Final Draft
of the original manuscript:

Zerbst, U.; Lunden, R.; Edel, K.-O.; Smith, R.A.:  
**Introduction to the damage tolerance behaviour of railway rails**  
DOI: 10.1016/j.engfracmech.2009.09.003
INTRODUCTION TO THE DAMAGE TOLERANCE BEHAVIOUR OF RAILWAY RAILS

Zerbst, U.\textsuperscript{a}, Lundén, R.\textsuperscript{b}, Edel, K.-O.\textsuperscript{c} and Smith, R.A.\textsuperscript{d}

\textsuperscript{a} GKSS Research Centre, Institute for Materials Research, Materials Mechanics, D-21502 Geesthacht, Germany

\textsuperscript{b} Chalmers University of Technology, CHARMEC, Dept. of Applied Mechanics, SE-412 96 Gothenburg, Sweden

\textsuperscript{c} University of Applied Science (FH), Magdeburger Str. 50, D-14770 Brandenburg, Germany

\textsuperscript{d} Royal Academy of Engng. Research, Dept. Mech. Engineering, Imperial College London, Exhibition Road, SW7 2BX, UK

ABSTRACT

Despite substantial advantages in material development and in periodic non-destructive inspection together with periodic grinding and other measures in order to guarantee safe service, fatigue crack propagation and fracture is still in great demand as emphasised by the present special issue. Rails, as the heart of the railway system, are subjected to very high service loads and harsh environmental conditions. Since any potential rail breakage includes the risk of catastrophic derailment of vehicles, it is of paramount interest to avoid such a scenario. The aim of the present paper is to introduce the most important questions regarding crack propagation and fracture of rails. These include the loading conditions: contact forces from the wheel and thermal stresses due to restrained elongation of continuously welded rails together with residual stresses from manufacturing and welding in the field, which is discussed in Section 2. Section 3 provides an overview of crack-type rail defects and potential failure scenarios. Finally the stages of crack propagation from initiation up to final breakage are discussed.

Keywords: Rail, fatigue crack propagation, fracture, damage tolerance.

1. INTRODUCTION

In October 2000 a high-speed train derailed less than one kilometre south of Hatfield station near London in UK. Four passengers where killed and more than seventy people were injured. The cause of the accident was fracture and subsequent fragmentation of the outer rail on a curved section. The investigations revealed numerous fatigue cracks at the running corner. When one of these penetrated into the web and foot of the rail it resulted in a knock-on effect leading to the extension of adjacent cracks and, thereby, to the fracture of a complete rail section \cite{1}.

Usually rail breakage will not have such drastic consequences. Nevertheless it is a paramount objective of every railway company worldwide to avoid rail breakages. Smith \cite{2}, collecting data from various sources, provides a comparison of the frequency of fracture events of various railway components in Great Britain at the end of the 19th and the end of the 20th century (Table 1). Although the list is certainly subject to some uncertainties, he is able to
conclude: “…it is clear that whilst failures of wheels and axles have been reduced by a factor of 20 over the last century, failures of rails per train kilometre have actually increased by a factor of more than 2”. Reasons behind this trend are heavier axle loads, increased volumes of traffic and axial tensile stresses at low temperatures due to continuously welded rails.

Table 1: Mechanical failures on UK railways in 1881-1890 and 1992-1994 (Rates per 10^6 train kilometres (according to [2]))

<table>
<thead>
<tr>
<th>Components</th>
<th>1881-1890</th>
<th>1992-1994</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheels or tyres</td>
<td>2.0</td>
<td>0.04</td>
</tr>
<tr>
<td>Axles</td>
<td>0.8</td>
<td>0.04</td>
</tr>
<tr>
<td>Rails</td>
<td>0.7</td>
<td>1.5</td>
</tr>
</tbody>
</table>

A thirty years statistics covering the time from 1969 to 1999 gives an average number of 767 +/- 128 broken rails per year on the British railroad network. Comparable railway systems worldwide experienced comparable numbers of rail breakages. Systems for heavy haul freight with their high axle loads are significantly more affected than systems with predominantly passenger transport [3]. In [4] the authors give an approximate figure of the economic costs of rail fracture and its avoidance: € 2000 million per year in the European Union alone.

A positive trend found with respect to the mentioned British network was that, whilst the detection rate of damaged rails which then had to be removed, has been increased continually throughout the period under consideration, the number of breakages was virtually constant (Figure 1). This fact shows that countermeasures, such as non-destructive inspection and periodical grinding, have brought a significant improvement towards failure prevention. On the other hand, the requirements on the networks such as an increased volume of traffic and higher axle loads etc. are permanently increasing. Therefore, fatigue crack propagation in rails remains an important issue with respect to both the quantitative understanding of the mechanisms and the development of analysis routines for practical application.

Figure 1: Long term trend of broken and defective rails removed in Railtrack railroad network (according to [3]).

The aim of the present paper is to present an overview of the damage tolerance behaviour of rails. In section 2, loads on the rails will be discussed. Section 3.1 provides information on
crack types and failure scenarios of rails and section 3.2 describes the stages of rail fatigue as well as the factors that have to be considered in modelling the damage behaviour of rails. The present review paper is aimed as an introduction to the technical papers in this special issue.

2. LOADING ON RAILS

2.1 General Remarks

The structure of a conventional railway track is illustrated in Figure 2. It contains elements such as the rails, the sleepers including rail pads and fastening elements, e.g. clips, the ballast bed and the subgrade.

![Figure 2: Basic elements of a railway track (dimensions according to [5]).](image)

According to Esveld [6] (see also [5]) the highest stresses occur at the running surface of the rail where the wheel-rail contact stresses typically can reach 1500 MPa for an axle load of 25 tonnes. The stresses between rail and sleeper are much smaller; in the order of 2.5 MPa and between the ballast and the subgrade only 50 kPa. The ballast causes damping of dynamic forces during train passages. Note, however, that some high-speed tracks do not use a common ballast bed. E.g. the German ICE track between Cologne and Frankfort is a slab track with a concrete foundation.

Methods of joining rail sections are shown in Figure 3. Up to the 1930s - and at some tracks even today – joining was performed by so-called fishplates which were attached to the rails by bolts (Figure 3a). Since there was a danger of fatigue cracks developing in the fishplates and in the rails initiated at the holes, frequent maintenance was required. The gaps between the rail segments contributed to wheel damage as well as to noise annoyance. Around 1930 continuous welding of the rails was introduced which solved these problems, however, it also created new ones. As in any weldment the material microstructure, and dependent on this the toughness, is inhomogeneous (for more details see [7]). The welding process creates residual stresses which contribute to the total stress levels and can also affect the straightness and alignment of the rail. Whereas the thermal stresses, at temperatures deviating from the rail neutral temperature at which the track was installed, in the old designs were concentrated at the fishbolts and rail ends as the weakest links, they now affect the entire rail length in welded tracks.

Sometimes, when track circuits are used for signalling purposes, insulation joints are used even in continuously welded tracks. In order to compensate the loss of strength due to the insulated block, side plates similar to the fishplates fastened by epoxy resin are used.
Figure 3: Rail joints. (a) fishplate with bolts; (b) continuous (aluminoothermic) welded rail; (c) fishplate joint of a broken rail; (d) defective isolation joint.

The propagation of a fatigue crack in a rail is driven by the contact stresses as well as bending and shear stresses arising from the load during wheel passage. These stresses are superimposed by further loading components such as residual stresses from manufacturing
and thermal stresses which depend on the ambient temperature. The complex longitudinal stress state in a rail is illustrated in Figure 4. Additionally, the rail is subjected to longitudinal and lateral forces as shown in Figure 5.

![Figure 4: (a) A wheel rolling on a continuously welded rail. (b) Contact stresses and longitudinal stress components.](image)

![Figure 5: Vertical, longitudinal and lateral forces introduced by a railway wheel on a rail.](image)
2.2 Rail Stresses Due To Wheel Loading

2.2.1 Bending stresses

Rail bending comprises a vertical and a lateral component. Based on the beam-on-elastic-foundation theory for plane bending applied to rails by Zimmermann [8] (see also [9,10]) the bending moments $M_y$ and $M_z$ can be determined by

$$M_y(x) = -\frac{F_V}{4\lambda_V} (\cos \lambda_V x - \sin \lambda_V x) \cdot \exp(-\lambda_V x)$$

(1)

and

$$M_z(x) = -\frac{F_L}{4\lambda_L} (\cos \lambda_L x - \sin \lambda_L x) \cdot \exp(-\lambda_L x)$$

(2)

with $x$ being the longitudinal position on the rail, $F_V$ and $F_L$ being the vertical and lateral wheel loads (Figure 6), and $\lambda_V$ and $\lambda_L$ being parameters which depend on the rail geometry and the vertical and lateral foundation stiffnesses $k_V$ and $k_L$. These are different for high-speed tracks, tracks for mixed traffic etc. Assuming $k_L = 0.85 k_V$ [10], $\lambda_V$ and $\lambda_L$ are obtained as

$$\lambda_V = \sqrt{\frac{k_V}{4E \cdot I_{yy}}}$$

(3)

and

$$\lambda_L = \sqrt{\frac{0.85 k_V}{4E \cdot I_{zz}}}$$

(4)

In Eqs. (3) and (4), $E$ is the modulus of elasticity of the rail steel and $I_{yy}$ and $I_{zz}$ are the second area moments of inertia with respect to the horizontal and vertical axis respectively, through the rail centroid. Values for $I_{yy}$ and $I_{zz}$ for some common rail sections are given in Table 2.

The formulae above are strictly valid only when the external loading and the support forces pass through the shear centre axis of the rail.
Figure 6: Rail head bending stress as a function of the wheel position (left); Lateral and eccentric vertical loading, $F_L$ and $F_V$, of a rail (right) (according to [10]).

