Metal Fatigue Principles and Analyses

A Compendium

Hans Andersson

November 2021

Metal Fatigue
Principles and Analyses
A Compendium

Hans Andersson

November 2021

ISBN 978-91-89385-90-0

Produced with support from
UTMIS – The Swedish Fatigue Network
and
RISE Research Institutes of Sweden

About the author

The author was employed as CTO and vice president of SP, the Swedish National Research and Testing Institute of Sweden, now being part of RISE Research Institutes of Sweden. He was also an adjunct professor at the Chalmers University of Technology. In this capacity he was a lecturer in courses on fatigue in Chalmers Master´s Programme.

Contact information:
Phone: +46(0)706506065
e-mail: hansof.andersson@telia.com

Contents

1 Scope and objectives

This text is an extension of notes from courses first given during the 1990-ies for engineers employed in industry, and then for master students in the international masters programme at Chalmers Technical University in Gothenburg 2001-2009. Hence, it covers most issues of fatigue analysis as taught on the master level in universities.

The micromechanical features of fatigue as studied in materials science constitute an area of research and education in itself. The descriptions here are simplistic. Some fundamental models are introduced just to give qualitative indications of the micromechanical background to fatigue phenomena. These models are useful for the understanding of principal effects found by experiments on the macroscopic level.

Some emphasis is given to the stochastic nature of fatigue. The use of statistical methods in design is described as well as uncertainty assessments in the evaluation of experiments to find material properties connected to fatigue. Most of the mathematical expressions used in fatigue design are not physical laws but suitable models suggested from curve fitting of experimental results, and then a critical statistical assessment is essential.

The aim is to supply an easy-to-use support to people working with fatigue in industrial environments or as teachers on the bachelor or master level in universities. It should give useful background and understanding to handbook formulas and requirements in standards.

If the text is used in courses it should be complemented by examples and problems to solve since only a few examples of a principal character are included. Likewise, extensive handbook data for e g stress concentration and stress intensity factors, and material properties, are necessary for use in practical design situations.

Computer programs for solving of various problems, as FEM codes or procedures to find load spectra are not included, since the field is changing quickly, and the Internet gives ample information.

The intention has not been to produce a reference work for scientific studies of fatigue. Therefore, literature references are few and only including works of authors having produced milestone contributions to the area, or well renowned textbooks, standards, and handbooks.

It is underlined that all experimental evidence shows that fatigue phenomena in real materials can by no means be precisely described by mathematical models and formulas. Models are sometimes developed in the text by sophisticated mathematical tools giving such an impression. But they are just models, often competing with one another, and they only constitute a framework for understanding and *approximate* data for design.

2 Introduction

It was first experienced during the 19th century that structures failed under cyclic load levels below those causing flow or fracture under static loading. Seemingly, the material in the structure was "fatigued" and the term fatigue was established for the phenomenon. Now we know that the cause is successive irreversible changes on the micro-structural level in the most severely stressed parts of a structure.

Our present knowledge is based on an enormous amount of experimental work. The fundamental understanding rests on microscopic studies, while quantitative effects and design rules are the results of extensive testing. Fatigue is a complex phenomenon influenced by many parameters and it is inherently of a stochastic character. This contributes to the need for a multitude of testing. Generally, it is not possible to extract exact formulas, like Ohm´s law, from the research. Instead, the experiments can be used to form mathematically simple relationships by curve fitting which are suitable for design. Some of the most important milestones in this development are mentioned below.

Since the area is so complex fatigue has been and still is a major cause of failures in industry and transportation. A few examples of spectacular accidents are related, where fatigue was found to be the cause. They illustrate some of the main causes where the design of a structure has been immature or not careful enough.

2.1 Milestones

The emerging industrialism during the 19th century meant that more and more components in machinery were exposed to a large number of cyclic loads.

Although some efforts to explore the phenomenon were made in the first half of the century, A. Woehler gave the first really significant contributions during the 1850-ies. He was employed by the Prussian railway company and had to face a number of sudden fractures in the axles of the coaches. He constructed equipment whereby tests could be made efficiently of test pieces in the rotating bending mode and performed systematic studies of such test pieces exerted to stresses of different levels.

The principal design of the machine is depicted in Figure 2.1. The cantilever test pieces are rotated and loaded at the free ends by weights or springs giving an easily varied bending moment changing sign for each half cycle of rotation.

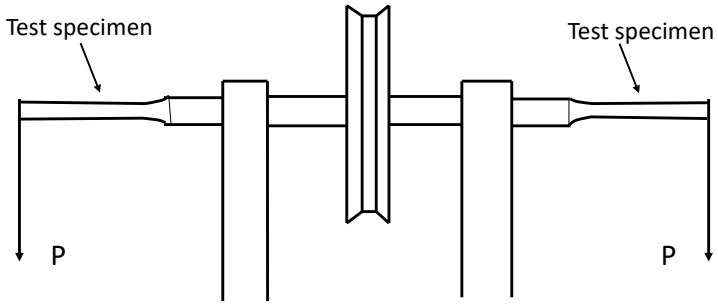


Figure 2.1 Device for testing of the type used by A. Woehler. (Principle from "Fatigue and Durability of Structural Materials"; ASM International (2006)).

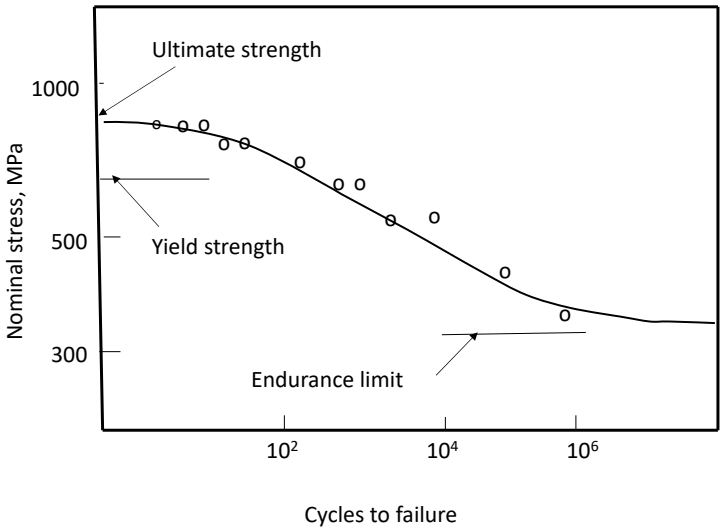


Figure 2.2. Sketch of the S-N-curve showing the principal features.

There were three main findings, see Figure 2.2. The figure shows the relationship between maximum bending stress, and cycles to fracture. This relationship is called the Woehler curve or the S-N-curve (Stress-Number-curve). One result, not unexpected, was that the mean life is longer the lower the stress level. With logarithmic scales on the axles of the diagram this sloping part of the diagram approximately becomes a straight line, which is an important result for analysis and design. Equally important was that there seems to be a specific stress level below which no fracture occurs and that this level is significantly lower than the yield stress. This level is now named the fatigue or endurance limit, s_e, and is a paramount design parameter. Finally, it was shown that there is a considerable dispersion in the lives for fixed stress levels. This dispersion is inherent in the fatigue phenomenon and means that consideration of stochastic effects is an important part of fatigue design.

Efforts to find the physical reasons for the fatigue behaviour involved both theoretical and experimental efforts by many researchers. Although it was found in early studies that thin bands were formed at the surface of cyclically loaded specimens, which subsequently developed into cracks and fracture, the reasons for the emergence of these bands were not clear. Models were developed based on irreversible motion and localization of dislocations into persistent glide bands, which then enabled the formation of extrusions and intrusions, microscopically small cracks at the material surface. The effects were studied by various types of microscopy, not least *scanning electron microscopy*, which became commercially available in the early 1960-ies. Also other mechanisms were studied, as inclusions acting as small internal cracks, and grain boundaries being obstacles for further development of microscopic crack growth. A sketch of a micrograph of a small crack-like defect formed from extruding glide bands is shown in Figure 2.3.

These micro-cracks, normally at the surface of the structure, can then cross grain barriers and develop into growing cracks and final fracture. The knowledge of the local nature of the start of fatigue damage is of great importance to explain such features of fatigue as the stochastic nature, the influence of material, specimen size, and surface treatment. However, the results give only principal information about differences between e g materials or surface treatments, and they do not give distinct quantitative results for design.

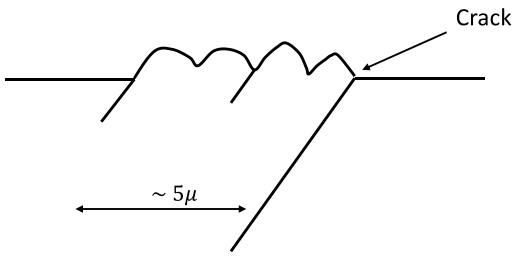


Figure 2.3. Sketch of a micrograph showing a developing microscopically small crack from a set of extrusions. Note the approximate 45-degree angle of the glide system. Micrograph from [2.1] where several similar examples of evidence can be found.

Not only the microscopic features on the grain level are of basic importance for fatigue. When small cracks have been formed, or in cases when they are present already from the start in a structure, as in welds, the understanding of cracks under cyclic load is essential. Here the establishment of *fracture mechanics* during the 1950-ies laid the ground for fatigue design with regard to cracks. According to linear fracture mechanics a parameter called the stress intensity factor, which can be calculated and tabulated from load and geometry, governs the conditions of material at a crack tip. From this Paris and co-workers in 1961 suggested that the stress intensity change during a load cycle could be used as a parameter related to the crack growth rate (growth per cycle). This principle was a break-through and has been used extensively for fatigue design, not least in the development of the fail-safe design principle for e g aircrafts.

Another important step was the introduction of *computer-aided servo-hydraulic testing systems*. They enabled the simulation in a laboratory environment of both load and deformation controlled cyclic loads, of non-sinusoidal waveforms and of irregular load sequences, such as those recorded from road transports or flights. The possibility to simulate virtually arbitrarily irregular sequences of load or strain has been and still is very important for development of theories and design principles where irregular loads are the cause of fatigue damage under multiaxial loading, e g in the car, airplane, and offshore industries.

Systems of several actuators, like the one shown in Figure 2.4, has been helpful in developing knowledge of fatigue loading. As seen, they can be used for testing not only of test specimens but also of structural components in full scale.

The possibility to control deformation, i e strain, by servo-hydraulic machines has been of importance for the understanding of fatigue where plastic strains are present, so-called low cycle fatigue. Such fatigue phenomena are typical for power plants and other applications with elevated temperatures.

Since the 1960-ies computerized stress and strain analysis, e g by *the finite element method*, has facilitated simulation of fatigue behaviour in complicated geometries and for materials with complex constitutive relationships in the non-linear range. With correct modelling of such things as kinematic hardening the material behaviour in

Figure 2.4. Full scale testing of a car with a servo-hydraulic system reproducing motions recorded from a service environment. (From Team Corporation UK Ltd.)

highly stressed parts can be obtained much better than with simple analytical models. The drawback is that modifications have to be made with regard to the large amount of data that has been collected for use in manual analysis. One example is the effect of stress concentrations.
Many other important results have contributed to the development of the field of fatigue. The choice of the ones mentioned above is the authors personal one.

2.2 Examples of major accidents due to fatigue

Even though fatigue has been known for nearly two hundred years and design methods have been developed and used systematically for around hundred years failures due to fatigue are still occurring frequently and causing loss of lives and severe economic consequences. This is due partly to the complexity of this seemingly narrow area of technology and partly to the stochastic nature of the phenomena. Here a few of the "classic" failures are related to demonstrate this and to show what can go wrong.

The Comet jet airliner
This was the first commercial jet airliner and had a considerable success 1952-1953 with many intercontinental flights although some crashes with fatal injuries occurred, which were thought to be caused essentially by human errors and adverse weather. During the spring 1954 two planes leaving Rome at different occasions, one in January and one in April, were lost in the Mediterranean. Naturally this caused an extensive investigation. It included the recovery and puzzling together of one of the crashed planes, and full-scale simulations in a pressure tank of a fuselage. It was found that fatigue cracks developed from the rivets around the corners of rectangular windows after around 3000 flight cycles (pressurizing and de-pressurizing of the cabin). The stresses at these corners were much larger than expected so when the crack tip reached the corner, rapid unstable crack growth ensued causing complete fracture of the cabin by several secondary cracks. The rivet holes had not been produced according to specification, which caused small initiation sites for fatigue cracks. So, the causes of these accidents were essentially

- *unsuitable design* of windows (rectangular instead of oval) causing increased stresses due to stress concentration

- *manufacturing* of rivet holes *in a non-intended way* that caused initiation sites for fatigue
- (possibly) *unexpectedly large load variations* in the pressurized cabin.

Each single one of these causes could perhaps have resulted in a still safe airliner, but together they became fatal. Detailed pictures of the situation can be found on the Internet, e g aerospaceengineeringblog.com, and Comet crash.

The Markham Mine accident
The transportation of miners up and down the 450-metre-deep mine shaft was accomplished by a winding system with two double deck cages at the ends of a wire, so that one cage went up when the other one went down. The winding drum was supplied with a DC motor and a brake shoe system braking by a compressed spring nest and opened by compressed air. Unfortunately, the force from the spring nest was transmitted by a brake rod loaded in tension.

Due to misalignment and insufficient lubrication large bending moments were introduced in the braking rod in addition to the (intended) uni-axial load. This meant so large total stress variations during braking/unbraking that a fatigue crack was initiated in the threads of the bolt at the end of the braking rod. When the final fracture occurred at a braking manoeuvre when 30 workers were transported down for a shift, there was no possibility to brake the system and the cage with the workers fell uncontrolled to the bottom of the mine. Eighteen were killed and eleven injured. The material in the rod was according to specifications and the uniaxial stresses, without the unintended bending stresses, would not have been dangerous. The causes of the accident were:

- Unintended bending stresses introduced through *unsuitable design and neglected maintenance* of the joint between brake rod and lever (a deceptively simple part of the system).
- Total dependence of a critical system on *one part loaded in tension*, without *regular inspections*. The system could as well have been designed so that the spring nest was working directly on the lever.

Details about the accident can be found in [2.2].

The Enschede train disaster.
In June 1998 a wheel in the first car of a high-speed train between Munich and Hamburg fractured near the town of Eschede. This would not necessarily have caused a major accident, but a series of events occurred making the fracture result in one of the most deadly train accidents ever, costing 101 lives and injuring 90 people seriously.

After the fracturing of a wheel the train de-railed at two consecutive switches causing the third and following coaches to enter a parallel track. These violent motions caused a collision with the foundation of a bridge over the tracks, which fell down and made the rest of the coaches stop in a zig-zag fashion that caused many deaths.

The wheels were of a new type using a rubber damping ring between the rail-contacting steel tire and the wheel body in order to give a more smooth and comfortable running of the passenger coaches.

Thorough investigations after the accident showed that the tyres were bent into an oval shape during rotation and rail contact. Particularly after some wear, making the tyres thinner, this induced periodic bend stresses large enough to initiate fatigue cracks. Before the accident no serious trouble had been experienced, although some concerns and warnings had been issued regarding the specific design with rubber rings and steel tyres. Before putting the wheels into operation only limited tests had been performed.

However, the main causes of this catastrophe were (author´s view):

- The new wheel design was *insufficiently tested* and verified for the high-speed service and wear conditions.
- *Warnings*, 1992 and 1997, *were neglected*; possibly partly because the new wheel design had improved the comfort for the passengers considerably.
-

Several pictures from the accident and of the wheel can be found on the Internet searching for "Eschede train disaster".

References

[2.1] Suresh S. "Fatigue of materials", Cambridge University Press (1992), ISBN 0 521 43763 6.

[2.2] Demaid, A.P.A and Lawley, A. "The Markham Mine Disaster", in ASTM STP 918 (1985)

3 Micro- and meso-mechanical background to the understanding of fatigue

3.1 Model for initiation of fatigue damage

Much knowledge has been obtained from the 1950-ies and onward by studying materials on the microscopic scale. The fundamental features found are *inhomogeneity* of the material, e g in the grain structure in metals or fibres in composites, and the presence of *imperfections*.

When a piece of inhomogeneous material is exposed to mechanical load the stresses and strains induced will also be inhomogeneous. Due to the production process, there are local, self-equilibrating stresses and strains present. This means that plastic, irreversible strains will occur locally, e g in grains, at stress levels well below the macroscopic yield stress.

For a discussion of fatigue initiation in metals, *dislocations* play an important part. Dislocations are line defects, and they are always present in the grain structure of metals. The main types are called edge and screw dislocations. In other types of materials than metals fatigue is initiated and leading to failure by analogous mechanisms, also building on accumulation of damage from in-homogeneities and imperfections on the microscopic level.

Metals normally have a body centred (BCC) or face centred (FCC) cubic structure. Due to the ways atoms are stacked there are planes and directions where slip deformation can occur more easily when a shear stress is applied, so-called preferred slip planes. The dislocations facilitate this deformation.

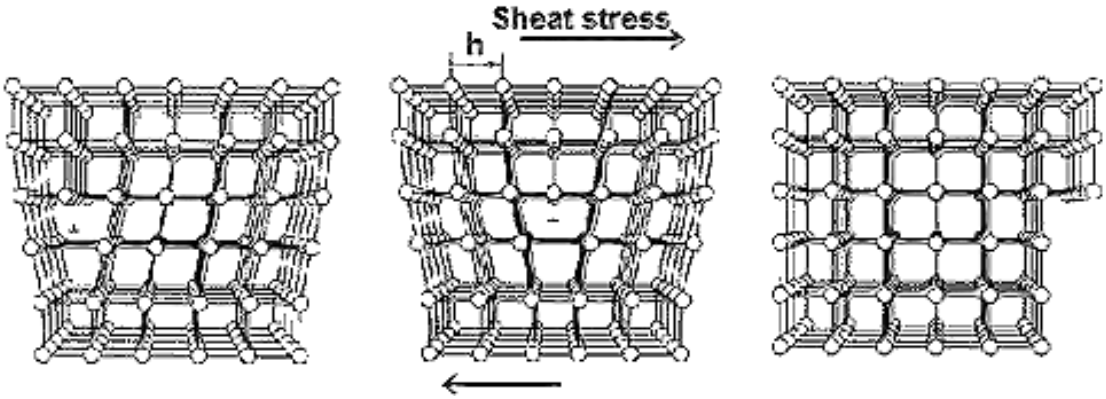


Figure 3.1 Edge dislocation in a cubic array, showing the motion through the array. The Burgers vector is denoted by h. (From chemistry.tutorista.com.)

An edge dislocation, which is the simplest form to use for demonstration of the principal features, is shown in Figure 3.1. It can be seen as an extra half plane of atoms. The end of this extra half plane of atoms is conventionally denoted by a "T" symbol, and the direction and unit step length of motion of the plane is called the Burgers vector, $\vec{b}$ (h in the Figure). At the edge of the dislocation the local deformation of the atomic array means a strong disturbance of the local binding forces, and that the dislocation can move easily along the slip plane if a shear stress of sufficient magnitude is applied as indicated in the Figure. For a monotonic load this will just lead to a homogeneous permanent displacement.

The general situation is much more complicated. There also exist so called screw dislocations, meaning that, instead of the extra half plane in Figure 3.1, atoms to the upper right are displaced one atom spacing perpendicularly to the plane of the figure, creating a "spiral ladder" structure along a line through the "T". Dislocations may create loops in the material appearing as both screw and edge type dislocations along different parts of the loop. Since there are twelve equivalent slip planes in an FCC structure, loops of edge and screw dislocations in the different plans may interact when the crystal is deformed.

When the external load is cyclic this interaction is complex and *irreversible* changes occur. New dislocations are created, and they may lock or annihilate each other when they interact in the different slip planes, resulting in e g hardening. Hence, it should be difficult to formulate a consistent and quantitatively precise theory for fatigue initiation under random, multi-axial external loads in a component. The models emerging serve as a basis for understanding and order of magnitude quantification.

The role played by dislocations in initiation of fatigue has been studied by shear loading of single crystals, often of copper. Detailed knowledge about the phenomena has been established by many researchers in material science, using SEM and TEM electron microscopy and by the development of dislocation theory. The state of the art can be studied in numerous textbooks as e g the one by Suresh [3.1], and its references. The knowledge is not yet complete. For an understanding of the major features of fatigue initiation it is useful to consider a simplistic, yet essentially correct model.

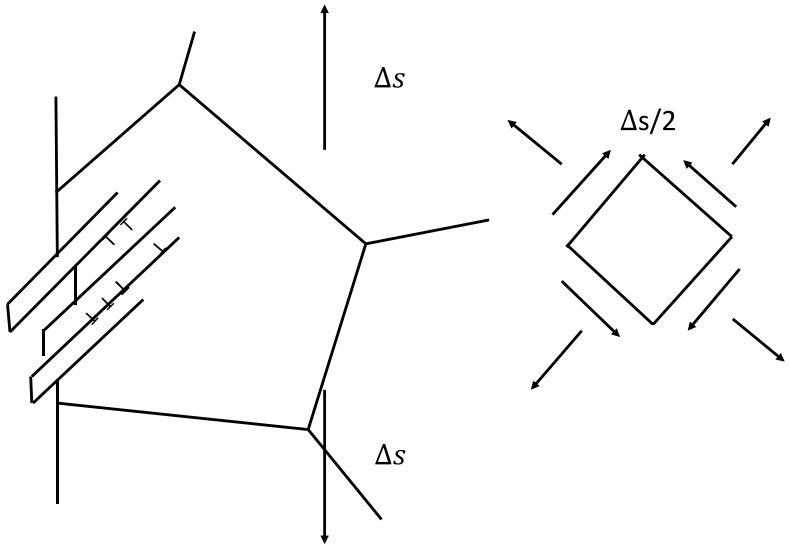


Figure 3.2 Formation of slip bands and extrusions/intrusions in a grain. The "T":s denote dislocations.

So, a conceptual model for initiation of fatigue damage in a component or test specimen is the following. Consider the surface of a material element exerted to reversed uniaxial loading, Δs, according to Figure 3.2.

At some positions, as in the Figure, there are grains with preferred slip planes at approximately 45 degrees to the load direction, i e where the resolved shear stress is maximum.

When the magnitude of the reversed stress is sufficient, there appears to build up a concentration, and subsequent *saturation*, of dislocations located at so called persistent slip bands with a thickness of some hundred atoms, and where the dislocation density is of the order of 10^{15} m^{-2}. Between those there will remain slices of material with a lower dislocation density around three orders of magnitude less, and some thousand atom layers thick. The way to saturation and the process for formation of persistent slip bands are not completely explored theoretically but the effects are amply verified experimentally by microscopy.

It is predominantly a surface phenomenon since material elements at the surface can deform more easily and are subjected to other conditions, as surface damage, than the interior. Exceptions from this is in contact mechanics, where sub- surface maximum shear stresses are present, and where faults like big inclusions cause severe local inhomogeneities inside the material.

Following the saturation of the persistent slip bands, extrusions and intrusions start to form during continued cycling. The grain can be completely damaged by a micro-crack the size of a grain, when the intrusions become large enough and the persistent slip bands break up.

Extrusions and intrusions form since the slices between the slip bands start moving irregularly in relation to each other under the reversed loading. A model for this development built on the random distribution of displacements between slip bands was presented by May, giving the frequency f of intrusions with the depth z and width w as a function of cycle number N as

$$f = F\exp\left[-\left(\left(\sqrt{z+w}-\sqrt{w}\right)^2/b\gamma N\right)\right] \qquad (3.1)$$

and where F is a slowly varying function of z and N of the order of unity, b is the magnitude of the Burgers vector and γ is the plastic shear strain. To get an impression of the meaning of (3,1) the following typical values are inserted. f/F is put to 0,1 (the value in the bracket does no change much for other choices). z is taken as a grain size of $5\ 10^{-6}$ m, and much larger than w, forming "grain size cracks", b is set to $3\ 10^{-10}$ mm, the atomic spacing of steel, and γ is $3\ 10^{-3}$ the size of strains at yield. Then N becomes $2{,}4\ 10^{6}$ cycles, i e there is a reasonable probability for a damaged grain. So, it is seen that near the limit of macroscopic yield, and with realistic data for steel, the initiation phase of a fatigue crack may take up a substantial part of the life of a fatigued component.

The model is analogous to the situation where a deck of cards is placed on a table and the top card is moved back and forth. Each motion is distributed randomly among the cards. Then, after a number of motions it is probable that some cards have extruded much more than others. Another analogy is an opened package of spaghetti straws on a shelf in a caravan. After some miles on a bumpy road the probability is high that some straws "extrude" quite a bit.

Although May´s model and other, later ones, are simplistic and neglect obvious effects, as the fact that total randomness is unlikely, they give a basic understanding of the mechanisms creating the first parts, the "birth" of a fatigue fracture.

The development of persistent slip bands and extrusions/intrusions of lamellae has been amply verified by experiments. Slip bands have been observed by light microscopy to develop as fine lines at the surface of copper specimens being plastically deformed in reversed tension for 10^4 – 10^5 cycles. After a light polishing they seem to disappear, but after continued cyclic

loading they quickly reappear, meaning that they existed in the material just some microns below the surface. Extrusions have been observed by many researchers, e g [3.2]

For the fatigue process to continue into the growth of a bigger crack the grain boundary barriers have to be overcome. Neighbouring grains may have their preferred slip planes in other directions and they may be exposed to smaller inherent stresses. The situation is also a three-dimensional one; there are grains both in the plane of Figure 3.2 and out of this plane. But if the external load is big enough the barriers can be overcome until a thumbnail crack with a radius the

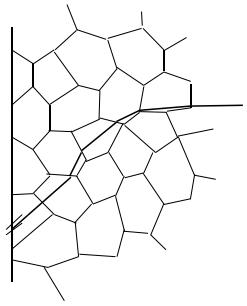


Figure 3.3. Microscopically small crack and its further growth.

size of several grain diameters is created. Finally, this very small increment of a crack grows and bends off, forming a real small crack according to the principles of fracture mechanics, see Figure 3.3.

Although the model does not reflect all the complex phenomena that may occur, the principal features are experimentally verified and may be regarded as a fundamental model. Some important conclusions may be drawn.

- Fracture of structures due to initiation and growth of fatigue cracks may occur at cyclic loading corresponding to stress levels well below the yield stress.

- Since initiation is local and on the microscopic level, and depending on orientation of grains, barriers between grains etc, it is natural that stress levels necessary for fatigue initiation, growth rates for subsequent fatigue cracks, and lives to structural fracture in terms of number of cycles, all have a statistical scatter which is much more prominent than in other areas of solid mechanics.

- This also means that size effect are to be expected, i e smaller specimens appear to be stronger than bigger ones with regard to cyclic stress since the probability for weak sites is bigger the bigger the specimen.

- The surface of a structure under cyclic loading plays an important role. The roughness, grain size, surface treatment, hardening and other properties interact with the formation of slip bands and intrusions. So, for hard materials, usually having smaller grains, a coarse surface from the production process plays a more important role than for low strength materials.

- A constant tensile stress superposed on the cyclic one lowers the cyclic stress necessary for initiation (called the endurance limit). It is natural to assume that the tensile stress facilitates

the slip in the persistent slip bands, and their breaking up into grain sized micro cracks. Also, the subsequent crack growth is favoured by the crack opening effect (see Figure 3.2 and 3.3) of a constant tensile stress. In the sequel the importance of such stresses (called mean stresses) will be extensively discussed.

- For stress levels above those necessary to initiate and grow a fatigue crack, the relationship between the number of cycles to initiation and the number of cycles for growth to fracture varies. For higher levels the initiation phase is shorter, and the main part of the life consists of crack growth. Near the limit for fatigue to occur most of the life may be due to initiation. For high stress levels, where macroscopic plastic strains occur (as in low cycle fatigue), the statistical effects are smaller, as are effects of surface treatment, since material damage is not localised but occurring in the bulk of the material.

As mentioned, initiation of fatigue may occur in other ways, and by other mechanisms. One such is the breaking up of grain boundaries in hard materials, where this mechanism is easier to activate than slip inside the grains. Still, the common features are the inhomogeneity of the material, the occurrence of faults on a microscopic scale, and some phenomenon of irreversible deformation during cycling, which creates accumulation of damage.

The mean of the stress level at reversed loading of polished specimen when fatigue can just occur is called the fatigue or endurance limit, s_e.

3.2 Physical background to crack growth after initiation

So, a small irregular thumbnail shaped crack may be created by slip essentially in 45 degrees to the uni-axial tensile stress in the material surface, and with an extension of several grains both along the surface and perpendicularly to it.

When a small crack-like geometry is present it is known that the continued growth should occur mainly through the influence of the tensile stress perpendicularly to the direction of the crack. Consequently, the small crack bends off as indicated in Figure 3.3 until it can grow under symmetric, so-called mode I conditions, i e perpendicularly to the largest principal stress. Note again that this is a three-dimensional process so that the thumb-nail crack has to adjust in a complex way to the new plane of growth. The conditions for the crack to bend off and continue to grow as a macroscopic crack to failure of the structure are obviously depending of several features on the micro-scale. This adds to the statistical scatter in the results of fatigue testing, both the fatigue limit and the life obtained for load levels above it.

Also, when a macroscopic crack has emerged, through the initiation process described above or directly as a consequence of manufacturing, e g welding, the growth to fracture under continued cyclic loading is a result of micro-structural features, now in grains at the crack tip. For the discussion of these phenomena, it is suitable to use some of the concepts of fracture mechanics, which are developed in more detail in Chapter 7.

At a crack tip nominal stresses and strains are highly intensified, meaning that yielding and plastic deformations occur locally. These phenomena may be summarized in terms of the stress intensity factor K, or for cyclic loading the variation of K, ΔK, during a cycle. The stress intensity factor is a coefficient, proportional to the nominal stress, of an expression describing the stresses and strains just at the crack tip. For each loading mode, see Chapter 7, as the

symmetric mode I of main importance for growth, these stresses and strains are always the same.

The physical features at the crack tip, which are the main issue here, are thus discussed in conjunction with the stress intensity factor. As opposed to the initiation phase, where the mechanism is sliding back and forth under *reversed* shear stresses, it is customary to treat crack growth in terms of *zero to tension* loading, nominally keeping the crack open during the process.

Different stages may be distinguished.

The threshold level
Just as in the initiation phase, where fatigue damage may not start and grow beyond the first grain for low stresses, there is a stress level at a crack tip below which the crack does not propagate. This is called the threshold level and the corresponding K-level is denoted by ΔK_{th}.

It is natural to assume that a connection exists between ΔK_{th}, characterizing the plastic deformations at the crack tip, and the grain size, and an indicative calculation can be made. Since ΔK_{th} is strongly influenced by crack closure, see below, ΔK_{th}-values for high values of the parameter $R = K_{min}/K_{max}$, when no closure occurs during the load cycle, should be chosen for the analysis. It appears that for many steels this value is around 2-3 MPa$\sqrt{m}$ and not much dependent on the yield stress of the material. The stress intensity factor corresponding to a plastic zone the size of one grain, d, is found from (see Chapter 7)

$$d = \frac{1}{\pi}\left(\frac{\Delta K_{th}}{2\sigma_y}\right)^2, \qquad (3.2)$$

where σ_y is yield stress. Re-arranging this becomes

$$\Delta K_{th} = \sqrt{4\pi}\sigma_y\sqrt{d} \qquad (3.3)$$

For a typical low strength steel with σ_y = 400 MPa d is around 5 microns ($5\ 10^{-6}$ m) which gives

$$\Delta K_{th} = 3{,}2 \text{ MPa } \sqrt{m} \qquad (3.4)$$

This seemingly precise agreement is a bit misleading since there is a considerable variation in material characteristics, and since a continuum analysis of plastic zone size has been compared with flow conditions in an inhomogeneous grain structure. Still, there is a clear indication that the threshold condition is connected to the fact that plastic deformations are limited to a region of the order of the grain size.

Further, there is a general, approximate, relationship between grain size and yield stress of the form

$$\sigma_y = A + Bd^{-1/2}, \qquad (3.5)$$

i e

$$\left(\sigma_y - A\right)d^{1/2} = const, \qquad (3.6)$$

indicating by comparison with the formula (3.3) for ΔK_{th} that there is, as expected from experiments, a weak dependence on threshold strength for wide groups of materials.

So, for existing cracks, ΔK_{th} plays a role similar to the fatigue limit s_e, determining whether cyclic loading will lead to failure of the structure. In some instances when initial cracks can be foreseen or detected, design with regard to the threshold value of the stress intensity factor is used.

Stage I crack growth
Just above the threshold level the extension of plasticity is a few grain diameters and then the cyclic loading causes formation of damage locally in grains. Again, it is sliding along slip bands that occurs, creating saturation and separation during a rather large number of cycles, until the grain is damaged. See Figure 3.4 a). When a grain is damaged, or slightly before, the barrier to the next one is broken, and the process continues in a zigzag fashion. It is to be observed that this occurs in many places along a crack front, and then the growth of the crack is uneven and in steps, giving a ragged crack surface.

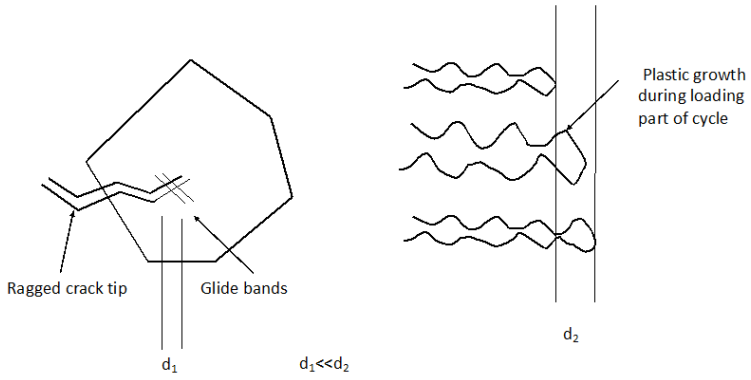


Figure 3.4. Sketch of formation of damage at a growing crack tip. a) for small load levels, b) for larger load levels (i e stress intensity factors). Note that d_2 is much larger than d_1, often notable parts of a grain diameters.

When the level of ΔK increases, the plastic sliding in the grains at the crack tip is more intense. A blunted crack tip results, now producing increments of crack growth in a complex way. This is called stage II crack growth. There is of course an intermediate stage, where both processes occur in parallel in different grains along the crack front that is not possible to visualize by a clear-cut model.

Stage II crack growth
When the cyclic stress intensity factor is high enough there is a significant plastic deformation in a zone at the crack tip, and a blunted crack tip is formed at the maximum loads of a cycle. A simple conceptual model is that a continuum mechanics situation is approached, where the reversed sliding of glide bands is superseded by plastic deformation of new areas in the grain in each cycle, see Figure 3.4 b.

Then crack growth takes place by the successive blunting and compression of the crack tip at each cycle of loading. The result of the process is visible as ripples, so called striations, which form wave patterns at the fractured grains. For even higher K-values the striations cover several grains and may be observed by a low-magnification microscope, or even by the naked eye for high loads.

To get an impression of the parameter values involved in such a model it is observed that an approximate relationship from fracture mechanics is

$$K^2/E = \sigma_y \delta , \qquad (3.7)$$

where K is the stress intensity factor at static loading, E is Young´s modulus, σ_y is yield stress, and δ is the crack tip opening, the blunting. Considering closure effects and cyclic conditions as hardening it has been suggested that the growth per cycle could be written (the notation da/dN is conventional, see Chapter 7)

$$\delta = \frac{da}{dN} = \beta \frac{\Delta K^2}{\sigma_y E} \qquad (3.7.b)$$

Here the cyclic yield stress is used and β is a compensating parameter. The fit is not very good, however, since comparisons with data for structural steels show that β is of the order of 0,05 - 0,1, while it should be expected to be more close to unity. Further, the slope of the relationship between da/dn and ΔK in a log-log diagram is 2, while many experimental data give a slope between 3 and 4, see Chapter 7.

More elaborated models consider accumulation of cyclic strain during the fatigue process when the crack grows. Material elements are successively approached by the crack tip under increasing strain changes during each cycle, and models for such processes lead to formulas of the type

$$\frac{da}{dN} = \frac{\Delta K^4}{\mu \sigma_y^2 U_*} \qquad (3.7.c)$$

Here μ is the shear modulus and U_* is a measure of accumulated plastic strain energy of the material up to the moment when separation of surfaces occurs. As seen this model predicts a slope in the growth curve of 4, which is more in agreement with experiments.

Although there are discrepancies between models for fatigue using microscopic findings, and experimental macroscopic data, the reasoning above may be useful for the principal understanding of the underlying phenomena.

Stage III crack growth.
Striations can be observed by the naked eye on surfaces of fractured specimens, see Figure 3.5. It should be observed that these shell-shaped patterns, which may be very helpful for the analysis of failures, are much larger. They are of the order of 0,01-0,1 mm produced during the very last percentage of cycles before failure, when the maximum K-value during a cycle

approaches the fracture toughness, or as effects of overloads during the process of fatigue loading.

A schedule for failure analysis from [3.3] is shown in Figure 3.5. Here, it is illustrated how it is possible to detect where the fatigue crack growth started, and hence where analysis by microscopy can be made to detect possible initial material defects. Further, it is possible to see whether the load has been low (small remaining surface of final fracture) or high, and whether the load has been in tension or (un-intentional) bending, etc. Such analysis is very useful to get a quick first understanding of the cause of fatigue failure.

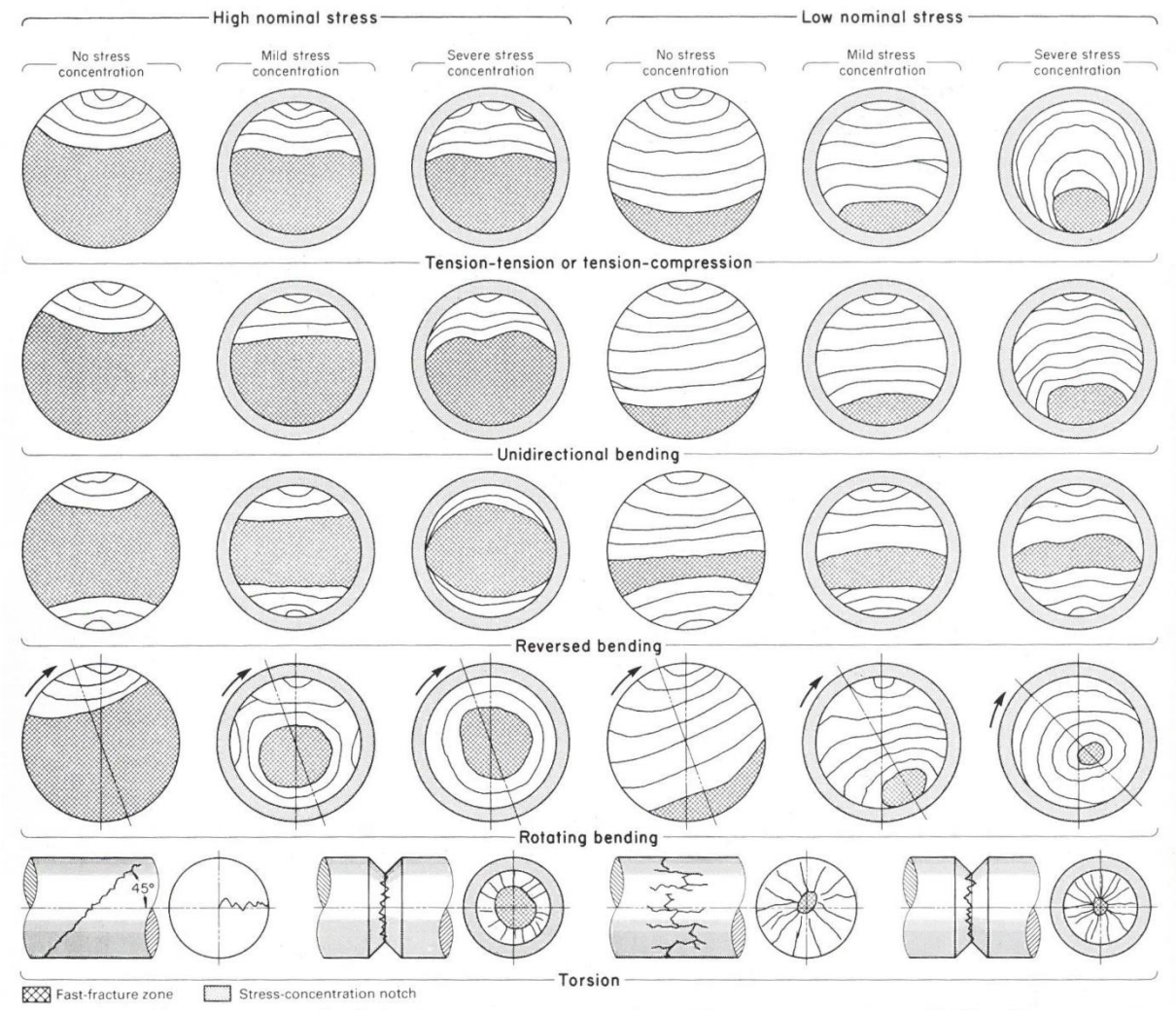


Figure 3.5 Map of striation patterns showing how they can be used for failure analysis [3.3].

3.3 Effects of crack closure

Just as in the case of initiation, the degree of mean stress in the load cycles influences the crack growth. Nominally a minimum nominal stress > 0 should keep the crack open during the whole cycle, whereas minimum stresses < 0 should mean that the crack is closed during part of the cycle, not influencing fatigue damage. This simple view has been shown in many studies to be wrong and in need of modification. There are several mechanisms giving crack closure before the minimum stress is reached in the crack tip region also when this is > 0. One important factor is that the plastic deformation during load cycling and incremental growth means that ripples of material emerge from the reversed blunting, which are mismatched and get into contact

during unloading. There are also factors like oxidation and other chemical influences that contribute to closure effects. This means that there is a mean value influence also on fatigue crack growth, which has to be considered in design, and this is discussed further in Chapter 7.

3.4 Pre-existing cracks or crack-like defects

There are several ways apart from initiation from slip bands in which fatigue cracks may be created in metals. Two of these are inclusions and welding.

Inclusions, as sulphides, may be relatively large as compared to the grain size and they have a different stiffness from, and are loosely connected to the surrounding grains. They are also irregularly shaped, from nearly spherical to more or less penny-shaped ones due to rolling. When cyclic loading occurs, the bonding is first dissolved so that the inclusion area effectively acts as a hole, with geometric stress concentrations, see Figure 3.6. Continued cycling means that small cracks may be created at the edges. When the conditions are severe enough these cracks-like defects may grow further in the fashion described above until final failure of the component. Inclusions may be the cause of sub-surface cracks, e g in rolling contact fatigue, where the shear stresses are large some distance below the contact surface.

Figure 3.6. Sketch of crack formation from an inclusion.

Due to thermal stresses from cooling and the very inhomogeneous material structure there are generally small cracks at welds, with lengths in the range of a few tenths of a millimetre. Their growth under cyclic load is very difficult to analyse but of great practical significance. Therefore, special methods have been developed in standards and recommendations for design of structures, as lifting devices, buildings and pressure vessels, see Chapter 13. The principle is to treat the weld as a "black box" the properties of which with regard to the relationship between life and nominal load have been determined by comprehensive testing. These results are then used together with design schemes including a number of safety factors and modifying factors for e g size, surface treatment and loading conditions.

3.5 Short cracks

There is a region of crack lengths, between the micro-structurally small ones in grains and those that are large enough to be treated by fracture mechanics, which is of interest to study. Some approaches that are frequent in the literature are the following.

The Miller representation.
In order to give a consistent and comprehensive description of fatigue phenomena Miller [3.4] presented the concept of micro-structurally small cracks, physically small cracks, and macroscopic cracks.

In terms of Figure 3.7 from [3.4] the fundamental types of cracks are cracks corresponding to the intrusions in single grains (10^{-6} – 10^{-5} m), cracks in the transition region (typically 10^{-4} m),

and cracks which can be treated by linear elastic fracture mechanics (larger than 10^{-3} m) respectively. This puts the relationship between grain size, crack length, life and cyclic stress amplitudes into perspective. High strength materials, for instance, with small grains correspond to smaller cracks and hence require higher stresses to develop physically small cracks through the barriers. This results in a higher endurance limit. On the other hand, once developed the macroscopic cracks grow faster under the high stresses.

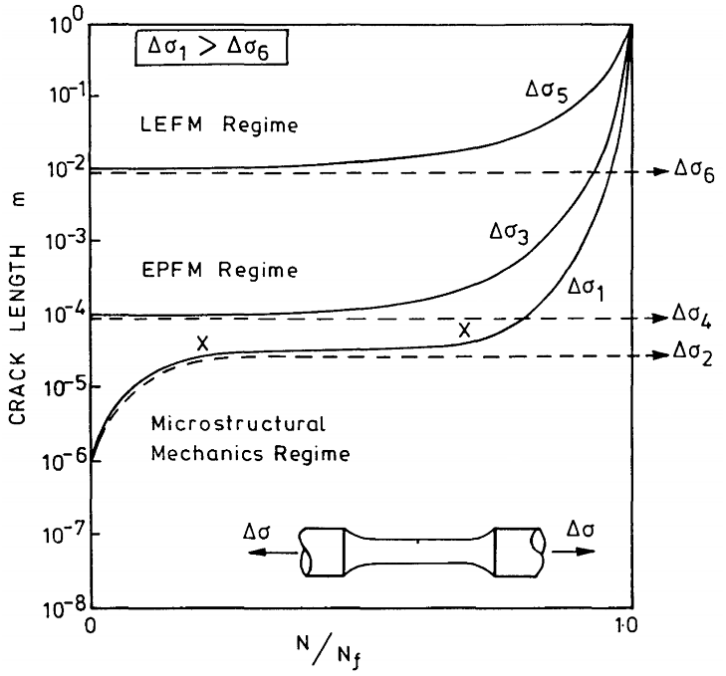


Figure 3.7. Representation of the fatigue fracture process according to [3.4].
(Note that N_f differs for the different stress levels.)

In Figure 3.7 stress amplitudes decrease from $\Delta\sigma_1$ to $\Delta\sigma_6$. Just above the fatigue limit, $\Delta\sigma_1$, a micro-structurally small crack grows by the sliding of glide bands through the first grain barriers, then struggles its way as a physically small crack for most of the life N_f, of the order of 10^6 cycles, and finally grows to structural fracture as a macroscopic crack. The slightly smaller level $\Delta\sigma_2$ is not sufficient and the crack arrests at some grain barrier.

A similar reasoning can be applied starting with a physically short crack (as at an inclusion or weld crack) for the stress levels $\Delta\sigma_3$ and $\Delta\sigma_4$. For these cracks it is not feasible to use linear fracture mechanics for analysis, since the amount of plasticity is too big as compared with the crack length.
The last pair of curves comprise the growth of macroscopic cracks with $\Delta\sigma_5 > \Delta\sigma_6$. Here, $\Delta\sigma_6$ corresponds to the threshold ΔK_{th}. They correspond to the crack growth curves utilized in Chapter 7.

The Kitagawa representation
(For comparability between stresses and stress intensity factors, stress entities in this section refer to R=0, instead of R=-1, changing the definition of nominal endurance limit Δs_e etc.)

Typical lengths of the different crack types and their characteristics can also be assessed

approximately by a reasoning using Figure 3.8 a), with log stress range and log crack length on the axles.

For macroscopic cracks linear elastic fracture mechanics, see Chapter 7, yields a straight line with slope -1/2 as a limit for crack growth from

$$da/dN = C(f\Delta s(\pi a)^{1/2} - \Delta K_{th})^n \ . \qquad (3.8)$$

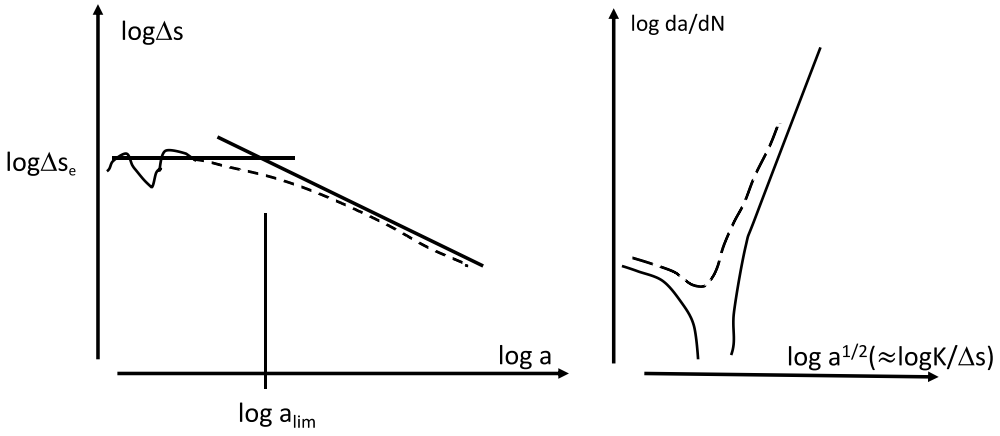


Figure 3.8. a) Relationship between stress level and crack length. b) Retardation and arrest of small cracks. (The Figure 3.8.a is generally named after Kitagawa).

Here da/dN is crack growth per cycle, and C and ΔK_{th} are material parameters. The shape factor f is of the order of 1. The straight line to the right is obtained by setting da/dN=0, thus meaning higher and higher crack growth rates to the upper right of the diagram. Hence, the line represents the border for crack growth.

The endurance limit, Δs_e , corresponds to very short cracks, grain initiation, near the vertical axis of the diagram. By inserting this in (3.8) for da/dN=0 a coarse measure of the border between physically small cracks and micro-structurally small cracks is obtained as

$$a_{lim} = \frac{1}{\pi} (\Delta K_{th}/\Delta s_e)^2 \qquad (3.9)$$

This can be combined with the formula for grain size related to ΔK_{th} (3.3) which gives an estimate of the relationship between grain size and size of physically small cracks as

$$\frac{a_{\lim}}{d} = \left(\frac{2\sigma_y}{\Delta s_e}\right)^2 \qquad (3.10)$$

With $\Delta s_e \approx 0.5\sigma_y$ it is seen that a_{lim} becomes some ten grain diameters. Significantly below this length the micro-structurally small cracks grow through barriers near the endurance limit, indicated by the wavy part of the curve, and for somewhat longer lengths the physically small cracks grow under non-linear fracture mechanics conditions, the slashed part of the curve. When the length becomes such that the linear part of the curve is approached the cracks can be

treated as macroscopic, and by linear fracture mechanics. See below under “Growth of short cracks”.

These effects are sometimes represented in terms of the Paris diagram (see Chapter 7), with crack length instead of stress intensity factor on the horizontal axis, Figure 3.8 b. For stress levels just below the endurance limit, the full line curve, the micro-structurally small cracks may retard and stop at barriers, while the longer ones may start and pick up speed above some threshold length in the vicinity of the border between physically small and macroscopic cracks, i e approximately when

$$K_{Ith} = f\sigma_e\sqrt{\pi a_{border}}\,, \qquad (3.11)$$

where, again, f is a geometry parameter of the order of unity. For stresses above the endurance limit, the dotted curve, there is just some retardation after which the crack gains speed and follows the Paris curve of growth.

Non-growing short cracks
There are cases at notches with high stress concentrations K_t, see Chapter 4.6, when cracks initiate and start growing, but then retard and stop. See Figure 3.9.

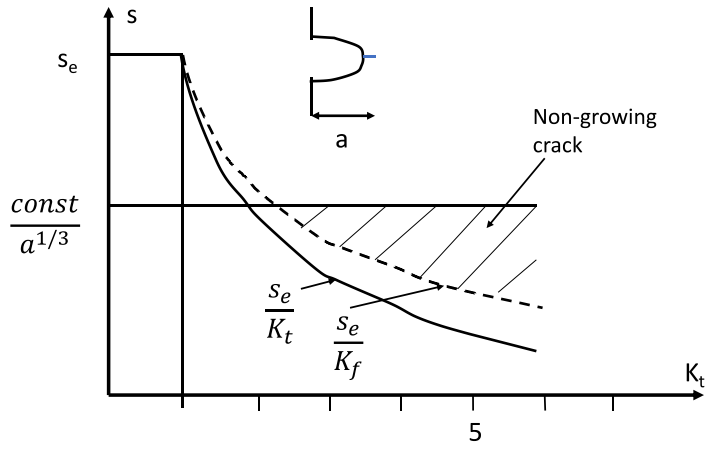


Figure 3.9. Diagram showing the region of non-growing cracks at high stress concentrations.

The expectation is that a crack initiates and starts growing to failure when the local stress amplitude exceeds the endurance limit s_e

s_e/K_t, or, in fact, s_e/K_f, dashed curve. (3.12)

Instead, there appears to be a level, defined by the notch depth, a, from

$$s = const/a^{1/3} \qquad (3.13)$$

for which a crack starts at the notch tip but arrests. Seemingly the stress concentration fades out above some value depending on the geometry. This is of some significance for structures with high stress concentration factors like screw threads and gear teeth.

The formula (3.13) stems from experiments. Actually, the arrest of very small cracks formed at the root of sharp notches can be qualitatively explained by assuming that the sharp notch plus the crack increment, see insert in Figure (3.9), acts as a crack of length a. Then is does not grow if the threshold value of the stress intensity factor (see Chapter 7), the stress level and the notch depth fulfil the relationship, for an edge crack,

$$s < \frac{K_{th}}{1{,}12(\pi a)^{\frac{1}{2}}} \approx \frac{K_{th}}{2a^{1/2}}$$

which can be compared to (3.13).

There is an apparent anomaly in such an assumption since it can be shown that the formal stress intensity factor for the crack increment increases, theoretically, for all values from zero.

Growth of short cracks from defects
The dotted part of Figure 3.8 a) should be seen as a bridge between initiation of microcracks at the endurance limit, s_e and the long crack growth area, where cracks are growing under the conditions where $K_I > K_{Ith}$ governs the growth, see Chapter 7. The parameters s_e and K_{Ith} are material parameters (dependent on R, here R=0). Two such bridging functions have been suggested by Murakami and Endo, and El Haddad et al.

Murakami and Endo made fatigue experiments with defects in the form of small, drilled holes and found that crack growth commences if the stress exceeds

$$s_{eM} \geq \frac{C}{Y(\sqrt{area})^{1/6}} \qquad (3.14)$$

Here, C/Y is 2,2(HV+120) for surface defects and 3,1(HV+120) for internal defects for several materials. Further, the square root of area plays the part of a fictive crack length calculated from the area of the defect in a plane perpendicular to the stress, see Figure 3.10 a). So, this means that a crack starts to grow from a defect if (3.14) is fulfilled, and according to Paris´ law (see Chapter 7) with a stress concentration factor based on the fictive crack length. The Figure 3.10 a) intends to visualize how a small crack initiates and starts to grow from the tip of a defect with extension also perpendicular to the surface of the Figure, i e some “half sphere”.

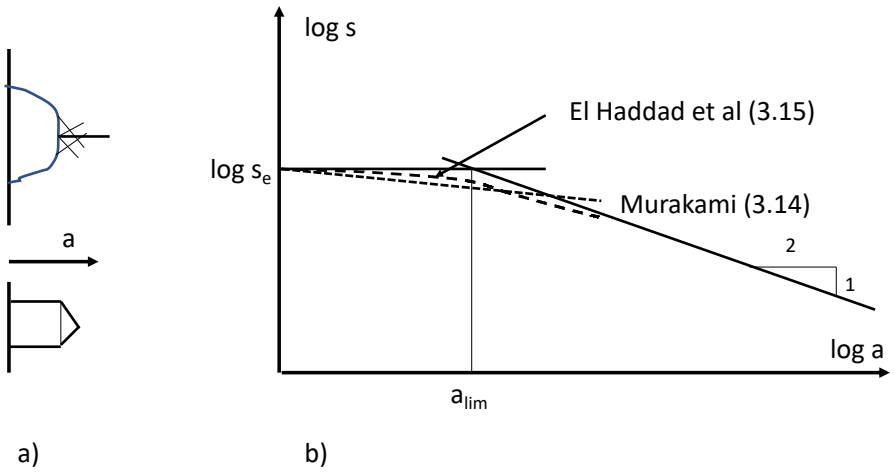


Figure 3.10 Sketch of a surface defect from which a real crack emerges and grows, and a drilled hole, a), and the two short crack bridging approximations, b).

El Haddad et al, based on experiments with artificially produced short cracks, found another expression for the bridging endurance limit

$$s_{eE} \geq s_e \left({a_0}/{(a + a_0)}\right)^{1/2} \qquad (3.15)$$

with a_0 according to (3.9) and a as the crack length. The two suggested bridging endurance limits are sketched in Figure 3.10 b).

References

[3.1] Suresh, S. "Fatigue of materials", Cambridge University Press (1992), ISBN 0 521 43763 6.

[3.2] (Boettner, R.C., McEvily, A.J., Liu, Y.C.: On the formation of fatigue crack. Phil. Mag. 10, 95 (1964)).

[3.3] ASM International, Metals Handbook, Vol. 12, Ninth Edition, "Fractography"(1987), ISBN 0-87170-007-7.

[3.4] Miller, K. J. "Materials science perspective of metal fatigue resistance", Materials Science and technology, vol 9, pp 453-462 (1993)

4 Fatigue properties observed on the macroscopic level

Studies of the causes of fatigue damage on the microscopic level in Chapter 3 explain why initiation and growth of cracks may occur for periodic stresses well below the yield limit. The analyses also indicate that size, mean stress, and surface conditions may influence the fatigue properties. In order to be of use for design purposes all those effects have to be quantified on the macroscopic level by experiments. The results of experiments then have to be structured into models enabling the formulation of design rules in the form of graphs or mathematical relationships.

The nature of fatigue phenomena means that the relationships have considerable statistical scatter; the properties are stochastic variables. Further, formulas used in fatigue are approximations built on best fits to experiments and they do not express true physical relationships. This adds to the uncertainty, particularly when a "rule" is used for conditions outside those for which it has been established.

In this chapter fatigue properties are explored that are traditionally used in design for long lives or for the avoidance of fatigue at all. They are put together in Chapter 5 outlining the classical principles for stress-based design.

4.1 Basic fatigue properties

As mentioned in Chapter 2 some of the first systematic experiments to understand fatigue on the macroscopic level were performed by Woehler to obtain fatigue properties for axle specimens exerted to rotating bending. The periodic maximum stress is called the *stress amplitude* at rotating bending, σ_{ar}. The limiting stress level for which fatigue damage does not occur defines a general material property for a set of well-defined conditions, the fatigue limit s_e. Since the conditions are idealised, with polished surfaces and test specimens with a diameter between 6 and 10 mm, modifications have to be made to adjust for such things as other load cases, sizes, and surface treatment in actual applications.

However, the experiments performed by Woehler and his co-workers had such an impact that much of the present nomenclature stems from them. The principal results are now known as Woehler or S-N- (Stress-Number-) diagrams, see Figure 4.1. They are still produced to describe material properties, but with more sophisticated equipment allowing more or less arbitrary load cases, temperatures and frequencies of loading. Conventionally, logarithmic scales are used in the resulting graphs. This appears helpful in giving results an analytical form, but it is also slightly deceptive since the graphs give a too smooth and deterministic impression of the physical reality they really represent.

These first results from rotating bending tests had a sinusoidal time function of the loading. From the underlying phenomena of formation and motion of glide bands on the microscopic level it is natural to assume that the shape of the time function is insignificant, and that it is only the maximum and minimum values that are of importance for fatigue damage. Also, the frequency of the cycling should be of virtually no importance if it is not very high, in the kHz range.

