A SunCam online continuing education course

Mechanical Fatigue of Metals

by

Gerald O. Davis, P.E.

Mechanical Fatigue of Metals
A SunCam online continuing education course

1. Introduction - Overview of Mechanical Fatigue of Metals

Among the several ways that metallic engineering materials may fail in service, mechanical fatigue is typically the most common. The percentages of the total mechanical failures that are due to fatigue vary by different sources from 50% to as much as 80 to 90 %. Some examples of applications in which fatigue failures occur include rotating shafts, bearings, gears, the wings and pressurized fuselages of aircraft, pipelines and tubes in shell-and-tube heat exchangers. There are many other examples. The variable loads and changing stresses that occur in fatigue have unique effects on the services lives of engineering alloys that don't apply in situations where the only stresses acting are static.

In situations where loading and resulting stresses don't vary, resistance to failures becomes relatively straight forward. They can be adequately addressed by assuring working stresses, i.e., allowable stresses, are less than the yield, or ultimate tensile strength or ultimate shear strength (depending on the application) of the metal selected after applying a factor of safety to cover unknown aspects. Mechanical fatigue is significantly more complex. It is essential to realize that metal mechanical properties generated from standard, constant load tests should <u>never</u> be applied directly to fatigue applications. Fatigue often produces failure and complete metal separation when the maximum, fluctuating stress is well below the yield or ultimate tensile strength of the metal.

Fatigue failures occur in three stages. These are crack initiation (currently crack nucleation is the preferred term), growth in crack size during its progression through the metal and final fracture. The latter stage is known as fast fracture because it occurs almost instantaneously when there is too little metal remaining to support the allied load. Tensile stresses nucleate and grow fatigue cracks. Compressive stresses resist nucleation and growth.

Fatigue analysis and procedures during design can be difficult for multiple reasons. First there are two types of fatigue – high and low cycle. In high-cycle (HC) fatigue there are many cycles of stress but the magnitude of variable stress is relatively low. As you would expect, in the HC case the metal used can sustain a large number of stress cycles (at least 10^6 and often many more) without failure. The maximum stress in HC fatigue can be less than the yield strength of the metal so that the amount of plastic (permanent) material deformation is zero or very small. For many metals this will permit a service life that is practically infinite. In contrast, low cycle (LC) fatigue (also known as strain-controlled fatigue) involves few stress cycles (generally less than 10^4 and often many less) and a short life before failure. This is because the level of variable stress is much higher in LC fatigue. In low-cycle fatigue the stress that produces failure can be

Mechanical Fatigue of Metals
A SunCam online continuing education course

close to the ultimate tensile strength of the given alloy and typically significant plastic deformation occurs.

In both HC and LC fatigue there are many potential considerations in conducting valid laboratory tests to generate data. Fatigue analysis is further complicated by the fact that generally there are three different approaches to designing to avoid failures. These are the S-N method, the strain-N method and a third method that uses linear elastic fracture mechanics (LEFM). S is a measure of the variable stress parameter acting at a given number of stress cycles, N. In both the S-N and strain-N methods the goal is to prevent the nucleation of a macro-size fatigue crack. Fracture mechanics assumes a detectable size crack pre-exists and monitoring of crack growth is used to prevent failure. Finally, as will be discussed, there are several factors that must be considered using any method.

The starting point for the S-N method is laboratory-generated plots of S versus N data. These plots define the fatigue limit (also known as the endurance limit) <u>and</u> the fatigue strength of different alloys tested. Fatigue limit and fatigue strength are different parameters that will be defined later. Laboratory-generated fatigue data cannot be used directly in design or analysis. This is because lab results are based on test conditions that most often are very different from actual service conditions. After making all necessary adjustments to laboratory-generated data and other variables, the engineer has a basis in theory that his/her design will be successful. In particular in the S-N method the given alloy's working stress must less than the final adjusted fatigue limit or fatigue strength of the alloy. However, additional testing of a prototype using the derived design is required for confirmation. This last step is very important. The S-N method is typically applied to high-cycle (HC) fatigue. ASTM Standard E466 provides information on laboratory test procedures.

The strain-N method is based on laboratory application of different levels of applied cyclic strain amplitudes (or cyclic strain ranges) until test specimens fail at different numbers of cycles. Resulting levels of strain of test specimens are recorded simultaneously during tests. Each recorded strain level includes both elastic and plastic deformation in the total strain. Deriving and using strain-based data is more complex than the S-N method. Strain-based data are particularly useful for low-cycle (LC) fatigue applications and when notches (stress concentration features) are simulated in lab specimens. As in the S-N method, confirmation testing is required. ASTM Standard E606 applies to the overall and general use of the strain-N method.

Both the S-N and strain-N methods are based on the assumption that an alloy's practical service life is over when a detectable size fatigue crack first appears. These methods are conservative.

Mechanical Fatigue of Metals
A SunCam online continuing education course

That is because, after nucleation, a fatigue crack has to grow to some critical size before final fast fracture and separation occurs. However, even though fatigue life predictions by the S-N or strain-N methods are very conservative they are still useful. The time required before fatigue cracks nucleate can be significant in HC fatigue with its low stress levels. Both the HC and LC methods can be used for making initial approximations for design and by alloy manufactures in the development of new materials.

The linear elastic fracture mechanics (LEFM) method takes an entirely different approach versus the other two methods. It is known as a damage-tolerant procedure. In this method it is assumed that a fatigue crack (of some detectable size) already exists. Certain essential parameters are necessary to access each application. These are the stress intensity factor (K) that drives crack growth and a critical property just before final fracture occurs, i.e., the alloy's fracture toughness, K_c. Use of these plus other parameters allows a calculation to be made of the largest crack size that can generated without final fracture. Calculation of the estimated rate of crack growth is a vital part of using the LEFM method. During the fatigue crack growth phase, before critical crack size is attained, the component is assumed to be safe. Extensive past development work has confirmed this assumption in many applications.

The LEFM procedure always includes use of non-destructive evaluation (NDE) techniques to inspect the metal and record the size of the growing crack versus time. Obviously the determination of critical crack size, crack growth rate and reliable NDE inspections are essential. The affected part or component is replaced or other corrective are taken before a crack reaches its critical size. The LEFM approach has been used successfully for several years in the aircraft, nuclear power and other industries. The ASTM Standards E647 and E399 apply to LEFM.

Mechanical fatigue is a broad and often complex subject. Attempting to cover all of the varied areas of fatigue and design or analysis of it is too ambitious for this short course. Therefore discussion will primarily be limited to the traditional S-N method for steels. This will include its application to high-cycle (HC) fatigue analysis and design for failure prevention.

The overall goal for this course is to provide an overview of the mechanical and, secondarily, the metallurgical factors that influence fatigue. The specific topics included are as follows**:** appearance and characteristics of fracture surfaces after fatigue failures, typical laboratory tests employed to generate S-N data and precautions in their use, the basic mechanism of fatigue crack nucleation and growth, and the several factors that influence the incidence of fatigue. Suggested steps for design to prevent HC fatigue failure using the S-N method are presented. Next a list of

fatigue characterization and control factors is provided. Finally, selected references that cover much more of the broad subject of fatigue than is possible here are provided.

2. Appearance and Characteristics of Fracture Surfaces after Fatigue Failure

Figure 1 illustrates schematically the typical <u>macro-scale</u> appearance of the fractured surface of a steel rotating shaft that in (a) has failed by low-cycle fatigue and in (b) has failed by high-cycle fatigue. The same alloy is assumed in both cases. As shown here the fatigue cracks grew from the same point on each shaft's surface in the lower right corner. The generally concentric circles (or portions of ellipses) that form from the lower right fatigue nucleation site towards the upper left are called <u>beach marks</u> (and also – arrest marks- among other names).

Cycles of fatigue may not be continuous or have the same stress level. Beach marks indicate locations at which the fatigue crack growth stopped for a short or extended period or when the magnitude of the stress changed significantly. These marks do not form when there is continuous application of the same cyclic stress and crack growth. When present and visible, they are definitive macro-scale indicators of a moving fatigue crack. The crack growth region (B or B') in each case is generally smooth while the final fracture region (C or C') is rough and irregular.
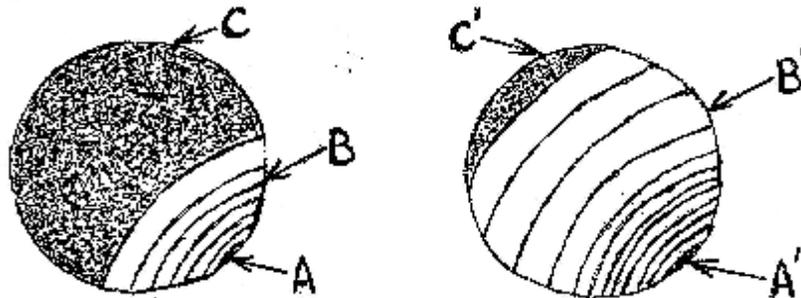
Figure 1 – Schematic representations of fatigue failure fracture surface appearances of a steel rotating shaft exposed to LC <u>compared</u> to HC fatigue. B & B' areas shown represent crack growth after crack nucleation. Locations of surface crack nucleation sites, A & A', are the same in both cases. Each surface has distinctive beach marks shown.

(a) Low-Cycle Fatigue (LC)   (b) High-Cycle Fatigue (HC)
Continuing from Figure 1:
A - Fatigue crack nucleation site   A'- Fatigue crack nucleation site

Mechanical Fatigue of Metals
A SunCam online continuing education course

B - Crack growth area – Small; Smooth* B'- Crack growth area – Large; Smooth*
C - Final fracture area – Large, Rough   C'- Final fracture area – Small; Rough

*If the failure occurred in a corrosive environment, the crack growth area may contain irregular, non-smooth corrosion product deposits visible on a macro-scale. In that case beach marks often are not visible. The partially separated cross-section permits corrosion as the crack advances.

The significant differences associated with low-cycle (LC) versus high-cycle (HC) results are the applied level of stress (S), the fracture toughness (K_c) of the given metal and the resulting number of cyclic stress cycles (N) before final fracture. In general, higher stress and lower fracture toughness lead to the LC case. In LC there are fewer fatigue crack growth cycles. These factors in LC fatigue produce a large final fracture area. By contrast, HC fatigue involves a relatively low level of cyclic stress but higher fracture toughness. After crack nucleation, HC fatigue produces a higher number of fatigue crack growth cycles before final fracture and a smaller final fracture region.