Table 2: Second area moments of inertia for some rail profiles (according to [11-13])

<table>
<thead>
<tr>
<th>Rail type (EN 13674)</th>
<th>60 E 1</th>
<th>56 E 1</th>
<th>54 E 3</th>
<th>Previous designation</th>
<th>UIC 60 BS 113lb BR Variant</th>
<th>DIN S54</th>
<th>JIS 60</th>
<th>136 RE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{xy}$</td>
<td>3038 cm$^4$</td>
<td>2321 cm$^4$</td>
<td>2074 cm$^4$</td>
<td>3083 cm$^4$</td>
<td>3950 cm$^4$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$I_{zz}$</td>
<td>512 cm$^4$</td>
<td>422 cm$^4$</td>
<td>355 cm$^4$</td>
<td>604 cm$^4$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to Eq. (1) the maximum tensile stress at the rail head surface occurs at a distance

$$x_o = \frac{\pi}{2\lambda_v}$$

from the wheel position (Figure 6), a phenomenon sometimes designated as “reverse bending”. Although the bending stresses induced by the lateral load contribute to fatigue damage, bending stresses induced by the vertical load dominate rail failure [14].

Note that Eqs. (1) and (2) refer to the continuous portions of rails and not, for instance, to the rail ends of a fishplate joint, although such a joint is known to be susceptible to fatigue damage. Possible treatment of rail ends is provided by Edel [15].

2.2.2 Shear stresses

Besides the bending stresses the wheel load also generates shear stresses in the rail section. These are the main cause of failures at bolt holes in fishplate-joined rails [14]. An example of the history of the shear force at a bolt hole during a wheel passage is given in Figure 7 [16]. Note that shear stresses play a major role in rails with cracks where they cause mixed mode loading conditions such as illustrated in Figure 8.
2.2.3 Dynamic effects

Dynamic effects are due to car and bogie motions but also due to the dynamic response of the track including its ballast bed and subgrade [17] which make the wheel loads vary at frequencies up to 10 Hz [10].

For a static axle or wheel load, the dynamic load $F_v(\text{dyn})$ can be modelled as a statistical distribution, the upper bound of which can be used as worst condition for design purposes. This is realised by multiplying the static load with a magnification factor $K_{\text{dyn}}$ as

$$F_v(\text{dyn}) = K_{\text{dyn}} \cdot F_v(\text{stat})$$ (6)
The magnification factor depends on the properties and quality of the train and the track and on the train speed. As an example, Deutsche Bahn recommends [18]:

\[ K_{\text{dyn}} = 1 + 3 \cdot n \cdot \varphi \]  

(7)

with \( n = 0.15 \) to 0.25 for different types of tracks and

\[
\varphi = \begin{cases} 
1 & \text{for } v \leq 60 \text{ km/h} \\
1 + 0.5(v - 60)/190 & \text{for } 60 \leq v \leq 300 \text{ km/h (passenger trains)} \\
1 + 0.5(v - 60)/80 & \text{for } 60 \leq v \leq 140 \text{ km/h (freight trains)} 
\end{cases}
\]  

(8)

Note, however, that irregularities in the rails (e.g. at sites damaged by spalling, at rail joints or, as shown in Figure 9a, corrugation) and in the wheels (e.g. out-of roundness, flat spots, Figure 9b) can significantly increase the dynamic effect, particularly at high speed. The effect of a wheel flat on both the peak rail-wheel contact force and the peak rail bending moment is illustrated in Figure 10 [19] (see also [20]).

**Figure 9:** Irregularities such as corrugation of the rails (a) or flat spots on the wheels (b) which increase dynamic load magnification.
2.2.4 Wheel-rail contact stresses

2.2.4.1 General

The forces arising between wheel and rail generate so-called contact stresses in a local volume of the two bodies. The most well-known calculation model is the Hertzian one which will be described in some detail. This model is important since it describes the local stresses with good accuracy for the most common wheel-rail contact problems. Further, it provides a good understanding of general contact phenomena. Although Hertzian theory is valid only for elastic contacts it can be useful far beyond that. Limited plastification will not affect the contact stresses very much. After a number of overrollings, even with severe plastification, residual stresses will build up which to a certain level may prevent further plastification, a phenomenon known as shake-down. The Hertzian theory will then again be valid and the calculated stresses can be superimposed on the residual stress state.

Also stresses at non-Hertzian contacts and elastoplastic contacts and the influence of surface roughness will be shortly described.

2.2.4.2 Hertzian contact

Important assumptions are linear elastic material, small contact area compared to the radii at the contact of the bodies and to other dimensions (semi-infinite bodies are assumed), and smooth surfaces at both macro and micro scale. Classical descriptions of Hertzian contacts are given in [21], [22] and [23] and more recent ones in [24] (Chapter 4) and [25] (Chapter 4).
Body 1 (wheel) is assumed to have principal radii $R_1$ and $R'_1$ and Body 2 (rail) principal radii $R_2$ and $R'_2$ at the contact, see Figure 11. The radii are counted positive for a convex curvature and negative for a concave curvature. The two bodies (wheel and rail) are here assumed to have the same material parameters: elastic modulus $E$ and Poisson’s ratio $\nu$. This means that locally identical stress fields will be induced in the wheel and the rail. The coordinate system is assumed to have its origin at the centre of the contact patch which will be elliptic with semi-axes $a$ and $b$, see Figure 12a. The contact pressure distribution becomes

$$p_z(x, y) = p_0 \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \quad -a \leq x \leq a$$
$$-b \leq y \leq b$$

where the maximum pressure $p_0$ is related to the total contact force $F_z$ as

$$p_0 = \frac{3F_z}{2\pi ab}$$

and where

$$a = \left[ \frac{6A^2E}{\pi} \times \frac{F_z(1-\nu^2)}{ER_{SN}} \right]^{1/3}, \quad b = \left[ \frac{6E}{\pi A} \times \frac{F_z(1-\nu^2)}{ER_{SN}} \right]^{1/3} \quad (11a-b)$$

Here $A$, $E$ and $R_{SN}$ are defined by

$$\theta = \frac{(A^2 + 1)E - 2K}{(A^2 - 1)E}, \quad A = a/b \text{ or } A = b/a, \quad k^2 = 1 - 1/A^2$$

with

$$\theta = R_{ST}/R_{SN}, \quad R_{ST} = \frac{1}{R_1} - \frac{1}{R'_1} + \frac{1}{R_2} - \frac{1}{R'_2}, \quad R_{SN} = \frac{1}{R_1} + \frac{1}{R'_1} + \frac{1}{R_2} + \frac{1}{R'_2}$$

$$\quad (13a-c)$$
Figure 11: Wheel in contact with rail. Radii $R_1$ and $R'_1$ of wheel and $R_2$ and $R'_2$ of rail. Forces $F_x$, $F_y$ and $F_z$ acting on wheel. Contact pressure distribution with semi-axes $a$ and $b$ corresponding to force $F_z$ is indicated. Counterdirected contact forces and pressure distributions are acting on rail (not indicated). Figure from [26].

Figure 12: Semi-infinite body used for calculation of local contact stresses in both wheel and rail. Co-ordinate system $0xyz$. (a) Rolling contact force components $F_x$, $F_y$ and $F_z$ (shown elevated over origin $0$) corresponding to Hertzian rolling pressure $p_z(x, y)$ with maximum $p_0$. Semi-axes $a$ and $b$ of contact ellipse. Resulting stresses $\sigma_x$, $\sigma_y$, ... at an arbitrary point $Q = (x_q, y_q, z_q)$ obtained from integration of influence from concentrated forces in Figure b. Frictional forces are taken as $p_x(x, y) = \mu_x p_z(x, y)$ and $p_y(x, y) = \mu_y p_z(x, y)$ where $\mu_x$ and $\mu_y$ are friction coefficients (as made use of). For a certain point $Q = (x_q, y_q, z_q)$, overrolling by Hertzian load together with frictional forces is simulated by letting $x_q$ vary while keeping $y_q$ and $z_q$ constant. (b) Components $P_x$, $P_y$ and $P_z$ of a concentrated force at point $P = (x_p, y_p, 0)$ on surface of semi-infinite body $z > 0$ giving static stresses $\sigma_x^*$, $\sigma_y^*$, $\sigma_z^*$, $\tau_{xy}^*$, $\tau_{yz}^*$, $\tau_{zx}^*$ at point $Q = (x_q, y_q, z_q)$. 

\[
p_z(x, y) = p_0 \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}
\]

\[
p_0 = \frac{3F_z}{2\pi ab}
\]
Further \( E \) and \( K \) are complete elliptic integrals with modulus \( k^2 \). Equation (12a) is solved for \( A \) by use of iteration. It is here assumed that \( R_1 \) and \( R_2 \), and also \( R'_1 \) and \( R'_2 \), are in the same plane, respectively. Further, it should be noted that

\[
A = a / b \text{ for } \frac{1}{R_1} + \frac{1}{R_2} < \frac{1}{R'_1} + \frac{1}{R'_2} \quad \text{and} \quad A = b / a \text{ for } \frac{1}{R_1} + \frac{1}{R_2} \geq \frac{1}{R'_1} + \frac{1}{R'_2} \tag{14a,b}
\]

There exist analytical solutions for evaluating the surface and subsurface stresses at Hertzian contacts caused by \( F_z \), see [21]. However, a numerical integration method is preferable today since it is quite straight-forward and offers the possibility to consider also contact shear loads \( F_x \) and \( F_y \). In addition, subsurface stresses from any known pressure and/or contact shear stress distribution can be evaluated in this way. The analytical solutions of Boussinesq’s and Cerruti’s problems for a semi-infinite body, see [27], are then employed giving the stress field \( \sigma^*_x, \sigma^*_y, \sigma^*_z, \tau^*_{xy}, \tau^*_{yz}, \tau^*_{xz} \) for concentrated loads \( P_x, P_y \) and \( P_z \), see Figure 12a. For example, the solutions for \( \sigma^*_z \) and \( \tau^*_{xz} \) are

\[
\sigma^*_z = \frac{3}{2\pi} \left[ -\frac{P_x x Z^2}{R^5} - \frac{P_y y Z^2}{R^5} - \frac{P_z z^2}{R^5} \right], \quad \tau^*_{xz} = \frac{3}{2\pi} \left[ -\frac{P_x y^2 z}{R^5} - \frac{P_y x^2 z}{R^5} - \frac{P_z x^2 y}{R^5} \right] \tag{15a,b}
\]

where

\[
R = \left[ x^2 + y^2 + z^2 \right]^{1/2}, \quad x = x_q - x_p, \quad y = y_q - y_p, \quad z = z_q - z_p, \quad z_p = 0 \tag{16a-e}
\]

