REPORT LR-360

LECTURE NOTES ON FATIGUE, STATIC TENSILE STRENGTH AND STRESS CORROSION OF AIRCRAFT MATERIALS AND STRUCTURES

part I: text

J. Schijve
ERRATUM

Report LR-360

It appears to be unavoidable to print a report of 195 pages (part I) and 117 pages (part II) without several mistakes and printing errors. A careful study by student Arnt Offringa indicated several ones. Those which might be disturbing are listed below.

J. Schijve
14 January 1983

Part I

Page 17 Eq. (2.19) In the numerator, replace (2-2b) by (a-2b).
In the denominator, replace (a-b) by (a-b)^2.

Page 31 Second part of Eq. (3.4) should read:
\[ K_\varepsilon = \varepsilon_{\text{peak}} / \varepsilon_{\text{nom}} (\geq K_t). \]

Page 60 Equation at the bottom, right hand part should read:
\[ D_{eq} = \sqrt{1.13} F_0. \]

Page 72 6th line from the bottom: replace "accept" by "except".

Page 83 Line 17: replace "and" by "at".

Page 96 6th line from the bottom: If the fatigue load is high ...

Page 117 In both line 8 and line 15 from the bottom: Replace "figure 8.8" by "figure 8.9".

Page 118 4th line from the bottom: replace "8.10" by "8.11".
14th line from the bottom: replace "is" by "in".

Part II

Figure 5.13 Replace \(\frac{1}{2.5}\) by 2.5

Figure 8.2 In Figure 8.2d only: Replace 200 and 400 along vertical axis by 400 and 800 respectively.
LECTURE NOTES ON FATIGUE, STATIC TENSILE STRENGTH AND STRESS CORROSION OF AIRCRAFT MATERIALS AND STRUCTURES

part I: text

J. Schijve

Delft - The Netherlands

October 1982
FATIGUE, STATIC TENSILE STRENGTH AND STRESS CORROSION OF AIRCRAFT MATERIALS AND STRUCTURES

by professor dr.ir. J. Schijve

Preface

The present course book was primarily written up as lecture notes for a summer course, given at the Technical University of Bandung (ITB) in August 1982. The summer course was part of a cooperation between ITB and the Department of Aerospace Engineering of the Delft University of Technology. The cooperation is a contribution to the TTA-79 Technical Assistance Project as agreed between the Governments of Indonesia and The Netherlands.

The subjects covered by the present course book are very much similar to those of the course "Aircraft Materials II", which is presented in the 4th year of the Delft curriculum. In the latter course some subjects are treated in more detail. Until now the Delft course was covered by a series of lecture notes ("volgnummers") in the Dutch language. Several notes were rather condensed and others were overlapping as a result of historical developments. For these reasons it was thought useful to print the lecture notes of the ITB summer course as a course book, which then can be used in Delft also. Moreover it may also be useful to other people involved in aircraft materials and structures, with the advantage that the English language makes it accessible to a larger group.

The subjects of the course are listed in the contents. The main emphasis is on fatigue (chapter 7 to 12), while stress corrosion (chapter 6) and static failure in tension (chapter 5) are covered each in a separate chapter. Stress concentrations at notches, stress intensities at crack tips and residual stresses are highly relevant to all topics. For that reason these concepts are treated first (chapters 2, 3 and 4).

The prerequisites for the present course can easily be indicated in the Delft curriculum, viz.: Applied Mechanics II, Aircraft Materials I and Aircraft Structures I,
all given in the 2nd year. In general terms, some basic knowledge should be available on: Stress and strain: Elastic stress-strain relations, inhomogeneous stress distribution, difference between plane strain and plane stress, statically indeterminate analysis.

Material concepts: Properties derived from a tensile test, grain boundaries, slipbands, anisotropy, dislocations, heat treatment of alloys, fibrous structure of materials, inclusions.

Aircraft structures: Function of the various components of an aircraft structure, joints.

For convenience the text and the figures have been printed in separate volumes. The author is grateful to Mrs. M.M.J. Grob-Bennett, Mrs. M. Schillemans-van Tuyl, Mr. J.A. Jongenelen and Mr. W. Spee, who took care of typewriting and drawing of figures in a very short time, and to student Wijnand Moonen for careful reading the manuscript and corrections.
# CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>FAILURE MODES IN AIRCRAFT STRUCTURES</td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Failure modes</td>
<td>3</td>
</tr>
<tr>
<td>1.3</td>
<td>Failure prevention</td>
<td>5</td>
</tr>
<tr>
<td>1.4</td>
<td>Failure analysis</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>Subjects of the present course</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>STRESS CONCENTRATIONS AT NOTCHES</td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>10</td>
</tr>
<tr>
<td>2.2</td>
<td>Definition of $K_t$</td>
<td>11</td>
</tr>
<tr>
<td>2.3</td>
<td>Analytical calculations</td>
<td>12</td>
</tr>
<tr>
<td>2.4</td>
<td>The effect of the geometry (shape) on $K_t$</td>
<td>18</td>
</tr>
<tr>
<td>2.5</td>
<td>Superposition of notches and similarity of notches</td>
<td>21</td>
</tr>
<tr>
<td>2.6</td>
<td>Methods for determination of stress concentrations</td>
<td>22</td>
</tr>
<tr>
<td>2.7</td>
<td>References</td>
<td>24</td>
</tr>
<tr>
<td>3</td>
<td>PLASTIC DEFORMATION AND RESIDUAL STRESS</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>25</td>
</tr>
<tr>
<td>3.2</td>
<td>Different sources of residual stress systems</td>
<td>26</td>
</tr>
<tr>
<td>3.3</td>
<td>Calculations on residual stresses at notches</td>
<td>30</td>
</tr>
<tr>
<td>3.4</td>
<td>Some additional remarks</td>
<td>32</td>
</tr>
<tr>
<td>4</td>
<td>STRESS INTENSITY FACTORS OF CRACKS</td>
<td></td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>34</td>
</tr>
<tr>
<td>4.2</td>
<td>Different types of cracks and crack opening modes</td>
<td>35</td>
</tr>
<tr>
<td>4.3</td>
<td>Definition of the stress intensity factor $K$</td>
<td>35</td>
</tr>
<tr>
<td>4.4</td>
<td>Examples of stress intensity factors</td>
<td>39</td>
</tr>
<tr>
<td>4.5</td>
<td>$K$-factors by superposition</td>
<td>43</td>
</tr>
<tr>
<td>4.6</td>
<td>Crack opening and the state of stress</td>
<td>45</td>
</tr>
<tr>
<td>4.7</td>
<td>Energy considerations and the Griffith criterion</td>
<td>47</td>
</tr>
<tr>
<td>4.8</td>
<td>Cracks with a curved crack front</td>
<td>49</td>
</tr>
<tr>
<td>4.9</td>
<td>Determinations of $K$-values</td>
<td>51</td>
</tr>
<tr>
<td>4.10</td>
<td>Crack tip plasticity</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>55</td>
</tr>
<tr>
<td>5</td>
<td>FAILURE IN TENSION</td>
<td></td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>56</td>
</tr>
<tr>
<td>5.2</td>
<td>The tensile strength of a material</td>
<td>57</td>
</tr>
<tr>
<td>5.3</td>
<td>The hardness of a material</td>
<td>61</td>
</tr>
<tr>
<td>5.4</td>
<td>The static tensile strength of a notched element</td>
<td>62</td>
</tr>
<tr>
<td>5.5</td>
<td>The fracture toughness $K_{IC}$</td>
<td>64</td>
</tr>
<tr>
<td>5.6</td>
<td>Cracks in sheets, behaviour under static loading</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>73</td>
</tr>
<tr>
<td>6</td>
<td>STRESS CORROSION</td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>74</td>
</tr>
<tr>
<td>6.2</td>
<td>The stress corrosion process</td>
<td>74</td>
</tr>
<tr>
<td>6.3</td>
<td>Stress corrosion tests, variables and purpose of tests</td>
<td>77</td>
</tr>
<tr>
<td>6.4</td>
<td>How to prevent stress corrosion in aircraft structures</td>
<td>83</td>
</tr>
<tr>
<td>6.5</td>
<td>Hydrogen embrittlement</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>86</td>
</tr>
</tbody>
</table>
12 FATIGUE OF AIRCRAFT STRUCTURES

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.1 Introduction</td>
<td>178</td>
</tr>
<tr>
<td>12.2 Design aspects</td>
<td>178</td>
</tr>
<tr>
<td>12.3 Predictions and calculations</td>
<td>183</td>
</tr>
<tr>
<td>12.4 Fatigue testing methods</td>
<td>185</td>
</tr>
<tr>
<td>12.5 Full-scale tests and scatter factors</td>
<td>189</td>
</tr>
<tr>
<td>12.6 Maintenance and inspection</td>
<td>191</td>
</tr>
<tr>
<td>References</td>
<td>193</td>
</tr>
<tr>
<td>Tabel 12.1</td>
<td>196</td>
</tr>
</tbody>
</table>
Chapter 1  FAILURE MODES IN AIRCRAFT STRUCTURES

Contents:
1.1 Introduction
1.2 Failure modes
1.3 Failure prevention
1.4 Failure analysis
1.5 Subjects of the present course

1.1 INTRODUCTION

The development of an aircraft is characterized by different phases. A global picture is given below.

Fig. 1.1

Aircraft design includes a large variety of topics, such as
- aerodynamic shapes and dimensions (windtunneling)
- aircraft performance and stability
- various systems (engines, undercarriage, control systems, electronics, hydraulics, etc.)
- structural lay out and material selection.

This list could be made much longer and certainly more detailed.

Aircraft design is a well developed engineering science. However, it cannot be done without tests. Testing is a complementary activity to analysis, calculations, predictions and estimates. This is generally recognized for windtunnel tests. It would be very unwise to develop an aircraft without a program of model tests in a wind tunnel. It may turn out that aerodynamic properties differ from expectations. Similarly, uncertainties about static strength and fatigue behaviour in the design phase can be covered by tests on specimens or components.
After the design has been fixed production can start. Some of the first aircraft will be used for testing. Nowadays modern aircraft development requires:
- flight tests to check the aircraft performance and stability
- static testing of a full scale structure
- fatigue testing of a full scale structure
- functional tests of systems.

Flight tests require a fully equipped aircraft. For static tests and fatigue tests the aircraft structure without the systems installed will be sufficient. Functional tests can be done on the relevant parts of the system (e.g. landing gear retraction system, flap operation system, etc.)

Why is this expensive testing necessary? First it has to be done to obtain accurate and reliable information on the properties of the aircraft. Second, testing can reveal hidden problems which were overlooked. We have to be sure that the aircraft will be safe and economical.

Third, quite often testing is requested by airworthiness authorities.

The aircraft in service should hopefully stay in good condition for many years without excessive maintenance. However, it also may turn out that certain problems become apparent later. Noteworthy in this respect are fatigue, corrosion, wear, damage etc. It is possible that such problems could have been prevented by an improved aircraft design or by better production techniques. The responsibility of the aircraft design office is quite large, but the production quality is also significant. In Figure 1.1 the dotted lines indicate feed back information which should help to improve the design.

Structures and materials

If we now restrict the picture of Fig. 1.1 to structures and materials it becomes

<table>
<thead>
<tr>
<th>Aircraft design</th>
<th>Aircraft production</th>
<th>Aircraft in service</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structural lay out</td>
<td>Production quality</td>
<td>Inspection</td>
</tr>
<tr>
<td>Material selection</td>
<td>Inspection</td>
<td>Maintenance</td>
</tr>
<tr>
<td>Production techniques</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1.2
The aircraft structure is supposed to be satisfactory "if nothing happens" in service, no premature failures, no frequent replacements, no excessive maintenance. This can be achieved only if we know what might happen and if successful preventive action is taken in the design phase. In this chapter a survey of possible failure modes of aircraft structures and materials will be given first. This will be followed by a short discussion on preventive measures. Finally a brief introduction is given to the subjects of the subsequent chapters.

### 1.2 FAILURE MODES

A survey is given below

| (Quasi) Static Failure modes | - failure under tension | - ductile failure |
|                            | - failure under compression | - brittle failure |
|                            | - failure under shear | - different types of buckling |
|                            | | - bolt and rivet failure |
|                            | | - buckling of webs |
|                            | - excessive plastic deformation |

| Time-Dependent Failure modes | - corrosion | - general corrosion |
|                            | - stress corrosion | - fretting corrosion (fatigue) |
|                            | - fatigue | - creep |
|                            | - wear |

*Fig. 1.3*

The static failure modes are all subject of strength calculation during aircraft design. Some of them are simply covered by allowable stress levels (shear failure of rivets, hole deformation by bearing pressure). Other ones require more sophisticated stress analysis. This applies to instability problems (buckling). However, especially in this case we know that the
buckling strength of a structure can differ from calculated values, in view of differences between a real structure and the streamlined model of the calculation. This is one important reason for carrying out static tests on full-scale structures.

Another complication is due to the presence of notches and cracks. Geometric notches are unavoidable in a structure (joints, holes, etc.). They cause stress concentrations. Fortunately the effect of these stress concentrations on the static tensile strength is small for a ductile material as we will see later in this course. However, this is not true for cracks. Although we try to avoid cracks, they still may occur. In a fail-safe structure cracks can be acceptable, but the residual strength in the cracked condition should be known. Calculations are possible but additional tests are frequently desirable. High strength materials with a low ductility can be very sensitive to cracks under static loads. The fracture toughness \(K_{IC}\) is the material property to indicate this sensitivity as will be discussed later.

Failure modes depending on time usually have a fully different character if compared to static modes. They can have a ductile appearance (creep) or show a quasi brittle fracture (fatigue, stress corrosion). However, it is essential that the failure requires time or a repetition of load cycles. In general if time is significant temperature and environment will also affect the failure. Whereas the resistance to static modes can be expressed in terms of "strength", the resistance to time-dependent modes is often indicated by "life time". In general the relation between time-dependent failure resistance and wellknown material properties, such as yield stress, tensile strength, hardness, elastic modulus, is not clear. Calculation methods for time-dependent failures are partly available, especially for fatigue life and fatigue crack propagation. Corrosion can be a large problem for aircraft operators. The corrosion resistance of an aircraft structure cannot be calculated, but the designer can do very much to prevent corrosion. Creep is significant for engines, but usually not for aircraft structures.
1.3 FAILURE PREVENTION

It is the responsibility of the designer to consider all possible means for preventing premature failures, cracks and corrosion later in service. What the designer can do is:
(a) qualitative analysis of various preliminary design options and preventive measures
(b) quantitative calculations on strength and other properties of the structure
(c) complementary testing to fill gaps or to prove assumptions and calculations.

The qualitative analysis (a) is the real design work. It requires imagination to see the various possibilities to build up a structure. All advantages and disadvantages of different options should be recognized. Criteria for selections will be related to weight, strength, life, easy production, durability, etc. Usually there will be no single optimal solution. A simple illustration can amplify this point. In Figure 1.4 three alternative skin-stiffener attachments are shown.

![Figure 1.4 Three different stiffener-skin attachments](image)

The advantage of integral machining is that cross section variations can be made if numerically controlled machining apparatus (expensive) are available. For static strength optimal designs are more within reach than for b and c. However, from a fatigue point of view integral machining is a vulnerable solution. If a fatigue crack arises (e.g. due to unintentional damage) it will not meet a natural barrier to crack growth.
In a riveted stiffener a crack will not immediately run into the skin and the same applies to the bonded stringer. Also if a crack starts in the skin it will not automatically run into the stiffener. The adhesive bonded stiffener then is still superior to the riveted stiffener. With respect to fail-safe strength (= strength if cracks are present) solutions b and c are also superior to a. In view of corrosion there is also a difference. Obviously integral machining cannot be done on clad material, whereas for the other two solutions it can be done. It should also be noted that repairs and replacements will be most easy for the riveted structure. Production capabilities and quality control are entirely different for the three alternatives.

Another option also included in the above example is material selection. If a high strength alloy is chosen the static strength will be high. However, in general the ductility of high strength material is lower and related to this, the fatigue notch sensitivity is higher. From the fatigue point of view a somewhat lower strength alloy will be superior.

1.4 FAILURE ANALYSIS

Unfortunately structural failures do occur in service quite frequently. The majority are "incidents" and not "accidents". Whatever they are, safety and/or economy come into play. Remedial action to prevent similar failures in the future is necessary. An easy conclusion is that the "material" has failed. However, we should be more careful in failure analysis. The following example will illustrate this point. Figure 1.5 shows part of an undercarriage where a crack occurred in a forged cylindrical part at the edge of a massive lug. The fracture surface indicated a number of very small fatigue crack nuclei. These cracks caused a large crack, from which the oil of the shock absorber could escape. It occurred during a landing (touch down). The landing gear could then be pushed inwards further than normal. The wing tip hit the landing track and the aircraft made a ground sweep. This is not an incident but an accident.
Fig. 1.5  A failure of an undercarriage to illustrate many different potential causes.

Possible explanations are:
(a) The radius of the corner was too small (wrong detail design).
(b) The stress level was too high (inadequate stress analysis or underestimation of loads in service).
(c) The material was too notch sensitive (wrong material selection).
(d) Corrosion pits caused fatigue crack nucleation (insufficient corrosion protection).
Four different explanations suggest four different solutions, which are improved detail design, a lower design stress level, a more fatigue resistant material, improved corrosion protection. However, other solutions has to be considered also. Whenever a failure occurs in service it is an important question whether the failure was a symptomatic one or an incidental one. Some possibilities to be considered in this case are:

(e) An incidental production error or material defect caused the crack (quality control).
(f) Incidental damage caused the fatigue crack (maintenance and inspection involved).
(g) Extremely high loads could occur due to exceptional conditions, outside the design requirements.

Clearly enough the responsibility of the production shop and the aircraft operator are involved. In the present case further analysis learned that unknown reasons had caused excessive clearance in the torque link mechanism. As a result shimmy could occur causing extremely high loads.

The above example illustrates that the designer has to be aware of the many different causes which can lead to a failure. It requires understanding of the mechanisms involved, the various factors which can have an unfavourable influence and the possibilities to design a high quality aircraft structure.

1.5 SUBJECTS OF THE PRESENT COURSE

Stress analysis of structures and stability problems in many cases require a minimum information on material characteristics. The elastic moduli and the yield limit are sufficient for many problems. However for static tensile failures, fatigue and corrosion the material quality cannot be described by just a few numerical data. Understanding of the failure modes involved requires more knowledge on the characteristic behaviour of the material. The main topics of the present course are:
- failure in tension if notches or cracks are present
- stress corrosion
- material fatigue
- fatigue under different types of loading
- fatigue of joints
- fatigue of aircraft structures.

In order to discuss these problems two concepts (tools) have to be discussed first:
- stress concentrations at notches
- stress intensity factors for cracks.
Chapter 2  STRESS CONCENTRATIONS AT NOTCHES

Contents:
2.1 Introduction
2.2 Definition of $K_t$
2.3 Analytical calculations
2.4 The effect of the geometry (shape) on $K_t$
2.5 Superposition of notches and similarity of notches
2.6 Methods for determination of stress concentrations
2.7 References

2.1 INTRODUCTION

The theory of strength of structures is largely based on the theory of elasticity. If the yield stress is exceeded plastic deformation becomes significant and the rather complex theory of plasticity has to be adopted. Fatigue and stress corrosion occur at relatively low stress levels and elastic behaviour may well be assumed to be applicable. The elastic behaviour of an isotropic material is characterized by three elastic constants: $E =$ elastic modulus or Young's modulus, $G =$ shear modulus and $\nu =$ constant of Poisson ($\nu = 0.33$ for Al-alloys and $\nu = 0.285$ for steel). There is one relation between these constants:

$$G = \frac{E}{2(1+\nu)} \quad (2.1)$$

In a structure geometrical notches, such as holes, cannot be avoided. The notches are causing an inhomogeneous stress distribution, see Figure 2.1, with a stress concentration at the "root of the notch". The (theoretical) stress concentration factor $K_t$ is defined by:

$$K_t = \frac{\sigma_{\text{peak}}}{\sigma_{\text{nominal}}} \quad (2.2)$$

The stress concentration is highly depending on the geometry. A designer should always try to reduce stress concentrations as much as possible in order to avoid fatigue problems. In the present chapter we will discuss various aspects of stress concentrations and the effect of the geometry (the shape) on $K_t$. This is a necessary tool for designing a fatigue resistant structure.
2.2 DEFINITION OF $K_t$

The strip with a central hole shown in Figure 2.1 is loaded by a homogeneously distributed stress $S$. The load $P$ on the specimen is

$$P = W \times t \times S$$

($W$ = width, $t$ = thickness). In the minimum section ($y = 0$) at the hole (also called the net section or the critical section) the nominal stress is higher because the cross section is smaller

$$\sigma_{nom} = \frac{P}{(W-d)t} = \frac{S}{1 - \frac{d}{W}}$$  \hspace{1cm} (2.4)

In this section the stress distribution is inhomogeneous with a peak stress ($\sigma_{peak}$) at the root of the notch. The stress concentration factor $K_t$ is defined as the ratio of $\sigma_{peak}$ and $\sigma_{nom}$

$$K_t = \frac{\sigma_{peak}}{\sigma_{nom}}$$  \hspace{1cm} (2.5)

For the definition it is assumed that all deformations are elastic. $K_t$ gives a direct indication of the inhomogeneity of the stress distribution in the critical section. Sometimes it is informative to see the ratio between the peak stress and the gross stress $S$, which will be indicated as

$$K_{tg} = \frac{\sigma_{peak}}{S}$$  \hspace{1cm} (2.6)

For Figure 2.1:

$$K_{tg} = \frac{\sigma_{nom}}{S} \quad K_t = \frac{W}{W-d} \quad K_t > K_{tg}$$  \hspace{1cm} (2.7)

$K_t$ and $K_{tg}$ are the symbols used by R.E. Peterson in his book Stress Concentration Factors (Ref. 1), which is the standard book on stress concentration factors.
For bending and torsion the definition of $K_t$ is the same as given in
Eq. (2.5), i.e. the ratio between the peak stress and the nominal stress
on the critical section. For the strip with side grooves in Figure 2.2 the
nominal stress is the bending stress, that would be present if no stress
concentration occurs. For the linear stress distribution without stress
concentration:

$$\sigma_{\text{nominal}} = \frac{M}{\frac{1}{6} t h^2} \quad (2.8)$$

For the strip in Figure 2.3 this is

$$\sigma_{\text{nominal}} = \frac{M}{\frac{1}{6} (w-d) t^2} \quad (2.9)$$

As a result of the definition of $K_t$ we can also say: The peak stress is
$K_t$ times higher than the nominal stress obtained from a simple calculation
that ignores the stress concentration.

$K_t$-values can be determined by different methods:
- calculations:  - analytical methods
  - finite-element methods
- measurements:  - strain gage measurements
  - photo-elastic measurements.

2.3 ANALYTICAL CALCULATIONS

The principles of a 2-dimensional problem will be briefly indicated.
Consider the strip with a hole in Figure 2.1. The displacements $u$, $v$ of a
point $x$, $y$ will depend on these coordinates

$$u(x, y) \text{ and } v(x, y)$$
For \( y = \text{constant} \): 
\[
\frac{u(x+dx) - u(x)}{dx} + \frac{\partial u}{\partial x} \frac{dx}{x} + \frac{\partial^2 u}{\partial x^2} \frac{dx^2}{21} + \ldots 
\]

(2.10)

(Taylor series)

In the "linearized" theory of elasticity the second and higher order terms are neglected. As a result:

\[
\varepsilon_x = \frac{u(x+dx) - u(x)}{dx} = \frac{\partial u}{\partial x} 
\]

Similarly: \( \varepsilon_y = \frac{\partial v}{\partial y} \) and \( \gamma_{xy} = \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \)

(2.11)

Because \( u \) and \( v \) are functions of \( x \) and \( y \) the same applies to the three strains \( \varepsilon_x \), \( \varepsilon_y \) and \( \gamma_{xy} \). There are only 2 independent displacement functions \( u(x,y) \), \( v(x,y) \). This cannot be compatible with 3 independent strain functions. There should be one relation between the three strains, which follows by elimination of \( u \) and \( v \) from equations (2.11). This extra equation is the compatibility equation:

\[
\frac{\partial^2 \varepsilon_x}{\partial y^2} + \frac{\partial^2 \varepsilon_y}{\partial x^2} = \frac{\partial^2 \gamma_{xy}}{\partial x \partial y} 
\]

(2.12)

Differentiations of (2.11) easily prove (2.12) to be correct. The three strain functions \( \varepsilon_x(x,y) \), \( \varepsilon_y(x,y) \), \( \gamma_{xy}(x,y) \) have to satisfy this compatibility equation.

Stresses are linked to the strains by Hooke's law, which for plane stress \( (\sigma_z = 0) \) is:

\[
\varepsilon_x = \frac{\sigma_x}{E} - \nu \frac{\sigma_y}{E}, \quad \varepsilon_y = \frac{\sigma_y}{E} - \nu \frac{\sigma_x}{E}, \quad \gamma_{xy} = \frac{\tau_{xy}}{G} 
\]

(2.13)
The three stress components have to satisfy equilibrium. From the figure:

\[
\sigma_x \left( \frac{\partial \sigma_x}{\partial x} \right) \ dy + \left( \frac{\partial \tau_{xy}}{\partial y} \right) \ dx = 0
\]

or

\[
\frac{\partial \sigma_x}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} = 0
\]

Similarly

\[
\frac{\partial \sigma_y}{\partial y} + \frac{\partial \tau_{yx}}{\partial x} = 0
\]

(equilibrium of moments: \( \tau_{xy} = \tau_{yx} \))

At this point the so-called stress function of Airy: \( \phi(x,y) \) is introduced. It is a still unknown function, but it has to be selected in such a way, that:

\[
\sigma_x = \frac{\partial^2 \phi}{\partial y^2}, \quad \sigma_y = \frac{\partial^2 \phi}{\partial x^2}, \quad \tau_{xy} = -\frac{\partial^2 \phi}{\partial x \partial y} \quad (2.15)
\]

This is a clever requirement because substitution in the equilibrium equations (2.14) shows that these equations are automatically satisfied. Substitution of the stresses (2.15) into the strain equations (2.13), followed by substitution of these strains in the compatibility equation (2.12) leads to:

\[
\frac{\partial^4 \phi}{\partial x^4} + 2 \frac{\partial^4 \phi}{\partial x^2 \partial y^2} + \frac{\partial^4 \phi}{\partial y^4} = 0
\]

(also written as: \( \nabla^4 (\phi) = 0 \) (biharmonic equation))

with \( \nabla^4 = \nabla^2 \nabla^2 = \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)^2 \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \)
The problem now is: Find a function \( \phi \) that satisfy (2.16). Then the

equilibrium equations and compatibility equation are also satisfied. If

such a function is found, it still will contain unknown constants. The

unknown constants are obtained because \( \sigma_x(x,y), \sigma_y(x,y) \) and \( \tau_{xy}(x,y) \) have
to satisfy some boundary conditions. The boundary conditions can also

consist of prescribed displacements \( u(x,y), v(x,y) \) at the edges of the

plate.

For the case shown in Figure 2.1 the boundary conditions are:

1. At the upper and lower edge: \( \sigma_y = S, \sigma_x = 0, \tau_{xy} = 0 \)
2. At the side edges: \( \sigma_x = 0, \tau_{xy} = 0 \)
3. On the edge of the hole: normal stress and shear stress equal to zero.

Infinite plate with an elliptical hole loaded in tension

With the analytical method outlined above the theory of elasticity
has provided a solution for an infinite plate with an elliptical hole
loaded in tension (Figure 2.4). Although this appears to be a very simple
geometry, the solution of the problem is not so simple. It requires the use
of elliptical coordinates and a complex stress function (Ref. 2 and 3).
The solution provides the stress distribution in the entire plate. The
tangential stresses along the edge of the hole are of great interest.
The maximum stress occurs at the end of the main axis, see Figure 2.4., where:

\[
\sigma_{\text{peak}} = S \left( 1 + \frac{a}{b} \right) \quad \text{or} \quad K_t = 1 + 2 \frac{a}{b} = 1 + 2 \sqrt{\frac{a}{b}} \quad (2.17)
\]

Apparently a small tip radius will give a high \( K_t \) value. A large tip
radius will give a low \( K_t \), see Figure 2.5.

In aircraft structures holes of various shapes are adopted for windows,
doors, access openings, etc. It will always be profitable to use large
radii in view of reducing the stress concentration. The structure is not
directly comparable to an infinite sheet with an elliptical hole. In a
structure we will reinforce the edge of the opening for a further reduction
of the stress level. The added stiffness of the reinforcement will affect the overall stress distribution and thus $K_t$-values (see Ref. 4). Nevertheless the qualitative lesson of long or flat holes in Figure 2.5 remains valid.

It is noteworthy that the tangential stress at the end of the other axis (y-direction in Figure 2.4) is compressive and equal to the tensile stress applied to the infinite plate. This result is valid for all ellipses, and thus also for a circular hole ($a=b=r$), see Figure 2.6. For a circular hole the $K_t$ value is equal to 3. This is a classical value. For an open hole in an aircraft component the $K_t$ value will be slightly lower, because the component has a finite width. However, it will only be slightly lower. Fatigue cracks have occurred at relatively small open holes, which were used during production for tooling purposes. Such holes, which have no purpose after production should be filled up with hard driven rivets.

**Infinite plate with a circular hole loaded in shear**

For a plate with a circular hole loaded in shear the $K_t$-value is readily obtained by superposition, see Figure 2.7. Pure shear can be split up in a tension case under $45^\circ$ and a compression case under $-45^\circ$. For these two cases the stresses at A and B follow directly from Figure 2.6. The summation implies:

$$
\sigma_{\text{peak}} = 4 \tau \rightarrow K_t = 4
$$

(2.18)

This is a rather high stress concentration. High shear stresses can be present in the shear web of a spar and a rib. If holes have to be made in these webs because of pipe lines or other equipment, reinforcements will be necessary to reduce the stress concentrations. (Usually it is also necessary for stability problems).

**Stress gradient**

Although the peak stress is of the utmost importance it is also interesting to know how fast the stress will drop off at some distance of the root of
the notch. For an elliptical hole in an infinite sheet the exact solution for \( \sigma_y \) along the X-axis is given by (with \( c^2 = a^2 - b^2 \))

\[
\left( \frac{\sigma_y}{S} \right)_{y=0} = 1 + \frac{a(2 - 2b)}{(a-b)} \frac{x - \sqrt{x^2 - c^2}}{\sqrt{x^2 - c^2}} \left( \frac{x^2 - c^2}{x^2 - c^2} \right) + \frac{ab^2}{(a-b)} x
\]

This reduces to a much simpler equation for a circular hole:

\[
\left( \frac{\sigma_y}{S} \right)_{y=0} = 1 + \frac{1}{2} \left( \frac{a}{x} \right)^2 + \frac{3}{2} \left( \frac{a}{x} \right)^4
\]

(2.20)

In Figure 2.8 the \( (\sigma_y) \) distribution is shown for an elliptical hole \( a/b = 3, K_t = 7 \) and \( y=0 \) a circular hole. Obviously the peak stress drops off much faster for the higher peak stress.

An equation for the stress gradient of \( (\sigma_y) \) at the notch root \( (x=a) \) is easily obtained by differentiation of \( (\sigma_y) \) Eq. (2.19) with \( K_t = 1 + 2 a/b \) (Eq. 2.17) the result can be written as:

\[
\left( \frac{d\sigma_y}{dx} \right)_{x=a} = - \left( 2 + \frac{1}{K_t} \right) \frac{\sigma_{\text{peak}}}{p} = - \beta \frac{\sigma_{\text{peak}}}{p}
\]

(2.21)

\[
\alpha = \left( 2 + \frac{1}{K_t} \right) + 2 < \alpha < 3
\]

(2.22)

Apparentely \( K_t \) does not have a large effect on the stress gradient coefficient \( \alpha \), and as a matter of fact for a variety of notches in finite-width situation \( \alpha \) is about 2 to 2.5 (Ref. 5).

The stress gradient gives an indication of the volume of the highly stressed material at the root of a notch. As a numerical example, let us consider at which depth \( q \) below the surface of a circular hole with a diameter of 5 mm,
the peak stress has dropped to \(0.9 \sigma_{\text{peak}}\). As a first approximation we assume a linear drop off. From equations (2.21) and (2.22) we obtain:

\[
\frac{d\sigma}{dx}_{x=a} \approx -\frac{\sigma_{\text{peak}} - 0.9 \sigma_{\text{peak}}}{q} = -2 \frac{1}{3} \frac{\sigma_{\text{peak}}}{2.5 \, \text{mm}} + q \geq 0.1 \, \text{mm}
\]

Since a characteristic grain size of an Al-alloy is 40 \(\mu \text{m}\) the depth \(q\) corresponds to just a few grains. Conclusion: Especially the grains at the notch root surface are the highly loaded grains. This is important for fatigue.

In view of the significance of surface grains the variation of the stress along the edge of the hole has to be considered also. Figure 2.9 shows lines of constant maximum principal stress. It shows that the highly stressed material is stretched along the edge of the hole. This emphasizes the significance of the "surface quality" for good fatigue properties.

2.4 THE EFFECT OF THE GEOMETRY (SHAPE) ON \(K_t\)

Exact analytical calculations are more easy for an infinite plate because that simplifies the boundary conditions. An infinite plate with an infinite row of equally spaced holes (diameter \(d\), pitch \(W\)), see Figure 2.10, was analysed analytically by Schulz. For \(d/W \to 0\) (that means \(d\) very small, or \(W\) very large) the classical result \(K_t = 3\) applies. For increasing \(d/W\) the \(K_t\) value drops to lower values and for \(d/W = 1\) the \(K_t\) values becomes 1 (question: is the latter limiting value in agreement with expectations?)

By making vertical cuts through the infinite plate of Schulz a strip with a central hole (specimen I) or a strip with two semi-circular edge notches (specimen II) is obtained. The lines along which the vertical cuts are made are not "stress free". Along these lines \(\tau_{xy} = 0\) (for reasons of symmetry) but \(\sigma_x \neq 0\). As a result the \(K_t\) for specimens I and II is different from \(K_t\) for the infinite plate. The small difference for specimen II is quite surprising. Larger differences occur for the strip with a central hole. This case was analysed analytically by Howland. The relation
\[
K_t = 2 + \left(1 - \frac{d}{W}\right)^3
\]  

(2.23)

is an approximation given by Heywood.

\(K_t\) for a strip with central hole is replotted in Figure 2.11, together with \(K_{tg}\). The first factor is a decreasing function, the latter one an increasing function of \(d/W\). For a constant load on a strip (constant \(S\)) it appears logical that \(\sigma_{peak}\) will increase if the hole is made larger for a constant width, which implies that \(K_{tg}\) will increase. However, the stress distribution becomes less inhomogeneous. The ratio between \(\sigma_{peak}\) and the average value of the stress distribution (\(\sigma_{nom}\)) becomes smaller: \(K_t\) will decrease.

According to Figure 2.11 the \(K_t\) value is depending on \(d/W\), that means on the ratio of the two dimensions, which specify the geometry. This is true in general. The dimensionless quantity \(K_t\) can be dependent on dimensionless ratio's only. In Figure 2.12 the ratio's \(d/W\) and \(H/t\) are exactly the same for both specimens. All dimensions of specimen II are 2 x larger than for specimen I. As a result of the geometrical similarity all displacements will also be 2 x larger. The relative displacements, however, are similar. That implies that the strains are the same. Consequently equal stresses will occur in both specimens. The same peak stress will be found and \(K_t\) is the same. (It should be noted that the non-dimensionless stress gradient is not the same. According to Eq. (2.21) it is inversely proportional to the root radius \(\rho\). Larger specimens have larger volumes of highly stressed material, which is significant for the size effect on fatigue.)

Many \(K_t\)-graphs for various shapes and different types of loading can be found in the book of Peterson (Ref. 1) while the ESDU Data sheets (Ref. 4) also give a collection. Some simple examples will be shown here to illustrate the effect of the shape on the stress concentration. The predominant effect of the root radius on \(K_t\) is shown in Figure 2.13 for fillets and edge notches. Thickness variations of a component are equivalent to fillets. The graph shows that a smaller radius gives a larger \(K_t\) and the same applies to the
edge notches. In some textbooks the occurrence of a stress concentration is explained by a picture showing the so-called principal stress trajectories (i.e. lines following the directions of the principal stress), see Figure 2.15. At the notch these lines have to bend rather sharply around the notch. The density of the lines at the notch root is large, which corresponds to stress concentration.

Another example is shown in Figure 2.14 for an axle with a stepped reduction of diameter. Also here a "generous" radius gives a much lower $K_t$. For this case large radii are not always possible in practice, for instance if a ball bearing has to support the axle at the change of the diameter. In such a case stress releasing grooves should be considered, see for example Figure 2.16 which is an effective increase of the root radius. Another example of stress relief is shown in Figure 2.17, where two extra holes with a smaller diameter "smoothed the stress flow". The stress concentration factor in this case is reduced by 20 percent.

A comparison between a T-head and a fillet is made in Figure 2.18. The $K_{t}$-values are significantly higher for the T-head. The reason is that the load application on the T-head occurs very close to the root of the notch. The thicker T-head (m=d) gives more bending stiffness and a lower $K_{t}$, but it is still much higher than for the fillet. Also the effect of the root radius is again obvious from Figure 2.18.

In Figure 2.19 the specimen is a so-called lug or lug head. In the literature it is also referred to as a pin joint or a pin loaded hole. In aircraft structures it is an important joint, a joint with a single bolt, where rotation in the joint should be possible. The lug in Figure 2.19 has a square head, whereas rounded heads are more usually in practice. However, $K_{t}$-values do not differ appreciably between square and rounded lug heads. For comparison the $K_{t}$-value of an open hole (curve from Figure 2.14) is also shown. Apparently the $K_{t}$-values of a lug for the same $d/W$ are considerably higher than for an open hole. This is another example of load application close to the root of the notch (AA in Figure 2.19). These high $K_{t}$-values imply that lugs can be fatigue critical, the more so since fretting corrosion inside the hole is another contributing factor. This will be discussed in chapter 10.
2.5 SUPERPOSITION OF NOTCHES AND SIMILARITY OF NOTCHES

If at the root of a notch a second smaller notch is added there is a superposition of notches. Examples are shown in Figures 2.20 and 2.23. In Figure 2.20 a small semi-circular notch occurs in the critical section of an open hole. The small additional notch can represent mechanical damage inside the hole. As a first approximation the $K_t$ factor will be equal to the product of the two stress concentration factors, $K_{t1}$ for an open hole and $K_{t2}$ for a semi-circular edge notch:

$$K_t \approx K_{t1} \times K_{t2}$$  \hspace{1cm} (2.24)

For Figure 2.20 this implies $K_t \approx 9$ which is extremely high. Nevertheless this estimate is reasonable if $r_2 << r_1$. If $r_2$ becomes larger there will be a tendency to overestimate $K_t$. In case 2 the small notch occurs in a semi-infinite plate to which a homogeneous stress is applied. However, at the edge of the hole the peak stress ($\sigma_{\text{peak}} = 3 S$) is present only locally, whereas it drops off around this point. For larger secondary notches Equation (2.24) will be pessimistic. Two cases shown in Figure 2.21 illustrate this argument. These cases are shown because exact solutions for an infinite plate (Ref. 8) are available.

In view of the predominant effect of the tip radius on $K_t$, ellipses have been drawn in Figure 2.21 with the same span (2a) and tip radius ($r_2$) as the holes with the hole edge notches. For these "equivalent" ellipses the $K_t$ is readily obtained from Equation (2.17):

$$K_t = 1 + 2 \sqrt{a/r_1}.$$  \hspace{1cm} Figure 2.21 shows that the $K_t$-values for the ellipses are fairly close to the exact values. Another example of the application of similar notch shapes is shown in Figure 2.22 for slots. Also here exact values are available for an infinite plate (Ref. 9). The $K_t$-values are very close to the exact values the differences varying from 1 to 3 percent. It should be noted that slots loaded in the direction transverse to the slot have high $K_t$-values.
Geometrical shapes to be similar in view of stress concentrations require a similar span (or depth) and a similar root radius. If the edge notches in Figure 2.15 are replaced by semi-elliptical edge notches with the same depth and the same root radius, the picture of the trajectories would hardly be different. Approximately the same $K_t$ will apply.

In Figure 2.23 another example of superimposed notches is shown. The small lubrication hole occurs in the critical section of the lug. $K_{t1}$ can be read from Figure 2.19 at the value $d_1/W$, and $K_{t2}$ from Figure 2.11 at $d_2/t$. Calculation of $K_t = K_{t1} \times K_{t2}$ will overestimate the real $K_t$-value. However, $K_t$ will be high anyhow because of adding a notch to the most critical section of the lug. Lubrication holes which enter the lug hole either at the top or the bottom of that hole are in a much better position.

A last illustration is shown in Figure 2.24. An open hole is reinforced by riveted doublers. Some rivet holes occur at a location where the stress level is still above the nominal one. The reinforcement may be beneficial for the edge of the hole, but rivet holes may impose a more severe stress raiser. Reinforcement of structural holes (doors, windows, etc.) should always be done carefully with due consideration to new stress raisers.

2.6 METHODS FOR DETERMINATION OF STRESS CONCENTRATIONS

As said before $K_t$-values can be determined by calculations or by measurements. In the preceding paragraphs several examples were drawn from analytically calculated results. It was also pointed out that similarity of notches (same root radius, same depth or span) may be helpful to estimate $K_t$-values. Sometimes interpolation between existing data is possible. Available data (e.g. Refs 1 and 4) should always be consulted first. For complex shapes analytical calculations become impracticable. If stress concentration factors or stress distributions are still urgently needed, calculations can be made by the Finite Element Methods. The geometry of the component has to be modelled as a large number of small elements, which are interconnected by specified element edge
conditions. More elements are required at places where stress gradients are high. This is illustrated by the model in Figure 2.25. The stress distribution in such a model can be calculated provided a computer of sufficient capacity is available. The solution is not an exact one, because the continuum material is replaced by a multi-element material. However, very satisfactory results can be obtained.