It should be understood that Figure 1 provides "typical" effects and results. Other service factors can make it more difficult to distinguish between LC and HC fatigue.

On a macro-scale both LC and HC fatigue fracture surfaces typically do not appear to have significant plastic, i.e., permanent, deformation. The exception would be for higher ductility metals that may display small plastic ridges projecting from the fracture surface at the edge of the final fracture region. These features are known as shear lips. A less ductile metal loaded in fatigue normally will not display any plastic deformation on a macro-scale. However, both LC and HC fatigue cracks grow (or propagate) through the metal by a stop-and-go plastic deformation process. This plastic deformation occurs only at the tips of an advancing crack and on a submicroscopic-scale.

When present, another indicator of failure by fatigue on the surface is the presence of striations. Unlike beach marks, striations can only be observed on a submicroscopic-scale, e.g., by use of a scanning electron microscope (SEM). Striations are many times smaller than beach marks. Each striation represents one cycle of variable fatigue stress and appears in a SEM image as a small ridge. It is not always possible to observe striations. In general they are more easily seen in higher strength aluminum alloys than in steels. Striations are often absent in both very soft and very hard metals.

Both beach marks and striations are not always visible for a variety of other reasons. Features on fracture surfaces may be obscured or destroyed by rubbing together of mating parts before final fracture separates them, by corrosion during or after the failure process, by damage during transport to the failure analysis lab or by initial "fitting together" of the two formerly mated parts by uninformed persons. When observable, beach marks and striations provide definitive evidence of fatigue failure in a metal.

Another useful type of macro-scale indicator of fatigue failure is ratchet marks.
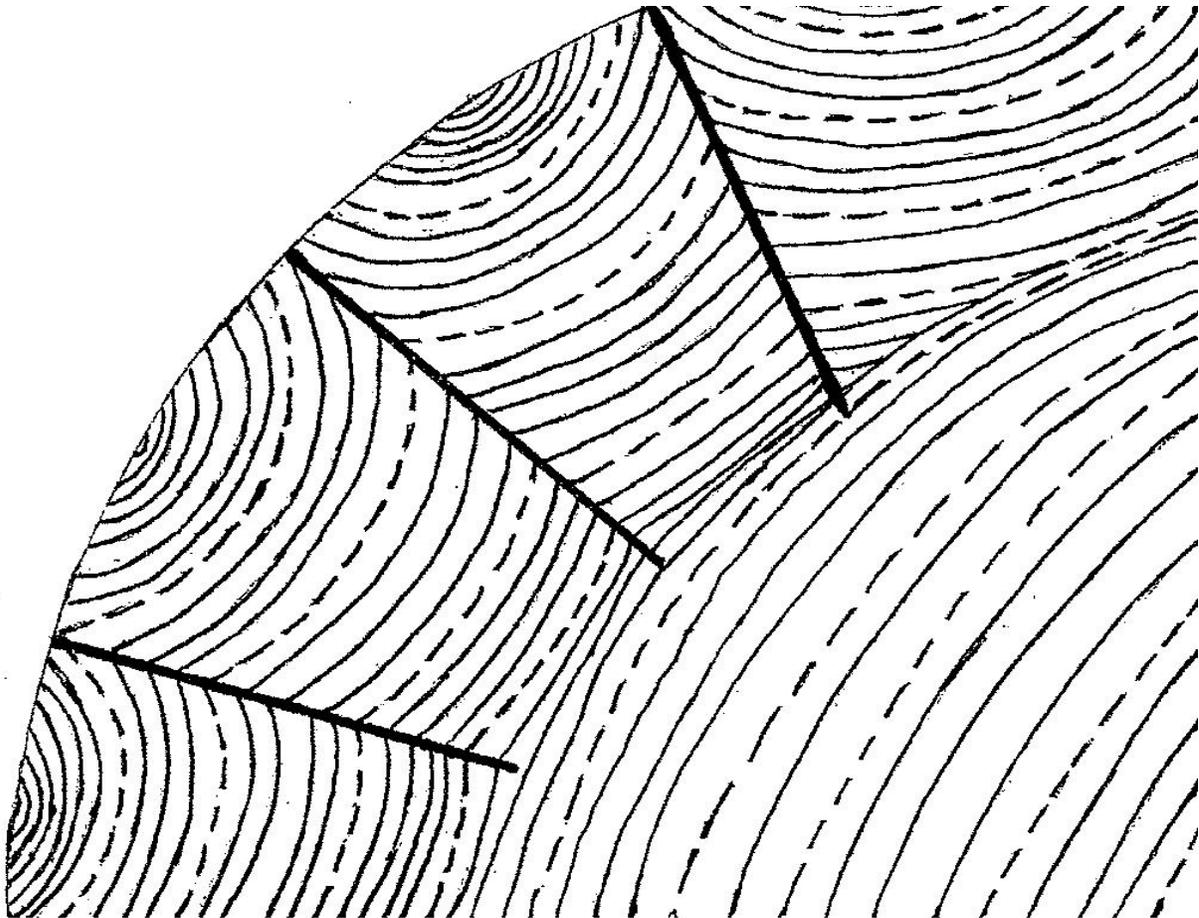
Figure 2 – Greatly enlarged and simplified schematic illustration of ratchet marks, beach marks and striations on a portion of a fatigue fracture surface of a circular shaft as viewed in a SEM. (After Reference 1, page 130)

Ratchet marks are raised ridges of metal perpendicular to the fracture surface. They are located so that they separate multiple crack origin sites and paths of growing cracks when those origins

Mechanical Fatigue of Metals
A SunCam online continuing education course

are in the same radial plane through the metal's outer surface. The outer surface may be straight (as on a flat metal plate) or curved (as on a circular shaft). Overlapping cracks from two adjacent fatigue origins generate one raised ridge that is one ratchet mark. This fact can generally be used to determine how many fatigue origins were present on a metal's outer surface.

Notice in Figure 2 that the ratchet marks (thick solid lines) are radial from the outer shaft surface towards the center of the shaft. In the figure it is seen that the multiple adjacent fatigue cracks eventually join together to form one large crack. This one crack then grows further through the metal thickness. Also note that the thin solid lines (striations) are located between bordering beach marks (thin dashed lines). In reality, because they are so small by comparison to beach marks, there would be hundreds or many more striations between two bordering beach marks.

3. Laboratory-Generated Fatigue Data

Many engineers are familiar with typical fatigue data expressed as the S-N solid- line plot shown below as Figure 3. The vertical axis indicates variable stress (S), typically the <u>stress amplitude</u> or the <u>stress range</u> as defined below, on a linear scale while the horizontal axis indicates the corresponding number of variable stress cycles (N) to failure on a logarithmic scale. The horizontal section of the solid-line plot indicates the <u>fatigue limit</u> (also known as the endurance limit) of the specific metal. This shape is typical of many ferrous-based alloys, i.e., iron-based, especially low alloy and moderate strength steels. Titanium alloys also display fatigue limits. The fatigue limit portions of the S-N curves of ferrous-based alloys receive much attention.

In a given application, if the maximum allowable alloy stress, i.e., working stress, is less than the cyclic fatigue limit stress (as cyclic amplitude or cyclic range) <u>after final adjustment</u>, then it is assumed that theoretically the number of cycles, N, to failure will be essentially infinite. It is essential that adjustments to lab-generated S-N fatigue limit data and other variables must be made before the lab values are used. Procedures to accomplish this are described later.

The dashed-line plot in Figure 3 indicates data for many metals including non-ferrous based metals such as aluminum alloys and copper-based alloys. That curve shows that for these alloys there is no horizontal section (fatigue limit) as defined above. Instead such alloys have <u>fatigue strength</u> values over their range of N cycles. Fatigue strength at approximately 10^8 cycles is often taken as the "substitute" for true fatigue limits for these alloys. As will be discussed later, the dashed-line data also schematically represents the joint action of a corrosive environment and fatigue loading on any alloy without corrosion resistance to the given environment.

The design engineer must confirm for any alloy selected that the working stress at the desired number of fatigue cycles is below the *final adjusted*, S-N data line. It is never correct to use material properties as defined in static stress tests, e.g., tensile tests, in designing for fatigue.

% of Applied Cyclic Stress (Stress Amplitude or Stress Range)
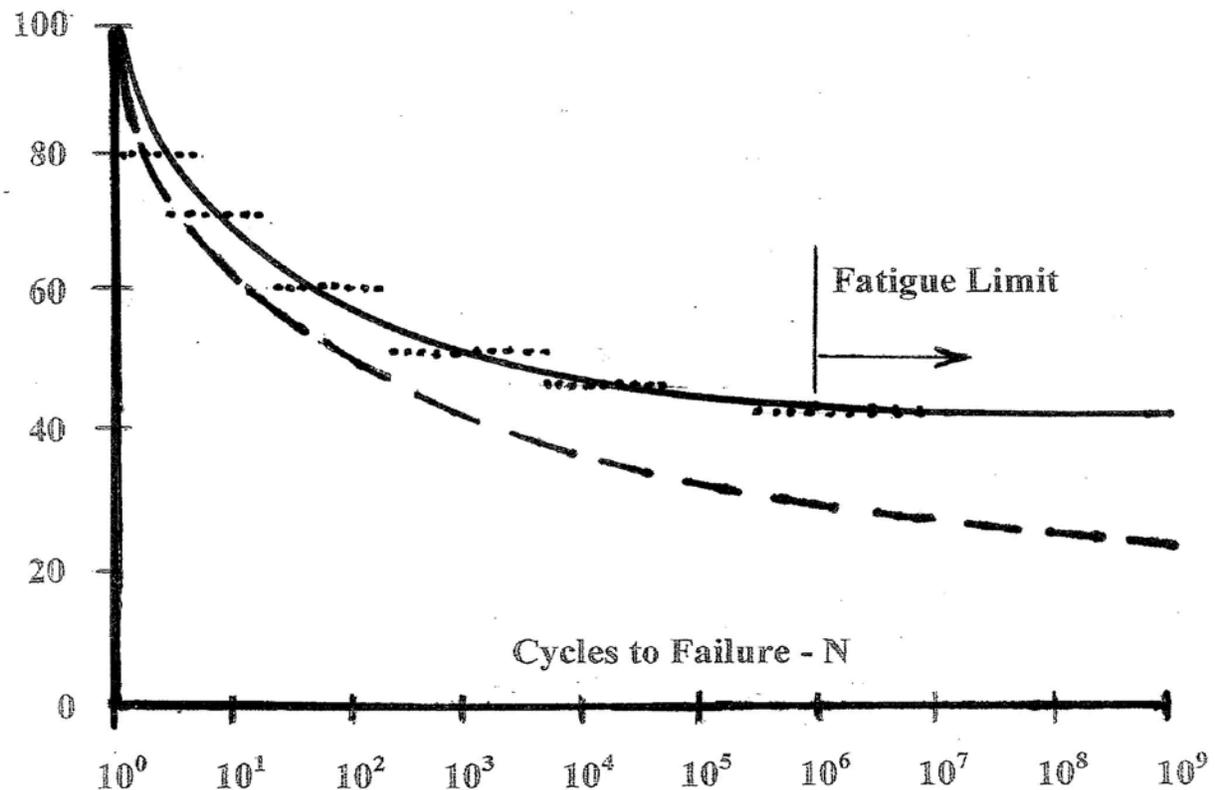
Figure 3 – Schematic representation of laboratory-generated S-N fatigue data. The solid-line curve applies to many low and moderate strength ferrous alloys plus titanium alloys. The dashed-line curve applies to non-ferrous alloys. *Note the scatter of data points with multiple test specimens at each given stress level. The test specimens are assumed to be identical.*

Data scatter is illustrated schematically here. Multiple data points are shown on the solid-line plot at various stress levels. In actual lab testing there would be more tests at a given stress level and more stress levels used than those shown in Figure 3. Similar scatter would apply to tests of non-ferrous alloys but that is omitted from the dashed line plot shown.