For the Hertzian contact problem and with fully developed friction forces one puts \( P_x = \mu_x, \quad P_y = \mu_y, \quad P_z = 1 \) in formulae (15a,b) and each of the stresses \( \sigma_x, \sigma_y, \sigma_z, \tau_{xy}, \tau_{yz} \) and \( \tau_{xz} \) (here denoted \( \sigma \)) is then evaluated through

\[
\sigma = \int \int_A \sigma^*_z p_z \, dx dy \tag{17}
\]

which can be performed by a standard numerical integration routine. When evaluating the stress at a location near the contact surface, convergence problems will arise because of small values of \( R \). Normally, this can be handled by choosing a reasonably small value of the coordinate \( z \) (instead of \( z = 0 \)), see Section 2.2.4.3.
2.2.4.3 Hertzian contact - numerical example

Calculations of the stresses by use of the integration method described in Section 2.4.4.2 will be demonstrated. A standard freight car, with axle load $2m = 22.5$ tonnes and wheel diameter $D = 920$ mm, slowly rolling in the $x$-direction (Figure 12) on a straight rail with railhead radius $300$ mm is chosen as an example and reference case. This means that the wheel load (contact load) is $F_z = mg = \left(\frac{22500 \text{ kg}}{2}\right) \times 9.81 \text{ m/s}^2 = 110.4$ kN (no dynamic magnification factor is considered). The longitudinal friction coefficient is taken as $\mu_x = -0.300$ (braking force here means $\mu_x < 0$).

For the calculation of the Hertzian contact pressure according to Section 2.4.4.2, the radii thus are $R_1 = 0.460$ m, $R'_1 = \infty$, $R_2 = \infty$ and $R'_2 = 0.300$ m. The material parameters are taken as $E = 210$ GPa and $\nu = 0.300$. The calculation yields the maximum Hertzian pressure $p_0 = 1270$ MPa and the semi-axes $a = 7.43$ mm and $b = 5.59$ mm, see Case 1 in Table 3. Note again that the Hertzian pressure, and the local stresses determined in the following, are the same for the wheel and the rail since these stresses are approximated by applying the calculated surface stresses on a semi-infinite body, see again Figure 12.

In Table 3 the contact force $F_z$ and the radii $R_1$, $R'_1$ and $R'_2$ are varied to study their influence. The wheel loads in Case 2 and Case 3 correspond to axle loads 20 and 25 tonnes, respectively. In Case 6 and Case 7 a concave (worn) wheel is studied by use of negative values of $R'_1$.

Table 3: Results of calculation of Hertzian stresses for wheel-rail contact at top of rail. Wheel load $F_z$ perpendicularly to contact patch. The material parameters are $E = 210$ GPa and $\nu = 0.300$. One rail radius is $R_2 = \infty$. Wheel radii $R_1$ and $R'_1$ and rail radius $R'_2$ are varied. Calculated semi-axes $a$ and $b$ and maximum contact pressure $p_0$ are given. Case 1 is the reference case.

<table>
<thead>
<tr>
<th>Case</th>
<th>$F_z$ [kN]</th>
<th>$R_1$ [mm]</th>
<th>$R'_1$ [mm]</th>
<th>$R'_2$ [mm]</th>
<th>$a$ [mm]</th>
<th>$b$ [mm]</th>
<th>$p_0$ [MPa]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>110.4</td>
<td>460</td>
<td>$\infty$</td>
<td>300</td>
<td>7.43</td>
<td>5.59</td>
<td>1270</td>
</tr>
<tr>
<td>2</td>
<td>98.1</td>
<td>510</td>
<td>$\infty$</td>
<td>300</td>
<td>7.14</td>
<td>5.37</td>
<td>1221</td>
</tr>
<tr>
<td>3</td>
<td>122.6</td>
<td>510</td>
<td>$\infty$</td>
<td>300</td>
<td>7.69</td>
<td>5.79</td>
<td>1315</td>
</tr>
<tr>
<td>4</td>
<td>110.4</td>
<td>510</td>
<td>$\infty$</td>
<td>300</td>
<td>7.81</td>
<td>5.49</td>
<td>1230</td>
</tr>
<tr>
<td>5</td>
<td>110.4</td>
<td>410</td>
<td>$\infty$</td>
<td>300</td>
<td>7.02</td>
<td>5.70</td>
<td>1316</td>
</tr>
<tr>
<td>6</td>
<td>110.4</td>
<td>460</td>
<td>-500</td>
<td>300</td>
<td>6.37</td>
<td>8.82</td>
<td>937</td>
</tr>
<tr>
<td>7</td>
<td>110.4</td>
<td>460</td>
<td>-400</td>
<td>300</td>
<td>5.85</td>
<td>11.05</td>
<td>816</td>
</tr>
<tr>
<td>8</td>
<td>110.4</td>
<td>460</td>
<td>$\infty$</td>
<td>200</td>
<td>7.89</td>
<td>4.53</td>
<td>1474</td>
</tr>
</tbody>
</table>
Figure 13: Calculated stresses $\sigma_x, \sigma_y, \sigma_z, \tau_{zx}$ (stresses $\tau_{xy}$ and $\tau_{yz}$ not given) in Figure 12 from Hertzian rolling pressure $p(x, y)$ with $p_0 = 1270$ MPa, $a = 7.43$ mm, $b = 5.59$ mm and with friction coefficient $\mu = -0.30$ (braking force here means $\mu < 0$) as functions of co-ordinate $x$ (co-ordinate $y = 0$). (a) Stress distribution at depth $z = 0.200$ mm. Extreme values of the stresses are $\sigma_z = -1267$ MPa and $\tau_{zx} = 347$ MPa (to be compared with surface stresses $\sigma_z = -1270$ MPa and $\tau_{zx} = 0.300 \times 1270 = 381$ MPa). (b) Stress distribution at depth $z = 3.00$ mm. Extreme values of the stresses are $\sigma_z = -943$ MPa and $\tau_{zx} = 290$ MPa.

In Figure 13a the calculated stresses $\sigma_x, \sigma_y, \sigma_z, \tau_{zx}$ for the reference case in Table 3 are given along an axis in the plane $y = 0$ just below the $x$-axis in Figure 12 with co-ordinate $z = 0.200$ mm. As expected the stresses $\sigma_x$ and $\tau_{zx}$ are found to closely follow the applied loadings $p(x, y)$ and $\mu_p(x, y)$. In the vicinity of $x = a$ it is observed that $\sigma_x > 0$ (tensile stress) which is induced by the contact loading in shear. The stresses $\tau_{xy}$ and $\tau_{yz}$ are both zero because of symmetry. The depth $z = 0.200$ mm is chosen to be reasonably close to the surface and leads to avoidance of the numerical problems arising for small values of the co-ordinate $z$.

In Figure 13b the corresponding stresses are given for the depth $z = 3.00$ mm. It is found that the stresses $\sigma_x$ and $\sigma_y$ are strongly reduced while $\sigma_z$ and $\tau_{zx}$ maintain a large part of their magnitude ($\tau_{zx}$ with a modified distribution) as compared to the stresses for $z = 0.200$ mm.

Since plastification is an important material response to contact loading it is of interest to study the induced effective stress. Here the von Mises effective stress

$$\sigma_{\text{vM}} = \left[ \sigma_x^2 + \sigma_y^2 + \sigma_z^2 - \sigma_x \sigma_y - \sigma_y \sigma_z - \sigma_z \sigma_x + 3\tau_{xy}^2 + 3\tau_{yz}^2 + 3\tau_{zx}^2 \right]^{1/2}$$

is chosen which is numerically easy to handle. The von Mises yield criterion states that plastification does not occur for $\sigma_{\text{vM}} < \sigma_y$ where $\sigma_y$ is the yield stress in uniaxial testing of the material. Contour plots of the von Mises effective stress are given in Figure 14 for the same problem as in Figure 13, although the friction coefficient is chosen to be either $\mu_x = 0$ or $\mu_x = 0.300$.

In Figure 14a the maximum effective stress is found at a depth of $z = 2.4$ mm. However, the tendency is that the stresses at the surface become higher with increasing values of $\mu_x$. For $\mu_x = 0.300$, see Figure 14b, the point of highest stress is still below the surface at depth $z = 2.1$ mm. Further calculations show that for $\mu_x = 0.35$ the highest effective stress will be at the contact surface $z = 0$. 
A typical initial yield stress level of ordinary rail materials (R260) is 410 MPa with deformation hardening up to 600 or 700 MPa. Thus plastification will occur for the stress levels in Figure 14. This will be further discussed in Section 2.2.4.6.

According to [22] the maximum shear stress under the contact pressure ($\mu_x = 0$) is $\tau_{\text{max}} \approx 0.31 \times p_0$ at depth $z_{\text{max}} \approx 0.25 (a + b)$ which means $\tau_{\text{max}} \approx 394$ MPa at depth $z_{\text{max}} \approx 3.2$ mm. The agreement with the calculation using von Mises stress is relatively good since the shear stress (Tresca yield criterion) should be compared with half the value of the von Mises stress.

![Figure 14: Contour plots of von Mises effective stress $\sigma_{\text{vM}}$ in section $y = 0$ for the same loading as in Figure 13 except for friction coefficient (braking force here means $\mu_x < 0$) which is (a) $\mu_x = 0$ and (b) $\mu_x = -0.300$.](image)

2.2.4.4 Conformal contact

In section 2.2.4.3 the wheel-rail contact patch is located on top of the rail. This means a contact between two convex bodies and is named non-conformal contact. However, for wheel-rail contact at the rail edge as may occur on curves, see Figure 15, the contacting part of the wheel is concave near the flange which means that the wheel and rail radii are near to “matching” each other. This is called conformal contact, see [23].