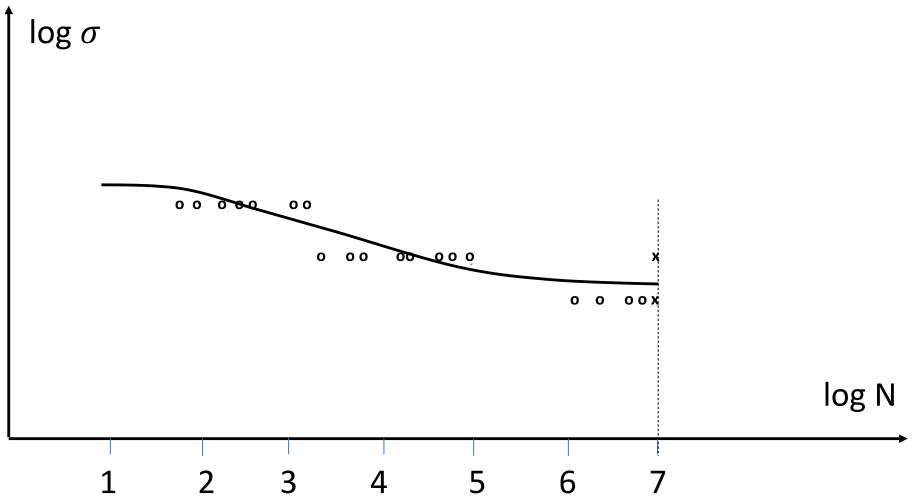


Figure 4.1 Sketch of a S-N-curve also indicating experimental results (rings mean fracture at the life indicated and crosses mean no fracture after a pre-determined life, here 10^7 cycles).

The main findings from S-N-curves were the following:

- As could be expected there is a monotonic relationship between stress amplitude and life, the number of cycles to failure.
- When plotted in a log-log diagram the data points of the sloping part of the curve can be approximately fitted to a straight line. The equation of this line is called the *Basquin equation,* and its parameters are given in handbooks as material data.
- For many materials, as most steels, there seems to be a stress level below which fatigue failure does not occur. This corresponds to local stress levels where the glide band formation is not capable to create micro-cracks extending across grain boundaries, and this is of course of great importance. This so called *fatigue* or *endurance limit*, s_e, is one of the most frequently used parameters in fatigue design. It is lower than the yield stress.
- The number of cycles when the sloping part of the curve turns off into the horizontal one is often in the range $2\ 10^6$ - 10^7. For aluminium and other alloys a fatigue limit does not, however, seem to exist. A continued decreasing relationship between load and life prevails but with a smaller slope. In recent years such an effect has been discovered also for steel for very high numbers of cycles, as high frequency testing devices have been developed.
- There is considerable statistical scatter in the data both in the sloping part of the relationship and at the fatigue limit. This can be understood by the fact that initiation of fatigue damage is of a local nature as described in Chapter 3. In order to evaluate the statistical properties of the scatter it is necessary to perform a substantial number of experiments (see Chapter 10, where testing strategies are discussed).

For shorter lives than around 1000 cycles the stress approaches the fracture strength. In this region, and in fact for stresses above the yield limit, with lives in the range of 10^4 to 10^5, the strain is really a more sensitive and relevant measure of fatigue damage. The damage due to plastic strain and initiation of cracks is not local any more. This type of fatigue is called low cycle fatigue (LCF) and is the subject of Chapter 6.

With more sophisticated equipment it is possible to apply other loading modes, as plane bending, axial loading and torsion, to test specimens. It is also possible to apply a mean load superposed to the alternating one. The effects of these extensions are discussed in Chapter 4.3 and 4.4.

4.2 Use of data from the S-N-curve

It appears that experiments performed in different loading modes give different results so it is ambiguous to define true material properties. The fact that fatigue test results depend on the mode of loading also means that one has to check the testing conditions when using data from the literature. Conventionally, the extra index letter r indicates rotating bending and the index b denotes plane bending. With no such extra index the underlying tests have been performed in axial loading. Stresses resulting from cyclic torsion tests are often denoted by τ instead of σ

In this section results from tests with reversed loading are treated, i e there is no superposed constant mean load.

4.2.1 The stress-life relationship

The sloping part of the S-N-curve gives an assessment of the life as a function of the stress level and therefore a simple analytical expression with few parameters is of interest. The Basquin equation is expressed on a general form as

$$\sigma = A(N)^b, \qquad (4.1)$$

where σ is the stress amplitude giving a mean life of N cycles, and where A and b are material parameters. It is represented by a straight line in a logarithmic diagram, see Figure 4.2.

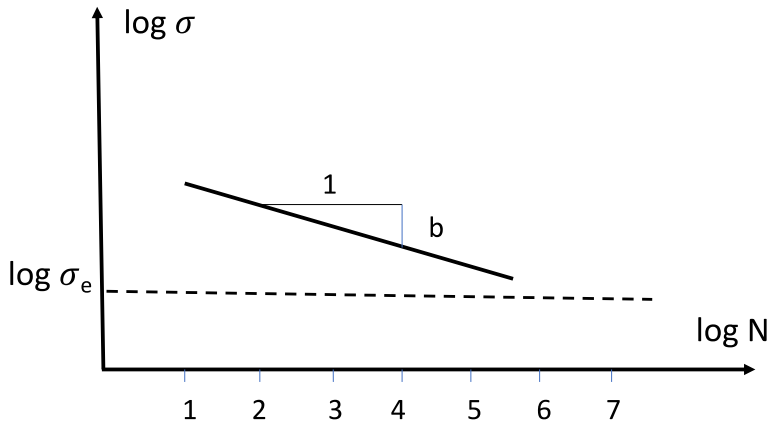


Figure 4.2 Linear approximation of the sloping section of the S-N-curve.

Some data from [4.1] are given in Table 4.1 from tests in axial loading to give an impression of typical values of the parameters. Here, the formula for the relationship is

$$\sigma_a = A(N)^b = \sigma_f'(2N)^b \qquad (4.2)$$

However, data are scarce and rules of thumb may be valuable, although it should be noted that they are crude. One such rule of thumb for structural steels, see Chapter 6, is to put $\sigma_f' = 1.75\sigma_u$ and b = -0.12, which can be compared with data in Table 4.1. The form of the last member of (4.2) is connected to its use in strain based fatigue design. The parameter σ_f' is sometimes approximated by the true fracture strength, and the exponent b is in many cases in the interval -0,09 - -0,12. As seen from these data and from Table 4.1 rules of thumb are not very accurate.

Table 4.1 Some data for the Basquin equation from |4.1]. Stresses are in MPa. σ_u is the ultimate tensile stress.

Material	σ_u	σ_f**'**	**A**	**b**
SAE 1015 (normalized)	415	1020	927	-0.138
Man-Ten (hot rolled)	557	1089	1006	-0.115
AISI 4142 (450 HB)	1757	1937	1837	-0.076

An example of the effect of loading mode is the low strength plain carbon steel DIN C10, see Figure 4.3. One gets A = 660 MPa, b = -0.076 in rotating bending and A = 620 MPa, b = -0.086 in uni-axial tension/compression with the notation of (4.2). (The endurance limits are around 220 and 180 MPa respectively.) The example illustrates the importance of checking the origin of data used in design and is a first indication that loading conditions are of importance.

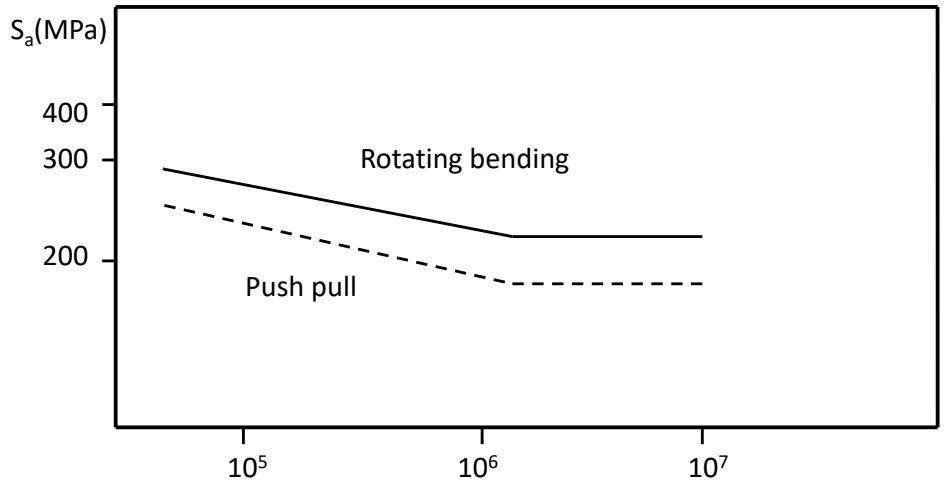


Figure 4.3 S-N-curves for C10 (DIN) steel for bending and for axial load on standard specimens (Approximate sketch from Figure 3.6 in [4.4]).

The scatter in the experimental results is significant and should be evaluated as well. This evaluation is described in Chapter 10. Experience indicates that the distribution in terms of cycles, N, to fracture for a certain stress level is approximately log-normal, i e the distribution of the logarithm of N is approximately normal. The standard deviation estimates can be used to assess the life for a certain level of risk for fatigue failure. Another distribution, which has been suggested as useful to fit to experimental data, is the Weibull distribution.

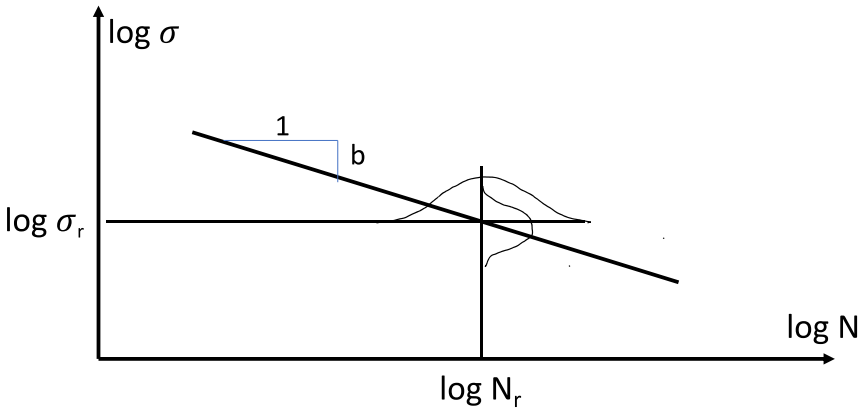


Figure 4.4 Relationship between standard deviations of the lognormal distribution with respect to life and stress.

The data can also be used to find the probability of fracture at various stress levels for a certain life. As an example consider Figure 4.4. For a life with mean N_r for the stress level σ_r the standard deviation of the life has been estimated to be s for an assumed lognormal distribution. Then for N_r the stress has a lognormal distribution with standard deviation sb, where b is the slope of the S-N-curve. With typical values set to s=0,1 and b=(-)0,1 the 95% interval of lives, corresponding to $\pm$ 1,96 standard deviations, becomes $(N_r\, 10^{-0,196}, N_r\, 10^{0,196}) = (0,64N_r, 1,57N_r)$, which is non-symmetric in N. The 95% interval of stresses for N_r correspondingly becomes $(\sigma_r 10^{-0,0196}, \sigma_r 10^{0,0196}) = (0,96\sigma_r, 1,05\sigma_r)$, which is almost symmetric.

The dispersion in data is normally smaller for shorter lives, since then fatigue initiation is not local but occurs due to plastic deformation on the macroscopic scale. It should also be noted that the total life is a sum of periods of cycles of initiation and of crack growth. For long lives the initiation phase with most inherent scatter is by far the longest.

4.2.2 The endurance limit

The endurance limit is the stress level where very small, intra-granular cracks can no longer cross the grain boundaries and continue growing to failure. It is clear from its physical nature, as described in Chapter 3, that it is a stochastic variable with significant scatter. Experiments to find the mean and the standard deviation of the fatigue limit are described in Chapter 10.

As mentioned, for some groups of material, as aluminium alloys, no fatigue limit is found. There is a decreasing relationship between stress and life, but with a smaller slope, even for cycle numbers up to lives of $10^8 - 10^9$ cycles. In recent years similar effects have been detected also for steels. This has raised questions about the reliability of structures, e g aeroplane fuselage used beyond a design life based on 10^7 cycles, and whether new initiation mechanisms can be activated. The area is presently subject to research. Such investigations have to be evaluated with care because for practical reasons experiments have to be performed on new types of small specimens and at high frequencies, also causing raised temperatures. Then the results may not be comparable with those for normal specimen.

Just as the parameters for the Basquin equation, the endurance limit for the different loading modes can be found for various materials in handbooks. Since the experiments to find reliable data are expensive and time-consuming available data are scarce and also here rules of thumb are desirable. One such is that s_e for forged or rolled steel is half the ultimate tensile stress if this is less than 1400 MPa, and 700 MPa for higher strengths. A scanning of several materials shows that the factor may vary between 0,4 and 0,6, so the rule is indicative only. Still such rules are of value since the ultimate tensile strength is generally available as a basic material parameter.

Tabulated endurance limits as well as most fatigue data are given for 50 % probability of failure. The typical standard deviation for the endurance limit is around 5 %.

Some illustrative examples of rules of thumb from [4.2] are given in Table 4.2. In addition to the data given in Table 4.2, data for more materials is given in Table 4.3.

Table 4.2. Approximate fatigue limits in MPa for groups of materials based on ultimate tensile strength, in MPa. (Z= RA, reduction in area in tensile test). Here $\sigma_e=s_e$.

	σ_{er} (bending load)	σ_e (axial load)	s (standard deviation)
Carbon and tempered steels (σ_u<1000 MPa)	$0.51\sigma_u$ -1.67(64-Z)	$0.8\sigma_{er}$	6% of σ_{er}
Ductile cast steels (σ_u<1400 MPa)	$53+0.44\sigma_u$ - $0.00017\sigma_u^2$	$0.8\sigma_{er}$	25 MPa

4.3 Influence of loading mode

It appears, see Figure 4.3 and Table 4.2, that the loading mode is of importance. Tests performed in axial loading give smaller values of the fatigue limit than those obtained from plane bending or rotating bending. In addition, one finds that tests on larger specimens give lower values than those on smaller ones. This is due essentially to stress gradients, and the portions of material exerted to near maximum stress.

The standardized specimens, see [4.3], used to obtain the S-N-curve or fatigue limit for a material have a diameter of 6 - 10 millimetres and a gauge length of less than 100 millimetres. In bend specimens there is a linear decrease of stress from the surface to the centre of the specimens, but this is not so under axial stress. This means that microscopically small cracks are more easily growing from the surface in axially loaded specimens.

When the size of bend specimens becomes larger, the gradient decreases and a high stress is maintained into the depth of the specimen. Then the endurance limit decreases from standard values for bending given in handbooks. Further, in larger specimens more grains at the surface are exerted to near maximum stress, increasing the probability that one of the grains most prone to produce a microscopically small will be present. These effects will be quantified in Chapter 4.5.

For tests in torsion of axial specimens the endurance limit, τ_e, appears to be around $0,6\sigma_{er}$. The shear mode also has a gradient effect with stresses decreasing linearly from the surface to the centre of the specimens, and should be comparable to the rotating bending mode. Presuming that the mechanism is local plasticity and that the material is obeying the von Mises flow rule,

an indication for the relationship should be that fatigue initiation occurs for the same fraction $\hat{\sigma}_o$ of the macroscopic yield stress in tension and torsion, i e from

$$\hat{\sigma}_o = \sqrt{\sigma_{er}^2 + 3\tau_e^2} \qquad (4.3)$$

it is incurred that $\tau_e = \frac{\sigma_{er}}{\sqrt{3}} = 0{,}57\sigma_{er}$.

4.4 Influence of a mean stress

Experiments verify the intuitive impressions from the micro-mechanical model that a constant positive mean stress superposed on the cyclically reversed stress has a negative influence. Since mean stresses are a rule rather than an exception their effect has to be included in fatigue design.

Therefore, stress levels conventionally are characterized by two parameters, mean stress and amplitude

$$\sigma = \sigma_m \pm \sigma_a \qquad (4.4)$$

A related parameter that is frequently used is the R-value

$$R = (\sigma_m - \sigma_a)/(\sigma_m + \sigma_a) \qquad (4.5)$$

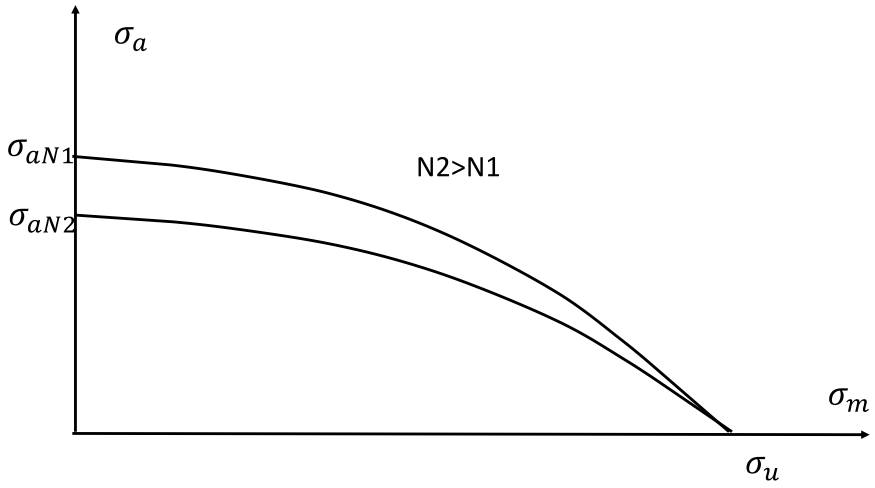


Figure 4.5 Typical relationships between amplitude and a superposed mean stress for different life levels N.

If fatigue experiments are performed with a number of mean to amplitude stress relationships it has been noted by e g Dowling [4.1] that the relationships for different levels of life, N, take on approximately the same shape. The curves are decreasing as functions of the mean stress and the amplitude goes to zero when the mean stress approaches the ultimate stress, as in Figure 4.5.

Then, and at least for relatively long lives N and moderate mean values, a suitable model would be to normalise the results by the amplitude stress for zero mean stress for a certain life, σ_{aN}, as a function of the ultimate tensile stress, σ_u, i e

$$\frac{\sigma_a}{\sigma_{aN}} = f\left(\frac{\sigma_m}{\sigma_u}\right) \qquad (4.6)$$

Such models enable approximate treatment of cases where complete data are lacking, e g when only the Basquin equation and the endurance limit for pure amplitude loading are known in addition to the ultimate tensile strength.

One is the parabolic representation

$$\frac{\sigma_a}{\sigma_{aN}} = 1 - \left(\frac{\sigma_m}{\sigma_u}\right)^2, \qquad (4.7.a)$$

for $S_m \geq 0$, named after Gerber which correctly represents the general features of the relationship, but does not give any mean value influence for small mean values.

Another is the linear Goodman representation

$$\frac{\sigma_a}{\sigma_{aN}} = 1 - \left(\frac{\sigma_m}{\sigma_u}\right), \qquad (4.7.b)$$

which has the correct trend for low mean values but descends too steeply. A modification suggested by Morrow is to use the value σ_f' from the Basquin equation instead of the ultimate tensile stress. σ_f' is approximately equal to the true fracture stress for many steels, also helpful if data are lacking, and since this value is bigger than the ultimate tensile stress it yields a better general fit to data, see Figure 4.6. It cannot be used for large values of the mean stress. This relationship is commonly expressed

$$\sigma_{aN} = \frac{\sigma_a}{1 - \left(\sigma_m / \sigma_f'\right)} \qquad (4.7.c)$$

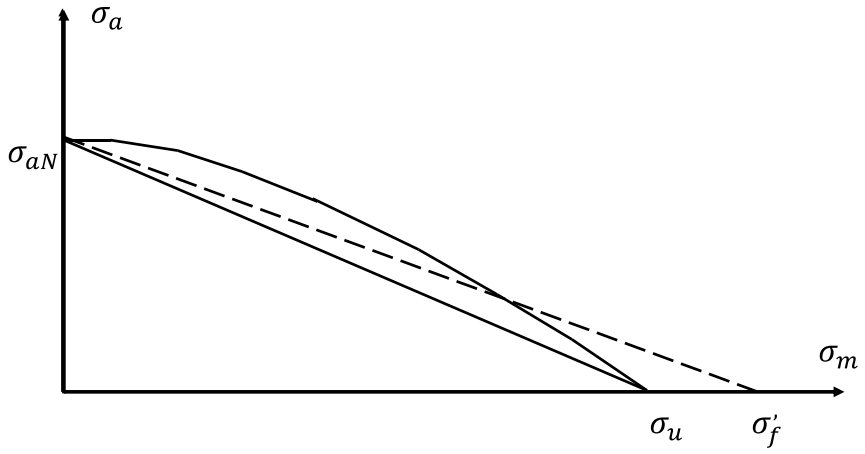


Figure 4.6 Comparison between the models (4.7.a-c) for mean stress influence.

A further model, based on experimental observations for groups of materials, is the one by Smith, Watson and Topper (SWT model). It has the advantage that it does not require any data but σ_{aN} for the life required but it has the disadvantage that it is not valid for large values of the mean stress. It is expressed as

$$\sigma_{aN} = \sqrt{\sigma_a(\sigma_a + \sigma_m)} \qquad (4.8)$$

Of course, a more accurate way to proceed is to use experiments including mean values in a standardized way and for the particular loading mode of interest, since the mean value influence is so important. One approach is to postulate that one additional set of data shall be used, in addition to the ultimate tensile stress and the stress amplitude for zero mean stress, R= -1. One such choice is when the amplitude and mean are equal (pulsating load, R=0, which also is easy to obtain experimentally), i e to use an approximation of the mean stress dependence by three experimental data instead of one (SWT) or two (Gerber, Goodman, Morrow).

Such data are normally given by an additional index p. So, for endurance limits and the basic loading modes the nomenclature is given in Table 4.3 a. Some examples of material data from [4.2] are given in Table 4.3 b. With this type of data the diagram in Figure 4.6 changes into the one in Figure 4.7. Corresponding data for life levels N are indicated with dotted lines. The diagrams in Figures 4.5 - 4.7 are called Haigh diagrams.

Table 4.3 a Nomenclature for endurance limit values.

	R= -1 (σ_m=0)	R=0 (σ_m=σ_a)
Axial loading	$\pm\sigma_e$	$\sigma_{ep}\pm\sigma_{ep}$
Plane or rotating bending	$\pm\sigma_{eb}$, $\pm\sigma_{er}$	$\sigma_{ebp}\pm\sigma_{ebp}$, $\sigma_{erp}\pm\sigma_{erp}$
Torsion	$\pm\tau_e$	$\tau_{ep}\pm\tau_{ep}$

Table 4.3 b Endurance limits including mean value influence for some materials (from [4.2]), in MPa.

Material	σ_u, $\sigma_{p0,2}$	$\pm\sigma_{er}$	$\sigma_{erp}\pm\sigma_{erp}$	$\pm\sigma_e$	$\sigma_{ep}\pm\sigma_{ep}$	$\pm\tau_e$	$\tau_{ep}\pm\tau_{ep}$
Carbon steel,norm.	490, ≥270	±240	210±210	±180	160±160	±140	140±140
Stainless 2324-02	>590,>410*	±350		±250			
Al 7075-T6	>540,>470*	±160		±150	130±130		

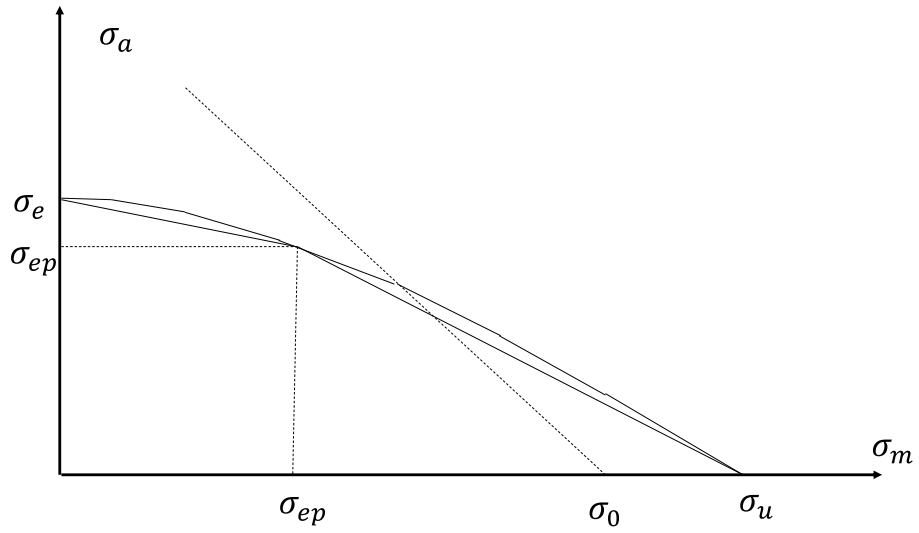


Figure 4.7 Haigh diagram for the endurance limit based on three data points, alternating load, pulsating load, and the ultimate tensile stress.

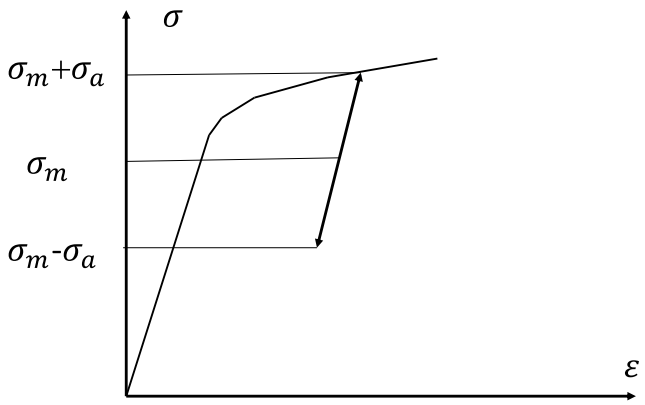


Figure 4.8. Situation with maximum stress in the plastic region but with elastic stress variations during a cycle.

It is seen from Table 4.3 b that the mean value influence in torsion is small for moderate mean stresses. This may be motivated by the fact that the mean shear stress does not induce an opening of the activated glide bands in the same way that mean tensile stresses do.

Intuitively, for stress based fatigue analysis one does not want the total stress $\sigma_m+\sigma_a$ to exceed the yield limit, σ_0, see the line indicating this in Figure 4.7. However, the amplitude stress that may initiate fatigue is still elastic in such cases, see Figure 4.8. Then cases with small amplitudes and high mean stresses may still be safe. This effect is of particular importance where stress concentrations occur at notches, e g at the threads in pre-stressed bolt joints. Effects of plastic stress states will be discussed further in Chapters 4.6 and 5.

4.5 Other factors influencing fatigue properties

As mentioned, the endurance limit, the S-N-curves and the mean value influence are obtained from carefully polished round bars with diameters 6-10 mm tested in a laboratory environment. These data serve as a general reference for a certain material. For use of the material in structural components influencing factors have to be evaluated. The most important ones are those related to dimension/size, surface treatment and environmental factors, including temperature and chemical composition of surrounding air or liquids. The influence is normally given as reduction factors tacitly assumed to be valid for rather wide groups of materials.

This way of working is practical and economic, since it would be more or less impossible to test all combinations of materials and influencing factors and have them tabulated and available. However, it will always introduce approximations and increased dispersion in the results. Often one is forced to accept that the data available is not verified for the exact material or environmental condition sought for. Then it is a question of sound professional judgement how to use it and how to assess the uncertainty introduced.

So, the influencing factors have to be evaluated for sample materials and then generalized for groups of materials (alumina, cast iron, stainless steels etc). Since fatigue damage in metals is so closely tied to microstructure the grain size would be natural to use as the primary parameter for generalization. Instead, the ultimate tensile stress is frequently used. It has the advantage that it is an easily available property for most material, and there is a monotonic although approximate relationship between ultimate tensile stress and grain size.

The reduction factors are related to the endurance limit. For limited lives they tend to have less importance. Still, for lives in the range $10^5 – 10^7$ it is recommended to maintain values obtained for the endurance limit. The use of influencing factors for limited lives is further dealt with in Chapter 5 and 6.

It is emphasized that the content of this section, 4.5, consists of samples from the textbook literature, and should be considered as an account for conventional methods rather than definite design rules.

4.5.1 Influence of size

Production technology (size, rolling direction etc.)
When plates and bars are produced with different sizes there is a larger risk that faults are present in large size components, since the process (forging, casting, rolling, composition) is more difficult to control. This means that the fatigue properties are prone to be inferior in components made from large plates or bars. Data for this effect is scarce and often design

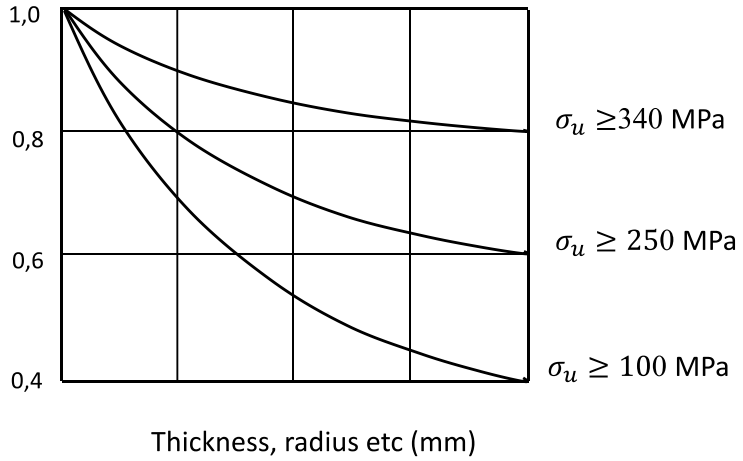


Figure 4.9 Example from cast iron of dependence of fatigue stress from production technology. codes neglect to take this effect properly into account. One of the scarce, and possibly doubtful, examples of reduction factors for the endurance limit for classes of cast materials from [4.2] is shown in Figure 4.9.

Another effect of production technology is that the microstructure is changed by the process, e g that rolling means that the grains become distorted and the fatigue properties may be different in different directions. However, it is stated in [4.2] that modern production methods are such that fatigue properties do not change.

Statistical effects
For materials produced by a specific process there will still be a distribution of weak spots of different kinds, as inclusions or clusters of different phases (ferrites. martensites etc.). This means that a longer bar will have a slightly less endurance limit than a shorter one.

A general way to take this effect into account is by the theory of the weakest link suggested by Weibull.

Assume that two elements of material are coupled in series, or in parallel. In both cases failure will occur if one of them fails under a homogeneous load. Now, generalize to a configuration of n elements, each with the same load and fracture probability P, and the assumption that the structure will fail if one element fails. This probability, P_n, is then calculated from

$$1 - P_n = (1 - P)^n \tag{4.9}$$

Introducing a general volume relationship V/V_e instead of n, and the assumption of the Weibull probability distribution for P as a function of stress

$$P(\sigma) = 1 - exp(-(\sigma/\sigma_0)^m)$$

where m and σ_0 are material parameters, yields

$$1 - P_V = \left[1 - \left(1 - exp\left(-\left(\frac{\sigma}{\sigma_0}\right)^m\right)\right)\right]^{\frac{V}{V_e}} = exp\left(-(\sigma/\sigma_0)^m (V/V_e)\right)$$

and

$$P_V = 1 - exp\left(-(\sigma/\sigma_0)^m (V/V_e)\right) \tag{4.10}$$

This means that for the same probability of failure, e g the 50% endurance limit, the relationship between the stresses σ_1 and σ_2 for the volumes V_1 and V_2 should be

$$\frac{\sigma_2}{\sigma_1} = (V_1/V_2)^{1/m} \tag{4.11}$$

which is the reduction factor when increasing the volume from V_1 to V_2. Since fatigue initiation is a surface related phenomenon for long lives the surfaces A_1 and A_2 should then replace the volumes for analysis of fracture probability. Values for m generally used are 30 or higher for rolled or forged steel and 20 for aluminium alloys, meaning that the distribution is rather narrow. If, for instance, $V_2=10V_1$ and m=30, then σ_2/σ_1 is 0,93.

If a component has a distribution of stresses σ(A) over the area A, with maximum σ_{max}, the probability for failure will be

$$P_A = 1 - exp\left(-\int \sigma(A)/\sigma_0\right)^m dA$$

The same probability would correspond to a smaller area, A_{red}, exerted to the uniform maximum stress calculated from

$$P_{A_{red}} = 1 - exp(-A_{red}(\sigma_{max}/\sigma_0)^m) \tag{4.12}$$

Putting these two expressions equal leads to the equivalent area

$$A_{red} = \int (\sigma(A)/\sigma_{max})^m dA \tag{4.13}$$

The expression (4.13) may be of use for structures with varying stresses above the endurance limit.

Geometry
Another dimensional effect, similar to the one due to stress gradients at notches, is the one occurring in bending or torsion, see Chapter 4.3. The stress decreases below the surface, more steeply the smaller the specimen. This means that the stress condition capable of driving a small crack is less favourable in small specimens. A formula to take this into account, by a reduction factor, δ, for smooth bars in bending or torsion, suggested by Kuhn and Hardrath, is

$$\delta = \frac{1+\sqrt{\alpha_n/r}}{1+\sqrt{\alpha_n/5}} \qquad (4.14)$$

where r is the radius of the bar in millimetres and α_n is the Neuber elementary radius, a function of the ultimate tensile stress, and hence the grain size. This parameter, α_n, a decreasing function of the ultimate tensile stress, is further described in Chapter 4.6 on notch effects, and shown in Figure 4.15. For r<5 mm δ is assumed to be unity. It is seen that δ is smaller than unity for larger r values and compares actual, larger structures with the standardized test specimen. The formula can be used in an indicative way for specimens in bending even if they are not round bars, setting the thickness to 2r. In cases with notches, see Chapter 4.6, the effect of the stress concentration is thought to supersede the dimensional effects and δ is set to unity.

The formula (4.14) indicates why bending is favourable to axial loading. In fact, if the radius of the bar in bending is very large the gradient is small and the stress condition at the surface is similar to that in axial loading. In the limit where r tends to infinity δ approaches values 0,85 – 0,90 for most steels. This compares not too badly with differences between endurance limits for bending and axial loading, particularly if the volume or surface effect according to (4.13) is considered.

There are several other methods and parameters in the literature to take size and geometry into account, e g in norms and standards for pressure vessels and welded structures, see Chapter 13. In many cases the reduction factor is set to be in the range 10-20%.

4.5.2 Influence of surface treatment

From the model established in Chapter 3 for initiation of fatigue damage it is clear that the surface conditions are important. Since the endurance limit is determined for specimens with polished surface and at ambient temperature and laboratory air one has to find correction factors for effects for the actual conditions.

The factor primarily taken into account is the surface roughness. Consider Figure 4.10.
There are three structural lengths to take into account, the surface roughness (characterized by roughness measures as the mean R_a or the profile depth R_t), the grain size (in the sense of a mean value), and the depth to which the surface region is different (e g by induced compression stresses). From the Figure it seems clear that high strength materials, with small grains should be more sensitive to surface roughness than low strength materials. This is also reflected in experimental findings.

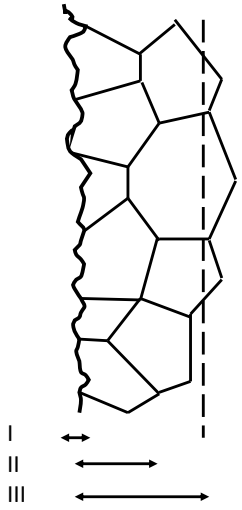


Figure 4.10 Sketch of influencing factors and dimensions near the surface, the surface roughness (I), the grain size (II), and the depth of surface treatment (III).

In Figure 4.11 the correction factor $1/K_r$ (m_s) is shown for various ultimate stresses and levels of surface roughness. The diagrams are based on experiments on certain groups of materials and hence the curves are indicative only when used generally for other types of steel. The Figure is just an example, surface treatments as grinding, polishing etc may be used as parameters in other diagrams.

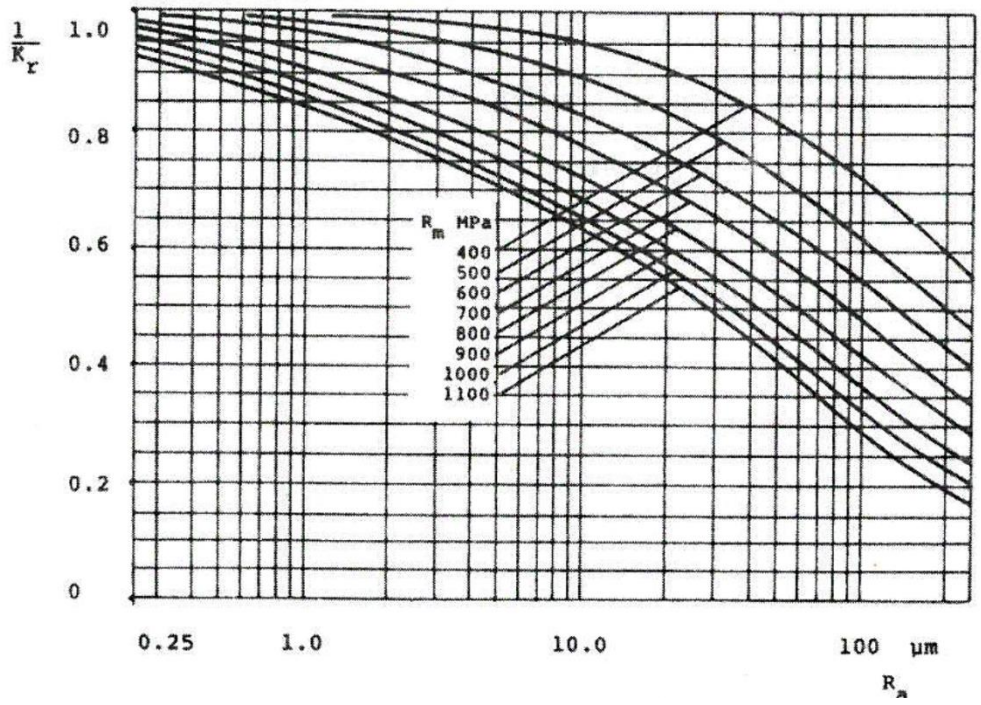


Figure 4.11 Example of reduction factors for various surface treatments (from [4,2]).

There are many other types of surface effect than roughness influencing the endurance limit in a more or less pertinent way. In most cases the correction factor will be smaller than unity, although there are exceptions. One such is the introduction of compression stresses to a certain

depth, by e g shot peening. Then the conditions are in fact improved since initiation and growth of micro-cracks is hampered.

Other effects are introduced by chemical treatment or by application of surface layers of other materials, as e g galvanising. In the literature there are many results available for such cases and for various industrial applications, as for example [4.4].

4.5.3 Temperature and environment

Effects of temperature, frequency, and environment may be significant, especially in combination. The reader is recommended to study the special literature available for various industrial applications, since the area is vast and it is difficult to find and describe general trends. Corrosion fatigue is a special scientific field, particularly in connection with crack growth, with important applications in salt-water applications. Corrosion fatigue is briefly treated in Chapter 12. The problem of irradiated materials in nuclear power plants is another specialized field.

4.6 Notches and fatigue

Due to the localised nature of fatigue initiation, changes in geometry and material properties as notches and welds play an important role in fatigue design since they create local zones of increased stress. The influence of notches is treated here in the fashion usual in textbooks. For other, even more complex issues, like fatigue in welds or contact phenomena, as rolling contact fatigue (RCF), a vast literature has developed. The treatment of welds in norms and standards is discussed briefly in Chapter 13. An introduction to RCF is given in [4.5]

4.6.1 An introductory example

Consider a long strip of width 2b made of an isotropic and linearly elastic material. It has a circular hole with diameter 2R and is loaded by a uniaxial stress according to Figure 4.12.

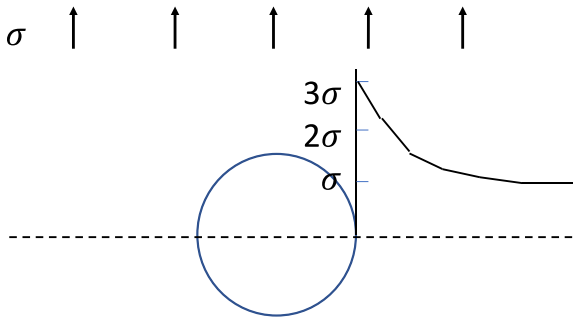


Figure 4.12 Stress concentration in a uniformly loaded sheet with a circular hole. The sheet has total width 2b and thickness t, and the total load is P.

Theoretically, the maximum stress at the hole is 3σ if $b>>R$, and it decreases continually to σ at the edge of the sheet. For other shapes of holes, or geometries as edge notches or notched shafts, the conditions are similar; there is a local concentration of stresses. The relationship between the maximum stress and the nominal stress is called the *stress concentration factor*, K_t, and in this example $K_t = 3$.

Exact values of the stress concentration factor can be found by analytical methods for simple geometries in linearly elastic materials. Approximate values for complex geometries are found by numerical methods or by experiments, e g photo-elasticity. Numerical methods include series solutions and finite element methods, FEM. Finite element methods must be used with care since they do not easily catch the detailed behaviour of stresses at the edges of notches. Rather dense meshes are required; partly since the integration points where stresses are calculated are positioned inside the elements.

K_t-values have been evaluated en masse and are collected in handbooks e g [4.6] for many geometries of use in engineering applications. They are available in the form of diagrams or approximating formulas with the geometrical relationships as parameters. The definition of nominal stress as a function of external load should be observed when using such diagrams or formulas. In Figure 4.12 for example, with P as the external load on the sheet, the nominal stress might be defined as either P/2tb or P/2t(b-R).

For a monotonically increasing load the stress concentration at the hole in the example would just mean that zones at the edges of the hole start yielding as a function of increasing nominal stress. The load limit for general yield in a perfectly elastic-plastic material with yield stress s_0 would be σ_0 (2b -2R) t, where t is the sheet thickness, i e approximately the same load as for the sheet without a hole if b>>R.

For a varying load, as a constant amplitude stress, however, the stress concentration would mean that fatigue initiation occurs locally at the notch root for a nominal stress $\pm\sigma_e/K_t$ instead of $\pm\sigma_e$, i e a hole or other notch means a considerable decrease of load bearing capacity in situations with cyclic loading.

4.6.2 Treatment of notches in practise

Significant deviations from theoretical predictions, like the ones in the example above, appear when experiments are made. In Figure 4.13 the S-N-curves are shown for an un-notched (curve A) and a notched (curve C) bar with the same nominal diameter. The curve B shows the result to be expected with the stress concentration factor K_t as a design parameter. For long lives the reduction in stress appears to be not quite as large as predicted by K_t. For shorter lives, 10^4-10^5 cycles, the deviation from K_t seems to be even bigger. (Data is fitted to the fatigue notch factor K_f explained below.)

These deviations from theory have to be explored in order to find methods for modifications of K_t in design situations.

The main reason why the reduction in nominal stress may be smaller than predicted by K_t is the nature of the process of initiation of fatigue damage and start of growth of small cracks. The theoretical maximum stress, defining K_t occurs at just one point, the notch root. Since the grains close to the edge of the notch have to reach the fatigue limit in average over a distance of several grain diameters this average stress is lower than the maximum one defined by K_t.

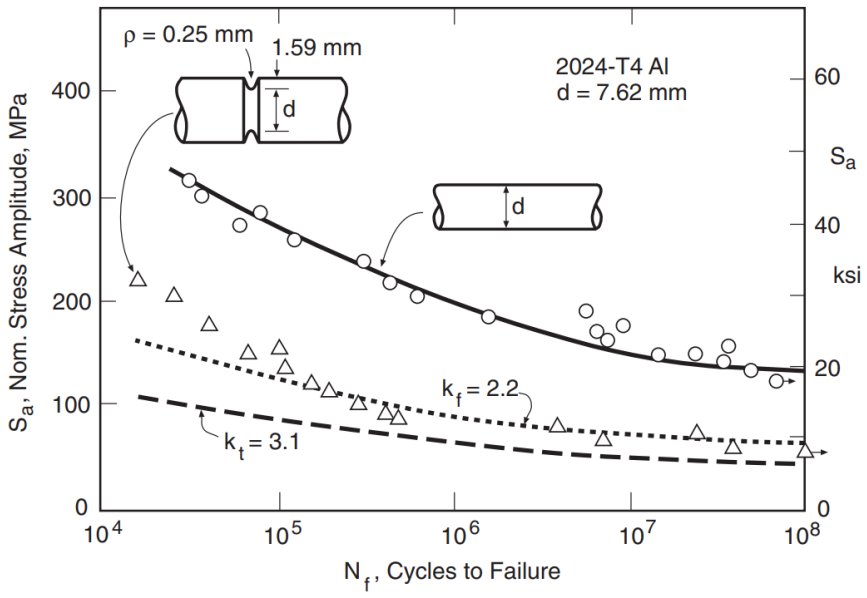


Figure 4.13 Experimental results showing the difference in notch sensitivity in relationship to theoretical prediction. (From Figure 10.2 in [4.1].)

Theoretical analysis shows that the size of the region with significantly increased stresses is generally of the order of the radius, ρ, of the notch root. Consequently, the gradient of the stress at the notch is steeper for sharper notches and then the deviation from K_t increases.

The stress gradient at the root of a notch can be written approximately as

$$\frac{d\sigma_y}{dx} = -\alpha \frac{\sigma_{max}}{\rho} = -\alpha \frac{K_t \sigma_{nom}}{\rho} \qquad (4.15)$$

where α is a factor dependent on other geometry characteristics of the notch than the root radius. This factor does not vary too much and a value of 2,5 may be used for a qualitative analysis. (For torsion it is about half that value, around 1,2.)

There is obviously a balance between the size of the heavily stressed region, essentially characterized by the notch root radius, ρ, and the size of typical microstructural dimensions, as grains or inclusions. For example, for a circular hole in a large sheet the bigger hole is more prone to fatigue initiation since a larger distance ahead of the notch root is exerted to large stresses. See Figure 4.14. There is also another dimensional effect in the fact that larger material volumes along the notch edge are exerted to large stresses.

Hence, the stress concentration factor must be replaced by another factor, which expresses the actual relationship between the nominal stress, $\sigma_{a,notched}$, causing fatigue in the notched structure and the stress, $\sigma_{a,smooth}$, causing fatigue in a smooth specimen. This factor is called the *fatigue notch factor*, K_f,

$$K_f = \frac{\sigma_{a,smooth}}{\sigma_{a,notched}} \qquad (4.16)$$

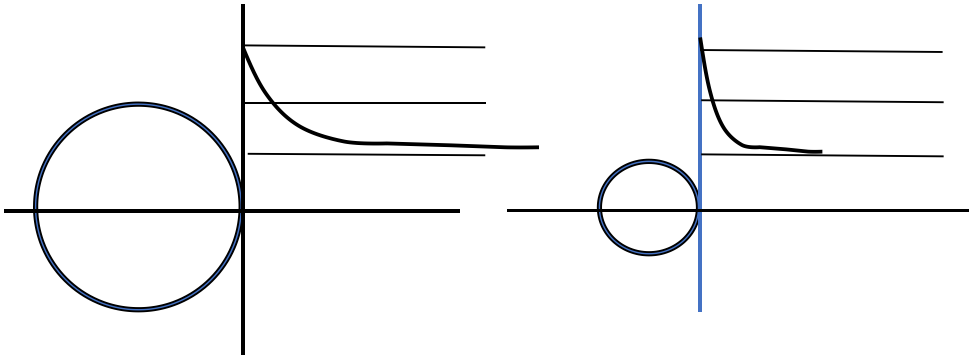


Figure 4.14 Comparison between the stresses at the edge of two holes of different sizes.

Since K_t is easily obtained from handbooks, it is attractive to try to find a model for a relationship between K_t and K_f. From the reasoning about the dependence of the stress gradient on the notch root radius ρ and the necessity of having large stresses over a distance the size of several grain diameters, it is natural to find a relationship based on the quotient between ρ and a length related to the grain size of the material. Based on large amounts of experimental findings expressions have been suggested on the form

$$K_f = 1 + q\,(\rho/\alpha_i)\,(K_t - 1) \quad i = p, n \qquad (4.17)$$

where $q\,(\rho/\alpha_i)$, the *notch sensitivity*, is a number between 0 and 1, and α_i is a material dependent length. When q is close to 1 the notch sensitivity is high, K_f is close to K_t and when q is small K_f approaches unity. Two common forms for q are the ones named after Peterson,

$$q = 1/(1 + \alpha_p/\rho) \qquad (4.18)$$

and Neuber,

$$q = 1/\left(1 + \sqrt{\frac{\alpha_n}{\rho}}\right) \qquad (4.19)$$

The role of structural length is played by the parameters α_p and α_n, which are fitted from experiments performed with different root geometries and various material properties.

Particularly for the Neuber formula, Kuhn and Hardrath established the Neuber elementary radius α_n (often denoted by A) as a function of the ultimate tensile stress in the form of the frequently used diagram in Figure 4.15. The diagram is intended for steel. For aluminium alloys α_n=0.5 mm is used. For other materials q is recommended in [4.2] according to Table4.4.

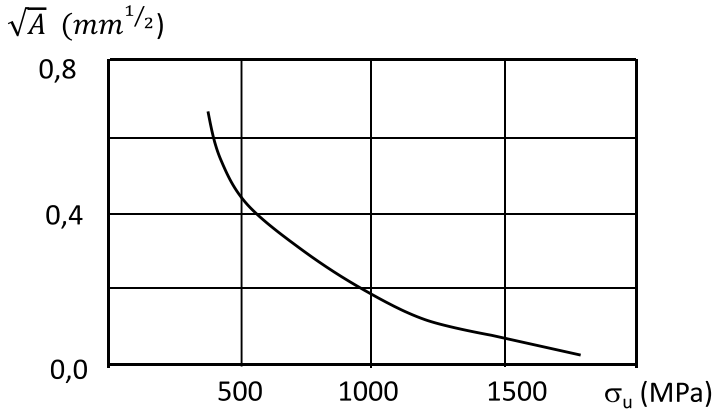


Figure 4.15 The Neuber elementary radius as a function of the ultimate tensile stress. (The Figure is only a sketch and not precise.)

Table 4.4 Notch sensitivity factors q for some types of materials.

Material	**q**
Cast iron	0,2
Austenitic steel	0,2-0,4
Austenitic steel, cold worked Cast steel	$\leq q_{steel}$
In doubt, and brittle materials	1

It is noted that the curve in Figure 4.15 fits nicely with the Hall-Petch relationship between yield strength and grain diameter d

$$\sigma_0 = \sigma_H + H/\sqrt{d} \qquad (4.20)$$

Comparisons with σ_H and H for mild steel, and with the A-values in the Figure indicate that A is of the order of 50 grain diameters. Although the comparison is rough, and H-values vary with type of material, there is a clear indication that the elementary radius A plays the role of a length to be compared with ρ, the typical length at the notch root where stresses are high.

The Peterson formula is also based on the stress gradient and the assumption of the stress level σ_e over a certain distance δ from the notch root necessary for initiation and growth of a fatigue crack, i e from (4.15)

$\frac{d\sigma}{dx} = -\alpha\frac{\sigma_{max}}{\rho} = \frac{\sigma_e - \sigma_{max}}{\delta}$, giving $\frac{\sigma_e}{\sigma_{max}} = 1 - \alpha\frac{\delta}{\rho}$. This leads to an expression for K_f

$$K_f = \frac{\sigma_e}{\sigma_{nom}} = \frac{\sigma_e K_t}{\sigma_{max}} = K_t\left(1 - \frac{\alpha\delta}{\rho}\right).$$

To bring this on the standard form for notch sensitivity q given above, α_p includes K_t on the form

$$\alpha_p = \alpha\delta\left(\frac{K_t}{K_t - 1}\right).$$

For a range of K_t –values and with α=2,5 as above this gives a_p-values from 3 to 5 δ.

Suggested values for α_p are

- 0,06 mm for quenched and tempered steels
- 0,25 mm for annealed and normalised steels
- 0,64 mm for Al-alloys.

For annealed steel this gives that δ is 0,05 mm, which corresponds to 10-50 grain diameters.

In Figure 4.16 the resulting q-values from the Peterson and Neuber approaches are compared for two steels. As expected q increases with notch root radius, since then high stresses act farther away from the notch tip, and then fatigue damage is more and more related to the full

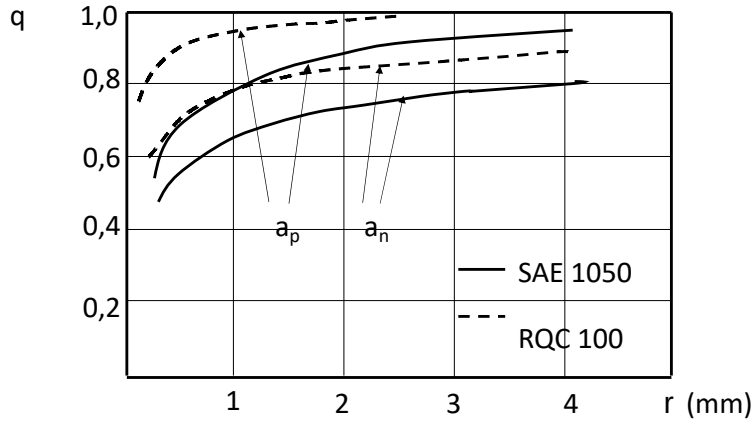


Figure 4.16 Comparison between the Neuber and Peterson expressions for notch sensitivity for SAE 1050 (normalized; σ_u=415 MPa) and RQC-100 (roller Q&T; σ_u=760 MPa).

stress concentration K_t. It is also seen that q may differ considerably both for mild steels and for high strength materials. This is reasonable due to the model character of the expressions, and the Figure 4.16 gives a feeling for the precision of estimates of K_f for design. The approximate character of q is also natural considering that the coupling between the length measures α_n and α_p, and ultimate stress and micro-structural properties as grain diameter distributions are by no means precise.

4.6.3 Notches and mean stresses

When a mean stress is superposed on the amplitude stress there are several suggested ways to extend the concept of a fatigue notch factor by modification of the Haigh diagram. For simplicity the Goodman representation, see Figure 4.17, is used as a basis for reasoning.

The conditions for initiation of fatigue without a notch are represented by the line ABC between the endurance limit and the ultimate tensile stress. Cyclic stress states above the line denoted by σ_0´, the cyclic yield stress, are partly plastic. When a notch with a fatigue notch factor K_f is introduced it is reasonable, see Figure 4.14, to presume that the amplitude and mean stresses at the small process area at the notch tip have the same effect as at a smooth

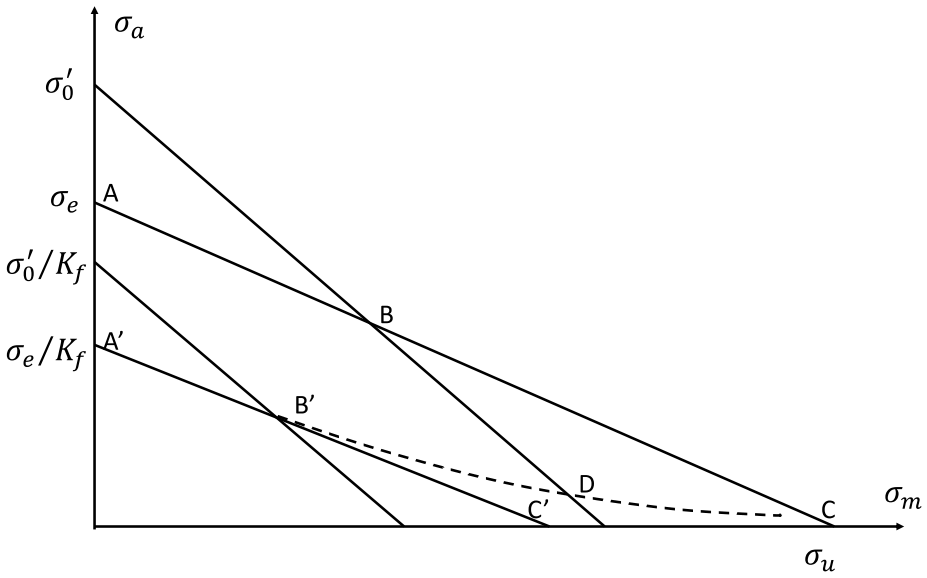


Figure 4.17 Modifications of the Haigh diagram with respect to notch effects.

surface. Then the nominal stress limit should be represented by the line A´B´C´, at least as long as no yielding occurs, to the point B´. When part of the cycle is plastic, see Figure 4.8, the influence of the mean stress gradually decreases, as for the line BC in the un-notched case. This would mean that the line B´C should represent the limiting stress states after commencing yield at the notch tip. When the mean stress increases in the region from B´ to D plasticity is spreading from the notch root to general yield and the influence of a mean stress decreases just as for the region BC.

The reasoning is simplistic and it is acknowledged that experimental evidence is necessary. Further, results for large mean stresses in un-notched cases are sparse and the line BC in the model may not be representative for the real Haigh diagram.

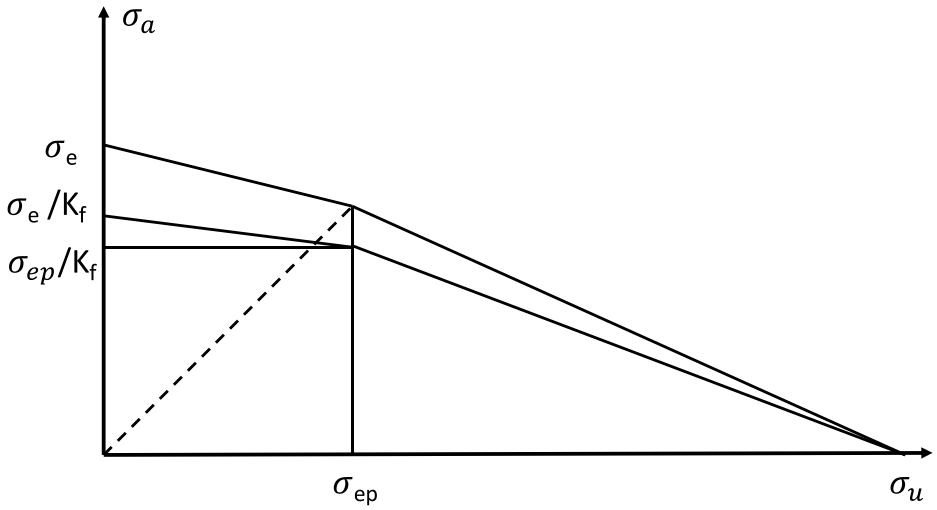


Figure 4.18. Example of Haigh diagram using only amplitude modification for the fatigue notch factor.

For moderate K_f –values the difference between A´B´C and a line A´C is not significant and hence it is suggested in many design codes that only the amplitude stress is reduced, and that

no plasticity should be allowed. For e g the three-point representation the Haigh diagram then becomes the one shown in Figure 4.18.

In e g threads the notch is sharp and the K_f – number becomes high. Then experiments show the general concave trend indicated by A´B´C in Figure 4.17. As a model approach for this situation the SWT relationship 4.8 could be used,

$$\frac{\sigma_e}{K_f} = \sqrt{\sigma_a(\sigma_a + \sigma_m)} \qquad (4.21)$$

In a general situation, where the notch gives a high K_f –value and the situation should be essentially elastic a conservative approach is to reduce the diagram with respect to both amplitude and mean stress, i e to use the curve corresponding to A´B´C´.

However, it seems that even for general yield situations in threads it is possible to have a stress amplitude, indicated by D in Figure 4.17, where fatigue initiation does not occur. As a matter of fact, design of screw joints is often based on tightening them to the nominal yield stress and that the stress variations are then small. The background to this effect is explained in Chapter 13.2.

The physical aspects of arrest of small crack increments at sharp notches were discussed in Chapter 3.5. It is not clear whether this phenomenon contributes or if it is just the fact that the stress variation that is elastic as shown in Figure 4.8.

4.6.4 Notch effects at limited cycle numbers

What has been said above deals with situations at long lives and the possible initiation of fatigue damage. From Figure 4.13 it appeared that K_f decreases even more for shorter lives. This is essentially due to the fact that local plasticity occurs at the tip of the notch for nominal stresses $S_a \geq \sigma_o/K_t$. Then the stress concentration, ideally, is some parameter $K_f' = \sigma_0 / S_a$ for increasing stresses up to general yield, where $S_a = \sigma_0$ and $K_f' \to 1$, see Figure 4.19 a). Experiments similar to the ones shown in Figure 4.13 give results according to Figure 4.19 b), for comparison (after Dowling [4.1]). Although the scales are in stress and fatigue life respectively, the effect is clear, the effective stress concentration factor decreases with increasing plasticity and shorter life. Therefore, design rules often comprise a diagram for correction of K_f at limited lives according to Figure 4.20 (from [4.2]).

