constant amplitude, constant frequency, and zero mean stress; constant amplitude, constant frequency, with a non-zero-mean stress level; random loading; and so on.

14.14.5 Servohydraulic Machines

Servohydraulically operated fatigue machines have become increasingly popular over the years. Figure 14.36 shows a line diagram of a servohydraulically operated closed-loop system. The load, applied through a hydraulic actuator, is measured by a load cell in series with the specimen. The amplified signal from the load cell is compared in a differential amplifier with the desired signal obtained from, say, a function generator. Thus, this system forms a closed-loop load control system. We can also have a displacement or strain control from a transducer or a strain gage on the specimen instead of the load cell. The actual value measured by a load cell, displacement transducer, or strain gage is continuously compared with the desired value and continuously corrected by the high-response electromagnetic servovalve. The energy is provided by a hydraulic power supply. The main advantage of such machines is a higher degree of flexibility. Larger specimen deflections are possible than are possible in electromechanical machines. Thus, we can test components involving large deflections, as well as conventional, stiff specimens. Another major advantage has to do with the versatility of the system in regard to the input

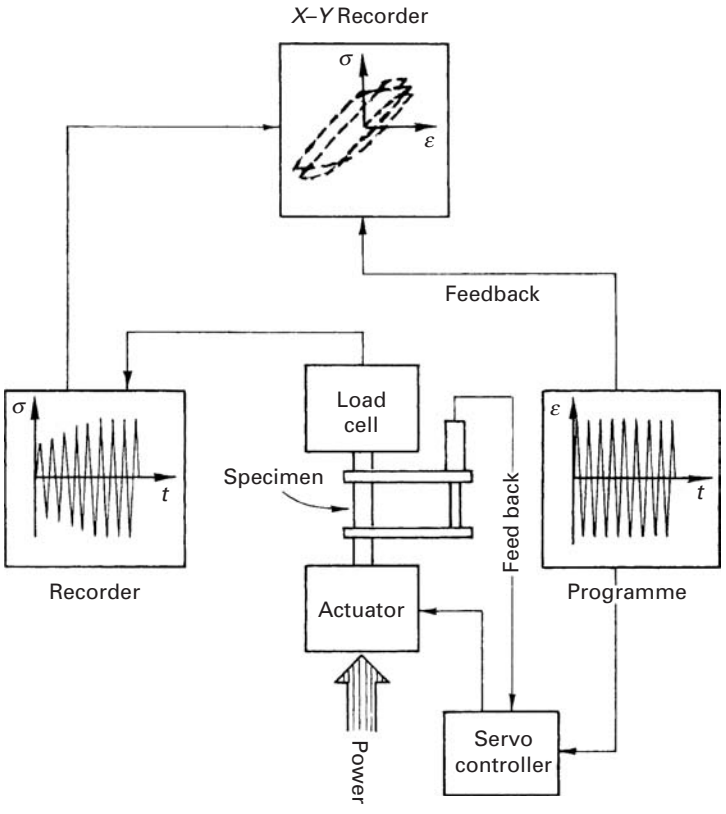


Fig. 14.37 Block diagram of a low-cycle fatigue-testing system.

signal that can be used. Virtually any analog signal from a function generator, magnetic tape, or a random noise generator is acceptable. This enables us to use not only constant-amplitude waveforms or block-program spectrum loading, but also random waveforms, such as those obtained from actual service conditions. The upshot is that the materials or components can be subjected to more realistic fatigue testing. The main disadvantage of servohydraulic machines is, of course, that they require much higher power consumption than conventional devices.

14.14.6 Low-Cycle Fatigue Tests

Under conditions of high nominal stresses (i.e., short lifetimes, less than 10^4 cycles), the constant-stress amplitude test gives only limited information. This is because rather large plastic strain components are involved in such cases. Under such conditions, the cyclic stress–strain curves obtained under strain control become more useful. Servohydraulic machines are generally used in a closed-loop mode. Figure 14.37 shows schematically a cyclic straining facility. Axial tension–compression is generally employed. We measure stress as a function of the number of strain reversals. Usually, stress and strain signals are fed to an *X*–*Y* recorder, and a complete hysteresis loop is obtained. The area of the loop is the plastic strain energy per cycle.

Cyclic Stress–Strain Curves

We can obtain cyclic stress–strain curves by linking the tips of a series of hysteresis loops obtained from equivalent specimens tested at different plastic strain amplitudes ($\Delta\varepsilon_p$). There are also methods of obtaining cyclic stress–strain curves from a single specimen. The hysteresis loop adjusts rather quickly following a sudden change in $\Delta\varepsilon_p$. Thus, we can obtain a cyclic stress–strain curve from one specimen tested at several strain amplitudes. This is called a *multiple-step test*. Another method is the *incremental step test* with one specimen. This method consists of gradually increasing the cyclic strain range until a cyclic strain of about $\pm 1\%$ is attained. The strain range is then slowly reduced, and the procedure is repeated until the material is stabilized.

14.14.7 Fatigue Crack Propagation Testing

As pointed out earlier, the process of fatigue failure consists of the following two stages:

1. A certain number of cycles N_i in which a small crack is initiated. Some people include in this stage early growth of the microcrack to a somewhat larger crack.
2. Propagation of a major crack. Generally, this occurs in such a way that we are able to describe the propagation behavior by some kind of standard relationship, say, the Paris–Erdogan relation. There is a substage of this propagation stage wherein the final rupture occurs, namely, when the crack has reached a certain critical length for the material, the applied stress, and the test piece or structural component.

Much attention has been paid to the crack propagation behavior of materials in fatigue. Fatigue crack growth rates under service conditions can be of great importance, especially in determining inspection intervals. For example, wheels on large aircraft may have an ample safe lifetime after the appearance of detectable cracks. What we want to be sure of is that these cracks will not grow to a size that is critical for the part during the time available before the next periodic inspection.

