9. TIME DEPENDENT BEHAVIOUR: CYCLIC FATIGUE

A machine part or structure will, if improperly designed and subjected to a repeated reversal or removal of an applied load, fail at a stress much lower than the ultimate strength of the material. This type of time-dependent failure is referred to as a cyclic fatigue failure. The failure is due primarily to repeated cyclic stress from a maximum to a minimum caused by a dynamic load. A familiar example of a fatigue failure is the final fracture of a piece of wire that is bent in one direction then the other for a number of cycles. This type of behavior is termed low-cycle fatigue and is associated with large stresses causing considerable plastic deformation with failure cycles, N_f, in the range of $<10^2$ to 10^4. The other basic type of fatigue failure is termed high-cycle fatigue and is characterized by loading which causes stress within the elastic range of the material and many thousands of cycles of stress reversals before failure occurs often with $N_f>10^5$ (sometimes $>10^2$ to 104).

Fatigue has been a major concern in engineering for over 100 years, and there is a very large amount of literature available on the fatigue problem. The importance of a knowledge of fatigue in engineering design is emphasized by one estimation that 90 percent of all service failures of machines are caused by fatigue and 90 percent of these fatigue failures result from improper design.

Fatigue failures of normally ductile materials in structural and machine members are very much different in appearance than failure under a static loading. Under quasi-static loading of the tensile test, considerable plastic flow of the metal precedes fracture and the fracture surface has a characteristic fibrous appearance. This fibrous appearance can also be noted in the ductile part of the fracture surface of Charpy impact specimens. A fatigue crack, however, appears entirely different. The crack begins at a surface, often at the point of high stress concentration. Once the crack begins, the crack itself forms an area of even higher stress concentration (also stress intensify factor), and it proceeds to propagate progressively with each application of load until the remaining stressed area finally becomes so small that it cannot support the load statically and a sudden fracture results. In fatigue failures, then, a characteristic appearance is always evident. The fatigue portion begins at the point of high-stress concentration and spreads outward showing concentric rings (known as beach marks) as it advances with repeated load. The final fracture surface has the same appearance as that of a ductile tensile specimen with a deep groove. The fracture is brittle due to constraint of the material surrounding the groove and has a crystalline appearance. The failure was not because the material crystallized as is sometimes supposed, it always was of crystalline structure.

Fatigue cracks, then, begin at a point of high-stress concentration. The severity of those stress concentrations will vary, even with carefully prepared laboratory specimens. In addition, the rapidity with which the crack propagates will vary. The cracks are irregular and will follow various paths around regions of stronger metal. A consequence of this manner of crack propagation is wide variation in time to failure of a number of seemingly identical test specimens loaded with the same load. For this reason a number of statistical procedures have been developed for interpreting fatigue data.

The basic mechanism of a high-cycle fatigue failure is that of a slowly spreading crack that extends with each cycle of applied stress. In order for a crack to propagate, the stress across it must be tension; a compression stress will simply close the crack and cause no damage. One way, then, of preventing fatigue or at least extending the fatigue life of parts is to reduce or eliminate the tensile stresses that occur during loading by creating a constant compressive surface stress, called a residual stress, in the outer layers of the specimen. We can picture how it is possible to induce a constant compressive stress of this type by a simple analogy. Consider a bar with a small slot cut in the surface. If we force a wedge tightly into this slot, the wedge and the material in the bar itself adjacent to the slot will be in a compressive stress state. If a tensile stress is now applied to the bar, no tensile stress can exist in the region of the slot until the compressive stress caused by the wedge is overcome. In other words, the tensile stress in the region of the slot will always be less than in a region removed from the slot by an amount equal to the magnitude of the residual compressive stress.

Favorable compressive stresses can be induced in parts by more practical methods than cutting slots. One of the most common methods is that of shot peening. Shot peening is a process in which the surface of the part is impacted by many small steel balls moving at high velocity. This process plastically deforms the surface of the part and actually tends to make it somewhat bigger than it was. The effect is the same as the wedge. A now larger surface is forced to exist on a smaller sublayer of material with the net result that a compressive stress is induced in the outer layer making it more resistant to fatigue failure.