For conformal contacts the use of Hertzian theory is questionable for several reasons. Geometrical requirements are not fulfilled, see [23]. For the wheel-rail contact the stresses often become very high which means that plastification will have a major influence on the resulting contact stresses. However, a calculation can still be valuable to study magnitudes and trends.

In Table 4 results of calculations of Hertzian stresses for the contact in Figure 15 are given. The same geometry, material parameters and loading as in Section 2.2.4.3 is used. The contact friction is assumed to be zero and the only load thus acts perpendicularly to the contact surface. The radius $R'_1$ starts at $-30$ mm and is then numerically reduced for an increasing conformity with the rail radius $R'_2 = 13$ mm. Note that the longest semi-axis $a$ is in the x-direction (along the rail) as for the non-conformal contact in Section 2.2.4.3. The calculated contact pressures become very high. Also the maximum von Mises stress $\sigma_{\text{vM}}^{\text{max}}$ becomes very high and is located very close to the surface.

A comparison of Hertzian stresses with results from an FE simulation is performed in [28]. The plastification in the contact zone means a larger contact area followed by reduced contact stresses and by residual stresses being induced in the rail.
Table 4: Results of calculation of Hertzian stresses for wheel-rail (conformal) contact at rail edge. Wheel load is $F_z = 110.4$ kN (acting perpendicularly to contact patch). Wheel radius is $R_1 = 0.460$ m and rail radii are $R_2 = \infty$ and $R'_2 = 0.013$ m. The material parameters are $E = 210$ GPa and $\nu = 0.300$. For four values of the wheel radius $R'_1 < 0$ (concave surface), the semi-axes $a$ and $b$, the maximum contact pressure $p_0$, the maximum von Mises stress $\sigma_{vM}^{\text{max}}$ and the location $z_{\text{max}}$ of the latter are given.

<table>
<thead>
<tr>
<th>$R'_1$ [mm]</th>
<th>$a$ [mm]</th>
<th>$b$ [mm]</th>
<th>$p_0$ [MPa]</th>
<th>$\sigma_{vM}^{\text{max}}$ [MPa]</th>
<th>$z_{\text{max}}$ [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>-30.0</td>
<td>10.00</td>
<td>1.44</td>
<td>3656</td>
<td>2071</td>
<td>0.85</td>
</tr>
<tr>
<td>-25.0</td>
<td>9.86</td>
<td>1.57</td>
<td>3394</td>
<td>1930</td>
<td>0.95</td>
</tr>
<tr>
<td>-20.0</td>
<td>9.59</td>
<td>1.86</td>
<td>2951</td>
<td>1692</td>
<td>1.08</td>
</tr>
<tr>
<td>-15.0</td>
<td>8.66</td>
<td>3.11</td>
<td>1957</td>
<td>1153</td>
<td>1.69</td>
</tr>
</tbody>
</table>

2.2.4.5 Non-Hertzian elastic contacts

A general theory for calculation of contact stresses and deformations in contacts between bodies was presented by Kalker in 1979, see [29], and the corresponding computer code is called CONTACT. However, the theory assumes that the material is linear elastic and that the bodies can be treated as semi-infinite when the local stress state is determined. A simulation using CONTACT resulting in non-Hertzian contact pressures was performed in [30] and a result is given in Figure 16. An interesting conclusion in [30] is that the loading for the worn rail showed less damaging impact on the rail than the Hertzian contact loading corresponding to a new rail.
2.2.4.6 Elastoplastic contacts

As mentioned above, wheel-rail contacts generally involve material plastification which also affects the contact geometry and the surface and subsurface stresses. Here the FE method is increasingly being used. The method enables analysis of complex geometries in three dimensions at the same time as advanced material models can be applied. However, in combination with analysis of rolling contact, the calculations become extremely large and time-consuming. The general analysis of problems related to wheel-rail rolling contact still awaits a major further increase in computer capacity and speed.

Elastoplastic calculations of stresses and deformations in rails are often performed using contact stresses that have been obtained from elastic calculations, see [30], [31] and [32]. In [33], contact stresses on the railhead and in the rail gauge corner are evaluated using Hertzian contact, the CONTACT software and an elastoplastic FE model. As expected, it was found that the calculated maximum contact pressure and maximum von Mises stress were considerably lower, especially at the gauge corner, when using the FE method. Further comparisons are available in [28].

In [34], a wheel (near flange) in contact with a manganese steel (initial yield stress 360 MPa with deformation hardening) crossing nose of radius 13 mm is analysed. The contact force is 111.5 kN. The calculated contact area and maximum contact pressure for Hertzian theory (65 mm², 2566 MPa), CONTACT software (67 mm², 2555 MPa) and elastic FE analysis (70 mm², 2561 MPa) agree well despite the small contact radius, while the elastoplastic FE analysis (128 mm², 1085 MPa) as expected gives a larger contact area and lower contact stresses.

An example of recent modelling work in the ongoing EU project INNOTRACK is given in [35]. Three-dimensional elastoplastic FE models are used to evaluate the contact and subsurface stresses at two-point contact in a switch, see Figure 17. The results will be used for calibrating a simpler and faster two-dimensional model which should give the same maximum von Mises stresses.
19

Figure 17: Result of FE calculation of a wheel in two-point indentation contact with switch rail (contact point P1 and left figure) and stock rail (contact P2 and right figure) of a switch. The assumed axle load 30 tonnes gives normal force at P1 of 114.2 kN and at P2 of 95.6 kN. In each contact point an area of 40 mm x 40 mm is modelled and double symmetry is used to include only one fourth of the modelled volume. The wheel material is elastic and the rail material elastoplastic with initial yield strength $\sigma_y = 410$ MPa. It is observed that the von Mises stresses exceed the initial yield stress which is due to material deformation hardening.

2.2.4.7 Influence of roughness

Surface roughness has been found to have an important influence on the wheel-rail contact stresses, see [36]. The roughness makes the contact stresses deviate from the Hertzian smooth distributions and very high local contact pressures will occur and cause local plastification. After many overrollings this will lead to very high plastic deformations in a layer near to the surface which, however, has only a thickness of a few tens of micrometers. This deformation is believed to be one of the major mechanisms behind rail fatigue. In [36] the roughness-induced stress field and its consequences are being analysed by computations, twin-disc experiments and field observations.

2.2.5 Thermal stresses

Thermal stresses in the rail develop due to the difference between the so-called neutral temperature and the service temperature. For service temperatures higher than the neutral temperature compressive stresses are built up and there is the danger that these may be released by buckling in the rail, with risk of train derailment, an effect on the rail which
sometimes is called “sun kinks”. At temperatures lower than the neutral temperature tensile thermal stresses arise which act as an additional static loading component together with the wheel loads and the residual stresses. Since the tensile thermal stresses reach their peak values in cold winter nights there is an increased risk of rail fracture at this time which is illustrated in Figure 18 [37]. In [38] within the present issue, the authors demonstrate how the thermal stress magnification in conjunction with the crack propagation pattern cause the highest fracture probability of rails in the beginning of the winter when the first really cold nights occur.

For continuously welded straight tracks the thermal stress $\sigma_T$ can be determined as

$$\sigma_T = \alpha \cdot E (T_N - T)$$  \hspace{1cm} (19)$$

with $\alpha$ being the coefficient of thermal expansion (in the order of $12 \times 10^{-6}/^\circ C$ at 20$^\circ$C for rail steels), $E$ the modulus of elasticity (Young’s modulus), $T_N$ the neutral rail temperature and $T$ the service temperature.

The neutral rail temperature is the temperature at which the longitudinal force in the rail is zero. As a first estimate this refers to the temperature at which the track has been installed, i.e. the rail is then anchored at the track and the delivered rail lengths are joined by welding. Note, however, that the neutral temperature is also affected by the in-service history which the track experiences after its installation. It is immediately clear that $T_N$ will be changed when a rail
segment is replaced or a rail is repaired after a fracture event. In addition there are further effects such as track disturbance by tamping at track installation, roadbed freeze-thaw cycles [13] or cumulative vehicle braking on certain track sections [39] which cause not only a change in $T_N$ with time but also a variation along the track. During the initial loading of a new rail the material close to the running surface is significantly deformed which also causes a modification of $T_N$ [40]. Two of the effects mentioned are illustrated in Figure 19.

![Figure 19: Measurements of the daily neutral temperature of a track in Pueblo (Colorado) (according to [40]).](image)

Different methods are available for *in situ* measurement of the neutral rail temperature comprising strain gauge techniques, rail uplift, ultrasonics, magnetics and – most promising – vibration methods. No detailed discussion on this topic will be given here, see, however, [41].

Note that Eq. (19) is only applicable to continuously welded rails. For rails joined by fishbolt plates the thermal stresses are smaller. Even more specific conditions exist in sections with rail portions containing switches [42].

### 2.2.6 Residual stresses

Stresses that exist in a rail without external loading are called residual stresses. They are introduced during the manufacturing process by heat treatment, roller straightening and welding of rail sections or by in-situ welding during installation in the field. Examples on longitudinal components of two residual stress fields due to roller straightening followed by service loading and due to welding are illustrated in Figure 20. They show quite different patterns. After straightening there are tensile residual stresses in the head and in the centre of the foot and compressive residual stresses in the web and at the foot ends, see, e.g. [43-47]. After a few wheel passages the residual stresses in the surface layer of the rail head change to compressive stresses due to plastic deformation (Figure 21) to a depth of 4 to 10 mm [48] (see also [47,50]. Note that the described pattern refers to the central parts along a rail. At the rail ends the residual stress distribution might be different due to a modified technology for rail straightening (pressing instead of rolling). For an investigation of residual stress pattern near the cut end, see [51].

