A Study on Contact Fatigue Mechanisms

Bo Alfredsson

Doctoral Thesis no. 44, 2000

Department of Solid Mechanics
Royal Institute of Technology
Stockholm, Sweden
Akademisk avhandling som med tillstånd av Kungliga Tekniska Högskolan i Stockholm framlägges till offentlig granskning för avläggande av teknisk doktorsexamen fredagen den 10 november 2000 kl 9:00 i Kollegiesalen, Administrationsbyggnaden, Kungliga Tekniska Högskolan, Valhallavägen 79, Stockholm.
Preface

The research presented in this thesis was carried out between December 1995 and September 2000 at the Department of Solid Mechanics, Royal Institute of Technology, Stockholm.

First, I would like to express my sincere gratitude to my supervisor Assoc. Prof. Mårten Olsson, for his excellent guidance and cooperation during the course of the work.

I want to express my appreciation to everyone at the department for creating a friendly and inspiring atmosphere. Among colleagues there are some I especially would like to thank for their contribution to this work: Mr. Hans Öberg for his valuable help when designing and conducting the experiments. Messrs. Per Lagercrantz and Bengt Möllerberg for manufacturing test equipment and test specimens. Dr. Per Nordlund for guidance with the Marc-programme. Mr. Mattias Widmark for helping me with the metallurgical investigation. Dr. Stefan Björklund for measuring surface profiles. Prof. Fred Nilsson for support with the statistical computations and the fracture mechanical investigation. Assoc. Prof. Jonas Faleskog for suggestions regarding the statistical treatment.

I also thank my wife Carina for her encouragement during these years.

Finally I would like to express my appreciation to Scania for funding and cooperation during this work.

Stockholm in September 2000

Bo Alfredsson
This dissertation contains a summary and the following appended papers:

Paper A

Paper B

Paper C

Paper D
Abstract

Surfaces subjected to rolling and sliding contacts may suffer from contact fatigue. This thesis deals with solid mechanic aspects of contact fatigue including the description and verification of explaining mechanisms. The new mechanism for surface initiated contact fatigue is based on tensile surface stresses from local asperity contacts. It is also realised that sub-surface initiated contact fatigue is the result of tensile residual stresses that emanate from plastic deformation below the surface. These mechanisms clearly show that contact fatigue cracks follow the same rules as ordinary fatigue cracks in hardened steel.

The thesis contains four papers that treat a new test procedure named Standing Contact Fatigue (SCF). The results of the test procedure have played an important role in the development and verification of the mechanisms for surface and sub-surface contact fatigue. The first part of the research work was experimental. In this part the SCF test properties was decided, crack results confirmed and crack detection methods developed. Here comparative studies were performed using some different materials and mechanical properties. It was verified that SCF could detect differences in contact fatigue resistance.

Next a finite element model of the SCF test was evaluated through the general-purpose program MARC. The model included graded material properties that originate from heat treatment. The residual surface deformation and surface compliance were verified against experimental results. Crack initiation was investigated in two ways. Firstly, the principal stresses at critical locations were computed and plotted in a Haigh diagram. The diagram showed that the cracks initiate in a direction perpendicular to the principal stress with the largest stress range provided that the principal stress is tensile sometime during the load cycle. Secondly, some high cycle multiaxial fatigue criteria, including the Haigh principal stress criterion, was evaluated against the SCF crack initiation results. The surface crack location was predicted by including statistical effects using a weakest-link criterion and a three-parameter Weibull distribution.

The SCF crack propagation was investigated by numerical evaluation of $J_1$ and $J_2$ integrals. The crack initiation and propagation phases were separated with a threshold criterion and a direction criterion. It was found that during crack propagation both surface and sub-surface contact fatigue cracks follow the direction with minimum mode II loading.