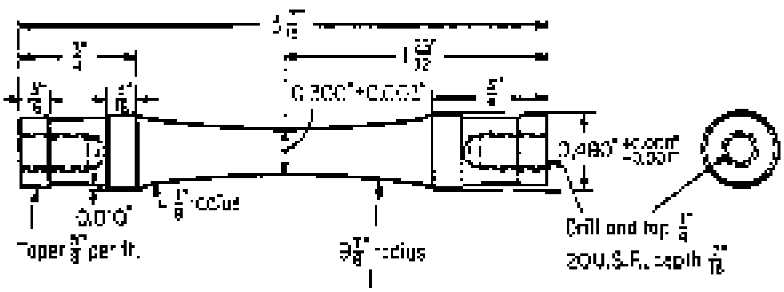

Figure 9.1 Typical fatigue test specimen

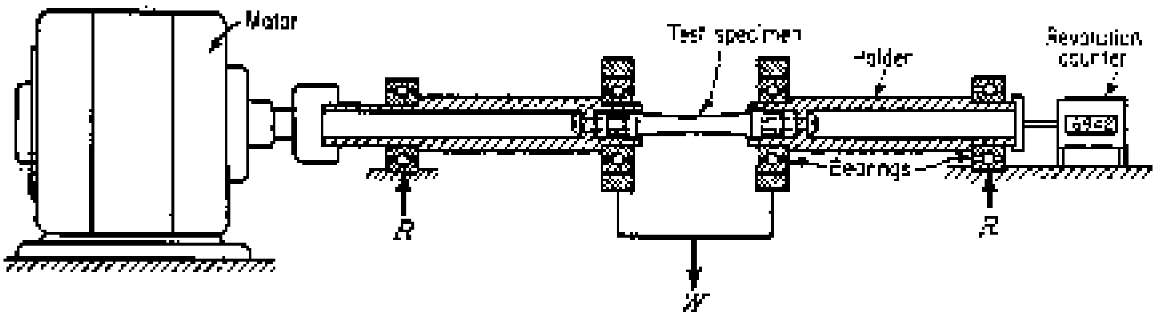

Figure 9.2 One type of rotating beam reversed stress testing machine (a.k.a. R.R. Moore rotating bending fatigue test machine)

A large portion of fatigue testing is done under conditions of sinusoidal loading in pure bending. The test specimen is shown in Fig. 9.1 and a schematic of the test machine (proposed by R. R. Moore) is shown in Fig. 9.2. This machine is called a rotating-beam type, and from Fig. 9.2 it can be seen that a constant moment of magnitude equal to one-half the applied load W multiplied by the distance d between the two bearings.

Since the specimen rotates while the constant moment is applied to the beam, the stress at any point in the beam makes a complete cycle from, say tension at a point on the bottom of the specimen to compression as the specimen rotates so the point comes to the top then back to compression as the rotation cycle to the bottom is completed. Thus, for each complete revolution of the specimen a point on the specimen experiences a complete stress cycle of tension and compression as shown on the left side of Fig. 9.3. The maximum and minimum values of this stress are equal and opposite at the specimen's surface, occurring at the peaks and valleys of the stress cycle, while the mean value of the stress is zero. This stress cycle is known as completely reversed loading and has a stress ratio, $R = \frac{\sigma_{min}}{\sigma_{max}}$, of -1. Other nomenclature and symbols for fatigue are shown in Fig. 9.4.