In contrast to the mechanically induced residual stresses, the welding residual stress fields are characterised by compressive stresses in a larger part of the rail head and in the centre of the
foot counterbalanced by tensile residual stresses in the web. This pattern becomes reversed at some longitudinal distance from the weld as illustrated in Figure 22 [52].

\[\text{Figure 20: Residual stress fields in rail sections (longitudinal components); (a) residual stresses due to roller straightening followed by service loading (according to [47], pre-processed in [38]); (b) welding residual stresses in a flash-butt-weld after grinding (according to [53]).}\]

\[\text{Figure 21: Effects of roller straightening and service loading on longitudinal residual stresses along the centre line of a rail (according to [48]; for the service loading effect, see also [6]).}\]
Residual stresses can be determined by finite element simulation taking into account the manufacturing process, the welding in the field and also the initial in-service loading (e.g. [43-45,53,54]). Alternatively, they can be determined by use of various measuring techniques such as the destructive saw-cutting and hole-drilling method [53] or the non-destructive X-ray and neutron diffraction [46,47] methods. For an overview of work up to the early 1990s, see the references in [55].

The magnitude of the residual stresses can be reduced, however moderately, by modifications in the manufacturing procedure [48,56]. Welding residual stresses depend on the employed welding method (flash welding, aluminothermic welding, etc.) [5] and on process parameters such as the cooling rate. An extreme example, comparing air cooling with water cooling, is shown in Figure 23 [5]. Although the effect will be significantly smaller under real conditions it should be kept in mind, e.g., for weld repairs. In [57] the authors demonstrate how the magnitude of the tensile residual stresses can be reduced by short-term reheating of the underside of a rail.
With respect to the rail head various authors (e.g. [5]) have shown that grinding substantially affects the surface residual stresses due to the heat supplied during the process.

Figure 23: Welding residual stresses at centre line (flash butt welding) obtained by air and water cooling (according to [5]).

Note that even sophisticated measuring methods for residual stresses have their limitations and shortcomings. For instance, the accuracy of the neutron scattering technique depends strongly on the volume of material sampled. In order to avoid extraordinary and time-consuming measurements, slices some millimetres thick have to be extracted from the rail which are then investigated as shown in Figure 24. Naturally, thin slices will disturb the complex three-dimensional residual stress state. The residual stresses at one point in the rail cross section as a function of the slice thickness are shown in Figure 25 [58] where the authors report that the residual stress state in the rail is preserved for a 12 mm thick slice.
Figure 24: Residual stress measurement: Typical test samples for neutron scattering.

Figure 25: Effect of slice thickness on the residual stresses at one point on the rail head determined by neutron scattering technique (according to [43]).
3. RAIL FATIGUE

3.1 Failure Scenarios

3.1.1 General remarks

Fatigue cracks in rails can be initiated at the rail head, at the web and at the foot. Their growth can cause spalling of material fragments which will affect the travelling comfort and noise and also dynamic load magnification for both track and rolling stock. If not detected in due time the fatigue cracks can also lead to fracture of the rail which in some cases may cause derailment. Therefore the possible failure scenarios including the nucleation and growth of potential fatigue cracks have to be known as a necessary basis for assessment of damage tolerance.

In this section a brief overview is given of the most important crack types and failure scenarios. It follows the most recent update (2002) of the UIC (Union Internationale des Chemins de fer) Catalogue of Rail Defects [69,70]. For a comparison with other international catalogues, see also [59].

3.1.2 Rail head cracks with surface origin

Typical cracks originating at the running surface are the so-called “head checks” and “squats”.

(a) Head checks

Head checks are groups of fine surface cracks at the running (gauge) corner of the rails with a typical interspacing of 0.5 to 10 mm. Their multiple occurrence makes them particularly dangerous as has been demonstrated, e.g., by the Hatfield accident on 17 October 2000 (see Section 1). The first crack that failed caused a knock-on effect: When also the adjacent pre-damaged rail sections failed the track damage became so extended that it caused derailment with the mentioned tragic consequences.

Head checking preferentially occurs at the gauge corner of the outer rail in curved tracks but is also found at switch or crossing rails. The reason is gross plastic deformation due to friction when the wheel passes. The cracks grow at a flat angle to the running surface in the traffic direction whereby lubrication plays an essential role [60]. They can cause spalling of pieces of material between the cracks (Figure 26) but also - after deviating at some millimetres of growth – cause transverse cracks leading to the eventual fracture of the rail (Figure 27). Transverse sections of head checks are shown in Figure 28.
Figure 26: Spalling originating at head checks.

Figure 27: Fracture of a rail with origin from a head check. The rail was broken up in the laboratory.

Figure 28: Early propagation of a head check (transverse sections).

(b) Squats

Like head checks squats are rolling contact induced defects. They occur in straight or slightly curved cracks, however, not at the gauge corner but at the running surface. In contrast to head checks they occur randomly at isolated sites. Squats and head checks have in common that their existence is not associated with any metallurgical fault but are caused by gross plasticity. Squats grow at a sharp angle with respect to the running surface until they turn into the transverse direction (Figure 29). They are visible at the surface as a widening of the rail/wheel
contact band together with a small depression at the surface sometimes called “dark spot” (Figure 30).

Figure 29: Early propagation of a squat (longitudinal section).

Figure 30: Squat: damage of the running surface (“hot spot” and crack nucleation).

One reason for more frequent occurrence of surface induced rail head cracks, particularly on high-speed tracks, are improvements in the wear resistance of modern rail steels. It should be noted that there is a competition between early fatigue crack propagation and metal removal due to wear. On the one hand improved wear resistance means a reduction of maintenance costs and an extension of the rail life due to the fact that the grinding interval may be increased. On the other hand there is the danger that small cracks which are not worn away may grow to a critical size.

3.1.3 Rail head cracks with internal origin

(a) Kidney-shaped cracks

In former times rail cracks with internal origin rather than surface induced cracks, were dominating the failure statistics. This type of cracks usually initiates from manufacturing defects, e.g. hydrogen shatter cracks, so-called “flakes” (Figure 31). The pre-existent flaw is the nucleus for a so-called “kidney-shaped” crack or “tache ovale” (Figure 32). Note, however, that sub-surface cracks can also initiate in virtually defect-free material.
Figure 31: Transverse section of a rail head containing hydrogen shatter cracks (“flakes”).

Figure 32: “Kidney-shaped” crack in a rail head.

(b) Longitudinal cracks

Special types of subsurface induced cracks are horizontal cracks beneath the gauge corner which can lead to breaking out of material (gauge corner shelling) (Figures 33 and 34) but also to subsequent transverse crack propagation (so-called detail fracture). The latter starts at one or both ends of a surface “shell” (Figure 35). The crack origin is usually about 10 mm below the surface and associated with a band of non-metallic inclusions.

Improved rail materials are the main reason why, at most railway companies today, rail head cracks with internal origin play a less important role compared to cracks with surface origin.
**Figure 33:** Gauge corner shelling at an early stage.

**Figure 34:** Development of gauge corner shelling: transverse section.

**Figure 35:** Transverse crack propagation (detail fracture) starting from gauge corner shelling. Left: damaged rail section of the dismounted rail; Centre: side view of the rail, now broken open; Right: View of the fracture surface. It can be seen that the transverse crack propagated as far as to the rail foot.
3.1.4 Rail web cracks

(a) Longitudinal vertical and horizontal cracks

Cracks in the web are usually caused by poor manufacturing. One example is the longitudinal vertical crack (Figure 36) also known as “piping”. A horizontal crack is shown in Figure 37, the branching at its end in Figure 38. Both types of web cracks will lead to rail fracture.

Figure 36: Longitudinal vertical web crack (“piping”).

Figure 37: Horizontal web crack.

Figure 38: Horizontal web crack. Detailed view of the branching crack tip.
(b) Cracks initiated at machined holes in the web

Machined holes can be fishbolt holes used for joining the rails instead of welding but also holes for insulation joints. Initially, the cracks usually grow at an angle of about $45^\circ$ to the horizontal, but can change their direction at further extension (Figure 39). When crack initiation is caused by vertical stresses due to fishplate restraint, cracks can also grow at an angle of $0^\circ$ as can be seen in the figure. Cracks initiating at holes are particularly dangerous when they occur near the rail ends as in the case of fishbolt holes.

![Figure 39: Web cracks originating from a fishbolt hole.](image)

3.1.5 Rail foot cracks

(a) Transverse rail foot cracks

Rail foot cracks can be transverse or longitudinal. Transverse cracks are usually initiated from galling due to wear and/or corrosion at the rail support. An example of a foot crack around a corrosion pit is shown in Figure 40. Since they are hard to detect they will frequently cause fracture.

![Figure 40: Rail foot crack starting from a corrosion pit at the foot underside.](image)
(b) Longitudinal rail foot cracks

In contrast to the transverse foot cracks the reason for the occurrence of this type of defects is poor manufacturing. Two types of longitudinal foot cracks can be distinguished with respect to their location in the foot. If the foot crack is away from the centre line of the foot it will probably cause a piece of the foot to break away. However, if it is near the centre part complete fracture of the rail can be the consequence. An example of a longitudinal crack is provided in Figure 41.

![Figure 41: Fracture due to a longitudinal foot crack. (a) Side view; (b) foot underside.](image)

3.1.6 Cracks at welds and switches

Meißner and Hug [61] evaluated statistical data of 65,000 kilometres track including 89,000 switches in Germany. They found that 34% of the rail fractures or detected cracks belonged to welds compared to 66% for the rest of the track and that 23% concerned switches compared to 77% for other track positions. The statistics shows that both welds and switches require special attention with respect to their damage tolerance behaviour.

(a) Cracks at welds

The most common used rail welding methods today are flash-butt welding and alumino-thermic or thermite welding. According to Skyttebol [62] the failure rate of thermite welds is 10 times as high as that of flash-butt welds. Nonetheless thermite welds are used world-wide for welding in the field. There are at least three effects which affect the nucleation and growth of cracks in or near welds.

- There is a modification of the material properties (particularly the toughness) on a local scale.

- Welding residual stresses are induced which act as loading components on a pre-existing crack but can also contribute to the nucleation of cracks during the joining process or later on.

- Residual distortions from the joining process affect the straightness and alignment of the rails which influence the dynamic load magnification during train passage.
Since the material at the running surface, and some millimetres beneath, is strongly deformed and compressional residual stresses are generated in that region, and since the global bending stresses are highest in the foot, fatigue cracks in welded rails tend to grow from underneath rather than from the rail head (Figure 42).