Flat-sheet specimens are commonly chosen for crack propagation studies. The starter notch can be a side edge notch, a central through-the-thickness hole, or some other shape appropriate to the form of defects observed in service. These notches can be cut by a mechanical saw, electrical discharge machining, and so on. Usually, crack growth measurements are made after a small initial propagation in which there is an atomically sharp fatigue crack. The crack length is measured as a function of the number of cycles, and subsequent analysis is carried out in terms of fracture mechanics concepts. Synchronized strobe lighting can be used to illuminate the surface of the sample in order to provide a stable, vibration-free crack-length reading capability, or, in more sophisticated cases, a movie record can be obtained of the increase in the length of the crack. Traveling

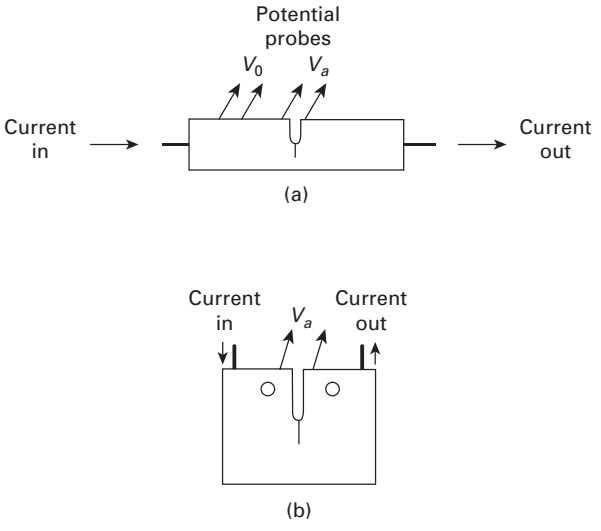


Fig. 14.38 Electric potential drop method for crack growth measurements. (a) Bend specimen. (b) Compact-tension specimen. V_a is the potential drop across the crack while V_0 is some reference potential. The crack length is obtained as function of the normalized potential (V_a/V_o). The sample is loaded by inserting pins in holes shown.

or stereo zoom microscopes are used in manual monitoring crack length. Such devices typically can read up to 0.01 mm. We may have scale markings photographically prepared on the sample or have a scale inserted in the ocular piece of the microscope. We can also use crack propagation gages, consisting of a series of 20 or 25 parallel, equally spaced resistance wires in the form of a grid. Crack length is measured by monitoring the overall change in resistance. In the electric potential drop method, a constant direct current is passed through the specimen containing a crack. The resistance of the specimen changes as the crack grows and is detected by measuring the potential drop across the mouth of the starter notch. Figure 14.38 shows the setup for a bend and a compact-tension specimen. As a crack is observed propagating, the number of cycles required for each increment is recorded, and a crack growth rate da/dN is computed from the curve of the crack length a versus the number of cycles N. The cyclic stress intensity factor at the crack tip (ΔK) may be computed from the crack length and the load. By plotting da/dN versus ΔK, we can obtain the fatigue crack growth characteristics of the material.

Suggested Reading

J. M. Barsom and S. T. Rolfe. *Fracture and Fatigue Control in Structures*, 3rd ed. West Conshohocken, PA: ASTM, 1999.

D. L. Davidson and J. Lankford, *Intl. Mater. Rev.*, 37 (1992) 65.

A Guide for Fatigue Testing and the Statistical Analysis of Fatigue Data. ASTM STP 91. Philadelphia, PA: ASTM, 1963.

R. W. Hertzberg, and J. A. Manson. *Fatigue of Engineering Plastics*. New York, NY: Academic Press, 1980.

J. G. Johnson. *The Statistical Treatment of Fatigue Experiments*. New York, NY: Elsevier, 1964.

R. O. Ritchie, *Mater. Sci. Eng.*, 103 (1988) 15.
S. Suresh. *Fatigue of Materials*, 2nd ed. Cambridge, U. K.: Cambridge University Press, 1995.

Exercises

14.1 Many operations, such as machining, grinding, electroplating, and case-hardening, may induce residual stresses in a material. Discuss, in general terms, the effect of such residual stresses on the fatigue life of the material.

14.2 A steel has the following properties:

Young's modulus $E = 210$ GPa,
Monotonic fracture stress $\sigma_f = 2.0$ GPa,
Monotonic strain at fracture $\varepsilon_f = 0.6$,
Exponent b (cyclic) $= 0.09$; c (cyclic) $= 0.06$.

Compute the total strain that a bar of this steel will be subject to under cyclic straining before failing at 1,500 cycles.

14.3 The low-cycle fatigue behavior of a material can be represented by

$$\sigma_L = \sigma_{UTS} N_f^{-0.1},$$

where σ_L is the endurance limit, σ_{UTS} is the ultimate tensile strength, and N_f is the number of cycles to failure. If σ_{UTS} for this material is 500 MPa, find its endurance limit. A sample of the material is subjected to block loading consisting of 40, 30, 20, and 10% of fatigue life at σ_L, $1.10\sigma_L$, $1.2\sigma_L$, and $1.3\sigma_L$, respectively. Use the Palmgren–Miner relationship to estimate the fatigue life of the sample under this block loading.

14.4 A microalloyed steel was subjected to two fatigue tests at ± 500 MPa and ± 300 MPa. Failure occurred after 10^3 and 10^5 cycles, respectively, at these two stress levels. Making appropriate assumptions, estimate the fatigue life at ± 400 MPa of a part made from this steel that has already suffered 10^4 cycles at ± 350 MPa.

14.5 The curve of crack growth rate da/dN vs. cyclic stress intensity ΔK for a material in the Paris regimen is shown in Figure Ex14.5. Determine the parameters C and m for this material. (*Hint*: Take any two points on the straight portion of the curve, and determine the slope m of the line.)

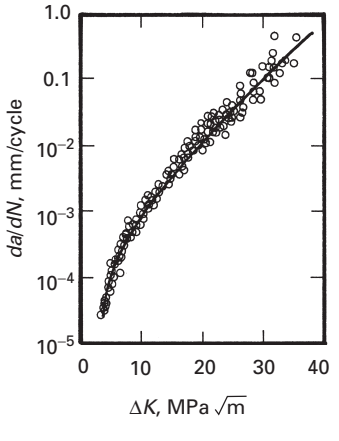


Fig. Ex14.5

14.6 A steel has the following properties:

Yield stress $\sigma_y = 700$ MPa,
Fracture toughness $K_{Ic} = 165$ MPa m$^{1/2}$.

A plate of this steel containing a single edge crack was tested in fatigue under $\Delta\sigma = 140$ MPa, $R = 0.5$, and $a_0 = 2$ mm. It was observed experimentally that fatigue crack propagation in the steel could be described by the Paris-type relationship

$$\frac{da}{dN} \text{ (m/cycle)} = 0.66 \times 10^{-8} (\Delta K)^{2.25},$$

where ΔK is measured in MPa m$^{1/2}$.