In the early days before the finite-element technique was available many $K_t$-values were obtained by measurements, especially by photo-elasticity (Refs 11, 12). Various graphs in Peterson's book are based on such measurements. Figure 2.19 is also based on photo-elastic measurements. Two pictures obtained by photo-elasticity are shown in Figure 2.26. Local stresses can be derived from the isochromatics. The advantage of this method is that an overall impression of the stress distribution is obtained. Moreover in model tests the photo-elastic model can be easily modified to see how improved stress distributions can be obtained. The accuracy of the method, however, is not so high.

An alternative measuring technique is to use strain gages. Strains can be measured fairly accurate. A problem is that the model should be large and the gage length should be small. This is especially true at notches, where the stress gradient is large. The gage will "average" the strain below the filament of the gage. Strain gages are extensively used for measuring nominal stresses in full-scale structures or components.
2.7 REFERENCES


4. ESDU Data sheets on Fatigue and on Stress Concentrations. Engineering Science Data Unit, London.


3.1 INTRODUCTION

Residual stresses can be present in a structure, or in a component, or even in a plane sheet, while there is no external load applied. In view of the absence of an external load the residual stresses are sometimes labelled as internal stresses. The background of the terminology "residual stress" is that a residual stress distribution in a material is left as a residue of inhomogeneous plastic deformation. Both residual tensile stress and residual compressive stress occur. Actually they have to occur together, because the residual stress system must be an equilibrium system, see Figure 3.1. Residual tensile stresses have to be balanced by residual compressive stresses, because there is no external load neither an external moment. Consequently in Figure 3.1

\[ \int_{-t/2}^{t/2} \sigma_x \, dy = 0 \]  \hspace{1cm} (3.1a)

\[ \int_{-t/2}^{t/2} \sigma_x y \, dy = 0 \]  \hspace{1cm} (3.1b)

We will refer to a residual stress distribution as a residual stress system. Such a system can be the result of inhomogeneous plastic deformation, but there are other possibilities to introduce a residual
stress system inside a structure or a component, to which no external load is applied. This is discussed in paragraph 3.2.

The significance of residual stresses stems from some practical consequences:

1. If a load is applied the residual stress has to be added to the external stress caused by the load:

   \[ \sigma = \sigma_{\text{external}} + \sigma_{\text{residual}} \]

Depending on the sign of the residual stress \( \sigma \) will be larger or smaller than the calculated stress \( \sigma_{\text{external}} \). This can have serious consequences for fatigue and stress corrosion. The material does not make any difference between external stress and residual stress. For that reason residual compressive stresses are favourable and residual tensile stresses are unfavourable. Favourable residual stresses are sometimes introduced in a component at those locations where it is needed (e.g. by shot peening). An unfavourable residual stress system may be introduced without knowing (e.g. by a heat treatment).

2. In a plate or a forging with a residual stress system machining will cause warpage. A simple example is shown in Figure 3.2. Removal of the bottom layer implies elimination of a part of the residual stresses. This upsets the equilibrium. The plate will automatically restore the equilibrium by becoming slightly shorter (difficult to observe) and by bending (more easily observed).

### 3.2 DIFFERENT SOURCES OF RESIDUAL STRESS SYSTEMS

Residual stress systems can be present in a material, a component or a structure because of:

1. Inhomogeneous plastic deformation at notches
2. Production processes
3. Heat treatment
4. Assembling.
Inhomogeneous plastic deformation

A simple theoretical model will be discussed first. In Figure 3.3 two tension bars are rigidly connected at the ends to load the bars in parallel. As long as the deformation is elastic the loads in the two bars will be equal \((P_1 = P_2)\). We now assume that one bar has a lower yield stress. The bars are loaded until an elongation \(\Delta l = \Delta l'\), which implies that bar 1 is still elastic, whereas bar 2 has deformed plastically. Unloading will cause "elastic spring back", and at \(P = P_1 + P_2 = 0\) residual loads will be present: a tensile load on the elastic bar and a compression load in the plastic bar. In other words: plastic deformation implies permanent deformation, and as a result bar 2 has become longer than it was before. Therefore it cannot fit anymore "stress free" between the rigid clamping in the unloaded condition. There will be residual loads: \(P_2^\text{res} = -P_1^\text{res}\).

One step further is to give both bars the same yield stress but a different load. This is done in Figure 3.4. Bar 2 is two times shorter than bar 1. If we now give the two bars the same elongation \(\Delta l\), bar 2 will carry more load than bar 1 \((P_2 > P_1)\), because it is stiffer due to its shorter length. There is a load "concentration" in bar 2. Plastic deformation will start first in bar 2. After unloading at \(\Delta l = \Delta l'\) again elastic spring back will occur until \(P = 0\) which implies that \(P_2^\text{res} = -P_1^\text{res}\).

A simple continuous model is a plate loaded in pure bending, see Figure 3.5. The maximum stress occurs in the outer fibres and plastic deformation will start in the upper and lower surface layer. If the bending moment is removed elastic spring back will occur and a residual stress system will remain. Equation (3.1a) will be satisfied automatically (symmetry), while spring back will occur until the moment equation (3.1b) is satisfied also. The upper surface layer which was loaded in tension by \(M\), will see a residual compressive stress. In terms of permanent deformation: The upper and lower surface layers were strained plastically. As a result the upper layer became longer than it was before and the lower layer became shorter. The plate got a curvature. The permanent deformations leave
the plate in a strained condition after unloading. Spring back is a well-known phenomenon of sheet metal bending operations, e.g. in the production of stiffeners. If sheet metal bending is done as a cold-working process residual stresses are present in the final product.

A final example of inhomogeneous plastic deformation is illustrated by Figure 3.6. If $\sigma_{\text{peak}}$ exceeds the yield limit a small plastic zone is formal at the root of the notch. As a consequence $\sigma_{\text{peak}}$ is smaller than $K_t \sigma_{\text{nom}}$. The peak stress is leveled off. In the plastic zone permanent plastic deformation has occurred. The plastic zone is larger than it was before. So in the unloaded condition it will be under compression. It does no longer fit into its elastic surrounding and it will cause a residual stress system. After unloading residual compressive stresses occur at the root of the notch (favourable for fatigue), which are balanced by tensile stresses at a larger distance.

**Production processes**

It was shown before that plastic bending will introduce residual stresses. This applies to almost all cold-working processes, such as rubber pressing, stretch forming, deep drawing etc. Forging usually is a hot-working process. Also rolling is done at higher temperatures, but rolling for sheet straightening is done at low temperatures. Anyhow, rolled sheets and plates may carry a residual stress system. A simple method to eliminate residual stresses from sheets and plates is to stretch the material to a small plastic strain. Figure 3.7 shows that the inhomogeneous residual stress distribution becomes much more homogeneous in the stretch operation. During elastic unloading the shape of the distribution (5 in Figure 3.7) will not change. As a result the original residual stresses are practically eliminated. Plates can be ordered from the aluminium industry in the prestrained condition. It is more expensive.

It is not always realized that machining operations can also introduce residual stresses. However, metal cutting implies a material failure process near the tip of the cutting tool. This failure process is preceded by plastic deformation. Depending on machining conditions residual stresses can be significant, although they do occur in a thin surface layer only.
Shot peening

Shot peening is a well-known process to introduce favourable residual stresses at the material surface of a component. In various cases it is applied to prevent fatigue or corrosion problems. The peening operation tries to make the surface larger than it is. As a result residual compressive stresses are introduced at the surface. Another result is that a plate shot peened at one side only will be curved. Dimensional distortions can be prevented only in a symmetric peening operation. The intensity of the shot peen operation can be measured by peening a so-called Almen strip from one side only, and measure the curvature of the strip afterwards (see Figure 3.8). The length of the strip is 76 mm (3"), the width 19 mm (0.75"), while different thicknesses are used.

Heat-treatment

Quenching is a very abrupt step of the heat-treatment cycle. Cooling occurs very fast at the outside of a component, see Figure 3.9, and slower at the inside. Due to the inhomogeneous cooling thermal stresses are introduced. The faster thermal contraction at the outside will cause tensile stresses at the outside, balanced by compressive stresses inside. If plastic deformation occurs, it will result in tangential stresses, which are compressive at the outside and tension at the inside. It is not essential that the section is hollow. However, if moisture can enter the section in service the residual tensile stresses may cause stress corrosion from the inside.

Residual compressive stresses at the outside of a component are favourable. Unfortunately many components have complex shapes and this makes it difficult to predict the pattern of the residual stress system. Tensile residual stresses at the outside surface may be possible. They can be reduced (or even reversed) by shot peening.

Assembling

During the assembling of aircraft components bolts are frequently used to join different parts. A schematic example is shown in Figure 3.10.

If tolerances on dimensions are inadequate it is possible that the two
mating surfaces at \( AB \) cannot really match. If the bolt is tightened it will lead to bending in the flanges of the small forgings. These stresses are sometimes referred to as "built-in stresses". They may well cause stress corrosion failures. Specified tolerances and production quality are involved.

3.3 CALCULATIONS ON RESIDUAL STRESSES AT NOTCHES

In view of the stress concentrations around notches it is quite well possible that loads in service will introduce residual stresses at the notch root (Figure 3.6). These stresses are significant for fatigue life and fatigue damage accumulation.

The calculation of stresses around notches if plastic deformation does occur is practically impossible with an analytical treatment of the problem. Such calculations can be done with finite-element technique, but the calculations are complex and expensive. Fortunately a reasonable estimate of \( \sigma_{\text{peak}} \) can be made with a rather simple procedure, based on a postulate of Neuber. (*).

As long as plastic deformation does not occur all stresses and all strains are proportional to the applied load. This is Hooke's law. The shape of the \( \sigma \)- and \( \varepsilon \)-distribution in the critical section is independent of the load applied. However, as soon as a plastic zone is formed at the notch root both the shapes of the \( \sigma \)-distribution and the \( \varepsilon \)-distribution will become different. The stress at the notch root (\( \sigma_{\text{peak}} \)) will be lower than the elastic prediction (Figure 3.11a) and the strain at the same place (\( \varepsilon_{\text{peak}} \)) will be higher (Figure 3.11b). In other words:

\[
\frac{\sigma_{\text{peak}}}{\sigma_{\text{nom}}} < K_t \quad (3.2) \\
\frac{\varepsilon_{\text{peak}}}{\varepsilon_{\text{nom}}} > K_t
\]

The postulate of Neuber implies that the fact of \( \sigma_{\text{peak}} \) being smaller than the elastic prediction is always related to the other fact of \( \varepsilon_{\text{peak}} \) being larger than the elastic prediction. This relation occurs in such a way that

the product of $\sigma_{\text{peak}}$ and $\varepsilon_{\text{peak}}$ is still confirming to the elastic prediction. In other words:

$$\sigma_{\text{peak}} \varepsilon_{\text{peak}} = K_t^2 \sigma_{\text{nom}} \varepsilon_{\text{nom}}$$  \hspace{1cm} (3.3)$$

Defining concentration factors $K_\sigma$ and $K_\varepsilon$ for the notch root by:

$$K_\sigma = \frac{\sigma_{\text{peak}}}{\sigma_{\text{nom}}} (< K_t)$$

$$K_\varepsilon = \frac{\varepsilon_{\text{nom}}}{\varepsilon_{\text{nom}}} (> K_t)$$  \hspace{1cm} (3.4)$$

the postulate of Neuber becomes:

$$K_\sigma K_\varepsilon = K_t^2$$  \hspace{1cm} (3.5)$$

Neuber proved that the postulate is correct for a hyperbolic notch under shear loading. He then assumed that it will be approximately true for other types of notches and loading. This has been confirmed empirically, provided the plastic zone is relatively small.

Substitution of $\varepsilon_{\text{nom}} = \sigma_{\text{nom}}/E$ into Eq. (3.3) leads to

$$\sigma_{\text{peak}} \varepsilon_{\text{peak}} = \left(\frac{K_t \sigma_{\text{nom}}}{E}\right)^2$$  \hspace{1cm} (3.6)$$

For a given load ($\sigma_{\text{nom}}$) and $K_t$ available, this equation gives us one relation between the unknown $\sigma_{\text{peak}}$ and $\varepsilon_{\text{peak}}$. A second relation is necessary for a solution and for that purpose the stress-strain relation as obtained in a tensile test is adopted. The graphical solution now is simple as shown in Figure 3.12. It is found at the intersection (A) of the two curves, where $\sigma_{\text{peak}}$ and $\varepsilon_{\text{peak}}$ satisfy both relations. Equation (3.6) is an orthogonal hyperbola which moves upwards for higher loads (i.e. higher $\sigma_{\text{nom}}$). If no plasticity had occurred the solution would not occur at A in Figure 3.12, but at B instead. The stress at B thus is $K_t \sigma_{\text{nom}}$. The residual stress after unloading is equal to $\sigma_{\text{peak}}$ minus the elastic spring back occurring at the notch root, which is $K_t \sigma_{\text{nom}}$. This implies a residual peak stress at the notch root (Fig. 3.6)

$$\sigma_{\text{res}} = \sigma_{\text{peak}} - K_t \sigma_{\text{nom}}$$  \hspace{1cm} (3.7)$$
as indicated in Figure 3.12. The Neuber postulate allows us to estimate \( \sigma_{\text{res}} \) after a high load has been applied to a notched element.

The discussion on Figures 3.6 and 3.12 explains how a residual compressive stress can be introduced at a notch. It should be clear that a high negative load can introduce in a similar way a residual tensile stress at the root of a notch. If an aircraft is pulled up out of a dive, a large upwards bending moment is applied on the wing. This can introduce favourable residual compressive stresses in the bottom structure of the wing and unfavourable residual tensile stresses in the upper structure of the wing.

3.4 SOME ADDITIONAL REMARKS

Residual stresses as discussed before occur on a macro scale. They have the same character as external stresses induced by loads applied to a structure. On a much smaller scale another type of residual stress can be indicated. Plastic deformation on a microscale is not a homogeneous process. It will be different from grain to grain (grain orientation, grain size) and even inside a single grain it may be concentrated in a few slip bands. As a result residual stresses varying from grain to grain will occur. Also in this case equilibrium requires that the sum of the micro-residual stresses is zero. The micro-residual stresses will be smaller in aluminium alloys than in steel, because Al has a low elastic anisotropy and a relatively large number of slip systems. The micro-residual stresses are significant for explaining the fatigue mechanism on a micro level and for the Bauschinger effect. Brief reference will be made to it in a later chapter.

The macro-residual stresses of the previous sections are the more important ones for the engineer in order to understand the prevention of fatigue and stress corrosion. It is rather unfortunate that we cannot see these residual stresses. There are no simple means to measure a residual stress. A non-destructive measurement can be done by X-ray deffraction techniques, but this is a fairly elaborate method, which certainly cannot be adopted on a routine basis. Destructive tests are possible. For instance it can be done by bonding a small strain gage on the surface where \( \sigma_{\text{res}} \) has to be measured.
The change of the output of the gage, after sufficient cuts are made in the material around the gage, will indicate the residual stress which was present before. The cuts are made to fully release the material under the gage from any interior loading. Another simple method to be used for plate and sheet only is to remove the surface material in small steps (by chemical milling) and to measure the curvature after each step. The change of curvature allows a calculation of the residual stress which was present before (Report VTH-155 of the Department of Aerospace Engineering, Delft).
4.1 INTRODUCTION

In aircraft structures fatigue cracks and stress corrosion cracks do occur in service. Stress corrosion cracks can be avoided by proper material selection, corrosion protection and other means, see chapter 6. However, the occurrence of fatigue cracks has to be accepted. If an aircraft should be designed in such a way that fatigue cracks can never occur the structure would be much too heavy. The design philosophy with respect to fatigue therefore is:

(a) Try to achieve a long crack free life.
(b) If fatigue cracks do occur (insufficient life, or crack initiation due to poor quality, incidental damage, etc.) it should be known where cracks may start (in view of inspections) and how fast they will grow (inspection periods).
(c) Fatigue cracks should have a limited effect on the static strength. Obviously (a) is related to "economy" and (b) and (c) to "safety". Also with respect to fatigue the designer can improve the quality by an adequate choice of fatigue resistant materials, good surface conditions, etc. Moreover, a structural lay-out with good fail-safe properties can be adopted (crack stopping properties, inspectability).

In view of the above philosophy the designer has to deal with fatigue cracks in relation to material behaviour and structural efficiency. For this purpose the stress-intensity factor $K$ is an essential tool. The stress intensity factor gives an indication of the severity of the stress field around the tip of a crack. The factor is dependent on the geometry of the component, the shape and the size of the crack, and the loads applied. This is discussed in the present chapter. Applications to fatigue crack growth, static strength if cracks are present and stress corrosion are dealt with in other chapters.
4.2 DIFFERENT TYPES OF CRACKS AND CRACK OPENING MODES

If a large fatigue crack is growing in the skin of a pressurized fuselage (large ≡ centimeters, if not decimeters in a Jumbo) the crack is growing through the full thickness of the skin sheet material. If a small crack is initiated at the edge of a hole in a thick forging it is not growing through the full thickness. In the literature reference is made to (see Figure 4.1):
- through cracks,
- part through cracks.

Part through cracks can be (approximately) quarter elliptical corner cracks or semi-elliptical surface cracks. Although most fatigue cracks start at holes they can be initiated at other places as well. For instance surface cracks can be nucleated at surface damage (nicks or dents), corrosion pits or fretting damage. After part through cracks have penetrated the full thickness they will become through cracks. This chapter will be largely restricted to through cracks because the analytical treatment of part through cracks is more complex.

Cracks under load can be opened in different ways. This is illustrated in Figure 4.2. In Mode I the crack is opened perpendicular to the fracture surfaces. In Modes II and III the displacements occur in the plane of the crack (shear modes). Because fatigue cracks have a strong tendency to grow perpendicular to the main principal stress the opening mode I is the more important one. This chapter deals with mode I cracks only. A mixed mode I/III occurs during shear lip formation where a fatigue crack is growing fast (chapter 7).

4.3 DEFINITION OF THE STRESS INTENSITY FACTOR K

The most simple case to be considered is a through crack in an infinite sheet loaded in tension, see Figure 4.3. For this case an exact solution is available if the (linearized) theory of elasticity is applicable. The assumed elastic behaviour is the same condition made in the previous chapter for the elliptical hole. Actually the solution for the crack is obtained from the solution for the elliptical hole, for which the minor axis is reduced to zero (b → 0). The flat ellipse becomes a crack then. The equation for the stress concentration of an elliptical hole is:
\( K_t = 1 + 2 \frac{a/b}{1 + 2\sqrt{a/\rho}} \)  \hspace{1cm} (4.1)

(with tip radius \( \rho = b^2/a \)). The obvious result for a crack is:

\[ b \to 0 \quad (\rho \to 0), \text{ then } K_t \to \infty \]

This is what we should expect for an infinitely sharp notch. Whatever the geometry of the component is, if there is a crack: \( K_t = \infty \). The impossibility of an infinite peak stress will be discussed later.

In chapter 2 an equation for \( \left( \sigma_y \right)_y=0 \) (stress distribution along the X-axis) was presented, see equation (2.19). This was for an elliptical hole. If we substitute \( b = 0 \) and \( c = a \) \((c^2 = a^2 - b^2)\) in that equation, and adopt polar coordinates \((r, \theta)\), see Figure 4.3) with the pole at the tip of the crack \((x=a+r\) for \(\theta = 0\)), it simplifies drastically to:

\[ \frac{\left( \sigma_y \right)_\theta=0}{S} = \frac{a + r}{\sqrt{2ar + r^2}} \]

This is still an exact solution. We now restrict our interest to a small zone around the crack tip, that means to:

\[ r \ll a \]

Equation (4.2) simplifies further to:

\[ \left( \sigma_y \right)_\theta=0 = S \frac{a}{\sqrt{2ar}} = \frac{S\sqrt{\pi a}}{\sqrt{2\pi r}} \]

(\( It \) is not necessary to add the \( \pi \) factor, but this is common practice in the international literature.) For \( \theta \neq 0 \) the exact solutions for the various stress components \((\sigma_x', \sigma_y', \tau_{xy}\) in Fig. 4.3) cannot be written up as explicit functions. However, under the restriction \( r \ll a \) it can be done. The resulting equations are very much similar to Eq. (4.4), because only a function of the polar angle has to be added:

\[ \sigma_{ij} = \frac{S\sqrt{\pi a}}{\sqrt{2\pi r}} f_{ij}(\theta) \]

(4.5)
The function \( f_{ij}(\theta) \) is different for the various stress components.

Fully written out:

\[
\sigma_x = \frac{S\sqrt{\pi a}}{\sqrt{2\pi r}} \cos \frac{\theta}{2} \left( 1 - \sin \frac{\theta}{2} \sin \frac{3\theta}{2} \right) - S \tag{4.6a}
\]

\[
\sigma_y = \frac{S\sqrt{\pi a}}{\sqrt{2\pi r}} \cos \frac{\theta}{2} \left( 1 + \sin \frac{\theta}{2} \sin \frac{3\theta}{2} \right) \tag{4.6b}
\]

\[
\tau_{xy} = \frac{S\sqrt{\pi a}}{\sqrt{2\pi r}} \cos \frac{\theta}{2} \sin \frac{\theta}{2} \cos \frac{3\theta}{2} \tag{4.6c}
\]

These equations are exact only for \( r \to 0 \), that means they are asymptotically correct. However, for \( r \ll a \) the equations are very good approximations of the exact solutions. The stress distribution shows a singularity at the crack tip because all stress components go to infinity for \( r \to 0 \).

Along each line through the tip of the crack (\( \theta = \) constant) the stress components increase with \( r^{-1/2} \) if we go to the tip. They become infinite for \( r = 0 \). The factor \( r^{-1/2} \) characterizes the "power" of the singularity.

In equation (4.6) the term \(-S\) in (4.6a) is the only non-singular term, which does not disappear for \( r \to 0 \). In most textbooks it is still dropped because it is so small as compared to the non-singular term.

However, it is of some interest that substitution of \( \theta = \pi \) gives

\( \sigma_x = -S \). Because \( \theta = \pi \) is the location of the crack edge it implies that there is a compressive stress \((-S)\) along the crack edge. As a matter of fact this is not only true for \( r \ll a \) but for the entire crack edge \((-a \leq x \leq a\) (for the ellips this result was only found at \( x = 0 \)).

The above equations can also be written as:

\[
\sigma_{ij} = \frac{K}{\sqrt{2\pi r}} f_{ij}(\theta) \tag{4.7}
\]

with

\[
K = S\sqrt{\pi a} \tag{4.8}
\]

The factor \( K \) apparently gives a complete indication of the "intensity" of the stresses around the crack tip. For that reason \( K \) was labelled
as the **stress intensity factor**. For the case of an infinite sheet under tension (Figure 4.3) there are two variables: the tensile stress $S$ and the crack length $a$.

The effect of both variables on the stresses in the crack tip area are fully accounted for by the factor $K$.

The dimension of $K$ is unusual:

$$ [K] = [\text{stress}] \times [\text{length}]^{1/2} $$

In modern units $[K] = (N/m^2) \times (m)^{1/2} = N/m^{3/2}$. This unit is very small and therefore it is more convenient to use $MN/m^{3/2} = 10^6 N/m^{3/2}$. Since $N/m^2$ is also replaced by Pa (Pascal) the dimension of $K$ is frequently indicated by MPa$\sqrt{m}$. The relation to other units is:

$$ 1 \text{ MPa}\sqrt{m} = 3.225 \text{ kgf/mm}^{3/2} = 910 \text{ psi}\sqrt{in} $$

An exact analytical solution for an infinite sheet with an infinite row of collinear cracks (Fig. 4.4) is also available (Westergaard solution). For the crack tip area equation (4.7) is valid again, but the equation for $K$ (eq. 4.8) requires a correction factor $C$:

$$ K = CS\sqrt{\pi a} \quad (4.9) $$

with

$$ C = \sqrt{\frac{tg(\pi a/W)}{\pi a/W}} \quad (4.10) $$

The factor $C$ is referred to as the geometry factor (or geometry correction factor). It is a function of the dimensionless ratio of the two dimensions $a$ and $W$, and so $C$ itself is also dimensionless. In figure 4.4 $C$ has been plotted as a function of $a/W$. It starts at $C = 1$ for $a/W = 0$ (equivalent to $W = \infty$) and it goes to infinity for $2a \to W$ (fully cracked).

It can be shown that for finite dimensions Eq. (4.9) is equally
valid, which implies that equations (4.6) can be replaced by:

\[
\sigma_x = \frac{K}{\sqrt{2\pi r}} \cos \frac{\theta}{2} \left( 1 - \sin \frac{\theta}{2} \sin \frac{3\theta}{2} \right) \quad (4.11a)
\]

\[
\sigma_y = \frac{K}{\sqrt{2\pi r}} \cos \frac{\theta}{2} \left( 1 + \sin \frac{\theta}{2} \sin \frac{3\theta}{2} \right) \quad (4.11b)
\]

\[
\tau_{xy} = \frac{K}{\sqrt{2\pi r}} \cos \frac{\theta}{2} \sin \frac{\theta}{2} \cos \frac{3\theta}{2} \quad (4.11c)
\]

or more generally:

\[
\sigma_{ij} = \frac{K}{\sqrt{2\pi r}} f_{ij}(\theta) \quad (4.12)
\]

with

\[
K = CS\sqrt{\pi a} \quad (4.13)
\]

A confirmation follows from the photo-elastic picture shown in Figure 4.20. The isochromatics in this figure are lines of constant maximum shear stress. The butterfly type lines at the crack tips \((r << a)\) are similar for all three cracks. It implies that \(f_{ij}(\theta)\) is the same for all cracks. The geometry affects \(K\) only, and it enters \(K\) through the geometry factor \(C\) (eq. 4.13). From isochromatics it also can be deduced that \(\sigma_{ij} \propto (r)^{-1/2}\) (eq. 4.12) but Figure 4.20 is insufficient detailed and accurate to check this.

4.4 EXAMPLES OF STRESS INTENSITY FACTORS

Stress intensity factors can be calculated or measured, as will be briefly discussed later. In view of the general equation \(K = CS\sqrt{\pi a}\) (eq. 4.13) the unknown factor is the geometry factor \(C\). In the literature there are three wellknown compilations of available \(K\)-factors:

Supplements to (1) and (2) have been issued. More solutions are published in the literature from time to time.

The fracture process at the tip of a crack is depending on the stress intensity in the crack tip area. This is true for fatigue, stress corrosion and static strength if cracks are present. For the characterization of the crack growth resistance of a material, crack growth results of tests on simple specimens can be used. \( K \) values of simple specimens are discussed below.

**Specimen with central crack** (CCS = center cracked specimen)

A sheet or plate specimen with a central crack (Figure 4.4) is frequently used for obtaining fatigue crack growth data. The crack growth rate \( (da/dn) \) is a function of the cyclic variation of \( K \) as will be discussed later.

If a specimen with a central crack is sufficiently long the \( K \)-factor is independent of the length. The two remaining dimensions are the crack length \( (a) \) and the specimen width \( (W) \). As a result the geometry factor will be a function of \( a/W \). An exact solution is not available, but a very accurate formula is the so-called Feddersen relation:

\[
C = \sqrt{\sec(\pi a/W)} \tag{4.14}
\]

(Note: \( \pi a/W \) in radians.) For \( W \rightarrow \infty \) equation (4.14) indicates \( C \rightarrow 1 \) as should be expected. For increasing crack length \( C \) also increases, see Figure 4.4, but the finite width correction is still less than 20\% \( (C < 1.2) \) for a 50\% cracked specimen \( (2a/W = 0.5) \). Equation (4.10) for the infinite sheet with co-linear cracks is compared to equation (4.14) in Figure 4.4. There is an increasing difference with the Feddersen factor. A finite width specimen is obtained from the infinite sheet by making cuts along the vertical lines 1 and 2 in Figure 4.4. However, these lines are not "stress free" \( (\sigma_x \neq 0) \), and consequently equations (4.10) and (4.14) cannot be identical.

**Compact tension specimen (CTS)**

The compact tension specimen shown in figure 4.5 is generally adopted to determine the so-called fracture toughness of a material. This
material property, usually indicated by the symbol $K_{IC}$ (I for opening mode I, and c for critical), is obtained in a static test. At the moment of failure the $K$-value calculated for the failure load is the critical value: $K_{IC}$. This is discussed in more detail in chapter 5. The advantages of the compact tension specimen are:

1. It is small and thus it does not require much material.
2. A relatively low load is sufficient for a high $K$-value, which can be significant in view of a limited maximum capacity of a testing machine.

The equation for the geometry factor $C$ is presented in figure 4.5. As shown by the graph in this figure the factor $C$ is significantly higher than for the center cracked specimen.

**Cracks at an open hole**

For an open hole with 1 or 2 edge cracks in an infinite plate loaded in tension exact solutions were given by Bowie. The geometry factor $C$ is presented in figure 4.6, for both uniaxial loading ($\lambda = 0$) and for biaxial loading ($\lambda = 0.5$ and 1, note that 0.5 applies to a pressure vessel). There are two definitions of the crack length in figure 4.6: (1) crack length $l$ from edge of hole to crack tip and (2) crack length $a$ (actually $2a$) which includes the hole.

There is one remarkable observation to be made in Figure 4.6. If $l$ exceeds about 15 percent of the hole radius ($l/r > 0.15$) the geometry factor (for uniaxial loading: $\lambda = 0$) is fairly close to 1, where $C$ is defined by:

$$K = CS\sqrt{a}$$

In other words: already for a small crack at a hole the stress intensity factor is approximately the same as for a much larger crack with a length that includes the hole diameter.

This is generally true for cracks originating at notches. If the width or the depth of the notch is added to the real crack length, see Figure 4.7, an effective crack length ($a_{eff}$) is obtained.

This effective length should be used for an approximate indication of the stress intensity at the crack tip. In Figure 4.7 a crack at the edge of a window thus can introduce a rather severe situation. Actually this is what happened during the Comet accidents in the early fifties.
Another illustration is given in Figure 4.8, which shows that edge cracks of holes in a finite width strip induce stress intensities of a comparable magnitude as central cracks of the same overall span 2a. If such a hole with cracks would occur in the tension skin of a wing structure it would be very dangerous. However, large holes such as access openings are closed by covers, bolted to the edge of the hole. Only for inspections the covers are removed. Proper connections between cover and edges are significant if cracks arise.

**Edge cracks**

A classical case is an edge crack in a semi-infinite sheet, see Figure 4.9. For this case:

\[
K = 1.12 \sqrt{\frac{a}{2}}
\]  
(4.15)

or \(C = 1.12\) (more accurate \(C = 1.1215\)). This is 12 percent higher than for a crack in an infinite sheet (Figure 4.3). Because the edge is free of stress no restraint on crack opening can be exerted along the edge. A somewhat higher \(K\)-value should thus be expected.

If a very small crack occurs at the root of a notch, see Figure 4.10, it may be considered to be a variant of Figure 4.9. That requires \(S\) to be replaced by \(\sigma_{\text{peak}}\), or:

\[
K = C \sigma_{\text{peak}} \sqrt{\frac{a}{2}}
\]  
(4.16)

with \(C = 1.12\) as a first approximation. It cannot be more than an approximation because \(\sigma_{\text{peak}}\) is present at the notch root only. The stress decays rapidly away from the notch and somewhat slower along the notch root edge. As shown in chapter 2 stress gradients are related to the notch root radius and \(\sigma_{\text{peak}}\) (eq. 2.21). This allows a better approximation for small cracks than \(C = 1.12\), which is (Ref. 6):

\[
C = 1.122 - 3 \left(\frac{a}{\rho}\right) + 4 \left(\frac{a}{\rho}\right)^{1.5} - 1.7 \left(\frac{a}{\rho}\right)^2
\]  
(4.17)

For cracks larger than some 20 percent of the root radius the \(a_{\text{eff}}\) approximation according to Figure 4.7 is more appropriate. If accurate information on \(K\) is required either calculations or measurements of
K have to be made.

Crack edge loading

In Figure 4.11 a few cases of crack edge loading have been indicated (P per unit thickness, N/m). It appears to be unrealistic to have loads which act on a single point of a crack edge. However, cracks starting from rivet holes or bolt holes come fairly close to this configuration. The load applied by the rivet or the bolt is a concentrated load if the crack is sufficiently large as compared to the hole diameter (case 3 in Figure 4.11).

For an infinite sheet with two loads on the crack edges along the center line an exact solution is available:

\[ K = \frac{P}{\sqrt{\pi a}} \]  

(4.18)

Because the dimension of \( P \) is N/m the dimension of \( K \) is still N/m\(^{3/2}\). Equation (4.18) seems to deviate from the general equation \( K = CS\sqrt{\pi a} \), but we can rewrite Eq. (4.18) as:

\[ K = \frac{2}{\pi} \left( \frac{P}{2a} \right) \sqrt{\pi a} \]  

(4.19)

If \( P \) would be spread over the full crack length 2a, then \( P/2a \) would be a bearing pressure and thus a stress, while \( 2/\pi \) would be the geometry factor.

The remarkable thing about equation (4.18) is that \( K \) decreases if \( a \) increases, see figure 4.12. This is also true for a finite width strip (case III in Fig. 4.12). However, it has been confirmed by a decreasing growth rate in fatigue crack propagation tests (see chapter 8).

4.5 K-FACTORS BY SUPERPOSITION

If a loading system 1 is inducing stresses \( (\sigma_{ij})_1 \) in the crack tip area and a second loading system 2 stresses \( (\sigma_{ij})_2 \), these stresses have to be added to obtain the stresses \( \sigma_{ij} \) if both loading systems are applied:

\[ \sigma_{ij} = (\sigma_{ij})_1 + (\sigma_{ij})_2 \]
or

\[
\frac{K}{\sqrt{2\pi r}} f_{ij}(\theta) = \frac{K_1}{\sqrt{2\pi r}} f_{ij}(\theta) + \frac{K_2}{\sqrt{2\pi r}} f_{ij}(\theta) + K = K_1 + K_2
\]  \hspace{1cm} (4.20)

In other words: superposition of loading systems implies that the corresponding K factors should be added. A simple, but interesting example is shown in Figure 4.13. Case 1 is considered as the sum of case 2 and case 3. Case 3 is a crack under an internal pressure \( p = S \).

In case 2 the crack edges are loaded by the stress \( S \), which is just sufficient to keep the crack closed. It implies that in case 2 a homogeneous stress \( S \) will be present in the full sheet. This means that there is no singular stress field at the crack tip.

Consequently \( K_2 = 0 \) and \( K_3 = K_1 \). The stress intensity factor for a crack under internal pressure is the same as for a crack in a sheet loaded in tension.

The same result can also be obtained by another superposition. The internal pressure can also be considered as a continuous row of point loads (\( Sdx \)) at a distance \( x \) from the center line. With the equation for two point loads (case 2 in Figure 4.11) we obtain

\[
K = \int dK = \int_{-a}^{+a} \frac{Sdx}{\sqrt{\pi a}} \sqrt{\frac{a + x}{a - x}} = S\sqrt{\pi a} \hspace{1cm} (4.21)
\]

Another example is shown in figure 4.14. It represents a plate with a stiffener, which is not continuous. The last rivet connecting the stiffener to the plate will transmit more load from the stiffener into the plate than the other rivets. The hole of this "end rivet" is fatigue critical. For a crack at this hole the loading system is modelled as case 1 in figure 4.14. This case is split up in two other cases 2 and 3. \( K_2 \) is known (Figure 4.4). To arrive at \( K_3 \) (case 3) we add case 3' obtained by rotating case 3 over 180°. The rotation does not change \( K \), thus \( K_3' = K_3 \). Summing 3 and 3' and splitting up the sum in case 4 and case 5 we arrive at two cases for which \( K \) is available (Figures 4.4 and 4.12). The final solution for \( K_1 \) is indicated in Figure 4.14.
4.6 CRACK OPENING AND THE STATE OF STRESS

A crack in a component loaded in tension will be opened (mode I). For an infinite sheet with a central crack (Figure 4.3) an exact (explicit) solution is available for the crack edge displacements. As should be expected the displacement is maximal at the center \((x = 0)\), for which the result is:

\[
(v)_{x=0,y=0} = 2 \frac{S}{E} a = 2\varepsilon_{\infty} a \quad (4.22)
\]

The total opening at \(x = 0\) is twice this amount. For \(S = 70 \text{ N/mm}^2\) (20% of \(\sigma_{\text{yield}}^*\) for 2024-T3), \(E = 70 000 \text{ N/mm}^2\) and a crack of \(2a = 20 \text{ mm}\) the opening is 0.04 mm, which is small indeed.

Near the crack tip an asymptotic solution \((r \ll a)\) can again be obtained in terms of the stress intensity factor \(K\). They are different for plane strain and plane stress. (Note: \(\sigma_{ij}\) is the same for both states of stress.)

\[
u = \frac{K}{G} \sqrt{\frac{r}{2\pi}} \cos \frac{\theta}{2} \left[1 - 2\nu + \sin^2 \frac{\theta}{2}\right] \quad \text{plane strain} \quad (4.23)
\]

\[
v = \frac{K}{G} \sqrt{\frac{r}{2\pi}} \sin \frac{\theta}{2} \left[1 - 2\nu + \sin^2 \frac{\theta}{2}\right] \quad \text{plane strain} \quad (4.24)
\]

\[
u = \frac{K}{G} \sqrt{\frac{r}{2\pi}} \cos \frac{\theta}{2} \left[1 - \nu + \sin^2 \frac{\theta}{2}\right] \quad \text{plane stress} \quad (4.25)
\]

\[
v = \frac{K}{G} \sqrt{\frac{r}{2\pi}} \sin \frac{\theta}{2} \left[1 + \nu + \sin^2 \frac{\theta}{2}\right] \quad \text{plane stress} \quad (4.26)
\]

Instead of \((r)^{-0.5}\) we now have \((r)^{+0.5}\). For the crack edges \((\theta = \pi)\) the horizontal displacement is: \(u = 0\) for plane strain and plane stress. The vertical displacement is:

\[
v = \frac{K}{G} \sqrt{\frac{r}{2\pi}} (2 - 2\nu) = \frac{4K}{E/(1-\nu^2)} \sqrt{\frac{r}{2\pi}} \quad \text{(plane strain)} \quad (4.25)
\]

\[
v = \frac{K}{G} \sqrt{\frac{r}{2\pi}} \left[\frac{2}{1 + \nu}\right] = \frac{4K}{E} \sqrt{\frac{r}{2\pi}} \quad \text{(plane stress)} \quad (4.26)
\]

The ratio of the two \(v\)-values is \(1 - \nu^2 \approx 0.9\). The crack tip opening is slightly smaller under plane strain. This result is based on an assumed elastic behaviour.
Plane strain is defined by $\varepsilon_z = 0$ (z is the thickness direction), and plane stress is defined by $\sigma_z = 0$. With the relation:

$$\varepsilon_z = \frac{\sigma_z}{E} - \frac{\nu(\sigma_x + \sigma_y)}{E} \tag{4.27}$$

the result for plane strain is:

$$\varepsilon_z = 0 + \sigma_z = \nu(\sigma_x + \sigma_y) \tag{4.28}$$

and with Eq. (4.11) $\sigma_z = \frac{2\nu K}{\sqrt{2\pi r}} \cos \frac{\theta}{2}$

For $r \to 0$ the result is $\sigma_z \to \infty$.

For plane stress we obtain:

$$\sigma_z = 0 + \varepsilon_z = -\frac{2\nu K}{E\sqrt{2\pi r}} \cos \frac{\theta}{2} \tag{4.30}$$

For $r \to 0$ the result is $\varepsilon_z \to \infty$.

What is the meaning of these results for a through crack in a plate or sheet, see Figure 4.15. Near the crack front $\sigma_x$ and $\sigma_y$ are extremely high. Consequently the material wants to contract in the lateral direction ($\varepsilon_z$). According to equation (4.30) an infinitely large contraction ($\varepsilon_z \to \infty$) should be expected for $r \to 0$ at the crack front. However, this highly stressed material is surrounded by material of a much lower stress level. The contraction in the surrounding material is relatively low, because $\varepsilon_z$ decreases with $1/\sqrt{r}$ for increasing $r$. In view of the continuity of the material the extremely high contraction for $r \to 0$ cannot occur, because it will be prevented by the surrounding material. As a result the material at the crack front will be approximately in a state of plane strain. In view of the prevented contraction a stress $\sigma_z$ is built up along the crack front in accordance with equation (4.28).

At the plate surface (ends of the crack front) $\sigma_z$ must go to zero. There will be a thin surface layer where plane stress ($\sigma_z = 0$) will prevail. If the sheet material is very thin a tendency for plane stress over a (relatively) larger part of the crack front will be
present. The thickness is too low to effectively prevent lateral 
contraction. Plane stress will be stimulated by crack tip plasticity 
to be discussed later on.

4.7 ENERGY CONSIDERATIONS AND THE GRIFFITH CRITERION

An uncracked plate will be considered first, see Figure 4.16. It is 
loaded by an increasing stress \( S \), which corresponds to a load \( P = Swt. \) 
The plate will become longer, the elongation being: \( \Delta H = He = HS/E. \) 
As a result the work done is (Fig. 4.16):

\[
A = \frac{1}{2} (Swt) \left( \frac{HS}{E} \right) = \frac{1}{2} \left( \frac{S^2}{E} \right) (Hwt) = \frac{1}{2} \left( \frac{S^2}{E} \right) \text{(volume)} \tag{4.31}
\]

The work per unit volume is \( 1/2 \frac{S^2}{E} \). It is important to realize that 
the work done is stored in the material as "strain energy" (like in a 
spring which is stretched).