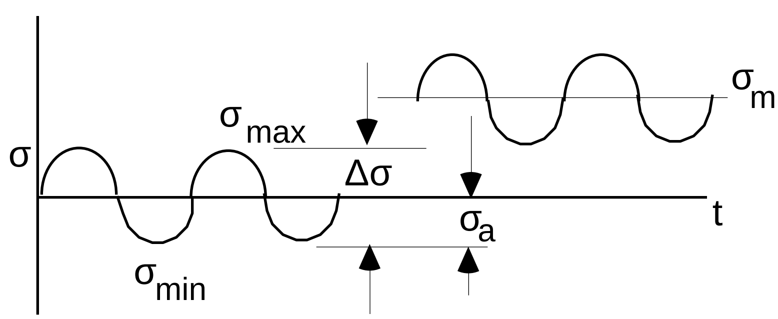

Figure 9.3 Examples of stress cycles for completely-reversed and tension-tension loading

σ_{max} = Maximum stress

σ_{min} = Minimum stress

$$\sigma_m = \text{Mean stress} = \frac{\sigma_{max} + \sigma_{min}}{2}$$

$$\Delta\sigma = \text{Stress range} = \sigma_{max} - \sigma_{min}$$

$$\sigma_a = \text{Stress amplitude} = \frac{\Delta\sigma}{2} = (\sigma_{max} - \sigma_m) = (\sigma_m - \sigma_{min})$$

Note: tension $= +\sigma$ and compression $= -\sigma$. Completely reversed $R = -1$, $\sigma_m = 0$.

$$R = \text{Stress ratio} = \frac{\sigma_{min}}{\sigma_{max}}$$

$$A = \text{Amplitude ratio} = \frac{\sigma_a}{\sigma_m} = \frac{1-R}{1+R}$$

Figure 9.4 Nomenclature and symbols for fatigue

Some test machines turn at 10,000 RPM. and are equipped with a counter that counts once for each 1000 revolutions. The machines are also equipped with an automatic cutout switch which immediately stops the motor when the specimen fails.

Fatigue test data are obtained in a number of ways. The most common is to test a number of different specimens by determining the time to failure for a specimen when stressed at a certain stress level. Figure 9.5 shows schematic representation of these test results on a type of graph called an S-N curve which shows the relationship between the number of cycles N for fracture and the maximum (or mean or amplitude or range) value of the applied cyclic stress. Generally, the abscissa is the logarithm of N, the number of cycles, while the vertical axis may be either the stress S or the logarithm of S. Typical test data are shown plotted in Fig. 9.6. in which the data can sometimes be represented by two straight lines.

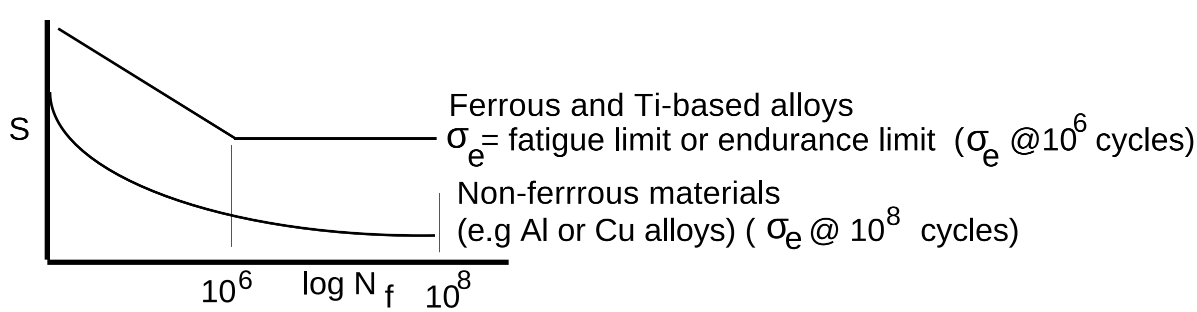

Figure 9.5 Schematic representation of S-N curves for ferrous and non ferrous materials.

The fatigue strength is then the stress required to fracture for a set number of cycles. The usual trend is for the fatigue strength to decrease rapidly in the cycle range of 10^3 to 10^6. In the range of 10^5 to 10^6 the curve flattens out and for some materials such as steel actually becomes flat, indicating this material will last "forever" in fatigue loading, if the applied stress is kept below a certain value. The stress corresponding to failure at an infinite number of cycles is called the endurance limit (or endurance limit) of the material and is often designated as S_e or σ_e. It should be recalled, in light of the previous discussion, that a considerable amount of scatter will exist and values such as endurance limit will be only approximate when a small number of test specimens are used.