![Figure 42: Fracture of an alumino-thermite weld due to weld defects in the foot.](image)

(b) Cracks at switches

Cracks at switches are similar to cracks on the straight track. Their origin and extension will, however, be affected by the geometrical features of the switch rails. Two examples are shown in Figures 43 and 44. The corner crack in Figure 44 is initiated at the site of the asymmetric profile of the so-called flexible zone of the switch rail where the bending stresses are highest or close to this location. Note that a preferred crack initiation site is the rail section where the stress induced by the lateral bending moment reaches its maximum due to the reduced rail foot width.

![Figure 43: Fracture of rail caused by a crack which initiated at a cutout at the foot underside.](image)
3.2. STAGES OF RAIL FATIGUE

3.2.1 Overview

The history of a squat-like rail crack can be subdivided into three sequential stages. At each stage different mechanisms control the crack propagation as illustrated in Figure 45 [63].

(a) At an early stage surface cracks initiate due to ratchetting, i.e. accumulation of plastic deformation under cyclic contact loading. During this early growth the crack propagation rate decreases which is typical for small cracks. A possible reason for this effect could be a gradual buildup of a crack closure mechanism which does not exist initially. It should be noted that wear can remove small surface cracks and in this way counteract crack propagation. However, the trend to optimise modern rail steel with respect to their wear resistance reduces this beneficial effect.

(b) After the crack tip has reached a certain depth, the exact value of which depends on the applied loading and on the ductility and yield strength of the rail material, the crack growth accelerates. The crack propagates by a combined opening (mode I) and sliding (mode II) mechanism. At this stage the crack is still predominantly driven by the wheel-rail contact stresses. However, when the crack becomes larger, and the crack tip moves away from the high stress region, the crack propagation slows down again. The growth pattern is assumed to be significantly affected by entrapped liquids (water or lubrication) [64-66]. The mechanism is mainly the reduction of friction between the crack surfaces by the liquid which in this way promoting mode II and III crack extension. Perhaps there could additionally be a minor effect due to the internal pressure in the crack caused by the liquid entrapped during wheel passage. The resulting effects on the stress intensity ranges $\Delta K_I$ and $\Delta K_{II}$ are illustrated in Figure 46.
Mode II fatigue cracks frequently tend to branch whereby a new crack is initiated at one of the crack flanks. This new crack then tends to propagate in the rail head upwards to the running surface or downwards in web direction. It has to be assumed that entrapped liquid also affects the generation of branch cracks since it causes reduction of crack face friction which prevents or at least reduces wearing at the crack faces.

During stages (a) and (b) the crack grows at an inclined angle with the running surface of about 10-20° (sometimes up to 40°).

(c) The last stage of fatigue crack propagation begins after the crack has branched. Note that the new crack usually does not grow in the vertical direction but again at an inclined angle with the running surface. However, this angle is in the order of 60-80° which is much larger than that of the original crack. The crack continues to grow under mixed mode conditions although the mode I component becomes more important than in stages (a) and (b). The crack is exclusively driven by the bending and shear stresses from the wheel-rail contact loading superimposed by the thermal and residual stresses (see section 2.2.5). The last stage of crack extension is terminated by the fracture of the rail when the stress intensity along the crack front becomes equal to or larger than the fracture toughness of the material. The critical crack size will show a certain scatter partly depending on a statistical variation of the fracture toughness.

Figure 45: Propagation stages of a squat-like fatigue crack.
Figure 46: Effect of a liquid entrapped in a squat-like crack on the mode I and mode II stress intensity ranges $\Delta K$ (simulation according to [65]).

To get an idea of the time or operation history of a subsurface induced crack (shelling) up to its development into a transverse crack (stage c) the reader is referred to Figure 47 [67]. Note that the numbers in the figure are obtained for a special application and should not be generalised beyond this. With respect to the transition between shellings and transverse cracks, simulations in [68] point to the tendency of short shellings to kink into transverse direction, while long shellings continue to grow in-plane.

Figure 47: Formation and development of a surface crack from an internal horizontal defect (shelling) as a function of service time and passed tonnage at a Japanese Shinkansen rail (schematic, according to [67]).
3.2.2 Crack initiation

3.2.2.1 Nucleation and initiation process

The squat-like cracks on top of the railhead considered here are caused by large and alternating plastic strains close to the railhead surface. The strains are induced by normal and shear contact stresses from the rolling wheel of the Hertzian type and by roughness-related local high stresses.

The fatigue initiation process starts with nucleation within the metal grains and is affected by the presence of grain boundaries, precipitates, impurities and inclusions, see [71] and [72]. Cyclic shear-induced dislocations in the grains cause plastic deformations at high enough stress levels. Continued plastic deformations give rise to slip bands surrounded by less affected material. Repeated deformation of such a slip band makes a crack form. A small grain size will suppress dislocations and will therefore contribute to a higher fatigue strength. However, material imperfections lead to stress concentrations that will nucleate cracks and thereby reduce the fatigue strength (see e.g. [73]).

A fatigue crack can be considered to have been initiated when the nucleation and formation of slip bands is completed. An initiated crack is physically quite short; in the order of 0.1 mm.

In numerical modelling work the complex process of initiation must be “summed up” into a macroscopic level. For the fatigue life models discussed in the following the initiation is thus described by material parameters obtained by employing standardized material testing methods.

It should be mentioned that the fatigue problem under contact loads is quite different from a conventional fatigue problem, e.g. in a railway axle. In an axle, bending stresses and surface characteristics are dominating factors and the fatigue cracks always initiate at the surface, normally at fillets with stress concentration, and propagate perpendicularly to the surface. In the rolling contact fatigue problem, important factors are large hydrostatic stresses and rotating stress fields. Residual stresses formed by plastification will suppress further plastification. The initiation could start at the surface or at a subsurface position depending on the loading, although squat-like cracks normally initiate at the surface and propagate into the rail at a shallow angle to the surface. The hydrostatic stresses (compressive) mean that material imperfections will have less influence on the fatigue life than in a conventional fatigue problem.

3.2.2.2 Material response and shakedown maps

The stresses in the railhead induced by rolling contact stresses will, depending on their level and nature, lead to different responses of the material, see Figure 48. For low load levels the response will be perfectly elastic and no permanent deformations will remain after repeated overrollings (cycles). At a certain load level some plastification will occur. However, after a few cycles induced residual stresses will suppress further plastification and subsequent cycles will occur elastically. This is called elastic shakedown. A further increased load level will cause repeated cyclic plastification to occur at each overrolling, which is called plastic shakedown. Above this load level ratchetting will take over which means an accumulation of strain in the material which continues until the material finally ruptures. The phenomena are illustrated in Figure 48 for a uniaxial loading.
So-called shakedown maps have been established for general non-conforming rolling contacts, see Figure 49. In the derivation of the curves the contact geometry, rolling friction, magnitude of normal contact pressure and ductility of the material are considered. The position of the fatigue damage is indicated as being either surface or subsurface. The load factor is $\lambda p_0/k$ (on the vertical axis) where $p_0$ is the maximum contact pressure, $k$ is the material yield strength in shear and $\lambda$ is a factor that depends on the shape of the contact ellipse. The contact friction (as made use of) is given as $\mu$ on the abscissa. Shakedown maps are useful for a first estimation of fatigue impact with expected location (surface or subsurface) for a specific contact load and geometry.

In summary there are three mechanism that contribute to shakedown in wheel-rail rolling contacts: (1) The surfaces develop more conforming contacts due to wear and plastification, (2) residual stresses develop at plastification and suppress further plastification, and (3) the material hardens at plastic deformation.

**Figure 48:** Principle of material response from repeated stress cycles having constant non-zero mean level. Figure from [71].

**Figure 49:** Shakedown map for general three-dimensional rolling-sliding contact. Figure from [71].
3.2.2.3 Fatigue life models

Some approaches to rolling contact fatigue modelling will be reported in the following.

In [74] elasto-plastic FE analysis of the contact problem is combined with a high-cycle fatigue method to predict fatigue at different positions of the rail surface. It is found that fatigue is likely to occur as a squat on top of the rail.

In [31] a tool consisting of two FE models was developed for simulation of residual stresses and plastic strain fields in railheads. Further, the tool can be used for estimating the location and orientation of initiated cracks and the number of loadings until they will occur. In addition to contact stresses also bending was incorporated but was shown to be of less significance. An advanced material model was employed for the ratchetting behaviour of the material. The tool was used in a case study of traffic at a Swedish test site. It was found that low-cycle fatigue rather than ratchetting would be the cause of fatigue initiation. The calculated angle (in the tangential plane of the rail) for surface cracks agreed well with observations from the test site.

In [76] the tool in [31] is further developed. A more advanced material model is used that includes a decaying ratchetting rate. Twin-disc experiments were employed to validate the fatigue evaluation strategy. Results from the tool showed good agreement with data from the Swedish test site regarding position for fatigue initiation, angle of the crack plane and number of wheel passages to initiation.

A method to study rolling contact fatigue and wear is presented in [77]. An earlier so-called brick model is further developed to calculate stresses, strains and damage accumulation at three-dimensional contacts. A method to predict crack initiation is included. Also the effects of microstructure and wear are explored by the method. In [78] the method is employed to study the effects of microstructural changes on fatigue initiation mechanisms. A twin-disc test program provided data on shear strain accumulation, material hardening and microstructural development. Especially, the influence of the content of pro-eutectoid ferrite on the fatigue life was investigated. To simulate the twin-disc experiment a brick model of the pearlitic steel was establish with brick elements of size $1 \mu m \times 1 \mu m$ in a hexagonal patterns with borders representing pro-eutectoid ferrite, see Figure 50. Shear strain accumulation and distribution is calculated for up to 20 000 cycles. The computed results on shear strain and hardness match well with twin-disc results. In [79] the method to model the pearlitic microstructure is further developed.
Figure 50: Hexagonal representation of a standard grade rail steel microstructure with pearlite regions surrounded by pro-eutectoid ferrite. (a) 63 μm /1 μm microstructure. (b) Close-up of 60 μm × 4 μm microstructure with 1 μm × 1 μm brick elements. Figure from [79].