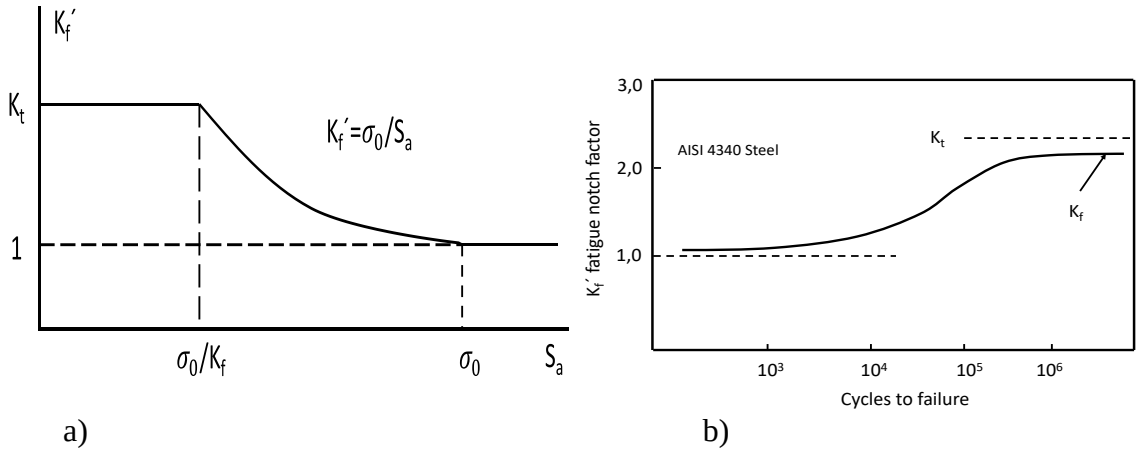


Figure 4.19 Theoretical model (a) and experimental results (b) for the decrease of K_f at higher stresses and lower cycle number to failure.

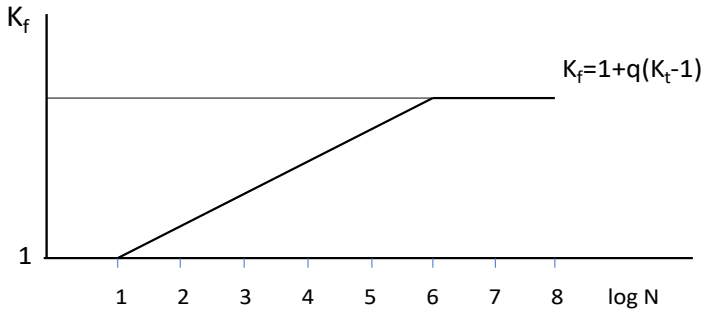


Figure 4.20 Design model for decreased K_f at limited lives, and higher stresses.

For moderate K_f -values the difference between A´B´C and a line A´C is not significant and hence it is suggested in many design codes that only the amplitude stress is reduced, and that no plasticity should be allowed. For e g the three-point representation the Haigh diagram then becomes the one shown in Figure 4.18.

When mean stresses occur for limited cycle numbers the situation becomes more complex for modelling, but the same principle as in Figure 4.20 could be used as a first approximation, at least for lives above around 10^5 cycles. Fatigue design including K_t and K_f, which is founded on the principles of initiation and growth under essentially elastic behaviour should really not be employed for short lives, normally including substantial plastic effects, although several design codes do not exclude them. In these cases the methods using strain as the governing parameter, low cycle fatigue (LCF), should really be used instead. See Chapter 6.

4.7 The Haigh diagram for compressive mean stresses

So far nothing has been said about the influence of compressive mean stresses, i e when R<-1 or for even larger negative mean stresses when the maximum stress during a cycle is in compression. This area is less well explored, partly because such stress states are not common in design. Some general trends may be mentioned, with reference to e g [4.4]. Consider Figure 4.21, which in principle is an extension of Figure 4.17.

For moderate negative mean stresses it is suggested that the trend of mean stress influence at zero mean stress is extrapolated linearly, curves AE and A´E. This may be somewhat non-conservative since the real slope at point A is smaller than the one indicated by the model.

To the left of points E and E´ there are no tensile stresses. It is argued then that start of crack propagation cannot occur and that the curve should change abruptly into E´EF. This is not completely true, for high stress levels cracks may in fact propagate even for negative stress states, particularly when yielding occurs. Further, there should be a smooth transition from one condition to another. Therefore, extrapolations should not be made to very extreme stress states and the principles demonstrated in Figure 4.21 should be applied with care. The Figure demonstrates the weakness of the linear Haigh diagram, where the mean stress dependence is exaggerated for small mean stresses. A more realistic model is obtained with the three-point model, Figure 4.18. Note also that the line to the left denoting the ultimate stress in compression, σ_{uc}, can be questioned.

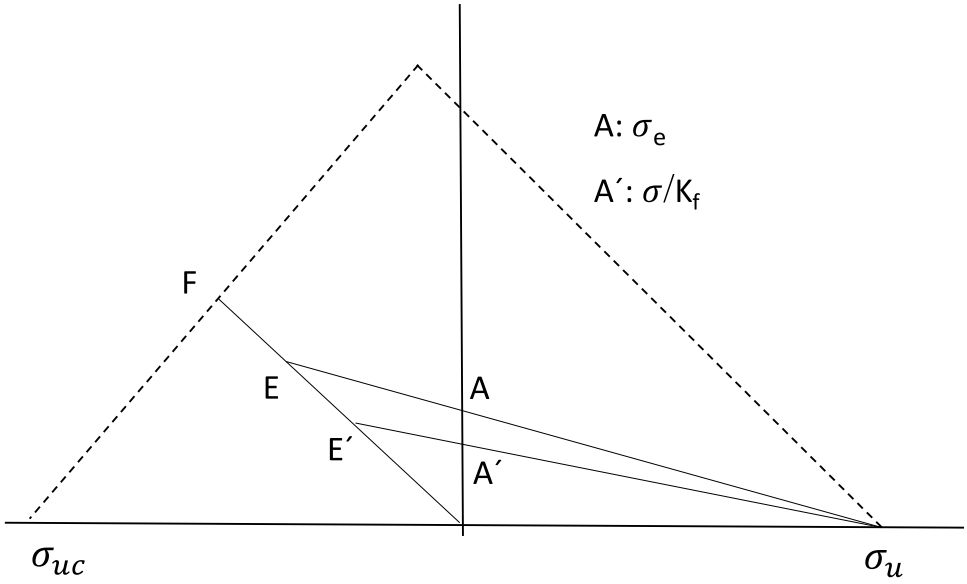


Figure 4.21 Suggested principle for estimation of mean stress influence for negative mean stresses.

References

[4.1] Dowling, N. E. "Mechanical Behavior of Materials", 3rd edition, Pearson Prentice Hall, 2007. ISBN 0-13-186312-6.

[4.2] "Handbok och formelsamling i Hållfasthetslära" KTH (in Swedish), The Royal Institute of Technology, Stockholm, 1998. Ed. Bengt Sundström.

[4.3] ASTM 466-15. "Standard Practice for Conducting Force Controlled Constant Amplitude Axial Fatigue Tests of Metallic Materials"

[4.4] SAE AE-10, "Fatigue Design Handbook". Society of Automotive Engineers, 1988. ISBN 0-89883-011-7

|4.5] Sadeghi, F. et al. “A Review of Rolling Contact Fatigue”, in Journal of Tribology, Vol 131 (2009)

[4.6] Pilkey, W.D. and Pilkey, D.F. “Peterson´s Stress Concentration Factors”, 3rd edition,Wiley (2008). ISBN 978-0-470-04824-5.

5 Fatigue analysis for long lives – the stress-based approach

The aim of fatigue analysis is to assess the probability that a structure can sustain an intended combination of cyclic loading and life. A reduction in load, a change of material, or a revised design of the structure may be the result of the analysis.

Different methods are used for different generic cases. In this chapter analysis is described for long or indefinite lives under constant amplitude loading in structures without inherent cracks. This so called stress-based approach uses the results from Chapter 4.

For shorter intended lives, sometimes involving high temperature applications, where cyclic plastic strain occurs, it is preferable to use a strain-based method, see Chapter 6. For structures where cracks may be present, as when welds are involved, a fracture mechanics based methodology should be used, see Chapters 7 and 13. The treatment of arbitrarily varying loads is described in Chapter 8.

A design strategy intended to guarantee the integrity of the structure for the entire life, as the stress-based approach is called a *safe life* design principle. Another main strategy is to guarantee the life under certain conditions, as redundancy or periodic inspections and repair. This is sometimes classified as a *fail-safe* strategy. The notions are not quite clearly defined in the literature.

5.1 The general picture

The constant amplitude properties of standardized test specimens of a specific material and a specific loading mode can be summarized according to Figure 5.1. Areas with negative mean stresses are not shown and for lives longer than 10^7 cycles the level of the curve CD prevails. The "roof" ABCD in the Figure is an envelope of the stress and life conditions where fatigue fracture does not occur. The curve AD is the S-N- or Woehler- relationship for zero mean stress, and the curve DHC denotes the mean stress dependence, the Haigh diagram, for infinite life.

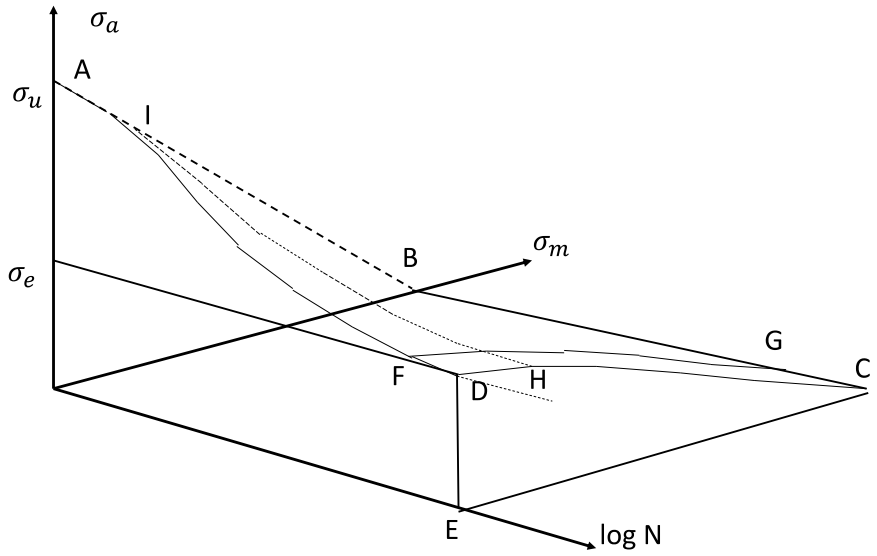


Figure 5.1. General view of the "roof" of stress and life conditions below which fatigue failure does not occur.

The curve IH indicates an S-N-curve with a certain proportion of mean stress.Further, the curve FG indicates the Haigh diagram for a specific, limited life, cf Figure 4.5.

In general, only some parts are explored experimentally, as the endurance limit, σ_e at D. Hence, approximations and extrapolations have to be used, as described in Chapters 3 and 4.

Data for the part AD in Figure 5.1 is normally given as the ultimate tensile stress, the endurance limit, and the approximating Basquin equation

$$\sigma_a = \sigma_f'(2N)^{-b} \tag{5.1}$$

These three parts are then used as a substitute for the real curve, according to the lines A-F-D in Figure 5.2. This would mean that at the point F

$$\sigma_e = \sigma_f'(2N_e)^{-b} \tag{5.2}$$

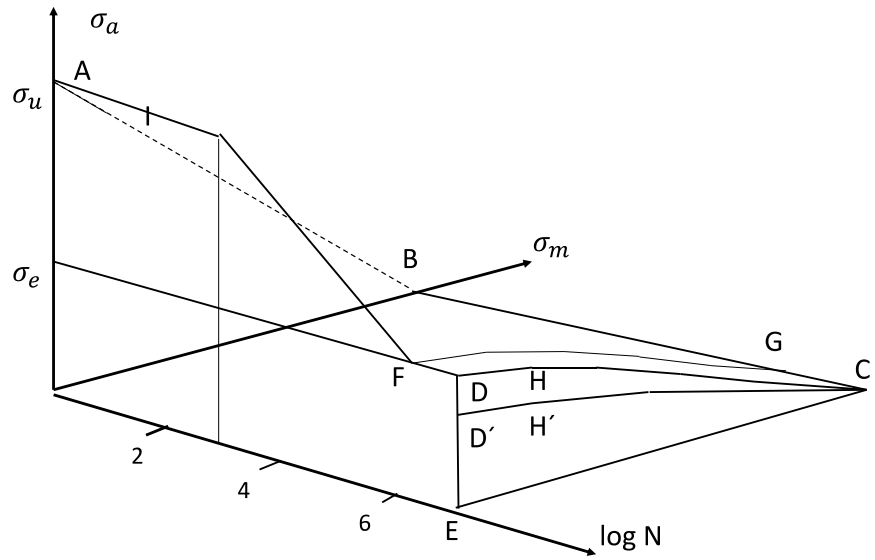


Figure 5.2. Approximations of the "roof" with commonly used data.

where N_e should have a value in the span $10^6 – 10^7$. Hence, instead of the smooth transition to the endurance limit level there will be a sharp corner, representing a conservative estimate. For lives longer than N_e, the roof is extended along the log N axis on the σ_e -level.

In many cases only the endurance limit and the ultimate tensile stress are available. Then it is customary to represent the S-N-curve with a straight line between the ultimate tensile stress at $N= 10^3$ and the endurance limit reached for $N= 10^6$, see Figure 5.2.

Next, the mean stress dependence is considered, represented by the curve DC in Figure 5.1. Depending on the data available the one-, two- or three-point approximations described in Chapter 4.4 and Figures 4.6 and 4.7 can be used, with their respective advantages and drawbacks. In Figure 5.2 the three-point approximation is shown as the curve DHC.

For limited lives one can use, as mentioned in Chapter 4.4, that there seems to be a conformity in the mean stress dependence at least for relatively long lives. Then, for instance, the stresses

at points D and H in the model represented in Figure 5.1 can be modified by the aid of the Basquin equation (5.1) and (5.2) for some desired life N_d by the factor

$$(N_e/N_d)^b > 1 \qquad (5.3)$$

at least for long lives ($> 10^4 – 10^5$), giving the curve FG.

That the conformity cannot be valid to very short lives is clear from the fact that the curve CD is convex and the line AB is straight. For shorter and shorter lives there is a gradual change from the actual non-linear relationship to the straight line denoting that at one half cycle the sum of the mean and amplitude stress is limited by the ultimate tensile stress.

The "roof" thus established for base data now has to be modified with regard to size, surface treatment etc. according to Chapter 4.5, and with regard to the presence of notches and stress concentrations according to Chapter 4.6.

This is indicated for the endurance limit by the line CH´D´ in Figure 5.2. The modification with regard to stress concentrations for large mean stresses is made according to the alternatives described in Chapter 4.6 and not shown in the Figure 5.2.

The influences of notches, size, surface etc, are smaller for shorter expected lives, and there are several ways to take this into account. A conservative way is to use the full reductions, particularly if rather long lives are considered. A more detailed procedure is described in Chapter 5.3

So, the Figures 5.1 and 5.2 give an overview and demonstrate the degree of approximation involved in various formulas used for the analysis. For short lives, normally involving gross cyclic plastic strains, the approximations are the coarsest ones and a strain-based approach should really be used, or at least compared with. For cases with large mean stresses the methods related in Chapter 4.6 should be considered.

In any case, the use of data and models has resulted in a surface, the modified "roof" that envelopes nominal stress states for which fatigue fracture should not occur. The simplest situation is, of course, when indefinite life is asked for since the material data and effects of stress concentrations are best explored for this situation and there is no life dimension to consider.

A more careful approach than just to determine a safety factor from experience, so called professional judgement, is to use the uncertainties in the assessments resulting in the roof, the strength, and in the assessed stress, in terms of standard deviation estimates. Then the risk for failure can be estimated in statistical terms. The background for such analyses is treated in Appendix 1. If the models for estimation of strength are built on simplified, and consciously slightly erroneous mathematical expressions, as the linear approximation for the Haigh diagram, a statistical approach is of course of value just to ascertain that the safety factor adopted is big enough.

5.2 Analysis with regard to indefinite life

This is the classical and the simplest task. The Haigh diagram with the three-point approach is used for demonstration. (The curves DHC and D´H´C in Figures 5.1 and 5.2.)

First, an analysis is made of the stresses resulting from the loads on the component to be evaluated to find the critical points or areas in the structure, i e where the stresses are biggest.

It has also to be decided if the stress state is mainly uni-axial and whether it is mostly bending, tension or torsion, in order to use the appropriate base data for the Haigh diagram, see Figure 5.3. The notation x stands for cases of bending, tension, or torsion respectively. In a general case, where stresses in the structure are evaluated by numerical methods, as FEM, tension data are recommended.

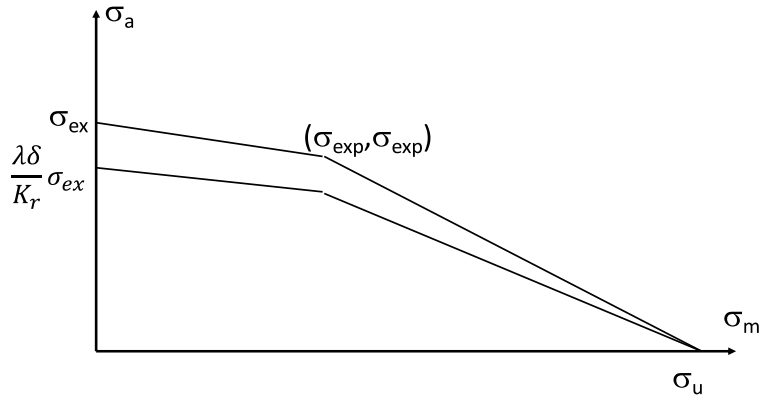


Figure 5.3. Haigh diagram for the endurance limit with modification for dimensional and surface effects.

The next step is to modify the data with regard to size, geometry and surface effects in the real structure according to Chapter 4.5.

The size effect of technology is to introduce a factor $\lambda \leq 1$ which is taken from Figure 4.9 for cast materials. For other materials λ is set to unity if there is no special information about e g in-homogeneities.

When there is no notch present, and for moderate sizes, geometry is considered for bending and torsion by the factor δ, see Chapter 4.5.1. If the heavily loaded regions of the component are much larger than test specimens the Weibull approach can be used for an assessment. For notched structures it is considered sufficient to use the factor K_f , from Chapter 4.6, which takes care of local effects, and which is thought to replace other geometrical size effects, as δ. i e δ is set to 1 in those cases.

Surface effects are considered by the factor K_r, according to Chapter 4.5. Other effects, as from surface treatments or temperature may be found in the literature, as handbooks like [5.1] or on the Internet. The special case of corrosion, which is a vast and complicated field for several important industrial sectors, is dealt with briefly in Chapter 12.

The modified Haigh-diagram in Figure 5.3 is then obtained by reducing amplitudes by the factor

$$\frac{\lambda\delta}{K_r} \tag{5.4}$$

The choice of reduction of the amplitudes is based on the assumption that initiation phenomena are connected to the variation in stress once the mean stress has been considered by the basic Haigh diagram.

The consideration of stress concentrations can in principle be made both as an increase of the load/stress and in reduction of the Haigh diagram for nominal stresses, which is the mostly used

approach. Then the diagram can be reduced uniformly by a factor K_f, see Figure 5.4. In some standards, see the discussion in Chapter 4.6, it is thought to be sufficient to make the reduction only with regard to amplitudes, i e to use the factor

$$\frac{\lambda\delta}{K_r K_f} \tag{5.5}$$

for amplitudes only. This alternative is shown in Figure 5.5.

It is customary to introduce a requirement that the maximum nominal stress must not exceed the yield stress σ_0, see Figure 5.4 and 5.5. Actually, yield occurs at a notch already for maximum nominal stresses σ_o/K_t , see Figure 4.17. The cyclic yield stress should be used for σ_0.

The contour thus obtained is taken as the design limit for 50 % risk for failure. The actual nominal stress is introduced as point O, (σ_m, σ_a) in the diagrams, see Figure 5.4 and 5.5).

Safety factors are obtained as relationships between the measures for margins in amplitude (O O_I), mean stress (O O_{III}), or proportional change in both mean and amplitude stress (O O_{II}).

The value of a required safety factor is normally given in design rules, see also Chapter 13. It should cover uncertainties in fatigue strength as well as the uncertainty in the determination of the nominal stress, in such a way that the risk is acceptably low for the intended application.

(As noted above a probabilistic approach is sometimes to prefer, e g in cases where no safety factor is given a priori. Methods for this are presented in Appendix 1. The principle is the following.

Uncertainties in the endurance limit, and in the correction parameters λ, δ, $1/K_r$ and K_f, are estimated as standard uncertainties (estimates of standard deviations). A combined standard uncertainty is then obtained by the Gaussian formula for combination of uncertainties. Then this measure can be used as the standard deviation of the fatigue strength.

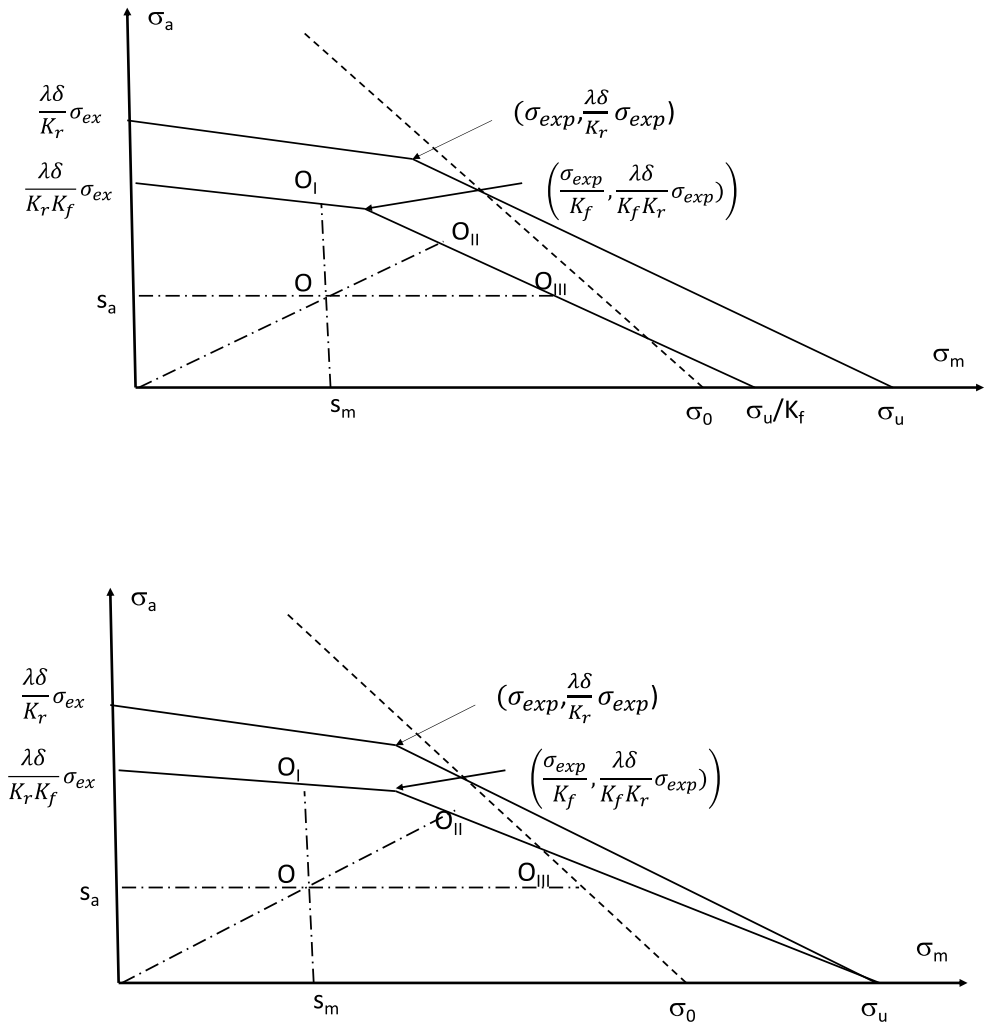


Figure 5.4 and 5.5. Haigh diagrams with different reductions with regard to notches.

Then the standard uncertainty in the nominal stress is assessed. This is in general a difficult task, but it is considered better to do this by available data for the service conditions of the structure than just prescribing a safety factor.

Finally the two distributions are compared by classical statistical methods and a risk for failure is obtained. The nominal stress can then be changed to obtain the desired/accepted risk level.)

There are several codes on the market for fatigue analysis using FE methods for the calculation of stresses. When FE methods are used solely in order to calculate the maximum stresses in a complicated structure for further theoretical analysis some care has to be taken.

The procedure of first reducing the K_t to a K_f and then using this in the Haigh diagram has to be reconsidered. The definition of a K_t from a nominal stress may be difficult. Further, the maximum stress from a FEM analysis is not the true maximum stress but the largest stress in an integration point of the element close to the notch root, and thus dependent on the element size in relationship to the root radius. For relatively shallow notches, a simple and sufficiently accurate way is to use the maximum stress obtained from the FEM analysis in a Haigh diagram without a K_f -correction.

5.3 Design with regard to a limited life

Consider Figure 5.1 again. In order to obtain a Haigh diagram for a not too short but limited design life N_d , corresponding to the curve FG in the Figure, two modifications are made of the diagrams 5.4 or 5.5 respectively.

First the basic amplitude data for the endurance limit are modified by (5.3). This means that σ_{ex} at point D is changed to

$$\sigma_{ex}\left(\frac{N_d}{N_e}\right)^{-b}$$

at point F in the Figure 5.6.

Then, with correction of only amplitudes also for notches, according to (5.5), and use of a modification analogous to the one suggested in Figure 4.20, gives a reduction factor

$$(\frac{\lambda\delta}{K_r K_f})' = 1 - (1 - \frac{\lambda\delta}{K_r k_f})\frac{\log N_d}{\log N_e} \tag{5.6}$$

for F´ instead of (5.5) for D´ in Figure 5.6.

Then a point for an arbitrary mean stress along FG is modified by the factor

$$(\frac{N_d}{N_e})^{-b}\left[1 - (1 - \frac{\lambda\delta}{K_r K_f})\frac{\log N_d}{\log N_e}\right] \tag{5.7}$$

The resulting curve is sketched as F´G in Figure 5.6.

Safety factors with regard to a load point may be decided in the same way as for the endurance limit in Figures 5.4 and 5.5.

It is underlined that both (5.3) and (5.6) are models, and that 5.6 just expresses the experimental fact that dimensional, surface and notch effects are reduced for shorter lives. Further, when b for some materials is a small number the modification (5.3) is small if N_d is reasonably large. As was mentioned before the stress based design method should not be used for short lives with considerable plastic effects.

Other methods have been suggested by Juvinall and by Shigley as described by Dowling in [5.2]. The principles are the same but other factors have been suggested for the modifications with regard to size, surface and notch effects.

The expression (5.7) can also be used to assess the safety margin of the life for a certain stress state with amplitude and mean stress, corresponding to the curve HI in Figure 5.1. If the endurance limit for a given mean stress is σ_{em}, a modified Basquin equation is

$$\sigma_{am}(N) = \sigma_{em}\left(\frac{N}{N_e}\right)^{-b}\left[1 - \left(\frac{\lambda\delta}{K_r K_f}\right)\frac{\log N}{\log N_e}\right] \ , N < N_e \tag{5.8}$$

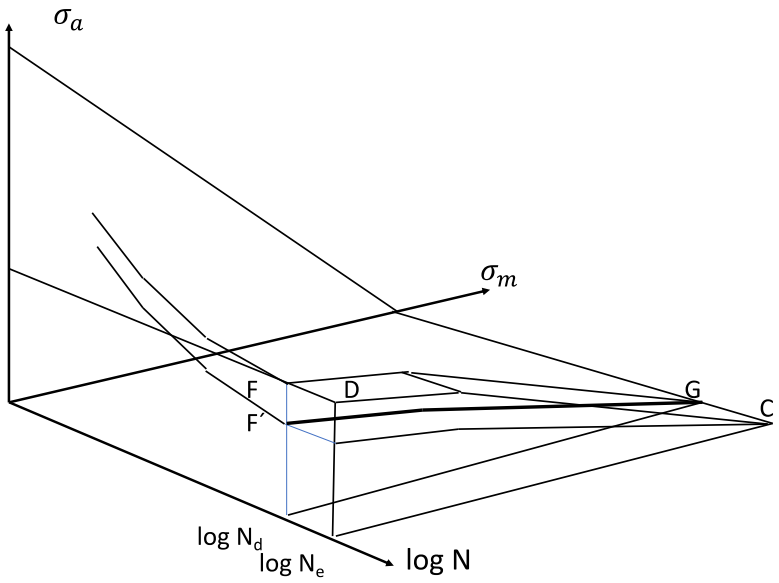


Figure 5.6 Modified Haigh diagram for a given stress state and assumed life.

for a mean and amplitude stress, σ_m and σ_a, and assumed life, N_d, see Figure 5.7 (depicting the plane HI in Figure 5.1), one can assess the life with 50 % risk, N. If the standard deviation of the Basquin equation is known, the probability for failure as a function of cycle number can be assessed as well, see also Figure 4.4 in section 4.2.1.

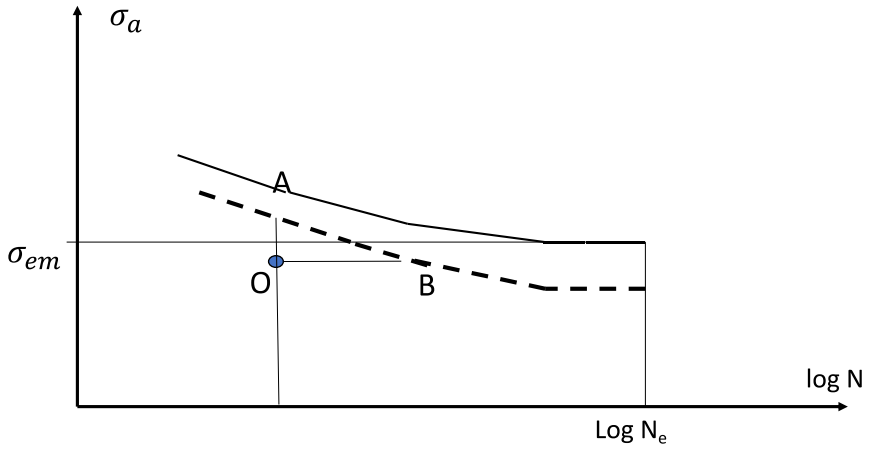


Figure 5.7 Assessment of life for a load point σ_m and σ_a. The life at B gives 50 % probability for fracture.

Example

Determine the pulsating stress σ giving a safety factor 2 with regard to risk for fatigue for $N=10^5$ in the component shown in the Figure 5.8 with R=5mm and 2R/W=0,3. The material is the Man-Ten alloy of Table 4.1. The surface conditions give $1/K_r=0,9$ and no corrections for dimension are considered necessary.

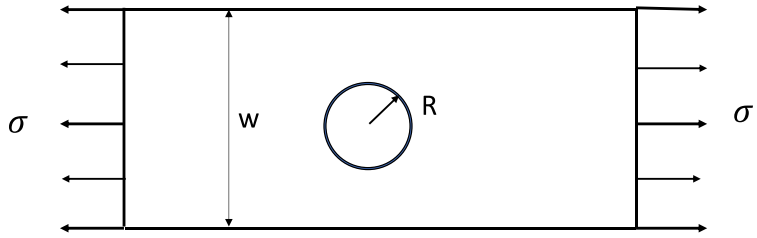


Figure 5.8 Component detail.

From the table A=1006 and b=-0,115 giving σ_a= 268 MPa for N=10^5. Choosing the simple expression (4.7.c) the unmodified Haigh diagram is as shown in the Figure 5.9 corresponding to the curve FG in Figure 5.1. The choice is motivated by the fact that the mean stress is moderate.

From handbook data the stress concentration factor is K_t=3,35. With the choice of the Neuber expression for q the notch sensitivity q becomes

$$q = 1/\left(1 + \sqrt{\frac{\alpha_n}{\rho}}\right)$$

From Figure 4.15 α_n=0,04 mm, and with R=ρ=5 mm q becomes 0,92. Then

K_f=1+q(K_t-1)= 3,16. Together with 1/K_r=0,9 the modification becomes 3,16/0,9=3,51 if no reduction for limited life according to Figure 4.22 is made. (If such a reduction is made the modification factor will instead be 2.81.)

The modified Haigh diagram is also shown in Figure 5.9, going from $\sigma_{a,mod}$ =268/3,51=76 MPa on the vertical axis to 1089 on the horizontal axis. The intersection of this line with the one for increasing pulsating loads, with the slope 1, is (71,71). Hence the pulsating load may be 35,5±35,5 MPa.

From the example it is seen that choices have to be made, but also that they do not influence the result too much with such a large safety factor.

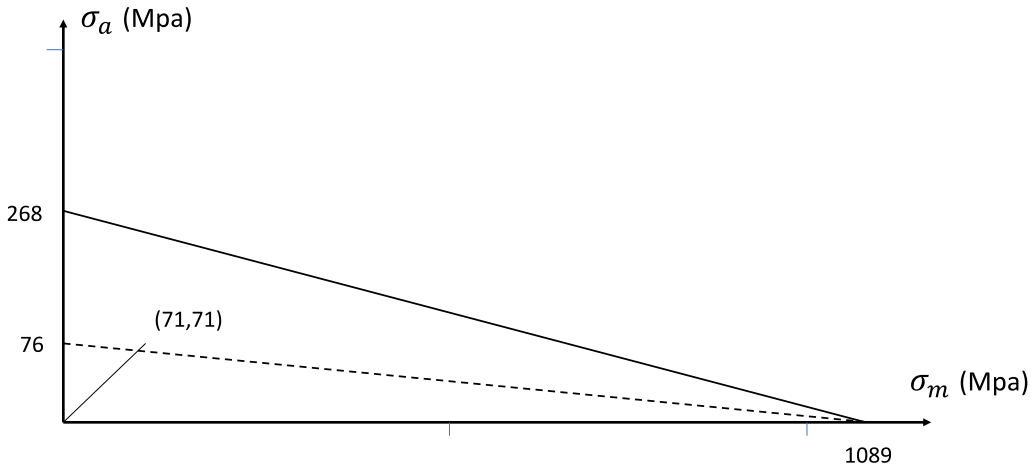


Figure 5.9 Resulting Haigh diagram for N= 10^5, showing the construction to find the allowed pulsating stress.

References

[5.1] SAE AE-10, "Fatigue Design Handbook". Society of Automotive Engineers, 1988. ISBN 0-89883-011-7

[5.2] Dowling, N. E. "Mechanical Behavior of Materials", 3rd edition, Pearson Prentice Hall, 2007. ISBN 0-13-186312-6.

6 Fatigue with influence of plasticity. The strain-based approach (LCF)

Strain rather than stress is found to be better as the governing parameter for fatigue studies when there is cyclic plastic behaviour on the macroscopic level in some part of a structure. Strain is more distinctly resolved in the plastic region of a material than stress. Strain is also likely to be directly connected to the damage processes since these are connected to the macroscopic plastic deformation in the plastically strained volume of material. The strain-based approach is often called LCF (Low Cycle Fatigue) analysis.

The strain approach is significantly different from the one based on stress analysis. As discussed before, the fatigue process under essentially elastic conditions starts by local plasticity in grains, where glide bands form and grow as small inter-granular cracks. Then effects of scatter, size and surface conditions can be expected to be considerable. When macroscopic plastic cycling occurs there is macroscopic damage, the strain amplitude becomes the significant parameter, and effects of scatter, size and surface are smaller.

A strain-based approach appears to be generally applicable for lives around 10^3-10^5 cycles, and shorter, to some extent depending on the material. Typical applications are found in the design of structures exerted to high temperatures, or when the design life is short, e g rocket engines, the number of emergency stops of power plants, and of course when irregular overloads occur, particularly at notches, in many types of components.

Much of the content in the chapter is based on the treatise by Dowling [6.1].

6.1 Model for cyclic material behaviour in the plastic range

For a strain-based analysis of fatigue it is desirable to establish mathematically simple but still realistic models for non-linear material properties. Modelling of the non-linear behaviour of structural materials, so called constitutive relationships, is a complex area, involving the theory of plasticity. Well-known treatises are the ones by Hill [6.2] and Johnson and Mellor [6.3].

For a basic treatment of fatigue just one standard model is introduced here, the Ramberg-Osgood model. The uniaxial form for monotonic loading, which can be consistently generalized to three-dimensional relationships between the stress and the strain tensors by an incremental or a deformation plasticity approach, (see e g [6.3]), is

$$\varepsilon = \frac{\sigma}{E} + \left(\frac{\sigma}{H}\right)^{\frac{1}{n}} \qquad (6.1)$$

This relationship, see Figure 6.1 a), can reasonably well represent the behaviour of many materials. The fact that there is a plastic component, ε_p, defined even from zero stress does not disturb much considering that n and H take on values which make this component small up to stress values corresponding to the yield stress, when it starts to grow quickly. The advantage of the model is that a single analytical expression can be used. The experimental assessment of the material constants E, H and n is described in Chapter 10.

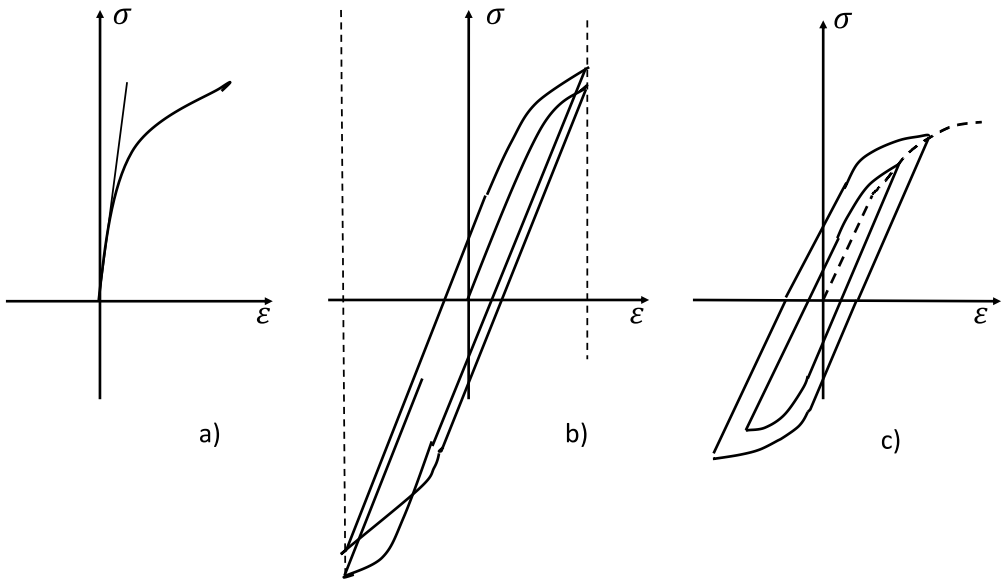


Figure 6.1 The Ramberg-Osgood model for monotonic loading, a), sketch of converging hardening of a material during strain-controlled cycling, b), and finding the cyclic stress-strain curve, c).

When the load is reversed, and then cycled, an additional phenomenon has to be considered. First, it is observed that the strain controlled cycling of a material shows that the material is either hardening or softening, until a stage is obtained where the stress is stable at the prescribed strain maximum, see Figure 6.1 b). Normally, materials in a hardened condition appear to be softening, whereas materials in an annealed condition tend to harden during cycling, as in the Figure, e g due to generation of intersecting dislocations that restricts further dislocation motion more and more until saturation occurs. This goes on for a limited number of cycles until a stabilized loop is obtained.

By connecting the end points of such stabilized stress-strain loops for a number of strain levels the cyclic stress-strain curve is obtained, see Figure 6.1 c). Experimental techniques to obtain such curves is described in Chapter 10.

The cyclic Ramberg-Osgood representation of this curve for corresponding stress and strain amplitudes, σ_a and ε_a, is conventionally written

$$\varepsilon_a = \frac{\sigma_a}{E} + \left(\frac{\sigma_a}{H'}\right)^{\frac{1}{n'}} = f(\sigma_a) \qquad (6.2)$$

where the prime denotes material parameters for cyclic conditions. This curve is the one to be used for strain-based fatigue analysis since it can be assumed that these stabilized conditions are reached rather soon during fatigue loading. The elastic and plastic parts are resolved according to Figure 6.2 a), i e intersection of the elastic line at zero stress defines the elastic and plastic strain spans, ε_{ea} and ε_{pa}.

Particularly when the plastic strains are large, and when the stress and not the strain level is governing, so called ratchetting may occur. This means that the strain is no longer cyclic between fixed end values but increases successively if there are no constraints, see Figure 6.2 b).

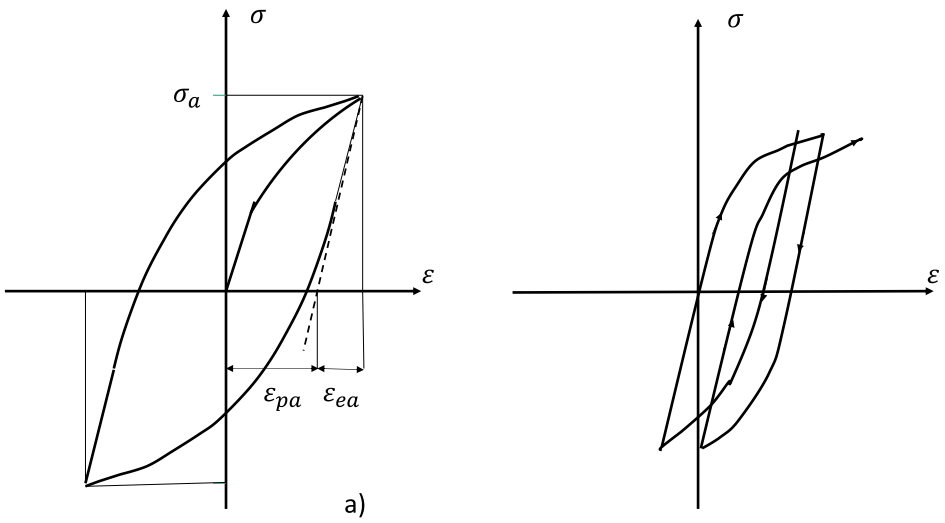


Figure 6.2 Determination of elastic and plastic parts of the stress-strain relationship, and definition of the cyclic model, a), and sketch of the ratchetting phenomenon, b).

It appears that the cyclic behaviour of many materials can be modelled with sufficient accuracy by curves uniform to the amplitude curve according to Figure 6.2 a), i e half the loop, from top to bottom is uniform to and twice the monotonic curve.

The modelling of such a material behaviour for general sequences of reversed cycling of the load is conform to an analogy with springs and sliders. The spring and slider model and its functioning are shown in Figure 6.3. The activation stresses of the sliders increase successively from σ_1 and onwards. First, when loading from A to B, only the first elastic spring deforms since the stress is smaller than the activation stress for the first slider. From B to C the first slider is activated and

$$\varepsilon = \frac{\sigma}{E} + \frac{(\sigma-\sigma_1)}{E_1} = \sigma\left(\frac{1}{E} + \frac{1}{E_1}\right) - \frac{\sigma_1}{E_1} \qquad (6.3)$$

giving the slope $1/(1/E+1/E_1)$ instead of E. For successively higher loads the process continues with more and more sliders activated. For a load between σ_i and σ_{i+1} the slope becomes

$$\frac{1}{\left(\frac{1}{E} + \frac{1}{E_1} + \ldots + \frac{1}{E_i}\right)} \qquad (6.4)$$

If an unloading now starts at point D only the spring E starts to contract between D and E since all sliders prevent displacement. At the point E, when the stress difference from the maximum is $2\sigma_1$, the slider in the element 1 is activated in the reversed direction until point F, where the stress has decreased a further amount $2(\sigma_2 - \sigma_1)$ and the slider 2 will be activated. If the stress is now reversed again from point F to point D the process is reversed again as depicted. First the slope is E until the stress has increased by $2\sigma_1$ to point G and the slider 1 is activated. The loop is closed at point D. Obviously further spring and slider elements could have been included in the loop in the same fashion, with each part of the loop twice the size of the corresponding one in cyclic stress strain curve ABC..DH.

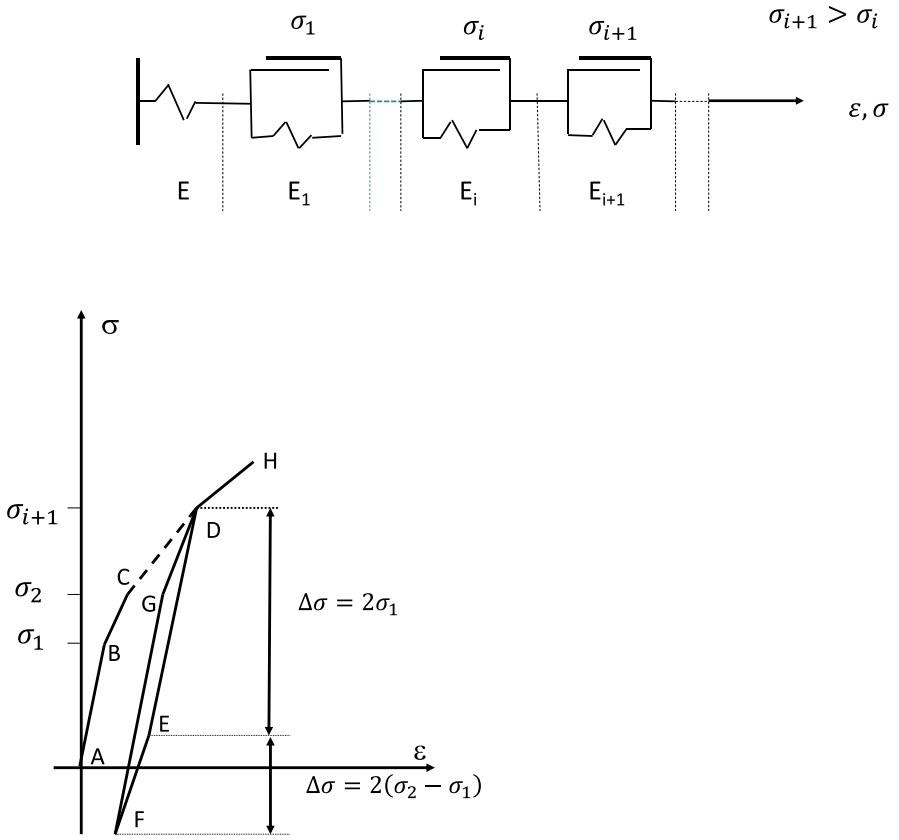


Figure 6.3. Spring and slider model of cyclic elastic-plastic material behaviour.

When the stress is increased further from the point D, where the stress is σ_{i+1}, the spring and slider element i+1 is activated and the original curve is resumed to point H and further.

The discrete model in Figure 6.3 can be generalized to one of continuous behaviour by increasing the number of spring and slider elements and diminishing the differences between the stress levels infinitely. The model is a simplistic analogy for a material saturated with systems of glide bands with obstacles of varying levels.

However, this model establishes two principles enabling a simple, but still realistic description of the stress-strain behaviour of an elastic-plastic material under irregular cyclic stressing.

1) For a reversed stress, represented by $\Delta\sigma$, all relations of the cyclic amplitude relationship, e g the Ramberg-Osgood one

$$\varepsilon_a = f(\sigma_a), \qquad (6.5)$$

are doubled giving

$$\frac{\Delta\varepsilon}{2} = f\left(\frac{\Delta\sigma}{2}\right) \qquad (6.6)$$

If one turning point is denoted by ε_A and σ_A the new state ε_B and σ_B is found from

$$\varepsilon_B = \varepsilon_A - 2\, f\left(\frac{\sigma_A - \sigma_B}{2}\right) \quad \text{(defining f(-x)=-f(x))} \qquad (6.7)$$

2) When a loop is closed the former stress-strain relationship is resumed due to the memory effect of the spring and slider model.

A sequence of varying stresses is transformed into a sequence of cycles, closed loops, each characterised by a strain amplitude and a mean stress. These can then be used with the methods for life calculation described in Chapter 6.2. In Figure 6.4 an example is shown for a simple stress sequence. Starting at one extreme of the stress levels is just for convenience and is conventionally used.

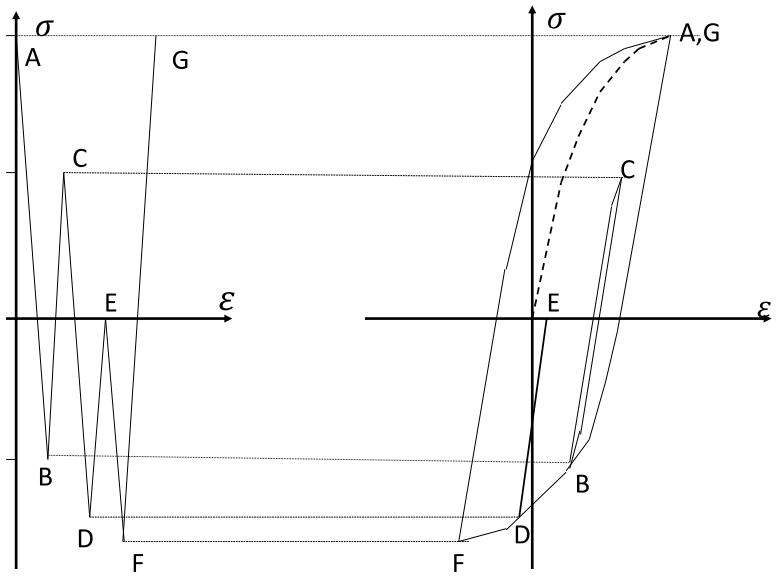


Figure 6.4 Demonstration of the method to produce stress-strain loops from a stress sequence. Three cycles are defined, viz. AF, ED and CB.

This way to proceed conforms to the rain-flow count method outlined in Chapter 8, and it is actually the origin of it, see [6.4]. It shows that rain-flow count defines physically

distinguishable cycles, the loops. This is not the case for other cycle counting methods used to produce so called load spectra, see also Chapter 8.

6.2 Relationship between strain and life

Experiments to find the fatigue behaviour under plastic cycling are normally performed with rather short, standardized, test specimens having an hour-glass shape, see [6.5]. This is to prevent plastic buckling and instability in the compression phase. Also, the testing machine has to be stiff and well lined up so that the specimen remains straight, and the grips have to be tight. The tests are performed in strain control by a displacement gage placed centrally on the specimen. The experimental technique is described in more detail in Chapter 10.

For each strain level chosen the number of cycles, N_f, to fracture and the stabilized stress-strain cycle are registered. The plastic and the elastic parts, ε_{pa} and ε_{ea}, of the strain amplitude can be distinguished, see Figure 6.2 a). It is then found that the relationships between N_f and the plastic and elastic strains respectively can be represented rather well by straight lines on a log-log plot of strain versus life.

The lines representing the plastic and elastic parts are written as

$$\log \varepsilon_{pa} = \log \varepsilon_f{}' + c \log 2N_f \qquad (6.8)$$

and

$$\log \varepsilon_{ea} = log \frac{\sigma_f'}{E} + b\ log 2N_f\ . \qquad (6.9)$$

This means that the line for the plastic part starts at $\log \varepsilon_f{}'$ on the vertical axis and then has a slope c. The elastic line is constructed analogously. See Figure 6.5.

The results can be written as

$$\varepsilon_{pa} = \varepsilon_f{}'(2N_f)^c \text{ and } \varepsilon_{ea} = \frac{\sigma_f{}'}{E}(2N_f)^b \qquad (6.10)$$

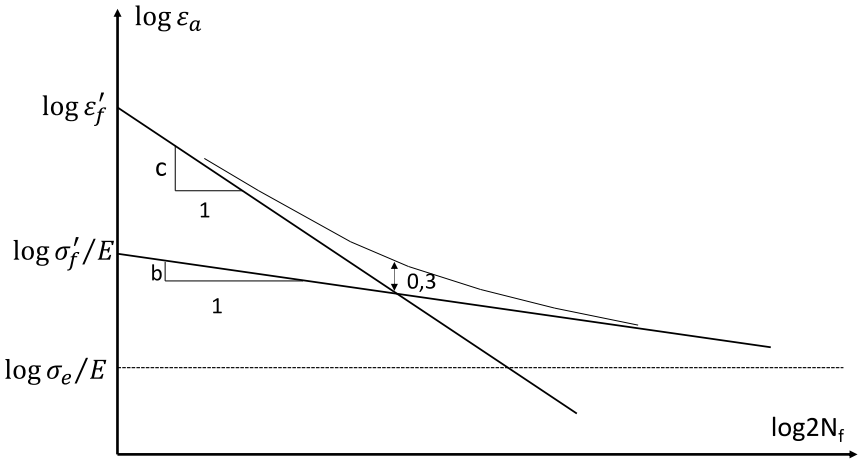


Figure 6.5 The Coffin-Manson relationship, with construction of the elastic and the plastic parts.

The elastic part is recalled as the Basquin equation, which takes over for long lives, observing that ε_{ea} E = σ_a for the uniaxial tension-compression mode.

Now, the total strain amplitude is

$\varepsilon_a = \varepsilon_{ea} + \varepsilon_{pa}$

so the complete relationship between life and strain amplitude becomes

$$\varepsilon_a = \frac{\sigma'_f}{E}(2N_f)^b + \varepsilon'_f(2N_f)^c \qquad (6.11)$$

This equation is called the Coffin-Manson relationship.

The curve is shown in Figure 6.5. For very long lives the endurance limit is approached and the curve bends off to be horizontal, the horizontal line in Figure 6.5. This part is seldom shown, and is mentioned here to point out the common basis for the two diagrams, where for short lives the representation in terms of strain gives a much better resolution as a basis for design.

At the intersection, where the elastic and the plastic strains are equal, the sum is twice the value, i e in logarithmic scale 0.3 units above the intersection. The life at the intersection is called the transition fatigue life and denotes a rule of thumb border for the use of a stress or a strain based analysis.

The technique for determining the material parameters σ'_f, ε'_f, b and c from experiments by making least squares linear regression for the two sets of data is illustrated in Appendix A.2. Since there is scatter in the results a reliable indication of the standard deviation at different levels requires many tests. Data given in tables and handbooks for various materials give the two straight lines for 50% risk for failure. The scatter is smaller for shorter lives since material damage under plastic conditions occurs in material volumes from the start, which is a more deterministic process.

Since data are scarce, and tedious and expensive to produce, Manson [6.6] using numerous experimental data has suggested an approximate expression only using the material parameters obtained from the tensile test:

$$\varepsilon_a = 1{,}75\frac{\sigma_u}{E}N^{-0{,}12} + 0{,}5D^{0{,}6}N^{-0{,}6} \qquad (6.12\text{ a})$$

where $D = ln\left(\frac{1}{1-\Psi}\right)$ and with σ_u as the ultimate tensile stress and Ψ the contraction $\Delta A/A$of the tensile test. These are readily obtainable in catalogues for structural materials. A more recent formula by Roessle and Fatemi using only the Brinell hardness HB is

$$\varepsilon_a = \frac{4{,}25HB + 225}{E}\left(2N_f\right)^{-0{,}09} + \frac{0{,}32HB^2 - 487HB + 191000}{E}\left(2N_f\right)^{-0{,}56}$$

(6.12 b)

Anther rule of thumb to approximate (6.11) is to set σ_f' and ε_f' to the true fracture stress and true fracture strain values respectively, and to try to find b from S-N-curves and use c=-0.6. If H´ and n´ values are available (6.10) can be used to get assessments by professional judgement. All such efforts should be recognized to be order of magnitude assessments.

The life for a given strain amplitude is obtained approximately from the diagram, or numerically from iteration by e g the Newton-Raphson iteration formula, i e

$$x_{i+1} = x_i - \frac{f(x_i)}{f'(x_i)} \quad \text{with} \quad x_i = (2N_f)_i \tag{6.13}$$

with

$$f(x_i) = \frac{\sigma_f'}{E}(x_i)^b + \varepsilon_f'(x_i)^c - \varepsilon_a \tag{6. 14}$$

Finally, it is of interest to note a relationship between the model material parameters used for the strain based analysis. From the plastic part of the Ramberg-Osgood equation one obtains

$$\sigma_a = H' \varepsilon_{pa}^{n'} \tag{6.15}$$

Further, if the life $2N_f$ is eliminated from the elastic and plastic parts

$$\varepsilon_{pa} = \varepsilon_f'(2N_f)^c \quad \text{and} \quad \varepsilon_{ea} E = \sigma_a = \sigma_f'(2N_f)^b$$

of the life equation, and σ_a is resolved, one obtains

$$\sigma_a = \sigma_f' \left(\frac{\varepsilon_{pa}}{\varepsilon_f'} \right)^{b/c} \tag{6.16}$$

Comparing the expressions for σ_a in (6.15) and (6.16) it is found that

$$H' = \frac{\sigma_f'}{\left(\varepsilon_f'\right)^{b/c}} \quad \text{and} \quad \text{n'} = \text{b/c} \ , \tag{6.17}$$

i e *there are only four independent parameters of the six used for the material and the life model equations*. If tests are performed for a real material and the results are made to fit the models it is improbable that the relationships for H' and n' will be fulfilled. Conversely, one could see the degree of agreement as an indication of how well the models can simulate the real behaviour.

6.3 Influence of mean stresses

General stabilized fatigue loops, where significant plastic strain takes place, can be characterized according to Figure 6.4. The parameters are the strain amplitudes, and the mean stress levels, ε_a and σ_m. It seems clear from how damage is produced that ε_m has little importance. A superposed constant strain around which there is a cyclic process producing damage by plastic deformation changes should not have much influence. However, a superposed stress could, as in the elastic case, facilitate the formation and growth of small cracks from the plastic damage, e g at inclusions or grain boundaries.

Hence a rational, still physically reasonable way to include σ_m in the relationship between strain amplitude and life should be established. A first approach is to assume that the mean value influence is of most importance in the elastic and near elastic situations, as before, when extension of locally formed glide bands into cracks is enhanced by mean tensile stresses. For parts of the curve with considerable plastic strains it is then assumed that there is only little influence from the superposed mean stress.

Use of the linear Morrow approximation (4.7.c) gives a simple expression. Introduction of

$$\sigma_{ea} = \frac{\sigma_{ea,m}}{\left(1-\frac{\sigma_m}{\sigma_f'}\right)}$$

where σ_{ea} is the stress amplitude with no mean stress and $\sigma_{ea,m}$ is the stress amplitude with a mean stress σ_m, into the elastic part of the Coffin-Manson relationship (6.10) written as

$$E\varepsilon_{ea} = \sigma_f'(2N_f)^b = \sigma_{ea}$$

readily gives

$$\sigma_{ea,m} = E\varepsilon_{ea,m} = (\sigma_f' - \sigma_m)(2N_f)^b$$

and the modified equation

$$\varepsilon_a = \frac{(\sigma_f' - \sigma_m)}{E}(2N_f)^b + \varepsilon_f'(2N_f)^c \qquad (6.18)$$

The only change is that the asymptote for the elastic behaviour is moved down and parallel to the original curve, see Figure 6.6 .