Data scatter occurs even though each test specimen is the same material and presumed to be the same size and prepared identically. In reality specimens differ. Each test specimen is somewhat

Mechanical Fatigue of Metals
A SunCam online continuing education course

different due to very small differences in dimensions, surface finish and especially due to internal metallurgical differences. These factors may appear to be insignificant but scatter in fatigue test data is always present.

The situation is addressed with statistical methods. However, it is impractical to test enough specimens at each stress level and to complete enough tests at different stress levels to gain high statistical confidence in results. Data curve fitting is used to generate plots that best fit the majority of the data points generated during limited testing. This creates finished S-N lines that pass through only a portion the data points. Stress levels that correspond to resulting fatigue limit lines have about a 50% reliability of being accurate. S-N lab plots are necessarily inexact. They should be treated as such.

Most S-N fatigue data are generated using small, polished cylindrical test specimens loaded in four-point loading in fully reversed, constant amplitude bending so that no shear stress is created. Each test specimen is rotated by a test machine <u>about a central zero-stress axis</u>, up to a given maximum tensile stress and down to an equal maximum compressive stress during each rotation. The number of cycles N before failure, i.e., complete separation, for each test specimen is recorded at each stress level for the given applied load. Data generated this way is the type most commonly available to the design engineer.

Figure 4 illustrates four-point loading, reversed bending fatigue test cycles with constant amplitude but with an important difference from the procedure just described. There the stresses do not cycle above and below a zero-stress level. Instead the applied cyclic stresses move above and below a constant level know as the mean (or average) stress, S_{MEAN}.

Mean stresses act continually and simultaneously with cyclic stresses. Their net effect is illustrated by the dashed line in Figure 4.

Mean stresses, S_{MEAN}, can be applied during lab testing but this is not commonly done. This is important because in actual service, mean stresses are very common. In service, net mean stresses include the algebraic sun of tensile (+) and compressive (-) stresses in the given application.

In Figure 4 the stress amplitude S_A, the stress range ΔS and the net mean stress S_{MEAN} values are each in the tensile stress range. This represents a severe fatigue loading effect. According to the test parameters shown in the figure and their definitions, the resulting stress ratio, **R** , would then be a positive fractional value between zero and one, for example., **+3/4.**

Mechanical Fatigue of Metals
A SunCam online continuing education course

S_A = Stress Amplitude = (S_{MAX} - S_{MIN}) / 2

ΔS = Stress Range = (S_{MAX} - S_{MIN})

S_{MEAN} = Mean Stress = (S_{MAX} + S_{MIN}) / 2

S_{MAX} = Maximum Stress = (S_{MEAN} + S_A)
S_{MIN} = Minimum Stress = (S_{MEAN} - S_A)

R = Stress Ratio = (S_{MIN} / S_{MAX})

A = Amplitude Ratio = (S_A / S_{MEAN})

Tensile and compressive stresses should be added algebraically in all definitions.
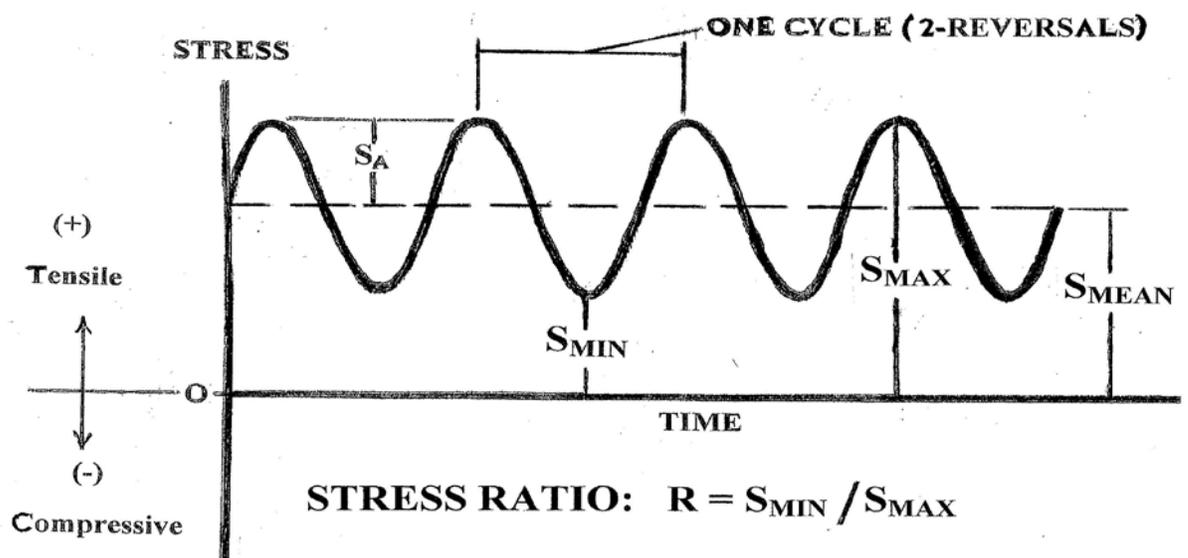

Figure 4 – Representation of constant amplitude fatigue stress cycles generated during a four-point, reversed bending test <u>with</u> a net tensile mean stress (S_{MEAN}) being applied. (After Reference 2, page 61)

Lab testing with a S_{MEAN} applied can be done but this is not the most common condition. Typically constant amplitude testing is done with S_{MEAN} = Zero; S_{MAX} = 1; S_{MIN} = -1 and, therefore, R = -1.

Mechanical Fatigue of Metals
A SunCam online continuing education course

The values of the <u>stress amplitude</u>, $S_A = (S_{MAX} - S_{MIN}) / 2$, and <u>stress range</u>, $\Delta S = (S_{MAX} - S_{MIN})$, are the most important indicators of the severity of fatigue loading for smooth test specimens used in lab tests. When both of these values are in the tensile range, test coupon fatigue life is inversely proportional to the value of S_A or ΔS. The magnitude of this effect can be different in service because of the presence of S_{MEAN} and notches.

Compressive stress always retards fatigue. Therefore if S_{MIN} is in the compressive region while there is a required minimum threshold value of S_{MAX} in the tensile region, fatigue can occur but fatigue life is increased versus the case of both stresses being tensile. If both maximum and minimum cyclic stresses are in the compressive stress region no fatigue failure can occur.

S_{MEAN} in actual service can be an externally applied constant stress <u>or</u> a residual internal constant stress left in the metal by a manufacturing process such as heat treatment, rolling or welding. Some of the practical effects on fatigue data generated in the lab without S_{MEAN} or notches versus actual service conditions where both net mean stress and notches likely exist are discussed in Sections 6 and 7.

S-N fatigue data are generally available only from four-point fully reversed bending tests, without shear stress, with constant amplitude cycles and S_{MEAN} equal to zero, i.e., $R = -1$. However, by the use of different types of testing equipment cyclic axial stress <u>or</u> cyclic cantilevered rotating bending stress <u>or</u> a combination of torsion and bending cyclic stress can be generated. Different values of one of these types of stress can then be plotted on the vertical axis of an S-N diagram versus the number of cycles to failure on the horizontal axis. Usually such data are not readily available to the design engineer.

As typically conducted, lab fatigue testing is useful for some purposes but it has multiple limitations. It does allow relative comparisons of the fatigue lives of potential new alloys with different compositions and heat treatments and thus different strengths. Alloy manufacturers can use these comparisons as initial steps in developing new materials. Engineers can use lab data in a first general step in design or analysis to compare the fatigue behavior of existing, alternative alloys. However, commonly available lab fatigue data is deficient because testing conditions often differ from service conditions in the following ways:

- Mean stresses, S_{MEAN}, occur frequently in actual applications but these are most often omitted in common lab tests.
- Test specimens are highly polished while real parts may have a variety of surface irregularities such as machining marks or scratches. Actual parts also may have <u>notches</u>, e.g.,

Mechanical Fatigue of Metals
A SunCam online continuing education course

small radii, drilled holes, machined grooves and other macro-scale features, that are stress concentration points. Fatigue cracks most often initiate <u>on the surface</u> of the actual part and polished surfaces without notches don't duplicate reality.

- Generally test specimens are small in size compared to actual parts. The stress distribution or gradient in small fatigue loaded test coupons is often different from much larger service parts. Larger actual parts also have internal metallurgical defects. The result is that fatigue limits for in-service components are less than found for lab coupons.

- There is no shear stress generated in the four-point fully reserved bending method that is most often used in tests. The actual part may be subjected to axial, bending and shear stresses or a combination of these.

- Both fatigue crack nucleation and growth are accelerated by raised or very low temperatures and corrosive conditions. Most available fatigue data is generated at room temperature and in clean, dry air where corrosion is minimal.

- The data scatter that always occurs in fatigue testing is addressed by a best-fit curve through the data points at each applied cyclic stress level. Typically this results in plotted S-N curves that are only 50% reliable at the different stress levels. Many applications require more than this level of reliability.