**Key words:** contact fatigue mechanism; spall; spalling; surface crack; sub-surface crack; elasto-plastic indentation; contact compliance measurement; mixed-mode fatigue; fatigue crack growth; $J$-integral; multiaxial fatigue; weakest-link.
# Contents

1. **Introduction** 
2. **The Contact Fatigue Damage**
   2.1. The Failure Process
   2.2. Surface Distress
   2.3. Surface and Sub-Surface Cracks
   2.4. Spalling Craters
3. **Contact Applications and Parameters**
   3.1. Hertz Pressure
   3.2. Positive and Negative Slip
   3.3. Elastohydrodynamic Lubrication
   3.4. Surface Topology
   3.5. Lubricant Contamination
   3.6. Residual Surface Stresses
   3.7. Mode I and II Crack Loading
   3.8. Material Inclusions
   3.9. Local Structural Changes below the Contact Surface
4. **Mechanical Modelling of Surface Fatigue**
   4.1. Existing Models
5. **A Mechanism for Rolling Contact Fatigue**
   5.1. Standing Contact Fatigue
   5.2. The Asymptotic Crack Angle
   5.3. Asymmetric Effect of the EHD Contact
   5.4. Crack Propagation
   5.5. Sea-Shell Opening Angle
   5.6. Sub-Surface Initiated Spalls
   5.7. Summary of the Spalling Mechanism
6. **Summary of Appended Papers**
7. **Suggestions for Future Work**
8. **References**

## Appended Papers

<table>
<thead>
<tr>
<th>Paper</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paper A</td>
<td>Standing Contact Fatigue</td>
</tr>
<tr>
<td>Paper B</td>
<td>Standing Contact Fatigue Testing of a Ductile Material: Surface</td>
</tr>
<tr>
<td></td>
<td>and Sub-Surface Cracks</td>
</tr>
<tr>
<td>Erratum</td>
<td>Paper B</td>
</tr>
<tr>
<td>Paper C</td>
<td>Initiation and Growth of Standing Contact Fatigue Cracks</td>
</tr>
<tr>
<td>Paper D</td>
<td>Applying Multiaxial Fatigue Criteria to Standing Contact Fatigue</td>
</tr>
</tbody>
</table>
1. Introduction

The subject of this thesis is surface contact fatigue. Surfaces subjected to repeated rolling and sliding contacts under high loading are exposed to contact fatigue. It is a major failure cause in mechanical applications, such as ball bearings, gears and cams with rolling members. These applications may suffer from a number of other failure modes, but as Tallian concludes in [1], if all of these are avoided, then the rolling contact application will eventually fail by surface contact fatigue. Thus, the number of affected applications alone makes every improvement in contact fatigue resistance important. The work presented here adds to the understanding of how and why contact fatigue develop, which when considered in the design process will lead to simplified design work, reduction of extensive test series and improved contact fatigue resistance of applications.

The research work focuses on a new test procedure named Standing Contact Fatigue (SCF). The study includes test analysis and implications for the understanding of rolling contact fatigue. A mechanism for rolling contact fatigue based on tensile stresses from axi-symmetric asperity contacts is reviewed and supported by the experimental findings. The results have also led to the understanding that there are two different mechanisms responsible for initiation of surface and sub-surface contact fatigue. First, however, the rolling contact fatigue damage, applications and influencing parameters are described.

2. The Contact Fatigue Damage

The surface damage created by rolling contact fatigue has been given various names in the literature. Following Tallian [2], the millimetre sized crater, that is the end result of rolling contact fatigue, is named a spall and the damage process is called spalling. Other names used in the literature are pitting, flaking, micro-pitting and surface fatigue. Spalling examples from gear, cam and bearing applications are displayed in Fig. 1.

2.1. The Failure Process

Although the damage process is not fully understood some events are confirmed [2]. It can be divided into three phases: 1) a brief initial phase of bulk changes in the material, 2) a long, stable phase when only micro-scale changes take place and 3) the final phase when a macro-crack grows to a spall. From a crack growth point of view, the third phase may be divided into early crack growth and final fracture. It is, however, assumed that the bulk of the application life is consumed during the first two phases.