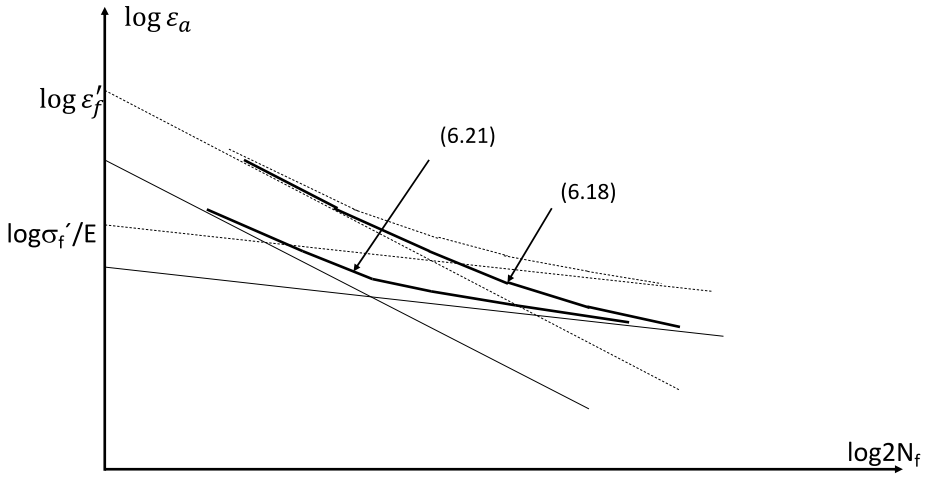


Figure 6.6. Coffin-Manson equation modified with regard to a mean stress. Non-modified curve is shown dashed for comparison.

The plastic part of the curve can be included in a consistent way in the following way. Re-writing the first term in the expression (6.18)

$$\frac{\sigma_f'}{E}\left[\left(1-\frac{\sigma_m}{\sigma_f'}\right)^{\frac{1}{b}}(2N_f)\right]^b \qquad (6.19)$$

it is seen that the original form of the term is retained with a transformed life N*

$$N^* = N_f\left(1-\frac{\sigma_m}{\sigma_f^{'}}\right)^{\frac{1}{b}}$$

and consequently, for both terms

$$\varepsilon_a = \frac{\sigma_f'}{E}(2N^*)^b + \varepsilon_f'(2N^*)^c \tag{6.20}$$

i e

$$\varepsilon_a = \frac{\sigma_f^{'}}{E}\left(1-\frac{\sigma_m}{\sigma_f^{'}}\right)\left(2N_f\right)^b + \varepsilon_f^{'}\left(1-\frac{\sigma_m}{\sigma_f^{'}}\right)^{\frac{c}{b}}\left(2N_f\right)^c \tag{6.21}$$

This means moving the values of $\sigma_f^{'}/E$ and $\varepsilon_f^{'}$ down on the vertical axis of the strain life diagram, see Figure 6.6. However, since $(1-\sigma_m/\sigma_f')<1$ and $c/b>1$ the reduction for the plastic term is larger than that for the elastic one, which is not generally in agreement with experimental findings.

A further possibility is to use the modified Smith-Watson-Topper (SWT) parameter $\sigma_{max}\varepsilon_a$ as a a parameter for the life relationship, i e

$$\sigma_{max}\varepsilon_a = F(N_f) \tag{6.22}$$

Since $\sigma_{\max} = \sigma_a$ for the special case $\sigma_m = 0$, and $\sigma_a = \sigma_f^{'}(2N_f)^b$ from before, the function $F(N_f)$ is found as

$$\sigma_{\max}\varepsilon_a = \sigma_f'(2N_f)^b\left[\frac{\sigma_f^{'}}{E}(2N_f)^b + \varepsilon_f'(2N_f)^c\right],$$

or, for comparison with the formulas above,

$$\varepsilon_a = \frac{\sigma_f^{'}}{\sigma_{\max}}\frac{\sigma_f^{'}}{E}\left(2N_f\right)^{2b} + \frac{\sigma_f^{'}}{\sigma_{\max}}\varepsilon_f^{'}(2N_f)^{b+c} \tag{6.23}$$

Here it is seen that the levels are adjusted slightly upwards, since $\sigma_f^{'}/\sigma_{\max}$ will in most cases be a number larger than unity. On the other hand the slopes will be steeper.

The various models illustrate efforts to include influence of a mean stress in the strain based approach. They should be seen just as models. It is the authors belief that formula (6.18) is the most realistic one.

In Figure 6.7 a life assessment is shown for a strain amplitude ε_a^* and a mean stress σ_m as in Figure 6.4 and with the use of the Morrow formula (6.18) for taking the mean stress into consideration.

For an arbitrary sequence of loading, as the one determined by rain-flow counting in the example in Figure 6.4, the damage can now be assessed by the Palmgren-Miner rule (see Chapter 8) from one of the equations (6.18), (6.21) or (6.23). For each cycle i the life N_{fi} is

determined from the strain amplitude and mean stress. Then the damage from one sequence becomes

$$D = \sum_i \frac{1}{N_{fi}}$$

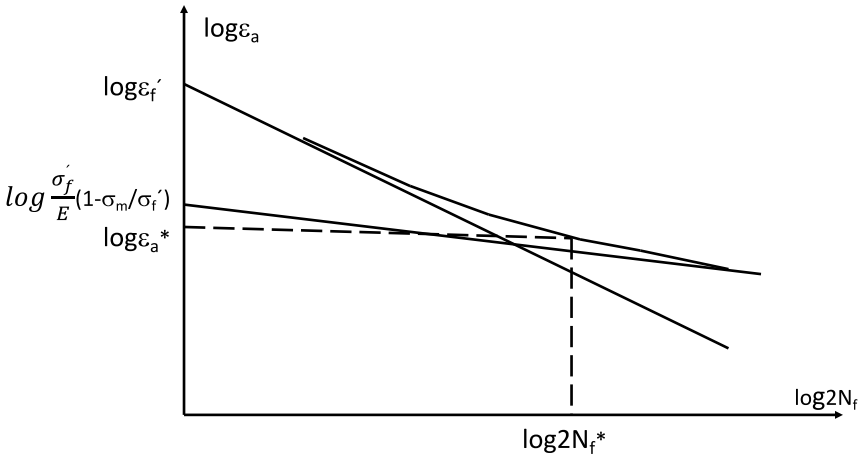


Figure 6.7 Determination of the life, $2N_f^*$ for a cyclic load with a strain amplitude ε_a^* and a mean stress σ_m.

and the life in number of sequences is 1/D. If the formula (6.23) is used the maximum stress of each cycle has to be used.

The process is tedious since a new life equation has to be used for each mean (maximum) stress. In cases where the mean stresses are small in relationship to the amplitudes, only the strain amplitudes are used together with life equations that are not modified with respect to mean stress. Then, when mean stresses are not included, the cycle counting is called range pair count, see Chapter 8.

6.4 Size, surface and other effects

Just as for fatigue in elastic cases one should in principle consider effects from size, surface treatment and other factors. As discussed in Chapter 4 and above in this chapter these effects are essentially connected to initiation of non-linear behaviour on the microscopic level in an otherwise elastic situation. In a strain formulation, and when the plastic strains are comparable to the elastic ones or even larger, there is immediate damage formation on the macroscopic level.

Then it would seem sensible to apply, if at all, the corrections for size and surface effects only to the elastic part of the Coffin-Manson relationship. From Chapter 4 it is recalled that the effect is diminished for larger stresses.

A simple way to include the effects of size and surface treatment, and other such effects then would be to reduce the elastic term in the Morrow-corrected formula (6.18) further by a term like (5.4)

$\frac{\lambda\delta}{K_r}$.

It might include corrections for other effects, of course. With (5.4) the modified equation (6.18) becomes

$$\varepsilon_a = \frac{\lambda\delta}{K_r}\left(\sigma_f' - \sigma_m\right)\left(2N_f\right)^b + \varepsilon_f'\left(2N_f\right)^c \qquad (6.24)$$

The importance of such corrections depends on the type of load sequence applied. In a case, where the load sequence consists of cycles with an even mix of large and small ones, and the large ones are plastic, it is easily seen from the structure of (6.24) and from Figures 6.6 and 6.7 that the life depends essentially on the large cycles. Then it is not thought necessary to include the correction terms according to (6.24).

However, in the not uncommon situation, where some equipment is running with a large number of small elastic load cycles and a few plastic overload cycles, the elastic term becomes significant.

6.5 Determination of strains at notches

Most applications of a strain-based design involve parts with notches, just as in stress-based design. Even modest nominal stresses may cause cyclic plastic straining locally at roots of notches, e g in threads or gears. The treatment is now complicated by the fact that the concentrations of strains and stresses differ. In the plastic region strains become larger and stresses smaller than what corresponds to the elastic stress concentration factor K_t. This is because the elastic-plastic cyclic stress-strain curve bends off. Approaching the limit of perfect plasticity the stress concentration tends to the value 1. Hence it is necessary to define separate stress and strain concentrations factors

$$\varepsilon_a = K_\varepsilon \varepsilon_{nom} \qquad (6.25)$$

$$\sigma_a = K_\sigma \sigma_{nom} \qquad (6.26\ )$$

where σ_{nom} is defined as before as the uniaxial nominal stress at the site of the notch, and where

$$\varepsilon_{nom} = \sigma_{nom}/E. \qquad (6.27)$$

For some particular notch geometries Neuber [6.7] found analytical solutions having the property

$$K_\varepsilon K_\sigma = K_t^2 \qquad (6.28)$$

where K_t is the elastic stress concentration factor. Since this relationship offers a simple way to analyse elastic-plastic situations at notches it has been tacitly assumed to be generally valid under the name Neuber´s rule. Multiplying the two expressions (6.25) and (6.26) one obtains, with the use of (6.27) and (6.28)

$$\varepsilon_a \sigma_a = \frac{K_\varepsilon K_\sigma \sigma_{nom}^2}{E} = \frac{K_t^2 \sigma_{nom}^2}{E} \qquad (6.29)$$

Here the right member, with K_t^2, is easily obtained from handbook data for many cases.

Now, the cyclic stress-strain relationship is also valid, e g the Ramberg-Osgood model

$$\varepsilon_a = \frac{\sigma_a}{E} + \left(\frac{\sigma_a}{H'}\right)^{1/n'} \qquad (6.30)$$

From the two expressions (6.29) and (6.30) ε_a (and σ_a) can be calculated graphically or numerically and used for insertion in the Coffin-Manson relationship to find an assessment of the life. The situation is shown in Figure 6.8. Note from the Figure that one point on the hyperbola (6.29) is given on the straight line from the origin by $K_t\sigma_{nom}$.

If the nominal stress is so large that elastic conditions do not apply, e g as in the case of a Ramberg-Osgood material model, actually, one can formally modify the Neuber rule and get

$$\varepsilon_a\sigma_a = K_t^2\sigma_{nom}\varepsilon_{nom} = K_t^2\sigma_{nom}\left(\frac{\sigma_{nom}}{E} + \left(\frac{\sigma_{nom}}{H'}\right)^{1/n'}\right), \qquad (6.31)$$

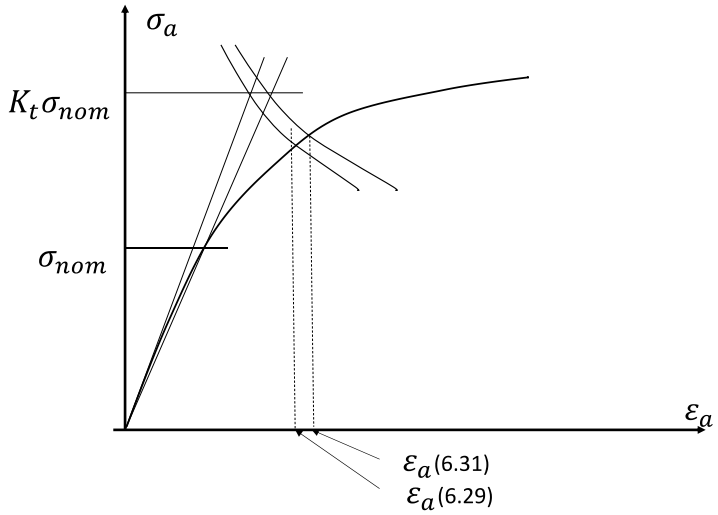


Figure 6.8 Graphical determination of the strain amplitude at a notch by the aid of Neuber´s rule with assumed elastic (6.29), and actual non-elastic nominal stress (6.31).

and proceed as before. The resulting equations for σ_a and subsequently for ε_a have to be solved numerically. The corresponding situation is shown graphically in Figure 6.8. For moderately non-linear nominal stresses the difference between the two approaches is not so big and the simpler formula can be used. Other cyclic stress-strain relationships can of course be used instead of (6.30).

6.6 Variable amplitude loading in the presence of notches

The treatment of notches in Chapter 6.5 was performed for a situation with a cycle of constant amplitude and zero mean stress.

For a general load sequence of variable loading, the Neuber rule can also be applied to find the stress and strain states at a notch together with the rules established in Chapter 6.1.

For each change in nominal stress ΔS the corresponding stress and strain changes $\Delta\sigma$ and $\Delta\varepsilon$ are now found from the Neuber rule (6.31)

$$\frac{\Delta\varepsilon}{2}\frac{\Delta\sigma}{2} = K_t^2\frac{\Delta S}{2}f(\frac{\Delta S}{2}) \qquad (6.32)$$

and the rule 1) from Chapter 6.1

$$\frac{\Delta\varepsilon}{2} = f(\frac{\Delta\sigma}{2}) \qquad (6.33)$$

as before. The rule 2) from Chapter 6.1 is applied as before, i e when a loop is closed the preceding one takes over due to the memory effect.

The process should start at a point where the absolute value of the nominal stress is maximal in order to simplify the counting.

Application for the whole load sequence results in a sequence of closed loops, which will give the strain amplitudes and mean stresses for each resolved cycle. The procedure is analogous to the example in Figure 6.4 of Chapter 6.1. In reality the nominal stresses in cases of notches are most often elastic so that

$$f\left(\frac{\Delta S}{2}\right) \approx \frac{\Delta S}{2E} \qquad (6.34)$$

Unfortunately the pair of equations (6.32) and (6.33) must be solved numerically. Often a graphical solution is made based on Figure 6.8. Then, instead of finding the intersection between the material property (6.33) and the Neuber hyperbola (6.32) for each cycle, it is practical to first find the relationship between strain amplitude and nominal stress amplitude

$$\frac{\Delta\varepsilon}{2} = F\left(\frac{\Delta S}{2}\right) \qquad (6.35)$$

by sampling a number of nominal stress amplitudes.

If the analysis can be made numerically from the beginning by e g FEM (finite element methods) the Neuber rule does not have to be used and (6.35) is obtained directly.

The whole procedure to find the stress and strain changes for a simple load sequence is described by the following example, illustrated in Figure 6.9.

Example
The nominal stress sequence is shown in Figure 6.9 a). The material is characterized in the form (6.30) by

$$\varepsilon_a = \frac{\sigma_a}{2 \cdot 10^5} + \left(\frac{\sigma_a}{10^3}\right)^5 \qquad (6.36)$$

shown in Figure 6.9 b) with σ_a in MPa, and the notch is assumed to have K_t=2,5. The function (6.35) is obtained graphically from Figure 6.9 b) by insertion of sample values of nominal stress amplitudes S_a, and the result is shown in Figure 6.9 c). The method is shown for the nominal stress amplitude 140 MPa denoted A in Figure 6.9 b) and c), giving ε_a=2,4 10^{-3}. The point A´ with the stress S_a K_t=350 MPa is positioned at the Neuber hyperbola as seen from (6.29). (A slightly different result would be obtained with (6.31), c f Figure (6.8).)

Then by insertion of stress amplitudes from Figure 6.9 a) into Figure 6.9 c) strain amplitudes are obtained. These are then used to find stress amplitudes from Figure 6.9 b) and finally the changes $\Delta\varepsilon$ and $\Delta\sigma$ as $2\varepsilon_a$ and $2\sigma_a$ respectively.

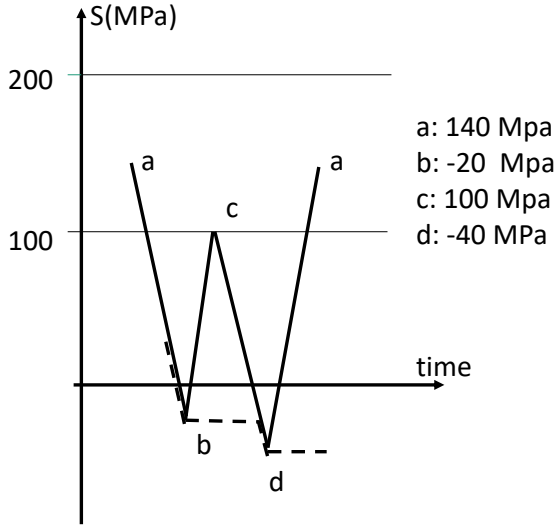


Figure 6.9. a) Construction of stress-strain loops at notches from an irregular load sequence. Load sequence.

The first point a is found from S_a=140 giving the first maximum as σ=260 MPa and ε= 2,4 10^{-3} (see the Figures 6.9 b) and c)) and introduced in Figure 6.9 d).

The first stress change ab is ΔS=160 MPa and has a half value S_a=ΔS/2=-80 MPa giving $\Delta\varepsilon/2$ as approximately -1,1 10^{-3} from Figure 6.9 c) (the signs are irrelevant).

Using this in the material curve Figure 6.9 b) gives $\Delta\sigma/2$ as approximately -175 MPa. The resulting curve with $\Delta\varepsilon$=-2,2 10^{-3} and $\Delta\sigma$=-350 MPa is shown as the part ab in Figure 6.9 d).

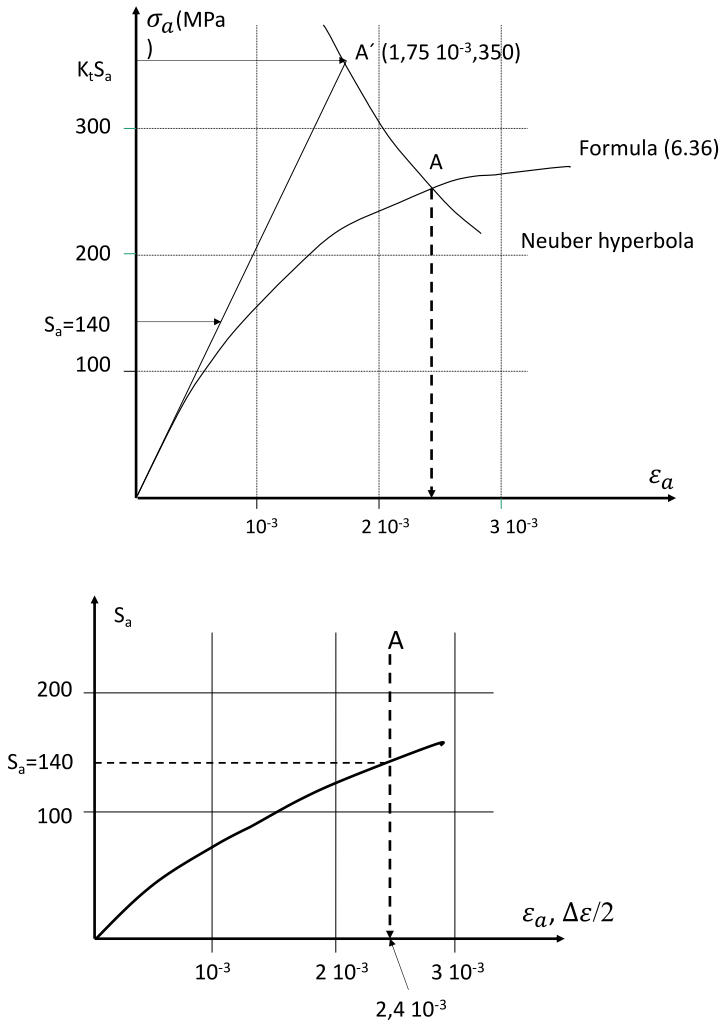


Figure 6.9.b and c. Construction of stress-strain loops at notches from an irregular load sequence. b) Material curve and Neuber hyperbola to solve stress and strain graphically, c) Help graph (6.35).

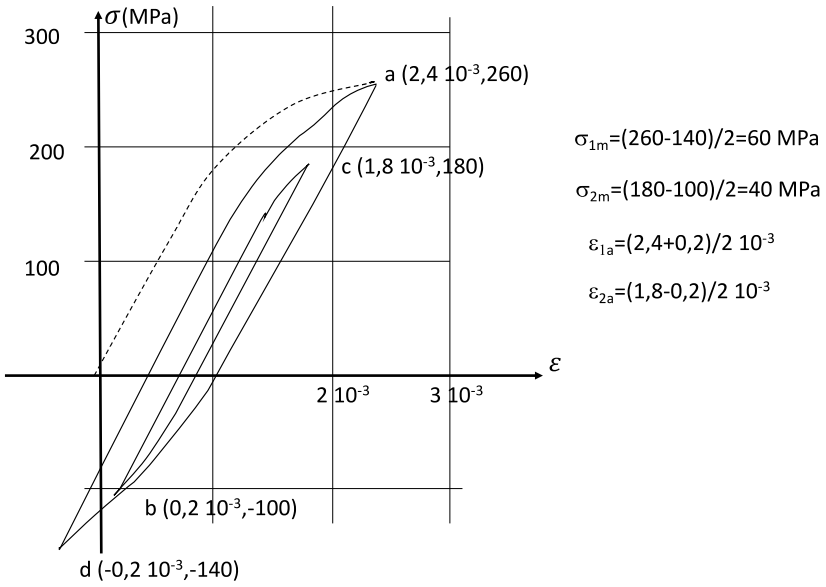


Figure 6.9.d Construction of stress-strain loops at notches from an irregular load sequence. Resulting cycles giving stress and strain data for life assessment.

Then the procedure goes on for the successive stress changes in Figure 6.9 a) resulting in the complete picture of Figure 6.9 d). The coordinates will be, approximately,

b (0,2 10^{-3}, -100)
c (1,8 10^{-3}, 180)
d (-0,2 10^{-3}, -140)

It is noted that the change cd means a return to the curve ab and this change, and similar ones, should be calculated from the full change ad due to the spring and slider model. (This is actually connected to the method for rain flow cycle counting, "cutting off", the closed loop bc, see the dotted line in Figure 6.9 a)).

The resulting stress-strain loops are shown in Figure 6.9 d). As is seen there will be two loops, with

ε_{a1}= 1,3 10^{-3} and σ_{m1}= 60 (MPa)

ε_{a2}= 0,80 10^{-3} and σ_{m2}=40 (MPa)

These can then be used for life estimates according to Chapter 6.3.

The rain-flow counting method, see Chapter 8, results in cycles with mean and amplitude stresses. It should be observed that these are not preserved in the case of inelastic stresses and strains at notches. Although the method presented here for analysis of such cases is built on a model, experiments show that it works out well in real life.

6.7 A note on the influence of strain states

Although the procedure in Chapter 6.5 is simple and intuitively attractive some comments are in place both from a materials and from a strain analysis point of view. The strains and stresses obtained are the ones just at the tip of the notch, and assumed to be the largest principal component, and intended to be compared with results from test specimens. In an elastic-plastic case the material damage occurs over material volumes and not by initiation in some grains. Further, the damage mechanism is most probably connected to strain rather than to stress. In the stress based design plane stress prevails at the notch, and the modification of K_t to K_f compensates for gradient effects.

Hence, in the strain-based approach the largest plastic strains should prevail over some range below the surface and that most of the life of the notched component is consumed when the material this range is completely damaged. This seems to function well in many cases when the contained region of plasticity at the notch is large enough, typically of the order of the notch radius.

Another point to consider is the strain state. The test specimens used to obtain the strain-life relationship (for the Coffin-Manson model) are in a uniaxial stress state, which is also present in thin, notched specimens. So, for these the use of material data for life calculations is consistent. However, most cases of notches are in thick specimens with constraints. At the notch root the stress and strain state is

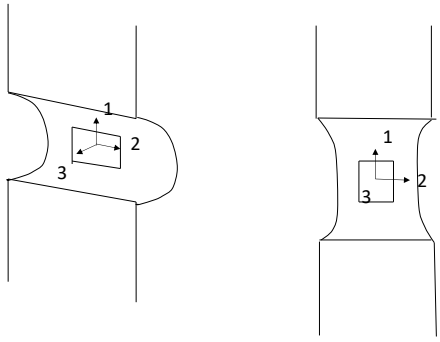


Figure 6.10 Stress and strain states at a notch and in a test specimen.

σ_1, σ_2, with σ_3 =0, and ε_1, ε_3 with ε_2=0, where the principal directions are 1, 2 and 3 with the notation of Figure 6.10. In a test specimen the non-zero stresses and strains are σ_1 and ε_1, ε_2, and ε_3.

In a non-elastic situation with proportional and reversed loading a deformation theory of plasticity gives approximately correct solutions. Here, effective strains and stresses, $\bar{\varepsilon}$ and $\bar{\sigma}$ are introduced giving for elastic and plastic parts

$$\bar{\varepsilon} = \bar{\varepsilon}_e + \bar{\varepsilon}_p = \frac{\bar{\sigma}}{E} + \bar{\varepsilon}_p = \frac{\bar{\sigma}}{E} + \left(\frac{\bar{\sigma}}{H'}\right)^{1/n'}$$

using the Ramberg-Osgood model for the non-linearity of the material. Here

$$\bar{\varepsilon_p} = \frac{\sqrt{2}}{3}\left[\left(\varepsilon_{p1} - \varepsilon_{p2}\right)^2 + \left(\varepsilon_{p2} - \varepsilon_{p3}\right)^2 + \left(\varepsilon_{p3} - \varepsilon_{p1}\right)^2\right]^{1/2}$$

and

$$\bar{\sigma} = \frac{1}{\sqrt{2}}[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]^{1/2}$$

with ε_{pi} and σ_i (i=1,2,3) denoting the plastic strains and the stresses in the principal directions.

With the usual notation i,j,k for permutation of 1,2,3, and since

$$\varepsilon_i = \varepsilon_{ei} + \varepsilon_{pi} = \frac{1}{E}\left(\sigma_i - \nu\left(\sigma_j + \sigma_k\right)\right) + \left(\frac{1}{E_T} - \frac{1}{E}\right)\left(\sigma_i - \frac{1}{2}\left(\sigma_j + \sigma_k\right)\right)$$

one can write

$$\varepsilon_i = \frac{1}{E_T}\left(\sigma_i - \hat{\nu}\left(\sigma_j + \sigma_k\right)\right)$$

With

$$\frac{1}{E_T} = \frac{\bar{\varepsilon}}{\bar{\sigma}}$$

$$\frac{1}{E_T} - \frac{1}{E} = \frac{\bar{\varepsilon}_p}{\bar{\sigma}}$$

And

$$\hat{\nu} = \left(\nu\bar{\sigma} + 0{,}5E\bar{\varepsilon}_p\right)/E\bar{\varepsilon}$$

Further calculations also give

$$\bar{\varepsilon} = \frac{1}{\sqrt{2}\,(1+\hat{\nu})}[(\varepsilon_1 - \varepsilon_2)^2 + (\varepsilon_2 - \varepsilon_3)^2 + (\varepsilon_3 - \varepsilon_1)^2]^{1/2}$$

Hence, a stress and strain state can be treated as elastic with the tangent modulus E_t and a modified Poisson´s ratio $\hat{\nu}$ being a weighted mean between the elastic one, and 0,5, which is valid for perfect plasticity where no volume change takes place. Further details about deformation plasticity theory can be found in textbooks on the subject e g [6.2].
With this tool-box at hand one can see, as mentioned above, that for a uniaxial stress in thin structures or in a test specimen the effective strain is retained, since insertion of σ_1, with $\sigma_2=\sigma_3=0$, and $\varepsilon_2=\varepsilon_3=-\hat{\nu}\varepsilon_1$ in the formulas for $\bar{\sigma}$ and $\bar{\varepsilon}$ give that

$\bar{\sigma} = \sigma_1$ and $\bar{\varepsilon} = \varepsilon_1$.

For the state of plane strain, with σ_1, σ_2 and σ_3=0, and ε_1, ε_3 and ε_2=0, however, one gets $\sigma_1 = \frac{\bar{\sigma}}{\sqrt{1-\hat{\nu}+\hat{\nu}^2}}$

and $\varepsilon_1 = \frac{\bar{\varepsilon}(1-\hat{\nu}^2)}{\sqrt{1-\hat{\nu}+\hat{\nu}^2}}$

For any pair of values $\bar{\sigma}$ and $\bar{\varepsilon}$ on the material curve

$$\bar{\varepsilon} = \frac{\bar{\sigma}}{E} + \left(\frac{\bar{\sigma}}{H'}\right)^{1/n'}$$

a value of $\hat{\nu}$ can be evaluated and then a pair of values σ_1 and ε_1.

These pairs constitute a curve which should be the one used to find ε_1 (and σ_1) in combination with Neuber´s rule. This means values of ε_1 which may be some 10 per cent lower than the ones without considering the plane strain effect.

Now, it is probable, a hypothesis though, that damage is more related to the plastic work than to ε_1, which might compensate for the difference.

Hence it can be concluded that the tacit use of the standard material curve in connection with Neuber´s rule gives sufficiently accurate values of ε_1 to be used in the Coffin-Manson strain-life relationship. The analysis demonstrates the approximate nature of the procedure, using a non-proven "rule" as well for the notch analysis as the same material relationship for uniaxial and plane strain in elastic plastic cases.

If the structure under consideration can be analysed with numerical methods (e g FEM) and using a relevant constitutive model for the material better strain values should be possible to obtain for use in the Coffin-Manson relationship.

6.8 Concluding remarks

The procedures described above to resolve strain cycles for use in strain based fatigue life assessments are useful but deceivingly precise. One must bear in mind that several model assumptions are used.

The assumption that the hysteresis loops have the same shape as the cyclic stress-strain curve is not physically self-evident. Experiments on many materials for the evaluation of the cyclic stress strain curve show, however, that this assumption can be used as a good approximation.

The validity of the Neuber formula was shown for certain cases and is tacitly assumed to be approximately valid in general for manual calculations. In numerical applications this uncertainty is avoided since the function $\varepsilon = F(S)$ is directly obtained.

It has been discussed before that the Palmgren-Miner rule cannot be valid. Still it is used in nearly all applications of damage accumulations.

(Also the models for inclusion of mean stresses in the strain life relationship are a bit artificial and extrapolated from experiments. In this connection it is mentioned that the importance of taking mean values into account is probably smaller when the strain amplitude is large and essentially plastic. This is due to the fact that large strains require negative stresses in order to obtain a wide hysteresis loop and the mean stress should not be so significant.)

Sometimes it is suggested that K_f be used instead of K_t, particularly for large stress concentrations, in order to get a better agreement between calculations and experiments. This may be motivated in the same way as before, that damage is created in a "test specimen" at the notch root and has to have some geometric extension, thus being in fact exerted to stresses and strains somewhat lower than those given by K_t.

Finally it is noted that FEM treatments are simplified by the assumption of uniform hysteresis loops. It is sufficient to perform just one elastic plastic run using a constitutive model of the material with the cyclic stress strain curve as the uni-axial relationship to obtain the function $\varepsilon = F(S)$, which can then be used for simplified calculations of strains changes as a function of the load sequence S(t).

References

[6.1] Dowling, N. E. "Mechanical Behavior of Materials", 3rd edition, Pearson Prentice Hall, 2007. ISBN 0-13-186312-6.

[6.2] Hill, R. "The mathematical theory of plasticity", Oxford University Press (1985) Paperback ISBN 019856 162 8

[6.3] Johnson, W. and Mellor, P.B. "Engineering Plasticity" Van Nostrand Reinhold Ltd (1983)

[6.4] Matsuishi, M. and Endo, T. "Fatigue of Metals Subjected to Varying Stress", Japan Society of Mechanical Engineering, (1968)

[6.5] ASTM STP 465 and ASTM E 606

[6.6] Manson, S.S. "Fatigue: a Complex Subject-Some Simple Approximations", Exp. Mech. (1965) 5(7) pp 199-226

[6.7] Neuber, H. "Theory of Stress Concentrations for Shear Strained Prismatic Bodies with Arbitrary Non-linear Stress Strain Law", J. App. Mech. (1961) pp 544-550.

7 The fracture mechanics approach to fatigue

7.1 The need for fatigue analysis of cracked structures

In the stress- and strain-based approaches to fatigue in Chapters 5 and 6, the analysis results in a total life assessment, presuming a safety factor in load to safeguard for statistical effects. Often “infinite life” is required in the stress based approach.

However, there is a number of situations where it is very useful, even necessary, to study the actual crack growth in order to assess a safe number of load cycles.

This is when a crack is present already when a structure is put into service, as is the case for welded parts. Due to the heating and cooling during the weld process small cracks with lengths in the range 0.1-0.5 mm are “always” present. Then the design is dependent on an analysis of how such cracks can grow under the intended life of variable loading. Due to the importance of welded structures their design is treated formally in rules and standards, as outlined in Chapter 13.

A more general approach when it is known that cracks will have a high probability to grow during service, from initiation at impurities or from short cracks, is to use fracture mechanics. The aim then is to keep control of the influence of cracks so that they do not cause accidents. An essential area of application is the aviation and space industry where weight conditions mean that the load bearing capacity of the material in the structure has to be used efficiently.

Th design principle is now based on two presumptions. One is that cracks exceeding some small length, a_0, can be detected with a high degree of probability by non-destructive methods, as X-ray or ultra-sound. The other is that the number of cycles of an applied load spectrum can be determined for the growth of the crack from a_0 to a critical crack-length, a_{cr}, for which catastrophic fracture occurs.

Inspection intervals are then determined in the following way. If the inspection of a component does not reveal any cracks, it is assumed that, at worst, a crack of length a_0 exists. Then the number of load cycles, or repeated sequences of cycles (e g flights) is determined that can at most grow the crack to a length a_{cr} s, where $s < 1$ is a suitably chosen safety factor. This number of cycles or sequences defines the inspection interval. If the following inspection does still not reveal any crack in the component, there was obviously no crack close to a_0 at the previous inspection and a new service period can be commenced. If a crack longer than a_0 is detected the component is repaired or replaced. In this way a structure can be given a long service life.

In order to find the necessary relationship between crack growth and sequences of variable load the use of fracture mechanics has evolved as a versatile tool. The basic principles are outlined in what follows.

7.2 Fracture mechanics preliminaries

Many serious failures are caused by brittle fracture due to the presence of cracks. Failure occurs due to an overload, as when a glass is dropped on the floor, or when the crack has grown under repeated service loading to a critical size when stability is lost.

Hence, the theoretical and experimental development of the knowledge of the behavior of cracks in loaded structures, *fracture mechanics*, has become an important branch of solid mechanics.

The foundation was laid down during the period 1900-1920. Kolosov, Inglis and Griffith gave important contributions. The theories were developed and adjusted for practical use in design by Irwin and others in the 1960-ies by the formulation of linear elastic fracture mechanics. The basics are outlined in this section, as a background for fatigue applications.

Later, during 1970-1980 and onwards the theory and application of fracture mechanics was extended into the range of non-linear material behavior by Rice and many others, driven not least by the safety requirements of the energy industry with nuclear power plants and pipe-lines.

A comprehensive description of the development of fracture mechanics including the works of above-mentioned authors can be found in many textbooks on fracture mechanics e g Kanninen and Popelar [7.1].

7.2.1 A fundamental problem and the essence of its solution

First, consider a sharp slit, a crack, of length 2a in an infinite linearly elastic sheet, which is loaded uniaxially by a uniform stress σ perpendicular to the direction of the slit, according to Figure 7.1.

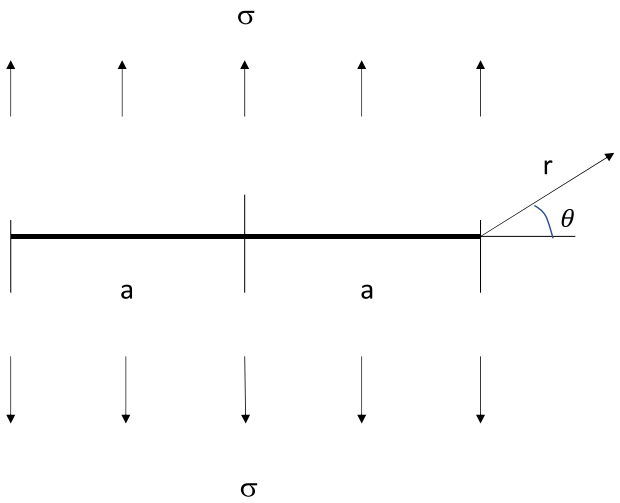


Figure 7.1 The elementary case for fracture mechanics analysis.

This constitutes a problem in the realm of the mathematical theory of elasticity that has a closed form solution. It is not trivial and the solution is best achieved by analytic function theory and conformal mapping. The main result is that stresses and strains at the tips of the crack can be written as

$\sigma_{ij} = \frac{K_I}{\sqrt{2\pi r}} h_{ij}(\theta)$ + an infinite series of non-singular terms in powers of *r*, and where i,j may denote polar (r, θ) or Cartesian coordinates with origin at the crack tip. For each stress component σ_{ij} the $h_{ij}(\theta)$:s are closed form trigonometric expressions, which may be found for either set of coordinates in any textbook on fracture mechanics. The expression means that the

conditions close to the crack tip, where the singular term dominates, are completely described by K_I.

The parameter K_I of the singular term can be determined from geometry and external loading and in the case in Figure 7.1 the solution is $\sigma\sqrt{\pi a}$. K_I is called the *stress intensity factor* and its dimension is $stress\sqrt{length}$, from which it is seen that it increases proportionally with external loads. (The case in Figure 7.1 is one of few having a closed analytical solutions.)

When the load, and hence K_I, is increased to a level K_{Ic} for which the stresses cause material separation on the micro-scale, the crack starts growing. K_{Ic} is a temperature dependent material property called the *fracture toughness*. This means that the problem of finding the conditions for fracture is reduced to finding, by experiments, the K_{Ic} of the material, and, by some analytical, numerical or experimental procedure, the K_I for the load and geometry in question and to see when the load reaches a level when $K_I=K_{Ic}$. The crack growth may be stable or unstable, "catastrophic", depending on the boundary conditions.

Due to the properties of analytic functions, the square root singularity is an inherent property of the stress field at arbitrarily shaped edges of sharp slits in loaded elastic bodies. Further, due to the principle of superposition in linear systems the stress field at a point of an arbitrarily loaded crack front can be described by three modes, see Figure 7.2. ; the symmetric one already treated, the sliding mode and the shearing mode, characterised by stress intensity factors K_{II} and K_{III} respectively.

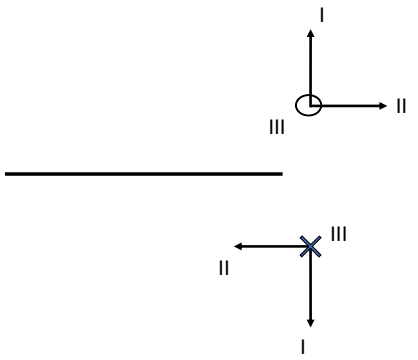


Figure 7.2 The three loading modes at a point of a crack front.

For K_{II} and K_{III} there are expressions similar to $h_{ij}(\theta)$ for the singular terms. For complex loads and geometries, where more than one mode is active, the criterion for fracture becomes more complicated and there are several suggestions for combinations of K-values that are to be compared with K_{Ic}, see Chapter 7.2.2. In the following applications for fatigue analysis the treatment is limited to crack growth under symmetric, K_I-loading, since growing fatigue cracks tend to adjust to this mode.

The determination of K_I can be made in several ways:

- Analytical methods, using analytical functions and conformal mapping, for some principal cases with simple geometries and boundary conditions, as the fundamental case giving $K_I = \sigma\sqrt{\pi a}$. This case is often used as a reference case, and stress intensity factors for other

cases are given in terms of this factor multiplied with a numerical expression including measures of the geometry of the structure,

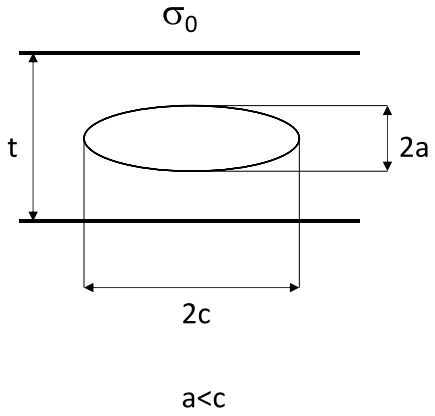
$$K_I = \sigma f(geometry)\sqrt{\pi a}$$

In many cases the expressions f(geometry) are of the order of unity, particularly when the crack is small in relationship to the structure.

- Series solutions. When r grows and the boundary of a (finite) structure is reached the complete series solution may, in principle, be obtained by fitting it to the boundary conditions. Use of this principle constitutes ways to find K_I for complex geometries.

- Experimental methods (Moiré and stress-optical methods)

- FE (Finite Element) or BE (Boundary Element) methods. FE methods (FEM) appear to give good results. This is due to the fact that FEM essentially is a way to minimise the potential energy of a discrete approximation of a loaded structure, and that the quantity K_I^2/E is the energy release per unit crack growth, see Section 7.2.4. Hence the change in energy between two FEM solutions with slightly different crack lengths gives a good measure of K_I. *Many commercial FEM codes have K_I-determination as an option.*



$$K_{Imax} = \sigma\sqrt{\pi a}\, f\left(\frac{a}{c}, \frac{a}{t}\right)$$

$$f\left(\frac{a}{c}, \frac{a}{t}\right) = Q^{-1/2}\left[M_1 + M_2\left(2\frac{a}{t}\right)^2 + M_3\left(2\frac{a}{t}\right)^4\right]$$

$$M_1 = 1$$

$$M_2 = 0{,}05/\left(0{,}11 + \left(\frac{a}{c}\right)^{3/2}\right)$$

$$M_3 = 0{,}29/\left(0{,}23 + \left(\frac{a}{c}\right)^{3/2}\right)$$

$$Q = 1 + 1{,}464\left(\frac{a}{c}\right)^{1{,}65}$$

Figure 7.3 Example from a handbook giving the maximum stress intensity factor for an elliptical crack embedded in a large plate with thickness t. Stress s_0 is perpendicular to the plane of the figure. (From Newman, J.C. and Raju, I.S. (1981)).

There is an abundant literature, e g [7.2], and of course the Internet, where methods and results may be found. Results are generally given as diagrams and/or polynomial formulas, with limitations on parameters for validity and accuracy. A few handbook solutions for K_I are given in Appendix A3. An example is given in Figure 7.3. It is noted that K_I varies along crack fronts, as in the Figure or along the front of an elliptic surface crack, which is common in real cases, a so called "thumb nail" crack. Then it is assumed that crack growth is initiated at points where K_I is maximum.

7.2.2 Mixed mode fracture conditions

As mentioned above the influence of a set of external loads on the crack tip in a component can be separated into the three basic modes at each point along the crack front. There will be material dependent limiting values K_{IIc} and K_{IIIc} also for K_{II} and K_{III} when cracks start to grow in the pure modes. Those critical values are not normally available.

The main issue is instead to find an expression in terms of K_I, K_{II}, and K_{III} which can be compared with the principal and mostly available material parameter K_{Ic} to assess the condition for crack growth. This has been the subject of considerable research [7.3], and no definitive conclusion has been reached. A common suggested formula is based on the energy release rate (see also below in Chapter 7.2.4), which leads to

$$G = K_I^2 / E^{'} + K_{II}^2 / E^{'} + (1+\nu) K_{III}^2 / E \qquad (7.1)$$

and the criterion for fracture

$G = G_{Ic} = K_{Ic}^2 / E'$, where $E' = E/(1-\nu^2)$ for plane strain and E for plane stress. Plane strain generally prevails for parts of crack fronts inside components.

When an initiating small crack is not aligned perpendicularly to the largest principal stress, e g a small interior flaw in the form of a flattened inclusion as in Figure 7.4, the successive crack growth will generally tend to align rather soon. The start in the form of a kinked crack tip may come rather quickly and be the consequence of the complicated conditions, stress fields and inhomogeneities, but then change to conditions calculable by using only K_I and the largest principal stress.

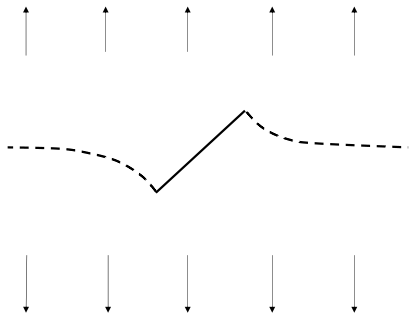


Figure 7.4 Originally slanted crack aligning perpendicularly to the stress field.

7.2.3 Additional ways to obtain stress intensity factors from elementary results

The principle of superposition in linear systems can be used for finding stress intensity factors for complex cases, which are not covered by handbooks.

For example, an arbitrary combination of tension and bending on a beam with an edge crack can be treated as soon as the cases of pure tension and pure bending are known. See Figure 7.5 a)

A further example is the observation that a crack can be considered as closed, "non-existing" by applying the correct set of crack face stresses. So, the bending mode of the beam in Figure 7.5 a) can in its turn be seen as the sum of two cases, see Figure 7.5 b). Here B is elementary, and C has only a local stress field, which makes a numerical solution by e g FEM faster since a less dense grid can be used in most of the structure.

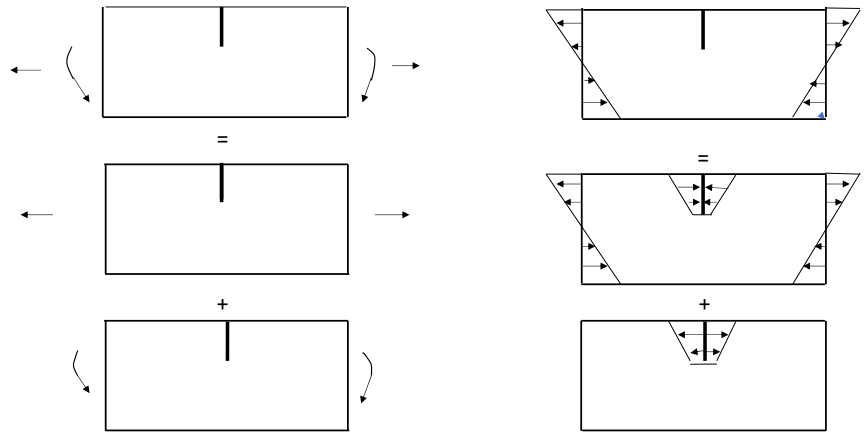


Figure 7.5 Examples of superposition.

The other principal feature is that linearity of a system implies symmetry of the tensor relating load and deformation quantities, leading to the principle of reciprocity. This constitutes a powerful tool to obtain solutions also in fracture mechanics. Basically, the principle says that for two points, i and j, in an elastic system the forces P_i and P_j, and the displacements d_i^j and d_j^i, in the directions of the forces, are related as

$P_i d_i^{\ j} = P_j d_j^{\ i}$ where d_i^j is the displacement at i caused by P_j and vice versa

Besides the direct use to obtain solutions for forces and deformations by combination of handbook cases, the principle can also be used to derive a useful relationship for stress intensity factors (see e g [7.1]), viz

$$K_I^{(2)} = \frac{4\mu}{(\kappa+1)K_I^{(1)}} \int_{C_T} T_i^{(2)}(s) \frac{\partial u_i^{(1)}(s)}{\partial a} ds \qquad (7.2)$$

Here, for plane cases, $\mu = E/2(1+\nu)$ and $\kappa = 3-4\nu$ for plane stress and $\kappa = (3-\nu)/(1+\nu)$ for plane strain. The formula means that if for the load case (1) with $K_I^{(1)}$ displacements $u_i^{(1)}(s)$ are known as a function of crack length a on the part of the body C_T where tractions T_i are prescribed, $K_I^{(2)}$ for a second case (2), where the tractions are $T_i^{(2)}(s)$,is obtained from the formula. The formula can be generalized to boundaries where displacements are prescribed and to body forces, as well as to three dimensions. The formula (7.2) hence is a simplified version in order to demonstrate the principle.

Example:

For a crack of length l in an infinite sheet and with a uniform normal stress $\sigma^{(1)}$ on the crack surfaces (equivalent to a uniform stress at infinity, due to superposition) the stress intensity factor is
$K_I^{(1)} = \sigma\sqrt{\pi l/2}$ and the displacement in the direction perpendicular to the crack is, from the Kolosov-Inglis basic solution,

$$u^{(1)}(s) = \pm\frac{\kappa+1}{4\mu}\sigma\sqrt{(l-s)s}\ , \qquad (7.3\text{ a})$$

where s is now a coordinate along the crack (see Figure 7.6 a)), giving

$$\frac{\partial u^{(1)}(s)}{\partial l} = \pm\frac{\kappa+1}{8\mu}\sigma\sqrt{\frac{s}{(l-s)}} \qquad (7.3\text{ b})$$

Insertion into (7.2) then gives a relationship to find the stress intensity factor for an arbitrary normal stress distribution $\sigma^{(2)}(s)$ at the crack surfaces, see Figure 7.6 a):

$$K_I^{(2)} = \sqrt{\frac{2}{\pi l}}\int_0^l \sigma^{(2)}(s)\sqrt{\frac{s}{(l-s)}}ds \qquad (7.4)$$

This formula can be used for example for calculating the stress intensity factor in a beam under a bending moment and with a small crack, substituting beam stresses with internal crack stresses by superposition according to Figure 7.6 b). The principle is useful for a large set of problems normally not covered in handbooks.

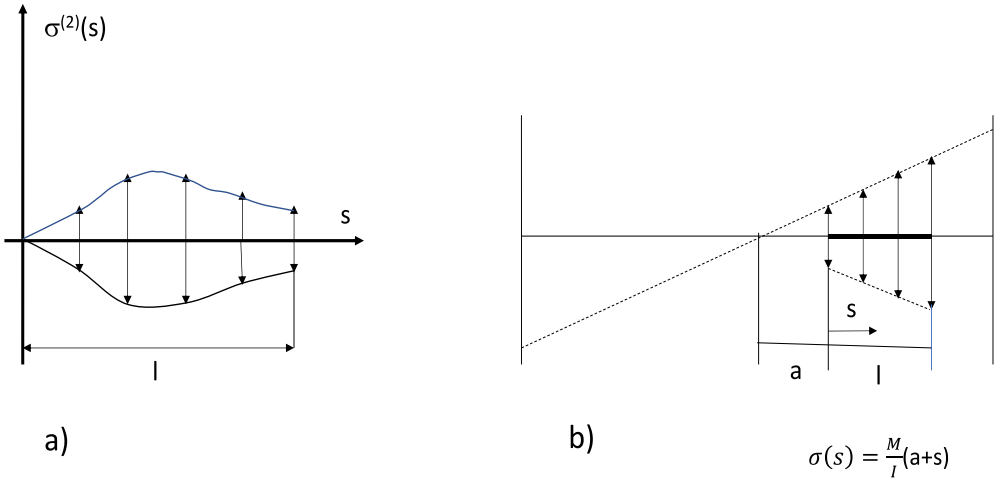


Figure 7.6 a) Coordinates and example loads for application of (7.4), b) Example with superposition of bending stresses in a cracked beam, bending moment M and moment of inertia I, demonstrating a case where (7.4) can be applied. (Strictly speaking, boundary effects must be considered but it is of some value to compare the trapezoidal distribution of stresses with a uniform one.)

7.2.4 An alternative fracture mechanics approach

Griffith approached the fundamental problem in a different way, using the potential energy, U(l) of an infinite sheet loaded by a uniform stress σ, as before, and containing a crack where l is the crack length (=2a). U(l) for the equivalent case of uniform internal stress along the crack surfaces, as in the example above, can be calculated from (7.3 a) and is for plane strain

$$U(l) = -\frac{\pi\sigma^2 l^2}{2E'} \qquad (7.5)$$

For a virtual growth of the crack tip dl there will be a release of energy $\frac{\partial U}{\partial l}dl$ per unit thickness. For the growth it is necessary to separate the crack surfaces which requires surface energy $2\gamma dl$ per unit thickness and where γ is the specific surface energy. If the two terms are equal, the

conditions for crack growth are fulfilled, and for smaller loads giving less potential energy and release rate $\frac{\partial U}{\partial l}$ the situation is stable. So, the criterion for crack growth will be

$$-\frac{\partial U}{\partial l} = 2\gamma = \frac{\pi\sigma^2 l}{E'} \qquad (7.6)$$

From comparison with the first member of (7.1) it is seen that the surface energy, the release rate at growth, is K_I^2/E'. (The two notations G and 2γ for the energy release rate have historical reasons.)

The energy approach enables assessment of the stability of the crack when the growth criterion is fulfilled.

Example.
Consider the double cantilever beam in Figure 7.7

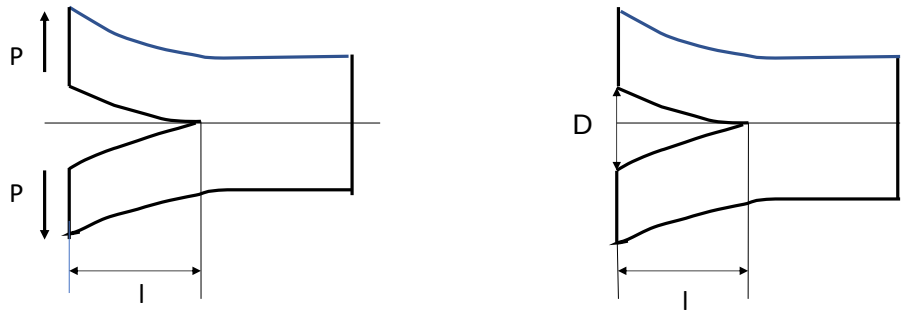


Figure 7.7 Example of assessment of stability conditions of crack growth for a double cantilever beam loaded by prescribed forces and prescribed displacements.

The separation of the two beams under a pair of loads P is

$\Delta = \frac{2Pl^3}{3EI}$. Then for prescribed loads (7.7)

$$U = -\frac{P\Delta}{2} = -\frac{P^2 l^3}{3EI}, \text{ and } -\frac{\partial U}{\partial l} = \frac{P^2 l^2}{EI}, \quad \text{and } -\frac{\partial^2 U}{\partial l^2} = 2\frac{P^2 l}{EI} \qquad (7.8)$$

i e, the change of energy release rate at growth is positive and then the situation is unstable, the crack continues to grow unstably.

For prescribed displacements

$$U_\Delta = \frac{P\Delta}{2} = \frac{3\Delta^2 EI}{4l^3}, \text{ and } -\frac{\partial U_\Delta}{\partial l} = \frac{9\Delta^2 EI}{4l^4}, \text{ and } -\frac{\partial^2 U_\Delta}{\partial l^2} = -\frac{9\Delta^2 EI}{l^5}. \qquad (7.9)$$

In this case the change of energy release rate is negative at growth, i e the situation where a wedge is forced in between the beams is stable, which seems natural. Note that, by using (7.7), $-\frac{\partial U}{\partial l}$ is the same in the two cases, so the equilibrium situation is the same.

The energy release rate, is generally connected to the stress intensity factors by (7.1)

$$-\frac{\partial U}{\partial l} = G = \frac{\kappa+1}{8\mu}(K_I^2 + K_{II}^2) + \frac{1}{2\mu}K_{III}^2$$

There is also a connection to the J integral valid for energy release rates at crack extension in non-linear elastic materials by G=J, see e g [7.1].

As mentioned, FE methods are suitable for determination of stress intensity factors. This is related to the fact that the crack driving force $G = K_I^2/E' = -\partial U / \partial l$ is the energy release rate, and that FE methods are in essence methods to find the minimum potential energy of a large system of discrete elements representing a continuum. Various techniques are available and implemented in commercial FEM codes. The simplest application is to release the node at the crack tip, with nodal forces P, an element length ahead of the crack da, and then see what crack opening D that results. Then, per unit thickness,

$$PD/(2da) = G = \frac{1-\nu^2}{E}K_I^2 \qquad (7.10)$$

This works with reasonable accuracy also for rather coarse meshes, which is in contrast to the case of stress concentration factors, for which it is more difficult to obtain accurate results. This is due to the stress gradients that have to be approximated by integration points of elements at sharp notch roots. Intuitively one would think that the conditions would be the reverse.

7.2.5 Conditions for linear elastic fracture mechanics to be valid

So far the reasoning has presumed purely elastic conditions. However
such a thing as a stress singularity cannot exist. This ideal model is close to real material behaviour only for glass and other brittle materials that have a small cohesive zone representing the surface energy γ.

In reality there is more energy consumption at a growing crack tip than the surface energy. Stresses cannot be infinite and there has to be irreversible plastic deformation, meaning energy consumption, at the growing crack tip. As long as the zone where this plastic energy consumption takes place is small in relationship to the crack length the concept of surface energy can be extended to all the non-reversible energy consumption rate required for crack growth, e g plastic energy in small zones around a growing crack tip, p, which gives $(2\gamma + p) = G$. A condition that the zone is small, and that linear fracture mechanics can be approximately applied using the stress intensity factor as the tool for analysis is found from the following reasoning.

Assume that the region where plastic strains occur, having the characteristic size $2r_p$, is embedded in a zone with radius R, fulfilling the conditions that the singular term of the stress distribution is dominating at R. If $R >> r_p$ the conditions at the crack tip, including plasticity, are thought to be described by K_I and the redistribution of stresses due to plasticity does not disturb the conditions at R significantly. See Figure 7.8 a). The reasoning is of a principal nature

and then it is necessary to investigate the quantitative conditions for validity, essentially the meaning of the condition $R>>r_p$.

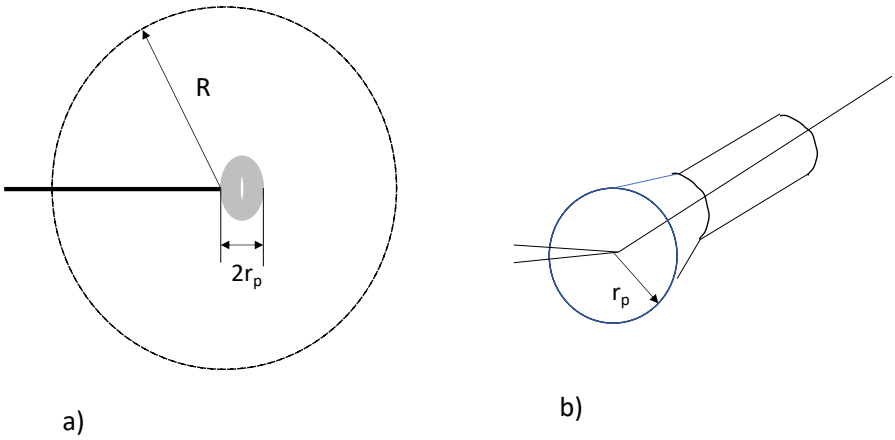


Figure 7.8 a) Relationship between sizes at a crack tip. b) Principal sketch of plastic zone size along a crack front from the surface to the interior of a plate. (See 7.11)

Analytical considerations, assuming the singular term solution show that the extension of the plastic zone at a crack tip is $2r_p$ where

$$r_p = \frac{1}{2\pi}\left(\frac{K_I}{\alpha\sigma_o}\right)^2, \qquad (7.11)$$

and where $\alpha = 1$ for plane stress and $\alpha = \sqrt{3}$ for plane strain, i e the plastic region is smaller for plane strain.

In a plate with a through crack the conditions are plane stress at the surface and plane strain in the interior. The depth of the region with predominant plane stress, the change is continuous, is of the order of the plastic zone radius. The situation is sketched in Figure 7.8 b)

Careful experiments to evaluate the fracture toughness for several geometries and materials show that the results become asymptotic to within a few per cent if crack length, a, thickness, t, (if applicable) and distance from crack tip to the nearest surface, W-a, of the component all fulfil the condition

$$\text{a, t , W-a} > 2{,}5\left(\frac{K_I}{\sigma_0}\right)^2, \qquad (7.12)$$

which has also been adopted as a standardised value for monotonic loading. Hence the factor 2,5 is an agreed value and not a physical constant (it could have been 2,3 or 2,6 or something similar), and should be considered as such.