Reference [80] reports simulation work on wheel-rail contact and damage in switches and crossings, see Figures 51a,b. Here the complicated geometry calls for vehicle dynamics simulations of wheel-rail contact forces, creepages and contact positions. An advanced material model described in [81], was calibrated such that the model can mimic the experimentally observed uni-axial and multi-axial ratchetting behaviour of the standard rail material R260, see Figure 51c. For the severe contact load FE simulations for a large number of load cycles are used to predict the irreversible plastic deformations and the work hardening of the material at rail cross-sections. Further, wear simulations at rail cross-sections were performed followed by a summation of the profile changes due to plastic deformation and wear. Rolling contact fatigue was predicted by use of a fatigue index method. A demonstration example shows that an increase of the axle load from 25 tonnes to 30 tonnes would increase the vertical profile change by 27 % and the index for rolling contact fatigue by 10 %.
Figure 51: Damage on switch components and material behaviour: (a) detached part of switch rail, (b) damage due to plastic deformation and wear on crossing noses and wing rail, (c) stress-strain relation for material R260 at cycles 1, 25, 225, 425, 625, 825 and 1023. The square-dotted curves show the experimental response, whereas the solid curves are the calibrated model response. Figure from [80].

In [82] the roughness-induced stress field and its consequences are being analysed by use of computations, twin-disc experiments and field observations. After many overrollings very high plastic deformations will occur in a layer near to the surface which, however, has only a thickness of a few tens of micrometers. This deformation is believed to be one of the major mechanisms behind rail fatigue. See further Section 2.2.4.

3.2.2.4 Improved fatigue life

A number of factors in the wheel-rail system could be modified and improved to increase the rail life. Change of the wheel-rail contact geometry is obviously a possibility regarding both initial geometries and practices for maintenance of wheel and rail.

An important aspect is the competing mechanisms between wear and RCF. Improved wear characteristics of modern rails have meant a higher risk of RCF since initiated cracks are not worn away before they start to propagate. A balance should here be sought for.

Improved RCF characteristics were aimed at in a research effort on surface coating of rails, see [83]. The investigation incorporated dynamic train-track interaction simulations, three-dimensional FE calculations, shakedown theory and laboratory and field tests. The stresses from the FE calculations were used to calculate shakedown diagrams. It was found that the two-material rails with a correct profile could be used to prevent rolling contact fatigue in the traffic situation studied.
3.2.3 Crack propagation

3.2.3.1 Early crack growth in rails

Repeated rolling contacts cause RCF and wear on the railhead. High tangential forces together with surface roughness induce uniaxial plastic deformations in a thin surface layer. Surface-breaking cracks are then initiated on the top of the rail (squat & rail fracture) or at the gauge corner. Once a crack has been initiated (with a length of, say, 0.1 mm) it will grow at a shallow angle (10-25°) from the surface in the direction of the plastically deformed anisotropic material until it reaches a critical length (of, say, 1-2 mm), see Figure 52. At this critical length the stresses and strains at the crack tip will govern the continued growth which can be upwards (spall) or downwards (squat). See further the description in [84].

Fluids (generally water) play an important role in crack propagation in rails. Fluids can lubricate crack faces and a trapped fluid may force the crack faces apart or prevent the crack from closing. For short cracks, lubrication is a necessary condition for growth, while it is unclear whether liquid entrapment phenomena will occur at all, see [84].

The railway accident at Hatfield in October 2000 had its origin in headchecks on the rail. Detection and removal of such cracks is necessary. Knowledge of the rate of growth of short cracks is therefore of great significance for both economy and safety. In the development of optimized maintenance strategies the combined effect of the competing mechanisms of wear and crack growth should be considered.

![Figure 52: Three phases of life of a (rolling contact) fatigue crack initiated at the surface of a rail. The tangential forces act on the rail from right to left as in Figure 53. Figure from [85].](image)

3.2.3.2 Modelling of short cracks

Reference [84] gives an extensive overview of the early crack propagation problem in railheads. It emphasizes that the early crack should be regarded as short from the fracture mechanics aspect. This means that the plastic zone near the crack tip is large in comparison with the size of the crack. Thereby the prerequisites for linear elastic fracture mechanics (LEFM) are violated and elastoplastic fracture mechanics (EPFM) should be employed. A two-dimensional FE model was established, see Figure 53. LEFM calculations were performed and verified against similar calculations in the literature. Results for EPFM were found to deviate strongly from the LEFM results, see Figure 54. The latter are generally given
as stress intensity factors in modes I, II and III. The short cracks in the railhead are mainly exposed to mode II stress intensities. However, in the elastoplastic analysis stress intensity factors cannot be defined. Instead the concept of Crack Tip Shear Displacement (CTSD) was employed. The results in Figure 54 clearly show that the EPFM analysis gives a much higher CTSD than the LEFM analysis.

In Reference [86] a two-dimensional EPFM analysis of propagation of early cracks is performed and compared with twin disc experiments from the literature. A so-called Pineau criterion was employed to establish the crack growth mechanisms. The influence of wear was included. The shortest cracks studied were found to grow in parallel to the surface while the longest (0.4 mm) deviated towards the surface. The strain ranges in the crack area were found to be very high, several times larger than the tensile strain range. The predicted crack paths agreed well with the twin-disc experiments but crack growth rates were overestimated. It was concluded that future research should involve many load passages; only a few could be studied in the referenced work.

**Figure 53:** FE model of a crack. (a) Notation and sign convention. The contact pressure shown corresponds to a combination of normal compressive and negative (driving) tangential traction. (b) Enlargement of FE mesh near the crack tip. Figures from [84].

**Figure 54:** Result for crack slope $\alpha = 20^\circ$, maximum contact pressure $p_0 = 1100$ MPa and friction coefficients $\mu_{\text{surface}} = 0.4$ and $\mu_{\text{crack}} = 0.2$. CSTD from (a) linear elastic and (b) elastoplastic FE calculations. Figure from [84].
3.2.3.3 Further studies on rail cracks

References [87] and [88] address the problem of liquid entrapment for railhead cracks of lengths around 3 mm and larger and use LEFM for the analysis. A recent three-dimensional LEFM analysis of railhead cracks is given in [89].

Elastoplastic computer simulations of the growth of railhead cracks are performed in [90]. Cracks of 3 mm length and larger are studied and competition between wear and crack propagation is considered. The crack growth rate for short cracks is postulated to decrease as their tip moves away from the contact stress field at the rail surface. At a critical length of the crack the growth rate increases since compressive closure forces become lower. It is claimed that the article constitutes the first detailed investigation of the crack driving forces accounting for elastoplastic deformation history. Results from an extensive computer simulation study are given.

Surface cracks in rails are discussed in [91]. The conditions of loading, heavy surface plastification and microstructure under which cracks of length 1 mm and larger propagate are described qualitatively and by use of LEFM concepts. The main question is whether a crack will branch into a direction that may cause the entire rail to break. It is suggested that this would only happen if the crack is fairly long and if high tensile stresses induced by low temperatures of the rail would make the mode I stress intensity larger than the mode I threshold value.

3.2.3.4 Ongoing research

The subject of early crack growth in rails is complex. The growth takes place under a significant influence of plastic deformation. Other influencing factors are crack face friction and the complex rotating stress field when the rolling contact load passes. Only limited research efforts related to rail cracks are presently focussing on this problem. Two examples from the Centre of Excellence CHARMEC, see www.chalmers.se/charmec, are here mentioned which continue the previous studies reported in [84], [86] and [92].

The CHARMEC project MU17 “Elastoplastic crack propagation in rails” is focussing on some basic issues before dealing with applications, see [93]. The so-called crack-driving force (generalized J-integral) is defined in the context of “material forces” which is a vectorial measure of the release rate due to a (virtual) variation of the position of the crack tip. A particular issue is the role of material dissipation at crack advancement. Also interaction between closely spaced cracks is studied.

In the CHARMEC project MU20 “Wear impact on RCF of rails” the interaction between wear and rolling contact fatigue is studied. Wear has an influence by removing incipient cracks (so-called track truncation) and via a change of the contact geometry. The project includes development of strategies for rail profile updating including automatic control of the prediction quality.

3.2.4 Crack branching and deviation

As mentioned, a squat starts growing at a fairly shallow angle relative to the top of the railhead. At a certain length it will either branch/deviate upwards towards the rail surface or
branch/deviate downwards [75]. At a crack length of, say, 5 mm branching/deviation upwards will create an uneven rail surface leading to larger contact forces and increased noise and vibration. Downward branching may lead to fracture of the full rail and thereby constitute a risk for derailment.

The mechanisms of crack branching and deviation are quite complex. A number of FE models have been developed to study the phenomena and the influences of different factors, see, e.g., [94] and [95]. A parametric study is performed in [95] by use of a two-dimensional FE model. Examples of results from [95] are that residual tensile stresses increase the propagation rate resulting in crack branching downwards and that tangential forces from braking with low crack face friction promote crack propagation and branching. The paper [90], as already mentioned in Section 3.2.3, stated that downward branching could occur at high tensile stresses induced by low temperatures in the rail which would make the mode I stress intensity factor larger than the mode I threshold value, a fact which agrees with the results in [95].

### 3.2.5 Final crack propagation

Fracture mechanics based analyses have been performed for a number of simplified crack geometries some of which are summarised in Figure 55. Many investigations (e.g. [96-100]) focussed on elliptical internal cracks (in North American sources also designated as “detail cracks” or “detail fracture”), Figure 55a. Other transverse crack types such as straight and semi-elliptical surface cracks in the rail head (e and d) and in the rail foot (f) and corner cracks at both sites (b, g, d) have also been investigated (e.g. [101-104]). Besides the transverse crack geometries shown in Figure 55, further crack configurations such as web cracks induced by bolt holes and weldments, vertical axial rail head cracks (splits) and cracks in switches have been investigated as well [105-111]. No detailed discussion of the various analyses listed above will be given here. Note that they differ significantly with respect to their model simplifications, e.g., neglecting thermal and residual stresses or mode II loading components.