- Most lab tests are run with constant amplitude stress cycles as in Figure 4. In reality many fatigue applications involve irregular, variable amplitude stress cycles.

Many of these issues can be overcome by making changes to common fatigue testing procedures. However, the added complexity and costs of necessary modifications mean the changes are not frequently accomplished. Thus more desirable data are rare and not usually available **Therefore, to the extent possible, it is essential that commonly available lab data be adjusted to account for the realistic effects found in each given application.** As shown later, adjustment procedures add difficulty to fatigue design and analysis but they are needed to obtain meaningful results.

4. Overview of the Fundamental Mechanism of Fatigue

As previously stated, the process of mechanical fatigue has three stages: crack nucleation, crack growth and final (fast) fracture. The last stage is not fatigue. It is an almost immediate, complete separation of the damaged metal due to the action of the last cycle of fatigue stress. At that point there is too little intact metal remaining to resist the applied load.

The process of crack nucleation most often begins inside the crystalline structure of metals. A crystal structure means that the atoms in each crystal (also called known as a grain) are arranged in a specific order relative to adjacent atoms in that same grain. However, when external stress causes even a

Mechanical Fatigue of Metals
A SunCam online continuing education course

small amount of inelastic, i.e., permanent or plastic, deformation of the metal, a small rearrangement or relocation of atoms occurs. Thus imperfections are formed in the micro-scale structure of affected grains. These imperfections are <u>dislocations</u>. Because of reoccurring applied stress cycles, the dislocations move on specific paths and directions depending on the type of metal and its particular crystalline structure. The movement is known as slip. Some dislocations move so that eventually microscopic projections and concave features are formed on the metal's outer surface.

During static loading and thus constant stress that causes some extent of plastic deformation, a few dislocations can emerge on a metal's surface. However, in fatigue loading with many stress cycles even with lower stresses than in the static stress case, many dislocations slip and emerge on the metal surface. The projections and concave surface features created by fatigue are called extensions and intrusions, respectively. They are much more numerous and somewhat larger than these features created by a one-time application of stress.

Extensions and intrusions are generally microscopic in size but they create areas of increased local concentrated stress so that very small fatigue cracks can nucleate. If the extensions and intrusions are nearby macro-scale geometric surface features, e.g., drilled holes, machined grooves, small radii on a stepped shaft, corrosion pits or even a rough or scratched finish, stress concentration effects are magnified. These larger outer surface features as a group are known as <u>notches</u>.

The level of stress produced by an applied force is most often highest (either tension or compression) on a metal's outer surface compared to the stress level elsewhere in its cross-section. This is true for loading due to bending, direct shear, torsional shear or a combination of these. Only in pure uniaxial loading is the resulting stress uniform across the metal cross-section. The combination of higher surface stresses, numerous surface extrusions and intrusions plus often unavoidable notches make the nucleation of fatigue cracks most probable on a metal's surface.

In high-cycle fatigue (HC) stress levels are relatively low. Therefore in the majority of HC applications the bigger portion of a metal's total fatigue life is consumed with nucleating the first micro-size cracks. A smaller period of the total fatigue life involves growth of macro-size cracks to the final (fast) fracture stage. In low-cycle fatigue (LC) the higher level of stress results in a very short period required to nucleate initial small cracks and a somewhat longer period for macro-size crack growth. However, the total life of a component subjected to LC fatigue is much shorter than that same component subjected to HC fatigue. This is because in the HC case the final fast fracture period (almost instantaneous) occurs over a large percentage of the metal cross-section.

Mechanical Fatigue of Metals
A SunCam online continuing education course

While most fatigue cracks nucleate directly on the surface of the metal that is not always true. Dislocation slip during fatigue loading can be slowed or stopped by internal or subsurface features of the metal. These include internal metal defects such as inclusions, pores or voids in welds or other cast metal, at the boundaries between grains OR at micro-size metal cracks that pre-existed before the fatigue loading started. When moving dislocations encounter these features, as multiple cycles of fatigue stress continue, small fatigue cracks may nucleate there just as they do on the outer surface. Then the second stage of fatigue, i.e. crack growth, can begin. Often these internal defects that act as crack nucleation sites are just a small distance below a metal's outer surface but they can also be deeper.

Eventually several micro-size cracks may nucleate at a small area on or nearby the fatigue loaded metal surface at about the same time. In the first portion of the crack growth stage these small cracks join together to form one or more larger cracks that become visible without magnification. Initially the micro-size crack growth follows a direction that is most favorable for movement by shear stress. That direction is 45^0 to the direction of the principal applied stress. After a brief period the direction of crack growth changes to its final direction that is most favorable to crack movement by tensile stress. That is 90^0 to the direction of the principal applied stress. This continues as the larger crack (or cracks) moves through the thickness of the metal.

As previously mentioned, high cycle fatigue often produces little or no macro-scale plastic deformation of the metal. The exception occurs if the metal has high ductility. By comparison, the higher stresses in low cycle fatigue often do produce macro-scale plastic deformation. However, on a microscopic scale plastic deformation occurs in both HC and LC fatigue as a growing crack is extended. Very localized high tensile stresses are created at the tip of an advancing crack and small zones of plastic metal deformation form there for each cycle of fatigue. This is true because the sharp tips of a growing crack provide significant stress concentration points. Compressive stresses retard or stop crack growth.

The crack growth process may continue in a generally stop-and-go manner, in some variation of that manner or not occur at all. This is true when consideration is given to the additive nature of the absolute values of the acting cyclic stress and acting mean stress, S<small>MEAN</small>, defined in Figure 4. If the minimum overall stress (cyclic plus mean stress) is tensile and is sufficient to continue opening the advancing crack, as in Figure 4, the crack growth rate may be relatively fast. If the additive minimum stress is not at a sufficient tensile level to advance the crack but the net maximum cyclic stress is sufficient, then the crack likely would advance but at a slower rate. The rate of crack growth depends on several material properties, alternative ways of loading the component and other service factors besides just the effects of stress. Net overall compressive stress always stops crack growth.

Mechanical Fatigue of Metals
A SunCam online continuing education course

An in-depth explanation of the mechanism of fatigue crack growth and in particular estimating the rate of that growth involves several aspects of fracture mechanics. As previously stated, use of the LEFM approach is based on the assumption that a detectable-size crack is initially present. With LEFM much is then devoted to defining and using multiple parameters that affect the rate of crack growth.

The high-cycle (HC) approach using the S-N method is much simpler. That is true because the S-N method is based on taking actions to minimize the probability of simply nucleating a macro-size crack. Analysis of the growth of an existing macro-size crack is not a part of the S-N method and so that area is not discussed here.

5. Quantifiable Effects on Fatigue – By the Marin Method

The following discussions provide summaries of five factors that can be closely estimated by specific methods plus the last factor that is based on judgement alone. The product of the six factors determines an initial adjustment to the fatigue limit of an alloy that was established in lab testing. The assumptions are that a steel alloy is being used with the S-N method and that safe operation for an infinite number of cycles is desired.

The objective is to allow the engineer to begin adjustment of laboratory-generated fatigue limit values to more realistic fatigue limit values given actual service conditions.

The procedures are based on early empirical work by the late Dr. Joseph Marin (Reference 3) regarding alloys that can have defined fatigue limits such as low strength steels and titanium alloys. Specific methods used in the Marin calculation procedure are provided in standard mechanical design textbooks. The specific resource is used here is Reference 4.

Reference 4 uses the symbol **k** with a different subscript for each factor depending on which effect is being considered. The six factors are metal surface roughness, size of the in-service component, the type of in-service stress loading, the in-service temperature, the desired reliability of the service component and a miscellaneous effects factor to cover influences not recognized by the other five factors. The value of the miscellaneous effects factor depends entirely on the judgement of the engineer. Knowledge of these factors (plus other effects discussed in other sections below) is vital in designing to avoid fatigue failures. Recognizing them is also very helpful when conducting a root-cause failure analyses in which fatigue may be involved.

Mechanical Fatigue of Metals
A SunCam online continuing education course

Factor for Metal Surface Roughness -

Most fatigue cracks nucleate at or near a metal's surface. Any pre-existing surface roughness or small irregularities act to encourage crack formation because of the stress concentration effects of the roughness. This becomes more detrimental as the number of acting fatigue cycles, **N**, become large. The effect is particularly harmful when the roughness is nearby macro-size stresses concentration features, i.e., notches. This effect is essentially absent in the mirror-polished surface condition of coupons commonly used for lab testing. Therefore lab-generated data must be adjusted.

Relative surface roughness has long been studied for steel alloys. Relevant data have been developed and have been incorporated into the Marin calculation as the first factor considered. The surface roughness factor is given the symbol k_A. These values for steel decrease and thus roughness becomes more pronounced progressively as the possible surface changes from fine-ground or commercial polished, to machined or cold-drawn, to hot-rolled or to as-forged. For a given type of surface condition, k_A values decrease directly as the hardness and strength of the steel alloy used increases. Available data for other alloys besides steels are rare. Obviously the goal should be to avoid a surface conditions that provides a small value of k_A.

Factor for Component Size -

Most laboratory test coupons are small in size but in-service components are typically significantly larger. Actual-size components provide lower fatigue limit values in service than smaller laboratory test coupons when other variables are equal. It has been concluded this is primarily because a larger service component has different stress distribution compared to a small test coupon. This is true all types of loading except for axial loading. The more typical loading is bending. In bending fatigue loading on a rotating shaft, a larger cross section means there is little stress at the neutral axis, centerline of the shaft compared to the stress near the metal's surface. In contrast a much smaller lab coupon in bending has a more gradual gradient of stress from neutral axis to the surface. The difference in fatigue limits provided is accounted for by a size factor in the Marin procedure. This factor is given the symbol k_B.

Factor for Different Types of Cyclic Loading-

Actual parts in fatigue service might be loaded in different ways, e.g., bending, axial or torsion. In contrast coupons for lab fatigue tests are most commonly loaded only by bending. This will provide different S-N lab data compared to actual service results with axial or torsional loading assuming

Mechanical Fatigue of Metals
A SunCam online continuing education course

other factors are equal. This is addressed by a load factor for either axial or torsional loading and that factor, given the symbol **k**c , is included in the Marin procedure.

Factor for Service Temperature –

Accounting for the effect of this factor may not be straight forward. Most lab testing is done at room temperature but some service environments vary from this value. In some applications the service temperature variation above or below room temperature may be extreme.