In fatigue, the situation is complicated by repeated reversed loading and continued crack growth through the plastic region, leading to crack closure and to the fact that the area with reversed plasticity is smaller. These effects are in some respects beneficial, and they will be discussed later on. However, the condition (7.12) is generally adopted as rule of thumb.

Comparing the condition (7.12) with the expression for the plastic zone radius in plane strain it appears that a > $2{,}5(6\pi)r_p \approx 50r_p$, i e the plastic zone $2r_p$ should be less than 4-5% of the crack length. It can also be noted that in many cases K_I is a function of $\sqrt{a}$ times a slowly varying shape factor f(a) of the order of unity and then the condition becomes

$$a > 2{,}5\left(\frac{\sigma f(a)\sqrt{\pi a}}{\sigma_o}\right)^2 = 2{,}5\pi a f(a)^2\left(\frac{\sigma}{\sigma_o}\right)^2 \text{ or } 1 \geq 2{,}5\pi f(a)^2\left(\frac{\sigma}{\sigma_o}\right)^2, \qquad (7.13)$$

giving a guiding relationship for the applied nominal stress vs the yield stress when the size condition is fulfilled.

In fatigue situations plane strain can generally be assumed. For realistic crack growth rates, the stresses and plastic zones are considerably smaller than the plate thickness even for rather thin plates. Further, in many cases the cracks are internal, “penny shaped” or surface cracks with a “thumb-nail” shape, also motivating an assumption of plane strain.

The concept of brittle and ductile fracture in the presence of cracks is illustrated in Figure 7.9 a). A strip with a crack is loaded by a uni-axial stress. When the crack length is small, fracture occurs by general yield of the ligaments between the crack and the edges, shown by the line A-B. For long cracks fracture is governed by the condition that the fracture toughness K_{Ic} is reached. According to (7.11) this corresponds to a certain plastic zone, and when this is considerably smaller than the crack length fracture occurs approximately according to the curve C-D

$$\sigma = \frac{K_{Ic}}{\sqrt{\pi a}}. \qquad (7.14)$$

For intermediate cracks, and higher required stress levels, the plastic zone is no longer small in relationship to the crack length. Further, the stresses become so high that there may be influence from the boundaries when general yield is approached. This bridging region between the clear-cut cases is shown by the dotted curve and represents cases that have to be treated by non-linear fracture mechanics. From the Figure 7.9 it is clear that the choice of the numerical factor for validity of linear fracture mechanics (7.12), the crack length when the true curve is decided to deviate “too much” from the hyperbola, is by no means a precise one, as mentioned above.

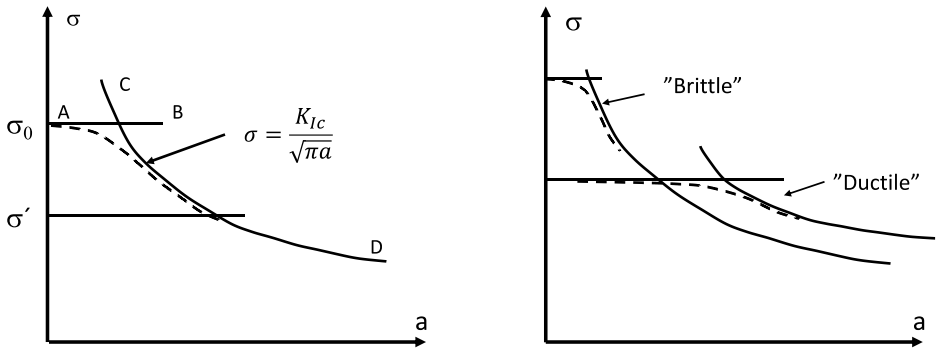


Figure 7.9 a) Bridging of region when crack grows from small under general yield to long where linear elastic fracture mechanics is valid (σ´ denotes asymptotic level according to (7.13)). b) Sketch of relationships between high strength, "brittle", and low strength, "ductile", materials. For "ductile" materials there is a longer intermediate region of crack lengths, and longer cracks, where non-linear fracture mechanics is applicable.

The relationship between the regions in Figure 7.9 a) is dependent on the material. In general, materials with high yield stress have low fracture toughness and vice versa, "the curse of the designer". The situation is shown in Figure 7.9 b). Modern research in materials technology has led to materials, which are both strong and have improved fracture toughness.

The criterion for linear fracture mechanics to be valid, as will be presumed for fatigue analysis in the sequel, thus includes requirements on the crack length. In some instances, the initial cracks may in fact be small, and the stresses high, as in welds with residual stresses of the order of the yield stress. Short cracks in general are discussed in Chapter 3.4 and the practical treatment is described in Chapter 13 on standardised design rules and analyses.

In most practical cases, however, the material properties for crack growth under variable loads mean that linear fracture mechanics can be presumed. Fatigue can be tacitly assumed to be governed by the stress intensity factor K_I for the tensile mode.

7.3 Fatigue analysis with regard to crack growth

7.3.1 The fundamental relationship – Paris´ law

If the conditions for linear elastic fracture mechanics are fulfilled, the stress and strain state, including the non-linearities at the crack tip, is uniquely determined by the stress intensity factor, as discussed in chapter 7.2. The plastic behaviour and damage mechanisms at the crack tip are material dependent. It is then natural to conclude that for cyclic loading with a nominal stress $\Delta\sigma$ of a component with a crack any small growth of the crack during the cycle should be some, material dependent, monotonic function of the corresponding change in the stress intensity factor, ΔK_I. Formally, taking the expression da/dN as denoting an increment of crack growth, Δa, per increment ΔN of cycles, one would then have

$$\frac{da}{dN} = f(\Delta K_I) = f\left(\Delta\sigma f(a)\sqrt{\pi a}\right) \qquad (7.15)$$

Experiments on cracked sheets, see Chapter 10 for details, where the apparent length of a crack on the surface of the sheet is registered by a low magnification microscope as a function of cycle number, result in diagrams like the one shown in Figure 7.10.

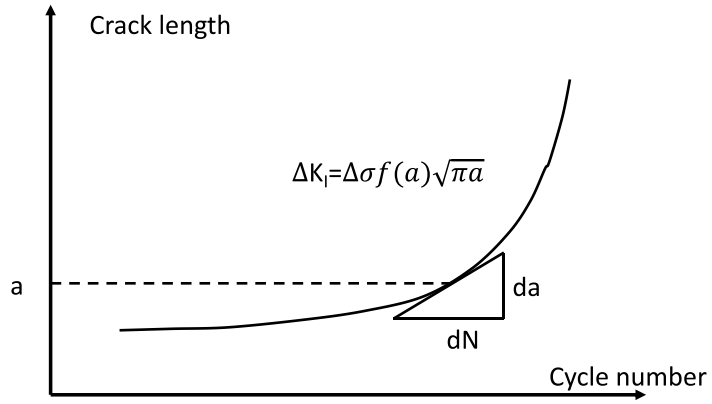


Figure 7.10 Derivation of data pairs (ΔK_I, da/dN) from experiments giving crack length as a function of cycle number.

For different points on the curve the crack length a, and the slope of the curve, da/dN, are registered. From a and the corresponding $\Delta\sigma$ values the variation in K_I can be determined,

$$\Delta K_I = \Delta\sigma f(a)\sqrt{\pi a} \qquad (7.16)$$

where f(a) is the shape function for the specific geometry to be found in handbooks.

Example
For a long strip with width 2W and a central crack with length 2a, and loaded by a uniform stress f(a) is approximately

$$f(a) = \left(sec\frac{\pi a}{2W}\right)^{1/2}\left[1 - 0{,}025\left(\frac{a}{W}\right)^2 + 0{,}06\left(\frac{a}{W}\right)^4\right]$$

which is close to 1 for a/W < 0,2 (then f=1,025).

It appears that for experiments on a large variety of materials the data pairs $(\Delta K_I, da/dN)$ form a sigmoidal curve in a log ΔK_I vs log da/dN diagram, which can be approximated by a straight line for large intervals of values of ΔK_I as in Figure 7.11, ie

$$\log(da/dn) = \log C + n\log(\Delta K_I) \qquad (7.17)$$

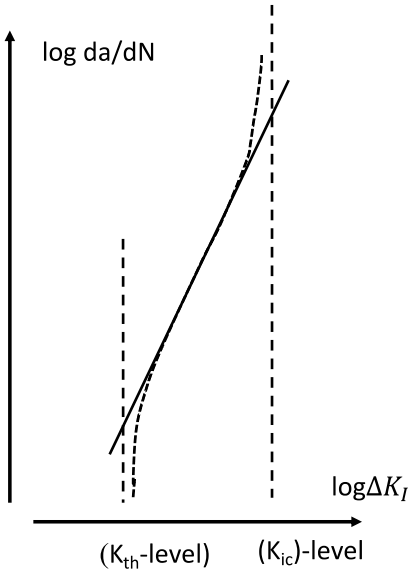


Figure 7.11 The Paris relationship between ΔK_I and crack growth rate da/dN. (Levels for threshold and fracture toughness values are indicated, see Chapter 7.3.3.)

From this the relationship

$$da / dN = C(\Delta K_I)^n , \qquad (7.18)$$

known as Paris´ law, is obtained, where C and n are material parameters. Experiments with different test specimen geometries and varying combinations of

Table 7.1 Crack growth data for some materials. (With da/dN in mm/cycle if ΔK is inserted in $MNm^{-3/2}$.) Data for different forms (hardening, etc) of the same nominal material can differ considerably.

Material	σ_0	$C/10^{-9}$	n	γ
A-533B	490	75	3,3	
RQC	780	0,08	4,2	
AISI 4340	1500 1250	7,2 0,5	3,2 3,2	 0,4
Al 7075-T6	620 520	11,2 27,1	3,9 3,7	 0,64

stress levels and crack lengths have shown that it is solely the stress intensity factor that determines the crack growth rate. This also means that it is not possible to determine the stress levels or crack lengths used to obtain a diagram like the one in Figure 7.11 separately. Data for some materials are given in Table 7.1

Just as the value R=-1 was taken as the standard value in stress and strain based analysis, mainly since experiments were easiest to perform, the value R=0 is the standard value for data in fracture mechanics based analysis. Again, this is due to
the experimental situation. When tests on specimens with cracks are performed it is natural to have the load span from zero, or a small positive value, to a maximum so that a totally closed

crack is avoided. As will be discussed further in Chapters 7.3.3 and 7.3.4 the crack growth rate is strongly dependent on the R-value, to a large extent due to closure effects. It is also seen from Figure 7.11 that there are deviations for low and high ΔK_I-values, and the simple formula (7.15) has to be modified. Efforts to model these effects result in formulas like (7.21) – (7.23).

The exponent n mostly lies in the range 3-4. A number of efforts to derive the exponent theoretically by modelling the plastic conditions at the crack tip, related by Suresh [7.4], give that n=4. The parameter γ in Table 7.1 is used in Chapter 7.3.5 to adjust (7.18) with regard to influence from R for various materials.

7.3.2 Modifications for high and low stress intensity factor levels

Detailed studies, where the stress intensity factor is lowered in steps, reveal a threshold effect. The curve has an asymptotic behaviour and crack growth tends to zero for a material specific value called the threshold value ΔK_{Ith}. This is analogous to the fatigue or endurance limit and the physical reason is the same; the local plasticity, here at the crack tip, is not sufficient to create enough damage in the form of glide bands in the grains at the crack tip to enable continued crack growth. The physical background is treated in Chapter 3.2, where it is also discussed how the R-value and resulting closure effects influence the apparent threshold stress intensity factor.

The experimental technique to find the threshold values is crucial due to plasticity effects creating retardation of growth. As discussed in Chapter 10 the stepping down in ΔK_I then has to be made in small steps and the crack must be allowed to grow through the plastic zone of the previous step. There is a significant risk that a too high threshold value is obtained. Alternative methods have been suggested and applied. They build on starting the cycling with a stress intensity factor well below the expected threshold value and then successively increasing it until crack growth is detected.

Threshold values for steels follow a general trend [7.5]. For R=0 values between 6 and up to10-12 MPa $\sqrt{m}$ are obtained, with the higher values for lower strength, larger grain materials. When R increases the values becomes lower, due to less closure effects (see Chapter 7.3.5), reaching a minimum of around $\Delta K_{Ith} = 2\,MPa\sqrt{m}$ for many materials when R is around 0,7. This indicates that such a high R-value is needed to avoid closure effects completely. This "true" threshold value for a completely open crack tip should in fact vary with the grain size, see Chapter 3.2, but the modelling is a coarse one and it can only be taken as indicative. There are other values for the threshold value for specific materials given in the literature. A relationship representing the behaviour of many steels (except high strength ones) is [7.5]

$$\Delta K_{th}=7{,}0(1-0{,}85R)\ \text{MPa m}^{1/2} \qquad (7.19)$$

When the load is instead increased so that the highest K_I-value in a cycle approaches the fracture toughness K_{Ic}, the da/dN vs ΔK_I curve naturally tends to a second vertical asymptote indicating that the crack grows significantly at each cycle viz.

$$\Delta K_{max}=K_{Ic}(1-R) \qquad (7.20)$$

Then the complete relationship attains the typical sigmoidal shape of Figure 7.11 for R=0, i e when all ΔK=K. The material parameters needed are C, n, $\Delta K_{Ith,R=0} = K_{Ith}$, and K_{Ic}. Still the curved "shoulders" between the Paris´ law region and the two asymptotes remain to be modelled. Formulas for this and parameter values for some materials are given in Chapter 7.3.3.

7.3.3 Comprehensive formulas for design

For design purposes the three lines in Figure 7.11 are not always considered sufficiently accurate. Often large parts of the crack growth occur in a region covering parts of the lower bend of the real curve close to the threshold level, and then this has to be modelled. Just using the Paris´ law from the threshold and upwards would be overly conservative, regarding that the scales are logarithmic.

The upper parts of the curve are generally not so important to model. In this region the crack grows fast, a small portion of the life remains, and the risk for fracture is considerable. Hence the K_{Ic}- value essentially serves the purpose of giving a basis for a safety factor S, so that the $K_{I,max}$ is not allowed, during crack growth, to exceed K_{Ic}/S.

Some suggested formulas are the following. The Forman equation

$$da/dN = \frac{C_F(\Delta K_I)^{n_F}}{(1-R)K_{Ic} - \Delta K_I} \qquad (7.21)$$

modifies with regard to R for the middle and upper region of the da/dN-curve. If the quantity $\left((1-R)K_{Ic} - \Delta K_I\right) da/dN$ is plotted, for various R- and ΔK_I -values versus ΔK_I, and if the formula is a good approximation for the material, it is expected that the points fall on a straight line from which C_F and n_F can be determined. These material parameters are different from C_0 and n in the Paris equation. Some data is given in Table 7.2.

Formulas such as

$$da/dN = C_P\left(\frac{\Delta K_I - \Delta K_{I,th}}{K_{Ic} - K_{I,\max}}\right)^{m_P} \qquad (7.22)$$

and

Table 7.2 Data for the Forman equation for crack growth rate. da/dN in mm/cycle if K-values in $MNm^{-3/2}$. (K_C- values are for plane stress situations.)

Material	n_F	$C_F/10^{-6}$	K_{Ic}	K_C*
INCONEL 718	2,8	4,3	132	-
Al 2024-T3	3,4	2,3	34	110
Al 7075-T6	3,2	5,3	29	78

$$da/dN = C_S\left(\frac{\Delta K_I^{m_S} - \Delta K_{I,th}^{m_S}}{(1-R)K_{Ic} - \Delta K_I}\right) \qquad (7.23)$$

both describe the full sigmoidal curve using just two parameters each for curve fitting, the C:s and the m:s, as soon as K_{Ic} and $\Delta K_{I,th}$ have been determined. The threshold value should be modified for R according to the above reasoning though.

There are also formulas using more parameters for fitting, but these may be considered with some doubt. It should be borne in mind that the formulas are approximations intended to facilitate design work, and they are by no means relationships bearing any physical significance.

7.3.4 Life assessment for a design strategy

Since the load parameter ΔK_I can generally be written as

$$\Delta K_I = \Delta\sigma f(a)\sqrt{\pi a}$$

where $\Delta\sigma$ is a nominal stress, f(a) is a shape function of the geometry and a is crack length, all the relationships of the types exemplified by (7.15) and (7.21) – (7.23) can be written as

$$da/dN = G(\Delta\sigma, R, a) \qquad (7.24)$$

where for constant amplitude loading only a is a variable. Formally they can then be separated to yield

$$\int_0^{N_f} dN = N_f = \int_{a_0}^{a_f} \frac{da}{G(\Delta\sigma, R, a)} \qquad (7.25)$$

and integrated numerically from a starting crack length a_0 to a final, allowed, critical crack length a_f. This is precisely what is needed for the design strategy, the number of cycles used in growing the crack tip from a detection limit a_0 to the longest safe value a_f.

The formula (7.25) is illustrated in Figure (7.12). Since the function $1/G(\Delta\sigma,R,a)$ is generally a smooth function, the numerical integration can easily be performed with standard formulas, e g Simpson´s rule.

For many practical purposes a very useful way is to use the formula for the Paris´ law region, where analytical integration can be applied. Replacing the lower shoulder of the sigmoidal curve with the straight line will, however, as mentioned give a conservative estimate.

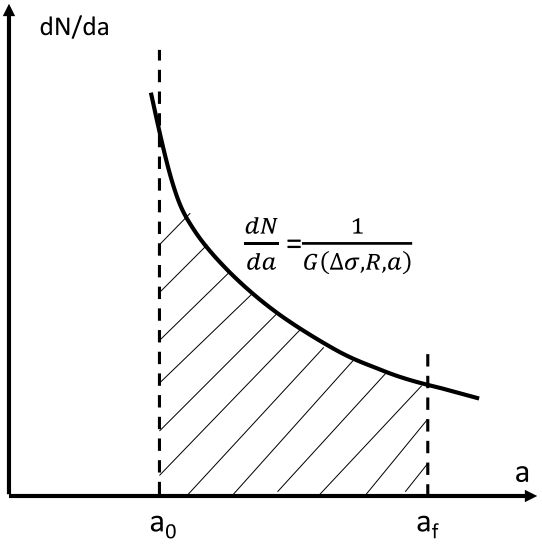


Figure 7.12 Illustration of the numerical integration of (7.25).

The R-dependence is taken into account by replacing C_0 by C_R, see (7.31). Further, for many geometries of practical interest the shape function f(a) is a constant or a slowly varying function for a-values of interest. This follows from the fact that rather small crack lengths are sufficient to be considered critical with normal load levels and geometries, and then f(a) tends to be nearly constant and equal to its value for a=>0.

Then the relationship becomes

$$da/dN = C_R(\Delta\sigma f(a)\sqrt{\pi a})^n,$$

and

$$N_f = \frac{1}{C_R(\Delta\sigma f(a)\sqrt{\pi})^n}\int_{a_0}^{a_f}\frac{da}{a^{n/2}} = \frac{(1-n/2)^{-1}}{C_R(\Delta\sigma f(a)\sqrt{\pi})^n}(a_f^{(1-n/2)} - a_0^{(1-n/2)}) = \frac{a_0(n/2-1)^{-1}}{C_R(\Delta K_{I,0})^n}(1-(a_0/a_f)^{(n/2-1)}) \quad (7.26)$$

This is the basic relationship used for design. In practical cases n is of the order 3 – 4 (for n=2 the integration results in a logarithmic expression). The second form may be practical to use since just the starting values for crack length and stress intensity factor, a_0 and $\Delta K_{I,0}$ are required for finding N_f or a_f as desired. If $a_0 << a_f$, i e if the starting crack is very small (7.26) is simplified further so that the life depends only on a_0.

7.3.5 Influence of the R-value.

There are several reasons why the crack surfaces close before the minimum load/stress intensity factor of a cycle is reached. As described in Chapter 3 the crack tip is moving through the material during cycling, creating crack surfaces that have traces of permanent plastic deformation and of remaining asperities from material separation. This gives rise to material

contact for closure values $K_{Icl}>K_{Imin}$ in the unloading portion of the cycle. The crack closes successively from the tip until K_{Imin} is reached. During a following loading the closure is released until the crack surfaces are fully separated at some value K_{Iop}. The process is sketched in Figure 7.13. Although the values for closure and opening are not exactly equal, and there is a continuing closure process, it seems sufficient to assume one value as characteristic, K_{Iop}.

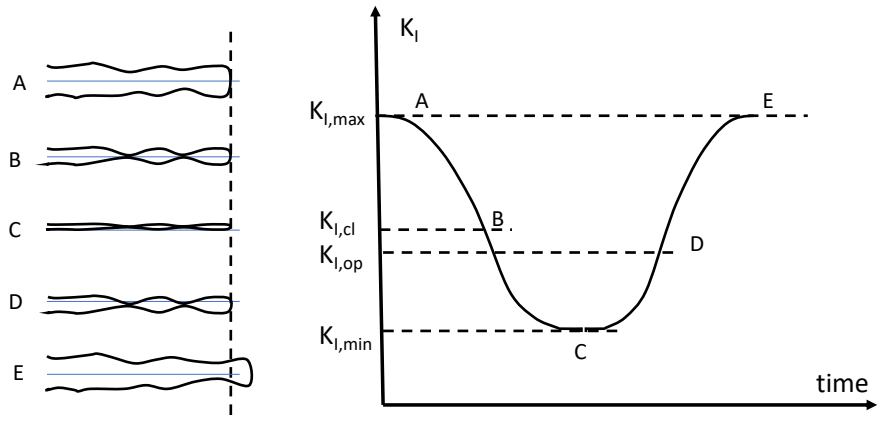


Figure 7.13 Sketch of the closing and opening of a growing crack.

Then the damage processes at the crack tip are shielded and only the upper part of the ΔK_I-interval is active in driving the crack. This occurs even when the minimum stress intensity factor is higher than zero. Careful studies using strain gage and optical methods (by e g Elber [7.8]) lead to suggested values of a closure function U(R) defined by

$$\Delta K_{Ieff} = U(R)\Delta K_I = K_{Imax} - K_{Iop}$$

with

$U(R) = 0{,}5 + 0{,}4R$ for -0,1<R<0,7 (Elber) (7.27 a)

$U(R) = 0{,}75 + 0{,}28R$ for -1<R<0,7 (Maddox) (7.27 b)

The relationships are shown in Figure 7.14 It is not so astonishing that the differences are significant regarding that there should be some material dependence. A theoretical analysis built on the Dugdale strip yield model (Budiansky and Hutchinson [7.6]) forR=0 gives U(0) = 0,52 and U(0) = 0,44 for closing and opening respectively. Still, the experiments and analysis demonstrate that closure effects are significant also for small R-values. If, intuitively one would assume that the crack is open as soon as $K_I > 0$ one would have

U(R) = 1/(1-R) for R<0
U(R) = 1 for R>0 .

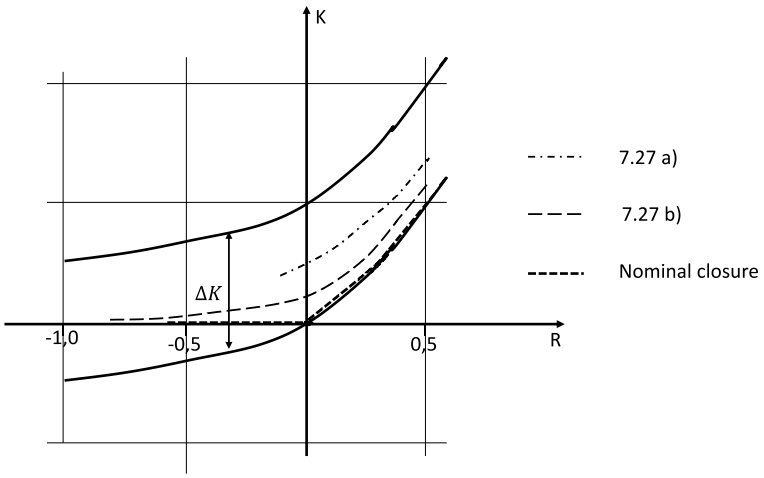


Figure 7.14 R-dependence on effective stress intensity factor.

The findings about the R-value influence for the complete curve in Figure 7.11 can be summarised by saying that it is shifted to the left for increasing R-values, the effective part of ΔK increases when closure decreases.

Several formulas have been suggested for R-value correction of each of the three regions.

For the threshold asymptote (7.19) can be compared with an equation of the Walker type:

$$\Delta K_{I,th} = \Delta K_{I,th,0}(1-R)^{1-\gamma_{th}} \qquad (7.28)$$

where the constants $\Delta K_{I,th,0}$ and γ_{th} are material specific. Values 7,0 MPa$\sqrt{m}$ and γ_{th} = 0,3 give a good fit to (7.19).

The Walker equation is also used for the Paris´ law region, instead of the above formulas for U, in the form

$$\Delta K_I = \Delta K_{I,0}(1-R)^{1-\gamma} \qquad (7.29)$$

where ΔK_I is the actual K_I-span for a loading with a certain R-value, $\Delta K_{I,0}$ is the corresponding span for the same growth rate at R=0, and γ is a material constant, different from γ_{th}. Then, denoting the constant in Paris´ law for R=0 by C_0 one gets

$$da/dN = C_0(\Delta K_{I,0})^n$$

and, by insertion of (7.26)

$$da/dN = C_0\left(\frac{\Delta K_I}{(1-R)^{1-\gamma}}\right)^n = \frac{C_0}{(1-R)^{n(1-\gamma)}}(\Delta K_I)^n \qquad (7.30)$$

This means that the relationship for R=0 retains its slope for all R-values but the constant C_0 is replaced by

$$C_R = \frac{C_0}{(1-R)^{n(1-\gamma)}} \qquad (7.31)$$

which, in the log-log representation gives a parallel line moved to the left, i e increasing growth rates for increasing R. Values for C_0 , n, and γ are given in tables as material parameters, obtained by curve fitting of experimental data. Some examples are given in Table 7.1 (from Dowling [7.5]).

It may be of some interest to compare (7.27), coming from crack closure investigations, with crack growth data leading to the model in (7.29). This would mean to compare the expressions (1+0,28R/0,75) (Maddox) and (1+0,4R/0,5) (Elber) with$1/(1-R)^{(1-\gamma)}$ for different values of γ. Results for γ–values 0,4 and 0,6, see Table 7.1, are shown in Figure 7.15). The comparison gives an impression of the correspondence between model expressions based on different experimental data.

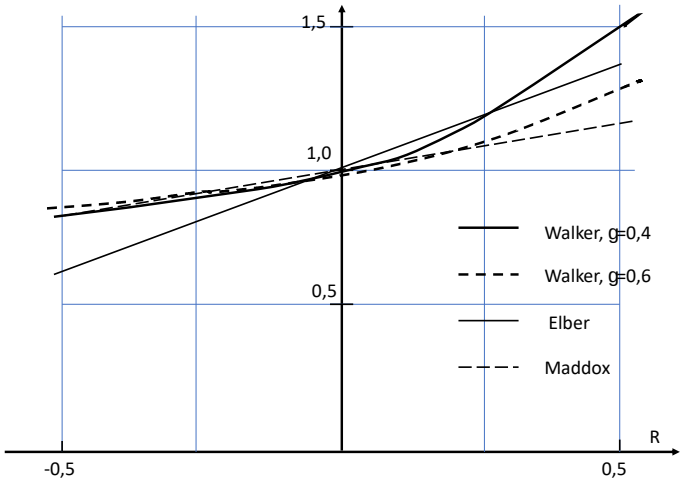


Figure 7.15 Comparison between models in (7.27) and (7.29) for R-effects.

Finally, the asymptote limiting ΔK_I so that $K_{I,max}$ does not exceed K_{Ic} is obtained from $\Delta K_I \le K_{Ic} - K_{I,\min} = K_{Ic}(1-K_{I,\min}/K_{Ic}) = K_{Ic}(1-R)$. The result is summarised in Figure 7.16.

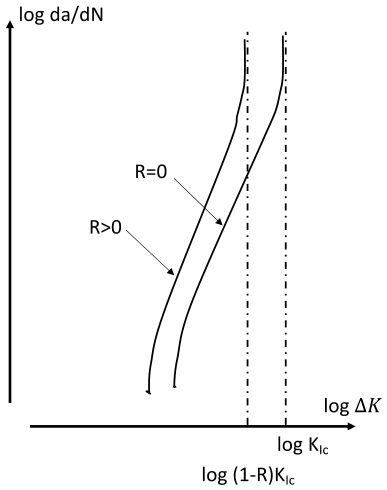


Figure 7.16 Principal sketch of the crack growth curve with modifications for the R-value.

7.3.6 Variable loading.

Just as in the previous treatments of stress and strain based fatigue the R-value is an intrinsic parameter for the fatigue process, and it was used in 7.3.5 for crack growth under constant amplitude loading to characterize closure and plasticity effects. The treatment of variable loads in connection with growth of cracks should also include consideration of such effects since they are so prominently affecting growth depending on the order of large and small cycles in the load sequence. The discussion starts with an elementary example.

Consider block loading with two K-levels and R=0. Both levels have the same closure ratio, see Figure 7.17 a) i e

$$K_{op,1} / K_{\max,1} = K_{op,2} / K_{\max,2}$$

When the load level is decreased the effective K-range is $K_{max,2} - K_{op,1}$, which corresponds to a decrease in the span of K active for incremental crack growth. This means that the growth rate will be smaller until the crack has grown so much that the process is stabilized on the second level, $K_{max,2} - K_{op,2}$. When the load level is then increased again the K-range will at first be $K_{max,1} - K_{op,2}$. This means that the growth rate will be even bigger that the original one until, again, the process is stabilized on the first level.

The changes are illustrated in terms of growth rates in Figure 7.17 b). The effects are called crack retardation and crack acceleration respectively. If the load is decreased sufficiently the crack growth may stop completely, if $K_{max,2} - K_{op,1}$ is below the threshold value. This effect can be used by the method of single overloads of structures. Then a few load cycles on the "1"-level, overload, are applied, and when the load is then decreased to the "2"-level, the service load, the crack does not grow at all.

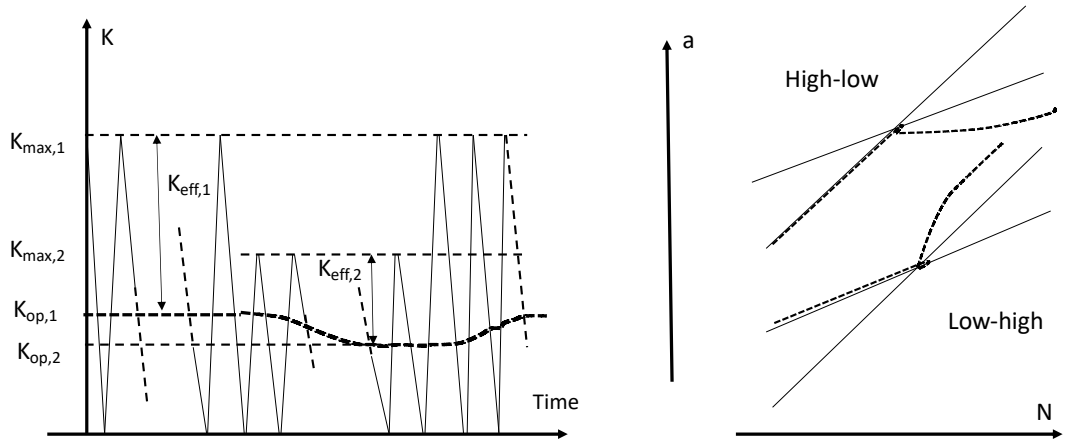


Figure 7.17 Retardation and acceleration effects due to crack closure when changing levels of block loading.

This is a simplified reasoning. In reality there are effects both of closure, due to roughness behind the moving crack tip, and effects of reversed plasticity ahead of it, and these effects are interacting. What is called closure level is just a way to define the effective span of stress intensity factor driving the crack tip. The general approach for defining cycles from an arbitrary, irregular, sequence of maxima and minima of loading is described in Chapter 8. Load spectra defined in this way are used also for fatigue crack growth although there are sequence effects connected to closure and crack tip plasticity.

Two main groups of loading can be distinguished. *One group is where the load is changing between levels of the same order of magnitude frequently and irregularly*. Such cases are treated in Chapter 8.2.3, together with variable loading in cases of stress and strain based analysis. This is the most commonly used approach, and it does not consider sequence effects.

The other group of loading, treated here, is typically where there are single or few overloads in a sequence of lower loads with many cycles, or where blocks of loads with different levels occur after each other. These cases are discussed below in terms of two models. The models are based entirely on the plastic zones ahead of the crack tip and have been suggested by Wheeler and Willenborg [7.8], [7.9] to quantify retardation after overloads, i e how to obtain a stabilized situation as depicted in Figure 7.17. Load sequences with a few large load cycles and many small ones, are typical for e g flight simulations. The models comprise parameters, which have to be determined by experiments.

The Wheeler model is illustrated in Figure 7.18. After a large load cycle, $\Delta\sigma_1$, giving the plastic zone size r_1, smaller load cycles, $\Delta\sigma_i$, commence, giving plastic zones r_i, where i is a counting parameter. Then, with the growth rule for constant amplitude loading

$$\frac{da}{dN} = Cf(\Delta K) \qquad (7.32)$$

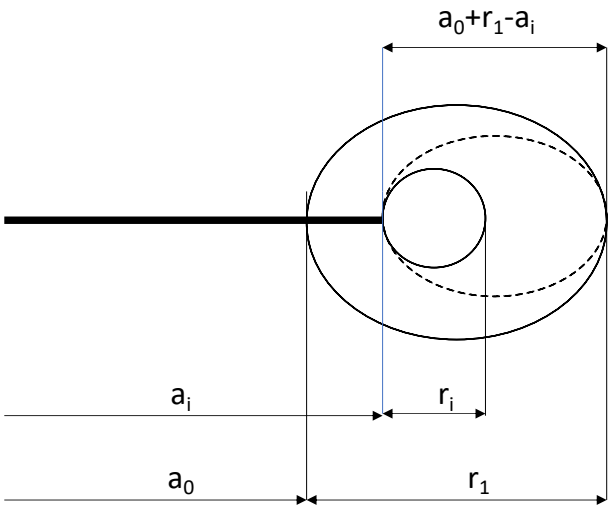


Figure 7.18 Measures used in the Wheeler and Willenborg retardation models.

a modified growth rate is taken as

$$\left(\frac{da}{dN}\right)_i = C_i f(\Delta K_i) \qquad (7.33)$$

and where

$$C_i = C\left(\frac{r_i}{a_0 + r_1 - a_i}\right)^m, \qquad (7.34)$$

until $a_0 + r_1 - a_i = r_i$, i e when the small plastic zone reaches the border of the large one. Further, $\Delta K_i = \Delta K_i(\Delta\sigma_i, a_i)$.

Here m is a material parameter, to be evaluated. The plastic zone sizes are calculated for the various loads, assuming plane stress, as

$$r = \frac{1}{\pi}\left(\frac{K_{max}}{\sigma_0}\right)^2. \qquad (7.35)$$

Then crack growth increments can be calculated to evaluate the influence of the overload.
It is noted that in realistic cases $a_0 \approx a_i$ and then ΔK_i and r_i can be taken as constant, and the main variation in C_i comes from the growth increments in a_i.

The iterative procedure for da_i is cumbersome with updating of C_i and ΔK_i for each step. It may be noted that for a meaningful fatigue procedure small scale yielding must prevail. Otherwise the crack growth is very fast. Then plastic zones are much smaller than the crack length and for a process when a sequence of constant low load cycles cause the crack to grow one plastic zone ahead, r_i may be taken as constant and equal to r. Further, $a_0 - a_i$ may be taken as the integration parameter so that for this case

$$\frac{dN}{dx} = \frac{1}{F(\Delta K,R)}\left(\frac{R_0-x}{r}\right)^m \qquad \text{for } 0<x<R_0\text{-}r \qquad (7.36)$$

This gives that the relationship between the number of cycles during the retardation phase N_R and the number of cycles without retardation N is

$$\frac{N_R}{N} = \frac{1}{(m+1)}\left[\frac{(R_0/r)^{m+1} - 1}{R_0/r - 1}\right]. \qquad (7.37)$$

The model suggested by Willenborg is also founded on the comparison of plastic zone sizes, but in terms of K-values instead of plastic zone sizes. With reference to Figure 7.18 a K-value is first calculated that produces a plastic zone just large enough to reach the front of the one from the overload, i e, from

$$K_{ref} = \sigma_0\sqrt{\pi(a_0 + R_0 - a_i)},$$

again assuming plane stress.

Then, for the low level cycles the maximum and minimum K-values are modified to effective ones from

$$K_{max,eff} = K_{max} - K_{red}$$
$$K_{min,eff} = K_{min} - K_{red}$$

where

$$K_{red} = K_{ref} - K_{max}$$

If $K_{min,eff}$ is negative it is put equal to zero. The effective values obtained are then used to calculate effective ΔK- and R-values, which are finally used in any of the formulas, 7.16, 7.21-7.23, for the crack growth increment of cycle i.

It is observed that, nominally, the ΔK is unaffected and only the R-value is changed. Further, if the level of the low loads are below half the level of the overload crack arrest is predicted, which is not true in all real cases. (This is seen by putting $R_0=4r$ for $a_i=a_0$, i e with double overload and for the first cycle with the low load. Then $K_{ref}=2K_{max}$ and $K_{red}=K_{max}$ giving $K_{max,eff}=0$.)

For both models it has been found necessary to add more parameters. When the parameters have been adjusted from experiments with one type of load sequence and then tried for others the results have appeared to be less successful, ranging in errors of the order of 20-40 per cent both on the conservative and non-conservative side. This can be expected since the models are inherently non-physical. It is also noted that the basic Willenborg model does not have to be calibrated by adjusting a parameter, as the exponent m in the Wheeler model.

More realistic models built on crack closure have been suggested, by e g Elber [7.10]. The simplest one is to assume that a constant closure level of K is obtained during cycles of several load levels, if these are changing reasonably frequently. This level can of course be assumed a

priori or calibrated from type spectra. This the basis for the approach used in Chapter 8.3. The more advanced models built on crack closure require some computational resources.

7.3.7 Short cracks

It should be observed that cases with small crack like defects, where the conditions for linear elastic fracture mechanics are not relevant, can be treated by the techniques in Chapters 7.3.4 – 7.3.6 taking the in Chapter 3.5 into account. There an equivalent crack length is calculated from the dimensions of the defect, and modified values of the endurance limit and threshold stress intensity factor are defined.

References

[7.1] Kanninen, M. F. and Popelar, C. H. *Advanced Fracture Mechanics,* Oxford University press (1985). ISBN 0-19-503532-1

[7.2] Y. Murakami, (Editor-in-chief), Stress Intensity Factors Handbook Volume 2, Pergamon Press (1987).

[7.3] Gdoutos, E.E. and Sih, George C. ”Problems of Mixed Mode Crack Propagation”, Kluwer Academic Publishers (1984) ISBN 9789024730551

[7.4] Suresh, S. ”Fatigue of Materials”, Cambridge University Press (1992)
ISBN 0 521 43763 6

[7.5] Dowling, N.E. ”Mechanical Behavior of Materials” Pearson Prentice Hall (2007) ISBN 0-13-186312-6

[7.6] Budiansky, B. and Hutchinson, J.W. ”Analysis of Closure in Fatigue Crack Growth”, J. Appl. Mech., 45, pp 267-276 (1978)

[7.7] Newman, J.C.”Finite Element Analysis of Crack Growth under Monotonic and Cyclic Loading”, ASTM STP 637, pp 56-80 (1977)

[7.8] Wheeler, O.E. ”Spectrum Loading and Crack Growth” Journal of Basic Engineering, 94, pp 181-196 (1972)

[7.9] Willenborg, J.E. et al. ”A Crack Growth Retardation Model Using an Effective Stress Concept” Air Force Flight Dynamics Laboratory Report AFFDL-TM-71-1-FBR (1971)

[7.10] Elber, W. ”Equivalent Constant-Amplitude Concept for Crack Growth under Spectrum Loading” ASTM STP 595, pp 236-250 (1976).

8 Fatigue damage from variable amplitude loading

In real life the load on a structure is often irregular, and not a constant amplitude variation. Hence, it is important to be able to assess the life of a structure under such conditions. They may be of very differing types.

One typical situation is *when a limited number of high loads must be tolerated in combination with a larger number of lower cycles under routine service conditions*. This occurs for example in operation of aeroplanes, where the start-landing cycle means higher stresses than the normal variations during flight, and in power plants, where planned and emergency stops give high thermal stress cycles as compared with those under service conditions. Another type of loads is *when the variations have the same order of level but a complicated statistical nature, as the loads from waves on ships or loads from uneven roads on cars*.

The analysis of fatigue under varying loads must be compatible with results from constant amplitude theory and tests. Even though modern technology now enables spectrum load tests on structures in laboratory environments, the fundamental study of fatigue damage is made in terms of constant stress or strain cycling. Also, in stress and strain based design, the whole process from creation of persistent slip bands to macroscopic crack growth is linked to data of constant amplitude fatigue life, as S-N-curves and Coffin-Manson data.

Two essential problems have to be approached. First, the irregular load has to be characterized in terms of a number of cycles. Second, a damage rule has to be established and verified, which can make use of constant amplitude data to compare and assess the damage of a sequence of cycles with different amplitudes and mean values.

The problem of characterizing an irregular load sequence is treated in various ways. One is to transform the time function into a number of groups of cycles, each with a mean value and amplitude. This can be done in several ways, which are more or less artificial and not so well suited to model the real situation. Another method is to characterize the load in terms of its stochastic properties and then use approximate methods to retrieve the content in terms of stress or strain cycles. In standards and norms simplified methods are found, where typified spectra adjusted to industrial branches are used for comparison with constant amplitude fatigue data, c f Chapter 13.

The damage accumulation from a set of cycles with different spans of stress or strain, often represented as a so-called load spectrum, should in fact include consideration of the various kinds of "memory effects" which are found to appear in the different phases of initiation and in the growth of small and large cracks, as discussed in Chapter 3 and 7. This is seldom possible to do, and hence the results may be conservative or non-conservative in different cases. Generally, a change from large amplitudes to small ones is found to have a retarding effect for following cycles, and vice versa. In practise, such effects are too complicated to map and the consequences have to be tacitly neglected, necessitating the use of safety factors, or in some cases compensated for by factors established from experience of types of loading in branches as car or airplane industry.

Finally it is observed that the time sequence of loads or stresses must of course be the one acting at points of interest in the structure. External loads may be filtered or amplified through the structure in various ways and a structural analysis may have to be performed as a starting point, e g by the aid of finite element methods.

8.1 Representation of load sequences

In this section some of the most frequently used principles of representation of time sequences of load or stress in terms of sets of cycles for damage analysis are presented and discussed with respect to usefulness, areas of application, and limitations.

8.1.1 Peak-valley and level crossing counting. The concept of load spectra

The oldest and most direct ways to represent a time sequence of stress in terms of something resembling cycles are those of peak-valley and level crossing counting. They are to some extent motivated by the ease to register and store electric signals from gauges before the era of computers.

The peak-valley counting data is obtained from the time sequence by registering how many times certain, pre-determined stress levels are exceeded and undercut by extreme values, i e peaks and valleys respectively. A similar method is the counting of level crossings, i e to register for the stress levels how many times the sequence crosses each level in the upward, or downward, direction. The resulting diagrams are shown for demonstration in Figure 8.1 a) for an overly simple time sequence. With so few levels, and in a model example, there are problems to define the peak and valley values. Here, they are defined as lying just outside their respective level values. It is seen that there are valleys above peaks for some of the low levels, and that peak and valley counting seems to envelope the level crossing count. The interpretation of the diagram is that corresponding maximum and minimum values define cycles. Hence, there is one cycle with mean 0 and amplitude 4, one cycle with mean 0,5 and amplitude 3,5, and so on. In the example it is seen that difficulties arise to define more than five cycles. This is due to the irregularity of the example sequence.

In real situations the sequence may have many hundreds of peaks and valleys and the stress levels can be many more. Then the diagrams will be smoother, as in Figure 8.1 b). It is common to use a logarithmic scale for the number of counts. Diagrams as the one in Figure 8.1 b) are often called load spectra, which may be misunderstood in relation to other representations.

Level crossing data are in fact enveloped by peak-valley data. There are minima above a level σ_0 and maxima below it. The number N_0 is the number of crossings of the σ_o-level upwards, and N_0/N_{tot} =k, is called the irregularity factor, see also Chapter 8.1.3. For the extreme k=1 all peaks are above σ_o and all valleys below it.

The classical interpretation of a load spectrum as the one in Figure 8.1 b) is as the one in the example. In a small interval of cycles $n_i=N_i-N_{i-1}$ where the stress variation is small, the maximum stress is $\sigma_i = \left(\sigma_{N_{i-1}} + \sigma_{N_i}\right)/2$ and the minimum stress is $\bar{\sigma}_i = \left(\bar{\sigma}_{N_{i-1}} + \bar{\sigma}_{N_i}\right)/2$.

Then for all cycle number intervals n_i , where $\Sigma n_i=N_0$, there are stress ranges defined that can now be used for damage analysis, i e how much of the life that is consumed for the load spectrum with N_0 cycles.

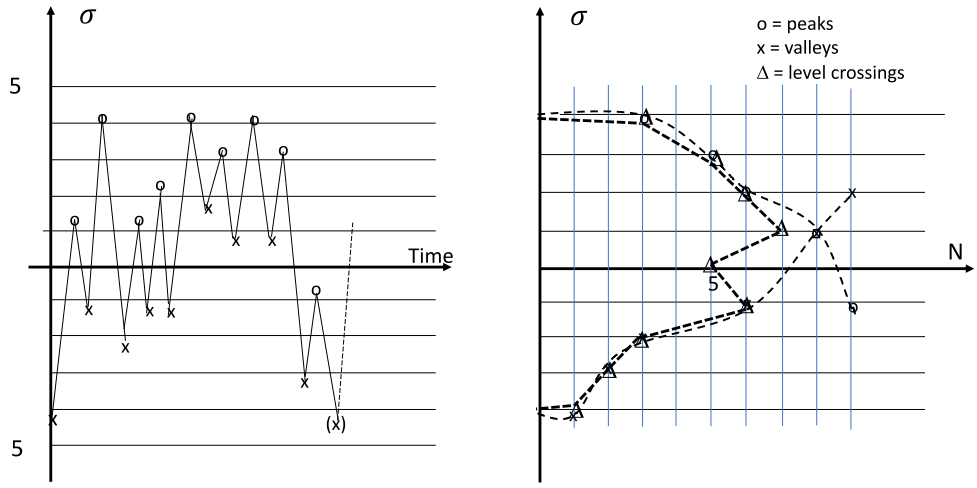


Figure 8.1 a) Peak-valley and level crossing spectra resulting from a short time sequence.

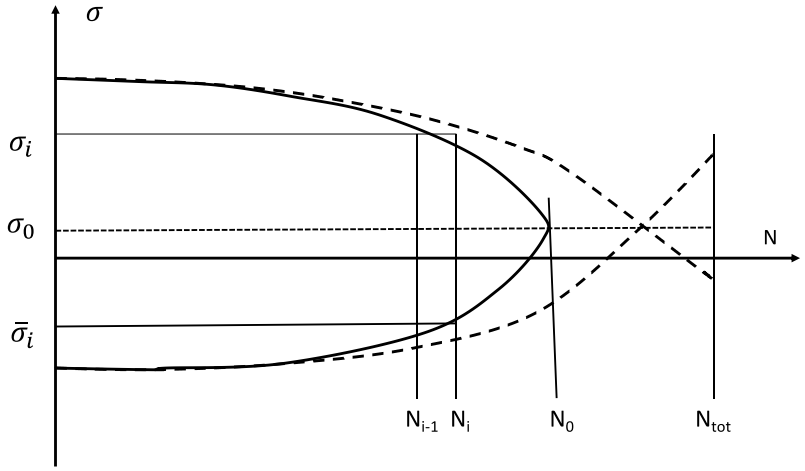


Figure 8.1 b) General representation of cycle counting with the peak-valley and level crossing methods.

An obvious disadvantage is that many small cycles are not included when there is an irregularity significantly smaller than 1. This may not be so serious since small cycles do not contribute much to the damage. There is no obvious one-to-one correspondence between maxima and minima connected to real cycles. This can be appropriate since the consistent grouping of the largest maxima with the smallest minima should be conservative, i e result in a smaller calculated life than the real one.

Another principal weakness, when measured spectra are to be used for experiments, is that the order of cycles, which is known to create retardation or acceleration effects, is not included in the data. However, the blocks of constant amplitude can be determined from experience to be applied in a special order to account as well as possible for memory, acceleration and deceleration, effects. An example of this is the so-called LBF spectrum, see [8.1].

Generally, data exceeding N_0 are truncated in peak-valley counting as well, since they represent comparatively small stress variations, adding little to the damage considering the logarithmic nature of the S-N or Coffin-Manson-curves.

8.1.2 Rain-flow count and range pair count

A method for counting cycles, which takes the connection between peaks and valleys into account in a systematic way, is the so-called rain-flow count method. Its features are most clearly demonstrated in LCF where elastic-plastic stress-strain loops can be distinguished. This is demonstrated in Chapter 6.

The method was originally suggested for this purpose by Endo and Matsuishi [8.2], but it has become one of the most used general methods for cycle counting. The method is attractive since it has been shown to be conservative under certain conditions for the loads, which can be assumed to be at hand in many cases. It is also very simple to apply and to program for automatic evaluation of long sequences of cycles. Although it solves the problem of finding connected peaks and valleys it does not take the sequential effects into account.

The method is most easily described in terms of Figure 8.2 with a short LCF sequence with loops of cycles (Figure 8.2 a)), and a corresponding time sequence of peaks and valleys of stress (Figure 8.2 b)). A similar procedure for strains can of course be developed. Note that the triangular shape is just for simplicity; the shape of the time vs stress curve is insignificant.

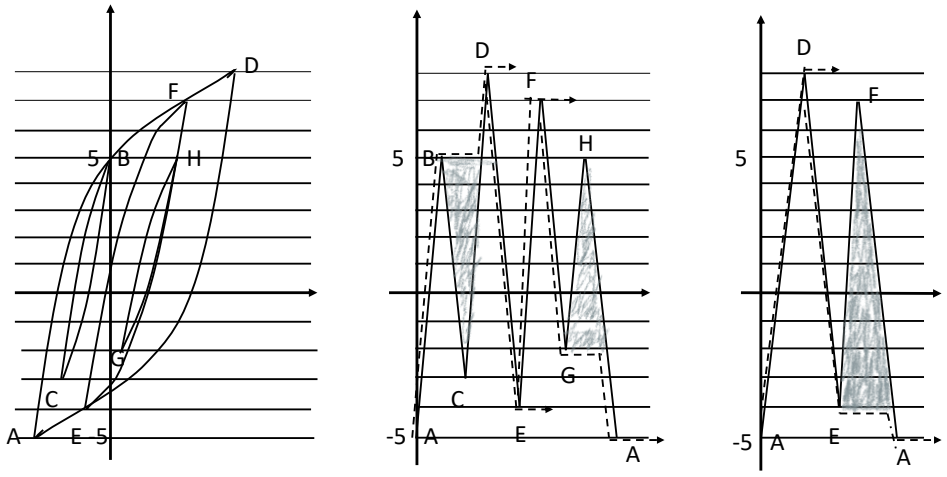


Figure 8.2 The background to rain-flow count. Representation of stresses from a LCF stress-strain sequence.

If the sequence does not start at its maximum or minimum point, as it does in the example, it is first re-arranged so that this becomes the case. Then the time sequence is followed "like a rain drop flowing on a pagoda roof" ("vertical" is to the right in the Figure). When the drop falls into free space, a new drop is started at that turning point (D, E, and F), but when the drop falls onto a part of the roof extending further in the next cycle (B, G) cycles are counted (B, C and G, H), with their mean values and amplitudes, and discarded from the sequence, which then becomes simpler. If there are cycles remaining at the end, the process is repeated, Figure 8.2 c) giving the cycle (E, F), until only the final main loop remains and is included in the set of cycles.

In terms of a point-by-point rule the procedure is as follows:

- Re-arrange the time sequence so that it starts at its absolute minimum or maximum.

- When, following the sequence, a local extreme value is larger or equal to the preceding one, a cycle is counted and registered with a mean value and amplitude. It is discarded from the sequence by connecting the two local extreme values at the level of the largest one, and the analysis of the sequence is continued.

- When, following the sequence, a local extreme value is smaller than the preceding one, the analysis of the sequence is continued.

- If cycles remain when the end of sequence is reached the procedure is repeated.

The result of the procedure is a matrix of cycle numbers characterized by their mean values and amplitudes, which can then be used for life assessment by a damage rule, see Figure 8.3. In complicated cases with many cycles it is practical to make a grouping of experimental data so that peaks and valleys falling into intervals of chosen size are given the same nominal value.

Amplitude > Mean	3,5	4,0	4,5	5,0	5,5	6,0	6,5
1,0		1(BC)					
1,5	1(GH)				1(EF)		1(AD)

Figure 8.3 Resulting matrix of rain-flow count data from the example in Figure 8.2, which can be used for damage analysis.

For applications to LCF it is often necessary to have cycles characterized by strain amplitudes and by stress mean values. This requires simultaneous measurements of both strain and stress. In several areas of application experience shows that the mean values can be omitted. Then a technique called range pair counting can be applied. The results coincide with the rain flow results from the same sequence if the mean values of the matrix are omitted.

8.1.3 Characterization from power spectral density

The theory of stochastic variables offers general methods to treat irregular time sequences. A comprehensive treatment can be found in e g [8.3].

The reason to include some of the results in connection to fatigue damage is that full scale measurements, and numerical simulations with FEM, of complicated structures, as offshore platforms, primarily give results in terms of the power spectral density. This is a representation of the intensity of a parameter (displacement, stress etc) as a function of its frequency, and it is an implicit description of the time sequence.

The power spectral density $S(\omega)$ is defined from the stress sequence $\sigma(t)$ by

$$S(\omega) = \frac{1}{2T}\int_{-T}^{T}\left[\frac{1}{2T}\int_{-T}^{T}\sigma(t)\sigma(t+\tau)\,dt\right]exp(-i\omega\tau)d\tau \qquad (8.1)$$

for a time interval 2T that ideally is infinite. In practise it should be sufficient that the process is stationary and repeated several times within the time T. Further, $\sigma(t)$ is an amplitude value, where the expected (mean) value is subtracted, and has to be retained for fatigue life assessments.

Since also phase information is lost in the transformation, the reversed process to estimate the probability density of peak stresses p(σ), necessary for damage analysis, when the base data is

in the form of $S(\omega)$ has to include assumptions and approximations. Methods employed normally make use of the first four moments M_i of $S(\omega)$, where

$$M_i = \int_0^\infty \omega^i S(\omega) d\omega \tag{8.2}$$

For example, the number of passages of the mean level per time unit is

$$\nu_f = \frac{1}{\pi}\left(\frac{M_2}{M_0}\right)^{1/2} \tag{8.3}$$

$\nu_f/2$ is called the virtual frequency. The number of peaks and valleys per time unit is

$$\nu_m = \frac{1}{\pi}\left(\frac{M_4}{M_2}\right)^{1/2} \tag{8.4}$$

The spectral width in the frequency domain is $k=\nu_f/\nu_m \quad 0<k<1$ (c f Chapter 8.1.1)

Several methods to find $p(\sigma)$ have been suggested. An overview is given in e g [8.4].

When $p(\sigma)$ has been estimated the damage D per time unit can be determined. The procedures in the literature invariably rely on linear accumulation of damage, as described in Chapter 8.2 and D becomes

$$D = \sum \frac{n_i}{N_i} = \nu_m \int_0^\infty \frac{p(\sigma)d\sigma}{N(\sigma)} = \nu_m A^{1/b} \int_0^\infty \sigma^{-1/b}\, p(\sigma) d\sigma \tag{8.5}$$

with the Basquin equation

$$\sigma = AN^b \tag{8.6}$$

representing the S-N-curve for all stresses. As mentioned before it is customary to do this when some of the stresses exceed the endurance limit. It is also necessary to adjust A, and possibly b for the estimated mean value of the time sequence.

A method to find $p(\sigma)$ which is viewed as giving among the best results compared to rain-flow count for experiments performed with a range of load spectra is the one developed by Dirlik. The formula is given as

$$p(\sigma) = \frac{1}{2\sqrt{m_0}}\left\{\frac{D_1}{Q}\exp\left(-\frac{\sigma}{2Q\sqrt{m_0}}\right) + \frac{D_2\sigma}{2R^2\sqrt{m_0}}\exp\left(-\frac{\sigma^2}{8m_0R^2}\right) + \frac{D_3\sigma}{2\sqrt{m_0}}\exp\left(-\frac{\sigma^2}{8m_0}\right)\right\} \tag{8.7}$$

where, with k as above and $x_m = \frac{m_1}{m_0}\sqrt{\frac{m_2}{m_4}}$,

$$D_1 = \frac{2\left(x_m - k^2\right)}{\left(1+k^2\right)} \qquad R = \frac{k - x_m - D_1^2}{1 - k - D_1 - D_1^2}$$

$$D_2 = \frac{1 - k - D_1 - D_1^2}{1 - R} \qquad D_3 = 1 - D_1 - D_2$$

$$Q = 1.25\left(k - D_3 - RD_2\right)/D_1$$

When $p(\sigma)$ is Gaussian, the distribution of maxima, which is of interest in fatigue will be normal for k=0 (an extremely broad banded time sequence, i e there is a high probability for maxima below the mean level). For k=1 (a narrow band sequence) a Rayleigh distribution is obtained, with no maxima below the mean (and no minima above it).

8.1.4 Full scale tests and type spectra

For various branches of industry the loads on structures may be possible to characterize in terms of typical load sequences. Vehicles on roads of different quality, aeroplanes in civil or military flight situations, tankers and oil drilling platforms in different marine environments, as the North Sea, parts of the Pacific etc, are examples where such so called type spectra are used.

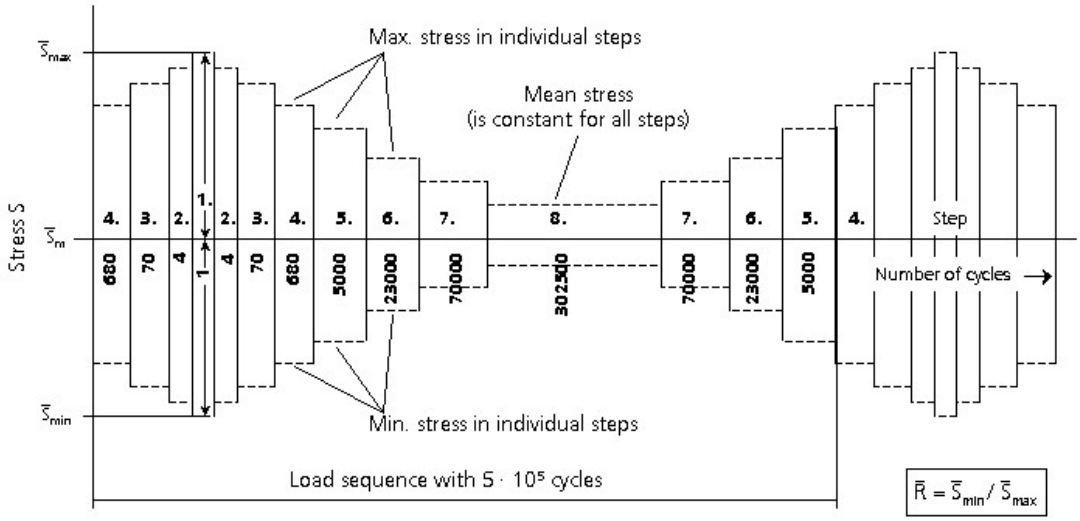


Figure 8.4 LBF block sequence for use in experiments.

In the automotive industry a lot of time sequences is recorded in the field and used for reproduction in full-scale component tests in the laboratory servo-hydraulic test equipment, see Chapter 10. This is expensive but gives realistic conditions and results. As an alternative, type spectra can be used. An example is the so-called LBF-spectrum [8.1], see Figure 8.4. This method can be used when the available test equipment has limitations as regards reproduction of the precise time sequence.