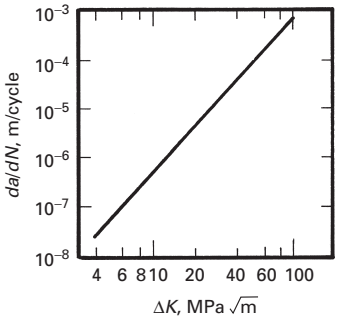


Fig. Ex14.7 Adapted from *ASM Metals Handbook*, 9th ed. (Metals Park, OH. ASM, 1986), p. 103.

(a) What is the critical crack size a_c at σ_{max}?
(b) Compute the fatigue life of the steel.

14.7 Obtain the parameters for the Paris-type relationship for the data shown in Figure Ex14.7 for the aluminum alloy 7075-T6.

14.8 For the 7075-T6 alloy in Exercise 14.7, determine the length of a crack after 10^5 cycles if the initial crack size was equal to 0.2 mm and the cyclic loading was such that, at the onset of fatigue, $\Delta K = 10$ MPa $m^{1/2}$.

14.9 Suppose that the fatigue life of a precracked specimen is totally occupied by crack propagation. If, in a certain case, the initial crack growth rate is given by

$$\frac{da}{dN} = C\,\Delta K^2$$

and $\Delta K = 2\alpha\ \Delta\sigma\sqrt{\pi a}$, show that

$$C = \frac{1}{4N\,\Delta\sigma^2\alpha^2\pi}\ln\frac{a_f}{a_0},$$

where α is a constant N is the fatigue life, a_0 is the initial crack length, and a_f is the final crack length, of the specimen.

14.10 Fatigue crack propagation in a polymer can be described by the relationship

$$\frac{da}{dN} = 0.5 \times 10^6 \Delta K^{3.5},$$

where da/dN is in m/cycle and ΔK is in MPa $m^{1/2}$. A sample with the following dimensions and a central through-the-thickness crack was subjected to fatigue under a maximum load of 200 N and a minimum load of zero:

thickness $B = 10$ mm,
width $W = 50$ mm,
crack length $2a = 10$ mm.

Using an appropriate expression for ΔK (see Chapter 7), calculate da/dN for this sample.

14.11 (Design Problem) Estimate the life of a hip implant made of 304L stainless steel if it contains initial flaws with length $2c = 200$ μm and a height $2a = 100$ μm. Assume that the force applied on the artificial hip is

walking: $3W$,
running: $7W$,

where W is the weight of the person. The fatigue response of 304L can be represented by

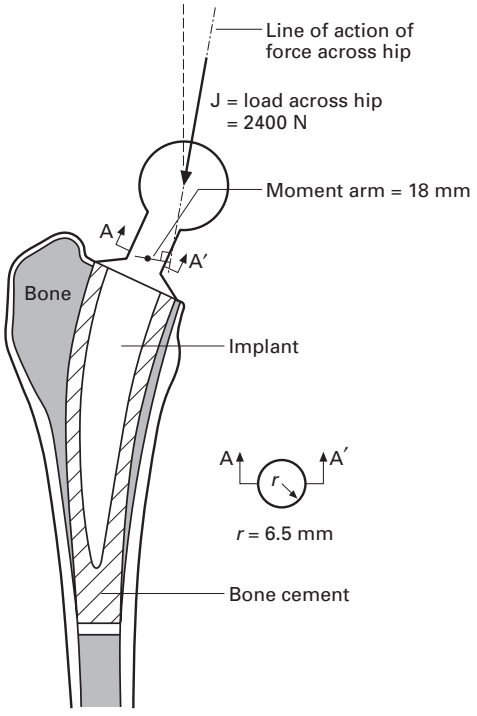
$$\frac{da}{dN} = 5.5 \times 10^{-9} \Delta K^3,$$

where da/dN is in mm/cycle and K is in MPa $m^{1/2}$. The person is assumed to

(a) walk 3 hours per day
(b) walk 3 hours and jog 20 minutes per day.

Make all necessary assumptions.

Fig. Ex14.11



14.12 Assuming that fatigue failures are initiated at the "weakest link," we may use the Weibull frequency distribution function to represent the fatigue lives of a group of specimens tested under identical conditions. We have

$$f(N) = \frac{b}{N_a - N_0}\left(\frac{N - N_0}{N_a - N_0}\right)^{b-1} \exp\left[-\left(\frac{N - N_0}{N_a - N_0}\right)^{b}\right]$$

where N is the specimen's fatigue life, N_0 is the minimum life 0, N_a is the characteristic life at 36.8% survival of the population (36.8% $= 1/e$, where $e = 2.718$), and b is the shape parameter of the Weibull distribution curve. Letting $x = (N - \mathrm{N}_0)/(\mathrm{N}_a - N_0)$, plot frequency curves $f(N)$ versus x for $b = 1$, 2, and 3.

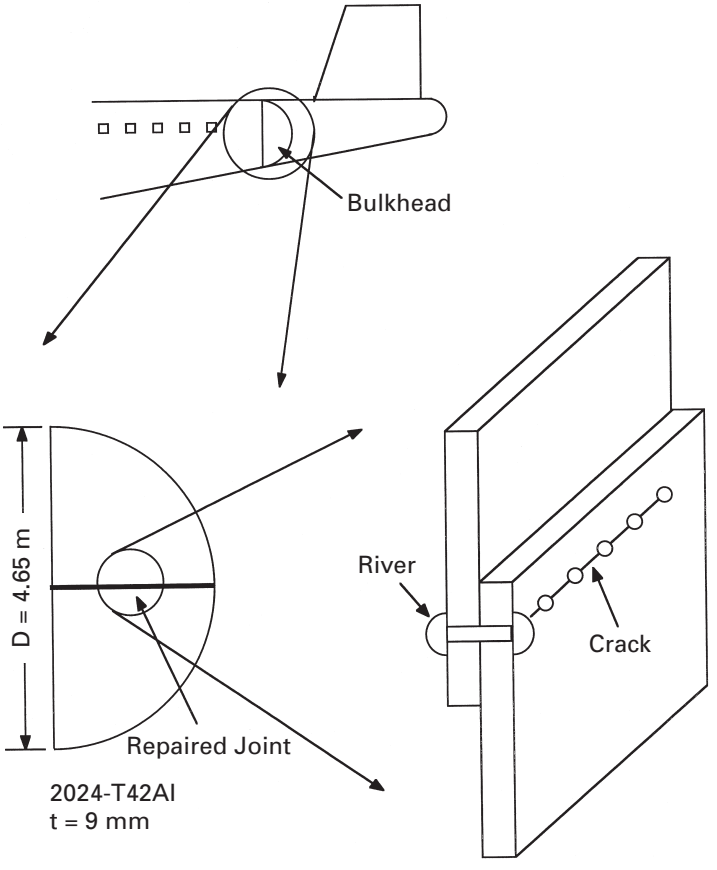
14.13 Fatigue data are, generally, analyzed cumulatively to determine the survival percentage. The Weibull cumulative function for the fraction of population failing at N is an integration of the expression for $f(N)$ in the preceding exercise. Show that this function is