Information on the effects on fatigue loading at low temperatures was discussed very little or not at all in the references consulted. The effect of alloy brittleness at very low temperatures with a notch present, as would be common in practice, was said to be a particular problem for high strength materials with low ductility. No specific guidance was found for using this relationship in design.

Temperatures moderately below room temperatures do not seem to have a detrimental effect on fatigue performance. In fact somewhat lower temperatures are associated with increased alloy strength and thus greater resistance to fatigue failure. This is particularly true for high-cycle (HC) fatigue at long lives as emphasized here.

Most of the information located on temperature effects covered high temperatures well above room temperature. High service temperatures can introduce considerable complexity to predicting alloy fatigue failures and uncertainty into how to adjust variables to prevent failures. This is because both creep and fatigue can occur together depending on the service temperature, other operational variables and the alloy being used. Fatigue and creep interact so that cyclic stress, time of exposure and other parameters affect results. The interactions become complex.

Creep is plastic strain deformation (often leading to creep rupture and separation) of a metal while a constant stress is acting at high temperature over a period of time. Definitive action of creep starts at about 50% of a metal's melting point. Based on their melting temperatures, different classes of alloys begin creep as follows: titanium alloys at about 1600^0 F, stainless steels at about 1320^0 F, carbon steels at about 1250^0 F and aluminum alloys at about 610^0 F.

To simplify this discussion, only more common raised temperature applications will be considered. It will be assumed that the maximum service temperature is 800^0 F, steel alloys are being used and <u>only fatigue without creep</u> is acting. Reference 5 states that carbon steels and low alloy steels may not have a fatigue limit at temperatures above about 800^0 F. That limit on temperature is used here. This

Mechanical Fatigue of Metals
A SunCam online continuing education course

is because adjustments to lab-generated fatigue limits of steels are the central objective of the Marin procedure.

As stated in Reference 4 and elsewhere, the room temperature yield strengths of common steels decrease as service temperatures increase. However, the ultimate tensile strengths (UTS) of these steels initially increase slightly from their room temperature values up to about 500^0 F but then continually decrease as temperature increases above 500^0 F.

The authors of Reference 4 state that fatigue limits of steels respond to raised temperatures similarly to the way that their UTS values respond. They propose that the Marin factor to show the effect of raised temperatures on steel fatigue limits, k_D, follow this correlation between steel UTS values and raised temperatures. That proposal is used here.

Specifically, Reference 4 state that the Marin procedure factor for the effect of raised temperatures on fatigue limits of carbon steels and low alloy steels should be as follows:

$$k_D = S_T / S_{RT} \; , \; \text{where} \;\; S_T \; = \; \text{the UTS of the selected steel at service temperature, and}$$
$$S_{RT} = \text{the UST of the selected steel at room temperature.}$$

Values of S_T and S_{RT} can typically be found in general engineering references. Reference 4 provides a table that lists values of the k_D ratio for temperatures from 70^0 to 1100^0 F. However, given the comment above about 800^0 being the approximate maximum for which fatigue limits exist for steels, it is recommended that use of this procedure for k_D be limited to a range of 70^0 to 800^0 F. Interpolation between different temperatures and UTS ratios likely will be necessary.

Factor for Reliability of Laboratory-generated Fatigue Limit –

As previously discussed and shown on Figure 3, there is considerable scatter in laboratory-generated fatigue data. The normal curve fitting procedure is to plot the S-N curve based on points at the means of the multiple data points at each of the cyclic stress ranges. The resulting plotted data provides an S-N curve, including a fatigue limit that has an estimated reliability of 50%. This value may or may not be acceptable to the user of the data. Therefore the Marin procedure, as described in Reference 4, defines a reliability adjustment factor, k_E that can be applied to the lab fatigue limit data to obtain one of several higher reliability values.

The value of k_E is established for higher values of data reliability by a simple equation. It uses derived values of a parameter associated with different desired reliabilities. The equation was

developed by probabilistic studies of lab-generated fatigue limit data and related standard deviations of 8%.

Miscellaneous-Effects Factor -

This factor, given the symbol, k_M, is intended to show the estimated effects on the laboratory-generated fatigue limit by metallurgical characteristics of the steel being used, potential corrosion in the service environment or other effects that that are present in the given application. General information on metallurgical effects and corrosion are discussed in next section. The value of k_M used is not quantifiable and therefore depends on the engineer's judgement.

In summary**:** The Marin equation seeks to make an <u>initial</u> adjustment (see NOTE) to the fatigue limit of a steel alloy defined in the laboratory when using the most common lab test conditions. Those conditions do not normally exist in service. Thus a revised fatigue limit is defined by the following equation (Reference 4, page 295):

$$S_e = (\ k_A\ \ k_B\ \ k_C\ \ k_D\ \ k_E\ \ k_M\ )\ S'_e$$

Where S_e = the initial adjustment to the lab-generated fatigue limit stress amplitude, S_A, based on the product of the values of the **k** factors defined above, and

S'_e = the lab-generated fatigue limit stress amplitude with common testing conditions

NOTE: No effects of mean stress, S_{MEAN}, <u>or</u> of notches are assessed by the Marin procedure.

6. **Other Effects on Fatigue – Not Directly Quantifiable in the Marin Procedure**

Acting Stresses –

Recognize that the stress indicated on the vertical axis of Figures 3 and 4 represent the stress amplitude $S_A = \Delta S/2$ <u>or</u> stress range $\Delta S = (S_{MAX} - S_{MIN})$**,** for loading either a lab test coupon or an in-service component to fatigue. This net stress value is the algebraic sum of the signed values (tensile or compressive) of both the applied cyclic stress and often stress that acts without variation. The latter is the mean stress, S_{MEAN}. The mean stress will include any constantly acting operational stress plus any residual internal stress that exists in the metal. Often residual stress, whether tensile or compressive, makes a significant contribution to the realistic fatigue life that occurs and may be larger than applied constant stresses.

Mechanical Fatigue of Metals
A SunCam online continuing education course

All tensile stresses are detrimental and all compressive stresses are helpful in promoting resistance to fatigue failure. Assuming other factors are equal, the longest fatigue life in service will occur for a given alloy when net tensile stresses are minimized, net compressive stresses are maximized and net S_A or ΔS are minimized. It is very important to note that the net stress acting at or near the metal's surface has a major effect on fatigue life.

Many mechanical design textbooks state that a good estimate of the fatigue limit for steel components in-service can be found based on a simple relationship to the ultimate tensile strength (UTS) or ultimate shear strength of the given alloy. For example, one common relationship indicates a good estimate of in-service fatigue limit is 50% of the alloy's UTS for alloys having a UTS less than or equal to 200 ksi (or for steels with a UTS more than 200 ksi a good estimated actual fatigue limit can be taken as 100 ksi). As stated in Reference 2, these approaches are not recommended. These estimates are approximately true only for data from lab testing using all conditions that typically apply to lab-generated data. Those conditions don't normally exist in practice.

As emphasized here, there are many factors that often apply to components in actual service that effect fatigue limit that are not assessed in traditional lab testing. The simple 50% of UTS as a means to define a fatigue limit is not valid as a final value for use in design or analysis.

Metallurgical Effects –

Recognize that both an alloy's ultimate tensile strength and its ductility are important to its resistance to fatigue. The ideal alloy for fatigue resistance and long life would maximize both material properties. The problem is that typically these two measures of an alloy are inversely proportional, e.g., very high strength usually means low ductility. Clearly high strength is important but ductility is also important because raised ductility allows local yielding at points of maximum stress such as notches. This allows the metal to slightly deform at that spot without generating a crack that could lead to fatigue failure.

A trade-off between strength and ductility usually is desirable in selecting an alloy. Variations in heat treatments and/or thermomechanical metal working processes can provide a wide range of strength and ductility results.

Generally high strength is more important to high-cycle (HC) fatigue resistance but high ductility is more important in low-cycle (HC) applications. In the case of LC fatigue higher ductility may be sufficient to allow only local yielding, without nucleating a crack even though stress levels are relatively high.

Mechanical Fatigue of Metals
A SunCam online continuing education course

Ductility is indicated in percent elongation or percent reduction in area as defined in traditional tensile tests. Alloys that have percent elongation at about 5% or less are brittle and are unsuitable for any type of fatigue service. Most useful steel alloys for fatigue service have a ductility of about 20 % elongation. Many high strength steel alloys can have ductility values less than 20% elongation. In all cases much depends on the heat treatment or mechanical working process applied during alloy manufacturing.

Internal discontinuities in metals such as inclusions, voids, porosity and laminations lower resistance to fatigue. This is because sub-surface fatigue cracks are more likely to nucleate at these types of defects in metal. These features are more common in the cast form of metals rather than the wrought form. Welds solidify in cast form. This is one reason that welds are more susceptible to fatigue failure than adjacent, wrought parent metal. Steel manufacturing using what are known as "clean steel" methods such as vacuum melting decrease the size and quantities of internal discontinuities. These methods provide materials with improved fatigue resistance.

Often larger components have more internal discontinuities than smaller ones simply because there is more volume in the larger size part. The differences in stress distribution between large and small parts was discussed above relative the Marin size factor, k_B, The probability of more internal discontinuities in a large, full-size service part presents another reason for the need to adjust lab-generated fatigue limit data based on use of small test specimens. This effect is another reason for having a k_B factor in the Marin procedure.

Metallurgical grain (or crystal) size in the alloy used is another factor to consider. In general a small grain size is desirable because this tends to resist crack nucleation up to moderately high temperatures. However, at high service temperatures at which fatigue and metallurgical creep can occur together a large grain size is desirable to retard creep.

As previously mentioned, creep typically starts at about 50% of the melting temperature of an alloy. For plain carbons steels this threshold for the start of creep is about 1250^0 F. That value is about 1320^0 F for stainless steels. It is expected that in most applications with fatigue the maximum service temperature would be much less than these values. Therefore in those more common situations a smaller (rather than larger) grain size would be a desirable feature to resist fatigue.

Certain metal cold working manufacturing processes like drawing or rolling tend to elongate grains (crystal shapes) of the metal in the direction of the applied force while compressing the grains in a direction perpendicular to the force. Fatigue stress loading parallel to the direction of elongated grains gives much better fatigue resistance than fatigue loading perpendicular to elongated grains. The latter

Mechanical Fatigue of Metals
A SunCam online continuing education course

direction is known as the short transverse direction. This difference should be acknowledged during design for and analysis of fatigue.