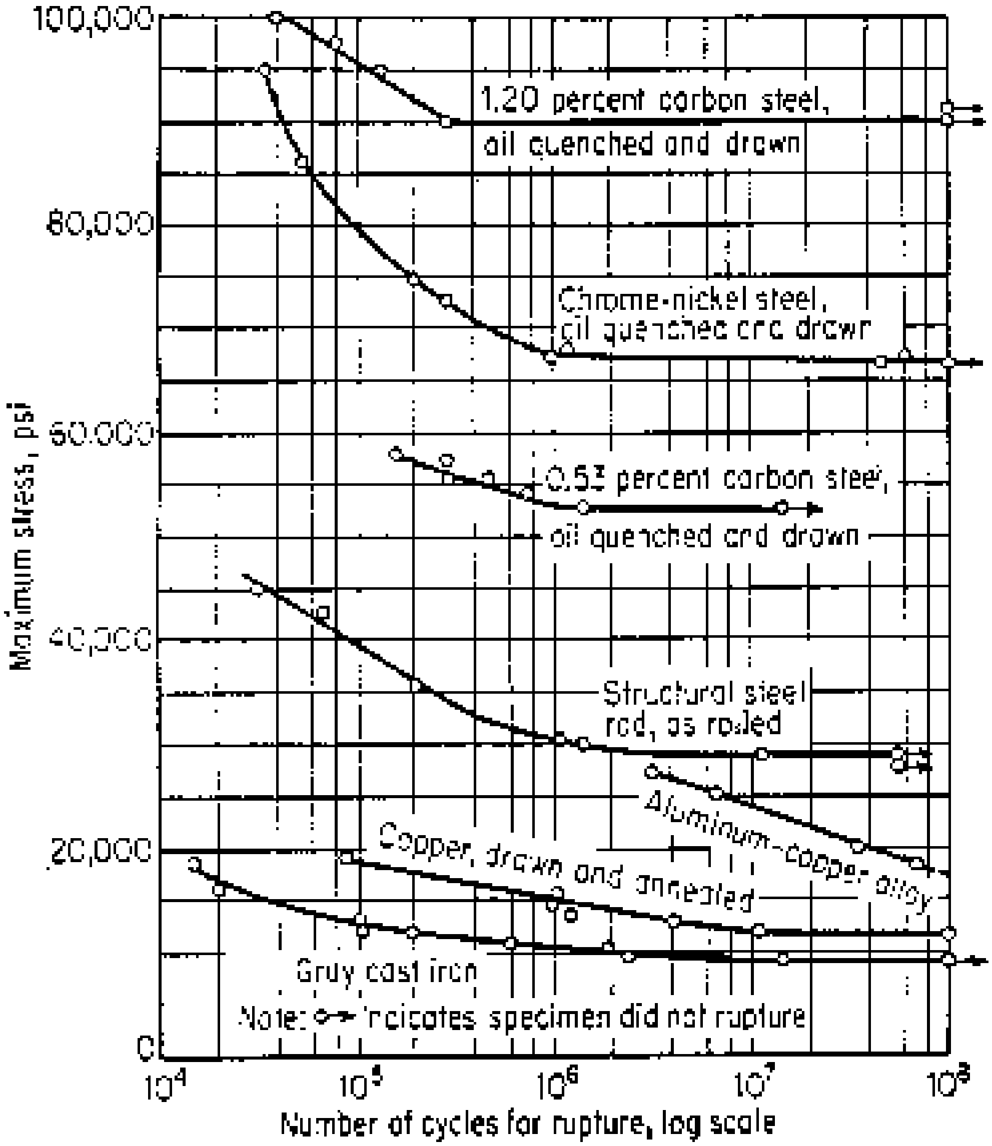

Figure 9.6 Typical S-N curves for determining endurance limits of selected materials under completely reversed bending.