During the first phase, bulk changes in the material structure occur in the highly stressed volume under the contact path. Changes may occur in hardness, residual stresses, austenite and martensite structure. The surface roughness is decreased through reduction of both asperity heights and asperity sharpness through wear and local plastic deformation, see for instance Martin et al. [3] and Voskamp [4]. The surface appearance changes towards that of a polished
surface with smooth low-frequency waviness and scratches remaining from the bottom of the deeper finishing marks.

In the second phase, deformation bands, described as white etching areas, are created by the micro-plastic flow. When created around a defect such as an inclusion or an asperity they may be designated as butterflies. Sometime during this stable phase micro-cracks are initiated at defect locations within the plastically deformed material. The micro-cracks usually occur at the contact surface or between that and the depth of maximum Hertzian shear stress. When numerous micro-cracks initiate at the surface the phase is named *surface distress*.

As one of these micro-cracks grows, a macro-crack is formed and the last phase of the spalling failure process starts. The macro-crack, driven by high Hertzian stresses, propagates down into the material to a level corresponding to the maximum Hertzian shear stress where it turns to a surface parallel direction. In bearing and gear applications the macro-crack branches upward to the free surface to create a spall, bounded by fracture surfaces. In railway tracks on the other hand, the crack branches down to a steep angle against the surface, eventually causing the railway track to fracture, see Bold *et al.* [5].

Figure 1. Contact fatigue damage. a) Spalling craters on the teeth flanks of a pinion or driving gear wheel. The craters on one tooth flank are all located next to each other at one radial position. b) Spalling damage on a cam surface and c) on the surface of a bearing roller. In all three examples the contacts have moved towards the top of the figures.
2.2. Surface Distress

Surface distress or frosting as it is sometimes called, is the result of the two first phases in the contact failure process. The surface is covered with micro-cracks. As the micro-cracks grow the surface becomes undermined in asperity dimensions and microscopic spalls may form. For the unaided eye the surface appears smooth but gray or “frosted”. When investigated through a microscope, either parallel cracks perpendicular to the rolling direction, butterfly formed cracks or micro-scale spalls may be recognised, see Fig. 2a. In perpendicular cuts through the frosted surface micro-cracks can be recognised, as is exemplified in Figs 2b and c. Some typical features of the micro-cracks are according to Alfredsson and Olsson [6] that:

• They are always inclined in the forward direction of the rolling contact with a shallow angle to the surface.
• If sliding is present in the contact, then the cracks are numerous if sliding is against the rolling direction but extremely few when sliding is in the same direction as the moving contact.
• Generally, they halt after 10–30 µm.
• For pure rolling the angle to the surface is between 18° and 28°.
• If sliding against the rolling direction is present, then the angle is larger: 41°–50°.

2.3. Surface and Sub-Surface Cracks

The last phase of spalling starts when some of the asperity scale micro-cracks continue to grow due to the cyclic loading into macro-scale cracks. Whether surface or sub-surface cracks are developed is believed to depend on the location of defects. The surface crack will thus evolve from a defect in the immediate sub-surface material or from a surface asperity. Sub-surface
cracks originate from defects located near but generally above the depth of the maximum Hertzian shear stress [2].

2.4. Spalling Craters

Depending on the initiation site the craters can be divided into surface and sub-surface spalls. The surface spall grows from an entrance crack in the surface in the same direction as the movement of the rolling contact point. Fig. 3 contains examples of surface initiated spalls. Characteristic features are according to Tallian [2] that:

- The entrance wall has a shallow angle to the contact surface, less than 30°. In [6] angles between 20 and 24° were measured. Bastias et al. [7] report similar values.
- Seen from above the spall grows in a V-shape from the starting point. The side walls that grow in straight lines often meet the surface at nearly perpendicular angles.
- The exit wall, also from above, has a semi-circular profile with centre at the entrance point. It meets the surface at a nearly perpendicular angle.
- The overall appearance with a starting point, V-shaped expansion and radial exit gives the crater a shape, which is often described as a “sea shell”. The “sea shell” is always oriented with the head against the rolling direction. Although this form is common it is by no mean typical. Thus other forms such as circular, long and thin, etc. do exist just as often.
- Just upstream of the initiation point a surface defect (nick, dent, furrow or pit) is often noticeable, often with a narrow bridge of unspalled surface between the defect and the spall.
- In some cases the starting point is not sharp but slightly convex [6].
- The spall bottom consists of a series of crack surfaces, often with small “wing” cracks, which have not continued to propagate, see Fig. 3b and Olsson [8]. The bottom runs parallel to the surface roughly at the depth of the maximum Hertzian shear stress.
• The length and width of the spall is often substantially larger than the depth.
• For gears the spalls often appear next to each other in bands crossing the gear tooth flank with starting points just below the pitch circle.