![Figure 55: Some crack configurations in rails for which fracture mechanics based analyses have been performed in the past.](image-url)
The simulation of fatigue crack propagation is based on the $\text{da/dN-}\Delta K$ curve whereby the crack growth at different points along the crack front has to be separately treated this way allowing for a change in the crack geometry over its growth history. This is essential because the assumption of constant crack geometry would cause significant errors in residual life of the component. Examples of $\text{da/dN-}\Delta K$ curves of rail steels for various ratios ($R = K_{\text{min}}/K_{\text{max}}$) are provided in Figure 56.

![Figure 56: Effect of R-ratio and scatter of crack propagation rate data for rail steels (according to [98]).](image)

Note that the R-ratio effect, at least partially, is an effect of plasticity induced crack closure and can be taken into account by an appropriate analysis, e.g., based on Newman’s crack closure model [113]. This is, however, restricted to mode I loading and certainly unrealistic if the mode II and III contributions to the crack extension are significant.

In order to account for the scatter band in $\text{da/dN-}\Delta K$ data, which is pronounced even within a single material batch, upper bound curves have to be used or, alternatively, statistical analyses should be performed. Upper bound reference curves for rail steels are proposed by a number of authors for both mode I and mixed mode loading (see, e.g. [64,104,113,114]).

In [38] in the present issue a squat-like transverse crack such as shown in Figure 55 (h) (the shape of which deviates from a semi-ellipse like (e)) is modelled. The authors provide analytical $K$ factor solutions for mode I, II and III loading and for thermal, residual and applied wheel loading. A special aspect is that variations in the contact patch, e.g. due to the so-called Klingel movement and its dependency on the wearing states of rail and wheel are taken into account. As has been schematically illustrated in Figure 45 the propagation rate of the fatigue crack increases after it has branched and deviates downwards. The authors show that the crack propagation rate strongly depends on the propagation direction which on its part is a function of the lateral position of contact point patch statistics (Figure 57).
Figure 57: Transverse fatigue crack extension depending on the statistics of the lateral position of the contact patch (which on its part depends on the wearing state of rail and wheel). The difference in the crack propagation rate is approximately 1 (upper figure) to 2.5 (lower figure).

Taking into account mixed mode loading the fatigue crack propagation is simulated analytically as illustrated in the flow chart of Figure 58 [38]. Mode I, II and III $\Delta K$ factors ($\Delta K = K_{\text{max}} - K_{\text{min}}$) are determined for three distinct points (A, B and C) at the crack front. This has to be separately done for the various loading components (thermal and residual stresses and wheel loading). Note that only the wheel loading causes a $\Delta K$. The thermal and residual stresses shift the mean value ($K_{\text{mean}} = (K_{\text{max}}+K_{\text{min}})/2$) or $R$ ratio which refer to $\Delta K$. Subsequently superimposed $\Delta K_{\text{I}}, \Delta K_{\text{II}}$ and $\Delta K_{\text{III}}$ factors are determined for points A, B and C. Note that it is important to separately superimpose the $\Delta K$ values of the different opening modes since with respect of $\Delta K_{\text{II}}$ and $\Delta K_{\text{III}}$ the algebraic signs have to be taken into account which would not be possible on the basis of a $\Delta K_v$. In a last step this equivalent $\Delta K_v$ value is then determined, e.g., by, see [116],

$$\Delta K_v = \frac{\Delta K_{\text{I}}}{2} + \frac{1}{2} \left( \Delta K_{\text{I}}^2 + 4 \left( 1.155 \Delta K_{\text{II}} \right)^2 + 4 \Delta K_{\text{III}}^2 \right)$$

A specific problem arises from the out-of-phase occurrence of the mode I, II and III loading cycles. In [38] the authors based their analyses on the respective minimum and maximum $K$ values, $K_{\text{min}}$ and $K_{\text{max}}$, which occurred at different wheel-crack site distances, but neglected the out-of-phase problem.
Figure 58: Basic scheme of fatigue crack propagation modelling of the squat-like transverse crack according to [38].

As indicated in Figure 58 the resulting $\Delta K_v$ has to be compared with a mixed mode $da/dN-\Delta K_v$ curve for consistency reasons. A number of authors (e.g. [104,109,117,118]) have determined such curves for both in-phase and out-of-phase mixed mode loading of rail steels. The question whether a $\Delta K_I$ based curve would always be a conservative upper bound to mixed mode results or not is not easily answered. Note that some synergy effects between mode I and II could occur: Superimposed mode I loading would cause a reduction of crack face friction this way promoting mode II crack propagation and mode II loading could influence mode I propagation by affecting crack closure [119]. Note that, at least for shorter cracks, liquids entrapped in the crack could also play a role (cf. Section 3.2.1).

A further question discussed in the literature is whether loading sequence effects can be expected to play a role in fatigue crack propagation in rails [120,121]. The idea seems reasonable because the orientation of the overload events is identical for all train passages. However, although there is obviously some sequence effect in mode I loaded specimens [122,123] the situation in rails may be more complicated. The loading sequence effect by its nature is a plasticity-induced crack closure effect well-known for mode I loading. However, as mentioned above, its significance for mode II and III loadings and the even more complex situation of out-of-phase mixed mode loading is not clear at all. Note that the low ductility of the material will also play a role. In [123] the authors found that even in the mode I case, uncertainties in the initial crack length and load levels were of greater influence on the residual life than the sequence effect.
3.2.6 Fracture

Fatigue crack extension is terminated when the rail breaks. This is the case when the cyclic loading $K_{\text{max}}$ is equal to or larger than the fracture toughness of the material which, in general terms, will here be designated as $K_{\text{mat}}$ which usually is specified as $K_{\text{IC}}$ for rail steels. Note that rail steels are rather brittle materials which tend to exhibit increased crack propagation rates at higher $\Delta K$ values due to mixed-in cleavage events, so-called pop-ins. This was, e.g., found in [124] and is also reported in [38] in the present issue. Following [125], in such cases, static mode crack extension contributions to the overall crack growth should generally be expected at $K_{\text{max}}$ values greater than 70% of $K_{\text{mat}}$.

Pop-in events, i.e. small amounts of abrupt crack extensions do not only occur in fatigue but also in monotonic testing. They are the reason for discontinuities in the load-displacement record as shown in Figure 59 for the rail steel investigated in [38].

![Pop-in event in a rail steel (R260, former 900A).](image)

**Figure 59:** Pop-in event in a rail steel (R260, former 900A).

Pop-in events on two rail steels have also been reported in [126] where the authors found them to be triggered by segregation bands of Chromium and Manganese. The crucial question with respect to pop-in events is whether the extending cracks can be expected to arrest or whether they would cause the complete failure of the component under service loading conditions. This depends, among other factors such as the crack arrest capability of the material, on the amount of the potential energy stored in the component at the moment of the pop-in. It is obvious that this amount usually is quite different in a displacement controlled test than in a rail at the track. If component instability cannot be excluded, fracture determination has to be based on the pop-in event, i.e., the critical K factor or J-integral is determined for the area under the load-CMOD curve up to the pop-in event.

The toughness of ferritic or bainitic rail steels usually belongs to the lower ductile-to-brittle transition range which means that some scatter has to be taken into account, see the example in [38] in the present issue. As a consequence statistical treatment of the data has to be performed, e.g., on the basis of the established Master Curve concept of VTT [112].
A further point to be mentioned here is the temperature and loading rate dependency of the rail steel toughness. However, this has a much smaller effect on residual lifetime than the stresses from varying temperatures as described in Section 2.2.5. Three examples of both the temperature and loading rate sensitivity of rail steels are shown in Figure 60 [111] (see also Figure 13 in [38]). In [49] the author reports on dynamic toughness values of a rail steel to be 30% lower than the quasi-static values.

![Figure 60: Temperature and loading rate sensitivity of the fracture toughness of three rail steels. The toughness values are given for one standard deviation. The loading rate at dynamic testing was 2.73 MPa m$^{1/2}$ s$^{-1}$ (according to [111]).](image)

### 3.2.7 Seasonal effects on fatigue crack extension and fracture

Seasonal effects comprise ambient temperature and moisture. Whereas the latter probably affects the occurrence of liquid entrapment in potential cracks with the effects discussed in Section 3.2.1, the ambient temperature will have an effect on the fracture toughness as mentioned above and, much more severe, on the thermal stresses as one component of the crack driving force.

The crack driving force is not controlled by the temperature level but by the temperature difference to the rail neutral temperature. Therefore, the effect of ambient temperature is more severe in regions with large seasonal temperature differences. In any case a large temperature difference, which usually refers to low temperature levels, increases both the fatigue crack propagation rate and the risk of fracture. The way this has to be implemented in a simulation is, however, different.

(a) The effect of low temperatures on the fatigue crack propagation rate

An increase of the temperature-induced stresses does not affect the stress intensity range $\Delta K$, but the mean stress $K_{\text{mean}}$ or the R-ratio. This, however, affects the crack propagation rate at identical $\Delta K$-values such as indicated in Figure 56. Since fatigue crack propagation is a cumulative process, its analysis should be based on average temperature, e.g., at a daily basis.
(b) The effect of low temperatures on fracture

In contrast to fatigue crack propagation, fracture is a non-cumulative process which, as mentioned above, occurs at $K_{\text{max}} > K_{\text{mat}}$. As a consequence, fracture analyses have to be based on the lowest temperatures which usually occur during night time. As a consequence the overall fracture probability is increased by two factors: a larger crack size as the result of previous fatigue crack growth and an increased $K_{\text{max}}$ value at a low temperature. The combination of the two effects is the reason why rails breaks usually occur during the first cold winter nights in Middle Europe (Figure 18) as the authors in [38] in the present issue show in their simulation. Note that the temperature dependence of fracture resistance as discussed in Section 3.2.6 is a further temperature effect to be considered in a damage tolerance analysis. The overall effect of this is, however, much smaller than the impact of the temperature.

Literature


[7] Paper on rail material properties relevant for damage tolerance (in progress, Zerbst, Drewett, Schöne)