If the plate has a crack the stiffness is lower and the plate can be 
extended more easily (Fig. 4.16). The work stored in the plate as strain 
energy will be less by an amount \( A \Delta \). This amount will be calculated 
for an infinite plate with a crack. As a result of the stress \( S \) the 

crack is opened, see figure 4.17. From the exact solution the crack 
edge displacements are known:

\[
v = 2 \frac{S}{E} \sqrt{a^2 - x^2} \tag{4.32}
\]

(Equation 4.32 represents an ellips with a major axis \( a \) and a minor 
axis \( b = v \) \( x = 0 \) = \( 2aS/E \), see Eq. 4.22.) We now apply an increasing 
pressure on the crack edges to close the crack. At full closure the 
work done per unit of length of the crack edge is \( 1/2. S \ t \ v(t = \text{thickness}). \) 
Integrated over both fracture surfaces the work done is:

\[
\Delta A = 2 \int_{-a}^{+a} \frac{1}{2} S \ t \ v \ dx \tag{4.33}
\]

Substitution of (4.32) leads to:

\[
\Delta A = \pi a^2 t \frac{S^2}{E} \tag{4.34}
\]

If we now start from an uncracked sheet under tensile stress \( S \) and a 

crack of length \( 2a \) is introduced this will lead to an energy release 
\( A \Delta \) equal to \( \Delta A \) just calculated:
\[ \Delta U = \pi a^2 t \frac{S^2}{E} \]  

(4.35)

In the literature \( \Delta U \) is usually expressed per unit thickness and (4.35) then reduces to:

\[ \Delta U = \pi a^2 \frac{S^2}{E} \]

If we now let the crack length a increase with a small amount \( da \) (at both tips) a further energy release will be:

\[ d(\Delta U) = dU = 2\pi a \frac{S^2}{E} \, da \]

and for each crack tip it is half this value, which implies:

\[ \frac{dU}{da} = \frac{\pi a S^2}{E} \]  

(4.36)

With \( K = \sqrt[4]{\pi a} \) the "energy release rate" is

\[ \frac{dU}{da} = \frac{K^2}{E} \]  

(4.37)

Equation (4.37) has been derived here for an infinite plate. It can be shown to be valid also for finite dimension, but the proof will not be given here.

Griffith in 1924 suggested that the small amount of strain energy \( dU \), which is available if the crack becomes longer by a small amount \( da \), is consumed by creating new surfaces at the crack tip. This requires a "surface energy" \( \gamma \) per unit of new surface, which implies an energy increment (per unit of thickness):

\[ 2 \, da \, \gamma \]

According to Griffith a crack will become unstable if more strain energy will be released than necessary for creating new crack surfaces:

\[ dU \geq 2 \, da \, \gamma \]  

(4.38)

If this is combined with equation (4.36) the result is:
\[
\frac{\mu a S^2}{E} \text{da} \geq 2 \text{da} \gamma + S \geq \sqrt{\frac{2\gamma E}{\pi a}} \tag{4.39}
\]

and the critical stress is:

\[
S_c = \sqrt{\frac{2\gamma E}{\pi a}} \tag{4.40}
\]

This is the Griffith criterion for unstable crack extension. It is based on elastic material behaviour. There is no energy dissipation by crack tip plasticity.

If Equation (4.38) is combined with Equation (4.37) instead of (4.36) the result is:

\[
\frac{K^2}{E} \text{da} \geq 2 \text{da} \gamma + K \geq \sqrt{2\gamma E} \tag{4.41}
\]

and the critical K value is:

\[
K_c = \sqrt{2\gamma E} = \text{constant} \tag{4.42}
\]

This is the same Griffith criterion, but now in terms of the stress intensity factor. As said before it assumes elastic behaviour, which implies that it is valid for brittle material. Griffith applied it to glass. For high strength alloys, which have a very low ductility, \(K_c\) is found to be a constant (= fracture toughness), but it is much larger than \(\sqrt{2\gamma E}\). The reason is that in this low ductility material some plastic deformation still occurs, and that requires much more energy than the surface energy \(\gamma\).

### 4.8 Cracks with a Curved Crack Front

In figure 4.1 the through crack with a straight crack front (which is perpendicular to the plate surfaces) is the more simple configuration. This type of crack was considered in the previous sections. In the same figure two examples are shown with curved crack fronts. For a corner crack the shape of the crack front usually approximates a quarter ellipse and sometimes even a quarter circle. For a surface crack it will be close to a semi-ellipse or a semi-circle. If the crack front is curved the problem has a 3-dimensional character. The K-factor varies along the crack front. It is no longer constant. The most simple case solved analytically by Sneddon is a circular crack in an infinite solid loaded intension, see Figure 4.18. For this
penny-shaped crack the K factor is still constant in view of circular symmetry.

If the circular crack of Figure 4.18 is replaced by an elliptical crack, see Figure 4.19 the K-factor varies along the crack front. Irwin derived the following equation:

\[ K(\psi) = \frac{S \sqrt{\pi E}}{\phi} \left\{ \sin^2 \psi + \left( \frac{b}{a} \right)^2 \cos^2 \psi \right\}^{1/4} \]  

(4.43)

The angle \( \psi \) determines the location \( D \) along the crack front in the way as shown in Figure 4.19. The symbol \( \phi \) represents a so-called complete elliptical integral of the second kind. It is depending on the aspect ratio \( a/b \) of the ellips:

\[ \phi(a/b) = \int_{0}^{\pi/2} \sqrt{1 - k^2 \sin^2 \theta} \, d\theta \quad \text{with} \quad k^2 = 1 - (b/a)^2 \]  

(4.44)

It can be found in tables, but a very good approximation is:

\[ \phi(a/b) = [1 + 1.464 \, (b/a)^{1.65}]^{1/2} \quad \text{(for } a \geq b) \]  

(4.45)

An illustration of the variation of K along elliptical crack fronts is given in Figure 4.19. Apparently K has a minimum at the end of the long axis (in A) and a maximum at the end of the short axis (in B). For a fatigue crack this will give a tendency to grow to a circular shape.

For quarter- or semi-elliptical cracks in finite dimension components the K-formula deviate from Equation (4.43). Solutions in the literature are usually presented as:

\[ K(\psi) = K_{\text{elliptical}} \cdot F \left( \frac{a}{b}, \frac{a}{r}, \frac{a}{W}, \frac{b}{t}, \psi \right) \]  

(4.46)

The corrections to be made for a finite width (W), thickness (t), the presence of a hole (radius r) if present, are contained in the correction function \( F \). The corrections are dependent on the aspect ratio \( a/b \) and the location angle \( \psi \) (Fig. 4.19). Exact solutions are not available. Formulas in the literature are partly based on clever estimates whereas calculated results obtained with the finite element technique are also available (Refs. 7-9).
4.9 DETERMINATION OF K-VALUES

If K-values are required the best approach is to consult handbooks first (Refs. 1-3). Sometimes it is possible to make clever estimates by evaluating available data for "similar" geometries. Considerations presented in the discussion of Figures 4.7 and 4.9 can be helpful. Sometimes superposition arguments can be used. However, when accurate K-values are required calculations or measurements have to be made. For calculations the finite-element technique can be adopted. It will be clear that a very fine mesh will be required around the crack tip, because the stress and strain gradients are very high. If stresses or displacements are calculated in the crack tip area the results are to be compared to \( \sigma_{ij} = (K/\sqrt{2\pi r}) f_{ij} (\theta) \) (eq. 4.7) or to \( u_1 = (K/G) \sqrt{r/2\pi} f_{ij} (\theta) \) (eqs. 4.23 and 4.24). From this comparison K has to be obtained. The accuracy is limited because in the equations the behaviour is singular for \( r \to 0 \). Special crack tip elements have been developed which have the singular stress behaviour \( (r^{-1/2}) \) as a property of the element.

A completely different calculation method is based on the strain energy release rate \( dU/da = K^2/E \) (eq. 4.37). It implies that \( U \) has to be calculated for some values of \( a \) in order to determine \( dU/da \). In a finite element model the crack length is increased by uncoupling of nodes along the crack line. Experience has shown that an acceptable accuracy can be obtained without the necessity to go to an extremely fine mesh.

K values can be determined by measurements, but it is not easy. The main reason is again that the gradients in the crack tip area are extremely high. This practically excludes the possibility of using strain gages. Photo-elastic measurements have been made, even for 3-dimensional cases, but again the accuracy is problematic in view of the high gradients. Figure 4.20 gives an illustration of a photo-elastic picture. It illustrates that similar stress distributions do occur at all three crack tips. A quantitative measurement, however, requires a much more refined photo-elastic measuring technique. An indirect method is to start from fatigue crack propagation results and to deduce the K factor from the crack growth rate. This method is discussed in chapter 8.
4.10 CRACK TIP PLASTICITY

According to the equation

\[ \sigma_{ij} = \frac{K}{\sqrt{2\pi r}} f_{ij}(\theta) = \frac{CS\sqrt{\pi a}}{\sqrt{2\pi r}} f_{ij}(\theta) \]  \hspace{1cm} (4.12)

the stress becomes infinite for \( r \to 0 \). This would be a disaster if the material is brittle. Although high strength structural materials very often have a low ductility they are not really brittle. As a result of the ductility a small plastic zone will be created. The infinite peak stress is leveled off. In the crack tip plastic zone the above equation can no longer be valid, because it was based on an assumed elastic behaviour. The question then is: will this small plastic zone completely destroy the meaning of \( K \)? Fortunately this is not the case. In figure 4.21 a circle has been drawn around the crack tip to give an approximate indication of the crack tip area, where \( K \) gives a meaningful indication of the stress intensity. Remember, that the derivation of Equation (4.12) was based on \( r \ll a \). If inside this area (radius \( r_e \)) a small plastic zone (size \( r_p \)) is formed it will cause a certain redistribution of the stress. However, if \( r_p \) is small as compared to \( r_e \) the redistribution will hardly affect the stresses at the periphery of the crack tip area. In other words \( K \) should still be a good indication for the severity of the stress field inside the crack tip area. In spite of the fact that the elastic equation (4.12) cannot be valid in the small plastic zone, the occurrence of the plastic zone and the size of this zone will be controlled by \( K \). We now consider two different specimens or components (same material) with different dimensions, different crack lengths and different loads. The same plastic zone at the crack tip will occur in both parts if the \( K \) values are the same. If some crack growth mechanism is associated with the plastic zone (e.g. in fatigue) the same crack extension will be found if the same \( K \) values apply. It is for this reason that \( K \) can be a useful parameter to correlate the material sensitivity for fatigue crack propagation, stress corrosion cracking and static strength if cracks are present. A comparison can then be made between a laboratory specimen and a component. If they have the same \( K \) the result of the laboratory specimen is used as a calibration result to predict the behaviour of the crack in the component.
From the requirements \( r_e < a \) and \( r_p < r_e \) it will be clear that large plastic zones will invalidate the usefulness of \( K \). Large plastic zones will be formed if the material has a low yield stress and if the load applied is relatively high. The size of the plastic zone is also dependent on the state of stress (plane strain or plane stress). A first estimate of \( r_p \) can be made as follows.

The stress distribution of \( \sigma_y \) along the X-axis follows from Equation (4.11b) by substituting \( \theta = 0 \):

\[
\sigma_y = \frac{K}{\sqrt{2\pi r}}
\]  
(4.47)

This hyperbolic relation is shown in figure 4.22. A first estimate of \( r_p \) for a plane stress situation is obtained from \( \sigma_y = \sigma_{0.2} \).

Substitution in equation 4.47 gives:

\[
\sigma_{0.2} = \frac{K}{\sqrt{2\pi r_p}} \rightarrow r_p = \frac{1}{2\pi} \left( \frac{K}{\sigma_{0.2}} \right)^2
\]  
(4.48)

This will be an underestimation because it ignores that the leveled off part of the stress distribution (shaded area in Figure 4.22) has to be carried also. It should be expected that \( r_p \) will be larger. The contents of the shaded area obtained by integration of equation (4.47) (\( r = 0 \) to \( r = r_p \)) turns out to be equal to \( r_p \sigma_{0.2} \). A better estimate of \( r_p \) then seems to be twice the value of equation (4.48) and that result is usually given in the literature:

\[
r_p = \frac{1}{\pi} \left( \frac{K}{\sigma_{0.2}} \right)^2 \quad \text{(plane stress)}
\]  
(4.49)

For plane strain an estimate can be made by adopting a yield criterion, e.g. the von Mises criterion. Because plane strain implies a constraint on lateral contraction the effective yield stress will be higher (substitute \( \sigma_2 = \sqrt{\sigma_x \sigma_y} \) in the von Mises criterion). As a result the plastic zone will be significantly smaller. The relation frequently quoted in the literature is:

\[
r_p = \frac{1}{3\pi} \left( \frac{K}{\sigma_{0.2}} \right)^2 \quad \text{(plane strain)}
\]  
(4.50)

The above equations do not recognize the fact that the plastic zone is not circular (as suggested in Figure 4.22), but has more a butterfly type shape. Nevertheless the proportionality:
\[ r_p \approx \left( \frac{K}{\sigma_{0.2}} \right)^2 \]

is approximately correct and the equations (4.49) and (4.50) can be used as a first approximation. A numerical example: Assume \( S = 100 \text{ N/mm}^2 \), \( \sigma_{0.2} = 400 \text{ N/mm}^2 \) (Al-alloy) and \( a = 10 \text{ mm} \).

\[
\begin{align*}
    r_p &= \frac{1}{(3)^\pi} \left( \frac{100 \sqrt{\pi} \cdot 10}{400} \right)^2 = 0.6 \text{ mm} \quad \text{(plane stress)} \\
    &= 0.2 \text{ mm} \quad \text{(plane strain)}
\end{align*}
\]

If \( r_p \) is much smaller than the material thickness (e.g. \( t = 10 \text{ mm} \)) lateral contraction at the crack front will be largely restrained and plane strain will apply at the above stress level. However in a thin sheet (e.g. \( t = 1 \text{ mm} \)) it will be largely in plane stress.
REFERENCES

Chapter 5  FAILURE IN TENSION

Contents:
5.1 Introduction
5.2 The tensile strength of a material
5.3 The hardness of a material
5.4 The static tensile strength of a notched element
5.5 The fracture toughness $K_{IC}$
5.6 Cracks in sheets, behaviour under static loading
References

5.1 INTRODUCTION

A component subjected to a monotonically increasing tensile load will ultimately fail. Such a static failure will generally start at a geometrical notch (e.g. rivet holes or bolt holes). If a crack is present failure can be due to crack extension. Suppose that time-dependent phenomena are not involved the failure can be labelled as a quasi-static failure. However, some time-dependent phenomena, such as creep, can occur during a very slowly increasing load. For technical materials creep does not occur at room temperature and for the aircraft structure it is not a significant failure mode. On the other hand, for several parts of a jet engine creep must be considered.

Very rapidly applied loads may cause dynamic effects which can affect the failure load. We have to make a distinction between two essentially different types of dynamic effects:

1. Effects due to the dynamic response of the structure.
2. Loading rate effects on material properties.

The first type of effects is due to inertia loads. For instance during a severe landing high inertia loads can occur at the moment of touch down. The occurrence of dynamic effects will be recognized in most cases, but the calculation of the dynamic loads may be difficult. Anyhow such calculations have to be made to arrive at the ultimate design loads.

The second type of dynamic effects is related to the material behaviour. Plastic deformation is supposed to be a loading-rate dependent phenomenon. However, for structural materials to observe any effect of loading rate (actually $\dot{\varepsilon}/\dot{t}$) on the yield stress, extremely high loading rates are required. These loading rates are outside the range of the loads on an
aircraft. The most well known example of a rate-dependent behaviour is shown by mild steel (C-steel with a low carbon contents), but this material is not an aircraft material. In addition mild steel has a reputation for cold-brittleness which also does not apply to aircraft materials (it is even unacceptable). The only thing to be recognized for aircraft materials is that the yield stress and the tensile strength are slightly higher at low temperature, whereas the elongation is somewhat lower. In other words the ductility of the material is slightly less at low temperature. However, from the point of view of dislocation mobility and diffusion inside the material there are no reasons to see significant rate-dependent effects on the tensile strength of aircraft materials.

Static tensile failures occurring in aircraft components are usually preceded by some plastic deformation before decohesion occurs. It should be expected that geometrical notches have an unfavourable effect on the static strength. This should be true even more if cracks are present. Before the static strength of notched and cracked elements is discussed, we first will deal with the significance of the "tensile strength" of a material, and how this is related to the ductility of the material.

5.2 THE TENSILE STRENGTH OF A MATERIAL

The stress-strain curve (σ-ε diagram) as obtained in a static tensile test is generally supposed to be characteristic for the elastic and plastic behaviour of a material. In practice a material is usually characterized by some properties derived from the stress-strain curve, which are the E-modulus (Young's modulus), the yield stress σ_{0.2} (in England also σ_{0.1}), the ultimate tensile strength σ_{U} and the elongation after failure δ. These properties are indicated in Figure 5.1. It is not always recognized that such a diagram is actually a load-displacement curve drawn by the tensile testing machine. The "engineering" stress σ in figure 5.1 is defined by:

\[ \sigma = \frac{P}{F_0} \quad (5.1) \]

where P is the load and F_0 is the original specimen cross section.
before the test starts. With this definition the tensile strength is:

\[ \sigma_u = \frac{P_{\text{max}}}{F_0} \]  

(5.2)

It is true that \( P_{\text{max}} \) is the highest load which can be carried by the tensile specimen. However, at the moment that \( \sigma = \sigma_u \) failure does not occur. At this stress necking of the specimen will begin, whereas failure defined by disruption of the material starts later. After reaching \( \sigma = \sigma_u \) the specimen can be elongated further by a decreasing load. If \( P_{\text{max}} \) is not removed an unstable failure will occur (plastic instability).

A better impression of the material behaviour is obtained by plotting the "true" stress instead of the engineering stress. The true stress is defined by:

\[ S = \frac{P}{F} \]  

(5.3)

where \( F \) now is the instantaneous cross section. During elastic deformation \( F \) will be hardly smaller than \( F_0 \), but during plastic deformation the cross section will be reduced. Consequently \( S \) will be higher than \( \sigma \). The true stress \( S \) is also plotted in figure 5.1 and now the point where \( P = P_{\text{max}} \) (start of necking) is no longer an extreme point on the \( S-\varepsilon \) curve.

Necking

Rewriting (5.3) as:

\[ P = F S \]

we find by differentiation:

\[ \frac{dP}{d\varepsilon} = F \frac{dS}{d\varepsilon} + S \frac{dF}{d\varepsilon} \]  

(5.4)

The first term \( F \frac{dS}{d\varepsilon} \) is positive because \( dS/d\varepsilon \) (tangent modulus) is positive as a result of strain hardening. However, the first term is decreasing because both \( F \) and \( dS/d\varepsilon \) decrease. The second term \( S \frac{dF}{d\varepsilon} \) is negative because \( F \) becomes smaller. The absolute value
of the second term increases, and hence at a certain moment \( \frac{dP}{d\varepsilon} \) should become zero.

Afterwards \( \frac{dP}{d\varepsilon} < 0 \), and further elongation can be done by a decreasing load. Apparently the strain hardening then is insufficient to balance the decreasing cross section \( \frac{dF}{d\varepsilon} \) and the related increase of \( S \).

After \( \frac{dP}{d\varepsilon} = 0 \) has been reached further plastic strain is a local phenomenon, known as necking. The local character is a consequence of slight inhomogeneities of the tensile specimen. In the weakest cross section of the specimen (lowest strain hardening \( \frac{dS}{d\varepsilon} \), or smallest cross section \( F \), or both) the condition \( \frac{dP}{d\varepsilon} = 0 \) is fulfilled first and necking will start there with a decreasing \( P \). The other cross sections could be given further plastic deformations, but it would require a load slightly higher than \( P_{\text{max}} \) and that cannot occur anymore. Plastic deformation is then concentrated in the necked zone.

During plastic deformation the volume of the material may well be assumed to remain constant. With the cross section \( F \) and a length of the test section of the specimen \( L \) the constant volume implies:

\[
FL = F_0 L_0 \quad (5.5)
\]

\( F_0 \) and \( L_0 \) are the original values before testing. With:

\[
L = L_0 (1 + \varepsilon)
\]

equation (5.5) becomes:

\[
F(1 + \varepsilon) = F_0
\]

Differentiation gives:

\[
dF(1 + \varepsilon) + F \frac{dF}{d\varepsilon} = 0 \implies \frac{dF}{d\varepsilon} = -\frac{F}{1 + \varepsilon}
\]

Substitution of this result in equation (5.4) leads to:

\[
\frac{dP}{d\varepsilon} = F \frac{dS}{d\varepsilon} - \frac{F}{1 + \varepsilon} S \quad (5.6)
\]
and at the moment that \( P = P_{\text{max}} \rightarrow \frac{dP}{d\varepsilon} = 0 \), equation (5.6) gives:

\[
\frac{dS}{d\varepsilon} = \frac{S}{1 + \varepsilon} \tag{5.7}
\]

This equation gives the slope of the \( S-\varepsilon \) curve at the moment that \( P = P_{\text{max}} \). The point where this occurs can easily be constructed with the equation as shown in Figure 5.2. From this construction we now conclude: The point of the \( S-\varepsilon \) curve, where \( P = P_{\text{max}} \) and necking will start, is a function of the shape of the \( S-\varepsilon \) curve. In other words: It is depending on the strain-hardening behavior of the material. It thus should be concluded that \( P_{\text{max}} \) and the classical material tensile strength \( \sigma_U \) (= \( P_{\text{max}} / F_0 \)) are depending on the strain-hardening of the material. They do not indicate a cohesive strength of the material. The relation between \( \sigma_U \) and strain hardening explains why there can be a relation between \( \sigma_U \) and the hardness of the material, see later.

Until necking starts the plastic deformation was homogeneous over the length of the test section of the specimen. After necking has started further plastic deformation is concentrated in the necked area. The total elongation after failure is thus composed of two parts: a homogeneous one and an inhomogeneous one, see Figure 5.1. In the USA and the UK the elongation after failure is usually measured for an initial gage length \( L_o = 2 \) inch. If the material has a high thickness the inhomogeneous part of the plastic strain will be relatively large, see Figure 5.4. If the thickness is low it will be relatively small. Consequently a higher elongation is found for a thicker material. This inconsistency is removed by taking \( L_o \) proportional to the thickness. Ratio's which are used for round bar tensile specimens (diameter \( D \)) are:

\[
L_o = 5 \, D \quad \text{or} \quad L_o = 10 \, D
\]

For tensile specimens produced from sheet or plate material, having a rectangular cross section, an equivalent diameter \( D_{eq} \) is calculated from:

\[
F_0 = \frac{\pi}{4} \, D_{eq}^2 \quad \Rightarrow \quad D_{eq} = 1.13 \, F_0
\]

The same ratio's (5 or 10) can then be applied. Although these "proportional"
test specimens are better defined, it is still very usual for aircraft materials to take $L_0 = 2''$ or $L_0 = 50$ mm.

**Cohesive failure in a tensile test**

In a tensile test decohesion starts in the necked area in the center of the specimen. On a micro level voids are formed, especially at inclusions. The void formation is not a consequence of the necking, but rather of the large strain and stress in the minimum section. The voids will grow and coalesce. The failure spreads outwards and a full separation will occur, see figure 5.4. The last part of the separation, when the central failure approaches the outside surface, occurs by shear (under ~ $45^\circ$). This gives the so-called cup and cone fracture.

The start of the void formation will be postponed to a higher strain (and a further necking down) if no inclusions (or if smaller inclusions) are present. The elongation after failure is thus depending on the cleanliness of the material. A better indication of the cleanliness, however, is obtained from the reduction of area:

$$\psi = \frac{P_0 - F_{\text{min}}}{P_0}$$  \hspace{1cm} (5.8)

Because void formation usually starts after $P = P_{\text{max}}$ has been passed $\psi$ does not give a useful indication of the cleanliness of the material.

In some "quasi brittle" materials failure in a tensile test occurs before necking starts. Grey cast iron is an example, but the ductility of cast metals usually is lower than for wrought alloys. In the real brittle materials there is no plastic deformation at all, e.g. in ceramics. Some plastics also behave elastic until failure, although the elastic behaviour is not always linear.

**5.3 THE HARDNESS OF A MATERIAL**

We will not discuss all the various hardness measurement methods, but briefly refer to the relation between hardness and $G_U$. In a hardness test a steel ball, a diamond cone, or a diamond pyramid is indented into the material. The hardness gives an indication of the material resistance against such an indentation. Since the indentation is a
manner of plastic deformation the hardness is depending on the resistance against plastic deformation, i.e. on strain-hardening. As we have seen before $\sigma_U$ is also depending on strain-hardening. Consequently it is not strange that a relation between hardness and $\sigma_U$ seems to exist. None of the two properties is related to fracture, but rather to the strain-hardening characteristics of the material. This qualitative explanation for the correlation between hardness and $\sigma_U$ is not a quantitative explanation for the generally quoted relation:

$$\sigma_U \approx \frac{1}{3} H_V$$  \hspace{1cm} (5.9)

($H_V$ is the Vickers hardness) although small material dependent deviations of this relation do occur. Actually it is surprising the relation is so simple. It is convenient at the same time, because a simple hardness test can give us a reliable indication of $\sigma_U$. Hardness measurements are frequently used to check material quality and heat treatment. Quite often it can be done on production parts, without making these parts unusable (non-destructive quality inspection).

5.4 THE STATIC TENSILE STRENGTH OF A NOTCHED ELEMENT

A strip with a central hole will be adopted as a simple prototype of a notched element. The stress distribution in a strip with a hole was discussed in chapter 2 for the elastic behaviour, and in chapter 3 for a small exceeding of the yield stress. We now have to consider relatively large plastic deformations. Stress distributions for increasing load steps are shown in figure 5.5. If the load is high enough the full net section will be in the plastic range (net section yield). The stress distribution becomes much more homogeneous than for the elastic behaviour. The question now is at which moment will failure occur, or in other words: what is the fracture criterion?

At first sight it appears obvious that failure should occur when the stress at the edge of the hole is equal to the tensile strength:

$$\sigma_{\text{edge}} = \sigma_U$$  \hspace{1cm} (5.10)

However, after the discussion in section 5.2, it is not at all obvious that equation (5.10) is a good suggestion. It was shown that $\sigma_U$ has
no relation to fracture, but rather to the onset of necking. For a
tensile test specimen that implies a plastic instability because the
load $P_{\text{max}}$ could not be carried any longer. Such an instability will
not occur in the strip with a hole. Local necking at the root of a
notch may locally reduce the load transmission. This would imply that
the other part of the cross section, still the major part, has to
carry more load. This is possible because the stress level is somewhat
lower in that part.

A better criterion appears to be that failure will occur when $\sigma_{\text{edge}}$
is equal to the true failure stress $\sigma_y$ (fig. 5.1). Because $\sigma_y$ can be
considerably higher than $\sigma_y$, the average net section stress may also
be above $\sigma_y$. It would imply that the net section strength exceeds the
material tensile strength. Actually, it is difficult to indicate a
good criteria. Although some necking can sometimes be observed at the
root of a notch the restraint on necking is anyhow different (more
efficient) as compared to the unnotched tensile test specimen. More-
over, stress distributions as shown in figure 5.5 can hardly be calcu-
lated, because they involve large plastic deformations. Under such
conditions we better have a look at experimental data.

Figure 5.6 shows in a qualitative way the effect of stress concentration
on the strength of two wellknown Al-alloys, 2024-T3 and 7075-T6 (see
also figure 5.8). For the notched specimens the strength is defined
as the nominal stress (= average stress) on the net section. It is
normalized in figure 5.6 by dividing it by the material tensile strength.
Two remarkable trends are to be observed:

(1) For moderate $K_t$ values ($K_t = 1$ to 3) the notch effect on the
static strength is small. For $K_t = 2$ the strength of the notched
element can even be higher than the tensile strength. This trend has
also been observed for high strength steels,
especially if round specimens are notched by a
circumferential groove as shown on the left.
The nominal strength exceeded the tensile
strength by 20 to 40%. In this type of specimen the
lateral contraction, also if plastic deformation
occurs, is highly restrained by the much lower
stressed material away from the minimum section. Large plastic
defformations are impossible and the strength is relatively high. As said
before this restraint is lacking in a tensile test specimen.
(2) In figure 5.6 and 5.8 the more ductile 2024-T3 alloy appears to
be more sensitive to moderate notch effects than the lower
ductility alloy 7075-T6. Only for very sharp notches and cracks the
2024-T3 alloy is superior.

As indicated in figure 5.5 the stress distribution becomes fairly
homogeneous shortly before failure. It depends on the strain-hardening
behaviour how homogeneous the distribution will be. The strain-hardening
is different for the two Al-alloys, see figure 5.7. Let us assume that
the strains \( \varepsilon_A \) and \( \varepsilon_B \) will occur at the edge of the hole and the edge
of the strip respectively. The corresponding stresses are hardly different
for the 7075-T6 alloy, whereas they are more different for the 2024-T3
alloy. This suggests that we will have a more homogeneous stress
distribution in the 7075-T6 material than in the 2024-T3. This qualitatively
explains why the 7075-T6 is less sensitive to moderate notch effects.
It has a more flat \( \sigma-\varepsilon \) curve, the strain-hardening is less than in the
2024-T3 alloy.

Figure 5.6 also gives a line for a real brittle material (e.g. ceramic
materials). It can not accommodate any plastic deformation and static
notch effects are large. Fortunately, high-strength low-ductility
structural materials allow some plastic deformation. An elongation
of 10% obtained in a tensile test is fairly low, but it is not a low
value for a strong aircraft material. Apparently it is sufficient to
accept "blunt" notches in view of static strength. In the static strength
analysis of an aircraft structure the blunt notch effect is usually
covered by "allowable" stresses. These allowables at ultimate design
load can be in the order of the yield stress. Usually the risk of any
blunt notch effect is more than satisfactorily covered then. The
allowable stress levels in the aircraft industry are to be found in
structural handbooks.

5.5 THE FRACTURE TOUGHNESS \( K_{Ic} \)

If cracks are present the static strength of a component may be
significantly smaller. The material property to indicate the
sensitivity of a material for cracks under static loading is the
fracture toughness \( K_{Ic} \). As discussed in chapter 4 the \( K_{Ic} \) value is
the critical value of $K$, which leads to static failure of a mode I crack. It was emphasized in section 4.10 that a $K$ value can only give a good indication of the stresses around the crack tip if the plastic zone is very small. Small plastic zones are obtained if the crack remains in the plane strain condition, which will occur as long as the plastic zone size $r_p$ is much smaller than the material thickness. A $K_{IC}$-test (fracture toughness test) should therefore be carried out on a sufficiently thick specimen. To avoid high testing loads (and have the advantage of a small specimen) the so-called compact tension specimen was developed, see figure 4.5 of the previous chapter. The ASTM (American Society for Testing and Materials) has standardized the dimensions as shown in figure 4.5. A value of $W = 50$ mm (as an example) can be chosen. The specimen has to be tested in fatigue first (see ASTM Standard E 399-78) before it is pulled to failure in a static way afterwards.

During the static part of the test the load $P$ is measured as a function of its displacement $v$. Possible records to be obtained are shown in figure 5.9.

Before $P_{\text{max}}$ is reached a deviation from linearity is observed. There are two possible causes:

1. Plastic deformation occurs at the tip of the crack.
2. The crack is growing in a stable way until it becomes unstable at $P = P_{\text{max}}$.

If this occurs plastic deformation at the crack tip must have occurred also.

According to the ASTM standard the non-linearity should be limited to less than 5% at $P_{\text{max}}$ (see figure 5.9a). If it is more (fig. 5.9b) $P_5$ should be used to calculate $K_{IC}$. The formula to calculate $K_{IC}$ as given in the ASTM standard is:

$$K = \frac{P}{t\sqrt{W}} \cdot F(a/W)$$  \hspace{1cm} (5.11)

with:

$$F(a/W) = \frac{2 + a/W}{(1 - a/W)^{3/2}} \cdot \left[ 0.866 + 4.64 \left( \frac{a}{W} \right) - 13.32 \left( \frac{a}{W} \right)^2 + 14.72 \left( \frac{a}{W} \right)^3 - 5.6 \left( \frac{a}{W} \right)^4 \right]$$
The same formula in a different form was previously presented in figure 4.5.

Because a large plastic-zone invalidates the meaning of $K$ (section 4.10) a restriction on the size $r_p$ has to be specified. The requirement specified in the ASTM standard is:

$$2.5 \left( \frac{K}{\sigma_{0.2}} \right)^2 < t \quad (5.11)$$

(the ASTM standard uses the symbol $B$ for thickness instead of $t$ used here). The equation for $r_p$ (plane strain) presented in section 4.10 is:

$$r_p = \frac{1}{3\pi} \left( \frac{K}{\sigma_{0.2}} \right)^2 \quad (5.12)$$

Combining the two latter equations gives:

$$r_p < \frac{t}{7.5 \pi} \approx t/25 \quad (5.13)$$

As compared to the thickness $1/25$ is a small ratio indeed. After a fracture toughness test has been carried out it has to be checked whether $r_p$ was sufficiently small to produce a "valid" $K_{IC}$. Minimum thicknesses required to obtain valid results can be illustrated by the following data:

<table>
<thead>
<tr>
<th>$K_{IC}$ (test results) (MN/m$^{3/2}$)</th>
<th>$\sigma_{0.2}$ (MPa)</th>
<th>$t_{min} = 2.5 \left( \frac{K_{IC}}{\sigma_{0.2}} \right)^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2024-T3</td>
<td>40</td>
<td>360</td>
</tr>
<tr>
<td>7075-T6</td>
<td>27</td>
<td>470</td>
</tr>
</tbody>
</table>

A valid $K_{IC}$ test cannot be made on a thin sheet specimen, especially for more ductile materials like 2024-T3. It should be pointed out, however, that the ASTM requirement on thickness ($t > 25 r_p$) is very restrictive. Somewhat thinner specimens still can yield useful indications of $K_{IC}$.
$K_{IC}$ - material structure and $\sigma_{0.2}$

A static test on a material with a crack will still imply a static fracture mechanism. In the discussion on the tensile test it was pointed out that fracture starts with void formation. This is also true for a crack causing a static failure, although the voids now start at much smaller inclusions. In Al-alloys many very small intermetallic inclusions (0.1 to 0.2 \( \mu m \)) occur. Si and Fe are usually found in the inclusions. They are not visible for the optical microscope, but they can be observed in the electron microscope. During the fracture voids are formed around the inclusions, and this can be seen on the fracture surface as so-called dimples, see figure 5.10. The intermetallic inclusions can be reduced considerably in numbers and size by limitations on the chemical composition, see the table below (weight percentages).

<table>
<thead>
<tr>
<th></th>
<th>Zn</th>
<th>Mg</th>
<th>Cu</th>
<th>Zr</th>
<th>Cr</th>
<th>Mn</th>
<th>Cr</th>
<th>Si</th>
<th>Fe</th>
<th>Ti</th>
</tr>
</thead>
<tbody>
<tr>
<td>7075</td>
<td>5.5</td>
<td>2.5</td>
<td>1.5</td>
<td>0.3</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>clean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>alloys</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7475</td>
<td>5.7</td>
<td>1.75</td>
<td>1.5</td>
<td>0.06</td>
<td>0.1</td>
<td>0.12</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7050</td>
<td>6.2</td>
<td>2.3</td>
<td>2.3</td>
<td>0.12</td>
<td>0.10</td>
<td>0.04</td>
<td>0.12</td>
<td>0.15</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>averages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>maxima</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The beneficial effect of a cleaner chemical composition is shown below:

<table>
<thead>
<tr>
<th>Heat treatment</th>
<th>$K_{IC}$ (MN/m$^{3/2}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T651</td>
<td>30</td>
</tr>
<tr>
<td>T7651</td>
<td>44</td>
</tr>
<tr>
<td>T7351</td>
<td>34</td>
</tr>
</tbody>
</table>

Because inclusions are significant for fracture it should be expected that $K_{IC}$ will be dependent on the orientation of the fracture in a material, especially so if a fibrous structure is present. A fibrous structure is present in extrusions, in forgings and to a lesser degree in rolled material. For a compact tension specimen six different
orientations can be specified, see figure 5.11. The first capital of the code indicates the loading direction, the second one indicates the crack growth direction. As a result of the material production there is a tendency for inclusions to be arranged in chains in the longitudinal direction of the deformation process. For that reason the loading direction perpendicular to this direction is unfavourable, especially if it is loading in the short transverse direction. The effect on $K_{IC}$ is shown below.

<table>
<thead>
<tr>
<th>Orientation</th>
<th>$K_{IC}$ (MN/m$^{3/2}$)</th>
<th>Fracture direction</th>
</tr>
</thead>
<tbody>
<tr>
<td>L-T</td>
<td>34</td>
<td>perpendicular to fibres</td>
</tr>
<tr>
<td>T-L</td>
<td>23</td>
<td>in fibre direction</td>
</tr>
<tr>
<td>S-L</td>
<td>21</td>
<td></td>
</tr>
</tbody>
</table>

Material: 7075-T6511 (AGARDograph No. 176)

For most alloys the mechanical properties are heavily depending on the heat treatment. A certain correlation between different mechanical properties should then be expected. A wellknown correlation is: increase of $\sigma_{0.2}$ and $\sigma_U$ reduces the ductility. There are more unfortunate correlations between high strength and other properties. High-strength low-ductility alloys usually have a lower resistance against fatigue notch effects and fatigue crack propagation. There is also a tendency to lower $K_{IC}$ values. A large amount of $K_{IC}$ data drawn from Ref. 1 were found to be scattered over the shaded area's in figure 5.12. Although for each alloy system (Al-alloys, Ti-alloys, steels) there is no unique correlation between $\sigma_{0.2}$ and $K_{IC}$ the general trend still is a reduction on $K_{IC}$ for higher $\sigma_{0.2}$. Apparently a higher strength to be used for a lower weight of the aircraft structure has to be paid for by more sensitivity for fatigue and for cracks. This can only be acceptable if the aircraft designer can improve the quality of the structural design.

Application of $K_{IC}$

$K_{IC}$ values can be used to compare alternative materials in view of their sensitivity for cracks under static loading. Strong materials with a low $K_{IC}$ value can be dangerous if fatigue cracks are initiated. A simple example can illustrate this point. Let us calculate the critical crack length $a_c$ that would cause failure at a stress level equal to 50 percent
of $\sigma_{0.2}$. For simplicity it is assumed that the geometry factor in $K = C_s V\pi a$ is approximately equal to 1. The value of $a_c$ then follows from:

$$K_{IC} = 0.5 \sigma_{0.2} \sqrt{2\pi a_c}$$

For four different alloys the result is:

<table>
<thead>
<tr>
<th>Alloy</th>
<th>$\sigma_{0.2}$ (MPa)</th>
<th>$K_{IC}$ (MN/m$^{3/2}$)</th>
<th>$a_c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2024-T3</td>
<td>360</td>
<td>40</td>
<td>15.7 mm</td>
</tr>
<tr>
<td>7075-T6</td>
<td>470</td>
<td>27</td>
<td>4.2 mm</td>
</tr>
<tr>
<td>Ti-6Al-4U</td>
<td>1020</td>
<td>50</td>
<td>3.1 mm</td>
</tr>
<tr>
<td>4340 steel</td>
<td>1660</td>
<td>58</td>
<td>1.55 mm</td>
</tr>
</tbody>
</table>

Small cracks in very strong materials can apparently have large effects. The situation is better if plane stress prevails at the crack tip. The plastic zone is larger and in the plastic zone the peak stress is leveled off much more. However, in thick sections we should be aware of the risks of cracks, and many thick sections occur in an aircraft structure, e.g. spar caps, forgings, undercarriages, and even a thick skin on the wing of a Jumbo aircraft. In those cases it may be questionable whether high-strength low-ductility materials should be applied or not.

5.6 CRACKS IN SHEETS, BEHAVIOUR UNDER STATIC LOADING

Fracture toughness tests carried out on specimens of different thicknesses indicate the trend shown in figure 5.13. In this figure $K_C$ is based on $P_{max}$ (fig. 5.9) also for the thinner material. Going from a large thickness to a lower thickness the $K_C$ value initially remains constant, even at values below the minimum requirement of the ASTM. However, at still lower thicknesses $K_C$ increases. Cross sections of the fracture surface show the so-called shear lips. The shear lip width is fairly independent of the material thickness. Actually the shear lips should be associated with the presence of a free surface (the two plate surfaces). At a free surface the crack front is (per definition) in plane stress and the restraint on fracture by shear is practically absent as compared to the center
part of the crack front.

For a lower thickness the shear lips occupy a relatively larger part of the thickness. If the thickness is low enough the full fracture will occur in shear under approximately 45°. At that thickness the maximum of $K_C$ is reached. For Al-alloys the thickness then is in the order of just a few millimeters. The failure is no longer a mode I crack but a mixed mode I/III crack. The higher $K_C$ values thus obtained do not satisfy anymore the ASTM thickness requirement, and we should speak about "invalid" $K_{IC}$ values. Since it is no longer a mode I fracture the symbol used now is $K_C$ instead of $K_{IC}$.

The compact tension specimen is not well suited for testing thin sheet material because it becomes unstable under load (buckling phenomena). A better specimen then is a sheet specimen with a central crack (fig. 5.14). Also in this case buckling of the crack edge is possible, because the crack edge is under compression in the X-direction (section 4.3). For thin sheets and long cracks it will try to curve outwards. However, this does not lead to unstable behaviour. Moreover in experiments it can easily be prevented if desirable.

During a static test on a sheet specimen with a central crack relatively large plastic zones develop at the crack tips. A plane stress situation through the entire thickness along the crack front will be obtained. At the same time the crack will grow in a stable manner under increasing load, see figure 5.14. After some crack extension it becomes unstable at a stress $\sigma_C$ and a complete failure occurs. Especially in ductile alloys (such as 2024-T3) the stable crack extension can easily be observed with the naked eye. We now may consider (1) two crack lengths and (2) two stress levels: (1) the initial crack lengths $a_0$ and the critical crack length $a_C$ at which the unstable failure occurred, and (2) the stress level $\sigma_i$ at which stable crack extension started and the critical stress level $\sigma_C$ at which unstable crack growth took over. Note that all stress levels considered here are gross stresses, calculated for the uncracked sheet panel width. In figure 5.15 results of a number of tests with different $a_0$ values have been plotted. Line 1 gives $\sigma_i$ as a function of $a_0$, line 3 gives $\sigma_C$ as a function of $a_C$. For a designer line 2 ($\sigma_C$ as a function of $a_0$) appears to be of more interest, because it gives the failure stress as a function of the
initial damage.