The sub-surface spall has many similarities with the surface initiated spall, in particular regarding size, characteristics of the bottom, sides and exit wall. However, it distinguishes from the surface initiated spall by the following features [2]:
• The sub-surface spall has a steep entrance wall, often with an angle above 45° to the contact surface.
• The originating defect may remain visible on the bottom or may have been lost with the spalled material.
• The typical overall “sea shell” form is less common.

3. Contact Applications and Parameters

Rolling contact fatigue has been found in contact surfaces of a wide variety of applications such as: bearings, gears, cams with rolling followers, rolls of steel mills and wheel to rail contacts. Since all fundamentals of rolling contact fatigue are present in a deep groove ball bearing or a spur gear, these applications are used to exemplify the application mechanics.

When the ball of a deep groove ball bearing rolls in the supporting raceway a combined rolling and sliding contact exists. The contact is spread across the raceway creating a wide but short elliptical contact that is often approximated as a line contact. Pure rolling conditions only exist along two lines, see Fig. 4. Due to the geometries of the ball and raceway, all other interfering surface elements suffer tangential sliding motions relative to each other. Elements in the centre of the raceway experiences friction forces against the rolling movement whereas elements on the raceway sides has friction forces and rolling in the same direction. Thus, for bearings the relative motion is constant through the interaction but variable across the contact width.

Figure 4. Local variation of surface velocity in the contact area of a deep groove ball bearing.
For the spur gear, the contact between two teeth follows the line of action, as is illustrated in Fig. 5a. On the driving wheel it starts at the tooth root and moves towards the top. For the following wheel the movement is reversed, i.e. from top to root. For the first half of the interaction, when the contact is located between the root and the pitch circle of the driving wheel, friction forces are directed towards the root of the driving tooth, counteracting the rotation, Fig. 5b. When the contact passes the pitch circle, the friction forces change direction and acts with the rolling motion. For the following tooth the friction forces are directed in the opposite direction. Thus for the spur gear, relative motion is constant across the contact width but variable through the interaction. The distribution of the relative velocity over the driving gear tooth is displayed in Fig. 5c.

Although the mechanics differs between applications it is realised that the rolling contact consists of some common geometric properties. The nominal contacts can be regarded as two-dimensional or line contact where Hertz contact conditions exists. Locally the contact is either purely rolling or is a combination of rolling and sliding. Furthermore, for the bearing, gear and cam contacts a lubricant is present in order to reduce friction and wear. In order to simplify testing and investigate the influence of different parameters Rolling Contact Fatigue (RCF) tests have been developed based on the common properties, see for instance Hoo [9]. The most frequently used RCF test consists of two rolling contact discs through which the effect of contact pressure, relative sliding motion, lubricant and other parameters can be studied individually.
3.1. Hertz Pressure

The Hertz theory for contact of two parallel elastic cylinders is used to derive the contact pressure distribution, stresses and displacements for the two dimensional line contact. The pressure distribution in the contact region shown in Fig. 6 is, for instance given in Johnson [10],

\[ p(x) = \frac{2P}{\pi a^2 L \sqrt{a^2 - x^2}}, \]  

(1)

where \( P \) is the force acting over the contact length \( L \) and \( 2a \) is the contact width. The maximum contact pressure at \( x = 0 \) is

\[ p_o = \frac{2P}{\pi aL}. \]  

(2)