$$F(N) = 1 - \exp\left[-\left(\frac{N - N_0}{N_a - N_0}\right)^{b}\right].$$

Transform F into a straight-line relationship by taking the logarithm of the logarithm of the equation. Show how this relationship can be used on log–log paper for graphically fitting the Weibull cumulative distribution and for graphically estimating the parameters b, N_0, and N_a.[17]

[17] See, for example, C. S. Yen, in *Metal Fatigue: Theory and Design*, A. F. Madayag, ed. (New York: Wiley, 1969), p. 140.

Fig. Ex14.14



14.14 (Design Problem) One of the worst single aircraft accidents in history resulted in the loss of 520 lives. It was produced by the growth of a fatigue crack in the back of the bulkhead of a Boeing 747 plane. (See Figure Ex14.14.) The fatigue fracture was caused by a repair that replaced a double row of rivets by a single row in certain places. The accident occurred after the plane reached an altitude of 7,200 m. The atmospheric pressure decreases by 12 Pa for every meter increase in altitude.

(**a**) Calculate the stress cycle to which the pressurized cabin and bulkhead were subjected in each takeoff–landing sequence of the plane.

(**b**) Establish the critical crack length for which catastrophic growth would occur.

(**c**) Assuming that fatigue failure started at one of the rivet holes (which had a diameter of 12 mm) and that it propagated through subsequent holes, calculate the number of cycles necessary to bring down the "big bird," given the following data:

Paris relationship constants, $C = 5 \times 10^{-8}$ and $m = 3.6$,

$\sigma_y = 310\,\text{MPa}$,

$\sigma_{\text{UTS}} = 345\,\text{MPa}$.

14.15 An alloy steel plate is subjected to constant-amplitude uniaxial tension–compression fatigue. The stress amplitude is 100 MPa. The plate has a yield

strength of 1,500 MPa, a fracture toughness of 50 MPa $m^{1/2}$ and an edge crack of 0.5 mm. Estimate the number of fatigue cycles to cause fracture if da/dN (m/cycle) $= 1.5 \times 10^{-24}\ \Delta K^4$, where the units of K are MPa $m^{1/2}$. Use $Y = 1$ in the fracture toughness equation.

14.16 A part is subjected to cyclic loading at 50 Hz. The stress intensity at the tip of a flaw is just above ΔK_{th}, and da/dN is equal to 10^{-8} mm/s. What is the length of crack after one week? Take $m = 3$.

14.17 The low cycle fatigue of a metallic alloy can be described by the Coffin–Manson expression relating the number of cycles to failure, N_f, to plastic strain, $\Delta\varepsilon_{pl}$, as follows:

$$N_f^{0.5}\Delta\varepsilon_{pl} = 0.4.$$

Compute the number of cycles to failure if the alloy fails at a plastic strain of 5×10^{-3}.

14.18 An alloy of titanium, Ti-6Al-4V, has a fracture toughness of 50 MP $m^{1/2}$. A plate of this material has a penny-shaped crack of diameter 10 mm. Take the yield stress of this alloy to be 980 MPa. What is the stress that can be applied in service to this plate without causing failure? Given: $K = 2\sigma(a/\pi)1/2$.

14.19 A 2024-T6 aluminum alloy used for the frame of an aircraft was tested in a Wöhler-type machine rotating at 400 rpm (sinusoidal stress variation, mean stress equal to zero). The following results were obtained:

stress range = 310 MPa; $N = 10^4$ cycles;
stress range = 230 MPa; $N = 10^7$ cycles.

Predict the life of the aircraft if it is in the air 16 hours per day and if it is subjected to the stress range of 180 MPa at the same frequency. The material obeys Basquin's law.

14.20 Explain the effect on fatigue life of the following design and environmental factors:

(a) A high polish surface finish.
(b) A rivet hole.
(c) Increasing the mean stress, but keeping the range constant.
(d) A corrosive atmosphere.

14.21 (Design Problem) On April 28, 1988, an Aloha Airlines Boeing 737, was flying from Hilo to Honolulu. As the airplane leveled at 8,000 m, both pilots heard a loud "clap" or "whooshing" sound followed by a wind noise behind them. The cockpit entry door was missing and that there was blue sky where the first-class ceiling had been. The plane was able to land. After the accident, a passenger stated that as she was boarding the airplane through the jet bridge at Hilo, she observed a longitudinal fuselage crack. The crack was in the upper row of rivets, about halfway between the cabin door and the edge of the jet bridge hood.

What is your interpretation on the cause of the accident? If the plane is pressurized, what is the critical crack length at an altitude of 8,000 m? Make all necessary assumptions and use sketches in your answer.

14.22 A polymer has cycles to failure N_f as 50 when the stress amplitude is 43 MPa, and cycles to failure N_f as 5,000 when the stress amplitude

is 15 MPa. Estimate the cycles to failure when the stress amplitude is 20 MPa.

14.23 An alloy has a Paris-type fatigue crack propagation relationship:

$$\frac{da}{dN}(m/cycle) = 0.8 \times 10^{-6}(\Delta K)^2.$$

If it is subjected to fatigue under a stress amplitude $\Delta\sigma = 100$ MPa, how long will the crack be after 1,000 cycles. Take the crack to be 0.5 mm.

14.24 Calculate the mean stress for a titanium alloy with UTS = 1,140 MPa, yield stress 1,075 MPa, by the Goodman, Gerber, and Soderberg relationship. The stress amplitude is 120 MPa and $\sigma_0 = 140$ MPa.