In the marine and airplane sectors there is also a number of type spectra for typical load situations. Examples are the TWIST and FALSTAFF spectra [8.5], [8.6] for transport and military aircrafts respectively. They are given primarily as time sequences for experimental use, but also transformed to load spectra for damage analysis. Both spectra have simplified versions. Often the FALSTAFF spectrum is used also for civil aircrafts. It is considered to give conservative results since it contains load cycles from more violent motions of the aircraft.

In codes for welded structures, as the Swedish BSK [8.7] (see also Chapter 13) other principles are used. If the load can be referred to certain, common service conditions, as lifting devices, construction equipment etc, the maximum stress is reduced with a factor ($\kappa < 1$) and then the damage is calculated from the total number of cycles in a service sequence with a constant stress modified by this reduction factor. If this is not the case but the load spectrum can be estimated, the reduction factor is determined graphically as a best fit to the spectrum, see Figure 8.5. These principles include truncation of cycles with small amplitudes. This and similar methods are connected to standardized ways of estimating life from weld classes, size and surface effects etc, i e it is a part of a design package and should not be used independently.

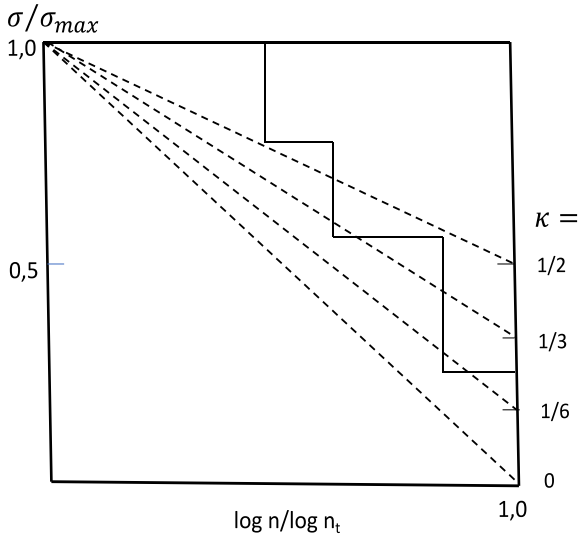


Figure 8.5 Demonstration of determination of reduction factor from a load spectrum. (From BSK, see Chapter 13) The spectrum is a level crossing spectrum, see Figure 8.1, but based on stress ranges only.

8.1.5 Experimental support for design and predictions

Thanks to the development of experimental equipment it is now technically possible to simulate recorded irregular loads in full scale on structures and components, as for the FALSTAFF and other spectra. With less sophisticated equipment it is of interest to perform tests with the cycles arranged in blocks of constant load derived from cycle counting methods, similar to the type loading sequence in Figure 8.4.

As mentioned, the order in which the cycles are arranged is of importance. When experiments are performed with cycles ordered in other fashions than the one in real life, the experimental results must be expected to differ from each other. This is analogous to the cases when life is estimated with cycles counted by different methods, e g the ones in Chapter 8.1.1-8.1.3.

The possibilities for life assessment are depicted in Figure 8.6. Experiments are used on different levels.

Experiments on a structure	With real load spectrum (8.1.4) (L)	With simplified spectrum, e g blocks (8.1.4) (L´)	
Experiments on material			Constant amplitude tests, giving material data for Basquin or Coffin-Manson equations

Predictions based on measure-ments of stress	Exceedance and level crossing count spectra (Ch 8.1.1) and (8.2) (L_1, L_2)	Rain-flow and range pair count spectra (Ch 8.1.2) and (8.2) (L_3, L_4)	Probabilistic spectrum p(□) (Ch 8.1.3) and (8.2) (L_5)

Figure 8.6 Background data and ways for design and assessment of fatigue lives L.

Of course, the direct way is to perform tests in full or component scale with the real spectrum, giving the true life L, possibly even with a measure of uncertainty, see also Chapter 14. If this is not possible, due to limitations of experimental resources, the real spectrum can be reduced to block spectra in different ways. The blocks are applied as sequences of constant amplitude load. The experiments give lives L´, which may differ considerably from L, and very much so depending on the order in which blocks with high and low load are applied.

If the only experimental data available is from constant amplitude material testing, suitably modified according to Chapter 4, predictions can be made with load cycle counting from service conditions according to Chapter 8.1.1-8.1.3 and a damage accumulation rule according to Chapter 8.2. Then life estimates L_1 – L_5 are obtained. This approach is useful e g in early stages of product development.

It is by no means sure that L´-values are better estimates than the L_1 – L_5 ones or that any of them are conservative. It can also be said that the rain-flow-based data L_3 – L_4 generally give the best predictions.

8.2 Assessment of life from a damage accumulation rule

8.2.1 Accumulation of damage. Damage rules

In a general analytical situation it is not possible to follow the evolution of damage, from initiation to fracture. In particular cases, as discussed in Chapter 7, there are models to follow the growth of an existing crack, with retardation and acceleration, although this becomes cumbersome and not very precise for complicated spectra.

In most situations one has to use the limited information of a load spectrum for analysis, as an exceedance or level crossing spectrum, as a range pair spectrum, or as a rain-flow matrix with

information also of mean value stresses. To facilitate the analysis the spectrum is split up in groups of cycles that are nearly identical, see e g Figure 8.3.

A straightforward approach is the Palmgren-Miner rule of linear accumulation of damage. It is named after Palmgren, who suggested it for design of ball bearings already 1924 [8.8], and Miner, who independently suggested it 1945 [8.9]. Even though it can easily be shown to have principal weaknesses it is still used more or less exclusively, e g in most of the computerized and FEM-based expert systems for fatigue design available on the market. This is so because more complex approaches also build on assumptions and in addition lead to the need for more, material dependent parameters, and not always give superior results.

The concept is simple enough. Each cycle, or block of n_i cycles defined as nearly identical, is assumed to produce a damage, consume a part of the life, equal to n_i/N_i, where N_i is the life for such cycles. For a sequence with a number m of sets of such identical cycles, i e a total of n_0 cycles

$$n_0 = \sum_{i=1}^{m} n_i \tag{8.8}$$

the damage D will then be

$$D = \sum_{i=1}^{m} \frac{n_i}{N_i} \tag{8.9}$$

irrespective of the order of the cycles, i e with no regard to the "memory" of the material. When the sequence is so long that D reaches the value of 1 the life is assumed to be spent and fracture occurs. In other words, the life is spent for a sequence with the length n_0/D and the same proportion of cycle types n_i.

It seems obvious then that the same load sequence represented with different counting methods will give different results. It is motivated theoretically, and found in comparisons, that the rain-flow method generally gives the best agreement with experiments.

Some of the reasons why the Palmgren-Miner rule must give somewhat erroneous results, sometimes conservative and sometimes non-conservative, are the following.

- The total life represented by the S-N- or Coffin-Manson curve is composed of an initiation phase and a phase of crack growth. It must be of significance how the order of large and the small cycles influence damage in the two phases.

As an example, take the case of just two levels of stress, first one σ_1, with a life of N_1, and then one σ_2 with life N_2. If the first level is applied for N_1' cycles, where N_1' is the number corresponding to the transfer from initiation to crack growth (suitably defined) then the number of cycles to fracture according to the Palmgren-Miner rule is

$n_2 = N_2\left(1 - \frac{N_1'}{N_1}\right) = N_2 - N_2'$, where N_2' is the number of cycles to the transition at the level with stresses σ_2.

But this means that

$$\frac{N_1'}{N_2'} = \frac{N_1}{N_2}$$

i e that there is a proportionality between the lives and the number of cycles to the transition at all levels of stress. This is known not to be the case. For higher stress levels the crack growth phase takes up a larger part of the life than at lower levels.

- In the crack growth phase it is known that there are retardation and acceleration effects for high-low and low-high transitions of cycle levels, meaning that the order of cycles of different amplitudes is of significance.

- The mechanisms of damage are different in the plastic, LCF, and in the elastic range. For sequences mixing cycles from both ranges there is a complicated interaction and a simple accumulation of damage is not realistic.

In order to adjust for the deficiencies of the Palmgren-Miner rule and still be able to make use of its simplicity one can learn from experience in different areas of application. In those the load spectra often have common characteristics regarding e g ranges of amplitudes and regularity. Hence they result in fracture for characteristic D-values different from1, for example for sea wave spectra, loads on wind power equipment etc. Typically, such so called Miner-sums range from 0.2 - 5.

A further note is that the fatigue limit loses its significance for variable loading as soon as parts of the load sequence exceed it. Since damage leading to crack formation can occur at the stress levels above the fatigue limit these cracks can grow at stresses well below the fatigue limit for initiation once they have reached a sufficient size. Then, in cases where sequences of variable amplitude occurs, as typified load spectra in standards, the S-N curve applied is continuing with a slope to lower stress levels for lives longer than those normally connected with the fatigue limit, see Chapter 13.

As mentioned in Chapter 8.1 it is customary to truncate small cycles, if they are not a dominating part of the load spectrum or exceed the fatigue limit. This is motivated by the fact that the S-N curve has a small slope. For example, for the slope b=0,15 in (8.10) a decrease in stress level of 30 per cent means an increase of life $(1/0,7)^{1/0,15}=9,88$, i e nearly a factor 10.

The example points to a general principle. If the S-N curve can be approximated by a straight line in a diagram with logarithmic scales with slope -b, the change of the levels in any load spectrum with a factor r will give a change in life with a factor

$$(1/r)^{1/b} \qquad (8.10)$$

The stochastic properties have to be dealt with also for variable loading. Of course there are uncertainties both in the load spectrum and in the strength of the component. In standards, as the example discussed in Chapter 13, this is managed by a combined set of factors and risk levels for different applications. In analytical situations one normally starts with a given load sequence and a S-N curve including surface, size effects etc for a 50 % risk for failure. With an additional assessment of the standard deviation of the S-N curve, see Chapter 4.2.1 and 10, a modified S-N curve can be chosen for the desired risk level. A general method to take probabilistic aspects in both load and strength into account is described in Chapter 14.2.

There are other methods for life assessment, both experimentally and analytically, in branches of industry depending on materials used, service loads and requirements on safety levels. They are often the results of experience and availability of technical resources. The descriptions in this chapter are therefore just examples of some of the most pertinent principles. The most significant omission is the one of the use of more complicated non-linear life assessment methods. The reason is, as mentioned, that they tend to be limited to special cases and that they

require extensive evaluation of material properties which are hard to come by for generalized situations.

8.2.2 Example of application of the Palmgren-Miner damage rule in stress based fatigue design

The cycle counting methods result is a table of m cycle types, C_i, each characterized by amplitude σ_{ai}, mean value σ_{mi}, and number of repetitions, n_i. For each such cycle type fatigue data for the respective amplitude and mean value give the life N_i for constant amplitude loading.

It is now assumed that the life relationship, with corrections for surface and dimension effects, can be represented by

$$\sigma_{ar} = AN^{-b} \ , \tag{8.11}$$

and the dependence of mean stresses by the model relationship

$$\sigma_{ar} = \frac{\sigma_a}{1-\frac{\sigma_m}{\sigma_f'}} \tag{8.12}$$

(see Chapter 4). Then one can write the life N_i for any of the C_i:s as

$$N_i = \left[\frac{\sigma_{ai}}{A\left(1-\left(\frac{\sigma_{mi}}{\sigma_f'}\right)\right)} \right]^{-1/b} \tag{8.13}$$

Then the damage from the n_i cycles of that C_i –type in the sequence is

$$D_i = \frac{n_i}{N_i} \tag{8.14}$$

and the total damage from the sequence

$$D = \sum_i \frac{n_i}{N_i} \tag{8.15}$$

The life assessment of the structure in terms of number of sequences is 1/D.

Of course, any analytical or graphical representation of the S-N-curve and mean value dependence can be used, e g to include also notch effects in various materials.

It is noted, see above, that for Miner sums M known to be significantly different from 1 for the load sequences to be analyzed, the life will instead be M/D.

8.2.3 Accumulation of damage in fracture mechanics based fatigue design

Chapter 7.3 gives a description of how the mean value of a cycle, expressed as an R-value, and a change of amplitude affect increments of growth of a crack from a fracture mechanics point of view, and in Chapter 7.3.6 some special methods to treat changes of amplitude are described.

When a load spectrum is given, methods are often used that are analogous to the principle of linear damage accumulation.

Those are the ones that are used in standards and norms and which are outlined in Chapter 13. Those methods normally give conservative results, and at least formally a risk level may be chosen. Even they are, however, depending on the judgement of the user, e g in the choice of the load factor for the actual sequence.

For a general analysis one can use a straightforward application of the linear damage accumulation rule for crack growth, which may seem reasonable for an irregular mix of loads of the same magnitude.

The load sequence is defined as cycles by the rain-flow counting method, or some of the other methods generally in use. This yields cycles, i, with stress spans $\Delta\sigma_i$ and mean values giving R_i. Taking, for F in

$$da/dN = F(\Delta K, R)$$

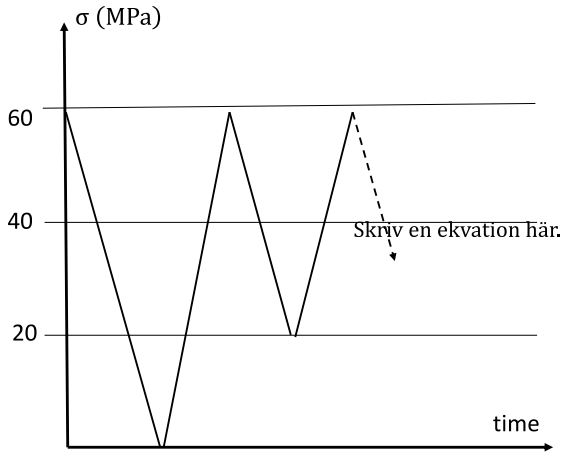
the Walker expression according to Chapter 7.3.6 gives

$$(da/dN)_i = C_0\left(\frac{\Delta K_i}{(1-R_i)^{1-\gamma}}\right)^n = \frac{C_0}{(1-R_i)^{n(1-\gamma)}}(\Delta K_i)^n \qquad (8.16)$$

for crack growth increments. This procedure does not take sequence effects into account, and is similar to the Palmgren-Miner assumption of linear damage accumulation. Still it retains some consideration of the influence of crack closure. More advanced methods to take sequence effects into account, as the FASTRAN one, is used in e g the aircraft industry.

Example

A component of Al 7075 – T6 (σ_0 = 620 MPa) is loaded so that the maximum stress is at an edge with possible cracks. The stress $\sigma(t)$ is shown in the Figure. Assess the number of sequences N that can be sustained if cracks longer than 4 mm are detected and if the maximum stress intensity factor must be smaller than half the fracture toughness K_{Ic} = 25 MPam$^{1/2}$. Use the Walker formula (8.16) to adjust for R-values other than 0.



There is one cycle with $\Delta\sigma$ = 60 MPa and R = 0, and one cycle with $2\Delta\sigma/3$ = 40 MPa and

R = 1/3. Material data from Table 7.1 are C=11,2 10^{-9} and n=3,9. γ is not given but is taken as 0,64. Then, from expanding (8.16),

$$da/dN = C\Delta\sigma^n\left(1{,}12\sqrt{\pi a}\right)^n\left[1+\frac{(2/3)^n}{(1-1/3)^{(1-\gamma)n}}\right] = C[1+0{,}364]\Delta\sigma^n\left(1{,}12\sqrt{\pi a}\right)^n$$

Now, the initial crack length a_0 = 0,004 m gives an initial $\Delta K_{I,0}$ = 1,12 60 $(\pi\ 0{,}004)^{1/2}$ = 7,53 MPa $m^{1/2}$. Further, the final crack length a_f is obtained from

$$a_f = \frac{1}{\pi}\left(\frac{12{,}5}{1{,}12\cdot 60}\right)^2 = 0{,}011\ m$$

Then the allowable cycle number N is obtained from (7.26)

$$N = \frac{a_0\left(\frac{n}{2}-1\right)^{-1}}{1{,}364C\left(\Delta K_{I,0}\right)^n}\left(1-\left(a_0/a_f\right)^{(n/2-1)}\right) = \frac{0{,}95\cdot 0{,}004}{1{,}364\cdot 11{,}2\cdot 10^{-12}\cdot 7{,}53^{3{,}9}}\left(1-\left(\frac{0{,}004}{0{,}011}\right)^{0{,}95}\right)$$
$$= 9{,}5\cdot 10^4$$

From the example it is seen that the extra cycle means a reduction of the life with little less than 30 per cent (the second term in the bracket), if no retardation effects are considered. The situation should be different if there are blocks of, say, ten small and ten big cycles instead. It might also be safer to require that the life should be half of that giving the maximum stress intensity factor as K_{Ic}.

The procedure according to (8.16) is rather cumbersome to perform cycle by cycle in more complicated situations than the example. One can then use the observation that the crack does not grow much during one repetition of the load sequence, as compared to the total life length. Then one can put a_i=a, and

$$\Delta K_i = \Delta\sigma_i f(a)\sqrt{\pi a}\ ,$$

and the expression for growth during one sequence becomes

$$\sum_{i=1}^{N}\left(\frac{da}{dN}\right)_i = C_0\sum_{i=1}^{N}\left(\frac{\Delta\sigma_i f(a_i)\sqrt{\pi a_i}}{(1-R_i)^{(1-\gamma)}}\right)^n \approx C_0\left(f(a)\sqrt{\pi a}\right)^n\sum_{i=1}^{N}\left(\frac{\Delta\sigma_i}{(1-R_i)^{(1-\gamma)}}\right)^n$$

(8.17)

The sum shall be equal to what would be obtained by N cycles with a constant stress span $\Delta\bar{\sigma}$ and R=0 which gives

$$\Delta\bar{\sigma} = \frac{\left[\sum_{i=1}^{N}\left(\frac{\Delta\sigma_i}{(1-R_i)^{(1-\gamma)}}\right)^n\right]^{1/n}}{N}$$

This stress span can then be used in ΔK for the complete analysis of crack growth.

It is mentioned here that one method mentioned in the literature for variable loading is to use the root mean square of the stress intensity factors to simplify the calculations.

8.2.4 Accumulation of damage in strain-based design.

This is treated in Chapter 6.3 and 6.6 since it comes naturally from the origin of the rain-flow count method in non-linear stress-strain loops. The use of linear accumulation of damage is tacitly assumed also in this area of application.

References

[8.1] Fraunhofer LBF, www.lbf.fraunhofer.de

[8.2] Matsuishi, M. & Endo, T. (1968) Fatigue of metals subjected to varying stress, Japan Soc. Mech. Engineering.

[8.3] Newland D. E. "An Introduction to Random Vibration … ", Longman, Burnt Hill, Harlow (1984) (Newer editions available at Amazon.)

[8.4] Matjaˇz Mrˇsnik, Janko Slaviˇc and Miha Bolteˇzar Frequency-domain methods for a vibration-fatigue-life estimation - Application to real data. International Journal of Fatigue, Vol. 47, p. 8-17, 2013.

[8.5] de Jonge, J. B., Schuetz,D., Lowak, H., and Schijve, J., "A Standardized Load Sequence for Flight Simulation Tests on Transport Aircraft Wing Structures", LBF-Bericht TB-106, Amsterdam, 1973.

[8.6] G.M. Van Dijk and J.B. de Jonge, "Introduction to a fighter aircraft loading standard for fatigue evaluation- FALSTAFF", Proceedings of eigth ICAF Symposium, International Committee on Aeronautical Fatigue, Lausanne, Switzerland, 1975.

[8.7] Bestämmelser för stålkonstruktioner, BSK 07, (2007), www.boverket.se.

[8.8] A.G. Palmgren (1924): Die Lebensdauer von Kugellagern. Zeitschrift des Vereines Deutscher Ingenieure (VDI Zeitschrift), ISSN 0341-7258, Vol 68, No 14, April 1924, pp 339-341.

[8.9] Miner, M. A., 1945, "Cumulative Damage in Fatigue," J. Applied Mechanics, 12, A159-A164.

9 Fatigue under multiaxial loading

Multiaxial loading is frequently occurring in mechanical design. The principal problem is, as an extension of the reasoning in Chapter 3, to get a physical understanding of how damage is accumulated locally under stresses and strains that change irregularly during a time sequence. Generalizations of the models for uni-axial analysis are then thought to be useful.

Fatigue initiation for uni-axial reversed loading is often due to the formation of persistent glide bands and extrusions/intrusions in planes with imperfections in a 45 degree plane relative to the uniaxial stress direction, I e maximum shear stress. Further, a superposed tensile stress on the plane means that the amplitude required to initiate fatigue decreases, i e it seems as if the formation of glide bands and microcracks is facilitated by tension across the shear plane.

These essential features of the uni-axial behaviour are assumed to be retained when a generalisation is made to fatigue initiation under multi-axial conditions. Changes in deformation caused by shear strains are taken as the main drivers for start and accumulation of local fatigue damage, and a tensile hydrostatic state of stress is a facilitating factor. Most models for fatigue initiation from multiaxial load sequences of varying complexity have been formulated with this as a starting point.

As in the uni-axial case elasticity prevails on the macro-scale near the fatigue limit, although the fatigue mechanisms are activated as non-reversible displacements in grains on the micro-scale. The models for multiaxial fatigue also presume isotropic material although highly stressed regions of a structure are anisotropic in many practical situations, due to forming, heat treatment etc.

One should distinguish between multi-axial loads and the multi-axiality of stresses in the highly stressed regions. Even if the external loads are multiaxial it may well happen that the stresses in points of interest are nearly uni-axial. Conversely, a single load on a complicated geometry may result in a, proportional, multi-axial stress state at the most severely stressed points. Very often fatigue initiation occurs at surfaces, and then the multi-axial stress conditions are two-dimensional, which simplifies the treatment.

Models for fatigue initiation under multi-axial stresses must, of course, comply with uni-axial criteria for uni-axial stress states. This fact can be used to calibrate the parameters in the multi-axial models. Modifications can be made for size, surface treatment etc, similarly to what is done in Chapters 4 and 5, and a desired risk level for fatigue can be chosen.

Models for long lives and particularly for fatigue initiation are treated in Chapters 9.1 and 9.2. First the special but important case of proportional loading giving fixed directions of the principal stresses is discussed in Chapter 9.1 as an introduction showing the principles but without computational complexities. The more general case, where the stress state varies arbitrarily during a time sequence of stress states is discussed in Chapter 9.2.

When the stress states during some part of a load sequence exceed those corresponding to the fatigue limit one must account in some way for equivalents to cycles in multiaxial states. For higher stress levels, corresponding to LCF conditions, multi-axial strain ranges have to be considered. These situations are treated in Chapters 9.3 and 9.4. Finally, crack growth under multiaxial stress states are briefly discussed in Chapter 9.5.

It is underlined that the models presented fit experiments with varying success, and that there are many investigations exploring more elaborated models. *Often it seems sufficient to assume*

that the largest principal stresses dominate the formation of fatigue damage, and they can form a basis for an assessment that is comparable in predicting capacity with a complete analysis using one of the models.

9.1 Fatigue initiation with proportional stress states

In the simplest case of multiaxiality there are only proportional, in-phase stresses with constant amplitudes, $\sigma_{ij,a}$, present. They may be caused by a single periodic load that results in a multi-axial stress state in the loaded structure.

It is desirable to find a scalar quantity, an equivalent stress, σ_{eq}, formed from $\sigma_{ij,a}$ which expresses the influence of shear stress variations activating fatigue mechanisms and which can be compared with the fatigue limit σ_e. Such a quantity is the von Mises equivalent stress that is a scalar measure of the *deformation* of a stressed element. The von Mises expression is, with tensor notation,

$$\sigma_{eq} = \sqrt{\frac{3}{2} s_{ij} s_{ij}} \tag{9.1}$$

where the deviatoric stresses are

$$s_{ij} = \sigma_{ij} - \frac{1}{3}\delta_{ij}\sigma_{kk} \ , \tag{9.2}$$

with the mean, or hydrostatic stress $\frac{1}{3}\sigma_{kk}$.

For a stress state with just amplitude stresses and proportional loading the principal stresses may be used, and then (9.1) can be expressed as

$$\sigma_e = \sigma_{eq} = \frac{1}{\sqrt{2}}\left[\left(\sigma_{1,a} - \sigma_{2,a}\right)^2 + \left(\sigma_{2,a} - \sigma_{3,a}\right)^2 + \left(\sigma_{3,a} - \sigma_{1,a}\right)^2\right]^{1/2} \tag{9.3}$$

In most cases fatigue initiation occurs at a surface and then the stress state is bi-axial, and the formula reduces to

$$\sigma_e = \sigma_{eq} = \left(\sigma_{1,a}^2 - \sigma_{1,a}\sigma_{2,a} + \sigma_{2,a}^2\right)^{1/2} \tag{9.4}$$

or in arbitrary Cartesian co-ordinates

$$\sigma_e = \sigma_{eq} = \left(\sigma_{x,a}^2 - \sigma_{x,a}\sigma_{y,a} + \sigma_{y,a}^2 + 3\tau_{xy,a}^2\right)^{1/2} \tag{9.5}$$

From this form it appears that for a case of pure shear, e g reversed torsion of a bar this would give a fatigue limit of

$$\tau_e = \frac{\sigma_e}{\sqrt{3}} = 0{,}58\sigma_e \tag{9.6}$$

This agrees well with experimental data for many ductile materials, indicating that the von Mises expression gives a reasonable equivalent stress. Yet, there are also materials, where

$$\tau_e \approx 0{,}5\sigma_e \tag{9.7}$$

a value that agrees better with the expression for maximum shear stress as a criterion for fatigue initiation.

Hence, another and perhaps more natural choice of equivalent stress is the maximum shear stress amplitude

$$\sigma_{eq} = \tau_{a,max} = \left|\sigma_{1,a} - \sigma_{3,a}\right|/2 \tag{9.8}$$

which for a bi-xial state in Cartesian coordinates is expressed as

$$\tau_{a,max} = 0{,}5\left[\left(\sigma_{x,a} - \sigma_{y,a}\right)^2 + 4\tau_{xy,a}^2\right]^{1/2} \tag{9.9}$$

Both the von Mises and the maximum shear stress criteria will be generalized for arbitrary multi-axial loading in Chapter 9.2. The advantage of the von Mises expression is that it is a scalar entity, in a way integrating the three-dimensional shearing behaviour, whereas the maximum shear stress criterion, which is more to the point, physically speaking, requires defined shearing planes.

Now, for proportional loading, stress states including mean stresses should be introduced as the next step.

Analogously to the von Mises criterion Sines [9.1] suggested for initiation of fatigue under proportional, in-phase loading

$$\sigma_e = \frac{1}{\sqrt{2}}[(\sigma_{1a} - \sigma_{2a})^2 + (\sigma_{2a} - \sigma_{3a})^2 + (\sigma_{3a} - \sigma_{1a})^2]^{1/2} + c_s(\sigma_{1m} + \sigma_{2m} + \sigma_{3m}), \tag{9.10}$$

where the hydrostatic mean stress, which is an invariant, has been taken as a measure of the normal stresses facilitating the formation of damage initiation. This gives an expression including mean stresses comparable to uniaxial cases.

The factor c_S, expressing the mean value influence, is obtained by insertion of data for the uni-axial case. If the linear mean value dependence from Chapter 4

$$\frac{\sigma_{1a}}{\sigma_e} + \frac{\sigma_{1m}}{\sigma_f'} = 1 \tag{9.11}$$

is used one obtains that $c_S = \sigma_e \big/ \sigma_f'$. For moderate mean values the parameter can also be determined from the uniaxial data for pulsating load,

$$\sigma_{ep} = \sigma_{1a} = \sigma_{1m}$$

giving

$$c_s = \sigma_e/\sigma_{ep} - 1 \tag{9.12}$$

The Sines expression is sometimes used in situations with non-proportional loading when reasonable values for amplitude and mean value stresses can be estimated.

The shear stress criterion is conceptually more clear-cut for proportional loading, since the maximum change in shear stress occurs in well-defined 45^0 shear planes according to (9.8). Then a criterion including mean values could analogously to (9.10) be

$$\sigma_{eq} = \frac{1}{2}|\sigma_{1a} - \sigma_{3a}| + c_{shear}(\sigma_{1m} + \sigma_{2m} + \sigma_{3m}) \qquad (9.13)$$

with calibration of the parameters from uni-axial data.

Now, the models for general cases discussed in Chapter 9.2 use the maximum hydrostatic stress instead of the mean values, probably, as will be seen, because of the difficulty to determine mean or mid-values from a general multi-axial stress sequence. So, instead of (9.13) the analogous expression would be

$$\sigma_{eq} = \frac{1}{2}|\sigma_{1a} - \sigma_{3a}| + \bar{c}_{shear}\big((\sigma_{1m} + \sigma_{1a}) + (\sigma_{2m} + \sigma_{2a}) + (\sigma_{3m} + \sigma_{3a})\big) \qquad (9.14)$$

giving, with the use of (9.11), $\sigma_{eq} = \frac{\sigma_e}{2}\left(1 + \frac{1}{\frac{\sigma_f'}{\sigma_e} - 1}\right)$ and $\bar{c}_{shear} = \frac{1}{2\left(\frac{\sigma_f'}{\sigma_e} - 1\right)}$ (9.15)

Obviously, there is some arbitrariness in the choice of criteria. The examples above are partly given as a background to the corresponding criteria for general multi-axiality where the concept of a mean or mid-value is less obvious to define in order to find a suitable measure for shearing and tension stresses contributing to fatigue initiation.

It is worth mentioning that only the amplitudes of stresses are used in standards for welded structures, due to the presumed prestressing of such structures. For multi-axiality, mostly bi-axial, equivalent stresses are formed in analogy with (9.9), i e as square root expressions.

9.2 Fatigue initiation with general periodic multi-axial loading

The principles of the von Mises (octahedral) stress, and of the maximum shear stress are used also for cases of general periodic loading, where the stresses $\sigma_{ij}(t)$ vary arbitrarily during a repeated stress sequence in the time interval $0 \le t \le T$. Measures of the hydrostatic stress are also retained to include the facilitating effect of normal stresses.

Octahedral stresses.

The formulation of the octahedral stress criterion, generalizing the one of Sines, is due to Crossland [9.2], suggesting that fatigue initiation occurs if

$$\sigma_{eC} \le \sqrt{\frac{3}{2}s_{ij,a}(t)s_{ij,a}(t)} + c_C\sigma_{h,max}(t) \qquad (9.16)$$

for any time t during some complex multi-axial stress cycling sequence.

Here $s_{ij,a}(t)$ means the components of the deviatoric stress tensor reduced by a mid value tensor $s_{ij,mid}$, i e

$$s_{ij,a}(t) = s_{ij}(t) - s_{ij,mid} \qquad (9.17)$$

so that some "amplitude" of the stress tensor is a basis for the maximum change of the von Mises stress during a cyclic period of stressing. Further, $\sigma_{h,max}(t)$ is the maximum of the hydrostatic stress during the time sequence

$$\sigma_{h,\max}(t) = max\left(\frac{1}{3}\sigma_{ii}(t)\right).$$

For proportional loading, using principal stresses and assuming the mean values for the mid value stress tensor and hence the amplitudes for the maximum values of $s_{ij,a}(t)$, it is seen that (9.16) is equal to the Sines expression (9.12) except that the maximum hydrostatic stress is used instead of the hydrostatic stress of the mean values.

For a uni-axial case with an amplitude and a mean stress, $\sigma_m \pm \sigma_a$, the expression (9.16) collapses into

$$\sigma_a + c_C\frac{1}{3}(\sigma_a + \sigma_m) \geq \sigma_{eC} \qquad (9.18)$$

Comparison with (9.11) gives the parameter values

$$c_C = 3\left[\frac{\sigma_e/\sigma_f'}{1-\sigma_e/\sigma_f'}\right] \qquad (9.18a)$$

and

$$\sigma_{eC} = \frac{\sigma_e}{1-\sigma_e/\sigma_f'} \qquad (9.18b)$$

The stresses in (9.17) could be regarded as a six-dimensional vector, due to symmetry of the tensor. The square root in (9.16) formed by the stresses is at any time the length of the shear stress vector in the octahedral planes, which as such rotate with t. The, quasi-physical, intention is to find half the maximum difference in the length of this shear stress vector, in order to find a measure comparable to the one formed by the amplitudes used for cases of proportional loading in Chapter 9.1. If, for example, there are large tensile stresses involved there will also be too large octahedral stresses, which do not mirror the changes sought for.

The reasoning is a bit hard to conceive since it is about finding the radius of the smallest super-sphere in a six-dimensional space that circumscribes the loop performed by the deviatoric stresses during the load sequence. So, the problem is to find the stresses $s_{ij,mid}$ minimizing

$$max\left(\sqrt{\frac{3}{2}s_{ij}(t)s_{ij}(t)}\right),\ 0 \leq t \leq T \qquad (9.19)$$

The determination of $s_{ij,mid}$ is not uncomplicated and requires a numerical algorithm.

The use of the hydrostatic stress is not ideal, since it is really the stress normal to the planes of large shear stress changes that are significant, but the hydrostatic stress may be seen as a useful compromise in complicated cases, partly because it is an invariant. The choice of the maximum hydrostatic stress has appeared to give reasonable results in comparisons with experiments.

Shear stresses.

The maximum shear stress amplitude at a critical plane has been suggested as an alternative criterion by Dang Van [9.3], in analogy with the shear stress criterion in Chapter 9.1. It states that fatigue is initiated if for some t during a time sequence of loading

$$\tau_a(t) + D\sigma_h(t) \geq \sigma_{DV}, \tag{9.20}$$

where

$$\tau_a(t) = \frac{1}{2}\Big(s_{1,a}(t) - s_{3,a}(t)\Big) \tag{9.21}$$

and where D and σ_{DV} are material constants, and $s_{1,a}(t)$ and $s_{3,a}(t)$ are the largest and smallest principal values of $s_{ij,a}(t)$. The mid value tensor $s_{ij,mid}$ can be estimated in the same fashion as before.

The Dang Van shear stress for one specific plane is illustrated in Figure 9.2. Here the vector $\tau_a(t)$ in the plane is shown during a multi-axial stress sequence. The origin is not the zero value of the shear strain vector $\tau_t(t)$, but as seen from (9.21) a value $\tau_\alpha(t)$ corresponding to subtraction of the tensor $s_{ij,mid}$ discussed in more general terms above. It has the property that the curve described by the tip of $\tau_a(t)$ is inscribed by a circle with the $\tau_a(t)$-origin as its centre.

A way to use (9.20) is to investigate a number of planes as the one in Figure 9.1 for a time sequence and see for which plane the criterion is fulfilled.

The critical plane approach assumes that initiation of fatigue occurs in one critically stressed plane during a repeated multiaxial load sequence. This may be questioned from a physical standpoint since there will be a complicated interplay between shear in different directions accumulating and cross-linking dislocations. Then a more integrating parameter as the von Mises stress in the Crossland criterion might be a more satisfactory description. Yet the Dang Van model is often possible to correlate well with experiments and is commonly used.

Since the Dang Van model should of course cover also uniaxial cases the parameters D and σ_{DV} can be found by insertion of endurance limit data for uniaxial cases. Such data may be found for e g pure shear, τ_e, uniaxial alternating load, σ_e, and pulsating load, σ_{ep}. The shear

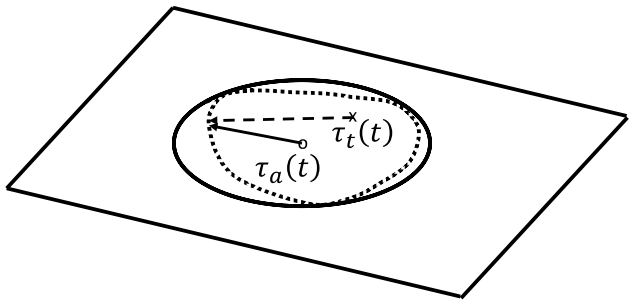


Figure 9.1 Illustration of the Dang Van shear stress in one specific plane for a load sequence.

value τ_e is actually more related to the bending data, σ_{eb}, (see Chapter 4 and Table 4.3). In a diagram with σ_h and τ_a as variables the corresponding pairs of maximum values, coordinates, for the three uniaxial cases will be $(0, \tau_e), (\sigma_e/3, \sigma_e/2)$ *and* $\left(2\sigma_{ep}/3, \sigma_{ep}/2\right)$, respectively. With, for instance, τ_e and σ_e chosen as the uni-axial data,

$$\sigma_{Dv} = \tau_e \tag{9.22.a}$$

$$D = 3\left(\frac{\tau_e}{\sigma_e} - \frac{1}{2}\right) \tag{9.22.b}$$

The situation is shown in Figure 9.2 with data for a real material, where both axial and bending data are included for comparison. It is seen that the uniaxial data do not fit very well to the shear stress τ_e, so this line is drawn more or less parallel to the one for bending data. This elucidates that the purely three-dimensional analysis above may be considered with some care.

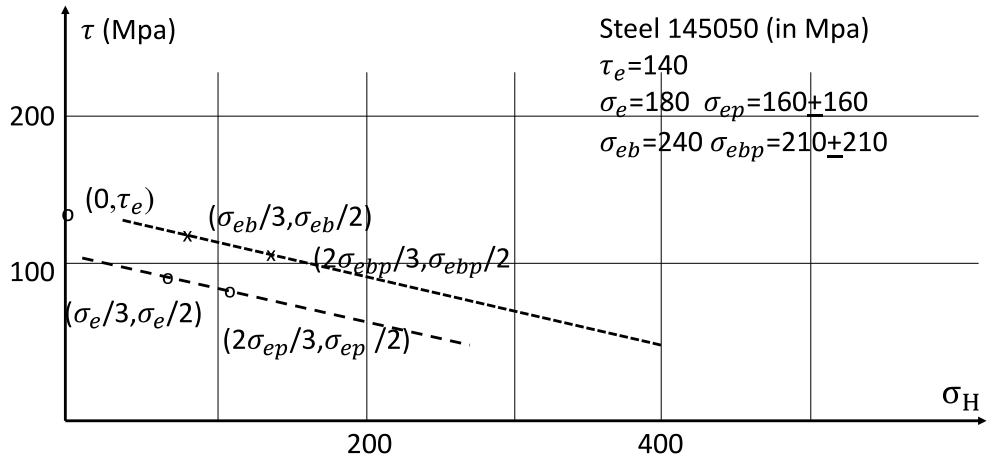


Figure 9.2 Diagram for evaluation of the risk for fatigue initiation according to the Dang Van criterion. The limiting line $\tau_a + D\sigma_h = \sigma_{DV}$ (9.20) is assessed from uniaxial data. If for any time t a stress state exceeds the limiting line fatigue is assumed to be initiated.

It is unlikely that the three coordinates will form a straight line as presumed by the model, and which is actually the case for the material chosen for Figure 9.2. This is a similar conflict that appears for the Goodman and Gerber models in the uni-axial case. Still, the model is useful and correlates well with multi-axial load experiments. Both the Crossland and the Dang Van criteria state that a linear combination of a shear amplitude expression and a hydrostatic stress shall exceed a value, a "fatigue limit", which is a material constant. This is in analogy with the uniaxial case and the Goodman equation for the relationship between amplitude and mean stress.

Example

The seemingly simple case with the surface of a bar loaded in tension and torsion is considered as an example. It gives a stress state in the surface with a uniaxial stress $\sigma_x(t)$ and a shear stress $\tau_{xy}(t)$. The time functions are chosen as

$$\sigma_x(t) = A\sin\omega t \tag{9.23}$$

$$\tau_{xy} = \frac{A}{2}\sin 2\omega t \tag{9.24}$$

Due to symmetry the mid value corresponding to $s_{ij,mid}$ is zero. The task is to find planes giving shear and hydrostatic stresses maximizing (9.20). This is done by investigating the development of $\tau_a(t)$ and $\sigma_h(t)$ with time. They will be

$$\sigma_h(t) = \frac{A}{3}\sin\omega t \qquad (9.25)$$

$$\tau_a(t) = \sqrt{\left(\frac{\sigma_x(t)}{2}\right)^2 + \left(\tau_{xy}(t)\right)^2} = \frac{A}{2}\sqrt{(\sin\omega t)^2 + (\sin 2\omega t)^2} = \frac{A}{2}\sin\omega t\sqrt{5 - 4(\sin\omega t)^2} \qquad (9.26)$$

The result is shown in Figure 9.3, with the Dang Van criterion sketched as the dashed line. It is seen that the point closest to the line occurs for ωt=0,9-1,0 radians, corresponding to 54 degrees and sin 0,95=0.81. The angle ϕ for the plane where this occurs is obtained from insertion in (9.23) and (9.24) and use of Mohrs circle, see Figure 9.4. The result is approximately 20 degrees. For this plane a trajectory for the development of the shear stress could be drawn in Figure 9.2. It would be just a section of a straight line describing the shear stress going up and down. (There is another maximum for ωt=2,2-2,3 radians, giving a plane with nearly the same risk for damage at an approximate angle ϕ=80 degrees.)

The example demonstrates the complexity of a detailed analysis even for apparently simple cases. The analysis would be even more cumbersome with phase differences and more uneven frequency relationships.

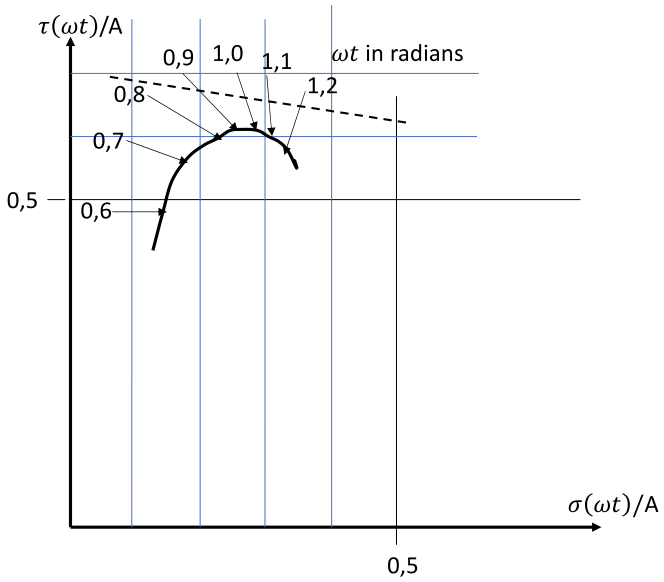


Figure 9.3 Development of the relationship between the hydrostatic and shear stresses in the example, for the Dang Van criterion.

As seen from Figure 9.3 the critical shear stress amplitude is close to 0,63 A, which should then be compared with a material value slightly smaller than τ_e. With the classical formula from e g VDI-norms, i e the first member of (9.26), one would obtain

$$\tau_a = \sqrt{\left(\frac{\sigma_{max}}{2}\right)^2 + \tau_{max}^2} = \sqrt{\left(\frac{A}{2}\right)^2 + \left(\frac{A}{2}\right)^2} = 0{,}71A,$$

which is not too much differing from 0,63A, and conservative!

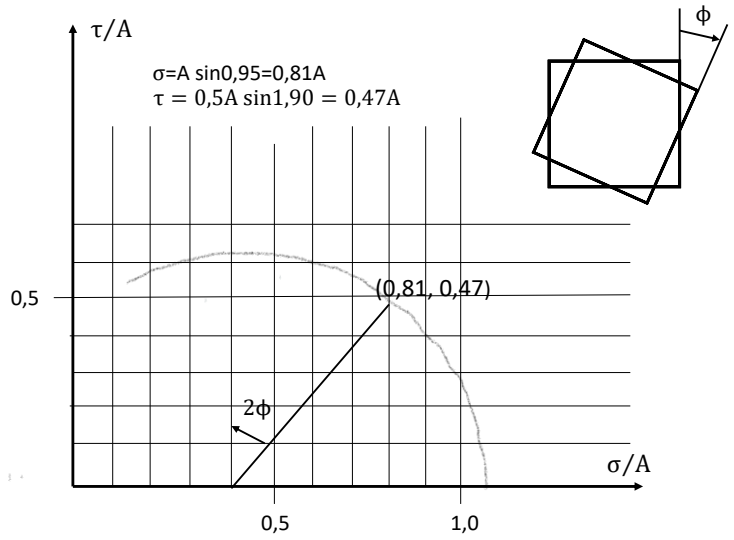


Figure 9.4 Mohr's circle to find the critical plane. The centre of the circle is $\sigma_x/2$.

9.3 Fatigue life assessment with multi-axial loading

If it s desired to find a limited life and not only the fatigue limit for a certain multi-axial load it is, formally, possible to calibrate the parameters in as well the Sines as the Crossland and Dang Van models to any life level by the Basquin equation for uni-axial loads.

Just as in the uni-axial case such extrapolations should be limited to long lives, where the initiation mechanisms are retained. For higher loads and shorter lives the damage is caused by plasticity on the macro-scale, and the assumptions of damage initiation in critical planes is no longer evident.

Then the strain levels should, again, be the loading parameters used for life assessments. This is treated in Chapter 9.4.

In practise, the problem with life assessments for long lives should not be a serious one. For most cases one would like to have a safety margin for initiation, and not laborate with long but limited lives and small stress differences. As demonstrated in Chapter 4 the change in stress between the endurance limit and a life of , say, 10^5 is small, of the same size as the uncertainties in the load and fatigue strength assessments.

9.4 Low cycle fatigue under multiaxial loading

In cases with proportional strain cycling, with a three-dimensional state of strain amplitudes instead of the uniaxial case treated in Chapter 6 it is natural to replace the uniaxial strain amplitude $\varepsilon_a = \varepsilon_{e,a} + \varepsilon_{p,a}$ in the Coffin-Manson model

$$\varepsilon_a = \frac{\sigma_f'}{E}(2N_f)^b + \varepsilon_f'(2N_f)^c \tag{9.27}$$

with a scalar strain measure. As before, the proposed choices are the equivalent strain amplitude

$$\varepsilon_{a,eq} = \frac{1}{\sqrt{2}(1+\nu_{eff})}\left[\left(\varepsilon_{a,1}-\varepsilon_{a,2}\right)^2+\left(\varepsilon_{a,2}-\varepsilon_{a,3}\right)^2+\left(\varepsilon_{a,3}-\varepsilon_{a,1}\right)^2\right]^{1/2} \quad (9.28)$$

and the maximum shear amplitude

$$\gamma_{a,max} = \left|\varepsilon_{a,1}-\varepsilon_{a,3}\right|. \quad (9.29)$$

The ν_{eff} is a weighted value of the Poisson´s ratio for strain states containing both elastic and plastic strains

$$\nu_{eff} = \frac{1}{2}-\left(\frac{1}{2}-\nu\right)\frac{\sigma_{eq}}{E\varepsilon_{eq}} \quad (9.30)$$

For a pre-dominantly plastic state ν_{eff} tends to the value ½.

When using the maximum shear amplitude, analogous to the Dang Van approach, it must be noted that the relation between maximum shear and axial strain in the uniaxial case is

$\gamma_a = (1+\nu)\varepsilon_a$, with $\nu = 1/2$ for plastic conditions. Then the Coffin-Manson expression becomes

$$\gamma_{a,\max} = (1+\nu)\frac{\sigma_f'}{E}\left(2N_f\right)^b + \left(1+\tfrac{1}{2}\right)\varepsilon_f'\left(2N_f\right)^c. \quad (9.31)$$

Both expressions have proved to give good correlation with experimental data in different investigations.

In this connection it is of interest to note how the uniaxial value ε in the Coffin-Manson model is generally used. The parameters are determined from tests on round bars giving a more or less uniaxial stress state resulting in principal strain amplitudes $\varepsilon_2, \varepsilon_3 = -\nu_{eff}\varepsilon_1$ and an equivalent strain $\varepsilon_{eq} = \varepsilon_1$. When applying the model to notched configurations the biggest principal strain is evaluated from FEM calculations or from Neubers formula and used to represent ε. However, in notches one must assume a condition of plane strain, giving principal strain amplitudes $\varepsilon_2 = 0$ and $\varepsilon_3 = -\nu_{eff}\varepsilon_1/\left(1-\nu_{eff}\right)$ and

$$\varepsilon_{eq} = \varepsilon_1\frac{\sqrt{1-\nu_{eff}+\nu_{eff}^2}}{\left(1-\nu_{eff}^2\right)} \quad . \quad (9.32)$$

This value is close to ε_1 for n_{eff} around 0,3 but takes on the value $2\varepsilon_1/\sqrt{3}$ for $\nu_{eff} = 0{,}5$, i e for predominantly plastic states of strain. (Cf Chapter 6.6.)

For more general states of multiaxial strain some critical plane approaches have been suggested. One, by Kandil, Brown and Miller [9.4], states that the parameter

$$C = \gamma_{a,max} + A\varepsilon_{max} \quad (9.33)$$

is a measure of the LCF life, where $\gamma_{a,max}$ is the half the maximum shear change on a plane and ε_{max} is the maximum strain normal to that plane during a straining sequence. This model is analogous to the Dang Van model for fatigue initiation in that it employs a linear combination of a shear/shear stress and normal strain/stress on a plane.

The parameter A and the function C=C(N_f) can in principle be found by insertion of results from uniaxial Coffin-Manson test results. It may be observed that with A=1 and a uniaxial test one obtains for the 45^0 plane that

$$\gamma_{a,max} = \frac{(1+\nu_{eff})}{2}\varepsilon_a \quad \text{and} \quad \varepsilon_{max} = \frac{(1-\nu_{eff})}{2}\varepsilon_a ,$$

$$\text{i e } C(N_f) = \varepsilon_a = \frac{\sigma_f'}{E}(2N_f)^b + \varepsilon_f'(2N_f)^c \tag{9.34}$$

Another, more recent approach, similar to the Smith- Watson- Topper model for uniaxial LCF, is the one suggested by Jiang-Sehitoglu [9.5]. This is again connected to energy spent rather than strain changes as the Kandil,Brown and Miller models. It is formulated as a fatigue parameter

$$JS = \frac{\Delta\varepsilon}{2}\sigma_{\max} + C_{JS}\Delta\gamma\Delta\tau \tag{9.35}$$

being maximum for any time and plane at the point of interest and where the first term is zero if $\sigma_{max} \leq 0$. The fatigue life N_f is then estimated from

$$(JS - JS_0)^m N_f = C, \tag{9.36}$$

where JS_o , m and C are material parameters.

On a plane where JS is evaluated, during a time sequence of loading,

- $\Delta\varepsilon$ is the maximum difference in strain normal to the plane

- σ_{max} is the maximum normal stress at the plane

- $\Delta\tau$ *and* $\Delta\gamma$ are the maximum shear stress and shear strain ranges in the plane that can be found in a loop describing the shear stress and shear strain vectors respectively during the loading sequence.

Also this model has been found to correlate well with experiments. The material parameters can in principle be assessed from data for uniaxial experiments.

In uniaxial cases of LCF there are several methods to include influence from mean stresses, see Chapter 6. In the two multi-axial approaches described this influence is implicit in the terms ε_{max} and σ_{max} respectively.

9.5 Fatigue crack growth under multi-axial loading

This is a complicated and relatively unexplored area. Only some guidelines for specialized cases are therefore discussed.

First, it is essential to distinguish between a crack that is growing under the influence of a stress sequence, and a virgin crack in a structure that may or may not start growing when exerted to stresses. This is so because the growing crack has already adjusted to the stress field and grows in the direction of the crack tip, whereas a crack in a new structure may well start growing in a "kink", a completely new direction. The criteria for these two cases should be different.

For the initiation of crack growth both the R value and the mixity of the modes should be significant due to closure effects. For the plane case with only mode I and II present an early study showed that cracks did not start growing for ΔK_I - and ΔK_{II} – values according to Figure 9.5 with proportional loading.

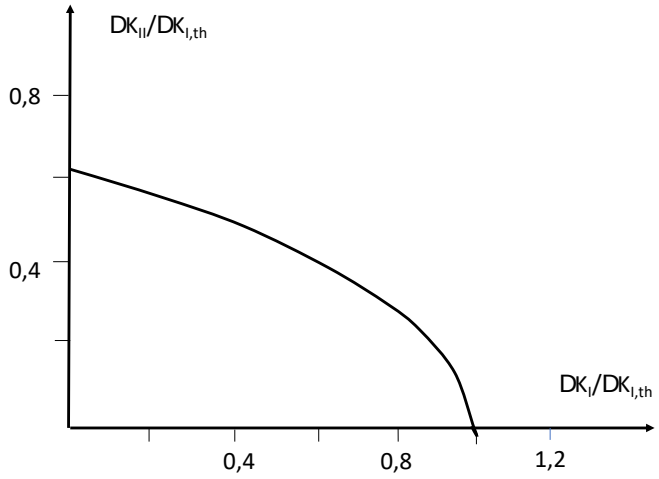


Figure 9.5. Sketch of experimental results from Gao et al showing the limit for crack initiation in a mixed mode with R=0,2.

From these results it is tempting to suggest a criterion

$$\left(\frac{\Delta K_I}{\Delta K_{I,th}}\right)^2 + \left(\frac{\Delta K_{II}}{0{,}6\Delta K_{I,th}}\right)^2 = 1 \qquad (9.37)$$

The factor 0,8 instead of 0,6 has also been suggested.

For non-proportional loading there are no reliable criteria, since the combination of closure effects are unsure when the relationship between K_I and K_{II} changes.

There is a vast but inconclusive literature also as regards the growth of a crack under multiaxial loading. A general analysis is missing. It seems that the crack adjusts to a direction suitable for the co-acting of the opening and shear modes. A formula for an equivalent ΔK_{eq} to be used in Paris law for crack growth has been suggested by Tanaka.

$$\Delta K_{eq} = \left(\Delta K_I{}^4 + 8\Delta K_{II}{}^4 + 8\Delta K_{III}{}^4/(1-\nu)\right)^{1/4} \qquad (9.38)$$

For plane cases it seems that (9.38) is similar to in its structure to (9.37). Probably a Paris law related to some mean R-value should be used.

9.6 Concluding remarks

The chapter describes the most usual methods to treat general multi-axial load and stress sequences. They may comprise several cycles of stresses in varying directions before they are repeated. As is seen the treatment is complicated and the stress analysis generally requires some numerical method to give $\sigma_{ij}(t)$, $0<t\leq T$, and then $s_{ij,a}(t)$ from $s_{ij,mid}$.

In most real situations the stress state is simpler. One major cause is that the most severely stressed region is at the surface of the component, giving a plane stress state, Another cause is that common load situations are bi-axial and/or in-phase, which simplifies both the stress analysis and the assessment of fatigue risk.

In fact, for practical reasons a majority of the experimental background in the literature consists of such simplified load situations. From these theoretical generalizations have been suggested. There is unavoidably some dispersion in the results, *and in the calibration of parameters from uni-axial data.* This means that, particularly when initiation of fatigue shall be avoided, the usual principle of a safety factor can be applied.

References

[9.1] Sines, G., "Behavior of Metals Under Complex Static and Alternating Stresses," in Metal Fatigue, G. Sines and J.L. Waisman. Eds., McGraw Hill, 1959, 145-169.

[9.2] Crossland, B. Institution of mechanical engineers. In Conf. on Fatigue of Metals, pages 138– 149, 1956.

[9.3] Dang Van, K., Griveau, B. and Message, O. in 'Biaxial and Miltiaxial Fatigue', EGF3, Mechanical Engineering Publications, London, 1989, p. 479

[9.4] Kandil, F.A. , Brown, M.W., and Miller, K.J. (1982) “Biaxial low-cycle fatigue fracture of 316 stainless steel at elevated temperatures, Vol 280, The Metals Society, London, pp 203-210.

[9.5] Jiang, Y. and Sehitoglu, H. (1999) “A model for rolling contact failure”. Wear, Volume 244, pp 38-49.

10 Evaluation of material properties

Reliable and commonly accepted data for material properties are of course vital for fatigue analysis and design. In order to achieve this, the experimental methods employed are often those described international standards, as ASTM, DIN, BS or AFNOR. In particular the ASTM standards have come to play an important role, and some of them are mentioned below. It is to be noted that standards are up-dated periodically, and that the latest version should be used. In any case the version used should be referred to. Many large companies also have established company standards, some of which have come to general use, as the ones of ABB or VOLVO in Sweden.

The sensitivity to the precise testing conditions, as temperature, size, surface treatment, environment etc. is an important aspect in the evaluation of fatigue properties. The fatigue properties are stochastic variables, for which the dispersion has to be assessed in addition to the mean value. This differs from the normal determination of solid mechanics parameters, as yield or ultimate tensile strength, modulus etc., where the dispersion is generally much smaller than the mean, and where a single value of this mean is tacitly assumed to be sufficient.

The necessity to evaluate statistical scatter by large test samples makes acquisition of fatigue data lengthy and expensive. Hence, there is a scarcity of data, which explains the many rules of thumb and empirical correction factors introduced in Chapter 4-8.

The scatter in material properties also has to be distinguished from two other effects, the scatter emerging from the experimental technique (determination of forces, strains etc.), and irregularities due to the fact that it is not at all sure that the assumed mathematical relationship (e g the assumption of a linear S-N-curve in logarithmic scales) is a physical reality. It is also worth noting, as has been revealed in several inter-laboratory comparisons, that small variations in performance of standardized methods can produce large differences in results. Such effects can be due to rather slight defects in the test equipment or to the established (mis-)practice of laboratory personnel. Here it is important to understand and use the concepts of *repeatability* and *reproducibility*.

It has become more common, with the development of more and more sophisticated experimental facilities, to evaluate fatigue properties directly for large structural components, hence incorporating effects of surfaces, notches etc. The methods described in this chapter are limited to those where basic material properties are evaluated for standardized test specimens under well-specified conditions.

10.1 Assessment of the S-N curve and the endurance limit

These tests are performed with well specified cylindrical specimens, polished and with diameter 6-10 mm, in order to give the base data, then used for further adjustments with regard to size, surface treatment etc., according to Chapter 4. A typical test specimen is shown in Figure 10.1.

The historically important rotating bend test described in Chapter 2 is taken as a reference and it is also the one most easily performed. The data obtained are normally given an extra subscript r or rb, as in [10.6]. With this method no mean value influence can be assessed.

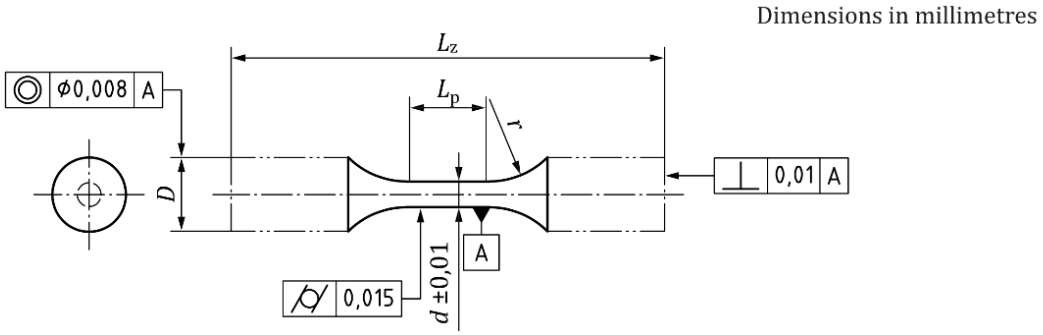


Figure 10.1 Fatigue test specimen. The requirements for surface finish given in the standard is R_a 0.2. Note the radii from the gripping part to the narrow test section.

With the modern testing systems described in Chapter 10.4 tests can be performed for bending, tension and torsion of specimens, including mean values. Some examples of results are given in Chapter 4.

It is noted, as also mentioned in Chapter 4, that the fatigue data obtained for the different modes are slightly different. In the rotating bend test all points on the surface are exerted to an alternating stress, contrary to the pure bend test which has maximum stress only at two points, and with a gradient effect. In tension-compression test all points are exerted to the same stress. This means that, generally, the fatigue properties seem to be slightly better for bending than for rotating bending, and significantly better than for tension-compression. For torsion tests again, the material mechanisms are different, leading to a low mean value dependence and specific shear stress data for fatigue.