It would be helpful if the empirical results of thin sheets would agree with some calculation rule. This is checked in figures 5.16 and 5.17 for test data of the National Aerospace Laboratory NLR (Amsterdam), the specimen width being 120, 300 and 540 mm respectively. The stable growth initiation stress \( \sigma_1 \) was defined as the stress for \( \Delta a = 1 \) mm because lower \( \Delta a \) values could not accurately be observed in the plastic zone. The results of 7075-T6 could well be represented by a single \( K_1 \) value, where:

\[
K_1 = C \sigma_1 \sqrt{\pi a_1}
\]

(with the geometry factor \( C = \sqrt{\sec \pi a_1 (W)} \)). In view of the width effect in \( C \) constant \( K_1 \) lines are different for different \( W \) values. The empirical results follow the lines fairly well. The same is true for the 2024-T3 results for \( W = 300 \) and \( W = 540 \) mm. However for \( W = 120 \) mm the test results are noticeably below the \( K_1 \)-line. Feddersen (Ref. 4) has contributed this result to net section yield. The dashed lines in figure 5.16 indicate at which stress net section yield will occur:

\[
\sigma = \sigma_{0,2} \left( 1 - \frac{2a_1}{W} \right)
\]

In the wide panels \( \sigma_1 \) is below the stress for net section yield. However, for the small width \( (W = 120 \) mm) and 2024-T3 it turns out that \( \sigma_1 \) is above the net section yield stress. If that is true small scale yielding is no longer applicable and it cannot any longer be expected that \( K \) can be meaningful for data correlation.

For the higher stress level \( \sigma_c \) (\( \sigma_c > \sigma_1 \)) even more deviations from a constant \( K_c \) behaviour should occur. For the low ductility alloy 7075-T6 the data point are still surprisingly well covered by a constant \( K_c \), at least for \( W = 300 \) and \( W = 540 \) mm. However, for the 2024-T3 alloy all \( \sigma_c \) values are close to the corresponding net section yield stress. \( K \) values are of little use under these conditions.

It is of some interest to recapitulate a few \( K \) data (MN/m\(^{3/2}\)):
Sheet material

\[
\begin{align*}
\text{Stable crack initiation } K_i & : \quad 73.7 \quad 57.7 \\
\text{Unstable cracking } K_c & : \quad - \quad 70.6 \\
\text{Valid } K_{IC} \text{ for thick material } K_{IC} & : \quad 40 \quad 27
\end{align*}
\]

Two things are clear: (1) Sheet materials with crack tips in plane stress are more crack tolerant than thick material with crack tips in plane strain. (2) The 2024-T3 alloy is more crack tolerant than the 7075-T6 alloy.

The Feddersen-diagram

For very small cracks a constant \( K \) for either stable crack initiation or final failure predicts an unrealistically high strength. For \( a \to 0 \) the result is \( \sigma \to \infty \). Even for a low ductility material large scale yielding will occur. Feddersen (Ref. 4) has suggested the solution shown in figure 5.18. A line of constant \( \sigma \sqrt{\pi a} \) is indicated. (For unknown reason Feddersen ignores the width correction on this line.) Then two tangent lines are drawn to this line, AB starting at \( \sigma = \sigma_{0.2} \) and \( a = 0 \), and DC starting at \( 2a = W \) and \( \sigma = 0 \). It is easy to prove that \( \sigma \) at B is equal to \( 2/3 \sigma_{0.2} \) and \( 2a \) at C is equal to \( W/3 \). The Feddersen diagram thus consists of the line ABCD.

The diagram can be drawn up if one data point of the line BC is available. A "valid Feddersen result" thus requires that \( \sigma_c \) (or \( \sigma_{0.1} \)) \( < 2/3 \sigma_{0.2} \) for \( 2a < W/3 \). The Feddersen diagram has no rigorous derivation as a background, but it is in approximate agreement with test results, accept for test panels with a small width. In the latter case \( \sigma_c \) is too close to net section yield for ductile alloys) and the Feddersen diagram would be too optimistic for wide panels. This can be seen in figure 5.16 for 2024-T3. Therefore Feddersen added a second requirement for a valid test result: \( \sigma_{net} < 0.8 \sigma_{0.2} \).

This is not satisfied by the 2024-T3 data for \( W = 120 \text{ mm} \).
REFERENCES

Chapter 6
STRESS CORROSION

Contents:

6.1 Introduction
6.2 The stress corrosion process
6.3 Stress corrosion tests, variables and purpose of tests
6.4 How to prevent stress corrosion in aircraft structures
6.5 Hydrogen embrittlement
References

6.1 INTRODUCTION

By definition stress corrosion is a failure mechanism which requires the cooperation of corrosion and a static tensile stress. The definition implies that corrosion alone would not have led to cracks. Also the presence of the tensile stress alone would not have induced a crack. However, only the simultaneous action of corrosion and a tensile stress produces a crack and ultimately a complete failure.

Also for stress corrosion it does not make any difference whether the tensile stress is a result of an external load or whether it is an internal stress (residual stress, built-in stress, see chapter 3). However, it is significant that the internal stresses are present all the time, whereas the external stress only when the component is under load. Most stress corrosion problems in service are caused by internal stress systems. For instance it occurred that a large landing gear of an Al-alloy forging spontaneously cracked over a very large length. The residual stresses remaining from the heat treatment caused the fracture. In figure 6.1 a stress corrosion crack is shown occurring in the eye head end of a stabilizer spar. The lug was holding the axle of an all-moveable tailplane. In the lug a bush was pressed and that set up tangential tensile stresses at the bore of the hole, which caused the stress corrosion crack.

In this chapter we will first discuss the stress corrosion mechanism. The principles of methods for stress corrosion tests are indicated afterwards. Finally the prevention of stress corrosion cracking is reviewed.

6.2 THE STRESS CORROSION PROCESS

The stress corrosion process occurs in two phases:
(1) crack nucleation period;
(2) crack propagation period.

In the nucleation period corrosion is the predominant factor. Corrosion starts at the material surface. It is recognized to be wet corrosion, i.e. an electrolytic process for which an electrolyte is essential. Both an anodic and a cathodic reaction should occur, which requires different electric potentials in the same material. Such differences can occur on a micro level. A well-known source is the grain boundary zone, which can develop a potential different from the potential of the matrix material in the grains. Grain boundary attack can be the result. Different potentials can also be caused by variations in the electrolyte (e.g. O₂-contents). This is supposed to apply in crevice corrosion and in pitting corrosion. Corrosion is a rather complex process, which moreover is depending on the type of material, material structure and the environment (electrolyte and wetting conditions). In the nucleation period corrosion is still a superficial process, but on a micro level penetration of corrosion into the material does occur. In many cases this will occur along the grain boundary. It is believed that in the nucleation period the effect of the tensile stress is not significant. However, it is possible that a tensile stress will affect the electric potential of the material.

As soon as the corrosive attack has reached the stage of a microscopic notch a different situation exists.
(a) The micronotch will induce a stress concentration at its tip. This may change the anodic potential at the tip and it can accelerate local corrosion. The tensile stress then becomes important.
(b) The tensile stress will also open the micronotch thus giving access to the electrolyte. The notch can act as a crevice and thus create even more unfavourable corrosion conditions.

If these conditions are present the superficial corrosion will proceed to a further penetration into the material. The crack propagation period has started. The tensile stress is important because the crack is growing faster at a higher stress level. The tensile stress opens the crack which allows the electrolyte to enter the crack until the crack tip. The tensile stress can also cause some plasticity in the crack tip area, which will enhance the anodic nature of the material. The main contribution of the tensile stress is to increase the crack growth rate. It is not fully clear how the tensile stress accelerates
the crack growth mechanism. In some materials a stress corrosion crack is growing more or less continuously, but in other materials it appears to occur in small jumps. Anyhow, the tensile stress implies a strain energy concentration at the crack tip, and this energy will stimulate the decohesion process.

An accurate definition of the end of the nucleation period and the beginning of the propagation cannot be given. However, it may be stated that the nucleation period is mainly corrosion controlled, whereas the propagation period is both stress and corrosion controlled. The difference between the controlling conditions is highly evident for some materials, which have a high resistance against surface corrosion, but a low resistance against stress corrosion cracking. This applies to some Ti-alloys. The reverse situation can also occur. A good resistance against stress corrosion does not guarantee that corrosive surface attack will not occur.

**Microscopic and macroscopic features of stress corrosion cracking**

On a microscopic level stress corrosion cracking occurs along the grain boundaries in several structural materials. This is especially true for Al-alloys and low-alloy high-strength steel (prior austenite grain boundaries). An example of the intergranular cracking is shown in figure 6.1c and d, where this behaviour is revealed both by optical microscopy (cross section) and electron microscopy (replica of fracture surface). A second microscopic feature is crack tip branching. Growing around a grain the crack will meet triple points. Growth can occur in two directions then, until the larger one becomes dominant. Growth of the shorter one stops because it is no longer under stress. Transgranular crack growth occur in some materials, e.g. stainless steel and Ti-alloys. However, the reputation of stress corrosion problems in aircraft structures is mainly associated with Al-alloys and steel.

Stress corrosion cracking can occur at low stress levels. As a result it usually occurs without macroscopic plastic deformation. As shown by figure 6.1 cracks can be quite long and sometimes they lead to complete failure. In other cases the stress corrosion crack releaves the tensile stress, for instance if it is a locally built-in stress.
Many stress corrosion cracks have occurred in Al-alloy forgings and extrusions. These materials usually have a pronounced fibre structure. Stress corrosion cracks are growing relatively easy in the fibre direction, see figure 6.1. The stress corrosion resistance for a tensile stress perpendicular to the fibres is significantly lower than for a tensile stress in the direction of the fibre. The stress corrosion resistance is highly anisotropic. For a crack growing along the grain boundaries the crack path is much more easily a continuous path in the fiber direction. In the transverse direction, a zig-zag path has to be followed which will make crack growth more difficult.

6.3 STRESS CORROSION TESTS; VARIABLES AND PURPOSE OF TESTS

Two different ways of presenting results of stress corrosion tests are shown in figures 6.2 and 6.3. In figure 6.2 unnotched specimens are subjected to a constant load in a corrosive environment in order to measure the stress corrosion life $L$. The life is plotted (usually logarithmic) as a function of the applied stress. Similarly to fatigue life curves there is a lower limit ($S_\theta$) below which stress corrosion failures do not occur. This threshold value may be considered to be characteristic for the specific combination of material and environment. Usually $S$-$L$ curves show a fairly abrupt bending to the horizontal asymptote $S_\theta$. The life $L_\theta$ at which this occurs can vary considerably. It may be as small as 10 hours, but in less aggressive environments it also can be 1000 hours.

In figure 6.3 a precracked compact tension specimen is carrying a constant load. The crack growth rate $da/dt$ (mm/hrs) is measured and plotted as a function of the stress intensity factor $K$. As should be expected larger $K$-values will give higher crack rates. The relation has a somewhat sigmoidal shape with two vertical asymptotes. The left one is the more important one because it is associated with the minimum $K$-value required for stress corrosion crack growth. The symbol for this value is $K_{ISC}$, where I refers to the mode I crack opening and SCC to stress corrosion cracking. Also this value can be considered to be characteristic for a specific material/environment combination. The second vertical asymptote is related to static failure, i.e. the $K_{IC}$. Between these two values a tendency towards a horizontal plateau of the $da/dt$-$K$ curve is sometimes observed. At this plateau $da/dt$
is independent of K. It is believed that crack growth in this area is
the result of a chemical "embrittling" process, which just requires
time independent of the stress intensity.

Apparently there are two different types of tests:
(1) Tests which include the nucleation period.
(2) Tests on precracked specimens, which eliminates the nucleation period.
In these tests stress corrosion crack growth is considered only.
Starting from the idea that stress corrosion should be avoided, the
two relevant characteristics are \( S_0 \) and \( K_{ISCC} \). Even then there are
several variables involved associated with (1) the material and (2)
the environment. Some aspects of both are listed below.

**Material**
- type of alloy and heat treatment
- material surface condition
- material structure, e.g. fibre orientation.

**Environment**
- type of environment - gaseous (e.g. humid air)
  - wet (water, sea water, solutions)
  - type of exposure - continuous (e.g. outdoors exposure)
  - interrupted (alternately wet and dry)
- temperature.

Apparently the number of variables which can affect the stress corrosion
behaviour is large. We therefore should seriously consider the question
which kind of information we want to obtain.

Practical experience has shown that stress corrosion can occur under
normal atmospheric conditions. We should realize, however, that water
vapour is always present and in view of condensation the normal
atmosphere cannot really be considered to be a dry environment. An
obvious approach then appears to be to perform tests in an environment
which is representative for severe service conditions. Unfortunately
corrosion is such a complex phenomenon that it is difficult to
specify a realistic test environment.

A second problem is related to the interpretation of the results of
stress corrosion tests. It is easy to state that the threshold levels
\( S_0 \) and \( K_{ISCC} \) should be sufficiently high to be sure that stress
corrosion problems will not arise in service. Unfortunately it is
difficult to indicate quantitatively what a sufficiently high level
should be. On the other hand, it is quite well possible to compare the stress corrosion behaviour of different materials by the results of stress corrosion tests (see e.g. Ref. 1). An illustration of different stress corrosion properties is given in the following table for Al-alloys.

<table>
<thead>
<tr>
<th>Alloy</th>
<th>Threshold stress level $S_0$ (MPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L</td>
</tr>
<tr>
<td>2024-T3 plate</td>
<td>250</td>
</tr>
<tr>
<td>T6 plate</td>
<td>&gt;350</td>
</tr>
<tr>
<td>T3 extrusion</td>
<td>&gt;350</td>
</tr>
<tr>
<td>T6 extrusion</td>
<td>&gt;400</td>
</tr>
<tr>
<td>7075-T6 plate</td>
<td>350</td>
</tr>
<tr>
<td>extrusion</td>
<td>420</td>
</tr>
<tr>
<td>forging</td>
<td>250</td>
</tr>
<tr>
<td>7075-T73 forging</td>
<td>&gt;350</td>
</tr>
<tr>
<td>7075-T6 forging</td>
<td>&gt;350</td>
</tr>
</tbody>
</table>

L = longitudinal direction = fibre direction  
LT = long transverse direction  
ST = short transverse direction, usually thickness direction

The most striking result of the table is the low threshold for several materials if loaded in the short transverse (ST) direction. For the 2024 alloy artificial aging (T6) is apparently better than room temperature aging (T3 and T4). For the 7075 alloy averaging (T7) gives good results also in the ST direction, whereas $S_0$ is low for the T6 aging condition. A comparison between materials can thus be made. However, it is clear that the loading direction with respect to any directionality of the material structure should be taken into account.

**Test specimens**

Tests which should include the nucleation period, still allow a variety of specimens. Examples are shown in figure 6.4. The C-ring specimen (fig. 6.4a and b) are loaded in bending by a bolt. Bolt tensioning in figure 6.4a implies a constant bending deformation. If cracks
are nucleated at A the bending stiffness is reduced and crack growth will go slower. Even crack arrest can occur, especially if there is crack branching. In figure 6.4b the load will remain approximately constant if cracking occurs, because there is a spring between the nut and the C-ring. The advantage of the C-ring specimen is that it can be made with a small outer diameter. This allows specimens to be taken from plate material, which still can be loaded in the ST-direction, see figure 6.4. It will be clear that corrosion between the bolt and the specimen is not relevant to the stress corrosion sensitivity of the material. For that reason electrical insulation between bolt and specimen has to be taken care of. It can be done by coatings or by plastic bushes, etc. A relation between bolt tightening or specimen deformation and the bending stress has to be established first. Formula's are available, but a calibration with strain gages can also be made.

The idea of the tuning fork specimen (fig. 6.4c) is similar to that of the C-ring specimens. However, it is still smaller. The bending stress is controlled by clamping the specimen on a close tolerance plate, which implies a constant deformation. Figure 6.4d shows a specimen loaded in tension by the elastic deformation of a ring. Cracking of the specimen will give some reduction of the load, but if cracking starts complete failure is to be expected. If we want to keep the load fully constant dead weight lever mechanisms can be used (figure 6.4g). A simple method to give sheet material or a hat stiffener a constant bending deformation is shown in figures 6.4e and f.

Specimens for tests on precracked specimens are characterized by:
(1) increasing K;
(2) constant K;
(3) decreasing K.

The compact tension specimen (figure 6.3) has an increasing K. For a stress corrosion test this implies that the test result is either a crack rate as a function of K, or more simple a crack growth life until failure. Although such data can be used for comparison it does not look very satisfactorily.

A specimen with a constant K has been proposed by Mostovoi, see figure 6.5. It is tapered in such a way that for a growing crack the K-value remains constant (for a constant load P).
If different specimens are tested at different loads $da/dt$ values are obtained for the corresponding $K$-levels. A threshold $K$-value can be obtained as the maximum $K$ for which crack growth does not occur.

The so-called double cantilever beam specimen (DCB-specimen) is an interesting specimen for obtaining a decreasing $K$ if the crack is growing longer (see figure 6.6). The specimen is loaded by two bolts, which are screwed into the specimen and which in this way cause a deflection. During the stress corrosion test the deflection ($v = \text{total deflection at bolt centerline}$) remains constant. If the crack is growing the stiffness of the specimen will decrease and the load between the bolts will also reduce. As a result $K$ will become smaller for longer cracks. The decreasing $K$ implies that the crack growth will also occur slower and slower, until crack growth stops. At that moment the $K$-value has reached the threshold value $K_{\text{ISC}}$.

To facilitate precracking the DCB specimen has a chevron notch. Precracking can be done statically by the bolts, which is a quasi-static way. A better method is to do it by fatigue. Subsequent crack growth under stress corrosion conditions can still give some problems. The crack length has to be measured from time to time. This can be done at the specimen surfaces only. The real crack can be longer if a curved crack front is present (see fig. 6.6). After crack arrest the specimen can be opened by static loading and the fracture surface then will show whether the last crack front was curved. Sometimes the crack growth path deviates from the horizontal line to follow a path indicated by $A$ in figure 6.6. The risk of such a deviating path will depend on the fibrous structure of the material and the orientation of it. The usual procedure to avoid such deviations is to provide the specimens with side grooves. The same procedure has been applied to Mostovoi specimens as well. The grooves are guiding crack growth in the desired direction.

Environmental conditions in the tests

There are several environments which have been used in tests:

1. Exposure to normal air.
2. Outdoors atmospheric exposure. Free access to specimens for rain, sun, air contaminants.
3. Full immersion into water, salt water, artificial sea water, other solutions (oxygen saturation, pH and $T$ as variables).
(5) Salt spray cabinet. Atmospheric exposure near the sea coast appears to be fairly realistic for severe service conditions. However, the tests may take a very long time. Acceleration of the test is possible by going to more severe environments. Both alternate immersion and the salt spray cabinet are severe. Alternate immersion appears to be more severe than full time immersion. A method applied for alternate immersion is using a Ferris wheel to which small specimens (tuning fork specimen) are clamped. The wheel is partly submerged. It rotates slowly, e.g. one revolution per hour, which can imply an immersion time of 10 minutes per hour for the specimen.

Instead of using a more severe environment another way to accelerate a stress corrosion test is to increase the load. An interesting development is to test with a slowly increasing load during the test, which ensures that failure will occur in a reasonable time. Usually this is done in such a way that the increasing load gives a constant strain rate. This procedure is referred to as the constant-strain rate method (Ref. 2). Small unnotched specimens are loaded in tension. The test result should give an indication of the sensitivity for stress corrosion. A good correlation with the result of constant load tests and of tests on C-ring specimens was observed in Ref. 2.

Test results
The results of tests for a certain material and a certain orientation of the structure of the material are depending on the environment used in the test. This can offer problems with respect to the interpretation of the results. In many cases stress corrosion tests are done for comparison with other materials in order to rank the materials in an increasing order of stress corrosion resistance. The ranking order can be different for different environments and the interpretation always should be done with due consideration to the real environment in service. This can be quite difficult. Sometimes it can be desirable to obtain either $S_o$ or $K_{ISCC}$ data. Assuming that crack nucleation is not always avoidable the more significant property is $K_{ISCC}$. Especially for very high strength materials it is desirable to know the sensitivity of the material for the presence of cracks (and unintentional flaws) in a corrosive environment.
6.4 HOW TO PREVENT STRESS CORROSION IN AIRCRAFT STRUCTURES

Preventive action against stress corrosion can be taken by:
(1) Material selection.
(2) Avoidance of permanent tensile stresses.
(3) Corrosion protection.
(4) Maintenance.

Material selection

It is almost trivial to say that materials with a low stress corrosion resistance should not be used. However, it has taken many years before it was clear to aircraft designers, that the strong Al-Zn alloys (7000 series) in the T6 condition (artificially aged to maximum strength) were very sensitive to stress corrosion. This was especially true if the material had a pronounced fibrous structure, and if machining caused a so-called end grain structure. Figure 6.7 shows a schematic picture of a die forging with the grain fibre orientation. The excess material along the parting line of the dies has to be removed anyway, but further machining is necessary to arrive and the final shape. This implies that the fibres which originally followed the contours of the dies are cut which gives an end grain structure, see figure 6.8 for an illustration.

The end grain structure facilitates the corrosion penetration along the grain boundaries considerably. Most stress corrosion failures in aircraft structures occurred in Al-alloy forgings, sometimes in extrusions starting at an end grain structure. Figure 6.1 is an example and another example is given in figure 6.9.

It is very surprising, but lucky at the same time, that overaging of the Al-Zn alloys (e.g. 7075, 7079) until the T7 condition has shown to improve the stress corrosion resistance considerably. In Al-alloys with a precipitation hardening the precipitation in the grain boundary zone differs from the precipitation in the matrix of the grains (see figure 6.10). These differences are making the grain boundary zone susceptible to corrosive attack. Apparently, if the aging treatment is continued beyond maximum hardness (T6) until some averaged condition (T7) the electrical potential differences between the grain boundary zone and the matrix become much smaller. The sensitivity to stress corrosion cracking is considerably reduced.

It should be realized that the static strength in the T7 condition is somewhat lower (overaged!):
<table>
<thead>
<tr>
<th></th>
<th>(\sigma_{0.2}) (MPa)</th>
<th>(\sigma_u) (MPa)</th>
<th>(\delta)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7075-T6</td>
<td>500</td>
<td>570</td>
<td>11%</td>
</tr>
<tr>
<td>7075-T7</td>
<td>430</td>
<td>500</td>
<td>13%</td>
</tr>
</tbody>
</table>

However, it is much wiser to pay this penalty if components will be designed as a forging. The same is true for extrusions if there is a possibility for permanent tensile stresses transverse to the longitudinal direction.

**Permanent tensile stresses**

As said before permanent tensile stresses should be avoided. In chapter 3 two important sources of these stresses were mentioned: (1) residual stresses of plastic deformation (production process or heat treatment) and (2) built-in stresses. The first source is usually recognized only after stress corrosion failures occur in service. An old remedy is to apply a local shot peening in those areas where cracks were found. In the design stage a better solution is to select a stress corrosion resistant material. At the same time it is worthwhile to consider the question whether residual stresses are introduced as a result of quenching forged products. Improved quenching techniques can sometimes be developed. Some attention to the grain flow inside a forging is also useful. By changing the shape of the dies and the position of the parting lines improved grain flow is possible.

**Design aspects**

The above aspects are material aspects. The avoidance of built-in stresses is a matter of design in some cases and tolerances on dimensions in other cases. An example of the latter was given in chapter 3 (fig. 3.10). An almost classical and simple example of a design error is shown in figure 6.11. Another example is the corner fitting in figure 6.12 with several machined mating surfaces. A slight mismatch between the fitting and the components, to which it will be connected, will cause built-in bending stresses during assembling. One flange cracked along its entire length as a result of stress corrosion.

For components with a pronounced fibrous structure, loading in the transverse direction should be avoided, both by external loads and by built-in stresses.

**Corrosion prevention**

Corrosion prevention is covering different aspects:
- corrosion resistant materials;
- surface treatments (anodic coatings, primers, painting);
- interfaying sealants;
- drainholes;
- avoidance of water collecting configurations (water traps).

These aspects will not be discussed here. The quality of the production is important for the success of preventive measures.

**Maintenance**

Poor maintenance can of course lead to corrosion problems. However, the resistance against stress corrosion of an aircraft structure is mainly the responsibility of aircraft design and production. Material selection, surface treatments, sealants, avoidance of permanent tensile stress, structural details, they are all settled in the aircraft industry. Good maintenance will help to indicate corrosion problems at an early stage when relatively simple solutions might still be possible. Stress corrosion cracks growing in the fibre direction are not always dangerous. For instance, in the spar shown in figure 6.7 a small crack growing in the fiber direction will hardly affect the strength of the spar. If the crack is small it still can be removed by carefully taking away the material around the crack, followed by surface treatments to prevent further stress corrosion action.

**6.5 HYDROGEN EMBRITTLEMENT**

Hydrogen embrittlement is a phenomenon with a bad reputation. It can occur in low-alloy high-strength steels which are used in aircraft structures. It may be considered to be a special case of stress corrosion because both tensile stress and corrosion are involved in hydrogen embrittlement. Crack growth is observed although it has been found to occur initially as a step by step process until final failure. As a result of surface corrosion hydrogen ions (H⁺) are formed. These ions can easily diffuse in steel. Depending on the hydrogen concentration inside the material it will be embrittled. For sharp notches and cracks, where high stresses are present, local failure will lead to crack initiation and crack extension. Corrosion can thus be dangerous for high-strength steels. Such steels are frequently cadmium plated as a corrosion protection. This is usually an electrolytic process and hydrogen ions can penetrate the steel during cadmium plating. It is necessary to drive out the ions afterwards by a baking cycle.
Hydrogen embrittlement is a potential risk if high tensile stresses are permanently present in a component. A well-known example is the pretensioned bolt, which can fail as a result of hydrogen embrittlement without any other additional loading.

REFERENCES


Textbooks on corrosion


Publications on aircraft corrosion

- The Theory, Significance and Prevention of Corrosion in Aircraft. AGARD Lecture Series No. 84, 1976.

Publications on stress corrosion


ASTM Standards on Testing

Chapter 7

MATERIAL FATIGUE

Contents:
7.1 Introduction
7.2 The fatigue process in different phases
7.3 The fatigue mechanism
7.4 Characteristics of a fatigue fracture
7.5 Factors which influence fatigue behaviour
7.5.1 Stress amplitude and mean stress fatigue curves and fatigue diagrams
7.5.2 Environmental effects
7.5.3 Some surface effects
7.5.4 Fretting
References

7.1 INTRODUCTION

Fatigue is a failure mechanism that occurs as a result of a large number of load applications. A single load cycle will do no harm to the material or the structure, because the load is far below the static failure load. However, if this is repeated many times a fatigue failure can occur. Fatigue is a failure mode, that has to be considered for structures and vehicles, which are subjected to many load cycles during their economical life. This applies to aircraft structures.

Fatigue in aircraft structures is an important problem for different reasons:
(a) Catastrophic accidents due to fatigue are still occurring in the present time, both in aircraft and helicopters.
(b) Nowadays aircraft are used more intensively than in the past (more hours per day). Moreover the economic life of an aircraft has increased also (say 20 years now and 10 years in the past).
(c) If fatigue cracks occur repairs or modifications are necessary. This is economically undesirable (also in view of down time).
(d) Aircraft operators are putting pressure on larger inspection periods (for civil aircraft one major inspection in two years instead of once every year).
(e) New trends in developments of lighter aircraft structures can partly be realized by using stronger materials. Unfortunately, a higher strength usually implies a lower ductility and an increased sensitive for fatigue and cracks.

The above list illustrates that contradictory aspects are involved. In general terms the contradiction is:
safety vs economy

The economy covers both aircraft production (industry) and aircraft utilization (the operator). However, it is in the design stage of the aircraft that the quality with respect to fatigue will primarily be determined. Consequently the aircraft designer has to pay due attention to the fatigue quality of the aircraft structure.

Aircraft fatigue involves a number of different aspects:
(1) material selection;
(2) surface treatments (crack initiation);
(3) structural lay-out (fail-safe design);
(4) detail design and joints (fatigue life);
(5) fatigue crack propagation and fail-safe strength (inspection periods and safety);
(6) fatigue loads in service;
(7) full-scale fatigue testing.

In order to deal with these kind of problems a basic understanding of the fatigue process is essential. This is a subject of the present chapter.

7.2 THE FATIGUE PROCESS IN DIFFERENT PHASES

If many load cycles are necessary for a fatigue failure it is a good question to see what happens inside the material during the fatigue life. We will divide the life in a number of phases:

- cyclic slip
- crack nucleation
- growth of microcrack
- growth of macrocrack
- final failure

Cyclic slip

At relatively low cyclic loads (i.e. low as compared to the static failure load) cyclic slip will occur in a small number of grains of a polycrystalline material. In these grains it is again a local phenomenon because it is concentrated in a few slip bands only. The reason is that only in a few bands the cyclic shear stress is high enough to produce slip (favourable crystal lattice orientation and grain shape).

If cyclic slip would not occur there would be no fatigue. The plastic
deformations involved are microscopically small, also because it is a cyclic deformation.

Crack nucleation

After a number of load cycles small microcracks can be observed in the slip bands. Microscopical investigations have revealed that crack nucleation occurs early in the fatigue life, possibly after just a few percent of the fatigue life has passed. Another observation is that crack nucleation is a material surface phenomenon. Why is this so? Apart from some trivial reasons to be discussed later (surface roughness, stress concentration, etc.) two significant aspects are:

1. Grains at the material surface do not have neighbouring grains all around. At one side there is the non-solid environment. For slip this implies a lower restraint on microplasticity. At the free surface slip can occur more easily over larger slip distances, because the restraint of a neighbouring grain is absent.

2. Surface grains are in contact with the environment. Even for normal air the presence of oxygen and water vapor can cooperate with cyclic slip to produce crack nucleation.

Sometimes an important exception to crack nucleation at a free surface is found, which is nucleation at inclusions. It can occur in Al-alloys at small intermetallic particles and in low-alloy high-strength steels at small inclusions. These impurities are numerous and microscopically small. There is still a tendency then for crack nucleation at inclusions near to the free surface. Possibly this is again a result of the first argument mentioned above (lower restraint near free surface).
Microcrack growth

Once a microcrack is present stress concentration at the crack tip will occur. Cyclic slip will be concentrated at the crack tip, leading to further crack extension. As shown again by microscopic studies the growth of a microcrack can cover a relatively large part of the fatigue life. In other words: at a relatively late moment of the life the crack becomes visible to the naked eye and will then be labelled as a macrocrack. This is illustrated by the results in figure 7.1, which shows that after a crack of 1 mm was visible only a few percent of the fatigue life until failure was left. These results apply to unnotched specimens. For notched specimens and other conditions macrocracks can appear earlier. Nevertheless it is still true to state that: A large part of the fatigue life is covered by microcrack growth at or near the surface of the material.

The statement implies that the local conditions at the material surface and the material quality at the surface are highly significant for the fatigue life to be spent in the microcrack region.

Macrocrack growth

The transition from microcrack growth to macrocrack growth cannot easily be defined quantitatively. The nominal definition of a macrocrack is that it is a crack visible to the naked eye. However, small cracks which are known to be present can be visible, whereas larger cracks which are not known to be present are not always found during an inspection. It was claimed in the past that cracks should be at least in the order of 1 cm (or 0.5 inch) in order of having a reasonable chance to be found during routine service inspections (visible inspection, sometimes employing magnifying glasses, good light should be available). Supplementary to what was said before about the fatigue life spent in the microcrack region, the following definition appears to be useful: Macrocrack growth applies if the crack growth rate is no longer dependent on the local surface and material conditions which were responsible for crack nucleation and microcrack growth. Different circumstances for crack growth are qualitatively indicated in figure 7.2. The vertical scale for crack length is logarithmic. The scale suggests a transition from micro to macro at about 1 mm, but this value should not be taken to literally. In the best case fatigue crack nucleation occurs at a high quality surface and the...
crack will turn up as a macrocrack relatively late in the life. If cracks start at an inclusion it occurs somewhat earlier. If crack growth starts from a macrodefect it is a macrocrack right from the beginning. In aircraft materials macrodefects should not be present, although sometimes they escape material quality control (F-111 accident). In other case unintentional surface damage has caused macrocrack growth early in the service life. In welded steel structures macrodefects are more common (weld defects).

Final failure

When the macrocrack is growing larger the remaining uncracked cross section (ligament) becomes smaller. Finally it will be too small to carry the maximum of the cyclic load. Then the final failure will occur in what is the last cycle of the fatigue life. The final failure usually is a quasi-static failure, showing significant macroplasticity as found in normal static tests.

7.3 THE FATIGUE MECHANISM

Microcrack initiation

Wood (1958) has proposed a simple and attractive model for the nucleation of the first microcrack, see figure 7.3. During the first upward loading (fig. 7.3a) slip will create a surface step. The following downward loading (fig. 7.3b) will cause slip in the opposite direction, not exactly on the same slip plane, but on a closely adjacent slip plane. The shear stress (τ) on this plane will be high for two reasons: (1) On a microlevel there is a shear stress concentration in this grain on these parallel slip planes. This was shown already during the first upward loading. (2) The microplasticity of the upward loading will leave microresidual stresses, which will help to induce reversed plasticity (local Bauschinger effect on a microlevel). We still have to explain why the reversed slip does not occur on exactly the same slip plane. If that would happen the original undamaged situation would be restored (and fatigue would be unknown). Two reasons can be mentioned: (a) Fully reversed plasticity implies reversible dislocation movements. In view of strain-hardening (pinning of dislocations, dislocation climb) that is a highly unlikely event. Exaggerating: a plane which has slipped once cannot slip in the reversed direction due to strain-hardening. (b) The second argument concerns the environment. The
slip step AB (fig. 7.3a) implies that new fresh material is exposed to the environment. All technical materials oxidize extremely rapidly and the oxide layers are strongly adhered to the base metal. Reversed plasticity requires that the oxide layer is removed but that cannot occur by reversed slip.

In vacuum this second argument cannot be applicable while fatigue in vacuum is still possible. Apparently the first argument is already sufficient to account for the irreversibility.

Figures 7.3c and d show a repetition of what occurred in first load cycle. A microcrack is formed as an "intrusion" into the material. The model explains a few things: (1) Crack nucleation can start right from the first cycle. (2) Crack extension can occur in every load cycle. (3) The very first part of microcrack growth should be expected along slip bands. Microscopical studies have confirmed these conclusions.

Figure 7.4 shows an optical microscope picture of pure Al. After 5000 cycles only three slip bands were visible and it was impossible to say whether cracks were present in these bands. The material was then loaded until macroplasticity occurred. That opened the slip bands and apparently they were microcracks. Note that additional static slip bands are terminating at the crack, forming small slip steps.

It may well be argued that reversed slip in figure 7.3b on an adjacent plane just below the original one (instead of above) would have led to an extrusion instead of an intrusion. Extrusions have been observed on pure aluminium. However, the intrusions introduce the microcracks, which will grow due to enhanced slip concentrations.

Crack growth and striations

In the preliminary stage of microcrack growth the lower restraint on slip as a result of the neighbouring free surface will promote slip band cracking for some time (Stage I crack growth according to Forsyth). After penetration inside the material the lower restraint will vanish. Moreover the stress intensity at the crack tip will increase. As a result slip on more than one slip plane orientation should be expected. A situation as sketched in figure 7.5 will occur. Even if two different slip systems are active at the crack tip, crack extension by slip steps is possible. The crack now will grow in a direction in between the orientations of the two contributing slip
systems. This is labelled stage II crack growth, see figure 7.6. It should be noted that a continuous crack front is a line through a number of adjacent grains. This continuity does not allow a free choice of crack growth directions in adjacent grains along the crack front.

If the crack becomes longer the stress intensity will increase and the crack extension in each cycle (= crack growth rate) will be larger. The microplasticity in each cycle will also increase and if the growth rate is large enough the microplasticity will be visible on the fracture surface in the electron microscope. Two different models how this can occur are indicated in figure 7.7. During load increase the crack will be extended and at the same time there will be crack tip blunting. During the subsequent unloading the crack tip is resharpened in one model and more or less "folded up" in the other one. Both models on a microlevel leave a kind of a plastic deformation ridges on the fracture surface. The ridges are generally referred to as striations. One striation corresponds to one cycle. An elegant proof of this is shown in figure 7.8, where a periodic load sequence corresponds exactly to the same sequence of striations.

Striations are hard to detect with the optical microscope. However, with the electron microscope they can be observed. The resolution is better in the TEM (transmission electron microscope, looking at a replica of the fracture surface) than in the SEM (scanning electron microscope, looking at the fracture surface itself).

7.4 CHARACTERISTICS OF A FATIGUE FRACTURE

It is most worthwhile to know the characteristics of a fatigue fracture for two purposes:

(1) For the interpretation of fatigue test results obtained in the laboratory valuable information can frequently be derived from a careful study of the fracture surfaces.

(2) For structural failures occurring in service fractographic analysis is a major key towards finding the cause of the failure. A distinction has to be made between micro and macrocharacteristics. The macrocharacteristics will be discussed first, because any analysis of a fracture surface should start by looking first with the unaided eye and a magnifying glass.
Macrocharacteristics

1. No macroplasticity and flat surface

In general the fracture surface of a fatigue failure shows two different parts: (1) The first part if the real fatigue failure caused by fatigue crack growth (see figure 7.9). (2) The second part is caused by the final failure. It is characteristic for the fatigue part that macroplastic deformation is practically absent. The fatigue part is usually also very flat. Both sides of the failed component or specimen can be nicely fitted together again (if the final failure part does not obstruct this). An example is shown in figure 7.10, which illustrates the difference between a fatigue failure and a static failure (see also fig. 7.13). It is not so surprising that there is no macroplasticity because fatigue crack growth is a result of microplasticity. Moreover it has a cyclic nature with a tendency to reverse the deformations.

A consequence of the absence of macroplasticity is that fatigue cracks in general are not easily found. Service inspections for fatigue cracks by visual observations are hard to carry out. Only if the inspector knows where to look for a crack and if he has seen similar cracks before, it can be reliable.

2. Concentric growth bands

Growth bands are shown in figures 7.11c and 7.12. The bands are also referred to as "tide", "beach", "clam-shell" or "oyster-shell" markings. The bands indicate how the crack has been growing. The different colours should be associated with variations of the magnitude of the cyclic load. Different degrees of corrosive attack can also cause bands, especially if cracks are dormant for certain periods.

3. Growing direction perpendicular to the main principal stress

Experience has shown that fatigue cracks are growing in a direction perpendicular to the main principle stress (provided the crack growth rate is not very high). For a cyclic tension load it implies that crack growth will be perpendicular to the loading direction. For cyclic torsion of an axle the main principal stress is under an angle of 45° with the long direction. This leads to a spiral crack growth, see figure 7.13.

If a fatigue crack is growing very fast, for instance in sheet material, shear lips are formed at the free surface in a similar way as for static
crack extension discussed in chapter 5 (see figure 5.13). The shear lip width will increase during faster growth until they cover the full thickness, see figure 7.14. The crack growth direction remains perpendicular to the loading direction. The transition from the tensile mode (mode I in figure 4.2) to the shear mode (mixed mode I/ mode III) is characteristic for fast fatigue crack growth in sheet material. In figure 7.14 the shear lips are parallel which leads ultimately to a single shear mode. However, if the two shear lips make opposite angles (+45° and -45°) it leads to a double shear failure.

4. Radial steps in the growth direction

Sometimes the fracture surface shows radial steps which run in the growth direction, see figure 7.12. Such steps can be the results of the crack growing on slightly staggered levels. In view of continuity a step must occur between two different levels. It is more or less typical for forged Al-alloys, if a pronounced deformation structure is present. There is no longer a fully random orientation of the crystal lattice from grain to grain (deformation texture). It leads to small local deviations from growing perpendicular to the main principal stress. The result is growing on slightly staggered levels and radial steps in between. In general the radial steps and the growth bands make it fairly easy to locate the point where the crack was initiated.

Similar steps (but non-radial) also occur if different fatigue cracks are overlapping. In figure 7.11c it can be seen at the notch on the left hand side. Two cracks, starting at the two corners of the notch, meet at mid thickness at slightly different levels and thus create a step. The step disappears later when the two cracks have merged into a single crack.

5. Number and size of macrocracks

If the fatigue load is low crack nucleation will occur at the weakest point only. If the fatigue crack is high crack nucleation will occur at more potential crack initiation locations. This is illustrated by figures 7.11b and c with only one dominant crack in figure 7.11b corresponding to a low stress amplitude and four cracks in figure 7.11c for a high stress amplitude. Another related feature is the area of the fatigue part. For the low stress amplitude it is large
because $\sigma_{\text{max}}$ is low. For the higher stress amplitude it is smaller because $\sigma_{\text{max}}$ was higher. It is more precise to say that the area of the fatigue part has some inverse relation to $\sigma_{\text{max}}$ in the last cycle of the fatigue life.

**Microcharacteristics**

6. **Transcrystalline crack growth**

Fatigue cracks in almost all materials are growing transcrystalline (= transgranular). They do not follow the grain boundary, contrary to stress corrosion cracks and creep failures. Because fatigue crack growth is a consequence of cyclic slip, it is not surprising that the fatigue crack is growing through the grain. Restraint on slip from the grain boundaries is minimal for a transcrystalline crack. The transcrystalline nature can easily be observed in the optical microscope.

7. **Striations**

As discussed before striations are remnants of individual load cycles. Sometimes it is possible to deduce fatigue crack growth rates from fatigue failures in service. An illustration is shown in figure 7.15 where striations occur in pairs. Further analysis requires information on loads in service. For this particular case of a landing flap beam it was not difficult to indicate two dominant loads in service. A small flap load during take off and a large flap load during landing (flap fully out). Each pair of striations thus corresponds to a single flight. By making replicas at different points of the fatigue part of the failure it could be deduced that the crack growth life was in the order of 5000 flights. That implies that it is not a dangerous crack.