Chapter 15

Composite Materials

15.1 Introduction

We can define a *composite* material as a material consisting of two or more physically and/or chemically distinct phases, suitably arranged or distributed. A composite material usually has characteristics that are not depicted by any of its components in isolation. Generally, the continuous phase is referred to as the *matrix*, while the distributed phase is called the *reinforcement*. Three items determine the characteristics of a composite: the reinforcement, the matrix, and the interface between them. In this chapter, we provide a brief survey of different types of composite materials, highlight some of their important features, and indicate their various applications.

15.2 Types of Composites

We may classify composites on the basis of the type of matrix employed in them – for example, polymer matrix composites (PMCs), metal matrix composites (MMCs), and ceramic matrix composites (CMCs). We may also classify composites on the basis of the type of reinforcement they employ (see Figure 15.1):

1. Particle reinforced composites.
2. Short fiber, or whisker reinforced, composites.
3. Continuous fiber, or sheet reinforced, MMCs.
4. Laminate composite.

Figure 15.2 shows typical microstructures of some composites: boron fiber/Al (Figure 15.2(a)), short alumina fiber/Al (Figure 15.2(b)), and NbC/Ni–Cr, an *in situ* (eutectic) composite (Figure 15.2(c)). Examples of microstructure of a silicon carbide particle (three different volume fractions) reinforced aluminium matrix are given in Figure 15.3. These were made by hot pressing of powders followed by hot extrusion. Note the preferential alignment of SiC particles in the extrusion direction.

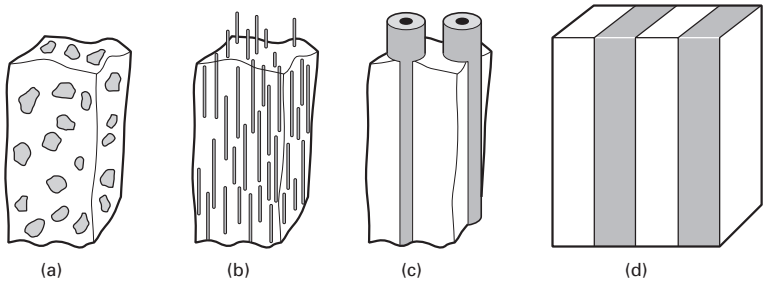
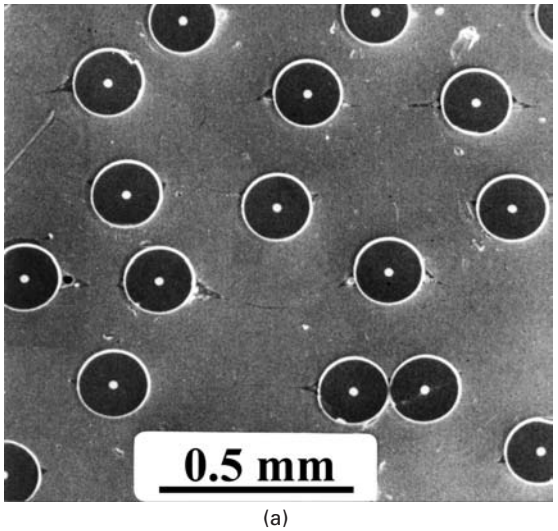
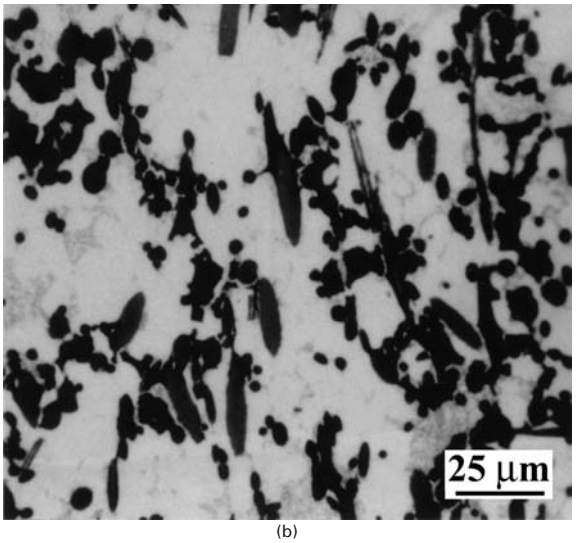


Fig. 15.1 Different types of reinforcement for composites: (a) particle reinforcement; (b) short fiber reinforcement; (c) continuous fiber reinforcement; (d) laminate reinforcement



(a)



(b)

Fig. 15.2 (a) Transverse section of a boron fiber reinforced aluminum composite. $V_f = 10\%$. (b) Section of a short alumina fiber/aluminum matrix composite. (c) Deeply etched transverse section of a eutectic composite showing NbC fibers in an Ni–Cr matrix. (Courtesy of S. P. Cooper and J. P. Billingham, GEC Turbine Gnerators Ltd, U.K.)

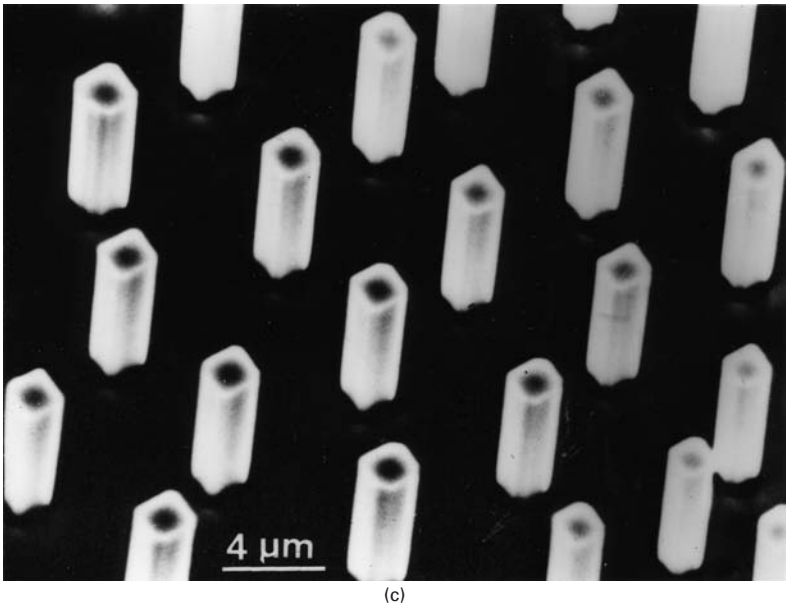


Fig. 15.2 (*cont.*)