Corrosion Effects –

Corrosion greatly lowers the fatigue resistance of several metals. It is one of the most important and difficult parameters to account for in fatigue design and analysis. Many alloys exposed to a corrosive environment will no longer have the clear fatigue limit that they previously had in a benign environment. The dashed data line in Figure 3 shows this effect schematically.

To various extents, all alloys exposed to corrosive service behave like the dashed line in Figure 3. When a fatigue limit is omitted, corrosion significantly lowers the fatigue strength data line especially at a high number of fatigue cycles. Alloy failure can then occur at a much reduced fatigue stress level.

In a strict sense no alloy can maintain a definitive fatigue limit while fatigue loaded in a corrosive environment. The degree of downward movement from the alloy's former fatigue limit and the extent of its loss of fatigue strength as the number of cycles, N, increase will depend on the corrosion resistance it has to the specific corrosion medium and other service conditions.

Corrosion is a not a distinct factor in the Marin procedure. However, like internal metallurgical discontinuities in a service component, the effect of corrosion can be roughly incorporated in the procedure as another input for selecting the value of the Marin miscellaneous effects factor, k_M. Selecting the value of this factor rests on the engineer's judgement and/or information from others. What is needed is knowledge of the corrosion resistances of different potentially useful alloys in the specific corrosion conditions of the specific application. Often alloy manufactures can very helpful.

Corrosion fatigue is complex because of the synergistic interaction of fatigue and corrosion. Each negative effect acts to make the other more damaging. The effects of corrosion fatigue are difficult to quantify by choice of the k_M factor alone. A more definitive but complicated and time consuming approach is to conduct special laboratory fatigue tests.

Corrosion fatigue tests should be completed using procedures that include the selected alloy, the expected fatigue loading plus the expected corrosion medium and service temperature. The objective is to establish the approximate minimum fatigue strength of the alloy tested at a desired maximum number of fatigue cycles, for example at 10^8 cycles. Multiple tests may be necessary. Some alloy

manufacturers have data from such testing of common alloys in frequently encountered corrosive media.

Corrosion is an electrochemical process requiring oxidation and reduction chemical reactions. Both reactions must take place simultaneously for corrosive to occur. Initially metal damage occurs on a microscopic scale. This magnitude increases with time and is accelerated by fatigue. Water, even water vapor (as moisture in humid air), or some other corrosive medium that contains water, plus oxygen are needed for the most common oxidation and reduction reactions to take place. Oxygen, and thus air, is essential for the most common reduction reaction. The overall process is accelerated by raised service temperatures.

During corrosion, small pit-like features are created on the metal's outer surface and grow in size with time. The pits act as stress concentration points to promote fatigue crack nucleation. Once a significant fatigue crack is growing through the cross-section of the metal, more corrosion occurs at the tip of the advancing crack. This corrosion combined with fatigue loading serves to accelerate the rate of crack growth. In addition exposed metal surfaces in the gradually opening split in the metal are attacked by corrosion. It often happens that indications of fatigue, i.e., beach marks and striations, on the final separated fractures surfaces are destroyed by corrosion or are obscured by corrosion products that form there.

The frequency of the acting fatigue cycles is important to the extent of metal damage that occurs. With lower fatigue cycle frequencies an advancing fatigue crack exposes "fresh" metal that has not yet been attacked by corrosion for a longer time per cycle. More corrosion can then occur compared to a higher cyclic frequency when fresh metal is exposed for a shorter time period per cycle.

Because of this effect, corrosive fatigue damage in a slower cyclic application is likely more severe than in an application with high frequency cyclic loading. For example, a rotating shaft running at lower RPM in a corrosive environment will likely suffer more corrosion fatigue damage than that same shaft in the same environment operating at a higher speed. The frequency of cyclic fatigue loading used in laboratory corrosion fatigue tests should always be reported as a required testing parameter.

Aggressive corrosives such as acid solutions or saltwater are not necessary for corrosion fatigue to occur. As previously mentioned, high humidity ambient air can be sufficient depending on the alloy used. Complete resistance to corrosion is only found in a vacuum where there is no water or oxygen. Then the two necessary electrochemical reactions for corrosion cannot occur. High susceptibility to moist air applies primarily to plain carbon steels and to a lesser extent to low alloy steels. Figure 5

illustrates the effects of some different environments on the fatigue resistance of a specific ferrous alloy:
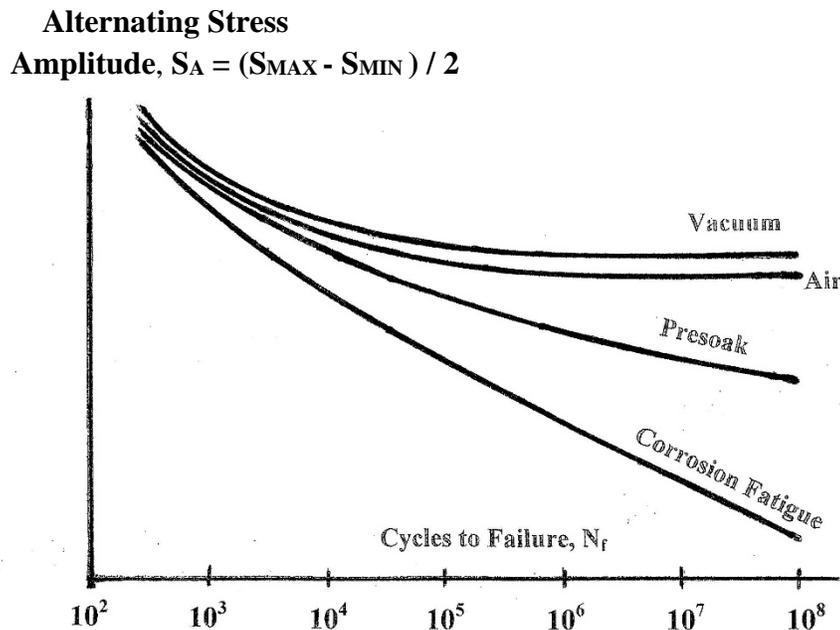
Figure 5 – Schematic Representation of Comparative S-N Fatigue Curves for Different Test Conditions (After Reference 2, page 347)

The schematic curves in the figure show the effects of different environments on constant amplitude S-N results generated at different stress amplitude values for the same hypothetical ferrous alloy. Approximate room temperature is assumed during each test. The relative relationships between the different curves and environments are said to be very realistic (Reference 2).

The top curve represents tests in which no water vapor or oxygen is present as in a vacuum. A fatigue limit is established. The next lower curve represents tests in ambient air that has a small amount of humidity and therefore a small amount of water vapor. A fatigue limit is defined here and this is the condition of most commonly used during laboratory fatigue testing. This fatigue limit is established but at slightly lower stress amplitude. The next lower curve included presoaking the metal in an aggressive corrosive without any fatigue loading followed by a full range fatigue test in ambient air without any aggressive corrosive medium. The lowest curve represents corrosion fatigue, i.e., the simultaneous exposure to the range of stress amplitudes levels plus an aggressive corrosive medium.

Mechanical Fatigue of Metals
A SunCam online continuing education course

As seen in Figure 5, the four curves are close together at a low number of fatigue cycles. This is because the variable corrosive effects represented have not had enough time to create significant differences in the S-N curves. It illustrates that the damage that occurs in corrosion fatigue of a given alloy in a specific corrosive environment is always time dependent. Less corrosion resistant alloys in more aggressive environments sustain damage quicker.

There are a few ways to minimize the damaging effect of corrosion fatigue. If viable, controlling the service environment to make it less aggressive is always desirable. This includes exposing the fatigue loaded equipment or component to less corrosive conditions by measures like use of barriers or enclosures. Another important measure, if possible, is to lower the service temperature. Corrosion is in part is a chemical process and as such the rate of corrosion is slowed at lower temperatures and accelerated at higher values.

Another corrosion fatigue control measure is to provide a protective coating on the alloy being used. Certain coatings can provide corrosion resistance for metals. Galvanizing a plain carbon steel provides good corrosion resistance with no detrimental effects on fatigue performance. Unlike the much softer zinc used in galvanizing, electroplating metals with chromium or nickel is very harmful to fatigue resistance. Those two elements introduce tensile stresses and micro-cracks in the surface of the metal and should not be used.

Perhaps the most common and practical measure to resist corrosion fatigue is to use an alloy that is resistant to corrosion in the specific environment. In general this means using a material that contains sufficient levels of alloying elements, e.g., chromium, nickel and molybdenum and other elements, in its composition to resist the specific corrosive medium. Plain carbon steels contain little or none of these elements. Low alloy steels have higher mechanical strength compared to plain carbon steels but not enough alloying to provide appreciably better corrosion resistance than carbon steels. Other alternatives may be copper or nickel-based alloys but their use may be limited due to inferior mechanical properties, cost or both.

One of the many stainless steel alloys is probably the most widely used material to resist corrosion fatigue. These alloys contain various levels of the alloyed elements mentioned earlier. Their corrosion resistance, mechanical properties, metallurgical properties and costs offer a variety of options. There are many alternatives in addition to the widely used 304 and 316 alloys. Sometimes this is not widely appreciated.

It is recommended that design engineers identify the specific details of their application and contact one or two alloy manufacturers for their advice on making a material selection. As previously stated,

those same suppliers may be able to offer laboratory corrosion fatigue data on some alternative stainless steel and related alloys.

7. Quantifying the Effect of Mean Stress on Fatigue Limit

Engineering researchers have done considerable work to define different procedures to predict the effect of non-zero mean service stresses on lab-generated fatigue limit values in which no mean stress was applied. In each of these procedures the lab-generated fatigue limit is first adjusted for the "**k**" factor effects included in the Marin procedure described in Sections 5 and 6.

One of the procedures shown here can then be used to predict a further reduced, <u>equivalent fatigue limit, S_{equiv}</u>. This derived value includes the effect of the given stress amplitude equal to S_e , i.e., the adjusted fatigue limit found from the Marin procedure, <u>plus</u> the effect a given calculated net mean service stress, S_{MEAN}. Several methods exist but only two of the many approaches are presented here.

Perhaps the best known and most often used of these procedures is the Goodman diagram that has long been in standard design textbooks and used by many generations of engineers. Two sources are References 2 and 4. The modified Goodman procedure provides for the effect of net mean stresses on the fatigue limit of the given alloy relative to its yield and ultimate tensile strengths. Generally this is done graphically as illustrated in Figure 6.