In failure analysis the various macro- and microcharacteristics can be a great help. Usually it is not difficult to discriminate between a fatigue crack and a stress corrosion crack. Although both do not show macroplasticity and although stress corrosion failures sometimes show growth bands, there are really characteristic differences.

Striations are found on fatigue fractures only. It should be noted that the absence of striations does not necessarily imply that it is not a fatigue crack. Not all materials show striations equally well. Some steels depending on the heat treatment hardly show any striations. Moreover, the visibility of striations is also depending on the crack
growth rate. At very low growth rates the striation spacing may be too small to detect anything at all. A most discriminating feature is the transcrystalline nature of fatigue cracks and the intercrystalline nature of stress corrosion cracks. Finally the fibrous structure affects the path of a stress corrosion crack, but it has hardly any influence on the direction of fatigue crack growth.

7.5 FACTORS WHICH INFLUENCE FATIGUE BEHAVIOUR

The fatigue life of a structure is depending on many factors. In general terms the following list can be made up:

1. External factors
   1a. the cyclic loads on the structure in service
   1b. the environments in which the structure should operate

2. Factors concerning the structure:
   2a. lay out of the structure
   2b. detail design and joints
   2c. material selection
   2d. surface quality, surface treatments.

Several practical aspects will be dealt with in other chapters. Here we will restrict us to a survey of some basic factors, which are:
- stress amplitude and mean stress;
- cyclic frequency and wave shape;
- environmental effects;
- surface effects;
- fretting corrosion.

It will be discussed for unnotched material only. The notch effect is treated in chapter 8.

7.5.1 Stress amplitude and mean stress, fatigue curves and fatigue diagrams

In this chapter fatigue under a constant stress amplitude ($\sigma_a$) and a constant mean stress ($\sigma_m$) will be considered only. This type of cyclic loading, see figure 7.16 is usually referred to as constant-amplitude loading (CA-loading), while it is tacitly assumed that $\sigma_m$ is constant also. Certain load cycles in service are approximately constant-amplitude loading, but for many other types of service loads the amplitude is not constant (and sometimes both $\sigma_a$ and $\sigma_m$ vary). This is referred to as variable-amplitude loading (VA-loading). The pressurization
cycle of the cabin of a civil aircraft can be considered to be constant-amplitude loading. The wave shape is not sinusoidal but trapezoidal with a rather long period. On the other hand gust loads on an aircraft wing are an example of variable-amplitude loading. VA-loading is a complication as compared to CA-loading. This is discussed in later chapters.

In figure 7.16 the stress ratio \( R \) is defined by:

\[
R = \frac{\sigma_{\text{min}}}{\sigma_{\text{max}}}
\]

(7.1)

\( R = 0 \) implies that \( \sigma_{\text{min}} = 0 \), while \( R = -1 \) implies \( \sigma_{\text{m}} = 0 \) (fig. 7.16). Apart from the frequency and the wave shape there are 5 quantities to define the stress cycle: \( \sigma_a \), \( \sigma_m \), \( \sigma_{\text{max}} \), \( \sigma_{\text{min}} \) and \( R \). Only two of them can be chosen independently. Usually a stress cycle is characterized by \( \sigma_a \) and \( \sigma_m \). However, from the point of view of the fatigue mechanism this is not so logical. It appears more appropriate to speak in terms of \( \sigma_{\text{max}} \) and \( \sigma_{\text{min}} \). At these stress levels the loading direction is reversed, and consequently the microplasticity is reversed. Slip comes to an end at \( \sigma_{\text{max}} \), and also at \( \sigma_{\text{min}} \) in the reversed direction. Also for the crack extension mechanism \( \sigma_{\text{max}} \) and \( \sigma_{\text{min}} \) are the milestones.

The effect of the cyclic stress on the fatigue life \( N \) can be represented by fatigue curves (also indicated as S-N curves or Wohler curves). An example is shown in figure 7.17. Tests were carried out at several \( \sigma_a \)-values, while for all tests \( \sigma_m \) was equal to zero. A number of specimens did not fail after 5 million cycles and testing was stopped then, assuming that they would have an infinite life if the test was continued. Apparently the fatigue curve has a horizontal asymptote. The corresponding stress amplitude is indicated as the fatigue limit \( \sigma_f \). For \( \sigma_a > \sigma_f \) there will be a finite life, for \( \sigma_a < \sigma_f \) the life is infinite. The fatigue curve has a second horizontal asymptote at the upper side, i.e. at \( \sigma_a = \sigma_U \) (or \( \sigma_a + \sigma_m = \sigma_U \) if \( \sigma_m \neq 0 \)). Apparently if the specimen does not fail in the first cycle (it would then be a tensile test) it can survive cycles in the order of 100 or more. This is a result of strain hardening, see figure 7.18. In the first uploading much plastic deformation occurs if \( \sigma_a \) is high. Afterwards the cyclic strain behaviour implies a much lower plastic strain amplitude.
Earlier we have split up the life in different phases. For the discussion of various influencing factors it is more efficient to divide the life in two periods:

\[
N = \text{crack initiation life} + \text{crack growth life} \tag{7.2}
\]

The crack initiation life is again defined as the life until the micro-crack is large enough to penetrate into the material. The effects of \( \sigma_a \) on the initiation life and the crack growth life is qualitatively shown in figure 7.19. At a high \( \sigma_a \) value the initiation period is relatively short. The amplitude is sufficiently high to produce effective crack growth at an early stage of the life. At low amplitudes the crack initiation period is relatively long. Apparently it is difficult to create a microcrack and let it grow through its preliminary microstage. If the crack can pass the initiation period the stress intensity at the tip has sufficiently increased to take care of further growth and final failure. However, if the crack initiation is unsuccessful the life will be infinite. We now give a new definition of the fatigue limit to be used later.

The fatigue limit \( \sigma_f \) is the smallest \( \sigma_a \) which still leads to crack initiation, or which is the same: \( \sigma_f \) is the largest \( \sigma_a \) which is still uncapable to produce crack initiation.

It is usual to plot the life \( N \) on a logarithmic scale. This is a distortion of a linear scale. Although the relative crack growth period becomes smaller at lower \( \sigma_a \) (see figure 7.19) its absolute value still increases. This is not directly evident from figure 7.19 due to the logarithmic scale.

**Scatter**

If fatigue tests are repeated it will be observed that the fatigue lives obtained are not the same. There is scatter of the results and fatigue has the reputation to show much scatter in fatigue life. It is a good question to ask whether there will be more scatter in the initiation life than in the propagation life. Experience has learned that scatter of the initiation period can be quite large, and this can be understood. Crack initiation is a matter of a highly localized fracture process. It therefore will be dependent on variations of a highly local nature. Incidental features (scratches,
large inclusions, etc.) can lead to a shorter initiation period than normal. Initiation is strongly depending on the conditions of the material in a thin surface layer, and local variations thus can cause much scatter. The situation is different in the crack growth period. Then the fracture process is depending on the bulk properties of the material. Local variations, now in a much larger volume of the material will have a smaller effect.

In the present discussion we will consider average fatigue curves only. Such curves are frequently presented on a semi-log scale, i.e. \( \log N \) as of function of \( \sigma_a \). However, it is not unusual to plot fatigue curves on a double log plot, i.e. \( \log N \) as a function of \( \log \sigma_a \). In many cases a fairly large part of the curve then becomes linear:

\[
\log N + \chi \log \sigma_a = \text{constant} \tag{7.3}
\]

where \( \chi \) is a constant, directly related to the slope, see figure 7.20. Equation (7.3) can be rewritten as the so-called Basquin relation:

\[
\sigma_a^\chi N = \text{constant} \tag{7.4}
\]

This relation does not have a horizontal asymptote at the fatigue limit. In figure 7.17 the sloping finite life part of the curve is approaching the fatigue limit at about \( N_\infty \approx 2 \times 10^6 \), a figure frequently quoted for steels. For Al-alloys \( N_\infty \) can be quite a bit larger and fatigue failures between \( N = 10^7 \) and \( N = 10^8 \) have sometimes been reported.

**Low-cycle fatigue**

It was discussed before that the fatigue curve has a second horizontal asymptote at low endurances. This becomes different if tests are carried out at a constant strain amplitude \( \varepsilon_a \), instead of a constant stress amplitude \( \sigma_a \). As long as the material behaves macroscopically elastic there is no difference between the two possibilities. However, at high \( \sigma_a \) and short endurances macroplasticity can be expected in every cycle, especially so if \( \varepsilon_a \) is constant. In figure 7.18a the plastic strain was large only in the first cycle. If we continue to have such a large plastic strain in all cycles the stress-strain hysteresis will remain to be wide. Very short fatigue lives will then be obtained. The relation between \( \log \varepsilon_a \) and \( \log N \) appears to be
approximately linear, see figure 7.21, which leads to the Manson-
Coffin relation:

$$\varepsilon_a N^\beta = \text{constant}$$  \hspace{1cm} (7.5)

The similarity with the Basquin relation is evident (substitute $\varepsilon = 1/\beta$).
However, for normal fatigue lives $\varepsilon$ may be in the order of 5 to 15, whereas $1/\beta$ is in the order of 2.
Low-cycle fatigue is of technical interest if the number of cycles in
service is very low, while each cycle may produce plastic deformation.
This can occur under fatigue at high temperatures. It will not be
discussed here any further.

Fatigue diagrams

For the fatigue curve in figure 7.17 $\sigma_m = 0$. If tests are carried out at
other $\sigma_m$-values different fatigue curves will be found, see figure
7.22a. Cross plots can now be made to arrive at a fatigue diagram with
lines of constant $N$, see figure 7.22b. Such a diagram illustrates the
effect of mean stress on the fatigue strength for a certain life $N$.
Some examples are given in figure 7.23. All lines for constant $N$ are
converging to $\sigma_a = 0$ for $\sigma_m \to \sigma_U$ which should be expected. The diagrams
suggest that the effect of $\sigma_m$ is not large, especially so if $N$ is high.
This is generally true: The stress amplitude $\sigma_a$ has a much larger
effect on fatigue than the mean stress $\sigma_m$.

In mechanical engineering the so-called Smith-diagram is frequently
adopted, instead of the fatigue diagram discussed before. In the
Smith diagram, see figure 7.24, $\sigma_{\text{max}}$ and $\sigma_{\text{min}}$ are plotted instead of
$\sigma_a$. The information, however, remains essentially the same.

7.5.2 Environmental effects

Fatigue is a matter of crack initiation and crack propagation. Both
can be accelerated by environmental effects, depending on the aggressiveness
of the environment in relation to the material. For the discussion three
different groups of environment are specified:
1. Dry inert gas or vacuum.
3. Humid gas, especially humid air:
3. Aggressive solutions, e.g. salt water.
The first group appears to be of academic interest only. Normal air in
which an aircraft is flying is supposed to be in the second category, but sometimes it is in the third one. At lower altitudes near the open sea the salt contents of the air is sufficiently high to collect salt condensation water in an aircraft structure.

In the literature corrosion fatigue (fatigue in a corrosive environment) is usually associated with aggressive environments in the third group. Electrolytic processes can both accelerate the initiation of a crack and crack growth afterwards. The contribution to initiation can be pretty much the same as described in chapter 6 on stress corrosion. Figure 7.25 shows corrosion damage of a material with an end grain structure. It also shows a similar geometrical notch, which according to the book of Peterson (Ref. 1 of chapter 2) has a stress concentration factor $K_c$ = 8. If this happens under a cyclic load with a stress amplitude below the original fatigue limit it should be expected that crack initiation now is possible again. The fatigue limit will be drastically reduced. Moreover crack propagation itself is also accelerated. Corrosive fatigue tests give the impression that a fatigue limit hardly exists any more, see figure 7.26.

If corrosion has such large effects it should be expected that the cyclic frequency should be important also, and perhaps the wave shape as well. It has been shown indeed that a lower frequency (more time per cycle) in general will give lower endurance (see figure 7.27)). Obviously corrosion prevention is most useful, not only for general corrosion and stress corrosion, but for corrosion fatigue as well. With respect to the effect of the wave shape there is some evidence available, which indicates that the time spent for raising the stress from $\sigma_{\text{min}}$ to $\sigma_{\text{max}}$ is the more important aspect, see figure 7.28. Qualitatively it can be understood by considering the cooperation between crack extension by fatigue (microplasticity) and the corrosive action. Crack extension on the atomic level is a matter of decohesion. The crack extension in a single cycle is not just a jump forward, but rather a progressive decohesion occurring only when the load is raised. This decohesion can be much more effective if it is assisted by some corrosion mechanism. Thus it is the loading rate (d$\sigma$/dt) and not the period of the load cycle (1/ν) which is the significant parameter.

In normal air the effect of the environment seems to be less spectacular.
Nevertheless, compared to fatigue in an inert environment fatigue in air will give lower fatigue curves and faster crack growth. Originally it was thought that oxygen is important in view of oxidizing the new fresh surface of fatigue cracks. However, there is not too much difference between fatigue results in vacuum and in pure dry \( \text{O}_2 \) provided the oxygen is very dry. The corrosive action of normal air compared to inert environments is mainly caused by water vapour. This is especially true for Al-alloys, but also for most technical steels. Illustrative results were collected by Bradshaw and Wheeler, see figure 7.29. Crack growth tests were carried out in such a way that the cyclic stress intensity factor variation \((\Delta K)\) remained constant. As a result the crack growth rate is also constant. Tests were carried out in environments with different amounts of water vapour, measured by the partial water vapour pressure. If there is almost no water vapour the crack rate is about \( 1.6 \times 10^{-5} \ \mu\text{m}/\text{cycle} \) and the environment appears to be inert. If there is ample water vapour the crack rate is in the order of \( 8 \times 10^{-5} \ \mu\text{m}/\text{cycle} \), i.e. 5 times larger. In between there is a transition which apparently depends on the test frequency. The pressure difference between the two transitions is about 100 times, which corresponds to the frequency ratio. This strongly suggests that the amount of water vapour has a dominant effect. It is now of practical significance to see how much water vapour can be present in the air. This depends on the temperature. At low temperature the saturation water vapour pressure is much lower than at room temperature. Since most fatigue loads in service will have a frequency between 0.1 Hz and 10 Hz there will be ample water vapour to reach the maximum crack rate. Under practical conditions a maximum water vapour effect will be found. Fortunately the same effect will be obtained in a laboratory test at a high frequency and room temperature. In other words within the limitations of service conditions and laboratory test conditions a frequency effect should not be expected. For fatigue in humid air a laboratory test can be representative for fatigue in service as far as it concerns corrosion fatigue. A similar convenient conclusion cannot be drawn for fatigue of a steel structure in salt water (e.g. off shore structures). An accelerated laboratory fatigue test will give longer lives and slower crack growth.
Case-history

Some 10 years ago the rear pressure bulkhead of the cabin of a transport aircraft failed at cruising altitude. All 85 passengers and the crew were killed. Figure 7.30 shows the attachment of the bulkhead to the fuselage frame and the skin. The catastrophic failure occurred because severe corrosion was found in cross section A, which had initiated fatigue crack growth which remained undetected. The fatigue load is the once-per-flight pressurization cycle. Note the following features:
1. The designer was aware that condensation water could be collected at this point. He prescribed the application of a (polysulphide) sealant and a drain hole.
2. The location of the crack is hard to inspect.
3. The spherical bulkhead was reinforced by a bonded strip for the connection to frame and skin joint. Unfortunately the strip can collect water at the upper edge (A). That started corrosion and subsequently fatigue.

Some simple measures could have prevented the accident:
 a. The sheet material of the bulkhead was rather sensitive to corrosion and fatigue. A more resistant material would have been much better.
 b. The upper edge of the strip should also have been sealed.
 c. If the strip had been bonded to the other side of the bulkhead collecting water would have been impossible.

7.5.3 Some surface effects

Surface finish

As said before the surface conditions are highly decisive for the crack initiation life. Surface roughness is one of those conditions. An old test series of De Forest (1936) is illustrative in this respect. He carried out rotating bending tests on unnotched specimens of a low-carbon steel.

Three different surface finish qualities applied: (1) coarse circumferential finish, (2) fine circumferential finish and (3) fine longitudinal finish.

A circumferential finish implies surface grooves on a small scale, with the grooves perpendicular to the bending stress. For a longitudinal finish they are in the direction of the stress, which is a more favourable situation. The results of De Forest, see figure 7.31 show indeed a significant effect on the initiation period (until a
surface crack length of 2.5 mm). The initiation period apparently is very sensitive to the surface quality. In contrast the crack propagation period is practically independent of the surface quality. It is more dependent on the bulk properties of the material.

The results of De Forest illustrate also another aspect, as shown by the table below:

<table>
<thead>
<tr>
<th>$\sigma_a$ (MPa)</th>
<th>coarse circumferential finish</th>
<th>fine longitudinal finish</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>345</td>
<td>90 000</td>
<td>288 000</td>
<td>3.2</td>
</tr>
<tr>
<td>276</td>
<td>436 000</td>
<td>2 720 000</td>
<td>6.2</td>
</tr>
</tbody>
</table>

Apparently the difference found between two surface finishes depends on the stress amplitude at which it is determined. If $\sigma_a$ is lower and the life longer the sensitivity for the surface finish increases. This seems plausible, because at lower amplitudes crack initiation becomes a more difficult process. It will be more sensitive to surface conditions. At a higher amplitude crack initiation is less difficult and therefore less depending on local conditions.

**Cladding layer on Al-alloys**

Sheet material of Al-alloys applied in aircraft structures is mostly provided with thin cladding layers at both sides (5 percent thickness at each side). The cladding layer is pure Al for 2024 and Al - 1% Zn for 7075. The cladding layers are anodic to the core material and thus provide a corrosion protection. That is the main purpose of these layers. Because they are pure or almost pure Al the cladding is soft and hardly contributes to strength: $\sigma_u$ and $\sigma_{0.2}$ will be somewhat lower (10 percent as a maximum) than for bare sheet material. The effect of the cladding layer on the fatigue limit of unnotched specimens is much larger, see figure 7.32. The fatigue limit can be about 50% lower. The reason is that crack initiation in the soft cladding layer is relatively easy. After traversing the layer penetration into the core material occurs. In figure 7.32 a lower limit of the scatter band of the bare material
is also indicated. The superior fatigue strength can be deteriorated by superficial surface damage. The clad material is not so vulnerable to surface damage because there is a low-fatigue-resistance surface layer anyhow.

7.5.4 Fretting

Fretting, or fretting corrosion as it is often called, is a surface phenomenon. It occurs as the result of small relative displacements in the contact area between two materials. The displacements have a cyclic nature and they are caused by a cyclic load. Fretting corrosion is a kind of wear. In an aircraft structure it frequently occurs that one material is clamped on another one. Actually that is what occurs in all riveted and bolted joints. Cyclic loading then causes surface damage due to fretting corrosion. This can have a disastrous effect on fatigue.

If materials are clamped together it is very difficult to prevent fretting if cyclic loads are present. Figure 7.33 should help to explain this. A bar is clamped at one end and loaded at the other end. Load transmission from the bar to the clamping occurs by frictional forces. In the unclamped part the bar will be elongated ($\varepsilon_x = 0/E$) and at the same time the thickness is reduced (lateral contraction: $\varepsilon_y = -\nu \varepsilon_x$). In the clamped part $\sigma = 0$ at the end, because the load is fully transmitted to the clamping parts. At the end $\varepsilon_y$ in the bar is negligible. As a consequence there is a varying thickness of the bar in the clamping, with the lowest thickness at AA, where the bar leaves the clamping. The result will be that the contact between the bar and the clamping will be slightly lost at AA. The bar will be pulled out of the clamping over a very small distance. This distance will be more a matter of a few microns rather than millimeters. If the load $P$ is now removed the opposite displacements will occur. In other words a cyclic load will induce very small cyclic displacements between the bar and the clamping. This can hardly be prevented. Increasing the pressure $p$ will make the displacements smaller. However, imagine that $P$ could be made large enough to prevent any displacement, then the bar and clamping would react like a single piece. The stress concentration at the points A would be extremely high, which again would induce a local contraction and subsequently a slight local movement of the bar out of the clamping. At the boundary between
the clamped and the unclamped area it is practically impossible to prevent microscopical displacements between two parts.

On a micro-scale the displacements are no longer small. Surfaces are never perfectly flat. There are higher and lower parts, see figure 7.34. If cyclic displacements have to be made local damage on a micro-scale will be done. Technical materials are covered by oxide layers. These layers will be disrupted and fresh material will be exposed to the environment. This will rapidly oxidize again, while some local welding between the two materials may occur also. In the latter case those bonds will be disrupted again. Small fragments of oxidized material (high hardness) will be produced in the contact area between the two materials. They will exert some micro-grinding effect on the surface in later cycles. This damaging process is known as fretting or fretting corrosion. The corrosion products in a dry environment look like a powder with a red-brown colour for steel and a black colour for Al-alloys. The significance of fretting for fatigue is that fretting damage, although it is superficial, has led to microscopically small cracks at the surface. The effect of fretting on the fatigue limit is large.

Some illustrative tests to show the large effect of fretting on fatigue were carried out by Fenner and Field, see figure 7.35. They clamped small blocks on sheet specimens to see the effect on the fatigue life. As shown by the table in figure 7.35 the blocks reduced the life by a factor of 80. The original life was $80 \times 10^6$ cycles, which implies that the cycle stress was close to the fatigue limit. Crack initiation is difficult then. However, if assisted by fretting it is apparently easy. Removing the blocks after a small percentage of the original life still produced a high reduction of the fatigue life. This means that the fretting damage is done early in the life. Figure 7.36 shows a fatigue curve without fretting and a similar curve with fretting. The effect of fretting is small in the finite-life area of the original curve. The reason is that here $\sigma_a$ is large enough to cause crack initiation without fretting. During the crack propagation period fretting at the surface will continue but this does not affect the crack growth rate. At stress amplitudes below the original fatigue limit cracks could not be initiated without fretting. However, with fretting it is possible and the fatigue limit ($\sigma_f$) drops to a significantly lower level ($\sigma_f'$)
For technical materials $\sigma_f'$ of unnotched specimens can be 2 to 3 times lower than $\sigma_f$. In view of the low $\sigma_a$ applicable then it takes a long time before the microcracks have grown to failure. The life $N_0$ moves to a higher value $N_0$.

For riveted and bolted joints it is evident that clamping forces are present between the materials joined together. However, in other cases it is less evident that fretting can occur. A classical case is a lug joint (see figure 7.37). The load is transmitted from a lug to a fork by a single bolt. In such a joint clamping is usually not applied. Unfortunately fretting can now occur inside the hole between the bolt and the wall of the hole. If the lug is loaded in tension a tangential tension stress will be present around the bolt in the lug head. This implies an elongation of the wall of the hole. It will be stretched over the steel pin. These small displacements are again responsible for fretting. The displacements are maximal at points A, where the fretting will be most severe. That is precisely at the points where the stress concentration has its maximum. The fatigue limit of a lug therefore is low. This will be discussed in more detail in chapter 10 on joints.

Another example is sketched in figure 7.38. The connection between a rib boom and a hat stiffener is apparently a joint of secondary importance. However, two parts are clamped together. During wing bending the hat stiffener will be under tension. Fatigue loads on the wing will cause a cyclic elongation of the stiffener. The stiffener is attached to the flange of the rib section, which will not elongate. Consequently small displacements between the two elements will occur and fretting damage to the hat stiffener can be done. A simple connection between a loaded part and an unloaded part may cause fretting corrosion, and thus introduce a fatigue crack in the loaded part.

**Prevention of fretting corrosion**

Unfortunately Al-alloy, steels and Ti-alloys used in aircraft structures are all prone to be damaged by fretting. The same applies to combinations of these materials, such as a steel bolt in an Al-alloy lug. The best method is to avoid metallic contact between elements. Usually that is not easy. Interfaying sealants are not very effective for this purpose.
Thin plastic lamina are generally unacceptable for various reasons. However, the designer if he recognizes the risk of fretting fatigue can sometimes alleviate the problem by improved detail design. This is discussed in chapter 10.

REFERENCES


Some books on fatigue of materials

Chapter 8  
FATIGUE UNDER CONSTANT-AMPLITUDE LOADING

Contents:
8.1 Introduction
8.2 The fatigue strength of a notched element under CA-loading
  8.2.1 Estimation of the fatigue limit of a notched element for $\sigma = 0$
  8.2.2 Estimation of the fatigue limit of a notched element for $\sigma_m > 0$
8.3 The notch effect for finite life
8.4 Fatigue crack propagation and the stress intensity factor
8.5 Fatigue crack propagation properties of a material and the stress intensity factor
8.6 Prediction of fatigue crack growth based on the stress intensity factor

References

8.1 INTRODUCTION

In the previous chapter fatigue was described in qualitative terms as a process occurring inside the material. A large part of the fatigue life was associated with the crack initiation period, i.e. the life until there is a very small crack. The second part of the life was covered by crack growth. In this chapter we will try to deal with fatigue in a more quantitative way. For that purpose some brief remarks on the general outline of aircraft fatigue will be given first.

1. CA-loading and VA-loading

In the previous chapter constant-amplitude loading and variable-amplitude loading were defined. Sometimes aircraft components are subjected to CA-loading, at least approximately. Frequently VA-loading is applicable. As an example figure 8.1 shows strain gage records which were made to measure the wing bending moment of two aircraft flying in turbulent air (gusts). The complex nature of VA-loading in service is dealt with in chapters 11 and 12.

2. Unnotched material ($K_t = 1$) and notched components

The effect of notches is to reduce fatigue life. Real structures always contain notches (holes, fillets, etc.), sometimes rather complex notches, viz. riveted and bolted joints.

3. Predictions of life and crack growth

For most aircraft components the requirement of an infinite fatigue life would make the aircraft too heavy. Consequently finite life problems have to be considered. If the life is defined as the number of cycles until there is a small crack, an aircraft safety requirement is that dangerous situations afterwards should not occur due to the presence
of a macro crack. In other words we have to bother about crack growth and residual strength (= strength when cracks are present).

4. Verification of predicted fatigue behaviour by tests

It will turn out that accurate fatigue predictions are difficult. If it is considered to be necessary predictions have to be verified by tests, in final situations by full-scale tests.

The above list indicates some major aspects of the problem to give quantitative information about the fatigue performance of an aircraft structure. Several aspects will be discussed in more detail in other chapters. The present chapter is restricted to the prediction of fatigue data under CA-loading for simple notched elements and for macro fatigue crack growth.

8.2 THE FATIGUE STRENGTH OF A NOTCHED ELEMENT UNDER CA-LOADING

8.2.1 Estimation of the fatigue limit of a notched element for \( \sigma_m = 0 \)

The estimation of the fatigue limit of a notched element for \( \sigma_m = 0 \) is the most easily defined problem in this chapter. It is relevant in several technical branches, where fatigue is not allowed in any case. A good example is crack shaft of an engine. It will see numerous load cycles, but crack initiation is not allowed. All load cycles should be below the fatigue limit and therefore the fatigue limit is the only relevant fatigue property of interest in such a case.

The fatigue limit for a notched element \( (K_t > 1) \) will be indicated by the symbol \( \sigma_{fk} \), while the fatigue limit for the unnotched material \( (K_t = 1) \) previously indicated by \( \sigma_f \) will now be represented by \( \sigma_{f1} \). It looks rather obvious to relate \( \sigma_{fk} \) to \( \sigma_{f1} \) and to assume that \( \sigma_{fk} \) can be simply obtained as:

\[
\sigma_{fk} = \sigma_{f1}/K_t
\]

We will give some more thoughts to the question whether this is really so obvious. However, before doing so it may be asked if \( \sigma_{f1} \) data are always available.

The fatigue limit for unnotched material \( (\sigma_{f1}) \) for \( \sigma_m = 0 \)

For most technical materials \( \sigma_{f1} \) data can be found in handbooks, at
least for \( \sigma_m = 0 \), frequently obtained as the result of rotating bending tests. If no data are available a rough estimate can be obtained from the correlation between \( \sigma_{f1} \) and the tensile strength \( \sigma_U \). Forrest (Ref. 1) compiled the data shown in figure 8.2. Although for each type of alloy there is an almost linear correlation between \( \sigma_{f1} \) and \( \sigma_U \), there is also a lot of scatter around the average trend

\[
\beta = \frac{\sigma_{f1}}{\sigma_U}
\]

For strong Al-alloys \( \beta \) is in the order of 30 percent. For many steels \( \beta \approx 50\% \), but for high-strength steel there is a tendency to lower values. This can be due to small inclusions, which are not harmful if the strength is lower and the ductility better. However, for high-strength steels applied in aircraft those "micro/notches" can initiate fatigue crack nuclei, and for that reason a stringent impurity control of high-strength steels is necessary.

The fatigue notch factor \( K_f \)

In figure 8.3 stress distributions present in an unnotched specimen and in a notched specimen are compared. As a prototype of a notched element a strip with a central hole is adopted. The comparison to be made is between the fatigue limits of the notched and the unnotched specimens. The second definition of the fatigue limit should now be recalled: it is the highest stress amplitude just incapable to initiate a small crack; or it is the lowest stress amplitude still capable to initiate a crack. It is the definition of the fatigue limit as the threshold stress level for crack initiation. For the comparison in figure 8.3 \( \sigma_{peak} \) at the notch is chosen equal to \( \sigma_{f1} \) in the unnotched specimen. The following conclusion then seems reasonable: If \( \sigma_{f1} \) is the threshold stress level for crack initiation in the unnotched specimen than \( \sigma_{peak} \) \((= \sigma_{f1})\) should also be the threshold stress level for crack initiation at the root of the notch. In other words \( \sigma_{peak} \) is the peak stress occurring at the fatigue limit \((\sigma_{fk})\) of the notched element:

\[
\sigma_{peak} = K_f \sigma_{fk} = \sigma_{f1}
\]

The fatigue notch factor (also fatigue strength reduction factor) is defined as:
\[ K_f = \frac{\sigma_{f1}}{\sigma_{fk}} \quad (8.4) \]

With this definition equation (8.3) implies:

\[ K_f = K_t \quad (8.5) \]

or in other words, the fatigue notch factor is equal to the stress concentration factor.

Fatigue tests in many cases showed that \( K_f \) was smaller than \( K_t \). If \( K_f \) is plotted as a function of \( K_t \) test series revealed the trends shown in figure 8.4. Apparently the deviations from \( K_f = K_t \) depends on the ductility of the material. Technically it is convenient that the notch effect is smaller than predicted by \( K_t \). However, it is useful to have some understanding why there is a deviation.

**Two reasons why \( K_f < K_t \) is possible**

1. **Local plasticity at the notch root**

The definition of \( K_t \) assumes that the material behaviour is (macroscopically) elastic. It is possible that \( \sigma_{\text{peak}} \) exceeds the yield limit, see figure 8.5. In that case:

\[ \sigma_{\text{peak}} < K_t \sigma_{fk} \quad (8.6) \]

and combining this with equation (8.3) gives:

\[ K_f < K_t \quad (8.7) \]

If \( \sigma_{\text{peak}} > \sigma_{0.2} \), equation (8.3) implies that \( \sigma_{f1} > \sigma_{0.2} \). This can sometimes be true, but usually for soft materials only. The second argument therefore is the more important one.

2. **Volume of highly stressed material**

The derivation of \( K_f = K_t \) started from the "similarity concept" that equal values for \( \sigma_{\text{peak}} (= K_t \sigma_{fk}) \) and \( \sigma_{f1} \), should have the same result with respect to fatigue crack nucleation. However, we then overlook that \( \sigma_{f1} \) is a stress which is present in the full prismatic part of the unnotched specimen, whereas \( \sigma_{\text{peak}} \) and stresses very close to \( \sigma_{\text{peak}} \) are present in a small volume of material at the notch root.
only. In the unnotched specimen crack nucleation can occur at many places, and it will do so at the location with the lowest fatigue resistance. However, in the notched specimen the location for crack nucleation is dictated by the presence of the notch. Nucleation has to occur at the notch root, although there may be locations with a lower fatigue resistance at other places. As a result the peak stress required at the notch root should be somewhat higher than $\sigma_{f1}$:

$$\sigma_{\text{peak}} (= K_t \sigma_{f_k}) > \sigma_{f1}$$  \hspace{1cm} (8.8)

Combination of this equation with eq. (8.3) leads again to: $K_t < K_L$.

**Statistical size effect**

The argument of locations with lower fatigue resistance used above, is associated with the "weakest link" concept. This concept generally serves to explain size effects on strength. It also implies that geometrically similar specimens which have the same $K_L$, but which have different size (see figure 8.6) can have different fatigue limits. This has indeed been found as illustrated by the NASA results (Ref. 2) below:

<table>
<thead>
<tr>
<th>2024-T3</th>
<th>Dimensions</th>
<th>$\sigma_{f_k}$ (N/mm$^2$)</th>
<th>$K_t$</th>
<th>$K_{IN}$ (eq. 8.9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_m = 0$</td>
<td>d (mm) W (mm)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d/W = 0.25</td>
<td>3.2 12.7</td>
<td>79</td>
<td>1.86</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>12.7 50.4</td>
<td>70</td>
<td>2.10</td>
<td>2.12</td>
</tr>
<tr>
<td></td>
<td>25.4 100.8</td>
<td>67</td>
<td>2.19</td>
<td>2.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Compare to</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma_{f1}$</td>
<td>$K_t$ = 2.43</td>
<td></td>
</tr>
</tbody>
</table>

The table shows that for a larger size $\sigma_{f_k}$ decreases and $K_t$ increases until 90% of $K_L$. Because the size effect is related to the volume of the highly stresses material at the notch root, it should be expected to be related to the stress gradient at the notch root. An equation for the stress gradient was presented in section 2.3, where is was shown that the width of the highly stresses area was a matter of a small number of grains depending on the notch root radius. Neuber
has tried to evaluate the stress gradient to a rational criterion for notch size effects. The equation which is finally proposed, however, was an empirically one. The usefulness of this equation was further evaluated by Kuhn and Hardrath (Ref. 3). The "Neuber" equation is:

\[
K_{FN} = 1 + \frac{K_t - 1}{1 + \sqrt{A/\rho}}
\]  

(8.9)

\(K_{FN}\) is the \(K_F\) value according to the Neuber equation, \(\rho\) is the notch root radius and \(A\) is a material constant. The size effect enters the equation through the value of \(\rho\). It is easily seen that for very large sizes \((\rho \to \infty)\) that \(K_{FN} \to K_t\). The material constant \(A\) has to be determined from experiments. Kuhn and Hardrath quoted:

2024-T3 and 7075-T6 (bare material): \(A = 0.5\) mm
steel: \(A\) is a function of \(\sigma_U\)

\[
\begin{align*}
\sigma_U & = 500 & 1000 & 1500 & \text{(N/mm}^2) \\
A & = 0.20 & 0.04 & 0.004 & \text{(mm)}
\end{align*}
\]

Note that a lower \(A\) value implies that \(K_{FN}\) comes closer to \(K_t\), that means an increased notch sensitivity for higher strength steels. A high quality steel is intolerant if notches are present. A low strength steel is much more tolerant. Designing in low alloy high strength steels has to be done very carefully to avoid stress concentrations as much as possible.

An illustration of the effect of \(\sigma_U\) on the notch effect is shown in figure 8.7, where test results are compared with predicted \(\sigma_{FK}\) values employing equation (8.9). For the alloy steel with \(\sigma_U = 1000\) N/mm\(^2\) (note: this is not yet a high value for an aircraft steel) the fatigue limit \(\sigma_{FK}\) is coming rather close to \(\sigma_{FL}/K_t\) if small root radii are disregarded (small root radii should be avoided anyhow to reduce \(K_t\)). As a matter of fact \(K_F = K_t\) should be adopted for high strength materials in order to be safe.

8.2.2 Estimation of the fatigue limit of a notched element for \(\sigma_m > 0\)

In the previous section we were dealing with the prediction of the fatigue limit of a notched element for \(\sigma_m = 0\). In this section the effects of a positive mean stress (\(\sigma_m > 0\)) will be considered. This case is of practical significance, because the risk of fatigue is larger at a positive mean stress. Very high \(\sigma_m\)-values do not occur in aircraft structures, and the more relevant part of fatigue diagrams may be
found between $\sigma_m = 0$ and $\sigma_m \approx 1/3 \sigma_U$. Negative $\sigma_m$-values are not of
great interest because the fatigue behaviour is much better then.
This is illustrated by the fatigue diagram in figure 8.8 for 2024-T3.
The explanation for the excellent behaviour at compressive mean stresses
is the very low crack growth rate at $\sigma_m < 0$. Crack nucleation does
occur at $\sigma_m < 0$, because it is mainly depending on alternating shear
stresses in slip bands. However, microcracks are not opened by a
compressive stress. Therefore the growth rate is very low. The low fatigue
sensitivity at $\sigma_m < 0$ is reflected in the fatigue behaviour of aircraft
wings. It is normal to have a positive mean stress in the lower skin
(tension skin) and a negative mean stress in the upper skin (compression
skin). As a result the lower skin of many aircraft is fatigue critical,
whereas the upper skin is not. The upper skin usually is stability critical
(buckling).

The fatigue limit for unnotched material ($\sigma_{f1}$) for $\sigma_m > 0$

Also in this section a relation between notched and unnotched fatigue data
will be looked for. The fatigue limit of unnotched material as a function
of $\sigma_m$ is sometimes available in a fatigue diagram, see figure 7.23.
In other cases rough estimates can be made starting from $\sigma_{f1}$ for $\sigma_m = 0$.
Two possibilities to estimate $\sigma_{f1}$ for $\sigma_m > 0$ are indicated in figure 8.8.
The linear relation:

$$\sigma_{f1} = (\sigma_{f1})_{\sigma_m = 0} \cdot \left(1 - \frac{\sigma_m}{\sigma_U}\right)$$  \hspace{1cm} (8.10)

(sometimes called: the modified Goodman diagram) is apparently a more
conservative estimate than the so-called Gerber parabola:

$$\sigma_{f1} = (\sigma_{f1})_{\sigma_m = 0} \cdot \left[1 - \left(\frac{\sigma_m}{\sigma_U}\right)^2\right]$$  \hspace{1cm} (8.11)

Both relations satisfy the condition that $\sigma_{f1} \rightarrow 0$ if $\sigma_m = \sigma_U$.
A comparison of figure 8.8 with figure 7.23 learns that in several
cases the fatigue limit will be reasonable approximated by the Gerber
parabola but not necessarily in a conservative way. A clear exception
is the 4340 steel heat treated to $\sigma_U = 1830 \text{ N/mm}^2$, which is high
indeed. In that case even the linear relation is not fully conservative.
A general trend is that materials with a lower ductility are more
sensitive to mean stress. This is illustrated here in figure 8.13
to be discussed later.
The fatigue limit for notched material \( (\sigma_{fk}) \) for \( \sigma_m > 0 \).

First we will assume that macro plastic deformation at the root of the notch does not occur. As a result all stresses at the notch root will be \( K_t \) times larger than the nominal values, see figure 8.10. Those \( K_t \) times higher stress levels should be compared to the same stresses in unnotched specimens. This is a simple procedure in a fatigue diagram, see figure 8.11a. Point B for the notched specimen is obtained from point A for the unnotched specimen by taking: \( OB = OA/K_t \). As a result both \( \sigma_a \) and \( \sigma_m \) for point B are \( K_t \) times lower than for point A. That implies that the stress amplitude and the mean stress at the root of the notch are exactly equal to \( \sigma_a \) and \( \sigma_m \) in point A. Since point A is a fatigue limit for \( K_t = 1 \) point B is a corresponding fatigue limit for the notched specimen. The full \( \sigma_{fk} \)-line is obtained from the \( \sigma_{fl} \)-line by a geometrical reduction of both \( \sigma_a \) and \( \sigma_m \) by a factor \( K_t \). In other words \( K_t \) is applied to both \( \sigma_a \) and \( \sigma_m \).

The above method may not be unreasonable for small \( \sigma_m \)-values, but it is definitely wrong for high \( \sigma_m \)-values. The method implies that the static strength of the notched specimen \( (\sigma_m \text{ for } \sigma_a \to 0) \text{ is equal to } \sigma_U/K_t \).

In chapter 5 we have seen that such a large reduction does not occur. On the contrary, for moderate \( K_t \)-values the static strength of the notched element is of the same order of magnitude as \( \sigma_U \). The explanation of the discrepancy is that \( K_t \) is an elastic concept, while plastic deformation is notched specimens at high \( \sigma_m \) values cannot be avoided.

Figure 8.12 indicates what happens if a small plastic zone is formed, which will level off the peak stress at the root of the notch. During unloading elastic spring back will occur. In view of elastic unloading the stress variation at the notch root will remain the same as in figure 8.10, viz. \( 2K_t \sigma_a \). The mean stress, however, will be lower than \( K_t \sigma_m \). Considering that \( \sigma_m \) has a much smaller effect on fatigue than \( \sigma_a \) an approximation is to apply \( K_t \) on \( \sigma_a \) only. This is done in figure 8.11b. Point A then goes to B' directly below A (same \( \sigma_m \)). Obviously \( \sigma_{fk} \) goes to point A now, because the \( \sigma_{fl} \)-line is reduced in the \( \sigma_a \)-direction only. The choice between the two methods of figures 8.10a and b should depend on the exceedance of \( \sigma_{0.2} \) at the notch root. This occurs if:

\[
K_t\sigma_{\text{max}} = K_t(\sigma_a + \sigma_m) > \sigma_{0.2} \tag{8.12}
\]
The limitation thus is:

\[ \sigma_a + \sigma_m = \sigma_{0.2}/K_t \]  \hspace{1cm} (8.13)

This line has been indicated in the two fatigue diagrams shown in figure 8.13. For the 7075-T6 alloy (fig. 8.13a) application of \( K_t \) to both \( \sigma_a \) and \( \sigma_m \) is theoretically relevant up to point D. The test results indicate this to be approximately correct. They anyhow show that application of \( K_t \) to \( \sigma_a \) only would be unsafe. However for the more ductile steel (fig. 8.13b) the other method (applying \( K_t \) to \( \sigma_a \) only) seems to be more appropriate.