Many times, in preliminary design work, it is necessary to approximate the S-N curve without actually running a fatigue test. For steel it has been found that a good approximation of the S-N curve can be drawn if the following rules are used.

1. Obtain the ultimate tensile strength S_{max} of the specimen (σ_{max} in a simple tension test).

2. On a diagram of S vs. log N, plot fatigue strength values of

   a. 0.9 S_{max} at 10^3 cycles
   b. 0.5 S_{max} at 10^6 cycles.

3. Join these points together to form a S-N diagram similar to that shown in Figs 9.5 or 9.6

Many factors affect fatigue failures. Not only do fatigue failures initiate at surfaces, but stress raisers create stresses which are greatest at surfaces. Fatigue factors
Recall the stress concentration factor:

$$k_t = \frac{\sigma_{LOCAL}}{\sigma_{REMOTE}} \tag{9.1}$$

where σ_{LOCAL} is the maximum local stress at the stress raiser and σ_{REMOTE} is the remote or net stress. The effect of the stress raiser on fatigue strength can be evaluated by first defining a fatigue strength reduction factor:

$$k_f = \frac{\sigma_e^{UN-NOTCHED}}{\sigma_e^{NOTCHED}} \tag{9.2}$$

where $\sigma_e^{UN-NOTCHED}$ and $\sigma_e^{NOTCHED}$ are the endurance limits for un-notched and notched fatigue specimens as illustrated in Fig. 9.7.

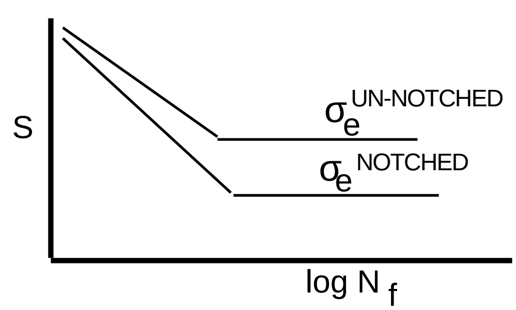

Figure 9.7 Illustration of endurance limits for notched and un-notched fatigue tests.

A notch sensitivity factor can be defined which relates the material behavior, k_f, and the component parameter, k_t such that

$$q = \frac{k_f - 1}{k_t - 1} \qquad (9.3)$$

where q=0 for no notch sensitivity, q=1 for full sensitivity. Note that q increases as notch radius, r, increases and q increases as S_{UTS} increases

Generally, $k_f << k_t$ for ductile materials and sharp notches but $k_f \approx k_t$ for brittle materials and blunt notches. This is due to i) steeper $d\sigma/dx$ for a sharp notch so that the average stress in the fatigue process zone is greater for the blunt notch, ii) volume effect of fatigue which is tied to average stress over a larger volume for blunt notch, iii) crack cannot propagate far from a sharp notch because steep stress gradient lowers K_I quickly. In design, avoid some types of notches, rough surfaces, and certain types of loading. As mentioned, compressive residual stresses at surfaces (from shot peening, surface rolling, etc.) can increase fatigue lives.

Endurance limit, $S_e = \sigma_e$, is also lowered by factors such as surface finish (m_a), type of loading (m_t), size of specimen (m_d), miscellaneous effects (m_o) such that:

$$\sigma_e' = m_a m_t m_d m_o \sigma_e \qquad (9.4)$$

Note that σ_e can be estimated from the ultimate tensile strength of the material such that: $\sigma_e \approx m_e S_{UTS}$ where m_e=0.4-0.6 for ferrous materials.

Note also that most fatigue data is generated for R=-1 and a mean stress of zero. Non zero mean stresses can also play a large part in resulting fatigue data. This is illustrated in Fig. 9.8 where the lines on the plot represent lines of infinite life. Note that when mean stress is σ_m=0, then the allowed amplitude stress, σ_a is the endurance limit measured for completely reversed fatigue test loading with R=-1. However, if σ_a=0 then the allowable mean stress is either the yield or ultimate strength from a monotonic test since the stress is no fluctuating when the stress amplitude is zero.