All <u>net</u> S_{MEAN} values shown here are assumed to be tensile stresses and therefore deleterious to fatigue life. They are 'net" values because each should be based on the algebraic sum of signed positive tensile stresses and signed negative compressive stresses. However, quantified values of compressive stresses usually are not known.

Omitting the favorable effect of compressive stresses – both applied and often the larger residual compressive stresses - provides a conservative prediction of S_{equiv} and simplifies the Goodman approach. That approach is used here.

All plotted stresses and strengths are first adjusted with a selected factor of safety. Alternating stress amplitudes, S_A, yield strength (here given the symbol S_y) plus the lab-generated fatigue limit S_e are plotted on the vertical axis. Ultimate tensile strength, here given the symbol S_{UTS}, and the S_y of the given alloy are plotted on the horizontal axis. Calculated values of net S_{MEAN} for the given service are also plotted on the horizontal axis. The scales of both axes are linear and here have units of (ksi or MPa).

Note: The many different textbooks and published articles on fatigue design use different symbols and subscripts in describing the same or very similar fatigue parameters and variables. All symbols used here are defined. As always, it is vital to carefully compare "apples-to-apples" when viewing information from different sources.
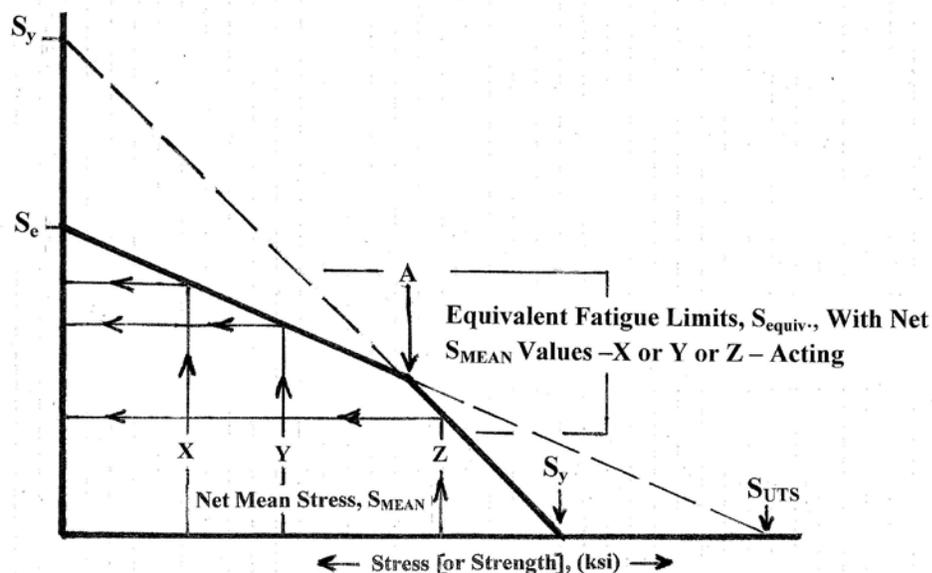
Figure 6 – Schematic Illustration of the Modified Goodman Method of Defining the Effect of Three Net Mean Tensile Stresses on the Fatigue Limit, S_e, with $S_{MEAN} = 0$, and the Resulting Reduced Equivalent Fatigue Limits (S_{equiv}). {after References 2 and 4}

This graphical method uses the thicker, solid bold line from S_y on the horizontal axis, to point **A,** to S_e. The intersections of vertical lines up from the S_{MEAN} values and the solid, bold lines projected over to the vertical axis establish the maximum allowable stress amplitudes of the three desired equivalent, reduced fatigue limits, S_{equiv}.

Any S_{MEAN} vertical intersection with the bold line from S_y on the horizontal axis to point **A** (here shown as S_{MEAN} line **Z**) would represent the maximum S_{MEAN} for a given alloy plus cyclic stress that could exist without failure by yielding. In the same way, any vertical intersections with the bold line from point **A** to S_e represent maximum S_{MEAN} values that can exist along with cyclic stress without failure by complete fracture occurring.

Mechanical Fatigue of Metals
A SunCam online continuing education course

Figure 6 shows that net means stress reduces (lowers) the fatigue limit value of an alloy. That is as S_{MEAN} increases the equivalent fatigue limit (here shown as S_e denoting the equal endurance limit term) is lowered along with the entire S-N diagram as shown in Figure 3.This result applies to all methods used to define the effect of net means stress. The figure also shows that the extent of the fatigue limit reduction is directly proportional to the magnitude of the given net mean stress.

In summary, as seen in Figure 6 for the Goodman method, the magnitude of the applied net mean stress and the degree of resistance to fatigue failure are inversely proportional. This is a general relationship that applies to all methods used to define the effect of net mean stress on fatigue limits.

In any application the working stress, i.e., the allowable stress, used for design must be less than the stress level corresponding to the final adjusted fatigue limit to produce an essentially infinite number of cycles without failure. That "infinite number" is generally taken as approximately 10^6 to 10^7 cycles. It is assumed that if a service component subjected to fatigue loading can function for that maximum number of cycles without failure it will do so forever.

While the Goodman method is the oldest and likely the most familiar to most engineers, several current researchers have shown it is inaccurate compared to other methods. For example, Dowling (Reference 7) compared Goodman to several other approaches in tests using a variety of alloys. He compared the predictions of S_{MEAN} effects by several methods to the actual performance of the different alloys he used in separate tests.

Dowling's conclusion was that a method by Walker (Reference 9) was more accurate than Goodman and also better than other proposed methods for a diversity of alloys. This is also discussed by Palmer (Reference 8).

Dowling stated that the equation developed by Walker to express his proposed method was best for different types of alloys because it contained an empirically developed constant that is specific to the type of alloy used. Palmer (Reference 8) further explained the deficiencies of the Goodman method and why the conclusion by Dowling is valid. The Walker equation is as follows:

Mechanical Fatigue of Metals
A SunCam online continuing education course

$$S_{equiv} = (S_e)^\gamma (S_e + S_{MEAN})^{1-\gamma},$$

where S$_{equiv}$, S$_e$ and S$_{MEAN}$ are as previously defined and
$$\gamma = 0.65 \; for \; steel \; alloys \; and$$
$$\gamma = 0.50 \; for \; aluminum \; alloys$$

Notice that the second factor in the Walker equation and its power with the empirical constant can be replaced with it equal, **S$_{MAX}$**. This is true because **S$_{MAX}$ = (S$_{MEAN}$ + S$_A$)** as defined in **S**ection 2**,** where here **S$_A$** is the cyclic stress amplitude at the fatigue limit, **S$_e$** with no **S$_{MEAM}$** acting. Therefore this equation may be used in either of the two forms.

Note: The effect of a macro-scale notch on fatigue limits is not accounted for in either the Goodman approach or by the Walker equation. Only predictions of the effects of different **S$_{MEAN}$** values are the objective of these two evaluations.

8. Quantifying the Effect of Notches on Fatigue Limit with Mean Stresses Acting

Notches are common <u>but</u> not present in all applications. If not present the following is not needed. Then only the Marin procedure and accounting for **S$_{MEAN}$** are needed to adjust lab fatigue data.

Notches on a component's surface often present larger negative effects on fatigue life than surface roughness. They create macro-scale stress concentration locations that are formed by features such as drilled holes, small fillet radii between different diameter sections on a stepped shaft, corrosion pits, machined grooves and screw or bolt threads. These common features reduce the fatigue limit of the given alloy as defined in lab tests. Applied tensile mean stresses also reduce fatigue limits.

The joint effects of notches and tensile mean stresses on fatigue limits can be assessed by manipulating certain relationships between the relevant variables and using empirically derived relationships. By use of this approach, a final adjusted fatigue limit can be estimated for use in design or analysis.

In cyclic stress applications a fatigue stress concentration factor is used to determine the extent that notches reduce fatigue life. It is given the symbol **K$_f$.** It is derived <u>in part</u> from the stress concentration factor that applies for the given notch in non-cyclic, constant stress applications. The

constant stress concentration factor is usually given the symbol K_t. It depends only on the geometry of the given notch and type of loading.

The value of the cyclic stress concentration factor K_f is dependent on the specific value of K_t <u>plus</u> the notch sensitivity of the given alloy and the mean stress, S_{MEAN}, that acts along with the given cyclic stress. Notch sensitivity is a material property that varies in value from zero to 1.0 and has the symbol **q**.

The fatigue stress concentration factor for an alloy is defined by the following ratio (Reference 2, page 196):

K_f = Fatigue Strength with No Notch / Fatigue Strength with A Notch,   Equation (1)

The value of the denominator in this ratio is the desired unknown quantity if a given mean stress S_{MEAN} is assumed to be acting with a notch present. That same S_{MEAN} value may also be assumed to be acting when no notch is present as is the case in the numerator of Equation (1). The **Sequiv** fatigue limit as derived by one of the two methods shown in Section 6 can be substituted for the numerator. Therefore if a value of K_f can be found by another method the desired unknown denominator, given the symbol S_{FINAL} here, can be easily found by rearrangement of the ratio in Equation (1) as follows:.

S_{FINAL} = Sequiv / K_f ,   Equation (1a)

Values of K_f for different alloys can be determined by use of the following relationships between notch sensitivity, **q**, K_f and K_t (from Reference 2 and others)**:**

q = (K_f - 1) / (K_t - 1),   Equation (2), by definition - or after rearranging,

K_f = 1 + (q)(K_t – 1)   Equation (2a)

The non-cyclic, constant stress concentration factor, K_t, in Equation (2a) can be found for the given notch (according to its geometry and loading used) in many mechanical design textbooks. Perhaps the best known one for this purpose is (Reference 11).

If multiple notches are present on the component being considered, the one that results in the largest value of K_t should be used.

Mechanical Fatigue of Metals
A SunCam online continuing education course

The only remaining unknown in Equation (2a) is **q**. This can be found by using an empirically developed relationship (also in Reference 11) that includes a material constant that depends on the type of alloy being used and the radius at the root of the notch. That relationship is as follows:

$$q = \frac{1}{\left((1) + \dfrac{\alpha}{\rho}\right)} , \quad \text{Equation (3)}$$

Where α = material constant
= 0.51 mm for aluminum alloys
= 0.25 mm for annealed or normalized lower
strength low carbon steels
= 0.064 mm for quenched & tempered higher
strength steels
ρ = radius at root of notch, mm

Now Equation (1a) can be used to find the desired final fatigue limit, S_{FINAL} – based on the presence of both a notch and mean stress effects - for the given alloy using the value of K_f found and results from Equations (2a) and (3).