For the diagrams in figure 8.13 \( K_t = 2 \), which is fairly low. For higher \( K_t \)-values the limiting line (eq. (8.13) moves to the left and point D goes to a lower \( \sigma_m \)-value. For \( 0 < \sigma_m < (\sigma_m)_D \) application of \( K_t \) to \( \sigma_a \) and \( \sigma_m \) is still justified. For \( \sigma_m > (\sigma_m)_D \) an approximation is a linear drop of \( \sigma_{fK} \) between point D and the point \( \sigma_a = 0, \sigma_m = \sigma_u \). Better approximations are possible by calculating \( \sigma_{peak} (> \sigma_{0.2}) \) by the method discussed in section 3.3 (fig. 3.12) but this will not be carried on here.

8.3 THE NOTCH EFFECT FOR FINITE LIFE

In the previous section 8.2 the notch effect on the fatigue limit was discussed. The fatigue limit is the fatigue strength for infinite life. The subject of this section is the fatigue strength for finite life, i.e. the estimation of the fatigue curve (S-N curve). For the estimation of the fatigue limit of a notched element the physical basis adopted was the similarity between crack nucleation in unnotched and notched material. Such a direct similarity cannot be indicated for finite life problems, because now both a crack initiation period and a crack propagation period are involved. Both periods should be predicted and the problem how to define the end of the initiation period and the beginning of the propagation period has to be solved. This is still beyond the present capabilities. Moreover for increasing stress amplitudes more plastic deformation at the notch root will occur. This further complicates the problem.

An estimation of fatigue curves is possible by relying on observed empirical trends. It was discussed in chapter 7 that a fatigue curve has
two horizontal asymptotes, see figure 8.14. The lower one is the fatigue limit $\sigma_{fk}$ estimated in the previous section. The upper one is related to $\sigma_U$ by $\sigma_{\text{max}} = \sigma_a + \sigma_m = \sigma_U$. The fatigue curve between these two asymptotes has now to be estimated. It was also pointed out in chapter 7 that for a large part of the fatigue curve the Basquin relation appeared to be applicable. In figure 8.14 it implies that an approximate fatigue curve could be drawn up if the fatigue lives $N_A$ and $N_B$ were known. Experience has shown that $N_B$ is in the order of $10^6$ or higher (for Al-alloys). A reasonably safe lower estimate seems to be $N_B = 10^6$. The value of $N_A$ can be in the order of $10^3$ for unnotched material, whereas for notched specimens lower values are applicable. A reasonably safe estimate here seems to be $N_A = 10^2$. Thus a rather simple method is obtained. It should be realized, however, that it gives a rough approximation only. It should not be a surprise if test results indicate fatigue lives 2 to 3 times larger (or even more, and sometimes smaller).

Similar to the definition of $K_f$ for the fatigue limit ($K_f = \sigma_f / \sigma_{fk}$) we can define a fatigue strength reduction factor for finite lives:

$$K_f(N) = \left( \frac{\sigma_{ak}}{\sigma_{ak \text{ same } N}} \right)^{\frac{1}{2}}$$

(8.14)

Figure 8.15 shows the trends that will be observed then. $K_f$ is maximal for the fatigue limit (high $N$) and minimal for low endurances. The notch effect is large if we consider high endurances, where the behaviour at the notches is predominantly elastic (macroscopically). The notch effect is much smaller at low endurances, where plastic deformations become significant. Statically ($N = 1$, or $N = \frac{1}{2}$?) the notch effect is really small provided $K_t$ is not too high, as discussed in chapter 5.

8.4 FATIGUE CRACK PROPAGATION AND THE STRESS INTENSITY FACTOR

A simple crack propagation test can be carried out in a sheet specimen with a central crack, see figure 8.16. The specimen is provided with a sharp central notch for rapid crack initiation, because we are now interested in crack growth instead of initiation. A crack will grow from both sides of the starter notch to form a single crack of length $2a$. In this type of specimen crack growth usually occurs fairly symmetric to the left and right hand side. During the crack propagation
tests the crack length is measured periodically as a function of the number (n) of applied load cycles. In this chapter we consider CA-loading only. Crack length measurements can be done visually. For that purpose a millimeter scale should be applied to the surface of the specimen. Automatic crack length recording systems have also been developed.

The first result obtained from a crack propagation test is a crack growth curve, see figure 8.17a. The first derivative of this curve is the crack growth rate: \( \frac{da}{dn} \). Note that \( \frac{da}{dn} \) is the crack extension per cycle, that means it is a crack length increment \( \Delta a \) occurring in one cycle. Figure 8.17b shows the crack growth rate as a function of the crack length. For a specimen with a central crack the crack will grow faster when the crack length increases. Different results are obtained for different cyclic stresses. Similar values of \( \frac{da}{dn} \) do occur at low and high cyclic stresses, but at different values of the crack length. It will now be shown that similar values of the crack rate do occur for similar values of the stress intensity factor.

A cyclic stress requires two values to define its magnitude (see figure 8.18a), e.g. \( S \) and \( S_{\text{min}} \), or \( S_{\text{a}} \) and \( S_{\text{min}} \), or \( \Delta S = S_{\text{max}} - S_{\text{min}} \) and \( R = S_{\text{min}} / S_{\text{max}} \). (Note: both symbols \( \sigma \) and \( S \) are used for stress. We prefer to use \( S \) if it represents an applied stress and \( \sigma \) for a stress inside the material or for strength, viz. \( \sigma_{0.2} \), \( \sigma_{U} \), \( \sigma_{f} \), etc.). As discussed in chapter 4 the stresses in the crack tip area can be described by:

\[
\sigma_{ij} = \frac{K}{\sqrt{2\pi r}} f_{ij}(\theta) \quad \text{with} \quad K = CS\sqrt{a} \tag{8.15}
\]

where C is the geometry factor. If there is a cyclic stress, there is a corresponding cyclic stress intensity factor variation, see figure 8.18b. We now should recognize what this means physically. In the crack tip area the stress intensity varies between \( K_{\text{min}} \) and \( K_{\text{max}} \). The crack extension in a cycle (\( \Delta a = \frac{da}{dn} \)) should be a function of \( K_{\text{min}} \) and \( K_{\text{max}} \). These quantities are the upper and lower limit of the stress intensity at the crack tip. In view of:

\[
\Delta K = K_{\text{max}} - K_{\text{min}} \quad \text{with} \quad R = K_{\text{min}} / K_{\text{max}} \tag{8.16}
\]
it may also be said that the crack growth rate will be a function of \( \Delta K \) and \( R \). We now compare two specimens, see figure 8.19, one specimen with a small crack and another one with a large crack. The cyclic loads are also different, i.e. a high cyclic load for the specimen with the small crack, and a low cyclic load for the specimen with the large crack. The loads and the crack length values are chosen in such a way that the same \( \Delta K \) and \( R \) apply to both specimens. Because the \( K \) variation will then be the same it should be expected that also the crack growth rate will be the same. In other words, the crack rate will be determined by \( \Delta K \) and \( R \):

\[
\frac{da}{dn} = f(\Delta K, R) \quad (8.17)
\]

Actually the above reasoning is another example of the similarity concept:

\[\text{similar conditions applied to: the same system will cause: similar consequences}
\]

\(\text{(same } K\text{-variation) + (same material) } \rightarrow \text{(same crack rates)}\)

In the previous section on the notch effect we have seen:

\(\text{(same stress cycle) + (same material) } \rightarrow \text{(same fatigue limit: } K_f = K_s)\)

The similarity concept is physically sound, but it should always be carefully examined whether the required similarity is satisfied. The similarity concept is tacitly applied to several predictions in Applied Mechanics. A full knowledge of a fracture mechanism going on in the system (crack tip extension in a fatigue cycle) is not required, since we do not state anything more than saying that the same results will be found if the conditions for the same material system are the same. The result is the crack extension \( \Delta a = da/dn \).

For crack propagation tests, carried out under different cyclic loads, but the same stress ratio \( R \), equation (8.17) may be written as:

\[
\frac{da}{dn} = f_R(\Delta K) \quad (8.18)
\]

Consequently the two tests in figure 8.17, if the same \( R \) applies, will
contribute to a single function \( f_R \). This is illustrated by the results in figure 8.20. Tests with different amplitude \( S_a \) but the same \( R \) value produced data points along a single curve. Only if \( R \) is changed another function is found.

If the \( da/dn-\Delta K \)-relation (usually plotted in a double-logarithmic graph) is extended by test results to still lower and higher \( \Delta K \)-values a sigmoidal curve with two vertical asymptotes is obtained, see figure 8.21. The left one has to be associated with a minimum \( \Delta K \) required to propagate a crack, again a threshold value: \( \Delta K_{\text{thr}} \). The right asymptote is related to extremely large crack extensions per cycle, as obtained in a static test (stable crack extensions, fig. 5.14). The corresponding \( K \)-value is thus related to \( K_c \). In region II there is an approximately linear relation between \( \log da/dn \) and \( \log \Delta K \) which implies a relation:

\[
\frac{da}{dn} = a(\Delta K)^b
\]  

(8.19)

This equation is known as the Paris relation. Originally it was thought that the effect of \( R \) on crack growth was small enough to ignore it. However, there is sufficient evidence to show that there is a systematic \( R \) effect (e.g. see figure 8.20). A well-known equation, that accounts for \( R \) is the Forman equation:

\[
\frac{da}{dn} = \frac{a\Delta K^b}{(1 - R)(K_c - K_{\text{max}})}
\]  

(8.20)

This equation also represents the asymptotic behaviour for high crack rates, because for \( K_{\text{max}} \to K_c \) the result is \( da/dn \to \infty \). The other asymptote is not included, but analytic relations which do include both asymptotes have been proposed in the literature. Some are mentioned in Ref. 4. It should be understood that the Paris relation, the Forman relation and other proposed relations are empirical relations, fitted to the trends as observed in crack propagation tests.

8.5 FATIGUE CRACK PROPAGATION PROPERTIES OF A MATERIAL AND THE STRESS INTENSITY FACTOR

As shown in the previous section there is a relation between the crack growth rate \( (da/dn) \) and the stress intensity factor \( (\Delta K) \) for a constant \( R \)-value. The relation can be determined experimentally by tests on specimens as shown in figure 8.16, but other types of specimens can be
used also, e.g. the compact tension specimen (figure 4.5). The relation between da/dn and ΔK, presented either in graphical form or analytically, is characteristic for the fatigue crack growth resistance of a material. Materials can be compared by comparing these relationships. Two examples are given in figure 8.22. Figure 8.22a indicates that the stronger alloy 7075-T6 has an inferior crack growth behaviour compared to the lower strength 2024-T3 alloy. This is one of the reasons to prefer 2024-T3 for fatigue critical aircraft components instead of the stronger 7075-T6. Figure 8.22b shows a similar trend. Artificial aging of 2024 (T8 condition) gives a significantly higher yield stress and a much lower elongation. This loss of ductility has led to approximately two times higher crack growth rates.

A comparison between an Al-alloy, a Ti-alloy and several steels is made in figure 8.23 (Ref. 7). The yield stress and tensile stress are highly different for these materials, which is also true for the specific weight (2.8, 4.5 and 7.85 respectively). For this reason the comparison has not been made for the same ΔK. Instead of the stress ΔS the relative stress ΔS/σ0.2 was introduced in ΔK. The graph in this way gives a first indication of the fatigue crack growth resistance of highly different materials. Apparently the high strength steel gives the highest crack rates. In the discussion on KIC it was already noted that steels with a very high strength are sensitive for cracks. The lower strength steels are much more tolerant, with the Al-alloy and the Ti-alloy at intermediate positions. The best results in figure 8.23 were obtained for stainless steel. However, this highly ductile material with a relatively low σ0.2 is not a structural material.

The presentation of crack growth data as da/dn versus ΔK can also be used to illustrate various effects on fatigue. Only one illustration will be given here, which is the effect of the environment on fatigue crack growth. Figure 8.24 shows crack growth data obtained in salt water, in normal air and in vacuum. Large differences in crack growth rates are found especially for lower ΔK-values. At high ΔK-values the general trend is that the effect of corrosion on fatigue crack growth becomes much smaller. At the high crack rate then obtained, the failure mechanism is an almost quasi-static phenomenon which is no longer so sensitive to corrosion.
8.6 PREDICTION OF FATIGUE CRACK GROWTH BASED ON THE STRESS INTENSITY FACTOR

The similarity concept opens a direct possibility for the prediction of crack growth rates in a structure. The $da/dn$-$\Delta K$ relation obtained with laboratory experiments is used as material data for the prediction. The second step to be made is the calculation of $K$ values for cracks in the structure. A first illustration is given in figure 8.25 for a lug (Ref. 9). For a cyclic load of $\Delta S = 80$ N/mm$^2$ ($R = 1/3$) the value of $K$ was calculated as a function of the crack length $(a)$ by adopting a solution available in the literature (fig. 8.25c). With these $K$ values the crack growth rate is read in the basic material diagram (fig. 8.25b). A plot can thus be made for the growth rate $da/dn$ as a function of $a$ (figure 8.25a). The result is compared with experimental data and the agreement is quite good for a fatigue prediction. A further step (not made in figure 8.25) is to derive a crack growth curve $(a-n)$ from the predicted $da/dn-a$ curve. This can be done by a simple integration procedure.

A second example is illustrated by figure 8.26 (Ref. 10). For a stiffened panel the $K$-value can be calculated for a crack in the skin, and the stiffeners still unbroken. The $K$-value goes up and down, which is caused by the unbroken stiffeners. Usually longer cracks have higher $K$-values. However, if the crack is growing towards a stiffener there will be a restraint of the stiffener on crack opening. As a result $K$ becomes smaller until it has passed the stiffener far enough. Then $K$ increases again until the next stiffener is approached. The relation between $K$ and the crack length was obtained by a finite-element calculation. A comparison between predicted and experimental crack growth rates is shown in figure 8.27. Again the agreement is considered to be good.

A last example shown in figure 8.28 is not highly relevant from the point of view of structural application, but it is an interesting confirmation of the validity of the similarity concept. Two types of specimens were tested, one specimen with a central crack, which gives an increasing $K$ for a growing crack, and consequently an increasing crack growth rate. The second type of specimen has crack edge loading, which gives a decreasing $K$ for a growing crack (see also figure 4.12), and as a result a decreasing crack growth rate. If the crack rate data of both types of specimen are plotted in the same graph of $da/dn$ versus $\Delta K$, the results are found to be in the same scatter band. The results for the second type could have been predicted from the results of the first
type of specimen.

Limitations on prediction possibilities

Until now the discussion was restricted to the crack growth rates of through cracks. In sheet material of a low thickness macrocracks usually are through cracks. However, in thick sections cracks frequently start as corner cracks or surface cracks as briefly discussed in chapter 4. The similarity concept leading to $da/dn = f(\Delta K, R)$ seems to work also for these types of cracks with curved crack fronts. The problem to obtain $K$ values, which vary along a curved crack front (section 4.8), is more difficult, because instead of a 2D-problem (2-dimensional configuration) we now have to solve a 3D-problem. Reliable $K$-values have to come from finite-element calculations. More solutions are coming available but a satisfactory development is not yet reached.

Another limitation is the question of very small cracks. As explained in section 4.10 (see figure 4.21) the unavoidable plastic zone should be considerably smaller than the crack length. There are several reasons why the validity of the $K$ concept will become difficult for very small cracks (Ref. 12). Actually it is surprising (and convenient) that the $da/dn = f(\Delta K, R)$ relation appears to hold for cracks in the order of 1 mm and even smaller. The validity also depends on the type of material. In this respect the high-strength low-ductility materials used in aircraft structures react more favourably than materials with a relatively low yield stress.

REFERENCES


Chapter 9  FATIGUE UNDER VARIABLE-AMPLITUDE LOADING

Contents:
9.1 Introduction
9.2 Fatigue damage, damage accumulation and interaction effects
9.3 Some results of tests with variable-amplitude loading
9.4 Macrocrack growth and variable-amplitude loading
9.5 Some concluding remarks on fatigue prediction for VA-loading
References

9.1 INTRODUCTION

In the previous chapter predictions of fatigue properties under constant-amplitude loading (CA-loading) were considered. In this chapter we will deal with fatigue under variable-amplitude loading (VA-loading). In service many fatigue loads do not have a constant amplitude and mean value. Thus predictions on the fatigue behaviour under VA-loading are technically relevant. In the present chapter the discussion will be restricted to fundamental aspects of fatigue damage accumulation under VA-loading, both with respect to crack initiation and crack growth. For this purpose simple examples of VA-loading will be considered only. The more realistic random load sequences occurring in service are the subject of chapter 11. A fundamental understanding of fatigue damage accumulation under VA-loading is important in view of highly practical questions discussed in chapters 11 and 12. Topics to be discussed here are:
- what is fatigue damage, and how does it accumulate?
- is fatigue damage accumulation similar for the crack initiation period and the crack growth period?
- what are the shortcomings of the simple Miner rule (Σ n/N = 1)?

9.2 FATIGUE DAMAGE, DAMAGE ACCUMULATION AND INTERACTION EFFECTS

If a specimen is fatigue tested until x % of its fatigue life, fatigue damage is obviously present in the specimen. The damage may still be invisible, but the life at the cyclic test load has been reduced to (100 - x) %. So there must be damage. In terms of the discussion in chapter 7 on material fatigue a first approach to identify the fatigue damage would be: a fatigue crack. A large crack would imply more fatigue damage than a small crack, but even microcracks are already real fatigue
damage. Fatigue damage accumulation thus appears to be equivalent to fatigue crack growth. A significant question now is whether the size of the fatigue crack gives a unique and sufficient indication of the fatigue damaged condition of the specimen. For the analysis of this question we will discuss crack growth under the most simple VA-loading, i.e. a cyclic load sequence in which the amplitude is changed only once, see figure 9.1.

In figure 9.1a the first stress amplitude $\sigma_{a1}$ after $n_1$ cycles is changed to a lower value $\sigma_{a2}$, and after an additional $n_2$ cycles failure occurs. For simplicity it is assumed that the size of the crack is fully described by one dimension (a). The crack growth curves for $\sigma_{a1}$ and $\sigma_{a2}$ under CA-loading are shown in figure 9.1b. Under the VA-loading of figure 9.1a crack growth will start at $\sigma_{a1}$ until a crack length $a_1$. With this crack length the test is continued at $\sigma_{a2}$. If the continued crack growth would be the same as for CA-loading, crack growth will occur along the line BC. It was Pålmgren (Ref. 1), who in 1927 assumed that $n_1$ cycles at $\sigma_{a1}$ consume a percentage $n_1/N_1$ of the fatigue life ($N_1$ is the total life at $\sigma_{a1}$ under CA-loading). Similarly $n_2$ cycles at $\sigma_{a2}$ consume a percentage $n_2/N_2$ of the life. Since there is full life of 100% the obvious result is:

$$\frac{n_1}{N_1} + \frac{n_2}{N_2} = 1 \quad (9.1)$$

or if more blocks of cycles are applied, the failure criterion becomes:

$$\sum \frac{n_i}{N_i} = 1 \quad (9.2)$$

Summing has to be done over all successive stress amplitudes (rank number i). This wellknown rule is frequently indicated as the Miner rule, sometimes as the Pålmgren-Miner, and sometimes as the linear damage rule. Miner published the rule in 1945 (Ref. 2), and before and after that year it was published a few times more. We will refer to it as the Pålmgren-Miner rule.

If we now return to the assumed crack growth history in figure 9.1b there is an interesting consequence if the Pålmgren-Miner rule is valid.
The same figure can be replotted with the crack length as a function of the relative number of cycles \( n/N \) instead of the absolute number \( n \). This has been done in figure 9.1c. Both curves for \( \sigma_{a1} \) and \( \sigma_{a2} \) under CA-loading should run from \( n/N = 0 \) to \( n/N = 1 \) at failure. In this figure \( n_1/N_1 \) and \( n_2/N_2 \) are also indicated. It will be clear that \( n_1/N_1 + n_2/N_2 = 1 \) requires that \( A' \) and \( B' \) coincide! Since this should be true for any \( a_1 \) the validity of the Pålmgren-Miner rule implies that the relative crack growth curve \( a(n/N) \) is the same for all stress amplitudes. In other words: the fatigue damage accumulation (i.e. crack growth) should relatively be the same process at all amplitudes. If we consider figure 7.1 it is clear that this is not true for low amplitudes. The same crack length is not reached after the same percentage of life.

Miner's definition (Ref. 2) of damage was not crack length but "consumed energy" (!?) and moreover he suggested a linear relation with \( n/N \) (see figure 9.2a). For that reason equation (9.2) is sometimes referred to as the linear damage rule. Shanley (Ref. 3) assumed fatigue damage to be equivalent to crack length and a single relative crack growth curve \( a(n/N) \) to be valid for all \( \sigma_a \). Such a non-linear curve (fig. 9.2b) (Shanley on doubtful assumptions arrived at a power function) equally well leads to the Pålmgren-Miner rule as shown before.

**Elementary shortcomings of the Pålmgren-Miner rule**

Three elementary shortcomings of the Pålmgren-Miner rule will be discussed by comparing the two simple load sequences in figure 9.3. The stress amplitude is changed only once in the test: an increase in figure 9.3a and a decrease in figure 9.3b.

1. **Cycles below the fatigue limit**

Assume \( \sigma_{a1} \) is below the fatigue limit. Then \( n_1/N_1 = 0 \) because \( N_1 = \infty \). In figure 9.3a no damage is done at point A (no crack nucleation because \( \sigma_{a1} < \sigma_f \)) and at \( \sigma_{a2} \) it should be found that \( n_2 = N_2 \) and consequently \( \Sigma n/N = n_2/N_2 = 1 \). However in figure 9.3b at point A crack nucleation has occurred already and a small crack is present. Although \( \sigma_{a2} \) is below the fatigue limit, the amplitude is large enough to propagate an existing
crack. In other words cycles below the fatigue limit can be damaging if cracks are created by cycles above the fatigue limit. The Pålmgren-Miner rule ignores this which can lead to unsafe results ($\Sigma n/N < 1$).

(2) Residual stress at notch root

Now assume that $\sigma_{a2}$ is large enough to produce some plastic deformation at the root of the notch. In figure 9.3a this will not occur before point A is reached. However, in figure 9.3b it has occurred before the amplitude is reduced to $\sigma_{a2}$. As a result of this plastic deformation a residual stress will be present at the notch root. This residual compressive stress is favourable for the cycles at $\sigma_{a2}$. The actual damage contribution will be less than $n_2/N_2$, because $N_2$ was obtained in a test with CA-loading in which the residual stress did not exist. Underestimation of the real damage will lead to a test result $\Sigma n/N > 1$. An illustration of this effect is shown in figure 9.4. In figure 9.4a the last maximum stress before changing the amplitude is positive, and as a result it will leave a negative residual stress at the notch. Consequently a longer life should be expected at $\sigma_{a2}$. However, in figure 9.4b the opposite occurs. The last maximum stress before changing the amplitude is a negative one and it leaves a residual tensile stress, which shortens the life at $\sigma_{a2}$. These trends are clearly reflected in the $\Sigma n/N$ values.

(3) Crack length at failure

Assume that a crack length $a = 2$ mm (as an example) is reached at point A in figure 9.3a, while complete failure would occur at $a = 20$ mm. So a significant part of the life for crack growth from 2 to 25 mm appears to be left. However, after changing over to the high amplitude $\sigma_{a2}$ the stress level may be high enough to cause failure in the first cycle at $\sigma_{a2}$. Obviously that would lead to $\Sigma n/N < 1$. In the reversed sequence, figure 9.3b, assume that the test at $\sigma_{a1}$ is continued to a moment slightly before failure at that stress level. Then $a$ is almost 2 mm and $n_1/N_1$ almost equal to 1. At the lower stress amplitude $\sigma_{a2}$ crack growth can be continued until $a = 20$ mm at failure. A value $\Sigma n/N > 1$ should be expected. These deviations of the Pålmgren-Miner rule are due to the fact that $N$ is
defined as the number of cycles until failure, irrespective of the question whether the fatigue crack at that moment is small (for high $\sigma_a$) or large (for low $\sigma_a$). The fatigue curve is not a line of constant fatigue damage, i.e. of constant crack length, see figure 9.5.

**Significance of the shortcomings**

The last shortcoming is a fundamentally correct objection, but it is not a very significant one, because the fatigue life spent in the macro range is relatively small. This shortcoming will therefore induce small errors only. However, the two other shortcomings are really significant. Errors introduced are highly depending on the load spectrum. Let us consider a lead spectrum with relatively few load cycles above the fatigue limit and numerous load cycles below that limit. The numerous last ones according to the Palmgren-Miner rule are not doing any damage. However, if the cycles above the limit take care of crack initiation the many cycles below $\sigma_f$ will cause a finite life. Values of $\Sigma n/N$ much lower than 1 are quite well possible. A life estimate based on $\Sigma n/N = 1$ could be highly unsafe. One drastic step to arrive at conservative life estimates is to extend the fatigue curve of the finite life area to lower amplitudes by a linear extrapolation (on double log scale, Basquin relation, slope $x$), see figure 9.6. A somewhat less drastic step was suggested by Haibach. Although these steps open a possibility to be more conservative it cannot be denied that such fatigue curve extensions are fairly arbitrary. In chapter 12 this point will be raised again.

The shortcoming related to residual stresses, as illustrated by figure 9.4, is significant both for predictions and for carrying out realistic service-simulation tests. The arguments discussed before can also be formulated as follows: The load cycles at $\sigma_{a1}$ will affect the damage accumulation at $\sigma_{a2}$ or in still other words the damage at $\sigma_{a2}$ is depending on the preceding load history. These effects are referred to as interaction effects, i.e. effects on damage accumulation due to load cycles of different magnitudes. During VA-loading such interaction effects can be very significant. The comparison between the two load histories in figure 9.3a and b indicated that the sequence of the two stress amplitudes clearly affect the fatigue life. There is a sequence effect as a result of interaction effects.
Definition of fatigue damage

It was suggested before that the fatigue crack is an essential part of fatigue damage. It will be clear that a second and significant aspect is the residual stress at the root of a notch, which is a consequence of local plastic deformations induced by the VA-loading. A formal definition of fatigue damage is that it includes a description of all changes in the material as caused by the fatigue load. Here it can be defined as:

\[
\text{Fatigue damage} = \begin{pmatrix}
- \text{amount of cracking} \\
- \text{induced plastic deformations} \\
- \text{resulting residual stress distribution}
\end{pmatrix}
\]

With this definition it is very difficult to arrive at a rational fatigue life prediction method. The main reason is that fatigue damage no longer can be specified by a single number, such as the length of a crack. A minimum of two is necessary at least, e.g. crack length and residual stress. Such a simplification ignores that there will also be strain hardening, and that a "residual stress" represents a residual stress distribution, which in itself cannot be fully described by a single number. Some modern prediction techniques try to calculate the local strain history at the notch root, which then includes local plasticity and residual stress. This notch root strain-history is then used for a life calculation employing \( \sum \frac{n_c}{N_c} = 1 \). The life \( N_c \) now is the fatigue life under constant strain-amplitude cycles, while \( n_c \) is the number of such cycles occurring under the VA-loading. Satisfactory prediction techniques for application to aircraft structure have not yet been achieved. For the time being we will have to live with the simple Pålmgren-Miner rule. Under such conditions we should know and understand the shortcomings of this rule in order to decide for which purposes the rule might be used and for which other purposes its use should be avoided.

9.3 SOME RESULTS OF TESTS WITH VARIABLE-AMPLITUDE LOADING

Many test programs carried out to check the validity of the Pålmgren-Miner rule are reported in the literature. Here only a few illustrative examples for simple VA-load sequences will be given. In figure 9.7 results of an
NLR-investigation on unnotched and notched specimens are summarized. In several test series the amplitude was changed only once, in other test series it varied periodically between two levels. The figure shows that for unnotched specimens large deviations of \( \Sigma n/N = 1 \) were not found, contrary to the results for the notched specimens. It should be expected that local plasticity at the notch will be significant only for notched specimens. In unnotched specimens a redistribution of the stress will not occur, at least not on a macro level. Consequently interaction effects due to plasticity are not likely in unnotched specimens.

Similar to the results in figure 9.4 it is also evident in figure 9.7 that a high-low sequence of \( \sigma_a \) produced the largest \( \Sigma n/N \) values, but they are much higher now than in figure 9.4. There is one difference between the two figures. In figure 9.4 \( \sigma_m = 0 \) whereas in figure 9.7 a positive \( \sigma_m \) applies (\( R = 0 \)). It should be realized that positive mean stresses increase the probability of exceeding the tensile yield stress (see figure 9.8) and decrease the probability of exceeding the compressive yield stress. The former exceedance gives residual compressive stresses and it is encouraging to know that these stresses are more likely to occur than residual tensile stresses. In view of the application of \( \Sigma n/N = 1 \) this trend is technically relevant.

Another set of test results is presented in figure 9.9 (Ref. 5). Heywood collected date of tests on various specimens, including a tail plane, which were fatigue tested under CA-loading, both with a high preload before the fatigue test and without such a preload. The results indicate that a positive preload can extend the life considerably. Note that the magnitude of the preload is given as a percentage of the 0.1% yield stress, and the fatigue life as a ratio between fatigue life with and without preload. Possible increases apparently can exceed a factor 100 if the preload is in the order of \( \sigma_{0.1} \). Beyond any doubt this is a result of plastic deformation in the notched elements, which then leads to residual compressive stress (Figures like this one have prompted ideas about preloading aircraft before they go into service). Figure 9.9 also shows that negative preloads cause the reverse effect, which again should be attributed to residual stresses, which in this case are unfavourable residual tensile stresses.
Program-fatigue test

Before the last example of test results will be discussed it has to be explained first what a program-fatigue test is. Figure 9.10 shows that the amplitude is systematically varied in a program-fatigue test, in this case in an ascending-descending order (also Lo-Hi-Lo sequence). The number of load cycles at each amplitude can be chosen in agreement with a load spectrum, for instance a gust load spectrum. If one period with all load cycles has been completed, it is repeated until failure occurs. This type of testing was introduced by Gassner (1939) in order to apply a more realistic load sequence than a CA-loading. A still more realistic random load sequence could not be applied by commercial fatigue machines until machines with closed loop systems were developed. A closed loop machine can apply any load sequence which can be generated as an electrical signal. With a small computer attached to a modern fatigue machine this is quite easy and a new generation of flexible testing equipment came thus available. However, useful information obtained in the older program fatigue tests can still be instructive. Figure 9.11 shows data from program fatigue tests with only 4 or 5 different amplitudes, applied in an ascending (Lo-Hi) sequence only. The fatigue life is not shown in cycles but in numbers of program periods. Since CA-data were also available, $\Sigma n/N$-values could be calculated.

The comparison to be made in figure 9.11a shows that a repetition of high loads is much more effective than just one cyclic high load. A single high load can extend the life, until a micro crack after an initial delay starts to grow again. A new high load will then restore delay effects by introducing plastic deformation at the crack tip, with apparently much success.

Figure 9.11b shows the comparison between program tests without and with periodic high load cycles. The effect of the high load cycles is negative, contrary to the results in figure 9.11a. The explanation is that the favourable effect of the positive peak is more than annihilated by the immediately following downward load. The last one of the two high amplitudes dominates the result. This is confirmed by the results in figure 9.11c where the positive high load follows the negative high load.
In figure 9.11 \( \Sigma n/N \) values are also indicated. It should be understood that one high load in each period gives a very small increase of the damage sum \( \Sigma n/N \) of a period, because it is adding only one cycle. If the Pålmgren-Miner rule were correct the addition of one high load per period should thus give a very small reduction of the fatigue life. In reality a large increase of life was found. Apparently the Pålmgren-Miner rule is fully incapable to indicate the damage of these extreme loads.

9.4 MACROCRACK GROWTH AND VARIABLE-AMPLITUDE LOADING

In chapter 7 crack growth was described as a fracture process with a small crack length increment \( \Delta a \) in each cycle. In principle it should be expected that crack extension in every cycle will also occur under VA-loading. Crack growth could then be described by:

\[
a = a_0 + \Sigma \Delta a_i
\]  

(9.3)

\( a_0 \) is some starting crack length, and \( \Delta a_i \) is the crack length increment in cycle number \( i \). As a first approach \( \Delta a_i \) in cycle \( i \) can be assumed to depend on \( \Delta K \) and \( R \) in cycle \( i \) only, and thus should be equal to the crack rate as obtained in a CA test at the same value of \( \Delta K \) and \( R \).

\[
\Delta a = f (\Delta K, R)
\]

(9.4)

If that would be true for all load cycles the prediction of crack growth is a simple summation according to equation (9.3) with the relation in equation (9.4) as basic material data.

In terms used before the above procedure would imply that interaction effects do not occur, because it was assumed that \( \Delta a \) does not depend on the preceding load cycles. Tests have learned, however, that the dependence can be quite large. Some results obtained in tests with simple VA-load sequences will be discussed to illustrate the interaction effects during macrocrack growth. There is a certain similarity with interaction effects discussed in the previous section with respect to fatigue life of notched specimens. There are also some essential differences.
Figure 9.12 shows three crack growth curves. Curve A was obtained under CA-loading. Curve C is the result due to applying a few high positive loads. Apparently such a high load gives a large crack growth retardation, while the delay is larger if the high load is applied at a higher crack length. When the delay is over the crack growth curve is again parallel to curve A. A high positive load introduce a plastic zone at the crack tip. The plastic zone has been elongated and it is larger than it was before. As a result of the residual plastic deformations residual compressive stresses are present in this zone. The crack has to grow through the zone, but the residual stresses will try to keep the crack closed, and that will retard crack growth considerably. If the high positive load is immediately followed by a high downward amplitude, see for curve B in figure 9.12, then the crack growth retardation is much smaller. However, the retardation did not fully disappear, and certainly there is no acceleration. Here there is an essential difference with notches. Negative loads can reduce the fatigue life of notched elements by introducing residual tensile stresses, see figure 9.9. Also a positive/negative load cycle can reduce the life, see figure 9.11b. Why is the situation different for a crack? If a negative load is applied the crack is closed and the two crack surfaces are pressed on one another. Load transmission through the crack surfaces is possible. Actually the crack is no longer a stress raiser. As a first approximation the material under compression behaves as if there is no crack. As a consequence it is impossible to create crack tip plastic zones in compression. The only thing which is possible is that a plastic zone introduced by a tensile load, will see some reversed plastic deformation during unloading to zero. A full reversion will not occur and for that reason curve B still shows some retardation of the crack growth.

Crack growth delay after high loads will depend on the size of the plastic zone. Larger zones will give longer delays, which is already evident from curve C in figure 9.12 (plastic zone size \( \sim R^2 \) (eq. 4.49) \( \sim a \)). As discussed in section 4.10 the plastic zone size is about 3 times smaller for plane strain conditions (in thick plates) as compared to plane stress...
(in thin sheet material). Delays after peak load should be expected to be smaller in thick material. This is confirmed by the test results in figure 9.13.

Figure 9.14 shows the effect of periodically repeated peak loads for one and the same alloyed steel, heat treated to three different strength levels. The reduction of the crack rate is high for the high ductility material (low $\sigma_{0.2}$). In this material after a peak load a larger plastic zone will occur than in a low ductility material ($R_p = 1/\sigma_{0.2}^2$, eq. 4.49). For the maximum strength ($\sigma_{0.2} = 1400$ N/mm$^2$) the reduction of the crack rate is fairly small. Figure 9.14 shows a few other points of interest. Without peak loads the crack rate increases for the higher strength material. At the same time $K_{IC}$ decreases. This is another illustration that high-strength low-ductility materials are more sensitive to cracks.

Crack closure

Earlier in this section crack growth delays were qualitatively explained by referring to crack tip plastic zones and residual compressive stresses in these zones. A more detailed explanation is based on the crack closure concept, developed by Elber (Ref. 9). It implies that a crack can be closed at a positive tensile load, because there is plastic deformation in the wake of the crack. This is what happens if a crack after a peak load is growing into the plastic zone. As soon as it enters the zone there is plastically strained material behind the crack tip. That keeps the crack closed (which is just another way of referring to residual compressive stresses as done before). It requires a tensile stress, $S_{op}$, to open the crack until the very tip. The effective stress variation is therefore reduced from $\Delta S = S_{max} - S_{min}$ to $\Delta S_{eff} = S_{op}$. Along these lines of crack closure and crack opening interaction effects during crack growth can be explained in a more detailed way. A survey is given in Ref. 10. Actually crack closure also occurs during CA-loading because the crack is then growing through its own plastic zones. There is plastic deformation in the wake of each macro crack. It was shown experimentally that crack closure
does occur during CA-loading at positive stresses. Crack closure has been successfully used to explain the R-effect on crack growth (also reviewed in Ref. 10), but it will not be discussed here any further.

Predictions

Non-interaction predictions should employ equations (9.3) and (9.4). It has to be realized that equation (9.3) implies a cycle-by-cycle calculation. For each cycle \( \Delta a_i \) has to be calculated and to be added to the existing crack length. The same procedure has then to be repeated for the subsequent cycle, etc. Usually such calculations can be made on computers only, unless simplifying assumptions can be made if justified. As explained before interactions for macrocrack growth usually will be favourable and we then predict a crack growth which is faster than in reality. It was also shown that high loads can induce very large favourable interactions by creating large plastic zones and significant crack closure effects. Because plastic zone sizes can be estimated (see chapter 4.19) crack growth prediction techniques have been proposed which do account for retardations induced by plastic zones. Equation (9.3) is then replaced by:

\[
\Delta a = a_0 + \sum (\beta_i \Delta a_i)
\]

(9.5)

with \( \beta_i \) being an interaction factor depending on existing plastic zones, introduced by previous cycles. Although much better predictions were sometimes obtained, this approach is still under development.

9.5 SOME CONCLUDING REMARKS ON FATIGUE PREDICTION FOR VA-LOADING

The possibilities for fatigue life predictions under VA-loading are not excellent. Rough estimates can be obtained with the Pålmgren-Miner rule. The risk of underestimating the damage contribution of low-amplitude cycles \((\sigma_a < \sigma_f)\) can be avoided by extending the fatigue curves to higher endurances. The beneficial effect of high loads on fatigue life cannot be accounted for by the Pålmgren-Miner rule, but fortunately it is a beneficial effect.

With respect to the prediction of macrocrack growth a non-interaction prediction will generally be on the safe side. However, it can be very far on the safe side. Consequently, if a prediction indicates an
unsatisfactory result (crack growth too fast), the real crack growth
behaviour may still be acceptable.

If we are in need of more accurate information on fatigue life and
crack growth the only alternative to approximate calculations is to
carry out realistic tests. The problem of carrying out such tests is
discussed in chapter 12.

REFERENCES

1. A. Fälmgren: The endurance of ball-bearings (in German). Z.V.D.I.,
2. M.A. Miner: Cumulative damage in fatigue. J. App. Mech., Vol. 12,
   1945, pp. A 159-A 164.
3. F.R. Shanley: A proposed mechanism of fatigue failure. Colloquium
   on Fatigue (ed. by W. Weibull and F.K.G. Odquist). Springer Verlag,
   1956, pp. 251-259.
4. J. Schijve and F.A. Jacobs: Fatigue tests on notched and unnotched
   Clad 24 S-T sheet specimens to verify the cumulative damage
5. R.B. Heywood: The effect of high loads on fatigue. Colloquium on
   Fatigue (ed. by W. Weibull and F.K.G. Odquist). Springer Verlag 1956,
   pp. 92-102.
6. J. Schijve: The endurance under program fatigue testing. Full-Scale
   Fatigue Testing of Aircraft Structures, 1st ICAF Symposium,
   1961, pp. 41-59.
   fatigue crack retardation in 2024-T3 aluminium alloy. Eng. Fracture
8. G.S. Petrak: Strength level effects on fatigue crack growth and


Some general references

(a) J. Schijve: The accumulation of fatigue damage in aircraft materials and structures. AGARDograph No. 157, 1972.


(c) Post Academic Course Book: Fatigue of Structures (in Dutch) Delft 1982.
Chapter 10  
FATIGUE OF JOINTS

Contents:
10.1 Introduction
10.2 Bolt loaded in tension
10.3 Fatigue of lugs
10.4 Symmetric joints with rows of bolts or rivets
10.5 Asymmetric joints with rows of bolts or rivets
10.6 Riveted joints
10.7 Adhesive-bonded joints
10.8 Some concluding remarks
References

10.1 INTRODUCTION

Fatigue cracks in aircraft structures in many cases occur in joints: riveted joints, bolted joints, lugs. The designer should pay careful attention to the various aspects of joint design, which can affect the fatigue properties of the joint. It will turn out that a quantitative treatment of the fatigue strength of joints is not easy, but a qualitative understanding has been developed. Qualitative understanding is of prime importance to come to good technical solutions.

There are several reasons why joints can be fatigue critical items:
(1) Usually a joint is a significant "discontinuity" associated with severe stress concentrations.
(2) In many joints excentricities occur in the load transmission line. This causes secondary bending, which is an additional source of stress concentration.
(3) In many joints fretting corrosion can occur and as a result low fatigue limits are a realistic possibility.

These aspects are discussed in the present chapter. Of course we cannot deal with all varieties of joints, because the variety is very large indeed. We will restrict the discussion to the principle elements of fatigue of joints. Subjects covered are:
- bolts loaded in tension
- lug joint (pin loaded hole, with the pin loaded in shear)
- bolted and riveted joints, effect of clamping
- bonded joints.
The bolt loaded in tension is discussed first, because it is a good example to introduce the concept of "pretensioning" as a means to improve the fatigue resistance. The lug is a simple and elementary joint, with several applications in aircraft structures and a number of characteristic features. Some procedures to improve the fatigue strength of lugs are technically relevant. Bolted, riveted and bonded joints also have characteristic aspects, one of them being the occurrence of secondary bending.

10.2 BOLT LOADED IN TENSION

If a bolt in a structure will be loaded by cyclic tension it should be realized that the shape of the bolt is poor from a fatigue point of view. There are several severe stress concentrations and moreover the load transmission to the bolt is unfavourable. Stress concentrations occur at (see figure 10.1)
- the transition from bolt head to bolt shank
- groove between shank and screw thread
- screw thread.