10.1.1 The S-N-curve

For each test specimen the life to fracture for a pre-determined stress level is obtained. The life is the sum of the inter- and trans-granular starting phases, and the growth phase. For lower stresses the initial phases take up most of the life and the statistical scatter is larger. A number of tests on different stress levels may give results looking like those in Figure 10.2 a).

The S-N-curve, i e the coefficients of the Basquin equation, and measures of scatter, can be assessed in different ways. One method, which gives a good opportunity to determine the statistical properties, is to use few load levels and many specimens on each level, see Figure 10.2 b). The basic assumption, verified by large numbers of experiments, is that the distribution of scatter is approximately log-normal, i e the logarithms of the lives are normally distributed. Alternatively, the Weibull distribution has been suggested.

So, for a set of data $x_i = \log N_i$, i = 1,n, where N_i are the lives of the n specimens on some load level, assessments of the mean m and the standard deviation s are obtained as

$$m = \sum_i x_i / n \qquad s^2 = (\sum_i (x_i - m)^2)/(n-1) \qquad (10.1)$$

If n is small the significance of the results is low.

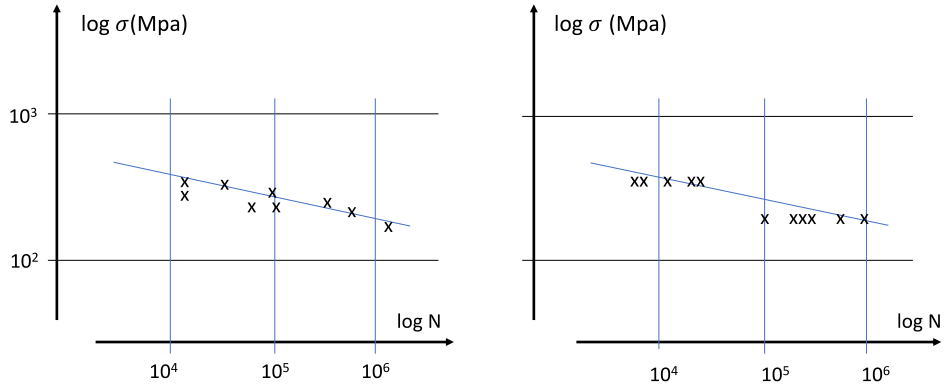


Figure 10.2 a) Example of results from fatigue life tests in general, b) results where the tests have been concentrated to only two stress levels.

Since m and s are estimates, a safe interval defined by a number of standard deviations k for a risk p for failure at a certain x, e g $p \leq 2.5\%$ for $x \leq m + k\,s$ where $k = 2$, can not be given.

	a=0,10		a=0,05	
n \ P	0,90	0,95	0,90	0,95
2	15,978	18,800	32,019	37,674
3	5,847	6,919	8,380	9,916
4	4,166	4,943	5,369	6,370
5	3,494	4,152	4,275	5,079
6	3,131	3,723	3,712	4,414
7	2,902	3,452	3,369	4,007
8	2,743	3,264	3,136	3,732
9	2,626	3,125	2,967	3,532
10	2,535	3,018	2,839	3,379
11	2,463	2,933	2,737	3,259
12	2,404	2,863	2,655	3,162
13	2,355	2,805	2,587	3,081
14	2,314	2,756	2,529	3,012
15	2,278	2,713	2,480	2,954
16	2,246	2,676	2,437	2,903
17	2,219	2,643	2,400	2,858
18	2,194	2,614	2,366	2,819
19	2,172	2,588	2,337	2,784
20	2,152	2,564	2,310	2,752
50	*1,916*	*2,284*	*1,996*	*2,379*
100	*1,822*	*2,172*	*1,874*	*2,233*
∞	*1,645*	*1,960*	*1,645*	*1,960*

Figure 10.3 P, a, k table (from [10.1]). Note that P denotes the whole coverage of the distribution, so that p=0,5P. Only n-values up to 20 have been included. (For n> 20 the results converge to the one for large n-values.

Instead, through statistical analysis, a k can be given such that the risk for failure p can be determined on a given level of confidence a. k is tabled, see Figure 10.3, as a function of n, p and a. The interpretation of this is that if a large number of samples, each with n specimens, are taken, in the long run the statement for p is correct for the percentage 1-a of the samples.

A visual impression of the results, e g in order to assess that the distribution is actually near log-normal, can be obtained by making a diagram for the normal distribution, where the vertical axis is scaled as the normal distribution function $\Phi(p)$ of the probabilities p on the vertical axis. The logarithms of the results are ordered from the smallest i = 1 to the biggest i = n. The probabilities are calculated as $p_i = (3i - 1)/(3n + 1)$ and the data pairs are introduced in the diagram, from which the mean and the standard deviation can be estimated directly. From the diagram it can also be seen how well the assumption of normality agrees with the data, i e how well the data fit to a straight line.

Example
On the nominal stress level 220 MPa nine experiments on a structural component have been performed. The results and the data for treatment are given in Table 10.1

Table 10.1

Test number, i	1	2	3	4	5	6	7	8	9
Life, N_i	1.42 10^5	1.66 10^5	1.80 10^5	2.45 10^5	3.64 10^5	4.02 10^5	4.56 10^5	5.04 10^5	6.40 10^5
$\log N_i = x_i$	5.152	5.220	5.255	5.389	5.561	5.604	5.653	5.702	5.806
$p_i = (3i-1)/(3n+1)$	0.071	0.179	0.286	0.393	0.500	0.607	0.714	0.821	0.929

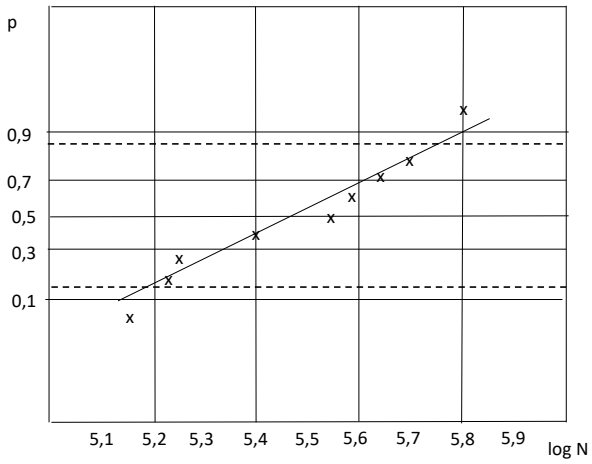


Figure 10.4 Example of how data can be assessed in a normal distribution diagram. (Standard deviation s is p=0,5 $\pm$ 0,34, dashed in the Figure.)

In a normal distribution diagram, the data and an eye-sight straight line are shown in Figure 10.4

From the diagram the assumption of a normal distribution can be assessed; here the data seem to form an s-shaped curve rather than a straight line, however the data are few. Approximate values of the mean and the standard deviation are found as $m = \log N_{mean} = 5.46$ and $\log s = 0.30$.

Treatment of the data by the formulas (10.1) for m and s above give m = 5.482 and log s = 0.235.

For a required p = 0.95 and 1-a= 0.95 a k = 2.97 is found in Figure 10.3 for n = 9, P=0.90 and a=0.05. Hence a 5 % risk for failure at the confidence level 0.95 (19 out of 20) is found from

$m - k\,s = 5.482 - 2.97\ 0.235 = 4.78$

giving $N = 10^{4.78}$ = 60 000 cycles, to be compared with the mean $10^m = 10^{5.482}$ = 303 000 cycles!

If the mean and the standard deviation were known instead of assessed from a sample of 9, the 5 % risk would be found from the normal distribution giving a k = 1.64 and

$m - k\,s = 5.482 - 1.64\ 0.235 = 5.096$

and a life $N = 10^{5.096}$ = 125 000 cycles.

The method described in some detail above is suited for analysis of data grouped according to Figure 10.2 b). A general way to find the parameters of a straight line describing a set of data, including their uncertainties, is described in Appendix A2. This may be useful in particular if test data do no come from a small number of load levels, as in Figure 10.2 a).

In addition, it is worth mentioning the method of maximum likelihood estimates. This is useful if the data is an unstructured set of lives and stress levels, but it requires computer treatment. A shape for the S-N curve is assumed, e g a straight line (2 parameters). Further, a standard deviation is assumed, which may be constant or stress dependent (1 or more parameters). The probability for each of the test results, as a function of the set of parameters, is calculated. Then the set of parameters is determined iteratively that maximizes the product of the probabilities (or the sum of their logarithms). The method includes the treatment of non-fractured tests, i e the experimental data is used slightly more efficiently.

The Basquin equation can be determined either by linear regression of data from Figure 10.2 a) according to Appendix 2, or by fitting the two mean values obtained from the analysis of data according to Figure 10.2 b) to a straight line.

10.1.2 The endurance limit

Many experimental results, and the reasoning about formation of glide bands and grain boundary barriers in Chapter 3, indicate that for large groups of materials there exists an endurance limit.

This endurance limit is also a stochastic material parameter appearing to have a considerable dispersion, which is natural considering the physical background. The assessment of the mean and standard deviation of the endurance limit has to be dealt with in a special fashion since the only experiment that can be performed is to run a test at a specified level and note the number of cycles to fracture, or that fracture does not occur within a specified life span representing "infinite life". Standardized lives for determination of the endurance limit are generally taken to be $2\ 10^6$ (e g welded structures) or 10^7 (smooth specimens). Obviously, running a full matrix of specimens, i e n specimens for each of m levels to find the probability distribution of failure vs non-failure would be very costly considering that both n and m should be of the order of

20-30. Hence, two methods that offer some simplification are usually employed. In both of them it is assumed that the distribution is normal.

In one of them, two stress levels are chosen from experience, e g from the lower load levels of the S-N-curve. They should encompass the mean value. Then a number of tests are run for each level, the probabilities for fracture are calculated and plotted in a normal distribution diagram (Figure 10.4), and a straight line between them is drawn, giving graphically an assessment of the mean and the standard deviation. Of course, the data can be treated analytically as well. The accuracy of the estimate depends on the number of tests, which should be in the order of 10-20 at each level. In many cases there is experience from the properties of similar materials enabling the number of tests to be reduced.

The other, more commonly used method is the one called the staircase method [10.2]. It uses a technique to concentrate successive tests around the mean value. Starting with a guessed value of the endurance limit, σ_e, and a fixed step for stress changes, d, which should be chosen from experience or checked afterwards to be a fraction (< 0.5) of the standard deviation, a series of tests is performed.

If fracture occurs within the number of cycles decided as the life for determination of the endurance limit (e g 10^7), the stress level for the next test is lowered with one step. If no fracture occurs, the level is increased one step. When the test series is completed, calculations are made, based on a maximum likelihood approach for the normal distribution. The procedure is as follows.

The least frequent occurrence of failure and survival is called an event. Indicate level numbers with i, and put i=0 for the lowest level where an event occurs. Denote the number of events at each level with J_i. Then put

$$J = \sum_i J_i \tag{10.2}$$

$$A = \sum_i i J_i \tag{10.3}$$

$$B = \sum_i i^2 J_i \tag{10.4}$$

From this, estimates of mean m and standard deviation s are obtained

$$m = \sigma(i = 0) + d\left(\frac{A}{J} \pm \frac{1}{2}\right) \tag{10.5}$$

(with + for the event survival and – for the event failure)

$$s = 1.62d\left(\frac{BJ - A^2}{J^2} + 0.029\right) \tag{10.6}$$

A check should be made that s > d according to the above. The quality of the estimate depends on the number of samples used, as in all sampling of assumed distributions.

10.2 Data for strain-based (LCF) design

Since repeated cycles with plastic straining have to be applied in these tests it is important to use a testing system that is designed to apply centred uni-axial displacements, and that the two grips do not impose any angular displacement. Further, since cycling goes from compression to tension no play in the grips can be tolerated.

Standard test specimens have to be designed so that the risk for plastic instability is minimised. This means that they have to be short and sturdy, often having an hour-glass shape, see Figure 10.5

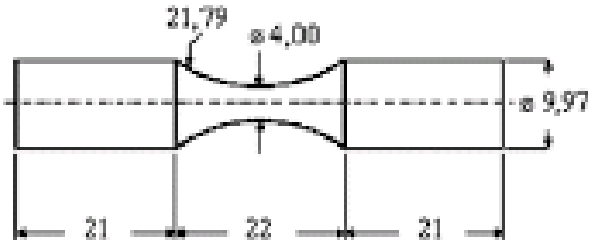


Figure 10.5 Example of test specimen for evaluation of LCF properties.

For the strain measurements a clip-gauge is normally used, based on strain gauges and with a measuring length of 10 – 25 mm. Although laser based devices can also be used, the strain gauges are advantageous since they can easily be adapted to a closed loop control system. The tests are run in strain control and most often in a servo-hydraulic testing system (see Chapter 10.4). Since strains have to be measured based on short gage lengths the gages have to be sensitive, well calibrated and with a good repeatability.

(Consider for example that the typical strain at half the yield stress in a medium strength steel is around 300/200 000 = 0,0015. With a gage length of 10 mm this corresponds to a displacement of 0,015 mm, and a measurement error of 1 % will correspond to a displacement error of 0,00015 mm = 0,15 mm, i e the order of a grain diameter.)

Often material data for strain-based design are required for elevated temperatures. There are different solutions, e g induction heating or temperature chambers, available for this purpose, as well as servo control systems and temperature adapted gauges.

There are several international companies offering complete testing systems including servo systems for running arbitrary strain-time functions, and performing stress measurements and temperature control, as MTS [10.3] and INSTRON [10.4].

Since the tests produce heat by plastic deformation there may be restrictions on the frequency by which tests can be performed, or cooling may be required. If the frequency is very low and plastic strains are considerable creep phenomena may disturb the measurements.

For loads in the plastic range most materials appear to be cyclically softening or hardening, i e for a constant strain amplitude the stress amplitude decreases or increases for the first number of cycles until an asymptotic steady state appears with a constant stress-strain loop. See Chapter 6.

In fatigue design it is these cyclic properties that are needed. The cyclic stress-strain curve is obtained as the curve resulting from joining the maxima of stress vs strain for a number of tests with different maximum strain, as discussed in Chapter 6. The curve is often fitted to the Ramberg-Osgood material model

$$\varepsilon_a = \frac{\sigma_a}{E} + \left(\frac{\sigma_a}{H'}\right)^{1/n'} \qquad (10.7)$$

The elastic term, from a part of the curve with low stresses, is separated and a fit is made by linear regression of the non-linear part on a log-log scale. Conventionally, the cyclic parameters are denoted by a prime sign $H' \text{ and } n'$ instead of H and n.

The material parameters in the Coffin-Manson strain-life relationship

$$\varepsilon_a = \frac{\sigma_f{}'}{E}\left(2N_f\right)^b + \varepsilon_f{}'\left(2N_f\right)^c \qquad (10.8)$$

are obtained by running a number of specimens at different constant strain amplitudes to fracture, often according to a standard, as the ASTM [5] one. A table of data points e_a, s_a, and $2N_f$ is obtained. The elastic and plastic parts are separated from each other, i e

$$\varepsilon_{pa} = \varepsilon_a - \frac{\sigma_a}{E}\,. \qquad (10.9.a)$$

Then the two sets of parameters, $\sigma_f{}' \text{ and } b$ and $\varepsilon_f{}' \text{ and } c$ respectively, are assessed by linear regression according to Appendix 2 on a log-log scale, i e from plots of log ε_{ea} and log ε_{pa} vs log($2N_f$):

$$\log \varepsilon_{ea} = \log\left(\sigma_f'/E\right) + b\log\left(2N_f\right) \qquad (10.9.b)$$

and

$$\log \varepsilon_{pa} = \log\varepsilon_f' + c\log\left(2N_f\right) \qquad 10.9.c)$$

The uncertainty is generally smaller for shorter lives, since the larger amount of plasticity generates overall damage, instead of locally at weak points as for fatigue initiation at elastic stress levels.

10.3 Data for fracture mechanics based design

The data needed are those necessary to approximate the sigmoidal relationship between crack growth rate and span in stress intensity factor, see Figure 10.6.

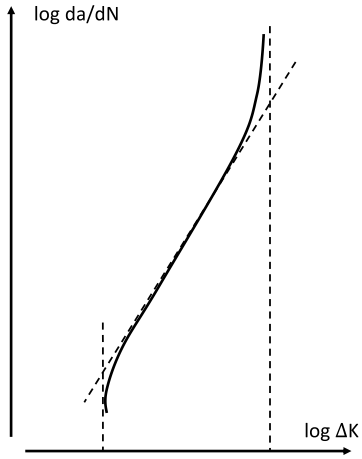


Figure 10.6 Sketch of the relationship between crack growth rate and stress intensity factor range. The dashed lines are to be assessed by experiments.

The two vertical asymptotes are found by different techniques. The fracture toughness, K_{Ic} , is obtained by a standardised procedure, most often the ASTM E399, also described in the Swedish "KTH Solid Mechanics Handbook" [10.6]. First, a pre-crack is produced from a notch by fatigue loading under specified conditions in bend or compact test specimens, see Figure 10.7 a). Then a monotonically increasing load is applied until fracture occurs while the force-displacement relationship is monitored and plotted, see Figure 10.7 b). The evaluation includes measuring of the crack length and its shape, and the linearity of the plot, in order to ensure that the presumption of linear elastic fracture mechanics is approximately valid. If the conditions of the standard are fulfilled, a K_{Ic} value can be defined from a certain value P_Q obtained from the plot and the relationship

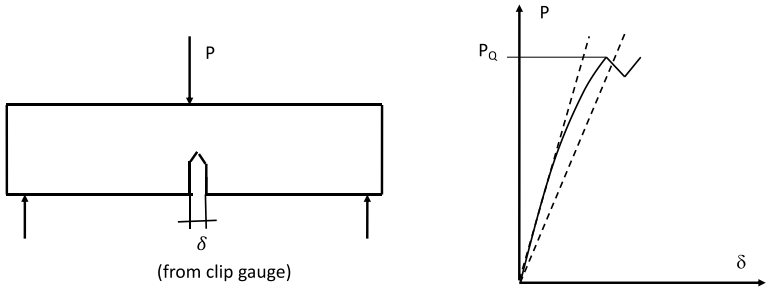


Figure 10.7 a) Sketch of specimen and setup for evaluation of K_{Ic}. b) Force vs crack opening displacement from fracture toughness test. The shape of the curve may vary, the P_Q taken for evaluation is the maximum within the linear elastic slope and a line with 95 per cent of that slope.

between P and K. (See e g Figure A3.3 for the compact test specimen how to find this relationship). Since the fracture toughness depends on local properties at the crack tip it is a stochastic material parameter with a dispersion, which is of limited significance, typically some

per cents. Further, K_{Ic} is coupled to brittleness and hence is strongly temperature dependent. In handbooks K_{Ic} values should be given for specified temperatures.

It is noted that the validity of the test is coupled to the size of the specimen. The more ductile the steel and the higher the K_{Ic} is, the larger the test specimen has to be. So, if the conditions for a valid test are not fulfilled, a larger test specimen has to be chosen. See also the discussion of validity of linear fracture mechanics conditions in Chapter 7.

The threshold value, K_{th}, is difficult to evaluate properly. In principle, a fatigue crack is grown in a plate under successively decreasing stress intensity factor, K_I, and the crack growth rate is measured optically or by an electric impedance method. The threshold value is assumed to have been reached when crack growth stops. The difficulties are essentially two. The first one is to decrease the K_I- value in sufficiently small steps and to measure very small growth rates, which requires sensitive equipment and many cycles between observations, which is tedious. The second difficulty is due to the fact that there is a crack retardation effect. When the load is decreased, particularly step-wise, the following cycles will have a smaller growth rate than the steady state one for the new K_I-value. This means that so many a cycles have to be run at the new level that the retardation effect is overcome. Hence, it may easily occur that too large K_{th}-values are obtained. A relevant standard describing the procedure is ASTM E 1681.

Although the threshold value and the lower, curved region of the diagram are really the most important ones to model, since so large proportions of the fatigue life are spent there, most of the practical application in design is concentrated on the, assumed, linear part, the Paris´ law region between the two asymptotes.

This determination of this line is based on measurement of the crack growth rate in a plate with a fatigue crack growing under constant amplitude loading, normally with $R \geq 0$ in order to avoid gripping problems, as play or buckling of the plate. The crack length is registered as a function of the number of cycles. This can be made manually by a low magnification microscope and a grid on the surface of the plate near the crack-tip, or automatically by LF or HF electric devices calibrated to give a signal related to the crack length. The result is a diagram according to Figure 10.8 a).

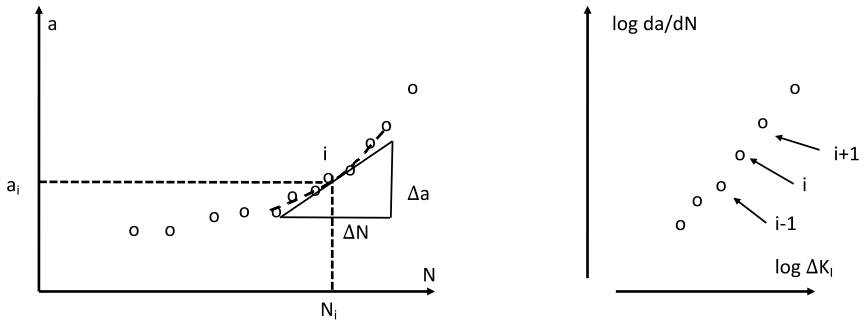


Figure 10.8 a) Sketch of results from crack growth experiments. b) Sketch of resulting data for da/dN and ΔK_I.

For a number of points, i, on this curve the slope, $(\Delta a/\Delta N)_i$, approximating $(da/dN)_i$, and the crack length, a_i, are tabulated. Since the slope is an important parameter and since there is scatter in the data points due to inaccuracy in registering the crack length (it is not easy to determine the position of a crack tip within a few per cent of a millimetre), it is generally determined by construction of a spline function for seven data points, the central one being the one to be

evaluated (indicated in Figure 10.8 a)). The whole procedure is standardised, in e g ASTM E 647

From the crack length, a_i, and the double stress amplitude Δs_{nom}, $\Delta K_{I,i}$ is calculated as

$$\Delta K_{I,i} = f(a_i, geometry)\Delta\sigma_{nom}\sqrt{\pi a_i} \qquad (10.10)$$

where f is the shape function for the plate geometry used.

Now, corresponding values of da/dN and ΔK_I can be plotted on a log-log scale, and the constants C and n in Paris´ law

$$da/dN = C(\Delta K_I)^n, or \log(da/dN) = \log C + n\log(\Delta K_I) \qquad (10.11)$$

can be estimated by linear regression. There is no physical evidence for the assumed linearity of the relationship; rather it is a useful model giving simple calculations and conservative results. Since different ΔK- and da/dN-values are obtained along the curve, in principle the whole relationship is obtained from one single test specimen. Yet, several tests are recommended in order to compensate for differences in the specimens and in order to get more data points. Typical results are shown in Figure 10.8b).

It is again worth to mention that material properties concerning cracks are dependent on the direction of the crack in anisotropic materials, where the grain structure has been changed by e g rolling or forging without subsequent heat treatment.

10.4 Test equipment.

One reason for the multitude of rules of thumb and approximate formulas in fatigue design is that experiments are expensive, partly since equipment is expensive. Tests must run for long times and due to the inherent scatter several specimens have to be used. Further, it is desirable to perform the tests for varying environmental conditions. Here, the principal properties of some of the most used types of test equipment are described.

In previous chapters it was mentioned that the original basis for S-N-curves was rotating bending tests due to their simplicity. The next step is to include mean values and uniaxial and bending modes for constant amplitudes. This can be accomplished by electro-dynamic resonance testing systems. The principle, taken from the supplier RUMUL [10.7], is illustrated in Figure 10.9 Machines are commercially available up to the 100 kN range.

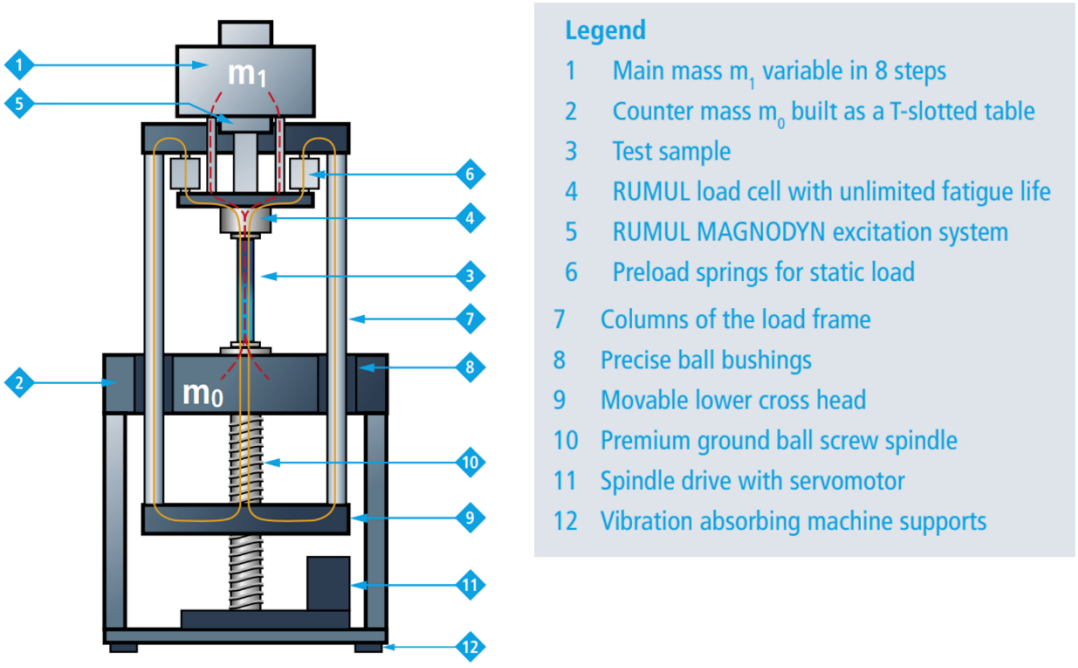


Figure 10.9 Sketch of RUMUL resonance testing system.

In the test specimen area, 3, grips and specimens can have varying geometries as soon as they are possible to centre uniaxially, from normal bar-shaped specimens to e g a set-ups with a gearwheel in a fixture. The machine consists of a base frame connected to the frame carrying the specimen, 7, 9 and upper crosshead, by a spindle and motor,10, 11. By changing the level of 9 the length of the test specimen area can be adjusted, and a static mean load can be applied through the springs 6.

The mechanical resonance system consists of the masses m_0 and m_1, where m_1 can be adjusted in steps to find a suitable frequency in combination with the stiffness k provided by the frame and the test specimen. The test specimen is by far the weakest component and defines k.

The system is driven by an actuator, 5, with an electromagnet, with inductance L, coupled to a battery of capacitors and a source of electric energy. By choosing capacitors iteratively, the electric LC circuit is tuned to be in resonance with the mechanical system:

$$\omega = \sqrt{LC} = \sqrt{k\left(\frac{1}{m_1} + \frac{1}{m_0}\right)} \qquad (10.12)$$

It is desirable to have as high a frequency as possible. The test setup and the machine stiffness give k. The mass m_1 is chosen to give the maximum ω that is practically obtainable, before starting to tune the circuit. Frequencies up to around 200 Hz are possible to obtain, practically obtainable values are generally somewhat lower.

The desired amplitude level is obtained by choosing the input to the electric circuit. The load, both the static mean load and the amplitude, is measured by a load cell, 4, based on strain gauges.

With this type of equipment it is possible to obtain uni-axial S-N-data near the endurance limit for materials and components of limited size at a low cost and within a reasonable time. At 200 Hz a specimen can be exerted to10^7 cycles per day. The machine is also provided by cycle counting devices and cut-off controls reacting to failure of the specimen.

For more demanding tests, with non-sinusoidal loads or with load sequences of arbitrary shape, servo-hydraulic testing systems are versatile. The principle is shown in Figure 10.10 a).

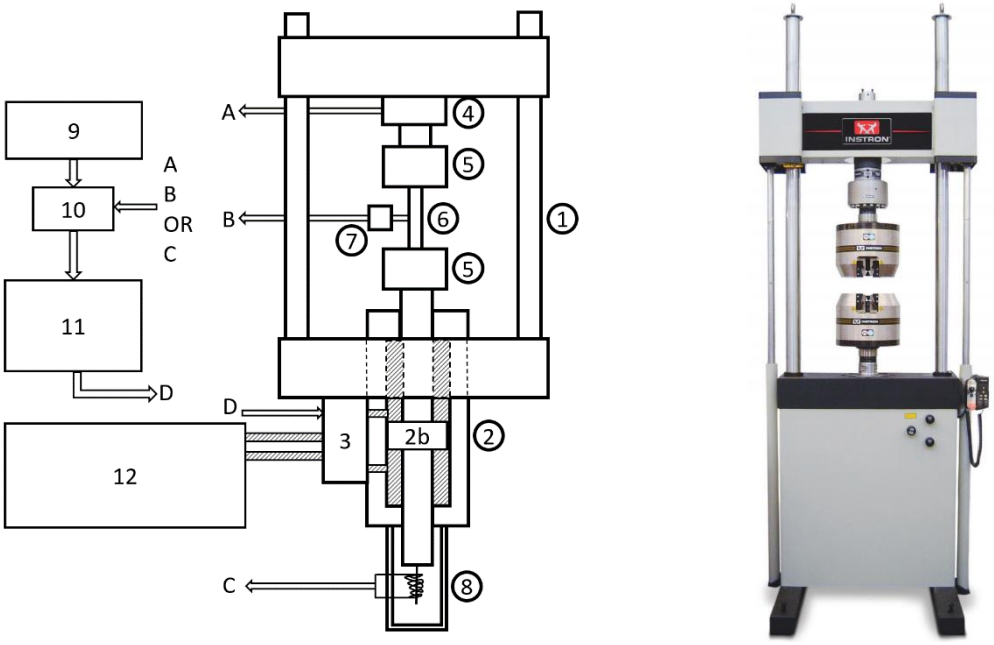


Figure 10.10 a) Principle of servo-hydraulic testing system. b) Photo of typical system (INSTRON).

A typical machine is shown in Figure 10.10 b). There are several suppliers and in most cases a testing system is put together from standard components in order to be optimal for the user, even though the mostly used systems for material testing are put on the market as ready-made so called universal testing systems. The capacities in terms of force range from 50 kN to 1000 kN.

The load frame, 1, is normally built so that the height can be adjusted for specimens of different sizes, either by screws or by hydraulics. The resulting stiffness of the frame, load cell and grips is an important parameter for the dynamic capability of the system, and the columns and the yoke should be rather stiff.

The grips, 5, are often hydraulic clamps. It is of utmost importance that the machine is lined up precisely so that no bending of the test specimen is induced. It has been shown that severe errors may occur in standard fatigue tests by the un-intentional bending moments or non-alignment introduced by the grips.

The servo-valve, 3, with the oil supply and high pressure pump, 12, is really the heart of the system, giving the possibility to produce a desired motion of the double acting piston, 2b, of the actuator,2, as function of time. Normally, the servo-valve has one electronic and two hydraulic steps, and the operating pressure of the pump an actuator is 200 bars. For more details of the properties of servo-valves the reader should consult e g Moog [10.8].

The functioning of the system is as follows.

The system can be run in three different modes, viz. displacement, load, and strain control, see 8, 4, and 6. The displacement is obtained from a linear voltage difference transformer, 8, in the actuator, the force is obtained from the load cell, 4, and the strain (optional) can be obtained from a strain measurement device, 6, (clip-on extensometer) attached to the test specimen.

Once one of the three modes has been chosen, the momentary gauge signal is entered (A,B, or C), conditioned and compared with a desired value given by a function generator, 9, in a conditioner/control unit, I. The function generator may give for example a replicate of a recording from field tests. The difference is given as input to a PID controller, 11, steering the servo valve (D). In this way any desired motion, load or strain can in theory be exerted on the test specimen/object, by a closed loop control system.

Since many different types of test specimens may be used, it is important to tune the system with regard to the test specimen or component to be tested by the aid of the PID controller. One technique for this is to optimize, through a specified methodology the response to a square-shaped wave as demand signal from the function generator.

The capability of the system is a function of the stiffness of the load frame, the stroke and effective surface (force) of the actuator, and the capacity of the pump and servo valve. For specific applications it is important to know this and to cost-optimize the composition of the system ordered. As mentioned, suppliers offer a number of load frames, grips, load cells, actuators, and servo valves as building blocks. For such a system the capability is often described graphically in a logarithmic diagram according to Figure 10.11.

For a harmonic motion with velocity

$$v = v_0 \sin\omega$$

the displacement d and acceleration a can be written

$$d = \frac{v_0}{\omega}\cos\omega = \mathrm{d}_0\cos\omega \text{ and } a = \omega v_0 \cos\omega = a_0\cos\omega$$

respectively.

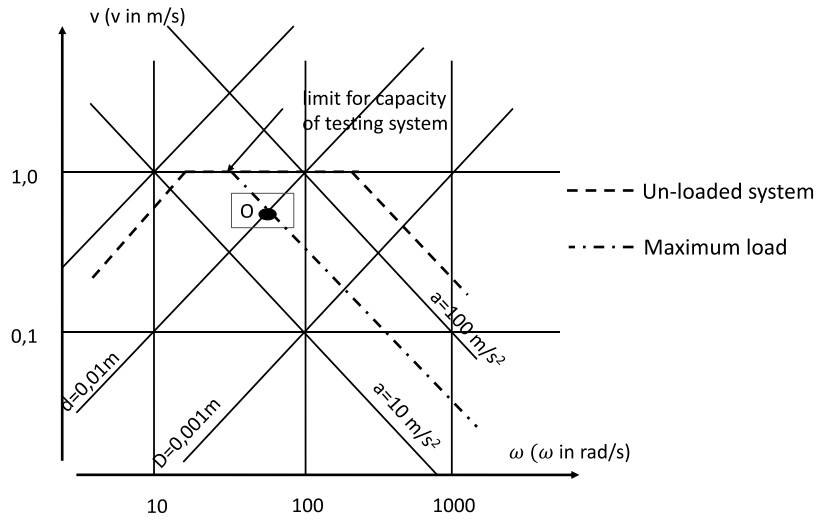


Figure 10.11. Diagram showing the capability of a testing system, dashed lines, with a load point O.

Then, since

$$\log v_0 = \log d_0 + \log\omega \text{ and } \log v_0 = \log a_0 - \log\omega$$

lines with constant displacement, log d_0, have the slope +1 in diagrams with log v_0 as a function of logω, and lines with constant acceleration, log a_0, have the slope -1.

The capacity of the testing system can now be illustrated in this diagram. In the example shown in Figure 10.11 the limit d_0 (0,05 m) is half the stroke of the actuator, the limit v_0 (1 m/s) reflects the capacities of the pump and the servo valve to feed oil to the actuator, and the acceleration a_0 reflects the relationship of the area of the piston and the oil pressure, i e force, and the mass of the piston/grip for an unloaded system. At maximum proof load of the machine the acceleration is considerably smaller, indicated by the dashed line. So, in the example, if a test is required to run at 10 Hz (63 rad/sec) and full load, only an amplitude of about 10 millimetres can be obtained (point O in the Figure). Therefore it is important to verify that complicated spectra with many frequency components can be reproduced accurately enough with the amplitudes or force levels required.

There are also four column machines and bi-axial systems available. The most common bi-axial machine is one where the piston in the actuator (see E in Figure 10.10 a)) is prolonged to a second pressure chamber. In this the piston can be exerted to an angular motion and a resulting torque on the test specimen. In such a machine a specimen can be tested with arbitrary combinations of load sequences in tension/strain and torque/angle. An example of a bi-axial machine for plane tests is also shown in Figure10.12.

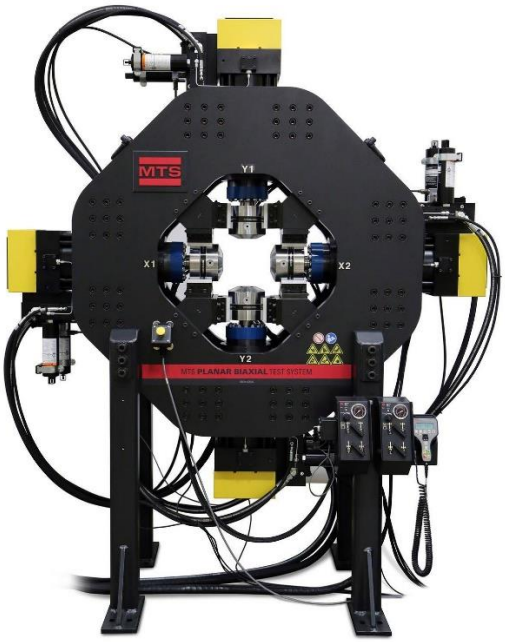


Figure 10.12 Picture of a bi-axial machine for testing of plane specimens.

References

[10.1] Statistics manual, by Crow, E.L. et al, Dover publications, inc., New York (1960).

[10.2] Little, R.E., Statistical Design of Fatigue Experiments. John Wiley & Sons, New York. (1975).

[10.3] MTS (www.mts.com)

[10.4] INSTRON (www.instron.co.uk)

[10.5] ASTM (www.astm.org) e g ASTM 465 and 606.

[10.6] ”Handbok och formelsamling i hållfasthetslära” (in Swedish), KTH, Stockholm (1998)

[10.7] RUMUL (www.rumul.ch)

[10.8] MOOG (www.moog.com)

11 Means to avoid or minimize risks for fatigue damage in the design process

Much of classic fatigue analysis in textbooks presumes a well-defined structure where the objective is to find levels of load for which fatigue does not occur or to find a safe relationship between occurring loads and required life. In real life fatigue analysis is introduced in early stages of design together with other factors of importance, as requirements from production technology, economy, environmental impact etc.

Such analysis must in most cases be verified by experiments in the prototype stage, particularly if the design is a new one or means a significant change of an existing one. If the structure or component is to be produced in great quantities, say gear boxes with a new material in the gear wheels, extensive full time tests may be required due to the complexity of the contact problem and the varying service conditions. Then the analytical stage is the first one in the complete process.

Structures of which only one or a few are produced, as bridges or pressure vessels in power plants pose a special branch of design problems. They are often strictly regulated, and the design is conservative as regards innovative materials or geometries. Safety factors are connected to risk levels. For instance, the risk levels are quite different for military and civil aircrafts. This chapter lists some of the most important issues of importance for fatigue to consider during the design process. It does not go into any detail. The subjects mentioned constitute research fields in themselves and should be studied as such.

11.1 Material

Choice of material

The choice of material or combinations of materials is a crucial task where professional judgement and traditions in industrial branches is blended with such things as requirements on strength, weight, corrosion and temperature on one hand, and production technology and economy on the other hand.

In several branches of industry, particularly in the transportation sector, one is searching for lightweight solutions enabling smaller environment impact, in combination with high crash-worthiness and good manufacturing economy. This can be accomplished by thin-walled structures of higher strength steels, or by the use of Al-alloys and composites, often with joints between different materials.

An inherent difficulty is the fact that low strength materials are normally less susceptible to fracture than high strength alloys, see Figure 11.1, and that the fatigue strength does not increase with yield strength for levels above 1400 MPa. Special alloys with extreme properties are available but they are generally more expensive than commonly available materials. In e g [11.1] several other relationships of importance for choice of materials can be found.

The production of materials and components may introduce anisotropy and in-homogeneity. Such effects may be necessary to consider e g in the choice between a cast component and a pressed steel profile. A notable technology is HIP, hot iso-static pressing, by which steel profiles with very high strength can be produced, but with considerable residual stresses. Then fatigue analysis should consider this analogously with the principles for welded structures.

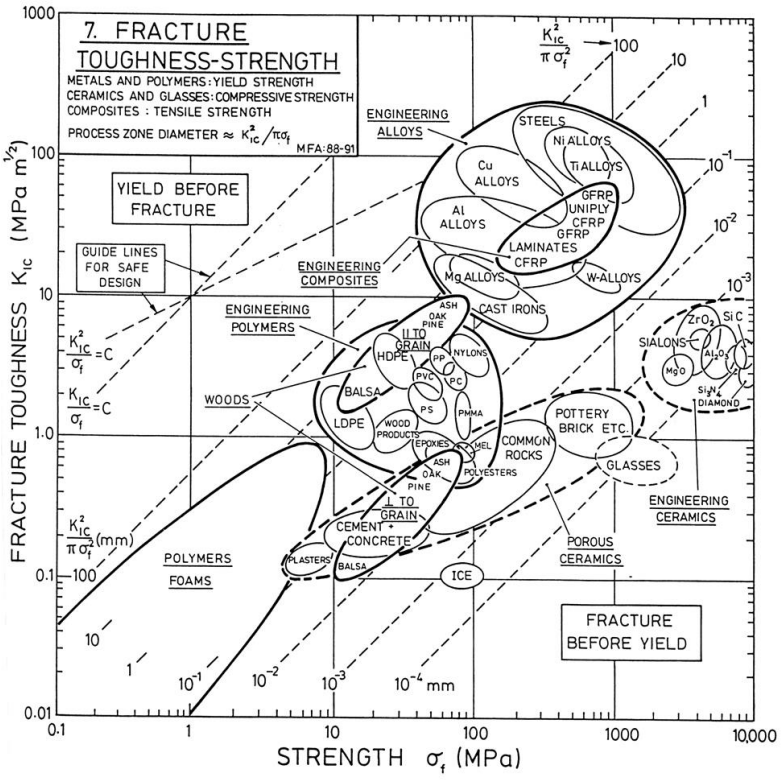


Figure 11.1 Picture showing relationships between strength and fracture toughness for several groups of materials (From [11.1]). A measure of the size of the process zone is also indicated (C f Chapter 3).

The choice of material clearly is an area where it is worth efforts to find an optimal solution by an optimization procedure. Some of the factors mentioned above relevant for the choice of material and production method are listed in Table 11.1.

Table 11.1 Basic factors to consider in the choice of material.

Requirements	**Properties/production**	**Economy**
Load and type of load Size and weight Environmental impact Restrictions in regulations	Strength Weight Workability Residual stresses Inhomogeneities Susceptibility to ***fatigue***, fracture and environment	Cost of material Cost of production Quality, inspection and maintenance costs

11.2 Geometry, fretting, wear, and surface treatment

Instead of analysing in late stages of design the effects of stress raisers, as holes and notches, it is well worth trying to minimize the impact of such stress raisers from the beginning. The introduction of nearly square windows in the Comet jets, related in Chapter 2, was a drastic example of an unnecessarily risky design.

Some simple rules are to avoid holes in areas of high stress, and to make radii as large as possible when changes of diameters are made. Examples from [11.2] are shown in Figure 11.2.

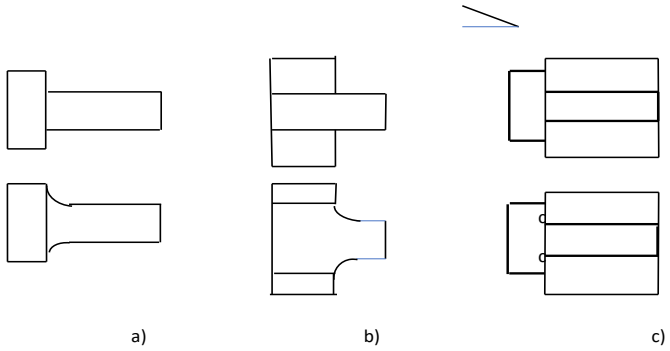


Figure 11.2 a) Tapering of change of radius of shaft, b) relaxation of stresses in shaft, c) change of bolt head design to diminish stress concentration.

A special type of stress raisers are those from contact phenomena. In the branch of bearings the problem is a basic one and has been treated successfully for a long time. The principles are related in Chapter 13. In e g railway mechanics the problem is one of on-going research due to the multitude of influencing factors. Cracks in rails and wheels may initiate and grow due to braking, heavy and irregular train loads, and environmental factors. Hence, the design rules and standards have to be combined with planned inspections and repair. The state of the art of the area is discussed in the literature and on the Internet under the name “RCF” – Rolling Contact Fatigue.

Welding comprises a number of principal problems. There is a heat effected zone (HAZ) between the base and the weld material where the material properties differ and may give rise to inferior fatigue properties. Residual stresses are present, which means that analysis has to start with the assumption that a stress amplitude acts with a high R-value, see Figure 11.3. There are always small cracks present, with a size of a few tenths of a millimetre. Finally, the geometry of the weld implies stress concentrations. Welded structures are treated in Chapter 13.

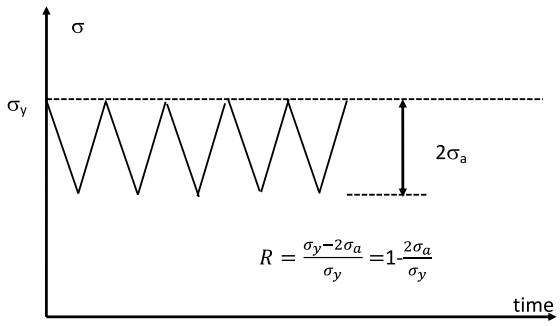


Figure 11.3 Resulting R-value of a stress amplitude applied on a weld with residual stress close to the yield stress.

Joints in general are weak areas with respect to fatigue. In addition to welding and bolted joints there are several other types being of interest, for example in light weight structures when these become of increasing interest for environmental reasons. One area is the combination of metals with polymers as fiberglass. Gluing is an upcoming technology where robotics is used to obtain

high quality. This is necessary since small defects may be harmful, considering e g that the materials have different moduli and temperature properties.

Fretting is a problem which is often overlooked. It is a type of wear due to small relative motions of surfaces in contact. Hard particles are formed from asperities, which then may accelerate deterioration of the surfaces and cracks may initiate. The problem is mitigated by avoiding design including such motions. See Figure 11.2 for examples, and by choice of materials.

11.3 Design

One principle, not always possible to use, is to have *critical parts loaded in compression* instead of tension. One example is from the Markham mine disaster related in Chapter 2, where it would have been easy to make the braking system independent of a rod loaded in tension. Another is railway cars, where the wheel brakes are spring loaded and can only be opened if the compressed air system is working.

The other main way is to have *redundancy*, i e to have several parallel elements carrying the load, so that one can fail without risking the complete structure. Examples are suspension bridges with many cables, and bolted joints. Normally the wheels of a car have five bolts. If one fails the other four are fully capable to keep the wheel in place. Here enters two other principal aspects, se below. One is service and inspection (how often does the car driver look at his wheels?). The other is quality in work (does the driver change wheels between summer and winter, and does the driver have the competence to mount the bolts in the right order and with the correct torque?)

11.4 Manufacturing and quality assurance

Since the risk for fatigue can change considerably for small changes in the properties of a component, it is particularly important that an efficient quality system is in place on the production site, may it be the welding of a large pressure vessel or the casting of thousands of components for the suspension of railway bogies.

There is an extensive number of internationally standardized quality procedures for different purposes. The most general one is ISO 9000. For specific branches company standards are merged, e g in the ISO 16949 for the car manufacturing industry. Personnel performing tasks critical for quality and safety normally undergoes a licensing procedure, as for certain types of welding, ISO 9606.

For those dealing with fatigue design directly or indirectly it is a significant part of the process to be aware of the quality of the production of the designed structures.

11.5 Inspection, maintenance and repair

When the design strategy presumes the existence and growth of cracks, inspection intervals are necessary. Inspection may be performed by the aid of non-destructive test methods, NDT, as magnetic, X-ray or ultra-sound procedures. Also in these cases the personnel may be required to have a license, and a formal education and periodic examinations for the task. The design gives intervals for acceptable crack lengths at inspection that have to be matched with the probability for inspectors to find cracks of these lengths. This was discussed in Chapter 7. The designer should of course know about the limitations and difficulties of non-destructive testing of the different kinds.

A common characteristic for NDT-methods is the probability of detection curve, POD, for the size of typical defects, which are potential risks for fatigue. An example of such curves for three different types of NDT-methods is shown in Figure 11.4. The statistical details are left aside here, as is the detection probability for different shapes and orientations f cracks. The intention of the Figure is just to show the dispersion of results, the order of size of cracks that can be detected, and the efficiency of different methods.

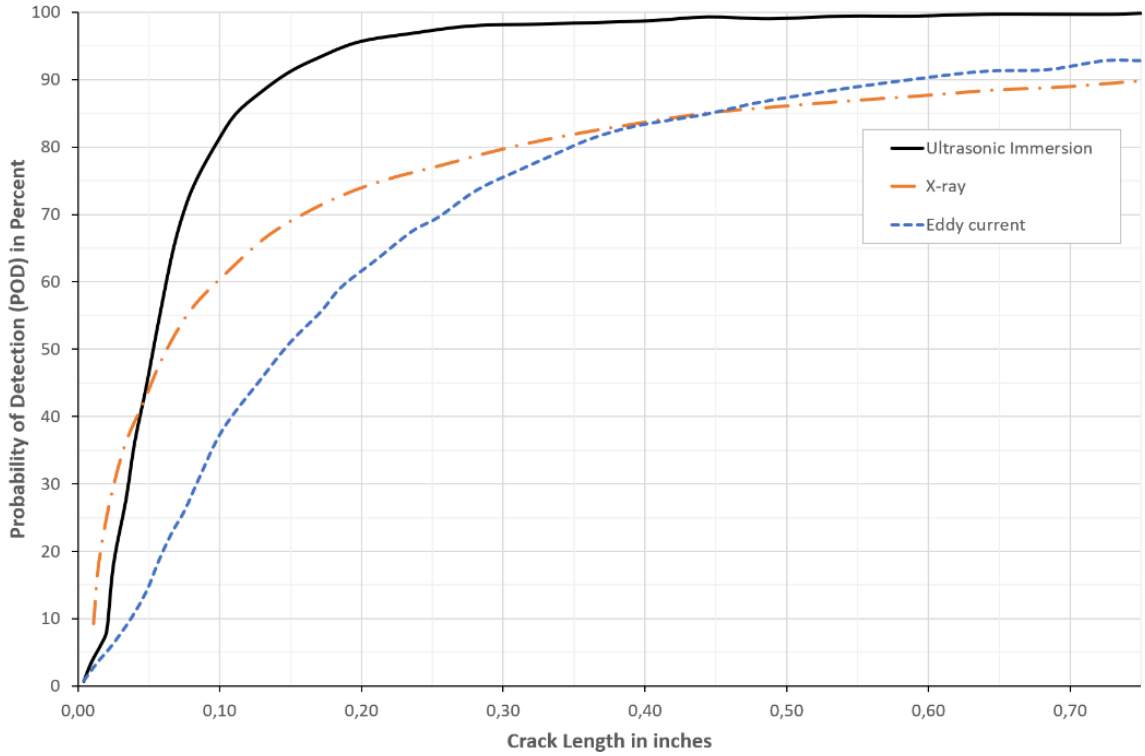


Figure 11.4 Example of POD-curves for some NDT methods.

It is notable that there is a significant interval where the detection is unsure, and that the size of safely detected defects is not insignificant when it comes to brittle materials. Methods may have improved significantly after the publishing of the Figure. It is shown essentially to emphasize the phenomenon.

References

[11.1] www.**lehigh**.edu/~intribos/resources.html

[11.2] H.J. Grover, ”Fatigue of Aircraft Structures”, NAVAIR 01-1A-13, Naval Air Systems Command, Department of the Navy, Washington DC (1966)

12 Corrosion fatigue

Fatigue properties published in handbooks and data sheets from suppliers generally emanate from experiments in laboratory air.
In aqueous environments, as water or moist air, as well as for other service conditions involving contact with chemicals, as chlorides for steels or ammonia for brass, these properties are no longer valid. Fatigue damage occurs for lower stress levels and leads to shorter lives.
The possible combinations of metallic materials and chemicals are more or less infinite in industrial applications. In addition, even rather small concentrations, and small changes in concentration, may result in considerable differences in detrimental effects. Hence, it is not meaningful to try to give rules for the quantitative effects. This is so also when it comes to preventive actions, as painting, galvanizing and other methods for surface protection. Only the principal features can be described here, and the vast literature on the subject has to be consulted for the cases of interest. Alternatively, specific experimental data have to be obtained.
Three common application areas are seawater, for marine applications, moist air for sensitive outdoor applications, and body fluids for implants. Some results for sea water are shown in Figure 12.1 for some materials. They illustrate that both the endurance limit and the S-N-curve are influenced.

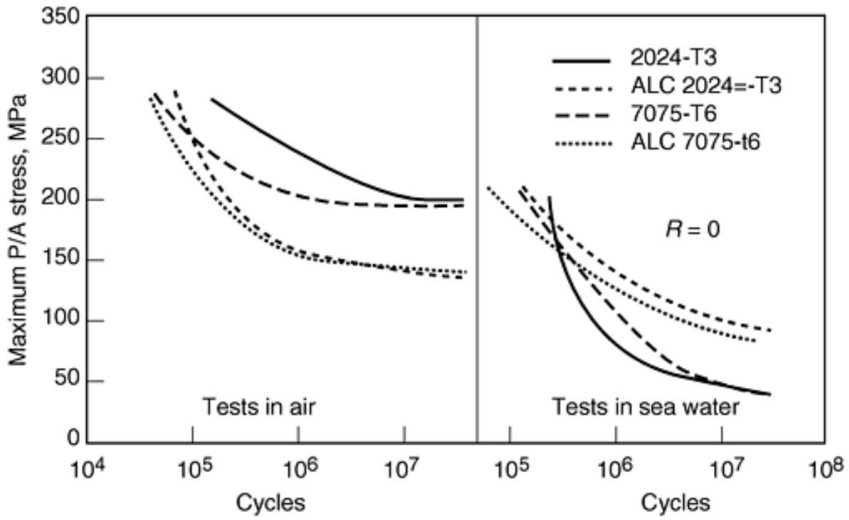


Figure 12.1 Examples of the general effects of corrosive environment on fatigue properties for axial fatigue strength in air and sea water.

12.1 Mechanisms

In both initiation and propagation of cracks there is interaction between substances in the environment and material elements during changes of the load. The mechanisms, as the ones described below, act at the surface of un-cracked material and in opened crack tip regions by disturbing glide band motion and development. The interactions mean formation of small surface particles, at crack tips, and deterioration of borders between glide bands, facilitating crack development and a more brittle behaviour on the macroscopic level. The life is shortened, and the relationship between initiation and growth changes so that a larger part of the life consists of growth.
One should distinguish between the corrosion effects under cyclic and under static loading. The latter case, where a static stress alone is sufficient to cause development of damage is called

stress corrosion cracking (SCC). SCC may interact with corrosion fatigue during fatigue crack growth.

12.1.1 Cyclic loading

There are several conceptual models for corrosion fatigue, see e g [12.2]. It is outside the scope of this text to go into details, partly since the models do not give any systematic or quantitative guidance for design.
One typical process of interest for the present objective of basic understanding, is that of hydrogen embrittlement. See Figure 12.2.

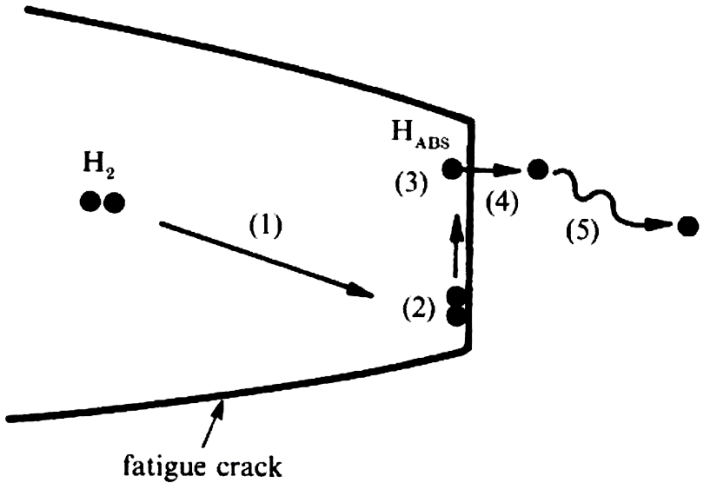


Figure 12.2 Idealized sketch of hydrogen embrittlement (from Suresh [12.2]).

Hydrogen molecules may dissociate to atoms and diffuse into the material and recombine to create increased local internal stresses resulting in embrittlement, i e facilitating the separation of internal bonds. Then microscopic cracks at grain boundaries are more easily formed. They may lead to initiation and continued growth of macroscopic cracks at lower stress levels than otherwise, and also to increased crack growth rate of existing cracks.
The other mechanism mentioned here is that of forming of oxides in the presence of water or acids dissociated into anions. In the case of varying loads, the oxides act locally as obstacles for glide bands and the breaking up of material elements is facilitated. This occurs at the surface or at freshly created glide band elements close to the crack tip. This lowers the endurance limit and increases the speed of damage formation and crack growth.
Contrary to fatigue with no corrosion effects, the frequency of the loading is of importance. This is natural considering that longer time of exposure in each cycle increases the damaging effects.
On the macroscopic level corrosion pits from oxidation are sites for facilitated initiation of crack growth, similar to other treatments coarsening the surface, like grinding or forging as discussed in Chapter 4.5.2. The effects of corrosion can be taken into consideration by experimentally found correction factors for the endurance limit.

12.1.2 Stress corrosion cracking

Stress corrosion cracking (SCC) is the formation and growth of cracks under the combined influence of environment and a constant stress state. The stress state may be a residual one caused by forming or welding. (One famous example of SCC is that of brass cartridges used in

wars in the old days when horses were extensively used, and hygienic conditions were not too good. It was found that cartridges cracked due to residual stresses from forming, and due to ammonia in the atmosphere. The remedy was to anneal the cartridges.)
SCC effects usually mentioned are those in aqueous or chloridic environments, hydrogen embrittlement, in ammonia, called season cracking, and in alkalis, called caustic cracking. They occur in alloys of carbon and austenitic steel, alumina and brass, and the time to deterioration and fracture is a combined function of the levels of stress, concentration of the active substance, and temperature.
One typical mechanism, in addition to hydrogen embrittlement, is the one denoted active path dissolution. Here, the mechanical stress opens susceptible grain boundaries so that the corrosive medium can penetrate more easily into the material.
Tri-axial states and extensive shear deformation occurs at loaded crack tips. Then there is an interaction between corrosion fatigue and stress corrosion cracking. Such interactions occur for many types of chemical environment, and this is the main reason why SCC is important to consider when designing with regard to corrosion fatigue. The growth often seems to be fast as soon as the maximum stress intensity factor exceeds a characteristic value for the material and chemical agent, K_{ISCC}.

12.2 Experimental features

Fatigue life from initiation
The early initiation of surface cracks in corrosion fatigue, resulting in a process where crack growth takes a relatively larger part of the life than initiation, means that the S-N-curve is shifted to the left, and "downwards", and that an endurance limit will not exist, see Figure 12.1.
As mentioned, the complicated and coupled physical processes the quantitative effects are sensitive to the precise environment and metal alloy. Therefore it is difficult to find general rules for the relationships between all the influencing parameters. The number of cycles per time unit also plays a part. The typical effects on the S-N curve is demonstrated in Figure 12.3. Below 10^3 cycles the effect of fatigue is small. The normal Basquin curve for air decreases typically by a factor 2 between 10^3 and 10^6 -10^7 cycles, i e the normal cycle number for endurance limit, s_e. In a corrosive environment the decrease is steeper, see the Figure, and with no apparent endurance limit. The decrease at 10^6 – 10^7 cycles may be considerably more than a factor 2.
This indicates that a change of material, or surface protection should be chosen if there are no experimental data for the precise application.