It should be emphasized that the value of notch stress sensitivity, **q**, from Equation (3) is an approximation because the derived material constants for different types of alloys were empirically developed after adjustment for agreement with test data. Therefore the final fatigue limit values found by the overall procedure are estimates. However, the predictions are suitable for practical engineering objectives.

9. A Summary of Sequential Steps for Designing to Avoid Fatigue Failures

The following presents a procedure that may be used in design based on the information provided in the previous sections.

Situations that use steel alloys that produce a defined fatigue limit and an essentially infinite number of desired cycles without failure, i.e., high-cycle (HC) fatigue, are the primary application for this sequence.

For alloys that do not produce a fatigue limit but have decreasing fatigue strengths (as the shown by the dashed line on the S-N diagram in Figure 3), these steps may also be useful. In that case it is

Mechanical Fatigue of Metals
A SunCam online continuing education course

generally assumed the minimum fatigue strength that exists at approximately 10^8 cycles and the corresponding cyclic stress level can act as a substitute for a true fatigue limit.

1) For the selected alloy, reduce the commonly available fatigue limit cyclic stress value as generated in common laboratory tests by identifying the values the various **k** factors and applying them by the Marin method equation in Sections 4 and 5. The resulting S_e value is thus established. However, S_e does not include the effect on fatigue life of net mean stresses, **S**MEAN, acting without notches on the service component <u>or</u> the joint effect of **S**MEAN and notches acting together to reduce the fatigue limit.

2) With the same alloy, reduce the fatigue limit S_e to account for the effect of a calculated value of **S**MEAN using one of the methods in Section 6 so as to obtain **Sequiv**. The Walker equation is recommended.

3) With the same alloy, reduce **Sequiv** to account for the joint effect of the calculated value of **S**MEAN <u>and</u> a notch on the component that produces the largest value of **K**f . Use the procedures in Section 7. The result is the final reduced value of the fatigue limit, **S**FINAL.

4) Chose an appropriate factor of safety, **FS**, for the specific application. **FS** values for fatigue loading are generally larger than used in static, constant stress applications because of the greater uncertainties with fatigue. **FS** values for fatigue typically are in the range of 2.5 – 4, or somewhat higher, when applied to cyclic stress. The engineer must use his/her judgement in picking this value based on the consequences of a premature failure and the any uncertainties that remain after accounting for all clear deleterious effects.

5) Determine the working stress (equal to allowable stress) for the application by dividing the selected alloy's ultimate tensile strength by the selected **FS**. In many cases the allowable stress will be dictated by a code.

6) If the working stress value is less than the **S**FINAL fatigue limit value the design is theoretically safe for the intended service conditions. However, if the working stress is more than **S**FINAL the engineer should make changes and go back through these seven steps. The changes may include a different alloy, different overall dimensions for the component, different treatment of notches (for example different root dimensions or other preventative actions) or incorporation of one or more of other fatigue minimization actions listed in Section 9.

Mechanical Fatigue of Metals
A SunCam online continuing education course

7) This final step is always desirable and may be mandatory. Complete fatigue tests with a full-scale prototype of the component as designed. In these tests use conditions as close as possible to the actual service conditions. This is the last, **best test** of the several assumptions and uncertainties inherent in the design.

10. Summary of Selected Metallic Fatigue Characteristics and Control Actions

Mechanical fatigue of metals is the most important mode of premature failure when considering the many applications in which it might occur. Meaningful fatigue design or analysis can be complex.

Macro-scale beach marks and microscopic striations on a fracture surface are definitive indications of fatigue. These features may be obscured for several reasons.

There are two types of fatigue: low-cycle (LC) and high-cycle (HC). LC includes a small number of cycles before failure but high stress. HC fatigue includes many cycles before failure but lower stress.

The objective of both the S-N and strain-N methods of fatigue design is to prevent the nucleation of macro-scale cracks. The LEFM method assumes that a detectable size crack already exists but seeks to prevent final, fast fracture.

Commonly available fatigue data generated in lab testing should not be used directly. Typical conditions used in lab tests differ in several ways from actual service conditions.

The majority of fatigue cracks nucleate on or near the surface of a metal. Attention to surface roughness and, especially, notches is essential. Notches can greatly lower fatigue resistance.

Control the effects of notches by: eliminating them when possible, making notch root radii as large as possible and modifying the geometry of components to make all size changes as gradual as possible.

Tensile fatigue stresses accelerate fatigue crack nucleation and growth. Compressive fatigue stresses retard both. Adding residual internal compressive stresses in a metal provides an effective fatigue control tool. Applied or residual internal tensile stresses always lower fatigue resistance.

Adding residual compressive stresses may be accomplished by shot peening, cold rolling and other methods. These actions are particularly beneficial in and nearby notches.

Mechanical Fatigue of Metals
A SunCam online continuing education course

Welds often form residual tensile stresses during cooling. This is one reason why welds can be particularly susceptible to fatigue failure. Post-weld stress-relief heat treatment is often mandatory.

The magnitude of cyclic stress amplitude, S_A, or stress range, ΔS when each is in the tensile range, is inversely proportional to possible alloy fatigue life when other variables are constant.

Operational S_A or ΔS may be minimized during HC fatigue design by selecting a higher strength alloy or by increasing metal cross-sectional dimensions.

Alloy strength and ductility are both important to the fatigue resistance. Strength is more important to achieving high-cycle (HC) fatigue resistance. Ductility is more important for low-cycle (LC) fatigue.

Brittle alloys are not suitable for use in fatigue service. Alloys having a percent elongation near to or less than 5% should never be used.

Many plain carbon steel and low alloy steels develop a well-defined fatigue limit at higher (N) cycles in relatively benign service environments.

Most non-ferrous alloys don't have a true fatigue limit but have only fatigue strengths that decrease significantly at high N values. For these alloys, a "substitute fatigue limit' at the fatigue strength of about 10^8 N cycles is often assumed.

The Marin procedure uses **k** factors to adjust lab-generated fatigue limit data to account for several conditions that often appear in service but that are not used in common lab fatigue testing. The Marin procedure doesn't account for either net mean stress, $S_{MEAN,}$ or the effect of notches.

$S_{MEAN,}$ is the net algebraic sum of constant (+) tensile and constant (-) compressive stresses. Either type may be applied stress or internal residual stress.

Ignoring the compressive component of S_{MEAN} during design provides conservative results, i.e., actual fatigue lives are longer than expected and fatigue resistance is greater than expected.

S_{MEAN} values, when assumed to consist of only tensile stresses, act with cyclic stresses to reduce the fatigue limits of metals below the values of their fatigue limits with no S_{MEAN} acting. This reduction can be determined by several methods. The Goodman diagram is traditionally used but it is inaccurate. The Walker equation is recommended.

Mechanical Fatigue of Metals
A SunCam online continuing education course

The effect of notches on fatigue limits (with S_{MEAN} acting) is not accounted for by either the Goodman diagram or the Walker equation.

The combined action of cyclic fatigue stresses, S_{MEAN} <u>plus</u> a notch causes a larger fatigue limit reduction compared to S_{MEAN} without a notch present. Section 8 provides a recommended procedure.

Not all applications have notches present. In those situations the procedure on Section 8 is not needed.

A corrosive environment typically has a significant negative effect on the fatigue behavior of alloys with poor corrosion resistance to the specific conditions.

The best control measure for corrosion fatigue is to select an alloy that offers good corrosion resistance to the specific environment. Usually one of the stainless steels is a good choice. The degree of corrosion resistance will dependent on matching the composition of the alloy to the requirements of the specific service conditions. Often alloy manufacturers can help.

The ultimate objective in HC fatigue design with steel alloys is to confirm that the working stress (equal to allowable stress) for the given component and alloy is less than the <u>final adjusted</u> fatigue limit of that alloy for a fatigue life of at least 10^6 cycles.

When maximum working stress is not dictated, it is the ultimate tensile strength of the selected alloy divided by a factor of safety (**FS**). **FS** values for fatigue are typically in the range of approximately 2.5 to 4.0 based on stress.

The ideal last step in design for fatigue resistance should be to confirm the performance of a full-size prototype of the designed component in fatigue tests that include all conditions to be encountered in actual service. Realistically this is only done in special situations.

Many topics in fatigue were necessarily omitted in this short course. These included the strain-N method, statistical aspects of fatigue data, stress cycles with non-constant amplitudes, the cumulative effect of different levels of fatigue loading, fretting fatigue, fatigue of welds and related codes, etc. If there is interest in these or other areas References 2, 6 and 10 are recommended.

Mechanical Fatigue of Metals
A SunCam online continuing education course

References

Note: The following sources were used for this course and are gratefully acknowledged.

1. Wulpi, Donald J., Understanding How Components Fail, ASM International, Second Edition, 2000.
2. Stephens, R.I., Fatemi, Ali, Stephens, R.R. and Fuchs, H.O., Metal Fatigue in Engineering, Wiley-Interscience, Second Edition, 2001.
3. Marin, Joseph, Mechanical Behavior of Engineering Materials, Prentice Hall, 1962.
4. Budynas, R.G. and Nisbett, J. K., Shigley's Mechanical Engineering Design, McGraw Hill, Tenth Edition, 2015.
5. Dieter, G. E., Mechanical Metallurgy, McGraw Hill, Second Edition, 1976.
6. Dowling, N. E., Mechanical Behavior of Materials, Prentice Hall, Second Edition, 1999.
7. Dowling, N.E., "Mean Stress Effects in Stress-Life and Strain-Life Fatigue" article F2004/51, Society of Automotive Engineers, Inc., 2004.
8. Palmer, D. E., "Design News – Guest Blogs" two articles: "Understanding Fatigue Failures", 4/9/12 and "Why the Goodman Diagram is Not So Good", 5/30/14.
9. Walker, K., "The Effects of Stress Ratio During Crack Propagation and Fatigue for 2024-T3 and 7075-T6 Aluminum", article in Effects of Environment and Complex Load History on Fatigue Life, ASTM STP 462, American Society for Testing and Materials, pages 1-14, 1970.
10. Collins, J.A., Failure of Materials in Mechanical Design, Wiley-Interscience,1993.
11. Pilkey, W.D., Peterson's Stress Concentration Factors, John Wiley, Second Edition, 1997 and Peterson, R.E., Stress Concentration Factors, John Wiley, 1974.