The bolt head is loaded in an unfavourable way, because it occurs close to the root of the notch. A similar loading case was shown in chapter 2 (figure 2.18) for a T-head. Here we have a circular cross section which will give somewhat lower $K_t$-values. Unfortunately the radius between bolt head and shank cannot be made large in view of assembling(hole edges). Values of $K_t$ in order of 3 and even larger are realistic.

The groove between shank and screw threads will generally not be critical because the radius is larger than the root radius in the screw thread.

The latter root radii are standardized, unfortunately at rather low values. Moreover, the load distribution between nut and bolt (figure 10.1b) is inhomogeneous. The first screw thread turn of the bolt carries a higher load than the other ones. Fatigue cracks usually start at the first screw thread turn.

There are several possibilities to improve bolt design with respect to fatigue. Two examples are presented in figure 10.2 but many more were proposed (see e.g. Ref. 1). The bolt head shown in figure 10.2a has a
groove in the head to increase the root radius and thus decrease $K_t$. The special nut design shown in figure 10.2b implies a change of the nut stiffness along its screw threads, which results in a more homogeneous load transmission to the bolt.

**Thread cutting or rolling**

For fatigue loaded tension bolts rolled thread is highly advisable, especially cold rolled thread. A cutting operation will cut through the fibrous structure, whereas rolling will force the fibrous structure to follow the thread profile. Moreover rolling leaves a better surface quality while cold rolling will also introduce residual compressive stresses. High quality bolts always have rolled thread. The fatigue strength, especially the fatigue limit, is much better than for bolts with cut thread.

**Bolt bending to be avoided**

If the material surfaces around bolt holes are not exactly parallel and perpendicular to the bolt hole axis, the support of either the bolt head or the nut (or both) will become asymmetric. As a consequence bending will occur in the bolt at locations where high stress concentrations are already present. This can lead to disastrous results as shown by the fatigue curves in figure 10.3. For cyclically loaded tension bolts flat-milled surfaces have to be provided for. An example how this was done in a most unfortunate way is shown in figure 10.4. It caused a very sharp corner, from which a fatigue crack started.

**Pretension in bolts**

During assembling it is usual to tighten a bolt by applying a prescribed torque moment to the nut. This introduces a pretension load in the bolt. If a cyclic load is now applied to the structure the pretension will increase the mean stress in the bolt, but it will reduce the stress amplitude. In view of the predominant effect of the stress amplitude significant gains of the fatigue strength can be obtained, see the fatigue curves in figure 10.5. The effect of pretensioning on fatigue is
sufficiently important to give a brief description about how it works.

In figure 10.6a two parts are clamped together by a pretensioned bolt. If the joint is still unloaded ($P = 0$) there is already a pretension load $P_o$ in the bolt. That implies that a similar compression load is present in the contact area between the two parts: $P_c = -P_o$. The elastic spring model in figure 10.6b leads to the equilibrium equation:

$$P = P_{\text{bolt}} + P_c$$  \hspace{1cm} (10.1)

The equation remains valid as long as there is contact between the two parts. Under load the bolt will become somewhat longer and $P_{\text{bolt}}$ will increase:

$$P_{\text{bolt}} = P_o + \Delta P_{\text{bolt}}$$  \hspace{1cm} (10.2)

Similarly for the contact area:

$$P_c = -P_o + \Delta P_c$$  \hspace{1cm} (10.3)

Substitution in equation (10.1) gives:

$$P = \Delta P_{\text{bolt}} + \Delta P_c$$  \hspace{1cm} (10.4)

which also could have been written up directly. According to the elastic-spring model all springs will become longer under load $P$. This implies $\Delta P_c > 0$ (it will be less compressive than in the unloaded state $P = 0$). With $\Delta P_c > 0$ it is clear from equation (10.4) that:

$$\Delta P_{\text{bolt}} < P$$  \hspace{1cm} (10.5)

In other words: The load increment on the bolt is smaller than the load on the joint. Consequently, apart from the bolt there must be another path for load transmission from one part to the other one. The other path goes through the contact area. Again in other words: As long as the contact between the two parts is maintained in the contact area, the two parts act as a single part. If the contact area had been a soldered joint the load transmission would have been the same.
How much will $\Delta P_{\text{bolt}}$ be smaller than $P$? This depends on the stiffness of the bolt ($C_{\text{bolt}}$) and the stiffness of the parts clamped between the bolts ($C_c$). For a bolt elongation $\Delta l$ we then can write:

$$
\begin{align*}
\Delta P_{\text{bolt}} &= C_{\text{bolt}} \Delta l \\
\Delta P_c &= C_c \Delta l \\
\text{Eq. (10.4)} \\
P &= \Delta P_{\text{bolt}} + \Delta P_c
\end{align*}
$$

elimination of $\Delta l$:

$$
\frac{\Delta P_{\text{bolt}}}{P} = \frac{1}{1 + \frac{C_c}{C_{\text{bolt}}}} < 1 \quad (10.6)
$$

A small $\Delta P_{\text{bolt}}/P$ requires a large $C_c$ and a small $C_{\text{bolt}}$. In general the designer cannot easily change the stiffness of the clamped parts, but for the bolts it is more easy. The diameter of the bolt is a variable, while bolts with partly waisted shanks are also a possibility. Bolts of Ti-alloys are becoming more popular in aircraft, because for the same strength weight-savings are possible. For pretensioning the lower $E$-modulus of Ti is significant ($E_{\text{Ti}}/E_{\text{steel}} \approx 0.55$). One aspect should not be overlooked. If $P$ becomes high enough the contact between the two parts will be lost. In figure 10.6c this occurs at $P = P_\ell$ and with $P = P_o + \Delta P_{\text{bolt}}$ and equation (10.6) it is easy to arrive at:

$$
P_\ell = P_o \left(1 + \frac{C_{\text{bolt}}}{C_c}\right) \quad (10.7)
$$

The equation shows that $P_\ell$ will be lower for a lower bolt stiffness.

Loosing contact between the two parts can be postponed by a higher pretension load $P_o$, see figure 10.7. The effect of bolt stiffness is also illustrated by this figure.

The general concept of pretension to be remembered is: there is another path for load transmission than the usual one without pretension. That implies a lower load along the usual path.
Installing pretension in a bolt

High strength steel bolts are pretensioned to high stress levels, even up to 70 percent of $\sigma_{0.2}$ ($P_0$ calculated for the minimum diameter of the screw thread). This also helps to come to a more uniform load transfer at the screw thread turns. We have to be sure that the pretension load is really in the bolt and will stay there. Usually the pretension is controlled by tightening of the nut with a calibrated torque wrench. The relation between the torque moment and the pretension load is strongly dependent on the friction between nut and washer and in the screw threads between nut and bolt. Calculations, confirmed by measurements (Ref. 4) have shown that good lubrication can give a three times higher preload than dry friction. For critical tension bolts there are special means to assure that the required pretension has been installed.

10.3 FATIGUE OF LUGS

A lug joint was already discussed in chapter 7 (see figure 7.37). A lug joint is an elementary and useful type of joint in aircraft structures, because:

1. it allows rotation around the pin or bolt and
2. because assembling is an easy and fast procedure.

From the fatigue point of view there are some disadvantages:

a. The stress concentration factor $K_t$ is relatively high (as compared to open hole), see figure 10.8.

b. Fretting can occur inside the hole between the pin and the wall of the hole as discussed in chapter 7 (page 7/22), with an adverse effect on fatigue life.

c. In many cases lug joints are not fail-safe and failure of the lug can lead to unacceptable consequences. For instance if an all-movable tail plane is supported by three lugs, failure of one of the lugs can be dangerous. Also, in many linkage systems (e.g. in undercarriages or control systems) a lug failure is unacceptable.

As a result of a high $K_t$ and fretting corrosion, the fatigue limit can be low. This is illustrated by the fatigue curves in figures 10.9a and 10.9b.
for two Al-alloys and in figure 10.10 for an alloy-steel. In chapter 8
the estimation of the fatigue limit \( \sigma_{f_k} \) of a notched element was based
on \( K_f = K_t \) with a possibility that \( K_f \) might be smaller than \( K_t \). For lugs
\( K_f = K_t \) can easily lead to unsafe prediction. Recall that \( K_f = K_t \) was
based on a similarity concept. This is not applicable to fatigue of an
unnotched specimen and fatigue of lugs. In the former case there is no
fretting, whereas for the lug fretting is present. A direct comparison
between an unnotched specimen and a lug cannot be made. Test results
indicate \( K_f > K_t \). In view of fretting, lugs are also sensitive to a
size effect which is not related to the volume of highly stressed material,
but to the amplitude of the fretting movements inside the bore of the hole.
From fretting corrosion tests on unnotched specimens (Ref. 7) it is known
that the amplitude has a significant effect on the fatigue limit. In lugs
a similar effect is found as shown by the results in figure 10.11. For a
large lug \( \sigma_f \) is about 1.5 times lower than for a small lug of the same
shape (same \( K_t \)). It may well be expected that the fretting amplitude is directly
proportional to the size of the hole. Figure 10.11 also shows that the
size effect disappears at lower endurance. As discussed in chapter 7
such a trend should be expected. At lower fatigue lives cracks can be
initiated also without fretting and fretting does not contribute to the
crack propagation. The occurrence of fretting then has a much smaller
effect on life, see figure 7.36.

Because a \( K_f = K_t \) basis does not exist, fully empirical methods were
developed for the prediction of the fatigue strength of lugs. The first
attempt was made by Heywood (Ref. 9). He proposed:

\[
\frac{\sigma_a}{\sigma_A} = \frac{2.5}{K_t} \left( \frac{2}{1 + \frac{d}{25}} \right) \quad (d = \text{hole diameter in mm}) \quad (10.8)
\]

\( \sigma_A \) is the fatigue strength of a reference lug with \( K_t = 2.5 \) and \( d = 25 \) mm.
\( \sigma_a \) is the fatigue strength to be calculated for a lug with stress concentration-
factor \( K_t \) and hole diameter \( d \). Further \( \sigma_a \) and \( \sigma_A \) apply to the same \( N \),
while the effect of mean stress is ignored. Heywood presented the \( \sigma_a - N \) curve
for the reference lug, which he obtained by analysing many lug fatigue data.
Equation (10.8) implies that $\sigma_A$ is inversely proportional to $K_t$, while the term $2/(1 + d/25)$ accounts for the size effect. For the two lugs in figure 10.11 equation (10.8) predicts a ratio ($d = 10$ and 25 mm respectively, same $K_t$) of $(1 + 1)/(1 + 0.4) = 1.43$, which is approximately the ratio found for the fatigue limit. According to Heywood this ratio should apply to all values of $N$, but figure 10.11 shows that this is not correct. As explained before it should be expected that the ratio decreases for shorter endurance.

Larsson (Re. 10) proposed an improved empirical relation, which accounts for ratio's depending on both $N$ and $\sigma_m$

$$\frac{\sigma_A}{\sigma_m} = 1 + \theta (k_1 k_2 - 1)$$

with $k_1 = \sqrt{\frac{ad}{c^2}}$ and $k_2 = \sqrt{\frac{10}{d}}$

(d in millimeters)

while $\theta = 0.25 \log N - 0.5$ for $10^3 < N < 10^6$
and $\theta = 1$ for $N \geq 10^6$

$\sigma_A$ and $\sigma_m$ are values for a reference lug with $a = c = d = 10$ mm. For the reference lug Larsson gave fatigue diagrams for 2024-T3 and for 7075-T6, and diagrams for other materials became available in the aircraft industries later on. In Larsson's formulas $\theta$ accounts for the N-effect, $k_1$ for the shape (stress concentration) and $k_2$ for the size effect. Predictions with Larsson's method are shown in figure 10.9a and b and the agreement is quite satisfactory.

It should be stressed that the Larsson approach is a fully empirical one. It implies that Larsson has proposed formulas which apparently agree with empirical trends of a large amount of test data. At the same time it should be emphasized that predictions of the fatigue strength should actually be considered to be estimates. Errors in the order of 10 %
in terms of stress are always possible and should be accepted. More accurate
data require relevant tests.

Lugs with $\sigma_{\text{min}} < 0$

If the load on a structure is reversed (opposite sign, but same magnitude)
we expect that the sign of all stresses in the structure is also reversed
(actually this is Hooke's law). For a trivial reason this does not apply
to a lug, see figure 10.12. If the tension load $P$ is reversed to be a
compression load, the bearing pressure of the pin on the bore of the
hole is not reversed. Instead of a bearing pressure on the upper part of
the hole we see a bearing pressure on the lower part of the hole. This
is a large difference for the critical section (AB) of the lug. If the
upper part of the bolt is loaded, this load has to pass the critical
section. However, this is no longer true if the lower part of the hole
is loaded. Strain gage measurements (Ref. 11) at A and B have confirmed
this. For a compression load $P$ there is even a small tensile stress at
A (instead a high compressive peak stress) because the bearing pressure on
the lower part of the hole tries to elongate the wall of the hole. If this
small tensile stress is disregarded the consequence is that only the positive
part of a cyclic load is a fatigue load for the critical section, whereas
the negative part can be disregarded. For the fatigue diagram this implies
that for $\sigma_{\text{min}} < 0$ ($R < 0$) the fatigue strength will be characterized by
$\sigma_{\text{max}} = $ constant. An illustration is given in figure 10.13.

Improvements of the fatigue limit of a lug

There are several means to improve $\sigma_f$ of a lug:
(a) slotted holes
(b) interference fit between pin and hole
(c) hole expansion to introduce residual stresses.

Slotted holes are not applied in practice, but it is instructive to see
how this solution improves $\sigma_f$ by preventing fretting corrosion. As shown
in figure 10.14 in a slotted hole a thin layer of the wall of the hole
is removed to prevent metal contact between pin and lug, especially at
the location A, where the maximum peak stress is present. A three
times higher fatigue limit is found. There is still fretting at E but
that occurs at an unloaded corner of the lug. Cracks now initiate at
B or F.

Interference fit

If a pin with a larger diameter than the hole (negative clearance =
interference fit) is pressed into the hole there are two reasons to expect
improvements. First, in view of the pressure between the pin and the hole
fretting movements will be more difficult and thus fretting damage may
be less. The second reason is a matter of prestressing. Due to the
interference fit there is a tangential tensile stress in the lug around
the bore of the hole. This is a prestress which will increase the local
mean stress, but reduce the local stress amplitude. The latter is more
important than the first aspect (as said before the favourable effect
of prestress is associated with an additional way for load transmission:
what is the other way?). The benefit of an interference fit is illustrated
by figure 10.15. It should be pointed out that a small interference is
not very effective. Large interferences give assembly problems. One way
to overcome this is the application of tapered bolts, which also gives
interference if sufficiently tightened. A more usual solution is the
application of steel bushes to be pressed into the hole of Al-alloy lug
elements, see figure 10.16. This preserves the rotation possibility of the
lug. The pin is loading the bush, and the bush is loading the lug hole.
The contact between the bush and the lug is the significant one for fatigue
of the lug. The bush should have a thickness of at least 10% of the hole
diameter for obtaining sufficient interference. There is another practical
reason for inserting a steel bush in a hole. It prevents the hole to be
damaged by scratches etc. during assembling.

Plastic hole expansion

The hole is first produced with a slightly undersized diameter (say 4% too small). A tapered pin is then pulled through the hole to expand the
hole. This can occur only if plastic deformation does occur around the
hole, see figure 10.17. As a result of the plastic expansion the plastic
zone around the hole is larger than it was before. The elastically strained material around this zone will exert a pressure on the zone, and as a consequence there are tangential residual stresses around the hole. These compressive residual stresses can be as high as the compressive yield limit. Small distortions of the cylindrical shape of the hole can be removed afterwards by reaming if necessary. This will hardly reduce the residual stresses. Figure 10.17 shows a commercially developed pulling apparatus, which allows hole expansion for a hole accessible from one side only. A lubricated split sleeve is used during the pulling operation to avoid scratching of the hole. The sleeve is removed after the hole expansion.

The large effect of plastic hole expansion is illustrated by the fatigue curves in figure 10.18. It should be pointed out that residual stresses do not prevent fretting corrosion, see the large damage inside an expanded hole in figure 10.19a. However, it prevents crack growth initiated by the damage, see figure 10.19b. Because the high residual stress keeps the cracks closed, the direction of micro crack growth is no longer guided by the main principal tensile stress. Some erratic growth along slip planes occurs because the cyclic shear stresses, independent of the residual stress, are very high.

As shown by figure 10.18 the improvement is higher for the 7075-T6 lug than for the 2024-T3 lug. The reason is that the residual stresses, being in the order of the compressive yield stress, are significantly higher for the 7075-T6 alloy.

10.4 SYMMETRIC JOINTS WITH ROWS OF BOLTS OR RIVETS

A simple double strap joint with two rows of bolts at each side of the joint is shown in figure 10.20. It appears to be a good design from a static point of view if both rows carry the same load. \( P_A = P_B \) (same bearing pressure in the hole, same shear stress in the bolts). This will be obtained if the straps have the same thickness as the plates:

\[ 2 t_2 = t_1. \]

From a fatigue point of view the two rows are in different positions now. The plate in row B is loaded by \( P_B \) only. In row A the same load \( P_A \) is introduced, but there is also a by-pass load coming from
row B. In row A the full load is present in the plate, in row B only half that load. If $S$ is the net section stress in the plate we may write up:

$$
\left(\sigma_{\text{peak}}\right)_{\text{row A}} = \frac{1}{2} S \left( K_t \right)_{\text{loaded hole}} + \frac{1}{2} S \left( K_t \right)_{\text{unloaded hole}}
$$

from by-pass load

$$
\left(\sigma_{\text{peak}}\right)_{\text{row B}} = \frac{1}{2} S \left( K_t \right)_{\text{loaded hole}}
$$

For $d/s = 1/5$ the two $K_t$-values are 4.6 and 2.4 respectively, and then

$$
\left(\sigma_{\text{peak}}\right)_{\text{row A}} = 3.5 S \text{ and } \left(\sigma_{\text{peak}}\right)_{\text{row B}} = 2.3 S
$$

This is an approximation because for the by-pass load $K_t$ of an open unloaded hole was taken. However the hole is not really open. Nevertheless it is still correct to say that the holes in the plate in row A are more severely loaded than in the row B. For the straps row B is more critical for exactly the same reason. Fatigue crack initiation will occur at either (1) or (2) in figure 10.20. Cracks at (1) are invisible during inspection.

For 3 or more bolt rows in each plate it is easily possible to have an inhomogeneous load transfer distribution between the rows. An example is shown in figure 10.21a. A simple elastic calculation will show that $P_A = P_C$, and $P_B = 0$. The bolts in row B do not take any part in the load transfer. Under static loading until failure the bolts in row B will contribute because plastic deformation of the holes will lead to a more homogeneous load distribution between the three rows. Also in the elastic case, which is much more relevant to fatigue, a better distribution can be obtained, e.g. by a staggered thickness variation of the strap plates, see figure 10.21b. Note that for $P_A = P_B = P_C$ it is again true that the
holes in the plate of row A will be the more critical ones. This is the so-called end row effect. The end row is more critical because of the maximum by-pass load. This effect can only be alleviated by reducing the load transfer of the end row. A somewhat exceptional example of an end row is shown in figure 10.22. It is a stiffener terminating on a skin panel, where the end row consists of one rivet only. Alleviation of the end row effect is obtained by tapering the stiffener. In spite of this fatigue cracks still should be expected to generate from the end rivet (they were observed in full-scale fatigue tests on two modern transport aircraft).

Clamping

Bolts can be clamped in a controlled way, while rivets will always give some clamping, depending on the type of rivet. Clamping changes the above picture, which is already true for the simple lug joint, see figure 10.23. Load transfer from one element to another one now also occurs partly by frictional forces. Moreover, the coefficient of friction, which statically may be in the order of 0.3 for a fairly dry assembling, can increase considerably under cyclic load as a result of fretting. In figure 10.23 if there is no clamping fatigue cracks will be nucleated at A. However, after sufficient clamping fretting movements inside the hole are suppressed. Crack nucleation shifts to locations B or B'. Clamping will increase the fatigue strength until this shift of nucleation location has occurred. More clamping will not give a further increase. Even in the optimal configuration (crack nucleation outside the hole) fatigue strength reduction factors can be pretty high due to local fretting at highly stressed locations. In aircraft structure modern high quality blind rivets and other fasteners can produce a considerable amount of clamping, in many cases sufficient to have crack initiation outside the hole, see figure 10.24.

10.5 ASYMMETRIC JOINTS WITH ROWS OF BOLTS OR RIVETS

Some examples of asymmetric riveted sheet metal joints are shown in figure 10.25. There are some reasons why asymmetric joints are in an unfavourable position in comparison to symmetric joints:
(a) The bolt or rivet is loaded in shear in a single cross section only. Double-shear joints are statically stronger.

(b) For a single-shear fastener there is only one contact surface between the two plates or sheets for load transfer by frictional forces. More contact surfaces are favourable.

(c) The fastener is asymmetrically loaded and it thus will try to rotate. This gives an inhomogeneous bearing pressure along the hole, especially if hole filling is poor.

(d) As a result of the asymmetric configuration a tensile load will induce a bending moment. This is most evident for a single row lap joint, see figure 10.26, but it also occurs in other asymmetric joints. This type of bending is referred to as secondary bending, because it is not an applied bending moment, but rather a by-product of a tensile load.

For the displacement $w$ of the neutral axis, see figure 10.27, the differential equation can easily be written up. The solution of the equation indicates that $w$ is a non-linear function of the load $P$. Secondary bending is a non-linear problem. For the simple case in figure 10.26 it is easily seen that the bending moment at the rivet row is equal to $Pt/2$ which leads to:

$$\frac{\sigma_{\text{bending}}}{\sigma_{\text{tension}}} = 3 \quad \text{(independent of P)}$$

This is a very large bending stress. For other types of joints the solution of the differential equation has shown that high bending stresses are found also, see figure 10.28. The maximum bending stresses unfortunately occur at the rivet rows. For the lap joint a larger overlap is apparently beneficial, which is confirmed by the test data in figure 10.29. Note in this figure that the best results are found for the symmetric double strap joint. The best but one is the single strap joint, for which the very thick strap largely suppressed bending deformations. Unfortunately this solution cannot be applied if more than one row of rivets is required in each plate. A very thick strap would then lead to a highly unfavourable load transfer distribution between different rows. For instance in the lower left configuration of figure 10.25 practically all load transmission occurs through row A with a negligible contribution of row B. The stiffness of the support M between rows A and B is too high.
A significant cause for secondary bending is the occurrence of excentricities, i.e. small shifts of the neutral axis. A trivial example can indicate how easily large bending stresses are obtained.

\[
\begin{align*}
\sigma_{\text{tension}} &= \frac{P}{bt} \\
\sigma_{\text{bending}} &= \frac{P_1}{bt^2/6}
\end{align*}
\]

Assume \( t = 2 \text{ mm} \) and \( e = 0.2 \text{ mm} \). According to eq. (10.12) there is a 60% extra bending stress. The high bending factors in figure 10.28 should be regarded as a serious indication that excentricities must always be kept as small as possible. The designer has to be careful in this respect, not only for simple joints, but also in more complex joints, such as between the tension skins of a central wing and an outer wing.

10.6 REVETED JOINTS

The number of variables of riveted joints is large, even if we consider only lap joints and simple strap joints. Wellknown variables are:

1. rivet diameter (d), rivet pitch (s), distance between rivet rows (m)
2. sheet thickness t
3. number of rivet rows, and pattern of rivet holes
4. type of rivet
5. method of riveting.

It is impossible to cover all this variables in a short discussion here. Many data are to be found in the literature with compilations in the ESDU Data Sheets (Ref. 18) and reports of the aircraft industries. Some general comments will be made here.

Larger rivets for the same pitch (increasing d/s) is generally improving the fatigue strength, but there are limitations coming from static strength requirements and riveting procedures. Values of d/s in the range
1/4 to 1/6 are fairly common. A real size effect, i.e. changing all
dimensions d, s, m and t in the same proportion seems to have no
systematic effect on the fatigue strength.
Two rivet rows are significantly better than one as shown before. More
than 2 rows gives minor improvements for fatigue. Increasing the overlap
of a single lap joint is favourable.
With respect to the rivet hole pattern a non-staggered pattern is
preferred if there are two rows in each sheet. If there are more than
2 rows it is not clever to have a lower number of rivets in the "end rows"
because this is already the more critical row. It would increase the hole
loading in the end row.

There are many types of rivets. Wellknown categories are countersunk and
non-countersunk rivets. The latter should be preferred if there are no
aerodynamic problems involved. Modern blind rivets can be quite good and
they are more expensive. Even at this time new types of rivets for
general purposes are being developed. The quality of the rivet should be
considered together with the riveting procedures:
- hammer riveting
- rivet squeezing
- single-blow riveting.
The first method is the oldest. A longer riveting time improves the
filling of the holes, with a favourable effect on fatigue. The second
method is normal for automatic and semi-automatic riveting. Filling of
the hole is less good. The single-blow riveting method is rather new.
One of the advantages as compared to hammer riveting is the low noise
level in a factory.

Finally it should be said that the fatigue strength of a riveted joint is
not high. This is not surprising if we realize that the stress concentration
factors are high and that fretting corrosion will occur between the sheets.
Cyclic stress levels for important riveted joints should be a matter of
serious concern. In figure 10.30 fatigue diagrams of lap joint (not an
optimal design) show that the fatigue limit can be low with a small effect
of $\sigma_m$. The estimation of the fatigue properties of a riveted joint of
arbitrary dimensions is a problem, for which calculation rules can hardly be given. Consultation of available data for similar joints is advisable.

**Spot-welded joints**

A spot-welded joint is not a riveted joint. It will not be discussed here in any detail, also because it is rarely applied in aircraft structures. If it is applied it is for non load carrying joints. Fatigue tests on spot-welded lap joints indicate poor results. There is no clamping, a very sharp notch at the spot-weld and a poor material structure as a result of the welding operation.

### 10.7 ADHESIVE-BONDED JOINTS

Adhesive bonding of metal sheets and stiffeners has been applied in the aircraft industries for many years. So far tensile loads on a adhesive bondline are avoided, but the static strength and the fatigue strength in shear is satisfactory. Some fatigue aspects will be mentioned briefly.

**Failure modes**

In figure 10.31 a cross section of a lap joint is shown with two failure modes. Under high cyclic loads it is possible that a bondline failure will occur. This failure then has a progressive character, it will grow as a function of the number of load cycles. The bondline failure can occur in the bondline (cohesion failure) or between the bondline and the metal (adhesion failure). This will depend on the quality of the pre-treatment of the metal surfaces, the adhesive and the curing cycle. The quality of the adhesives and the bonding techniques has been much improved and as a result the sheet metal failure is the predominant fatigue failure mode.

**Notch effect at the end of an overlap**

The sheet metal failure is the result of stress concentration at the end of an overlap. Apparently there is an abrupt change of thickness at the end of the overlap, which should give a most severe concentration. However, there is a bondline with a very low elastic modulus (order of magnitude 3000 N/mm² as compared to 72000 N/mm² for Al-alloys). As a
result the notch effect is low, even in spite of the small bondline thickness (order of magnitude 0.2 mm). The fatigue strength reduction factor of the symmetric specimen in figure 10.32a (thickness increase from 2 to 4 mm, ignoring bondline thickness) is no more than 1.25. For the asymmetric case in figure 10.32b the change in thickness is lower, but $K_I$ is higher. Reinforcing with bonded strip at one side only introduces excentricities with secondary bending, and that explains the larger fatigue strength reduction.

An adhesive-bonded joint is a continuous joint

In a riveted joint load transfer occurs in a number of concentrated areas, i.e. the rivets. These concentrations are avoided in an adhesive-bonded joint, which has a continuous character. There is another advantage: fretting corrosion in the joint is impossible because the bondline prevents metal contact. (Moreover on adhesive bondline is air tight and fuel tight.) Results of comparative fatigue tests on riveted and adhesive bonded joints are shown in figure 10.33. The tests were made in view of application to a pressurized fuselage. The bonded joint was clearly superior in this comparison. Tests were also carried out on bonded specimens with rivets, but that hardly changed the fatigue curve. That is a logical result if metal failures occur at the end of the overlap.

10.8 SOME CONCLUDING REMARKS

In the present chapter the complexity of the fatigue behaviour of various types of joints has been discussed. The prediction of the fatigue life or the fatigue strength is a difficult problem. In most cases the best method is to start from fatigue data of comparable joints. On the other hand the discussion has shown that the effects of various aspects of joint design can be understood in a qualitative way. This understanding is of great importance to the designer in order to develop good joint design for the aircraft structure. Avoidance of excentricities is of utmost importance. The number of variables is large and there are several means to improve the fatigue strength of joints. In cases of insufficient information fatigue tests can be necessary.
REFERENCES


12. See Ref. 8.


Chapter 11  
FATIGUE LOADS ON AIRCRAFT STRUCTURES

Contents:
11.1 Introduction
11.2 The load history
11.3 Counting methods
11.4 Deterministic and stochastic loads
11.5 Mission analysis
11.6 The purpose of mission analysis
11.7 Some concluding remarks
References

11.1 INTRODUCTION

Information on fatigue loads is required at different times of the aircraft development and life. A schematic indication is given in figure 11.1, which is an extension of the first figure of chapter 1. Already in the aircraft design phase predictions have to be made on the fatigue loads which will occur when the aircraft is in service. Obviously speculation will be involved, i.e. predictions on how the operator will use the aircraft. However, it is the responsibility of the design team to minimize the speculative aspects as much as possible by analysing the various missions for which the aircraft will be used. This requires a so-called "mission analysis". If missions have been defined the next question is to indicate the various loads which will occur during a mission. Numbers and magnitudes of these loads are then required to build up a full spectrum of all cyclic loads applied to the aircraft. Before these aspects can be discussed, some definitions on fatigue loads in service have to be considered first.

11.2 THE LOAD HISTORY

If a load \( P \) is varying as a function of time \( t \) the function \( P(t) \) is indicated as the load history, or load-time history. Two samples of a load-time history are shown in figure 11.2. They look quite different, which will be discussed later. The question to be raised here is: in which detail and how should such load histories be described for application to fatigue calculations or testing? Such questions cannot be answered without considering the process of fatigue damage accumulation.
From the discussions in the previous chapters it should be understood that the characteristic points of \( P(t) \) are the maxima and the minima. \( P_{\max} \) and \( P_{\min} \) values are the load levels that matter for growth of micro-cracks and macrocracks. In other words: what we should know about \( P(t) \) is a series of successive maxima and minima:

\[
P(t) \neq P_{min} \ P_{max} \ P_{min} \ P_{max} \ P_{min} \ldots \ P_{min} \ P_{max} \ldots
dotdotdot
eq P_{min} \ P_{max} \ldots
\]

Because this would lead to extremely long series of numbers there are a few obvious questions:

1. Should very small load variations be considered also? Ignoring small variations would reduce the amount of data.
2. Is it allowed to replace the sequence of maxima and minima by a statistical representation? The sequence of maxima and minima could then be replaced by a statistical distribution function. That would lead to a drastic reduction of the data. At the same time it should be expected that the original sequence of the successive maxima and minima will be lost in such a statistical representation. It was pointed out in chapter 9 on fatigue under variable-amplitude loading, that sequence effects can have a significant influence on fatigue damage accumulation. That implies that it is possible that statistical representations will lack essential information on sequence.

Both questions have to be kept in mind when we now discuss different "counting methods" to arrive at statistical data on fatigue load histories.

11.3 COUNTING METHODS

A sequence of minima and maxima is shown in figure 11.3a. The range of possible \( P \)-values is subdivided in a number of intervals with \( P \)-levels as boundaries between the intervals. It is a simple procedure to count the number of \( P_{\max} \) values and \( P_{\min} \) values in each interval, which then can be plotted as a histogram, see figure 11.3b. Histograms are associated with statistical density function. Instead of the number of peaks in intervals we can also count the number of \( P_{\max} \) values above level 1 and the number of \( P_{\min} \) values below level 1. This is also done in figure
11.3a. The number of exceedings is plotted in figure 11.3c. These curves are associated with statistical probability functions.

It should be noted that figures 11.3b and c give full information about how many peaks occurred in figure 11.3a. However, information on the sequence in which they occurred is lost. Because fatigue is a matter of load variations, it might be more instructive to know how many times load variations from one minimum value to a subsequent maximum occurred. For this purpose the number of positive load ranged \( \Delta P = P_{\text{max}} - P_{\text{min}} \) was counted in figure 11.3a, see the series of values on top of figure 11.4 (for simplicity peak values were supposed to be in the middle of the intervals in order to get integer values for \( \Delta P \), but this is not essential for the present discussion). The numbers of \( \Delta P \)-values can then be presented in a histogram (see fig. 11.4a) or in an exceedance curve (see fig. 11.4b). Although each \( \Delta P \) value is the difference between two successive peak values, information about the sequence of \( \Delta P \)-values is also lost in this case.

The counting results discussed so far are one-dimensional counts. Each load occurrence is characterized by a single value, i.e. a peak value in figure 11.3 or a \( \Delta P \) value in figures 11.4a and b. However, a load variation is characterized by two values, e.g. \( P_{\text{min}} \) and \( P_{\text{max}} \), or \( \Delta P \) and \( P_{\text{mean}} \). In a two-dimensional count system this extra information can be given, but instead of a histogram or an exceedance curve, one extra dimension is required for data presentation. This can be done in a matrix, see figure 11.4c. Each number in the matrix indicates how many times a load variation occurred between the corresponding \( P_{\text{min}} \) and \( P_{\text{max}} \) values. Obviously such a 2d-presentation is more informative than a 1d-presentation. However, again information about the sequence is lost (Note: in Fig. 11.4c only the positive \( \Delta P \) values have been counted. Of course the same procedure can be applied to the negative \( \Delta P \) values).

In figure 11.3a all \( P_{\text{max}} \) values were above a reference level \( (P_{\text{ref}}) \) and all \( P_{\text{min}} \) values were below that level. This appears to be applicable to the \( P(t) \) sample in figure 11.2a, but not to the sample in figure 11.2b. The peak counts and range counts can still be made in the same way as
illustrated by figures 11.3 and 11.4., but it is not evident how this will affect the meaning of the counting information obtained. Should all small load variations lead to meaningful counts? This question will be discussed by defining some counting methods which are used in practice.

Mean-crossing peak-count method

In the small sample shown in figure 11.5 four small load variations occur (3-4, 6-7, 9-10 and 12-13). If all peak values are counted (fig. 11.5a) the maxima 4 and 6 and the minima 9 and 13 may be associated later with equal amplitudes relative to the mean load $P_m$. Combining maximum 4 with minimum 9 (and 6 with 13) would lead to a much more severe load cycle than the small variations in the sample. This does not occur in the mean-crossing peak-count method, defined in figure 11.5b. Between each pair of successive mean crossings only one peak value is counted, which is the most extreme one. This method filters out the small variations. In the USA many load records (vgh-records) were analysed with this counting method.

Level-crossing count methods (Fatigue meter)

Figure 11.6a shows the principle of the level-crossing count method. Each time that the load passes a positive level ($P > P_m$) in the upward direction a count is made. If level $i$ is passed $m_i$ times and the next higher level $i+1$ is passed $m_{i+1}$ it might be expected that the number of maxima between the two levels is equal to $m_i - m_{i+1}$. This is not fully correct. In figure 11.6a it is true for levels 2 and 3 (there are 2 maxima equal to $m_2 = m_3 = 2$, but not for the levels 1 and 2 (there are 2 maxima, whereas $m_1 - m_2 = 1$. The discrepancy is a consequence of a small variation with a maximum and a minimum between the two levels. A more significant nuisance can occur if many small load variations occur around a counting level. One example is AB in figure 11.6a. This problem has been solved in the counting accelerometer by introducing a second counting requirement. A crossing of level $i$ is counted only if the load has return to a significantly lower level $i'$, see figure 11.6b. Also this method removes small load variations from the counting result. The method was developed in England in the
counts to be made. The residue should be counted by a simple peak count. Algorithms have been developed, (e.g. Ref. 1) which in a single run through the peak values will count all range pairs, also in a two-dimensional matrix form as discussed before (figure 11.4c).

In the past most data on local statistics came from the mean-crossing peak-count method and the fatigue-meter count method. In the future more data of the range pair count method may be expected. Although there is no strict proof to state that the latter method is superior, it is plausible that it gives more credit to the fatigue damage accumulation process. In the first two methods information about sequences is lost. The range pair method retains some information. From the definition given before (B, C between A, D for counting) it is clear that each range pair count can have had smaller ranges in between, but never larger ranges. Moreover, a 2d-count method (which will give more information than a 1d-method) is meaningful only if ranges are considered.

11.4 DETERMINISTIC AND STOCHASTIC LOADS

The discussion in the previous section tacitly assumed that the load history was characterized by a kind of a reference load. For gust loads and manoeuvre loads on a wing this should be the 1g stress level in flight. However, the 1g load on a wing can be highly different in flight and on the ground. For a further discussion a few more definitions on types of loads are necessary. The following types can be recognized:
(a) deterministic loads
(b) stochastic loads
   (b1) stationary stochastic loads
   (b2) unstationary stochastic loads.

Deterministic loads

Cyclic loads are deterministic if each load can be considered to be a separate occurrence, while the magnitude of the load is known before. The magnitude is not a matter of probabilities. A pressure vessel, which is always pressurized to the same working pressure is loaded by a
deterministic load. The cyclic load is fully planned and predictable. On the contrary, a lift in a building, which will carry highly variable numbers of people is not loaded in a deterministic way. For an aircraft several loads have a deterministic character, because they are routine procedures. Taking off and landing loads and the pressurization of a fuselage are obvious examples. In general manoeuvre loads for a commercial transport should be deterministic. For a training aircraft, at fighter aircraft or an aerobatic aircraft all manoeuvres are not predictable, and these loads have a stochastic nature.

**Stochastic loads**

For a stochastic load the magnitude and the occurrence can only be indicated in a probabilistic way. Stochastic variables per definition are to be described by probability functions. There is a certain probability (chance) that certain things will happen. In the previous example of a lift in a building the number of people in the lift cannot be predicted exactly, but it can be indicated in a statistical way by a distribution function. The load on the lift is a stochastic variable. We might hope that the distribution function of a stochastic load can be derived from an analysis about how the structure is supposed to be used. If this is not possible the distribution function should be obtained by long-term measurements.

A stochastic load with a distribution function that always remains the same, is a stationary stochastic variable. If the statistical properties do not remain the same it is an unstationary stochastic load. Manoeuvre loads on a fighter aircraft can be supposed to be a stationary stochastic variable, assuming that the aircraft is always flown in a more or less similar way. Gust loads obviously have a stochastic nature. The gust severity is high during bad weather and low during good weather. The statistics of gusts are depending on the type of weather. They also depend on the flying altitude. There are much less gust loads at a cruising altitude of 10 km than at a low altitude.
Some types of stochastic loads are caused by a stochastic random process. Turbulent air, in which the aircraft is flying, is supposed to be such a process. The same applies to random noise of a jet engine. For these random processes it is assumed that they are a stochastic Gauss-process (or random Gauss-process), which implies that the relevant variables have a normal distribution function (i.e. Gaussian distribution). For random loads, which are caused by a stochastic Gauss-process interesting calculations can be made on the relation between the random loads applied to a structure and the random loads which then will occur in the structure. This is outside the scope of these lectures (see e.g. Ref. 2). However, some comments will be made on the description of a Gaussian random load by the power spectral density function (PSD).

Random load

For a Gaussian random load the power spectral density function, \( \Phi(\omega) \), fully describes its statistical properties. It should be considered as a density distribution function with the frequency (\( \omega \)) as a variable, see figure 11.10. To evaluate this concept a bit further, we will consider a Fourier series with a very large number of terms, with very small differences (\( \Delta \omega \)) between the frequencies of successive terms, and coefficients \( A \) being a function of \( \omega \), see figure 11.11. In an oscillation the energy is proportional to the square of the amplitude, which implies:

\[
\Phi(\omega) \propto [A(\omega)]^2
\]

The sum of such a Fourier series is approximately similar to random noise. It becomes real random noise if \( \Delta \omega \to 0 \). Some examples of random noise and the corresponding power spectral density function are shown in figure 11.12. In figure 11.12a the energy is concentrated in a narrow frequency band and as a result the load-time history is somewhat similar to an amplitude modulated signal, in this case with a random modulation. This narrow band random loading is typical for resonance systems, which predominantly respond at one single resonance frequency if activated by some external random process over a broader frequency range (see also figure 11.2a).
In figure 11.12c the spectral density function of the random signal covers a much wider frequency band and the corresponding broad band random loading shows a higher degree of irregularity (see also figure 11.2b).

Because \( \Phi(\omega) \) fully characterizes random load some characteristic properties can be derived from \( \Phi(\omega) \) (Ref. 2), which are the distribution function of the peak values, the number of mean-crossing per second \( (N_0) \), the number of peak values per second \( (N_1) \) and the irregularity factor \( k = N_1/N_0 \).

For narrow band random load \( k \ll 1 \), whereas for broad band random load \( k > 1 \) (see figure 11.12). Several formulas are summarized in Ref. 4.

11.5 MISSION ANALYSIS

Mission analysis will be discussed here by dealing with the example of a civil transport aircraft. From the design goals of the aircraft some "typical" flights should follow. These flights are typical for certain missions for which the aircraft can be used. Obvious differences between different typical flights will be related to the flight distance and the amount of payload and fuel carried during the flight. Let us assume that three typical flights can be specified, e.g. a short flight, a long flight and an intermediate one. For each type of flight a flight profile has to be developed. An example for a short flight is shown in figure 11.13.

The flight time is split up in a number of time intervals. Each interval in figure 11.13 is characterized by speed, altitude and flap position. The quasi-static load on the aircraft can then be calculated for each interval. A schematic example for the wing bending moment is given in figure 11.14a. These loads are purely deterministic. The bending moment flight profile is one large quasi-static load cycle, also referred to as the ground-air-ground cycle (GAG) or the ground-to-air cycle (GTAC). During the flight the bending moment can change as a result of fuel consumption.