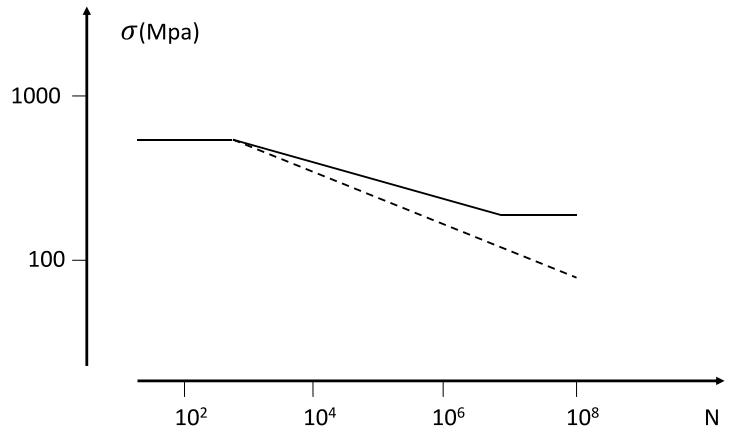


Figure 12.3. Illustration of the principal effect of a corrosive environment on the fatigue properties.

Crack growth
When it comes to the principal features of fatigue crack growth in a corrosive environment, there is a combined effect of corrosion fatigue and SCC, see [12.2]. With the presumption that the effects from corrosion fatigue and from SCC can be considered as essentially separated the picture would be as in Figure 12.4 b) and c).

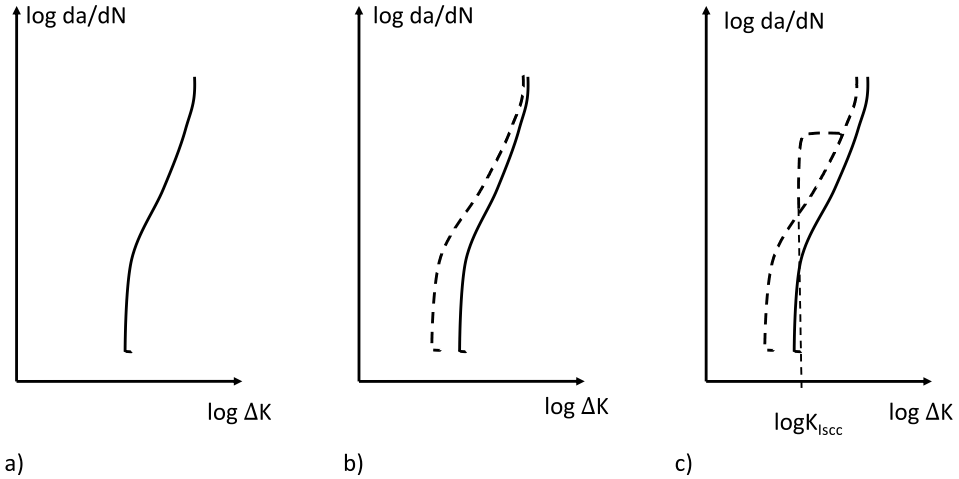


Figure 12.4 Combined action of corrosion fatigue and stress corrosion cracking (R is near 0 in the Figure).

Compared to the crack growth curve from ambient conditions in Figure 12.4 a), see Chapter 7.3, the aggressive environment gives a higher crack growth rate due to cycling, exposing new material surface elements in each load cycle. This means that the curve is moved to the left, Figure 12.4 b). When the maximum K_I exceeds a value typical for the combination of material and environment, the K_{ISCC}- value, there is a drastic increase in crack growth rate, see Figure 12.4 c). This SCC effect appears to dominate as soon as it becomes effective, but there also seems to be a saturation level for crack growth da/dN (≈ da/dt for moderate intervals of frequency). Note the relationship

$$K_{\mathrm{I}max} = K_{ISCC} = \Delta K/(1-R) \qquad (12.1)$$

meaning that the curve is shifted to the left for increasing R-values.
So, when K_{Imax} approaches K_{Ic} the corrosion effects are smaller, since then the pure fracture mechanics based crack growth supersedes. Both the K_{ISCC}-level and the nearly constant level of crack growth, da/dN that is obtained are significant from a design point of view. Both are sensitive to the precise combination of alloy and environment, and as mentioned to frequency and waveform.

12.3 Design

For structures without cracks and with well-defined material, loading (including frequency and waveform) and environment, the corresponding S-N-diagram, Figure 12.1 should be possible to use for design, analogously to the procedures in Chapters 5 and 6, particularly if an endurance limit can be defined. Since these conditions seldom are fulfilled, and the effects of stress concentrations are not well explored, one is often limited to use combinations of material and chemical agent, where the corrosion fatigue effects are small.

The same is generally valid for design that involves the presence of cracks. Here the K_{Iscc} is the governing parameter. The strategy should be to avoid or control the growth of cracks so that the length a_{max} does not violate the relationship

$$K_{ISCC} \geq S\sigma_{max} f(a_{max}) \sqrt{\pi a_{max}} \qquad (12.2)$$

where S is a safety factor.

References

[12.1] Fatigue Design Handbook (AE-10), Society of Automotive Engineers, SAE (1988), ISBN 0-89883-011-7.

[12.2] Suresh, S. "Fatigue of materials", Cambridge University Press (1992), ISBN 0 521 43763 6.

13 Fatigue design in standards

For several types of structures, it is desirable to have standardized procedures for fatigue design. Material and geometry may have features, which are complicated to treat over and over again by general design principles. Further, the safety level may be determined by government requirements or by international agreements. Examples of structures are pressure vessels and structures with welds, bolted joints and bearings.

Such design codes are highly formalized, with prescribed safety factors, definitions of loads etc. A drawback is that the underlying technical principles are to some extent obscured, which may be disadvantageous in cases which are not clear-cut or where new materials or design principles are introduced as a consequence of the technical development. By nature, codes and standards have to be well-proven and negotiated, and hence they tend to be slightly out-dated already from the beginning, particularly with regard to innovative structures and components.

In order to demonstrate the background of design codes in the general design principles described so far, welded structures are treated in some detail in section 13.1. Then, short comments are made on the design principles in codes for bolted joints and bearings. For actual design the relevant codes and standards have to be consulted.

13.1 Structures with welds

Many such structures are critical for safety, as ships, bridges, or lifting devices. They are conventionally made of plates and beams of a well-known group of structural steels, which are welded together. The main elements of design, load analysis, stress-life relationship, and consideration of such features as size, material and surface treatment, are analogous to the ones in general fatigue design, but described in a very formalized fashion. In order to demonstrate the techniques, the procedure used in the Swedish design code BSK [13.1] is related here. Other such codes are e g the ASME- and Eurocodes [13.2], [13.3]. It is noted that such codes are updated from time to time, and it is important to use the latest edition for such design work that shall be approved by authorities.

The special features of welded structures are the heating and cooling during the weld process, and the combined use of weld and base materials. They result in:

- Material phases with other micro-structural properties in the weld zones than in the bulk of the structure.

- Residual stresses in the neighbourhood of the welds, meaning that it must be assumed that the structure even in a load free condition has residual stresses of the order of the yield stress in these regions. This means that load amplitudes can be used without recourse to mean stresses, and assuming high R-values.

- Ubiquitous small cracks in the weld zones, of the order of some tenths of a millimetre.

With these features in mind the design rules follow the general principles for fatigue design, but in a formalized, flow-chart fashion.

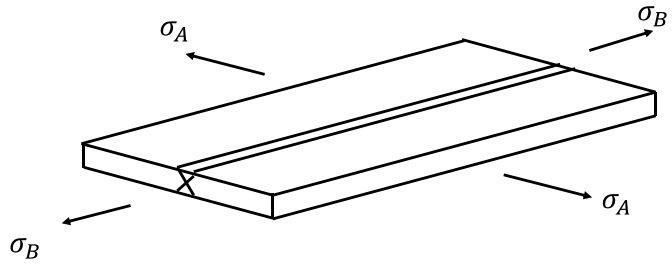
The description in 13.1.1 – 13.1.3 is a principal one, to demonstrate the similarity to stress based design. The standards themselves are more detailed and should be used in actual design work.

13.1.1 Relationships between stress level and life

Since the type of weld is so decisive for the design, S-N-curves are used for a large selection of typical welded components instead of pure materials. They have been obtained by large numbers of tests of different structural parts, as butt welds, spot welds, fillet welds etc, with different geometries and ways of loading. As for the general S-N-curves for pure materials, see Chapter 4, there is also an assumption of a standard plate thickness (in BSK 15 mm) in order to normalize for size of structures.

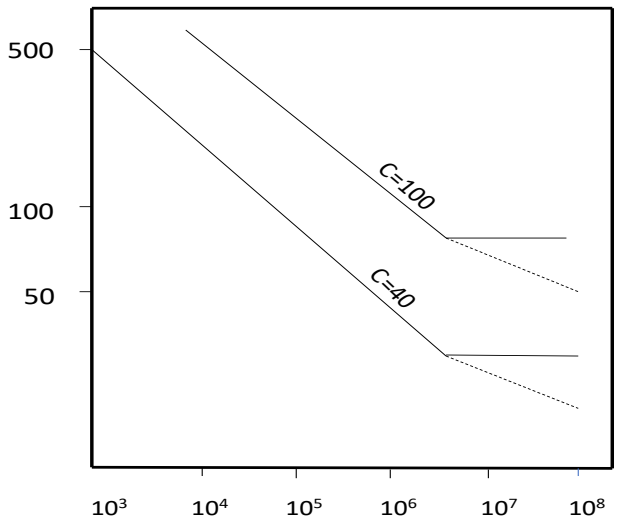
Groups of such structural parts with similar fatigue properties, i e relationships between nominal stress range and life, are said to belong to the same weld type clasification, C. This gives a set of design curves in a stress-number diagram, see Figure 13.1.b. The number of a weld class denotes the nominal stress range (in MPa) at 2 10^6 cycles of constant loading. Further, these curves are for 1% risk of failure, as compared to the S-N curves for general fatigue design, which denote a 50% risk of failure. The general assumption of straight lines in a log-log-diagram is maintained when establishing the curves.

The relationship between weld class, and type of weld and loading, is given in tables and diagrams, see the example in Figure 13.1.a. Here W denotes weld quality, and the different C:s denote loading in transverse and parallel directions.



Quality	C_B	C_A
WC	80	63
WB	100	90
WA	112	112

a)



b)

Figure 13.1 a) Example from BSK of weld types (nr 10 of around 80), and b) relationship between life and stress span for different weld type classifications C (only two of the set of 16 curves from C=35 MPa to C=195 MPa are shown, and the figure is not to scale).

The curves in Figure 13.1.b have a steeper slope, around -0.3, than S-N (Woehler) curves, which having slopes around -0.1. This is explained by the fact that welds contain small cracks (0.1-0.5 mm long), which means that the life is spent entirely in crack growth, and with virtually no initiation phase. Then the principal behaviour can be understood by applying the life calculation from Chapter 7 to fatigue design according to fracture mechanics, as follows.

The crack growth relationship is

$$\frac{da}{dN} = C(\Delta K)^m = C(\Delta\sigma f\sqrt{\pi a})^m \tag{13.1}$$

where C and m are material constants, a crack length, N cycles, and $\Delta\sigma$ the stress span. Assuming that the shape function f is a slowly varying function for realistic crack lengths in the weld geometry, separation and integration yields

$$N = \frac{1}{C(\Delta\sigma f\sqrt{\pi})^m}\left[\left(\frac{1}{a_o}\right)^{m/2-1} - \left(\frac{1}{a_{cr}}\right)^{m/2-1}\right]\frac{1}{(m/2-1)} = \frac{a_o}{C(\Delta\sigma f\sqrt{\pi a_o})^m}\left[1-\left(\frac{a_o}{a_{cr}}\right)^{m/2-1}\right]\frac{1}{(m/2-1)} \tag{13.2}$$

With a_o as a "standard weld crack length" always present, and a critical crack length a_{cr} at fracture such that $a_o/a_{cr} << 1$, making the bracket approximately equal to 1, it is seen that for any weld geometry (essentially defined by f and heat treatment etc) the relationship between stress span and life is

$$N \propto \Delta\sigma^{-m} \; or \; \Delta\sigma \propto N^{-1/m} \tag{13.3}$$

In most cases m= 3 – 4 making the slopes in the Figure 13.1 likely to be -0,25 - -0,33.

In some cases there are other stress concentrations than welds present. In those cases a K_f (fatigue notch factor) is calculated from the stress concentration factor K_t. This value is then transformed to an equivalent weld class C. In case of more than one stress concentration with mutual influence present, a lower weld class is defined from the different separate weld classes according to a given formula.

With the assumption of the presence of small cracks, and load spectra containing many amplitudes, it is also natural to abandon the concept of fatigue limit. Instead, the weld class curves are conventionally given smaller negative slopes for assumed longer life spans, as seen in Figure 13.1.b (dotted parts).

So, the curves for different weld classes in Figure13.1.b constitute the primary design design tool, corresponding to the S-N-curves including consideration of stress concentrations, welds or others, adjusted for K_f in Chapter 4 and 5. (As mentioned above and shown in Figure 11.3 the R-value is high due to residual stresses so the effect corresponding to the mean stress effect is inherent in the curves.)

13.1.2 Treatment of load spectra

The basic design curves have been obtained using constant stress ranges. In cases with load spectra, where varying stress ranges occur, this assumption is retained, but the influence of the variation has to be taken into account. This is made by introducing a parameter, κ , to

characterise the spectrum in a simple way. When κ is smaller than one (the value for constant amplitude load) the life becomes longer than if all cycles have the amplitude of the largest one. This is reflected in Table13.1, where data from Figure 13.1.b are modified to take the effect of κ into account. Only part of the table is shown; note that for $\kappa = 1$ the data correspond to the ones obtained from Figure13.1.b), for corresponding C- and n_t – values.

Two principal ways are accepted in BSK to obtain κ. One is simply to use experience from typical loadings on usual structures, as bridges, vehicles or lifting devices to assign a κ value.

Table 13.1 Part of table of modified stress levels for other κ-values than 1. The actual table covers several n_t-values and many C:s in intervals of 5 MPa.

κ	n_t	C= 35 195
1	10^3 . . . $5\ 10^6$	441 ..... 1410 25,8 154
5/6	10^3 10^8	
2/3		

The other way is to construct a load spectrum and compare it graphically with a set of type spectra representing different values of κ to find the κ giving the best fit.

The spectrum is defined in the following way. By making a level crossing or range pair count (note that this is equivalent to the rain-flow count, just omitting the mean values) for a representative sequence of cycles the spectrum is defined. The number of cycles at each level is multiplied by the relationship between the total intended life and the number of cycles, N, in the sequence. Then, conventionally, the 100 largest cycles are truncated. (They must not cause yield, however.) The resulting sequence consists of n_1 cycles for the highest stress level, n_2 cycles for the second highest one and so on until n_k cycles for the lowest stress level. The sum

$$\sum_{i=1}^{k} n_i = N - 100 \qquad (13.4)$$

Then, for each stress level S_j introduce the points

$$\log\left(\sum_{i=1}^{j} n_i\right)/\log(N - 100)\ ,\ \sigma_j/\sigma_1 \qquad (13.5)$$

into the diagram with type spectra for κ-values according to Figure 13.2 and make the best fit by eyesight to find the κ-value representing the spectrum, observing that the fit should be best for lower stress levels, to the right in the diagram. The κ-value obtained is then used together with C, N, and the maximum stress S_1 to define the equivalent stress value.

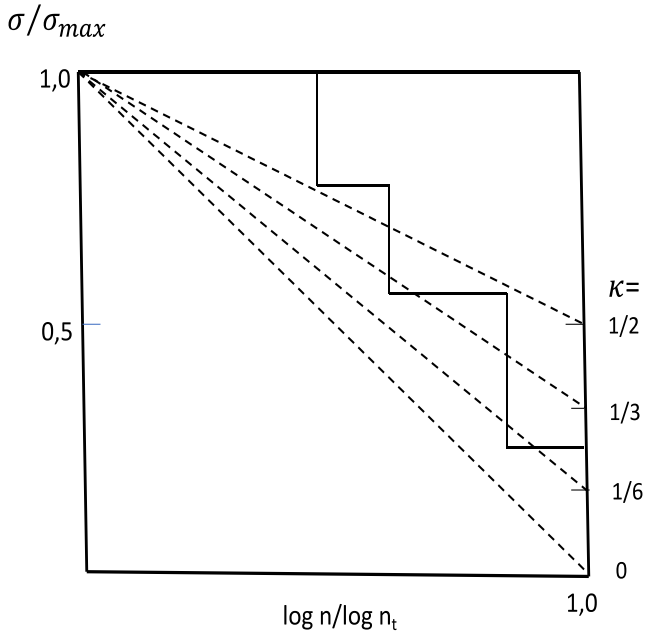


Figure 13.2 Example of determining a k-value from a load spectrum.

The method to use κ-values is synthetic and sometimes cumbersome. Therefore the normal way of calculating the damage by the Palmgren-Miner rule from the original load sequence and the curves in Figure 13.1.b may also be used.

13.1.3 Additional factors, and the design principle

As is seen so far, the general approach in fatigue design has been followed, with a stress- life relationship and a load spectrum analysis. The main difference is that *the virtually impossible definition of "stress concentrations" from welds have been included in a set of life relationships represented by weld classes*, and that provisions have been taken for *the presence of residual stresses when establishing the load spectrum*.

Now, the analogy with conventional fatigue analysis is completed by some further modifications.

- Surface effects are normally not included, since the presence of welds is thought to override any other surface defects; since cracks are present. (For parts of a structure with no welds, the standard gives a diagram for a material factor φ_m as a function of the yield stress and surface conditions.)

- Size effects. Here the thickness, t (in mm), of the (plate) material is the dimensioning factor, and a Weibull-like formula (c f Chapter 4)

  $\varphi_t = (t_0/t)^w$

  is employed, with t_0=15 mm as the reference thickness (as compared with approximately 10 mm before). The exponent w is taken as 0,10 for butt welds and 0,15 for fillet welds.

- In cases where R<0 (and residual stresses known) it is not considered acceptable to neglect the influence of mean stresses, and a factor

  $\varphi_R = (1 - R)/(1 - 0{,}7R)$

  is introduced.

Finally, a factor, γ_{mn}, taking care of the wish to have different safety levels of the design is introduced, see Table 13.2. In essence this factor is based on the dispersion found in the experiments made to find the S-N-curves in Figure 13.1.

Table 13.2 Factors γ_{mn}.

Consequence of failure	γ_{mn}	**Risk**
Negligible	1,0	10^{-2}
Less serious	1,1	10^{-3}
Serious	1,21	10^{-4}
Very serious	1,32	10^{-5}

Now, the relationships given in Table 13.1 for N, C, $\kappa$and allowable stress maximum stress span σ_{all} can be used in two ways. If the originally intended life is to be maintained the stress σ_1 has to be changed, and a new stress span

$$\sigma_{red} = \sigma_{all}\frac{\varphi_m\varphi_t\varphi_R}{\gamma_{mn}}$$

is calculated, to be compared with the original maximum stress σ_1.

If the stress level σ_1 shall be maintained a new N has to be found. First a design stress is calculated as

$$\sigma_{dim} = \sigma_1\frac{\gamma_{mn}}{\varphi_m\varphi_t\varphi_R}$$

and then from Table 13.1 σ_{dim}, κ and C give a new allowed N-value.

13.2 Bolted joints

Bolted joints are used in many structures, and there are handbooks and standards for the design and applied torque in different applications and for different types and material qualities of bolts. The common mechanical feature is the sharp stress concentration in the threads, which has to be considered.

Consider a configuration according to Figure 13.3. If a varying load P/2±P/2 is applied, and if the stiffness of the bolt is k_1 and the joint stiffness of the bolted parts is k_2 it is found that the variation of force in the bolt is f/2±f/2, where

$$f = \frac{Pk_1}{k_1 + k_2} \qquad (13.6)$$

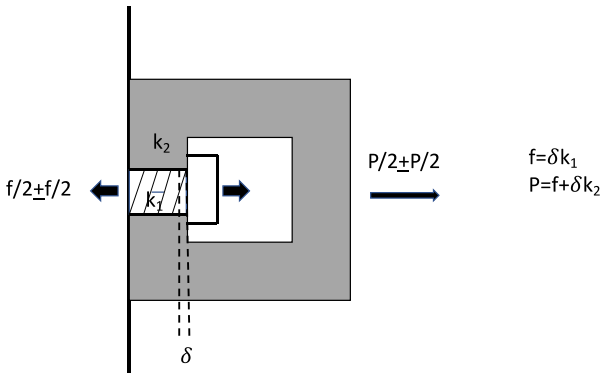


Figure 13.3 Illustration of a bolted joint.

It is seen that f<<P if $k_1<<k_2$, which is often the case. This relationship is valid until

$P \gg F(1+k_1/k_2) \approx F$, where F is the pre-load in the bolt from the torque. If P becomes larger than F, e g if the original torque is low or if F has diminished due to the tightening of other neighbouring bolts, the full variation of P is transferred to the bolt, possibly resulting in stresses giving fatigue damage in the thread roots. The conditions for such cracks and their arrest are described in Chapter 4.6 (Figure 4.6)

The relationship $\Theta=k_1/(k_1+k_2)$ is approximate and changes if the joint is eccentrically loaded, e g by bending of the joined parts. In handbooks, Θ-values are recommended as well as allowable stress amplitudes, f/2A, where A is the bolt area. They are given as functions of bolt material quality, normally in the range 35-50 MPa for lives of 10^7 cycles.

These stress levels can in principle be found by application of the following reasoning by (Gunn K W), see Figure 13.4

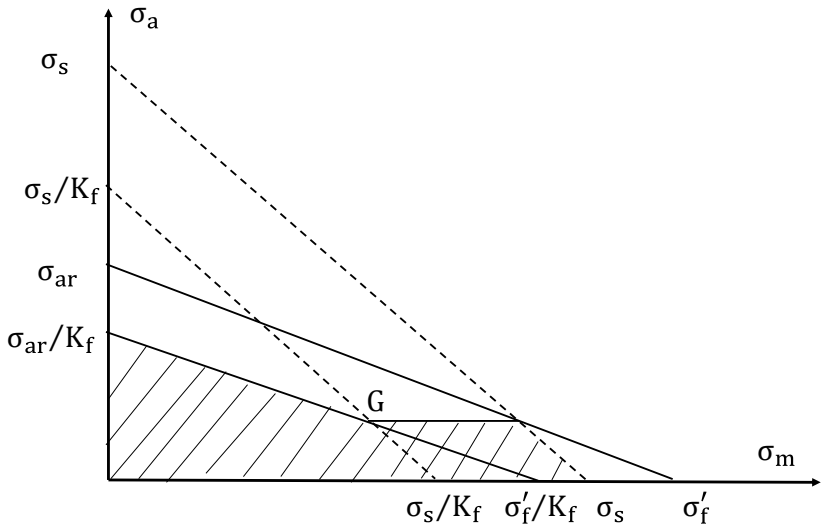


Figure13.4. Derivation of stress amplitude levels at high mean stresses.

The point G in the diagram corresponds to a situation, where the material is at the flow limit at maximum amplitude, and where the amplitude is not causing fatigue. It is then assumed that this amplitude is safe also for larger mean stresses, corresponding to high tightening loads, up to nominal stresses at the yield limit. The amplitude at point G is found to be

$$\sigma_a = \frac{\sigma_{ar}}{K_f}\left(\frac{\sigma_f' - \sigma_s}{\sigma_f' - \sigma_{ar}}\right) \qquad (13.7)$$

Rule of thumb tables are frequently encountered for common applications, e g when the lubrication is such that the torque can be related to the bolt force reasonably well. In Figure 13.5 such a table is shown for metric standard bolts (in Swedish).

Thread	Åtdragningsmoment (Nm)			Förspänningskraft (kN)		
	8.8	10.9	12.9	8.8	10.9	12.9
.	.	.	.	.	.	.
.	.	.	.	.	.	.
M8	23	32	39	17	24	29
M10	46	64	77	27	38	46
M12	80	110	140	40	56	67
M14	120	180	220	54	76	91
.	.	.	.	.	.	.
.	.	.	.	.	.	.

Figure 13.5. Example of rule of thumb table for design of bolted joints. (Mx is metric bolts with diameter x mm, 8.8, 10.9 and 12.9 are the material qualities, "Åtdragningsmoment" is torque, and "Förspänningskraft" is bolt force.) Sample values of diameters only from the original table.

13.3 Bearings

Bearings constitute a group of very well defined structural elements exerted to dynamic loads. Hence, standardized methods based on extensive experimental evidence are natural to use.

Experience led e g SKF to establish a simple formula for design,

$$L_{10} = (C/P)^m \qquad (13.8)$$

Here L_{10} is the life in million cycles, using the Weibull distribution to establish a 10 per cent risk for failure at that life. C is a characteristic parameter for the size and type of bearing, and m is 3 for ball bearings and 10/3 for roller bearings. P is an effective load calculated from a spectrum of variable loads P_i according to the Miner rule,

$$P = \left[\frac{\sum_i n_i P_i^m}{\sum_i n_i}\right]^{1/m} \qquad (13.9)$$

Later, more precise investigation has led to refined formulas, established from 1989-1990 in both in the SKF catalogue and in the ISO standard. In the SKF formulation the expression is

$L_{10aa}=a_{SKF}(\kappa,\eta_c P_u/P)_{brg}(C/P)^p$, (13.10)

where a_{SKF} is a parameter depending of the bearing, the lubrication conditions and the relationship between the effective load and a fatigue limit load. Hence, infinite life can be predicted for small loads. The SKF rolling bearings catalogue [13.4], for instance, should be consulted for actual design.

In the ISO standard 281[13.5] the formulation is

$L_{na}=a_1 a_2 a_3 (C/P)^p$ (a_1= 1 if n=10) (13.11)

Also in this case the basic reference, the standard, has to be consulted for actual design.

The formulas (13.10) and (13.11) summarize complex developments e g regarding lubrication leading to tables and diagrams for finding the parameters.

References

[13.1] Boverkets handbok om stålkonstruktioner, BSK 07, (in Swedish), Boverket (2007)

ISBN 978-91-85751-58-7 (www.boverket.se)

[13.2] Boiler and Pressure Vessel Code, ASME, (2013).

[13.3] Eurocode 3, EN 1993-1-9:2005.

[13.4] SKF rolling bearings catalogue. (2013)

[13.5] ISO 281:2007(E).

14 Probabilistic assessment of the risk for fatigue failure

14.1 General approach

The load on a structure and the fatigue strength of the structure are not exact values. They are stochastic variables belonging to statistical distributions. The general problem is to find the probability that the load exceeds the strength, see the examples in Figure 14.1

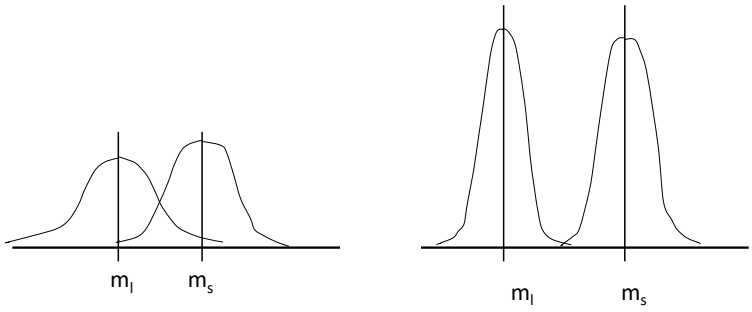


Figure 14.1 Two examples of load and strength distributions.

In the deterministic approach the structure should be safe as soon as the load, represented by m_l, is lower than the strength, represented by m_s. For the cases in the Figure, where the load is instead a variable with mean m_l and a standard deviation σ_l, and the strength is a variable with mean m_s and a standard deviation σ_s, the situation is different. If the means are close and the standard deviations, describing the widths of the distributions, are large, as in the example to the left, there is a considerable probability that the load exceeds the strength. If the mean values are instead wide apart in relationship to the standard deviations, the probability is obviously small, as in the example to the right. This example approaches the deterministic situation.

A practical example may be the following. Lorries contain a welded component in the suspension system, which may have a variable strength depending on the quality of the weld. The lorries are sold to users with varying demands, from heavy duty in a mining environment to transportation of light loads on high quality roads.

In the deterministic approach this is acknowledged by introducing a safety factor, i e to require a difference between the calculated or assessed values of m_l and m_s that is big enough to separate the "tails" of the distributions so much that the probability of failure is satisfactorily small. In cases where experience is considerable, the standard deviations, or realistic limits of variation, are well known and it is sufficient to use them to determine a safety factor. Important reasons for adopting a more detailed procedure are that safety factors may be overly conservative, that new materials or design principles are introduced, or that loads with less well-known load variations have to be considered.

The desired probability in such a case is found by deriving a distribution function $G_Z(z)$ of the stochastic variable z=y-x, that is the difference between stochastic variables load, x, and strength, y. If the distribution functions of these are $L_X(x)$ and $S_Y(y)$ respectively, the general expression is

$$g_z(z) = \int_{-\infty}^{\infty} l_x(x) s_Y(x+z) dx \qquad (14.1)$$

or

$$G_z(z) = \int_{-\infty}^{\infty} l_x(x) S_Y(x+z)dx \,, \qquad (14.2)$$

where g, l, and s represent the density functions, e g

$$g_Z(z) = G_Z'(z)$$

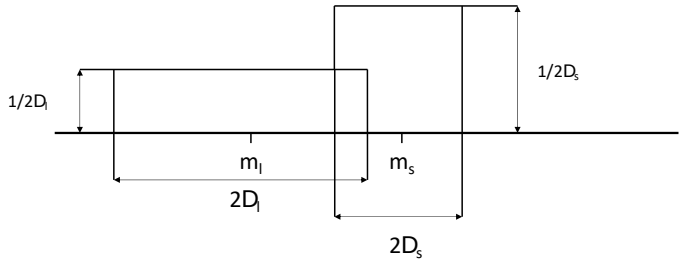
Since $G_Z(z)$ is the probability that the difference y - x is smaller than or equal to z, $G_Z(0)$ is the probability that x exceeds y, i e that failure occurs.

An illustrative example is when the two density functions are uniform, i e

$$l_X(x) = 1/2\Delta_l \quad for \quad m_l - \Delta_l \le x \le m_l + \Delta_l$$

and

$$s_Y(y) = 1/2\Delta_s \quad for \quad m_s - \Delta_s \le y \le m_s + \Delta_s$$



Putting z=0 then gives the desired probability

$$G_Z(0) = P_{fracture} = \frac{\left[m_l - m_s + (\Delta_l + \Delta_s)\right]^2}{8\Delta_l\Delta_s} \qquad \text{, if}$$

$$m_s - m_l \le \Delta_s + \Delta_l \quad and \quad m_s - m_l \ge \left|\Delta_s - \Delta_l\right|$$

As an example, set $\Delta_l=\Delta_s$ and a 10 per cent overlap, $m_l-m_s=-1.8\Delta_l$. This becomes

$$P_{fracture} = \frac{\left[-1.8\Delta_l + 2\Delta_l\right]^2}{8\Delta_l^{\,2}} = \frac{0.04}{8} = 0.005$$

This is smaller than what would be anticipated from just looking at a graph of the overlapping density functions, see Figure14.2.

Figure 14.2 Two uniform distributions for load and strength, for demonstration

A useful distribution is the Gaussian one, called the normal distribution. The reasons are that distributions of many physical parameters are approximately normal, and that this distribution has many attractive features, described in standard textbooks on statistics. It is conventionally written

$$\Phi\left(\frac{z-m}{\sigma}\right) \qquad (14.3)$$

The normalized form

$$\Phi(z) \tag{14.4}$$

and its density function

$$\phi(z) = \Phi'(z)$$

are tabulated in many books and available as functions in computers.

With normal distributions used for $L_X(x)$ and $S_Y(y)$ the integration to find $G_Z(z)$ is tedious, but the result is astonishingly simple and attractive,

$$G_z(z) = \Phi\left(\frac{z-(m_s-m_l)}{\sqrt{\sigma_s^2+\sigma_l^2}}\right) \tag{14.5}$$

The probability for fracture is found as before for z=0,

$$P_{fracture} = \Phi\left(\frac{m_l-m_s}{\sqrt{\sigma_l^2+\sigma_s^2}}\right) \tag{14.6}$$

This result is very useful. There are, however, two principal weaknesses.

One is that the distribution is defined for infinite arguments. This means that the "approximate normality", and usefulness, is limited to parameter values within a few standard deviations from the mean value. This is natural since the experiments proving the approximate normality must be limited in number, and hence they cannot be used to fit very small probabilities. Further, real physical entities have limited distributions of some form.

The other weakness, connected to the first one, is that the formula presumes that the standard deviations are known, while they must in reality be assessed from experience and experiments, and this has to be considered. There are, however, methods and rules of thumb for the assessment of standard deviations that are useful for practical purposes. See Chapter 14.3 below.

14.2 Resistance v s Load Reliability Analysis

A type of welded component in e g a car or railway bogie is studied. Its fatigue strength is evaluated by constant amplitude tests in a laboratory so that an S – N curve is obtained. In this way influences of size, geometry, surface treatment etc are included. Instrumented field tests are also performed for different service conditions in order to obtain representative load spectra.

It is assumed that the Palmgren-Miner linear damage rule can be used. Further, since fatigue is due to growth of weld cracks, it may be assumed that only amplitudes are relevant. Hence, range pair counting may be used to obtain the load spectra. (See also Chapter 8 on variable loading and Chapter 13 for details about welded structures.)

With the S – N curve as

$$N_i = C\Delta\sigma_i^{-k}, \tag{14.7}$$

where N_i is the mean life for a typical load cycle $\Delta\sigma_i$, the Palmgren- Miner rule at fracture becomes

$$\sum \frac{n_i}{N_i} = \sum \frac{n_i}{C\Delta\sigma_i^{-k}} = \frac{1}{C}\sum n_i \Delta\sigma_i^{k} = 1.$$

The parameter C can be seen as a stochastic parameter characterizing the strength in terms of the life of the component, from (14.7), since N_i for a certain load level is proportional to C.

Now the quantity

$$D = \sum n_i \Delta\sigma_i^{k} \tag{14.8}$$

is taken as a parameter characterizing the load for a certain number of cycles $\sum n_i$. D is also a stochastic variable expressing the variation in service load exerted on the components.

In a deterministic concept C=D at fracture and a safety factor can be obtained by choosing the loading time shorter by some factor. In order to pursue a probabilistic approach C and D have to be given some distribution functions. From (14.7) it is seen that since N_i is generally assumed to be log-normal the same can be assumed for C, i e log C has a normal distribution. Further, the mean, $M_{\log C}$, and the standard deviation s_c can be assessed from the laboratory tests. If it is also possible to fit log D_0 for some number of cycles N_0 to a normal distribution, with mean $M_{\log D}$ and standard deviation s_D, the probability for fracture for a multiple ν of this number is

$$P_{fracture} = \Phi\left(\frac{M_{\log \nu D} - M_{\log C}}{s_c^2 + s_D^2}\right) = \Phi\left(\frac{\log \nu + M_{\log D} - M_{\log C}}{s_c^2 + s_D^2}\right) \tag{14.9}$$

Note that the standard deviation s_D is not changed by the multiple ν of cycles. Hence a suitable risk level can be easily chosen from the last member of the formula.

The assumption of log-normal distributions is not necessary, but as is seen it is very handy. Arbitrary distributions can be used with the general formula for $G_Z(z)$ above. A useful choice in many cases is the Weibull distribution. For general choices the calculations are normally cumbersome, or "impossible". The use of log-normal distributions give a good hint about the risk even if the fit is not too good. It is noted again that reasonable assessment of the standard deviations requires many experiments, or a good experience of the type of structures and materials treated.

14.3 Assessment and use of standard deviations

This is a vast area in the theory of mathematical statistics. The following is, for convenience of reading, the description also given in Appendix 1. The term standard uncertainty used in the Appendix is replaced by standard deviation estimate to conform with the usual nomenclature.

Estimates of standard deviations can be made in essentially two ways. The first one, denoted as Type A in Appendix A.1, is by making a number of experiments and assessing the standard deviation in the usual way. Let q_j, j=1,n be n measurements of a stochastic variable q. Then

$$s^2(q) = \frac{1}{n-1}\sum\left(q_j - \bar{q}\right)^2 \text{ , where } \bar{q} = \frac{1}{n}\sum q_j \qquad (14.10)$$

The standard deviation estimate is s(q). When n is limited, the estimate should be corrected by the aid of the Student t distribution, tabulated in standard textbooks and with an excerpt given in Appendix 1. For practical purposes, and in particular when combined with several other estimates, it is considered that s(q) can be used for n>10.

The other way, leading to what is called a Type B estimate is when no experiments are available, which is a very common situation. This is to some extent an art. One can use:

- professional judgement and experience
- measurement data from similar situations
- manufacturer´s specifications
- data provided in calibration certificates
- values given in handbooks

If no standard deviation value is given, a rectangular distribution with width 2d is assumed covering what is judged to be the total dispersion. Then the estimate of s is $s = \frac{d}{\sqrt{3}}$. Type B estimates are assumed to be based on a large number of experiments performed by others or being assessed to be "on the safe side". Hence no correction for a limited number of samples is used.

In many cases, as in fatigue, one is interested in the uncertainty of an output quantity Y depending on several input quantities X_i in a functional relationship

$$Y=f(X_1, X_2, \ldots, X_n), \qquad (14.11)$$

For the estimates of the parameters we write

$$y=f(x_1, x_2, \ldots, x_n) \qquad (14.12)$$

Now, if the X_i are uncorrelated (i e independent of each other, a common situation in fatigue), the standard deviation estimate of y is

$$s_y^2 = \sum_{i=1}^{n} c_i^2 s_{x_i}^2 \qquad (14.13)$$

where s_{x_i} are the standard deviation estimates of x_i and

$$c_i = \frac{\partial f}{\partial x_i} = \frac{\partial f}{\partial X_i}$$

It should be noted here that the combination of several distributions tends to make the result more and more "normal-like". Hence it is thought acceptable to use rectangular estimates and still use the formula for a normal distribution in the end.

An important special case is when

$$Y = C \prod X_i^{p_i} \tag{14.14}$$

Then with $w_y=s_y/abs(y)$ and $w_{x_i} = s_{x_i} / abs(x_i)$, i e w are percentages,

$$w_y^2 = \sum p_i^2 w_{x_i}^2. \tag{14.15}$$

If there are quantities X_i and X_j that are correlated with correlation coefficients $C_{i,j}$ expression (14.13) is extended into

$$s_y^2 = \sum_{i=1}^{n} c_i^2 s_{x_i}^2 + 2 \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} c_i c_k \, s_{x_i,x_k} , \tag{14.16}$$

where

$$s_{x_i,x_k} = s_{x_i} s_{x_k} r_{x_i,x_k} \quad i \neq k \tag{14.17}$$

and

$$\left| r_{x_i,x_k} \right| \leq 1$$

is the correlation coefficient between x_i and x_k.

Example. Estimate of the risk for fatigue failure at the endurance limit and constant amplitude loading.

From Chapter 5 the mean of the strength is found as

$\sigma_{e,red} = \frac{\sigma_e \lambda \delta}{K_r K_f}$, where corrections are made for size, surface and geometry with factors taken from general purpose tables and thus having uncertainties for specific cases.

Putting

$Y=\sigma_{e,red}$, $\sigma_e=X_1$, $\lambda=X_2$, $\delta=X_3$, $1/K_r=X_4$, and $K_f=X_5$ in (14.11) the percentage standard deviation estimate w_y is found from

$$w_y^2 = \sum w_{x_i}^2$$

Assume now, as an example, that the mean of σ_e is 400 MPa and that the handbook value of the standard deviation is 20 MPa, i e w_{x1} =0,05. Assume further that only $1/K_r$ and K_f are relevant in the actual case and that they are both estimated from rectangular distributions to be max $\pm$ 5 per cent, giving

$$w_{x_4} = w_{x_5} = 0{,}05/\sqrt{3}$$

Finally

$$w_y = 0{,}05\sqrt{1+1/3+1/3} = 0{,}065$$

s_y= 0,065 400= 26 MPa.

The load is assumed to have a normal distribution with a mean of 300 MPa and a standard deviation of 30 MPa. Then the probability for fatigue failure is approximately

$$P = \Phi\left(\frac{300 - 400}{26^2 + 30^2}\right) = \Phi(-2{,}5) = 0{,}006$$

The example is a bit artificial in assuming normal distributions for the uncertainties of the correction factors with values using the "tails" of the distributions. In a simplified analysis one normally compares the values 300 and 400 giving a "safety factor" of 1,33 to be compared with some norm or standard requirement. The value of a detailed analysis is more evident when one is forced to accept e g larger uncertainties in the load.

Appendices

A1 Uncertainty of measurements

The following is a summary of the principles applied in assessments of measurement quality in accreditation of laboratories [A1.1]. They may be of use also for general assessment of uncertainties in parameters measured in experiments, e g in fatigue design.

All quantities measured in experiments are stochastic variables, and the standard deviations of their distributions are the usual measures of the uncertainties in the measured values. The concept *standard uncertainty* is an estimate of the standard deviation of the distribution.

(It may be useful to mention that uncertainties have global and local parts, defined by the concepts *reproducibility* and *repeatability*. In one laboratory one can repeat a measurement many times with a high precision, the local uncertainty is small, i e the *repeatability* is good; still the values found may be severely wrong due to systematic errors. By making measurements in a representative number of laboratories these errors are included in the *reproducibility*. Therefore, calibrations with standards or other comparisons common for all laboratories concerned are important to obtain the reproducibility, or the total uncertainty. *Comparisons between a number of laboratories in the comparatively simple measurement of tensile strength have revealed surprisingly low reproducibility*. This subject is not pursued further here, but the note serves as a reminder for those assessing fatigue parameters from experiments performed in only one laboratory, and perhaps just sporadically.)

Generally one is interested in the uncertainty of an output quantity Y depending on several input quantities X_i in a functional relationship

$Y=f(X_1, X_2, ….., X_n)$,

where f may be a complex relationship, described as a formula, a computer algorithm etc.

(A simple example is when the stress amplitude, σ, in a circular bar in tension is calculated from the diameter D of the bar, measured by callipers, and the force amplitude, F, measured by a load cell. Then

$Y=\sigma=4F/\pi D^2=f(F,D)$,

where F and D have uncertainties to be assessed.)

For the estimate of the standard uncertainty of Y the standard uncertainties of X_i, i=1,n, have to be assessed.

They are generally classified in two ways, as Type A and Type B standard uncertainties.

A Type A standard uncertainty is found by making a number of experiments and assessing the standard deviation in the usual way. Let q be one of the X_i:s with n measured values q_j. Then

$$s^2(q) = \frac{1}{n-1}\sum\left(q_j - \bar{q}\right)^2 \qquad \text{(A1.1)}$$

$$s^2(\bar{q}) = \frac{s^2(q)}{n} \qquad \text{(A1.2)}$$

and the standard uncertainty u for the estimate of the mean $\overline{q}$ of q to be used is

$u = u(\overline{q}) = s(\overline{q})$. When n<10 the estimate is unreliable.

A Type B standard uncertainty is an estimate that has to be made when no experiments are available, which is very common. This is an "art" depending on professional judgement. One can use:

- experience and literature studies
- previous measurement data from similar situation
- manufacturer´s specifications
- data provided in calibration certificates
- uncertainties given in handbooks

If no distribution is available, or if the uncertainty is given as a $\pm$ d value, a rectangular distribution is assumed covering a reasonable dispersion. Then

$$u = \frac{d}{\sqrt{3}}. \qquad \text{(A1.3)}$$

Now, if the X_i:s are uncorrelated (i e independent of each other, which is a common situation), the standard uncertainty of Y is

$$u_y^2 = \sum c_i^2 u_{x_i}^2 \qquad \text{(A1.4)}$$

and where u_{x_i} are the standard uncertainties of X_i and $c_i = \frac{\partial f}{\partial X_i}$.

(If some of the X_i are correlated, as for example when several lengths are measured with the same callipers, the formula (A1.4) has to be replaced by

$$u_y^2 = \sum_{i=1}^{n} c_i^2 u_{x_i}^2 + 2\sum_{i=1}^{n-1}\sum_{k=i+1}^{n} r_{x_i x_k}(c_i u_{x_i})(c_k u_{x_k}) \qquad \text{(A1.5)}$$

where $r_{x_i x_k}$ are the correlation coefficients between X_i and X_k. They may be difficult to estimate but could be set to unity for conservative results.)

An important special case is when

$$Y = C\prod X_i^{p_i} \qquad \text{(A1.6)}$$

Then with $w_y = u_y/\mathrm{abs}(Y)$ and $w_{x_i} = u_{x_i}/abs(X_i)$, i e w are percentages

$$w_y^2 = \sum p_i^2 w_{x_i}^2 \qquad \text{(A1.7)}$$

Normally one wants to use u_y as a measure for the uncertainty of Y with a certain coverage factor k. For a normal distribution a common coverage of 95% means two (1,96) standard deviations, and hence one writes that the 95 % coverage of Y is

$Y \pm 2u_y$ (A1.8)

This reasoning presumes that u_y is determined with an infinite number of samples (degrees of freedom). Otherwise k must be chosen larger for the 95% coverage interval (and correspondingly for other coverage levels).

An effective number of degrees of freedom can be found from

$$\nu_{eff} = \frac{u_y^4}{\sum \frac{\left(c_i u_{x_i}\right)^4}{\nu_i}} \qquad \text{(A1.9)}$$

where ν_i are the degrees of freedom for the estimates of u_{x_i}. These may be difficult to assess for Type B uncertainties. When the rectangular distribution is assumed to cover all reasonable values for a variable x_i, ν_i is taken to be infinity. A table for 95% coverage factors (Student´s t) is given, as

ν_{eff}	5	10	20	50	¥
k	2,65	2,28	2,13	2,05	2,00

Reference:

[A1.1] EA-4/02 “Expression of the Uncertainty of Measurement in Calibration”, European Accreditation (2013)

A2 Linear regression determination of parameters in approximate analytical expressions from experimental values

Many of the expressions used in fatigue design are of the form

$$U = A\, V^{B}, \qquad \text{(A.2.1)}$$

where A and B are parameters to be found from sets of experiments measuring U and V, and where U and V are stress, strain, life etc. The determination is made by first rewriting the formula in logarithmic form

log U = log A + B log V, or in the form Y= a + b x ,

demonstrating a linear relationship.

The most common way is then to use linear regression in order to get a fit of a and b to an irregular set of n data pairs x_i, Y_i; i =1, n, as described in textbooks. The model is that the, slightly uncertain, values Y are measured from precisely known x-values viz

$$Y = a + b\, x + \varepsilon \qquad \text{(A2.2)}$$

where ε is a dispersion in Y, which is assumed to have a normal distribution. (In a real situation both x and Y have some dispersion, x mainly from measurement errors. In fatigue the measurement errors of load, strain, etc are much smaller than the dispersion in material properties, and hence Y can be assumed to carry the whole dispersion term.)
A principal problem is that it is in general not at all sure that the mathematical model for the behaviour is physically sound. It is just an expression evaluated by curve fitting in order to obtain a suitable "toolbox" for analysis. In the literature on statistics various methods are described for testing the assumption of the linearity of the model and the normality of the dispersion.

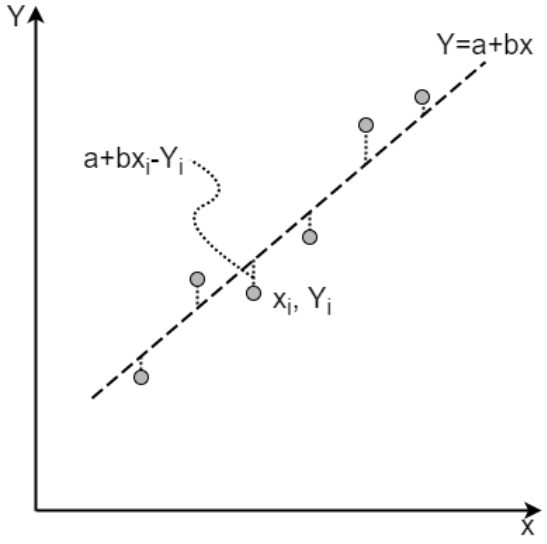


Figure A2.1 Sketch of data points and linear regression line.

So, with linear regression the least squares error to be minimised is, see Figure A2.1,

$$E = \sum [\, Y_i - (\, \mathbf{a} + \mathbf{b}\, x_i\,)]^2 \qquad \text{(A2.3)}$$

where **a** and **b** are the best fit coefficients to be found and Y_i (x_i) is the set of measured data. Minimising for **a** and **b** by partial derivatives results in

$$\mathbf{b} = \frac{\sum Y_i (x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \quad \text{and} \quad \mathbf{a} = \bar{Y} - b\bar{x} \qquad \text{(A2.4)}$$

where $\bar{x}$ and $\bar{Y}$ are the means of x_i and Y_i.

Normally the **a**- and **b**-values obtained from measurements and the model calculations are tabled as material data in handbooks without further discussion about their uncertainty, or whether material data in fact are stochastic variables having a dispersion.
Some useful indications about the dispersion ε can be found from the residuals

$$e_i = Y_i - (\mathbf{a} - \mathbf{b}\; x_i)$$

by calculating the variance

$$s^2 = \frac{1}{n-2} \sum e_i^2 \quad \text{and} \qquad \text{(A2.5)}$$

$$\text{Var}\,\mathbf{a} \approx s^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2} \right] \text{ and Var } \mathbf{b} \approx s^2 \left[\frac{1}{\sum (x_i - \bar{x})^2} \right] \qquad \text{(A2.6)}$$

i e approximate measures of the squared standard deviations of **a** and **b**.

Finally, a value Y calculated from an x-value and the parameters **a** and **b** has the approximate uncertainty

$$s = t_{n-2,a}\, M(x), \qquad \text{(A2.7)}$$

where $t_{n-2,a}$ is the Student t-distribution value for n-2 degrees of freedom and the percentile a desired, often a=95 %. Here

$$M(x) = \sqrt{\frac{(x - \bar{x})^2}{\sum (x_i - x)^2} + \frac{1}{n} + 1} \qquad \text{(A2.8)}$$

which for practical purposes can be taken to be 1 for reasonably large n. Note that the term depending on x gets smaller when many x_i are far from the mean, i e *it is favourable to have*

most measurements at each end of the interval to be modelled! This is seen also in the expressions for Var **a** and Var **b**.

A3 Stress intensity factors

Three typical examples of handbook results are shown in order to illustrate how results for various geometries are presented. Some of the widely used handbooks are [A3.1], [A3.2]. Ample results for many geometries may be found on the Internet.

The first geometry, Figure A3.1, is the extension to sheets of limited size from the infinite one discussed in Chapter 7. It is seen that both a graph and a formula are given, and that the formula is based on the generic expression for the infinite sheet

$$K_I = \sigma_0\sqrt{\pi a} \qquad \text{(A3.1)}$$

multiplied by a shape function f(a/W, h/W). For many realistic cases a is much smaller than both h and W, and then f is close to unity.

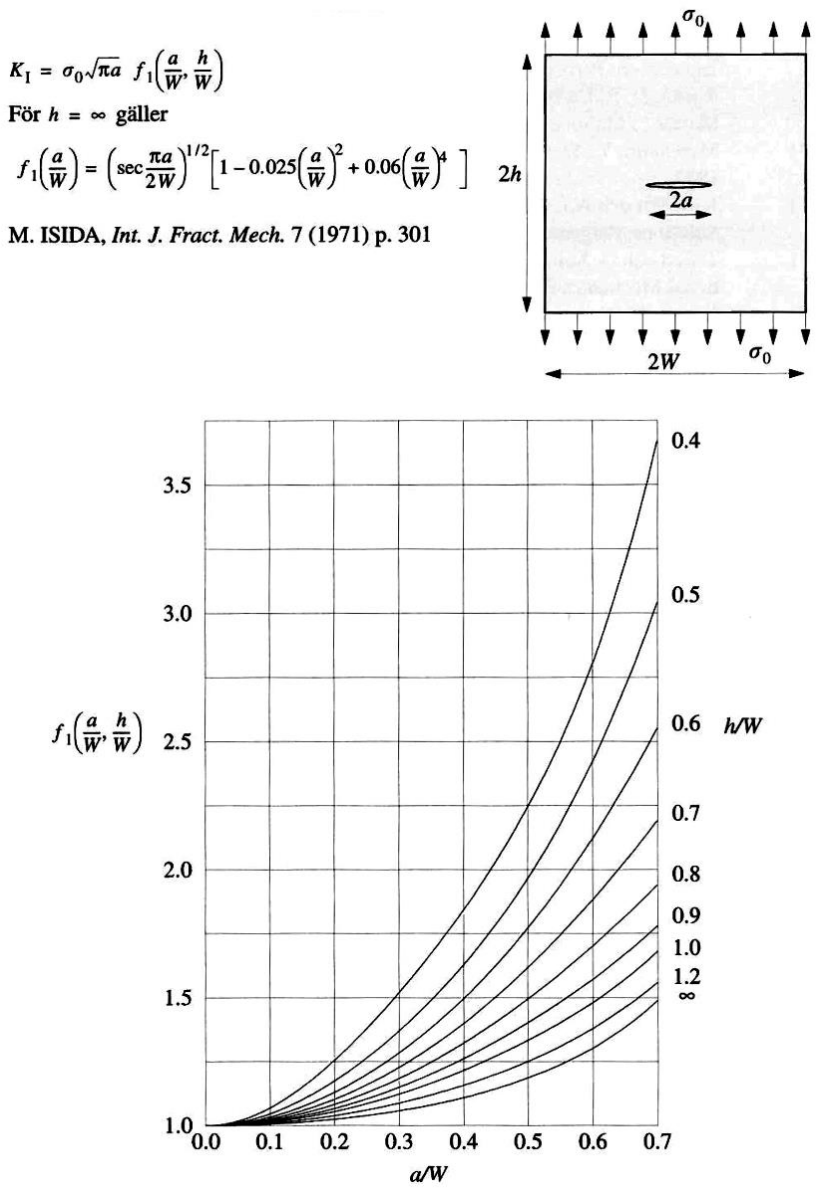


Figure A3.1 Stress intensity factors for a sheet with limited size.

A general and useful result is that the shape factor for small edge cracks is 1,12. The increase can be understood qualitatively by imagining the sheet in Figure A3.1 cut into two vertically giving two edge cracks. Then the constraints from the material removed become smaller and it is natural to imagine that the stress intensity increases.

This appears also in the second type example in Figure A3.2, showing stress intensity factors for a crack emanating from a circular hole.

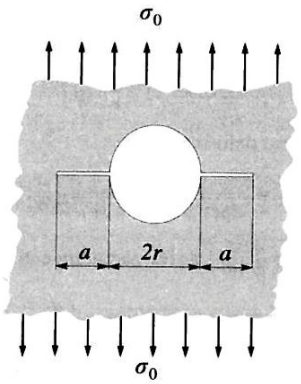
From Chapter 4.6 it is known that the stress at the edge of a circular hole is $3\sigma_0$. For small edge cracks, multiplying with the factor 1,12 gives 3,36, the value at the top left of the diagram in Figure A3.2. This factor changes gradually and when a becomes significantly larger than the hole radius, s>0,5, the stress intensity factor approaches the one for a crack in an infinite sheet (A3.1), i e presence of the hole does not influence the result very much. The result, 1,12, for small edge cracks can be generalized to notches of other shapes than circular holes.

$$K_I = \sigma_0\sqrt{\pi a}\ f_3(s)$$

$$f_3(s) = 0{,}5(3-s)[1 + 1{,}243(1-s)^3];0 < s < 1$$

$$s = \frac{a}{r+a}$$

O.L. BOWIE, *J. Math. and Phys.* 35 (1956)
H. TADA, P.C. PARIS, G. R. IRWIN,
The Stress Analysis of Cracks Handbook,
Del Research Corp, 1973



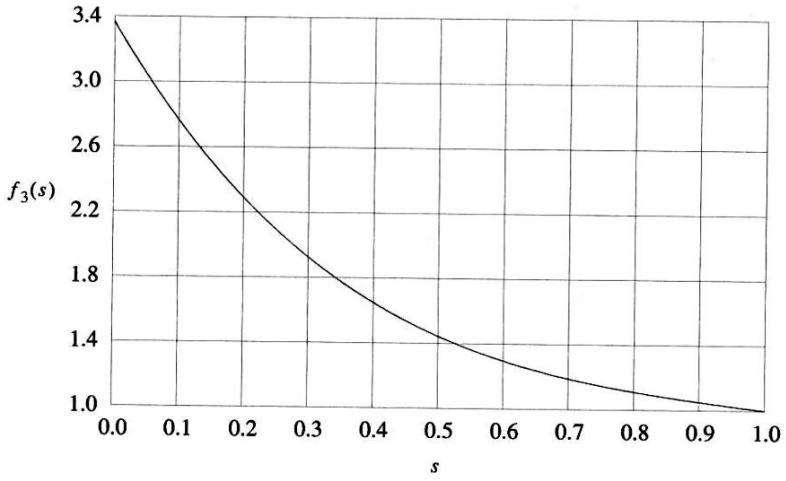


Figure A3.2 Stress intensity factors for a crack at a circular hole.

The third type example, in Figure A3.3, shows stress intensity factors for one of the standardized test pieces for fracture mechanics testing. It illustrates a situation where the shape function changes significantly for crack length intervals of importance. The

complex shape of the test specimen is chosen in order to save material when test pieces have to be extracted from parts of structures, e g pipes, where the thickness v s radius means limitations. In addition, the specimens have to be as big as possible in order that the conditions for linear fracture mechanics shall be satisfied.

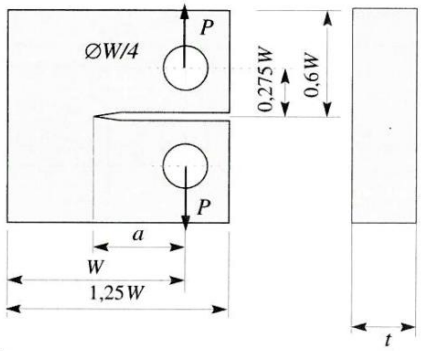
The three examples are thought to be illustrative and give general guidance.

Finally, it may be mentioned that for small internal, penny-shaped cracks, the shape factor is $2/\pi$ in comparison with cracks in plane sheets. A factor smaller than one is to expected, since the penny-shape means more constraints on the displacement than the geometry with a plane sheet.

$$K_I = \frac{P}{tW^{1/2}} f_{10}\left(\frac{a}{W}\right)$$

$$f_{10}\left(\frac{a}{W}\right) = \frac{\left(2+\frac{a}{W}\right)}{\left(1-\frac{a}{W}\right)^{3/2}}\left[0{,}886 + 4{,}64\left(\frac{a}{W}\right) - 13{,}32\left(\frac{a}{W}\right)^2 + 14{,}72\left(\frac{a}{W}\right)^3 - 5{,}6\left(\frac{a}{W}\right)^4\right]$$

För standardprovstaven rekommenderas
$t = W/2$ $a = W/2$
ASTM E 399,
American Society for Testing and Materials



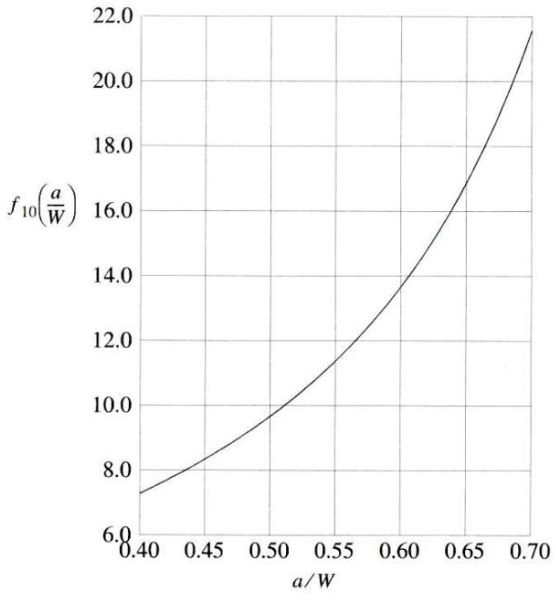


Figure A3.3 Stress intensity factors for a typical test specimen.

References.

[A3.1] Tada, H., Paris, P.C., and Irwin, G.R. “The Stress Analysis of Cracks Handbook.” (2nd ed.) Parsi Productions, Inc., St Louis, 1985

[A3.2] Murakami, Y. “Stress Intensity Factors Handbook” Pergamon Press, New York, 1987