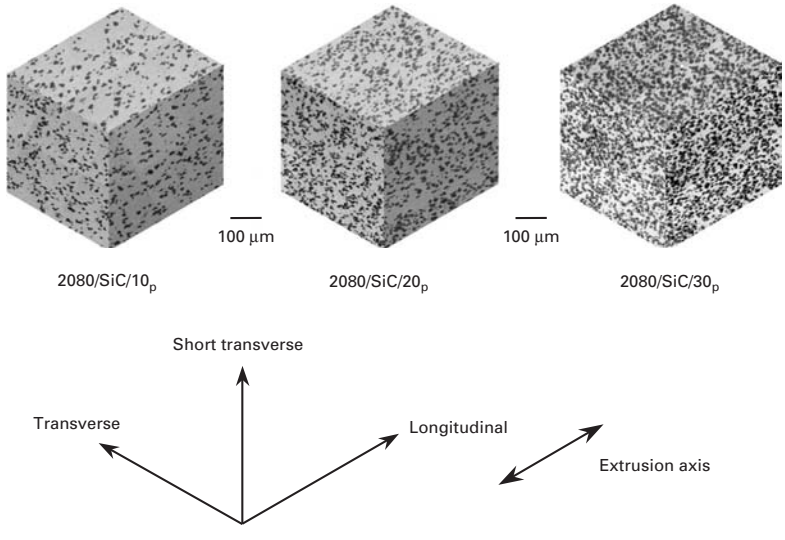


Fig. 15.3 Microstructure of a silicon carbide particle (10, 20, and 30%, three different volume fractions) reinforced aluminium alloy (2080) matrix composites made by hot pressing of powders followed by hot extrusion. Note the preferential alignment of SiC particles in the extrusion direction. The number and subscript p indicate the volume fraction of SiC particles in the composites. (Courtesy of N. Chawla.)

15.3 Important Reinforcements and Matrix Materials

Many reinforcement materials are available in a variety of forms: continuous fibers, short fibers, whiskers, particles, etc. Some of the important ones are listed in Table 15.1, along with a summary of their salient characteristics. Reinforcements include organic fibers such as polyethylene and aramid, metallic fibers, and ceramic fibers and particles.

Table 15.1 | Properties of Some Important Reinforcement Fibers

Materials (Fibers)	Tensile Modulus (GPa)	Tensile Strength (GPa)	Density (g/cm^3)
Alumina	350–380	1.7	3.9
Boron	415	3.5	2.5–2.6
SiC	300–400	2.8	2.8
E-Glass	71	1.8–3.0	2.5
Carbon P100 (pitch-based)	725	2.2	2.15
Carbon M60J (PAN-based)	585	3.8	1.94
Aramid	125	3.5	1.45
Polyethylene	110	3	0.97

A variety of materials – polymers (thermoset and thermoplastic), metals and their alloys, intermetallics, glasses, glass–ceramics, and crystalline ceramics – can be used as matrices. Most polymer matrix composites consist of cross-linked thermoset polymers such as epoxy, phenolic, and polyester resins. Cross-linked thermoset polymers have an amorphous structure. (See Chapter 1.) Phenolics have the advantage of being cheaper than epoxy and polyester resins. Their main disadvantage is that toxic by-products are liberated during the curing process. Cross-linking of polymer chains during curing in polyester and epoxy occurs by an addition mechanism, without any by-product produced. Glass fiber reinforced phenolic, polyester, or epoxy have been in use in a variety of automotive components. Epoxy resins have the added attractive feature that they can be partially cured to make fiber/epoxy prepregs, which are subsequently consolidated into a component, usually in an autoclave. A *prepreg* is a thin lamina of unidirectional (or sometimes woven) fiber/polymer composite protected on both sides with easily removable separators. A typical unidirectional prepreg comes in the form of a roll that is 300–1,500 mm wide, 0.125 mm thick, and 50–250 m long. Typically, the polymer content is approximately 35% by volume. It is not uncommon to use 50 or even more such plies in a component.

To a lesser extent than thermosets, thermoplastic resins such as poly(ether ether ketone) or PEEK and poly(phenylene sulfide), or polyamide, are used as matrix materials. PEEK is a high-performance semicrystalline thermoplastic that has been used as a matrix for carbon fibers. It is attractive as a polymeric matrix material because of its superior toughness and impact properties, compared to those of epoxies. Such properties are a function of the crystalline content and morphology of the thermoplastic.

There are many other important matrix materials. Among metallic matrix composite, we have aluminum and its alloys, mainly because of their low density and excellent strength, toughness, and resistance

to corrosion, as well as titanium alloys, magnesium alloys, copper, etc. Among intermetallic and ceramic matrix composites are a variety of intermetallic compounds such as molybdenum disilicide and aluminides of nickel and titanium, silica-based glasses, glass–ceramics, and crystalline ceramics such as alumina and silicon carbide.

Example 15.1

Carbon black is frequently used as a particulate filler in polymers, both thermoplastic and thermoset. Describe some of the important effects of the addition of carbon black to polymers.

Solution: Carbon black is stronger than the polymer matrix; thus, we get a stronger and harder composite. Carbon black is also thermally more stable than the polymer matrix; therefore, its addition results in a thermally stable composite – that is, improved creep resistance. In addition, carbon black leads to an enhanced dimensional stability. (It has a higher modulus and lower expansion coefficient than the polymer.)

15.3.1 Microstructural Aspects and Importance of the Matrix

The differential thermal expansion between the reinforcement and the metal matrix can introduce a high dislocation density in a metallic matrix, especially in the near-interface region of the matrix. This high matrix dislocation density, as well as the reinforcement–matrix interfaces, can provide high diffusivity paths in a composite. A semicrystalline thermoplastic matrix can have its crystallization kinetics altered by the presence of a reinforcement such as a carbon fiber. In the case of a ceramic matrix composite, the brittle matrix can undergo cracking in response to such thermal stresses. Thus, the characteristics of a matrix material are changed by the very process of making a composite. Such is not commonly the case with the reinforcement, however; only in rare instances of very high temperature processing, as, for example, in the case of a CMC, can the reinforcement also undergo a change in its microstructure. Hence, the matrix is much more than a mere medium or glue to hold the reinforcement, be that fibers, whiskers, or particles. Accordingly, it should be chosen after due consideration of its chemical compatibility and thermal mismatch with the reinforcement. Processing-induced chemical reactions and thermal stresses can cause changes in the microstructure of the matrix. These microstructural changes in the matrix, in turn, can affect the mechanical and physical behavior of the composite. The matrix strength in the composite (the *in situ* strength) will not be the same as that determined from a test of an unreinforced matrix sample in isolation, because the matrix is likely to suffer

several microstructural alterations during processing and, consequently, changes in its mechanical properties.