There are several other loads to be considered for each typical flight. Figure 11.15 indicates some categories of loads with characteristic periods for one load cycle and an indication on possible numbers of load
fifties. The acceleration (Δn) in the center of gravity of the aircraft was measured and automatically counted by this method. The acceleration gives a quantitative indication of the load on the aircraft. The relation between Δn and the load on the aircraft depends on the aircraft mass distribution and on its dynamic response, which can be approximated by calculations. Because the development was made in view of fatigue life problems the counting accelerometer was also called "Fatigue meter", but this is a somewhat misleading term. Apparatus are commercially available. They have a number of counting levels for both positive and negative loads (relative to the mean load, n = 1). Useful information can easily be obtained. An example is shown in figure 11.7 for an F-28. It shows the counting accelerometer data after a long period of service experience. The data are normalized as numbers per flight. Apparently one operator (frequently crossing high mountains) is very close to the original design analysis data. If the aircraft is used as an executive aircraft predominantly flying under good weather conditions, the load spectrum is much less severe.

**Range-pair count method (rain flow count method)**

In figure 11.8 we consider the decomposition of a load variation ABCD in load ranges. Two possibilities are indicated. With the first method the three successive load ranges are counted separately. This method does not recognize the fact that ignoring the interruption BC will leave a large range AD. Thinking in terms of fatigue damage accumulation it should be expected that the range AD should be counted anyhow if BC is small, and perhaps also if BC is not so small. The preference should therefore go to method II in figure 11.8. This is the basis of the range-pair count method, originally proposed in England, but better known from a publication in Japan under the name "rain flow method". The method is illustrated by figure 11.9. Consider four successive peak values, such as A, B, C, D in figure 11.8. A count is made only if peaks B and C are in between peaks A and D. The load variation BCB' is then counted as range pair (BC and CB') and B and C are removed from the series of successive peak values. In figure 11.9a four counts are made and in figure 11.9b another count is possible. The residue left in figure 11.9c does not allow any further
cycles during a life time of the aircraft. For the wing bending gusts, manoeuvres and landing loads including taxiing can be relevant. For a civil transport manoeuvre loads are considered to be low as compared to gust loads. In figure 11.14b it is indicated schematically how gusts, touch down and taxi loads will change the picture of figure 11.14a. It should be recalled that gusts are a non-stationary type of loads. So gust loads will be different from flight to flight. There is statistical information about average gust spectra (including the non-stationarity) as a function of flying altitude (see Ref. 7). For each interval of the flight in figure 11.13 the average gust spectrum can be derived. The spectra can be added to cover all intervals of the flight.

It is of practical interest to know that such an overall gust spectrum, normalized to numbers of gusts per flight, is not very much different for a short flight and a long flight. The explanation is rather simple. Most gusts occur at low altitudes, i.e. during climb and descent. During cruise at a high altitude the weather is very calm, usually without significant gusts. As a consequence cruising time hardly affects the number of gusts per flight. As a result the fatigue damage is much the same for a short flight and for a long flight. Fatigue life should thus not be expressed in flying hours, but in numbers of flight!

The touch down load will not be the same in all flights. It depends on the sinking speed of the aircraft, the dynamic response of aircraft and landing gear and the type of air strip. There are statistics about sinking speeds, but clearly enough this should be dependent on landing procedures, pilots and runway qualities. These aspects have to be analysed as part of the mission analysis. The same applies to taxiing loads, which are highly depending on runway roughness. Different types of airfields have to be considered. Anyhow the flight profile of the bending moment is going to be a mixture of deterministic ground-air-ground cycles and stochastic loads both in flight and during taxiing.

The above discussion was related to the wing of a civil transport aircraft. Mission analysis can also be done for the fuselage, the tailplanes and the undercarriage. It will turn out that flight profiles will be quite
different. For the fuselage the pressurization cycle is a fairly dominant load cycle. In addition fuselage bending by gusts, manoeuvres and landing can also be significant. This should follow from calculations. For the horizontal tail plane deterministic manoeuvre loads during landing can be important. For a landing gear the analysis can be complex and not only because the sinking speed is a stochastic load. The landing gear will be loaded in three directions. The vertical loading direction is an evident one. Horizontal loads due to spin up and braking are also fairly evident. However, sideways loads can occur also, while ground manoeuvring, including towing should not be overlooked. Several of these loads do not occur together, and if they occur together they still need not be fully synchronized. Only careful mission analysis will reveal relevant load-time histories for the various components of an aircraft.

11.6 THE PURPOSE OF MISSION ANALYSIS

Mission analysis can be done for different categories of purposes:
(1) to obtain general design data and to indicate problems, for which insufficient data are available
(2) to arrive at load spectra for fatigue calculation
(3) to compile information on load-time histories for fatigue tests (e.g. full-scale fatigue tests).

Mission analysis should not be made for fatigue purposes only. It has also to be done to reveal extreme loading cases, which must be considered for static design calculations and tests.

Before comments are made on fatigue calculations we will first discuss the third category: fatigue tests based on load-time histories obtained after mission analysis. The results of fatigue life calculations will always be influenced by several weak points in the calculation methods (e.g. assumptions to be made). The need of experimental verification will be recognized at various stages of the aircraft design, ultimately at the very end by full-scale fatigue test on the complete aircraft structure or on large componentes of the structure, such as the wing, the fuselage, the empenage. Full-scale tests are expensive and for that reason such a test should give relevant and reliable information. The information
expected from a full-scale test is discussed in chapter 12. Here it may be said that good information can only be achieved if the load history to be applied in the test is a realistic simulation of load histories to be expected in service. Such a test load history can be composed if a mission analysis has been made before.

Flight-simulation tests

Returning to figure 11.14 it will be clear that an exact simulation of the load-time history in service with respect to time would take years. Fortunately time compression appears to be allowed. This is supported by the discussion in chapter 7 (section 7.5.2) on corrosion fatigue. Fatigue cracks do only grow if the load is varying. During cruise the load is constant and this part of the flight can be dropped from the flight simulation. It was also shown in chapter 7 that frequency effects for Al-alloys are negligible in humid air, which implies that the gust cycles may be simulated at a faster loading rate. Comparative crack growth tests with a flight-simulation loading carried out at 10 Hz, 1 Hz and 0.1 Hz have confirmed this (Ref. 8). A further simplification which appears to be justified is to drop the small taxi load cycles. They occur at a low mean stress and this will be hardly effective in contributing fatigue damage. In this way the flight profile of figure 10.14b is reduced to the simulated profile in figure 11.14c. In a full-scale test such a flight can be simulated in just a few minutes, and a test until a sufficient aircraft life time can be completed in less than one year. A similar simulation of a flight on a specimen in a modern fatigue machine will be a matter of some 10 seconds.

It was pointed out before that the gust environment can be different from flight to flight, depending on the weather conditions. This has to be simulated in a realistic test. Figure 11.16a shows a sample of a load history, similar to the flight-simulation applied in the full-scale fatigue test on the F-28 wing. It illustrates that different types of flights were applied. Actually 10 different types were simulated. The most severe one (type A: heavy storm) occurred only once in 5000 flights, the two least severe flights (J and K: nice weather) occurred about 3700 times in 5000 flights. The different types of flights were applied in a random
sequence. The overall gust spectrum was in agreement with the required spectrum. The spectrum is shown in figure 11.17. In the flight-simulation it was approximated by a stepped function for experimental convenience but this is not essential.

**Truncation of high loads**

One delicate aspect of the flight-simulation test is the maximum load to be applied in the test. According to the gust spectrum (see figure 11.17) there are rarely occurring very high loads. The discussion in chapter 9.3 has shown that such loads can be beneficial for the fatigue life. We now should recall that gusts are a stochastic load. The load spectrum in fact is a statistical distribution function. That implies that we are not sure that all aircraft in service will meet the rarely occurring high loads. Some aircraft will not and others will see more of them. If we then apply these extremely high loads it may flatter the results of the test for those aircraft, that will not meet the high loads. For that reason it is usual to set a maximum limit to the amplitudes to be applied, which frequently is done at the level to be anticipated 10 times in the life time of the aircraft. This is a somewhat arbitrary choice, but it is believed to be a level that will be reached by the large majority of the aircraft fleet. In figure 11.17, the anticipated life being 50000 flights, the truncation level is thus found at the amplitude occurring once in 5000 flights. Gust loads with a still lower probability of occurrence are reduced to this truncation level.

Some examples of the effect of the truncation level on fatigue life and crack propagation are shown in figures 11.18 and 11.19. The results in figure 11.18 are not fully systematic, but the predominant trend is a longer life for a higher truncation. For macrocrack growth the effect can be even more dramatic as shown in figure 11.19.

For a manoeuvre spectrum the truncation of high loads usually is less critical because the spectra are more flat for the rarely occurring high loads, see figure 11.20. Nevertheless we always have to be careful with flatter test results by extremely high loads. Sometimes high fail-safe loads are applied during a full-scale test in order to prove that the structure is still capable of carrying high loads. Unfortunately the
the growth of small cracks in the structure, not yet found so far, can be effectively arrested. The indication of critical locations can thus be prevented, which violates one of the purposes of the test (see chapter 12).

Mission analysis for fatigue life calculations

Fatigue life calculations for an aircraft structure are difficult for several reasons, to be recapitulated in chapter 12. Here one problem has to be added, which is the load cycle definition. In figure 11.14 we have seen that an actual flight load profile is a superposition of deterministic and stochastic loads. Prediction methods for fatigue life and crack growth require load spectra to be defined in cycles, characterized by a minimum and a maximum load. The question of cycle definition was discussed in the present chapter as part of counting methods. There was a preference for the range-pair count method. It is interesting to see what this counting method will do if applied to the flight-simulation load sequence in figure 11.16. After counting the smaller cycles it will combine with the maximum stress of a flight to form the most severe cycle counted in that flight. Actually this seems not unreasonable. The picture of the real flight profile thus can help to arrive at a reasonable break down to cycles if separate load spectra for the deterministic and stochastic loads are available. Fully rational methods for this purpose do not exist, but with some understanding of the real load-time history and fatigue damage accumulation reasonable choices can be made.

11.7 SOME CONCLUDING REMARKS

In the previous section mission analysis and associated problems were discussed as part of the work to be done during the design phase of an aircraft. This work is completed by full-scale fatigue tests. It has to be recognized that the load spectra adopted are based on predictions of the anticipated aircraft utilization in the future. Typical flights and missions will have been analysed and usually a severe, but realistic utilization of the aircraft will be adopted for life calculations and fatigue tests. It is still quite well possible that the aircraft in service
will be used in an essentially different way as expected. How should we
know that, and who should care about it? From figure 11.7 it is clear
that different operators are using the same aircraft in very different
ways. In that figure the deviations of the assumed spectrum appears to be
on the safe side. This need not always be so. Two examples are:
(1) Civil transport aircraft used for pilot training are subjected to a
much more severe load spectrum than the aircraft used for normal air line
services. Significant manoeuvre loads are added. A similar difference
appears to military aircraft as well.
(2) Aircraft used for aerobatic exhibitions also have a more severe
load spectrum.
Although such difference are rather obvious it is not easy to quantify
the differences unless measurements are made. Flight load monitoring
is the concept of measuring loads in flight, either for a whole fleet,
or for a number of representative aircraft. It can be done with counting
accelerometers or by other strain gage based measuring techniques. If no
measurements are made on what is supposed to be normal service we have no
information. If measurements are made there is a data basis for further
analysis and for action if necessary.

REFERENCES

2. J.S. Bendat: Principles and applications of random noise theory.
3. J. Kowalewski: On the relation between fatigue lives under random
   loading and under corresponding program loading. Full-Scale Fatigue
4. J. Schijve: Four lectures on fatigue crack growth. Delft Un. of
5. E.J. van Beek: Fatigue testing of the F-28 "Fellowship".

7. ESDU Data Sheets, Engineering Sciences Data Unit, London.


Chapter 12  
FATIGUE OF AIRCRAFT STRUCTURES

Contents:
12.1 Introduction
12.2 Design aspects
12.3 Predictions and calculations
12.4 Fatigue testing methods
12.5 Full-scale tests and scatter factors
12.6 Maintenance and inspection
References
Table 12.1

12.1 INTRODUCTION

In the introduction of chapter 7 several arguments were mentioned why an aircraft designer should give serious attention to fatigue aspects of the aircraft structure. The previous chapters 7 to 11 have shown that a number of problematic aspects are involved. As an illustration three such aspects are mentioned below:

- fatigue damage accumulation under variable-amplitude loading is a complex problem.
- the fatigue behaviour of riveted and bolted joints is affected by clamping forces, fretting corrosion and secondary bending.
- service load histories can be combinations of deterministic and stochastic loads which makes a fatigue load cycle definition a difficult question.

Fortunately a qualitative understanding of these problems has reasonably well developed, but accurate quantitative answers in many cases cannot be given. As a result it was said several times that accurate quantitative indications will usually require fatigue tests or measurements, especially focussed on the particularly questions of interest.

In order to see how and where the aircraft designer should spend his efforts to arrive at structures with good properties, a survey of the various phases involved is given in figure 12.1. This figure will be used as a guide line for the present chapter. Special attention will be paid to different methods of fatigue testing.

12.2 DESIGN ASPECTS

An aircraft structure is designed according to certain design specification with respect to dimensions, load carrying capacity, special structural items such as doors, windows etc. special connecting elements in view of fixing equipment and systems and also for assembling purposes. There is a large variety of alternative
structural design possibilities and the designer has to make various choices concerning the general lay out of the structure, the joints and detail design, material selection, alternative production methods. There is no single optimal solution, because there are many different criteria to be considered. Some of these criteria are related to fatigue, structural safety, (stress) corrosion, inspectability, ease of maintenance, but even for these criteria there is no single solution. Most choices have to be made based on experience and qualitative analysis. For this purpose a qualitative understanding of the material behaviour in an aircraft structure is of prime importance. Some simple examples of elementary comparisons between alternative possibilities are indicated below.

**Fail-safe or safe life?**

When the general lay out of a structure is designed the question whether the structure is going to be either "fail-safe" or "safe-life" has to be considered. From the fatigue point of view the fail-safe structure should be preferred. In several cases the choice is between a multiple-load path structure and a single load path structure.

An example of a multiple-load path structure is the short wing in figure 12.2a with four spars, each connected to fuselage frames. If fatigue cracks are nucleated in one of the spars the stiffness will be reduced very locally only. The other spars will hardly carry more load. Roughly 3/4 of the strength will be maintained. If we now consider a single-spar wing a crack in the spar can be disastrous. As a matter of fact in the past at least four different aircraft types have crashed due to fatigue cracks in a single spar. Nowadays larger aircraft have at least two spars, but the main reason is that the structure can then be lighter for the same static load bearing capacity. It also can improve the fatigue resistance considerably.

Single-load path structures are sometimes more simple in production (economic aspect), but in view of fatigue high safety factors are required, which implies a lower allowable design stress and a heavier structure.

Lug type attachments between structural parts are attractive for assembling and disassembling (compare to an integral connection). Fatigue failures in such an attachment can be dangerous. It has been proposed to
the fatigue life until there is an obvious macrocrack. In this respect
detail design and the design of joints is of great importance.
Especially with respect to joints the number of possible solutions can be
large and it requires a thorough analysis of alternative solutions to
arrive at a better one.

Material selection

For fatigue critical elements materials with good fatigue properties
should be selected. The classical fatigue critical elements are the
tension skin of a wing and the fuselage skin, but fatigue cracks have
occurred in many other components as well. It is a matter of the load
spectrum on a component to decide whether it might be fatigue critical.
A second criterion is the question whether the component is fail-safe
or not. Evident examples of components which usually are not fail-safe
are helicopter blades and aircraft propellers. For such components
great attention has to be paid to detail design, material surface
quality, etc. Moreover, stress levels are kept low. If such a component
is made of a material with good fatigue properties and if the stress
level is low it is possible that incidental cracks will grow
extremely slowly. Such a "slow growth component" with a good inspection
procedure can again be a safe item, even if cracks arise. The inspection
then should be reliable.

Before we have referred to materials with good fatigue properties.
This should cover:
- high fatigue limit and fatigue strength
- low fatigue notch sensitivity
- low fatigue crack propagation rates
- high fracture toughness ($K_{IC}$).

From the point of view of the operator it can be indicated as shown
in figure 12.3. Obviously he prefers a long crack initiation life and
subsequently a slow crack growth in order to have ample time for crack
detection during inspection. In this respect the $K_{IC}$ (or $K_C$) has some
meaning for the maximum allowable crack length $a_C$. A high fracture
toughness is desirable. Low notch sensitivity, low fatigue crack
growth rates and a high fracture toughness are three properties which
are better for a high ductility material and worse for a low ductility
material. High ductility materials have a relatively low $\sigma_{0.2}$ and low
ductility materials a high $\sigma_{0.2}$. For reasons of $\sigma_{0.2}$ and $\sigma_Y$ low ductility
make such a lug as a quasi multiple-load path element as shown in figure 12.2b. This is not really a good proposal. If a crack starts to grow in one of the three lugs the stiffness of that lug will be reduced considerably (a crack in a small element). Consequently that lug will not carry much load. The other two lugs have to carry more and fatigue cracks will then occur in those two lugs. The result is that the three lugs will crack more or less together, instead of one after the other one.

Crack stopping elements

In figure 12.2c a schematic picture of a wing structure with 4 planks between front spar and rear spar is presented. The planks are connected by span wise strip joints. The idea is that if one of the planks fails the crack will be contained to that plank. A crack running in the chord direction will not automatically pass the span wise joint, because there are no longer crack tips when one plank has failed. An aircraft once had a hidden fatigue crack in one plank, which became unstable under a gust load. The plank fully cracked, but the aircraft could make a safe landing. Without the span wise joints the wing would have failed.

Another type of crack stopping elements are the crack stopper bands applied to pressurized fuselages of large aircraft. The idea is sketched in figure 12.2d. Such extra bands are connected to the fuselage skin between the skin and the frames (sometimes between the frames instead). The predominant fatigue load is the pressurization cycle (once per flight) which can induce fatigue cracks perpendicular to the hoop stress. It was such a crack that became unstable in the "famous" Comet accidents, but unfortunately also in another aircraft type later. If a crack is approaching such a band, or even run partly under the band crack opening is significantly restraint (see the discussion on figures 8.26 and 8.27) which will stop the crack.

Detail design and joints

The design aspects discussed so far are illustrative examples to show that the designer already in the early design phase of the structure can make decisions towards a more fatigue fail-safe structure. Clearly enough he should not only think about large cracks but also consider
materials are superior. For fatigue high ductility materials have better properties. A classical choice of the Al-alloys is between the more ductile 2024-T3 ($\sigma_{0.2} \sim 350 \text{ N/mm}^2$) and the less ductile 7075-T6 ($\sigma_{0.2} \sim 480 \text{ N/mm}^2$). For fatigue critical components 2024-T3 should be preferred. Various types of civil transport aircraft have 2024-T3 in the tension skin of the wing (sometimes with 7075-T6 stiffeners) and 7075-T6 in the compression skin (buckling is the critical failure mode and high $\sigma_{0.2}$ is favourable). Also the skin of pressurized fuselages is frequently made of 2024-T3 sheet material.

Because a significant part of the weight of an aircraft structure is depending on fatigue considerations, there is a continuous pressure on the development of materials with superior fatigue properties. Modified alloy compositions and heat treatments have not really been successful so far. Promising results on fatigue crack growth resistance were obtained in fatigue tests on laminated sheet material. Figure 12.4 shows the crack growth life for through cracks and part through cracks both for monolithic material and laminated sheet material, obtained by adhesive bonding of a number of thin sheets. Especially for part through cracks the gain in life is spectacular, and many cracks in aircraft components start as part through cracks. If a crack is present in one sheet but not yet in the other ones the latter sheets will restrain crack opening, thus reduce the stress intensity factor and the growth rate. A restraint on crack opening was already shown to be profitable for a crack in a stiffened skin (see figures 8.26 and 8.27). Recently the fatigue properties of laminated sheet material were significantly improved by introducing strong aramide fibres in the bond lines (Ref. 3). The material is called ARALL (Aramid Reinforced Aluminium Laminate). The fibres are very effective in reducing crack opening, even for very small cracks. Another promising material is carbon fibre reinforced plastic (so-called advanced composites). This is an essential different material, which requires new production techniques and design philosophies. It is supposed to be fatigue insensitive.

Production

Alternative production techniques have to be considered by the designer. Several components can be made as a single piece, such as machined forgings or numerically controlled machining of thick plates. On the other hand such components can also be made by combining a number of
smaller elements and to adopt riveting or bonding techniques. Decisions can depend on the available production machines and cost effectivity. However, fatigue and corrosion arguments can also be involved. In figure 1.4 of the first chapter (page 1/5) three different stiffener-skin attachments are shown. For fatigue the adhesive bonded attachment is the best one, the riveted connection is also good, but the crack stopping properties of integrally machined design are poor. Fatigue cracks run directly from the skin into the stiffener. Also the fail-safe strength (static strength in the cracked condition) is inferior. Machining of forgings also introduces the risk of corrosion problems in view of the corrosion susceptibility of end grain structures (chapter 6.4). A designer has to be aware of the various advantages and disadvantages of production techniques, not only so for fatigue and corrosion, but also for production problems in general, including economic aspects.

12.3 PREDICTIONS AND CALCULATIONS

After the design work on a structure has been completed the first steps to a fatigue analysis can be made. As shown in figure 12.1 and discussed in chapter 11 mission analysis should lead to possible fatigue load spectra and flight load profiles. It is important to know whether the fatigue damage accumulation rate should be considered as damage per flight or damage per hour. Aircraft operators usually think in terms of flying hours. Orders of magnitude can be very different. A commercial transport aircraft is used about 3000 hours per year (there are 8760 hours in one year), and for a fighter it is no more than about 200 hours per year. Some transport aircraft make long flights (e.g. wide body aircraft), say an average of 4 hours per flight, resulting in less than 1000 flights per year. Smaller ones, such as the F-27, used on short distances may come to 5000 flights in a year.

The fatigue load environment as revealed by mission-analysis has to be translated to load spectra for the various components of the aircraft. A considerable amount of analysis and calculations will be involved. Special problems can arise if the dynamic response of the aircraft has to be taken into account. Finally calculations should lead to stress spectra for those structural elements, that might be fatigue critical. These stresses will still be nominal stresses in certain elements. The nominal stresses will give a first indication
about possible fatigue problems and the desirability of additional design actions.

If stress spectra have been obtained life predictions can be made. It was pointed out in chapter 9 that the Pålmgren-Miner rule was not very reliable, but it was also said that there hardly is an alternative way. If we realize the shortcomings of this rule we may expect rough indications on the order of magnitude of the fatigue life. One shortcoming was the fact that load cycles below the fatigue limit do not contribute to damage according to the Pålmgren-Miner rule. It was also indicated in chapter 9 that extensions of fatigue curve can be adopted. We will illustrate the calculation by a fictitious example. A fatigue curve which is supposed to be relevant to the example is shown in figure 12.5 with the two extensions discussed before (figure 9.6). Two load spectra (H and H') are shown in the same figure. The spectra are given as the number of exceedings in 5000 flying hours. The life calculation is carried out in table 12.1, by first splitting up the S_a-scale in intervals, followed by reading H values in the graph. From these values the number of cycles (n = ΔH) in each interval is obtained. Then the life N for CA-loading at S_a in the middle of the interval is read from the S-N curve. This n/N can be calculated. Summing this values lead to Σn/N in 5000 flights. The estimated life at Σn/N = 1 then is:

\[
\frac{1}{\Sigma n/N} \times 5000 \text{ flights}
\]

The fatigue lives thus calculated for the three S-N curves (A, B and C in figure 12.5) are recapitulated below, together with the results for the other load spectrum H'.

<table>
<thead>
<tr>
<th>S-N curve (fig. 12.5)</th>
<th>Calculated life (hours)</th>
<th>Relative life</th>
<th>Ratio H/H'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spectrum H</td>
<td>Spectrum H'</td>
<td>H</td>
</tr>
<tr>
<td>A</td>
<td>37000</td>
<td>66000</td>
<td>1(*)</td>
</tr>
<tr>
<td>B</td>
<td>29000</td>
<td>24000</td>
<td>0.78</td>
</tr>
<tr>
<td>C</td>
<td>23000</td>
<td>7000</td>
<td>0.62</td>
</tr>
</tbody>
</table>

(*) per definition taken equal to 1.

Some comments have to be made:

(1) For spectrum H the calculated life decreases somewhat when going
from S-N curve A to B and to C. For spectrum H' the effect of the choice of S-N curve is considerably larger. The reason is that spectrum H' has many cycles below the fatigue limit, whereas for spectrum H it is much less. Therefore the result for H' is very sensitive to our assumption about the extended S-N curve.

(2) Comparing the severity of the spectra H and H' it appears that with the S-N curve A spectrum H is more severe than H' (H/H' = 0.6) and with the S-N curves B and C it is less severe (H/H' = 1.2 and 3.3). Only one of the two possibilities can be correct and here we are again faced to the limitations of our assumptions on fatigue damage accumulation.

(3) The reliability of such a calculation never can be better than the reliability of the S-N curves and the load spectra. Uncertainties involved were indicated in previous chapters.

The conclusion of this calculation example should be that the validity of the Palmgren-Miner rule is not the only weak point of our calculated result. At best we will obtain rough estimations about what the life might be. Better information will require relevant fatigue tests to be discussed later.

There is an alternative approach based on past experience. Two methods are indicated in figure 12.6. Starting from an older structure with good service fatigue records a new structure can be designed with at least the same standard of fatigue quality. Also, from past experience allowable stress levels can be adopted, which combined with careful design will also predict a satisfactory behaviour of the new structure. The confidence in these comparative methods can be increased by comparative fatigue tests between old and new design features.

12.4 FATIGUE TESTING METHODS

It has been said several times before that realistic information on fatigue properties of a new design will require supplementary fatigue tests. Let us now consider different types of fatigue tests and different purposes we can have in mind for fatigue tests. A survey is given in figure 12.7. It should be recalled that several decades ago only the conventional fatigue test (with CA-loading) could be carried out. Later program tests could be done, and still later random load tests (stochastic loading) and flight-simulation tests (mixture of deterministic and random loads), when computer controlled closed-loop fatigue machines were developed.
Conventional fatigue tests (CA-loading)

This is the most simple type of fatigue test. It is still abundantly used for providing fatigue curves and crack growth data on one hand, and for experimental comparisons on the other hand. Data of the first category can be used for fatigue predictions. Data of the second category should provide us with information on the effects of certain variables as discussed in chapter 7 and 8 (e.g. surface quality, notches, mean stress, corrosion etc.). However it was already pointed out that the result of comparisons can depend on the stress amplitude selected, see e.g. the discussion on figure 7.3 (surface quality) and on figure 7.36 (fretting corrosion). This can lead to test interpretation problems if we compare different design solutions for the same component. If two designs A and B have fatigue curves as shown in figure 12.8a comparative tests at $S_{a1}$ would indicate design B to be slightly better than design A. At stress level $S_{a2}$ however, design B would be highly superior to design A. It is even possible that S-N curves of two designs will intersect, see figure 12.8b. Comparative tests at $S_{a1}$ and at $S_{a2}$ would give opposite indications. The real question is that we want to know the differences between A and B for the load history that our aircraft will see in service. The apparently trivial solution is that this question requires comparative flight-simulation tests. There have been cases, for instance in comparing different fasteners, where tests with CA-loading indicated large life improvements for certain types of fasteners, whereas flight-simulation tests revealed relatively small improvements only. Although conventional fatigue tests can give useful information the results require further interpretation for other load-time histories.

Program tests and random load tests

Nowadays program tests are considered to be tests with a variable-amplitude loading, but a rather artificial load cycle sequence. Moreover it was found that program tests and random load tests with the same peak load distribution function did not necessarily give the same fatigue life. In other words a random sequence and a programmed sequence can lead to different fatigue lives and crack growth rates (Ref. 4). Also here the preference should go to the realistic sequence, because this now can be done with the present experimental techniques.
Flight-simulation tests

In the previous chapter (section 11.6.) it was discussed how flight-simulation tests can be a realistic representation of load-time histories to be expected in service. Examples were shown in figure 11.16. From the discussion above it now appears that flight-simulation tests can be carried out for different purposes:

(a) comparative tests to study effects of general interest.
(b) comparative tests for design studies.
(c) full-scale tests on components and structures.

There are many topics of general and practical interest, e.g. the effects of different surface qualities, different types of alloys, different fastener systems, etc. Realistic flight-simulation tests on unnotched specimens in many cases will not be a good proposition. The unnotched specimen is not a realistic structural element. Therefore a flight-simulation test on an unnotched specimen is combining a realistic fatigue load with an unrealistic specimen. In the past we sometimes did the opposite: a conventional fatigue test on a specimen, representing a joint in a structure. That is combining a realistic specimen with an unrealistic fatigue load. Both specimen and fatigue load should be realistic if we want information directly relevant to practical problems. For fatigue design problems for a particular aircraft the load spectrum of that aircraft should be adopted in the flight-simulation load sequences. For general evaluations, not directly related to a particular aircraft there are two standardized flight-simulation load sequences, viz. TWIST and FALSTAFF. TWIST (Ref. 5) is based on gust dominated wing bending (TWIST = Transport Wing Standard). FALSTAFF (Ref. 6) was designed for manoeuvre dominant wing bending of fighter aircraft (FALSTAFF = Fighter Aircraft Loading Standard For Fatigue evaluation).

Full-scale flight-simulation tests on a structure

If a flight-simulation test is carried out on a small component, for instance on a joint, the main purpose is to obtain direct life indications. However, if a complete structure, or large parts of the aircraft structure (e.g. the wing), will be tested under flight-simulation loading the goals are much wider.

1. Indication of fatigue critical elements and design deficiencies.
2. Determination of fatigue lives until visible cracking occurs.
4. Measurements on residual strength.

The full-scale test is the final proof of the fatigue quality of a new aircraft structure. It is possible that some structural fatigue problems were not well recognized during previous analysis. Moreover a separate component is not just similar to the same component as a part of a larger structure. Load transfer to the component may be different. After all it is also possible that certain structural details were erroneously supposed to be non-fatigue critical, or were simply overlooked as such. As an example open tool holes have caused fatigue cracks in full-scale fatigue tests. One of the main purposes of the full-scale tests therefore is to give a most realistic proof of a satisfactory fatigue performance, and to modify the structure if unexpected fatigue cracks turn up. Also the proof of sufficient fatigue resistance as part of airworthiness requirements becomes much easier if a full-scale fatigue test has been done.

If cracks arise during a full-scale test it will not always be necessary to modify the structure. That depends on the life until the crack was detected, on possibilities for replacement or repair in service, and on the growth rate of the crack. For the latter reason the crack will not be immediately removed or repaired. First observations will be made to see how fast the crack will grow. It is possible that only minor cracks arise during a full-scale test. In that case a number of artificial cracks (usually some cuts) will be made in order to obtain information about crack growth rates. This information is needed because cracks can start in service due to unintentional damage or other incidental reasons. We have to know whether the structure will be fail-safe in service, and which inspection periods have to be used. As suggested by figure 12.3 a safety factor of 3 is frequently adopted to arrive at safe inspection periods. The full-scale fatigue test is also a learning test to develop inspection procedures for the future operators of the aircraft. If small cracks can be found only by X-ray techniques or eddy current methods the initial experience comes from the full-scale test and will be documented in the maintenance handbooks for the operator.

At the end of the full-scale fatigue test the structure can be used for static fail-safe tests. For civil transport aircraft there is a tendency to require a residual strength equal to the limit load (LL)
which is 2/3 of the ultimate design load. Figure 12.9 shows crack growth curves of four artificially started fatigue cracks in the F-28 wing (Ref. 7). Note that the fastest crack propagated about 8 cm in 30000 flights. After simulating 150000 flights in the full-scale test (2.5 times the target life) the wing was loaded three times to limit load which caused small amounts of stable crack extension. The test was then continued as a research program, which amongst other things made it clear that the limit load applications gave a most significant crack growth retardation. This is the retardation discussed in chapter 9.4. It will be clear from figure 12.9 that the static fail-safe tests should be done at the end of the full-scale test. If limit loads are applied earlier they may retard small existing cracks, not yet found at that moment. As a consequence they never may be found. Indications on critical locations are lost and reliable information on crack growth is not obtained. Even if there are no cracks a limit load can cause local plasticity at various notches in the structure and change load transmission in joints. This can easily lead to longer fatigue lives than obtained without a limit load application. Unconservative information is then obtained.

A full-scale fatigue test on an aircraft structure is an expensive test, because it requires a complex test set-up with many hydraulic jacks, an extensive control system, electronics, etc. For that reason a careful planning of the load-time history has to be highly recommended in order to get the best and most relevant information of the test. Shortly after the Comet accidents in the early fifties several aircraft were tested with a "simplified" flight-simulation loading, see figure 12.10. All flights in the test were equal, and the gust cycles were calculated to give the same damage as the full gust spectrum. This calculation was done with the Palmgren/Miner rule. It was insufficiently recognized that this rule is fully unreliable to give indications of equal damage (see the shortcomings of this rule discussed in chapter 9). Later comparative tests between realistic flight-simulation tests and those simplified tests indicated that the latter ones may be expected to give conservative results. The explanation is that the benefit of higher loads (notch root plasticity) is not obtained in the simplified tests.

12.5 FULL-SCALE TESTS AND SCATTER FACTORS

If a crack is found in a full-scale test it may be questioned whether
the same crack would be found in service after the same number of flights. Fatigue lives are known for scatter, and a full-scale test is just one test, whereas there is a large number of the aircraft in service. Some aspects have to be recalled now.

1. The load spectrum in the full-scale test was chosen to be representative for a severe but still realistic use of the aircraft (see figure 11.7). Most aircraft should be expected to have longer lives.

2. Although a full-scale test is a single test most components will occur at least twice in a structure (left and right) if not more. If a crack is found in the test the other component should be watched for similar crack growth development. Of course this will give a primitive indication of scatter only, and the fatigue resistance of different aircraft can show more scatter for several reasons (e.g. variations in material and production quality). For calculations on scatter the life distribution function has to be known. Unfortunately we do not know it. Experience has learned that a normal distribution of \( \log N \) seems to be a reasonable approximation, with 0.12 as the order of magnitude of the standard deviation for a structure. Such information can be used to estimate how much earlier the first cracks can arise in service as compared to the full scale test. In view of the statistical assumptions to be made this is a rather delicate problem. Another conclusion to be drawn here is that a full-scale fatigue test on a structure should be carried on significantly longer than the target life of the aircraft. At least a duplication of the life is desirable. New cracks may then turn up which otherwise would have been missed.

3. Do we need a safety factor in view of scatter? For a fail-safe structure this is not really necessary. However, on economic grounds it is not a good aircraft if early cracks have to be expected. An acceptable minimum life is desirable.

The problem of scatter is fully different for safe life structures. A helicopter blade is a good example. A fatal accident will follow if a blade is lost. It is clear that substantial safety factors will be required. For aircraft components a safety factor of 6 on life was proposed in the past for safe life items, but it is questionable whether that is a good proposal for a helicopter blade. The blades in service are subjected to extremely high numbers of bending load cycles
superimposed on a high centrifugal load. All these cycles have to be below the fatigue limit because failure is not allowed. If a helicopter blade is tested until $10^9$ cycles without failure the test result does not provide much confidence that nothing will happen in service. It only proves that the blade was tested under its own fatigue limit, see figure 12.11. Another blade with a lower fatigue limit could fail after a relatively short life. It could have a lower fatigue limit due to some superficial surface damage (e.g. corrosion pit), or simply because of scatter of the fatigue limit. Under such conditions a safety factor on life is meaningless. What really matters is the scatter of fatigue strength and not of fatigue life (which should be infinite). Safety factors should then be applied to stress levels and not to life. The full-scale test on the blade can only indicate a reliable safe life if the safety factor is applied on the fatigue load in the test.

12.6 MAINTENANCE AND INSPECTION

Maintenance and inspection are two important aspects of the safety and the economy of operating aircraft. In Ref. 8 Holshouser and Mayner have given some most illustrative examples which will be briefly mentioned below.

a. Wing spar failure of a small transport aircraft.

A fatal accident was caused by a fatigue failure of a wing attachment fitting of a high strength steel.

The investigation revealed:
- at both sides of the critical hole a large part of the section was weakened by fatigue;
- the element was not fail-safe and it has a specified safe-life of 10000 hours after which replacement was mandatory. The aircraft failed after 9383 hours of service time;
- the fitting was chromium plated and this has reduced the safe-life to 10000 hours;
- the operator made considerable shorter flights at higher speeds and lower altitudes than the standard flight assumed by the aircraft manufacturer.

b. Wing of a transport aircraft.

The 7075-T6 wing structure completely failed during severe to extreme clear-air turbulence. The accident investigation revealed:
- the failure started from 2 fatigue cracks at either side of an access door, crack lengths being 3.25" and 2.5" respectively;
the critical location was covered by X-ray inspections, since it was known that cracks might originate in that area. Three sets of X-ray pictures from previous inspections indicated the two cracks and the aircraft had flown with the cracks for more than one year. The maintenance records did not indicate that the cracks had been detected.

C. Helicopter main rotor spindle.

This part was made from 4340 steel. A fatigue failure in a fillet caused a fatal accident. Investigation of the failure revealed among other things:
- crack growth had been slow and the crack was probably present during the last magnetic particle inspection 2 months before the accident;
- the crack nucleus started in an area with very small shallow pits. Moreover the Rockwell C hardness in that area was well below 28 as compared to the specified minimum of 34 Rc;
- the fillet area had not been properly shot peened.

There are several lessons to be learned from these accidents. The last example showed that production errors were additive in causing the accident. Apparently the errors could pass the inspection after production, while a crack, probably being present, could pass a service inspection.

The second accident actually proved the structure to be fail-safe, but this feature is meaningless if it is not backed up by an effective inspection.

The third example illustrates the risk of the safe-life philosophy. Unfortunately a practically similar accident occurred to the same aircraft type one year later. It then turned out that not all information on this critical topic had reached the inspectors, who carried out the periodic inspections on this lug. The communication of relevant information may be another weak link.

Holshouser and Mayner surveyed 230 failed components and in 60 percent the mode of failure was fatigue. Their general conclusion was: The most frequently identified cause was improper maintenance, including inadequate inspection, while fabrication defects, design deficiencies, defective material, and abnormal service damage also caused many fatigue failures.
CONCLUDING REMARKS

In the present chapter it has been emphasized that there are many design aspects involved in obtaining an aircraft structure with good fatigue properties. Most of these aspects cannot be treated (or not fully be treated) by analytical design rules and mathematical calculations. First comes that all those fatigue design aspects are properly recognized. Then a qualitative understanding will be major tool to arrive at good solutions and to indicate when analytical work in depth is justified. It will also indicate when insufficient knowledge has to be backed up by fatigue tests. It was made clear that in most cases "realistic" fatigue tests should be done, realistic with respect to both the load-time history and the test specimen (preferably full size). It finally should be completed by a flight-simulation test on the full-scale structure. If a new type of aircraft is designed it is the responsibility of the aircraft industry to develop a safe and fatigue resistance aircraft structure, which will not cause serious problems to the operator.

The aircraft operator is responsible for maintenance and inspection. Some case histories have emphasized the necessity that the operator should be fully aware of the possibility of fatigue cracks, also unexpected cracks. It is not always easy to maintain reliable inspection techniques, especially if tiny fatigue cracks are to be found. Another aspect discussed in the previous chapter is the question whether the operator is using the aircraft in agreement with the original design assumptions. Load monitoring systems to be carried out by the operator in cooperation with the aircraft manufacturer are very much desirable.

REFERENCES


General references on fatigue of aircraft structures

ICAF Symposium Proceedings issued by:

10th Symp. 1979: Aeronautics Administration, Belgium (ICAF Doc. No. 1169): Structural fatigue as a design factor.

AGARD Lecture Series (LS), Conference Proceedings (CP) and AGARDoGraphs (AG)


CP-141: Specialist meeting on design against fatigue.

AG-157: The accumulation of fatigue damage in aircraft materials and structures (1972).
AG-176: Fracture mechanics of aircraft structures.

ASTM STP publications

STP 486: Damage tolerance in aircraft structures (1971).
STP 714: Effect of load variables on fatigue crack initiation and propagation.
Table 12.1: Fatigue damage calculation for spectrum H and three S-N curves shown in figure 12.5

<table>
<thead>
<tr>
<th>Data of load spectrum</th>
<th>S-N data</th>
<th>Damage calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_a$</td>
<td>$H$ cycles in</td>
</tr>
<tr>
<td></td>
<td>number of exceedings</td>
<td>interval</td>
</tr>
<tr>
<td>250</td>
<td>100</td>
<td>450</td>
</tr>
<tr>
<td>230</td>
<td>550</td>
<td>700</td>
</tr>
<tr>
<td>210</td>
<td>1250</td>
<td>1150</td>
</tr>
<tr>
<td>190</td>
<td>2400</td>
<td>1900</td>
</tr>
<tr>
<td>170</td>
<td>4300</td>
<td>3200</td>
</tr>
<tr>
<td>150</td>
<td>7500</td>
<td>6000</td>
</tr>
<tr>
<td>130</td>
<td>13500</td>
<td>11500</td>
</tr>
<tr>
<td>110</td>
<td>25000</td>
<td>21000</td>
</tr>
<tr>
<td>90</td>
<td>46000</td>
<td>49000</td>
</tr>
<tr>
<td>70</td>
<td>95000</td>
<td>125000</td>
</tr>
<tr>
<td>50</td>
<td>220000</td>
<td>480000</td>
</tr>
<tr>
<td>30</td>
<td>700000</td>
<td>4300000</td>
</tr>
<tr>
<td>10</td>
<td>5000000</td>
<td></td>
</tr>
</tbody>
</table>

H for 5000 hours

Life estimation with $2n/N = 1$

- curve A: $\frac{1}{0.1353} \times 5000 = 37000$ hrs
- curve B: $\frac{1}{0.1748} \times 5000 = 29000$ hrs
- curve C: $\frac{1}{0.2150} \times 5000 = 23000$ hrs