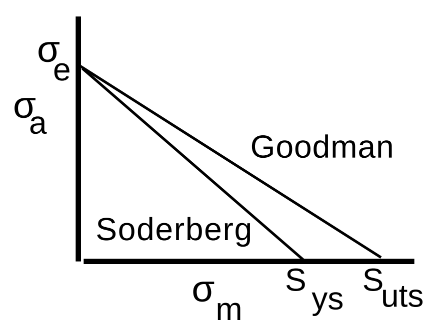

Figure 9.8 Illustration of effect of mean stress on allowable amplitude stress.

The mathematical expression for this effect is the equation of the line such that:

$$\sigma_a = \sigma_e \left(1 - \frac{\sigma_m}{S_{UTS}}\right) \qquad (9.5)$$

which is known as the Goodman line. If S_{UTS} is replaced with S_{ys}, then Eq. 9.5 is known as the Soderberg relation.

If a factor of safety, FS, and / or fatigue factors, k_f, are used then the following expressions results for either brittle or ductile materials.

Brittle
$$\sigma_a = \frac{\sigma_e}{FS \bullet k_f}\left(1 - \frac{\sigma_m}{(S_{UTS}/(k_f \approx k_t)FS)}\right) \qquad (9.6a)$$

Ductile
$$\sigma_a = \frac{\sigma_e}{FS \bullet k_f}\left(1 - \frac{\sigma_m}{(S_{UTS}/FS)}\right) \qquad (9.6b)$$

The previous understanding of fatigue behavior assumed constant amplitude or constant mean stress conditions and infinite life. Sometimes it is necessary to be able to understand the effect of variable amplitude about a constant mean stress as illustrated in Fig. 9.9 for non infinite life. Often a linear damage model (Palmgren-Miner rule) is applied such that:

$$\frac{N_1}{N_{f1}} + \frac{N_2}{N_{f2}} + \frac{N_3}{N_{f3}} = \sum \frac{N_j}{N_{fj}} = 1 \qquad (9.7)$$

where N_1 and N_{f1}, etc. are the actual number of cycles and number of S-N fatigue cycles at stress, σ_{a1} etc. (see Fig. 9.10) respectively.

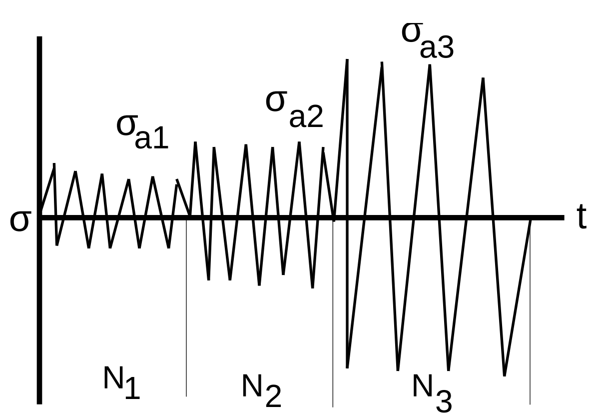

Figure 9.9 Variable amplitude loading with a constant mean stress

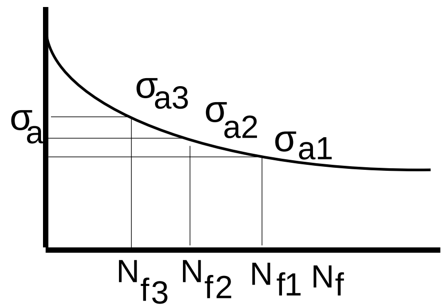

Figure 9.10 Illustration of cycles to failure on S-N curve for different stress amplitudes

S-N curves are generally used to design for infinite life in fatigue. That is, the fatigue data are used to chose stresses such that fatigue cracks will never develop. However, the fatigue mechanism is such that is possible to analyze the growth of fatigue cracks using linear elastic fracture mechanics. This is useful in extending the service life of components when a fatigue crack is noticed, rather than discarding a part which might still have useful life.