The final matrix microstructure is a function of the type, diameter, and distribution of the fiber, as well as conventional solidification parameters. For example, Mortensen et al. observed normal dendritic structure in the unreinforced region of the matrix of a silicon carbide fiber/Al–4.5% Cu matrix, while in the reinforced region, the dendritic morphology was controlled by the fiber distribution.[1] Second phase appeared preferentially at the fiber–matrix interface or in the narrow interfiber spaces. In short, the microstructure of the matrix in the fiber composite is likely to differ significantly from that of the unreinforced matrix material processed in an identical manner.

Porosity is a critical defect that is likely to be present in the matrix. Porosity can be highly deleterious to the overall performance of a composite. The main sources of porosity are any gas evolution, shrinkage occurring upon solidification, and, in the case of CMCs, incomplete elimination of any binder material. In a composite made by liquid infiltration of a preform, a high volume fraction of reinforcement may impede the flow of the liquid and inhibit any "bulk movement" of the semisolid matrix material. The desirability of having a low porosity in a PMC may be appreciated by the fact that the final stage in any PMC fabrication is called *debulking*, which serves to reduce the number of voids. A low quantity of voids is necessary for improved interlaminar shear strength. In the case of a CMC, made by sintering of glass or glass–ceramic powder and fibrous reinforcements, the reinforcements can form a network that impedes the transfer of mass required for sintering. Depending on the thermal expansion coefficients of the components, there is also the possibility of developing hydrostatic tensile stresses in the matrix that will counter the driving force for sintering. The following are some of the common structural defects in composites:

- Matrix-rich (fiber-poor) regions.
- Voids.
- Microcracks (which may form due to thermal mismatch between the components, curing stresses, or the absorption of moisture during processing).
- Debonded regions.
- Delaminated regions.
- Variations in fiber alignment.

15.4 Interfaces in Composites

The interface region in a particular composite has a great deal of importance in determining the ultimate properties of the composite,

[1] A. Mortensen, M. N. Gungor, J. A. Cornic, and M. C. Flemings, *J. Met.*, 38 (Mar. 1986) 30.

essentially for two reasons: The interface occupies a very large area per unit volume in a composite, and, in general, the reinforcement and the matrix form a system that is not in thermodynamic equilibrium. We can define an interface as a boundary surface between two phases in which a discontinuity in one or more material parameters occurs. According to this definition, an interface is a bidimensional region across which a discontinuity occurs in one or more material parameters. In practice, there is always some volume associated with the interface region, and a gradual transition in material parameters occurs over the thickness of this interfacial zone. Some of the important parameters that can show a discontinuity at the interface are the elastic moduli, strength, chemical potential, coefficient of thermal expansion of the composite, and others. A discontinuity in chemical potential is likely to cause a chemical interaction, leading to an interdiffusion zone or the formation of a chemical compound at the interface. A discontinuity in the thermal expansion coefficient means that the interface will be in equilibrium only at the temperature at which the reinforcement and the matrix were brought into contact. At any other temperature, biaxial or triaxial stress fields will be present, because of the thermal mismatch between the components of a composite. Thermal stresses due to a thermal mismatch will generally have an expression of the form

$$\sigma = f(E, a, b, r)\,\Delta\alpha\,\Delta T, \tag{15.1}$$

where $f(E, a, b, r)$ is a function of the elastic constants E and the geometric parameters a, b, and r; $\Delta\alpha$ is the difference in the expansion coefficients of the components, and ΔT is the change in temperature of the material. The term $(\Delta\alpha\Delta T)$ is, of course, the thermal strain. (Detailed expressions for thermal stresses in composites can be found in textbooks on composites.[2])

The reader can easily show that, for a given diameter and volume fraction of reinforcement, a fibrous composite will have a larger interfacial area than a particulate composite. The important point, however, is that the interfacial area in a composite increases with a decreasing reinforcement diameter. It is easy to visualize the interfacial area becoming very large for reinforcements less than 10–20 μm in diameter. Since chemical and/or mechanical interactions between the reinforcement and the matrix occur at interfaces, an extremely large area of interface has an enormous importance in determining the final properties and performance of a composite.

15.4.1 Crystallographic Nature of the Fiber–Matrix Interface

In crystallographic terms, ceramic–metal interfaces in composites are, generally, incoherent and high-energy interfaces. Accordingly, they can act as very efficient vacancy sinks and provide rapid diffusion

[2] See, for example, K. K. Chawla, *Composite Materials: Science and Engineering*, 2nd ed. (New York: Springer, 1998).

paths, segregation sites, and sites of heterogeneous precipitation, as well as sites for precipitate-free zones. Among the possible exceptions are some eutectic composites and the XD™ type of particulate composites. The *in situ* or eutectic composites do show semicoherent interfaces, that is, the lattice mismatch between the matrix phase and the reinforcement phase is accommodated by creating a network of dislocations. In general, interfaces in most composites are incoherent.

15.4.2 Interfacial Bonding in Composites

Some bonding must exist between the reinforcement and the matrix for load transfer from matrix to fiber to occur. Neglecting any direct loading of the reinforcement, the applied load is transferred from the matrix to the reinforcement via a well-bonded interface. However, the degree of bonding desired in different types of composites is not the same. In general, one would like to have a strong interfacial bonding in the case of PMCs and MMCs, with which one aims at exploiting the high stiffness and load-bearing capacity of a fibrous reinforcement. In CMCs, on the other hand, one would like to have a weak interfacial bonding, so that an advancing crack gets deflected there rather than passing through unimpeded. This is because the main objective in CMCs is to enhance their toughness instead of their strength. Crack deflection, crack bridging by fiber, and fiber pullout lead to an increased toughness and a noncatastrophic failure.