The fatigue process (see Fig. 9.11) consists of 1) crack initiation, 2) slip band crack growth (stage I crack propagation) 3) crack growth on planes of high tensile stress (stage II crack propagation) and 4) ultimate failure. Fatigue cracks initiate at free surfaces (external or internal) and initially consist of slip band extrusions and intrusions as illustrated in Figure 9.12. When the number of slip bands reaches a critical level (saturation) cracking occurs (see Figure 9.13). These slip band cracks will grow along directions of maximum shear for only one to two grain diameters (d_g), when the crack begins growing along a direction normal to maximum tensile stress. Although, fatigue striations (beach marks) on fracture surfaces represent successive crack extensions normal to tensile stresses where one mark is approximately equal to one cycle (i.e. 1 mark≈1N), beach marks only represent fatigue cycles during crack propagation and do not represent the number of cycles required to initiate the crack. Thus, summing the beach marks does not represent the total number of fatigue cycles.

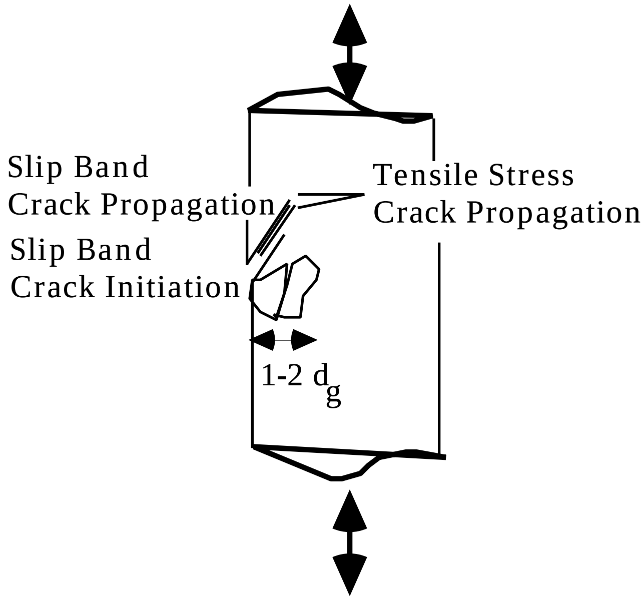

Figure 9.11 Illustration of fatigue process

During fatigue crack propagation (stage II may dominate as shown in Fig 9.14) such that crack growth analysis can be applied to design: a) cracks are inevitable, b) minimum detectable crack length can be used to predict total allowable cycles, c) periodic inspections can be scheduled to monitor and repair growing cracks, d) damage tolerant design can be applied to allow structural survival in presence of cracks.

The most important advance in understanding fatigue crack propagation was realizing the dependence of crack propagation on the stress intensity factor. The mathematical description of this is known as the Paris power law relation:

$$\frac{da}{dN} = C(\Delta K)^m \tag{9.8}$$

where da/dN is the crack propagation rate (see Fig. 9.15), $\Delta K = F(\Delta\sigma)\sqrt{\pi a}$ in which F is the geometry correction factor, a is the crack length, and C and n are material constants found in stage II of the da/dN vs ΔK curve (see Fig. 9.14)

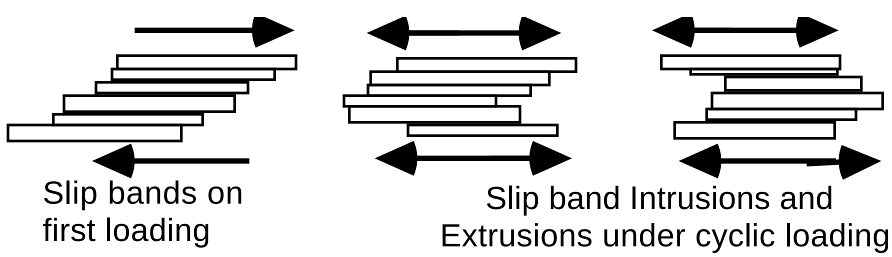

Figure 9.12 Woods model for fatigue

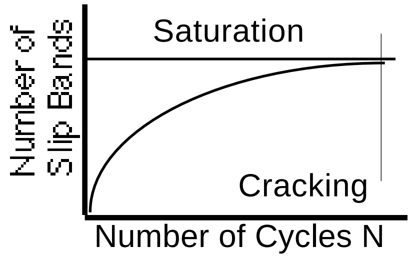

Figure 9.13 Slip band saturation and the onset of fatigue cracking

To predict the crack propagation life of a component, Eq. 9.8 can be rearranged such that:

$$\int_{N_i}^{N_f} dN = \int_{a_i}^{a_f} \frac{da}{C(\Delta K)^m} = \int_{a_i}^{a_f} \frac{da}{C(F\Delta\sigma\sqrt{\pi a})^m} \qquad (9.9)$$

Assuming that F can be approximated as nearly constant over the range of crack growth and assuming that m and C are constant, then:

$$N_f = \frac{a_f^{(1-(m/2))} - a_i^{(1-(m/2))}}{C\left[F(\Delta\sigma)\sqrt{\pi}\right]^m\left[1-(m/2)\right]} \qquad (9.10)$$

where a_i is the initial crack length which is either assumed (~2 d_g) or determined by non destructive evaluation and $a_f = \frac{1}{\pi}\left(\frac{K_{Ic}}{F\sigma_{max}}\right)^2$.

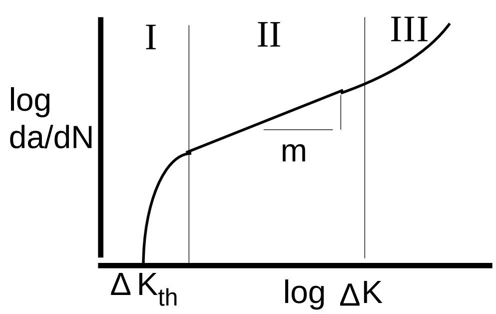

Figure 9.14 Crack growth rate versus stress intensity factor range.

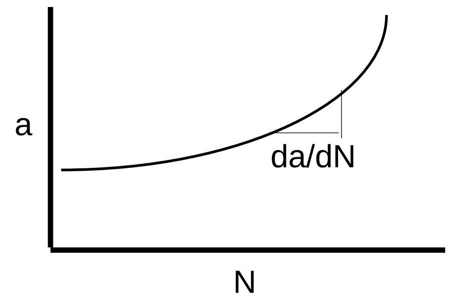

Figure 9.15 Crack length vs. number cycles relations from which crack propagation rate is found.

If F is a function of crack length, i.e. F(a,W, etc.), then numerical integration must be used such that:
.

$$\int_{N_i}^{N_f} dN = \int_{a_i}^{a_f} \frac{da}{C(\Delta K)^m} = \int_{a_i}^{a_f} \frac{da}{C\left[F(a,W,etc)\Delta\sigma\sqrt{\pi a}\right]^m} \qquad (9.11)$$

Crack propagation rates are also highly sensitive to R ratios primarily because crack propagation only occurs during tensile loading. Thus, the longer the crack stays open (i.e. R ≥0) the more time out of each cycle that crack propagation occurs. Thus, the more negative R, the more tolerant the crack is of ΔK and vice versa (see Fig. 9.16)

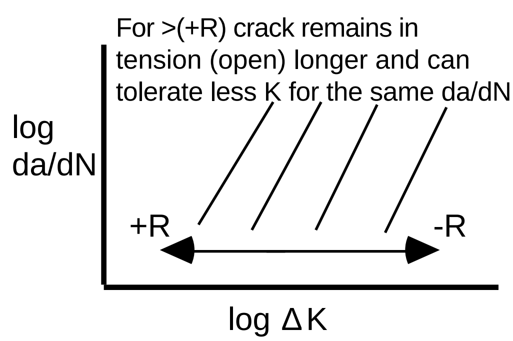

Figure 9.16 Effect of R-ratio on crack growth rates.