An important parameter in regard to the interface is the wettability of reinforcement by the matrix. *Wettability* refers to the ability of a liquid to spread on a solid substrate. Frequently, the contact angle between a liquid drop and a solid substrate is taken as a measure of wettability, a contact angle of 0° indicating perfect wettability and a contact angle of 180° indicating no wettability. *Wettability is only a measure of the possibility of attaining an intimate contact between a liquid and a solid*. Good wetting is a necessary, but not sufficient, condition for strong bonding. One needs a good wetting even for purely mechanical bonding or weak van der Waals bonding; otherwise voids may form at the interface. Besides wettability, other important factors, such as chemical, mechanical, thermal, and structural factors, affect the nature of the bonding between reinforcement and matrix. As it happens, these factors frequently overlap, and it may not always be possible to isolate their effects.

In PMCs, the surfaces of fibers are generally treated to promote chemical or mechanical adhesion with the matrix. For example, glass fiber, a common reinforcement for a variety of polymeric resins, invariably has a treated surface. The treatment is called *sizing*. The *size* is applied to protect glass fiber from the environment, for ease of handling, and to avoid introducing surface defects into the material. Common sizes are starch gum, hydrogenated vegetable oil, gelatin, and polyvinyl alcohol. These sizes are removed before putting in resin matrix by heat cleaning at approximately 340 °C and washing. After cleaning, organometallic or organosilane coupling agents are applied. Organosilane compounds have the chemical formula, R–SiX_3, where X

typically is Cl and *R* is a resin-compatible group capable of interacting with hydroxylated silanes on the glass surface.

A mechanical keying effect between two surfaces can also contribute to bonding. Chawla and Metzger observed such mechanical bonding effects in alumina/aluminum, a metal-matrix composite system.[3] Bonding due to mechanical interlocking at a rough interface can be equally important in PMCs and CMCs. Carbon fibers are given an oxidation treatment to provide, among other things, a rough surface that aids in bonding with the polymer matrix.

15.4.3 Interfacial Interactions

As mentioned earlier, most composite systems are nonequilibrium systems in the thermodynamic sense; that is, there exists a chemical potential gradient across the fiber–matrix interface. This means that, given favorable kinetic conditions (which, in practice, means a high enough temperature or long enough time), diffusion and/or chemical reactions will occur between the components. The interface layer(s) formed because of such a reaction will generally have characteristics different from those of either one of the components. At times, however, some controlled amount of reaction at the interface may be desirable for obtaining strong bonding between the fiber and the matrix, but too thick an interaction zone will adversely affect the properties of the composite. Metal and ceramic matrix composites are generally fabricated at high temperatures, because diffusion and chemical reaction kinetics are faster at elevated temperatures than at low temperatures.

A very important factor in regard to reinforcement–matrix compatibility has to do with the mismatch between the coefficient of thermal expansion of the reinforcement and that of the matrix. This thermal mismatch can lead to thermal stresses large enough to cause plastic deformation in a soft metallic matrix and cracking in a brittle ceramic or polymeric matrix. Plastic deformation in the metallic matrix leads to the introduction of defects such as dislocations, vacancies, etc., in the matrix, especially in the region near the interface. The introduction of such defects can and does affect the phenomena responsible for chemical reactions at the interface, as well as characteristics of the matrix such as the precipitation kinetics. Chawla and Metzger showed, in a definitive matter, the importance of thermal stresses in composites.[4] They used a large-diameter tungsten fiber, (225 μm)/copper single-crystal matrix and low-fiber volume fractions. An etch-pitting technique was employed to observe dislocations in the matrix. The researchers observed that the dislocation density near the fiber was much higher than the dislocation density far away from the fiber. The enhanced dislocation density in the copper matrix near the fiber arose because of the plastic deformation

[3] K. K. Chawla and M. Metzger, in *Advances in Research on Strength and Fracture of Materials 3* (New York: Pergamon Press, 1978), p. 1039.

[4] K. K. Chawla and M. Metzger, *J. Mater. Sci.*, 7 (1972) 34.

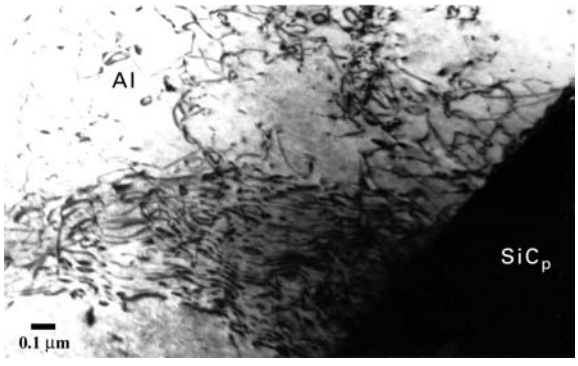


Fig. 15.4 TEM micrograph showing dislocations in aluminum in the region near a silicon carbide particle (SiC_p).

of the matrix, in response to the thermal stresses generated by the thermal mismatch between the fiber and the matrix. The existence of a plastically deformed zone containing a high dislocation density in the metallic matrix in the vicinity of the reinforcement was confirmed by transmission electron microscopy by a number of other researchers, both in fibrous and particulate metal matrix composites. Figure 15.4, a TEM micrograph, shows dislocations in an aluminum matrix near a particle of silicon carbide. Such high densities of defects in the metal matrix will, of course, lead to a different set of properties of the matrix. In nonprecipitation-hardening metals, this will simply cause a strengthening due to a higher dislocation density. In precipitation-hardenable matrix alloys, such as aluminum–copper, one would expect faster aging kinetics. Preferential precipitation at the reinforcement–matrix interface in an age-hardenable matrix has been observed by many researchers.

Example 15.2

Ultrahigh-molecular-weight polyethylene (UHMWPE) fibers such as the Spectra fiber are very hard to bond with most matrix materials. Why?

Solution: High-modulus polyethylene fibers have a very highly oriented and extended chain structure. This high degree of crystallinity and a lack of polar surface do not allow good bonding of the fibers to most matrix materials. Indeed, the surface of such fibers must be treated with plasma to provide roughness or else a copolymer should be used to provide a polar surface, both of which approaches allow for bonding with the matrix. A similar problem exists, to varying degrees, with aramid and carbon fibers as well.

15.5 Properties of Composites

We next describe some of the important properties of composites. In particular, we present expressions that allow us to predict the