The contact width \( 2a \), the reduced radius of curvature \( \bar{R} \) and reduced Young’s modulus \( \bar{E} \) are given by

\[ a = \frac{4P\bar{R}}{\pi L \bar{E}}, \]

\[ \frac{1}{\bar{R}} = \frac{1}{R_1} + \frac{1}{R_2}, \]

\[ \frac{1}{\bar{E}} = \frac{1 - \nu_1^2}{E_1} + \frac{1 - \nu_2^2}{E_2}. \]  

(3)

The indices in Eq. (3) refer to body 1 and 2, respectively. The normal stresses \( \sigma_x, \sigma_z \) and the maximum shear stress, \( \tau_1 \), on the \( z \)-axis (\( x = 0 \)) are drawn in Fig. 7. The shear stress reaches a maximum, \( (\tau_1)_{max} = 0.30 p_o \), for \( z = 0.78a \). In the literature \( (\tau_1)_{max} \) is often referred to as the maximum Hertzian shear stress. Note that the normal stresses along the symmetry line in Fig. 7 are compressive everywhere. For the line contact, this is the case for every location inside the bodies, except at the surfaces outside the contact, where all stresses are zero.
For an axi-symmetric Hertz contact the contact radius and contact pressure are given by
\[ a = \left( \frac{3PR}{4E} \right)^{1/3} \quad \text{and} \quad p(r) = p_o \sqrt{1 - \left( \frac{r}{a} \right)^2}, \] where \( p_o = \frac{3P}{2\pi a^2} \).

The maximum shear stress is \( \tau_{1\text{max}} = 0.31 p_o \) for \( z = 0.48a \), when \( \nu = 0.3 \). In this case the radial surface stress is tensile outside the contact, with the maximum
\[ \sigma_{rr} = \frac{(1-2\nu)p_o}{3} \]
at the contact radius, \( r = a \).

Typical contact pressures for bearings under field conditions range, according to Zwerlein and Schlicht [11], from 1.4 to 2.5 GPa. For gears the contact pressures are typically lower.

3.2. Positive and Negative Slip

The local relative tangential motion between surface elements in contact is called slip and is defined as
\[ s_1 = \frac{v_1 - v_2}{v_1}, \]
where $v_2$ and $v_1$ are the local velocities at the contact point of the two bodies. Note that Eq. (6) gives the slip of body 1. If locally $v_2 > v_1$, then a negative slip exists on body element 1.

The presence and sign of the slip will influence the number of cycles required for surface distress to appear. Negative slip shortens the number of load cycles needed for surface distress to initiate and increases the amount of surface distress for a given load level. Positive slip on the other hand counteracts the initiation of contact fatigue cracks. Pure rolling or negative slip also gives different behaviour of surface distress micro-cracks.

3.3. Elastohydrodynamic Lubrication

The theory of elastohydrodynamic (EHD) lubrication takes into account the elastic deformation of contacting bodies as well as the changes of lubricant viscosity with increasing hydrodynamic pressure, see for instance Stachowiak and Batchelor [12]. The lubricant film separating the contacting surfaces is formed under three effects: 1) the hydrodynamic film formation, 2) the modification of film geometry from elastic deformation of the bodies and 3) the transformation of the lubricant viscosity and rheology under high hydrostatic pressure. The viscosity of oil and many other lubricants increases substantially with pressure.

The elastic deformation of contacting bodies results in a central region between the entry and exit with almost parallel surfaces. The 2-dimensional Hertzian contact pressure profile is also modified as is illustrated in Fig. 8. The modifications of the Hertz solution may shortly be summarised as follows:

- The length of the contact increases slightly.
- At the entry region the hydrodynamic pressure is lower than the Hertzian value.
- The opposing surfaces within the contact are almost parallel. The film thickness in this region is described by the central film thickness, $h_c$.

![Figure 8. The two dimensional contact pressure under elastohydrodynamic lubrication. The contact moves from the left to the right.](image-url)
• The lubricant experiences a substantial rise in viscosity as it enters the contact followed by an equally large but sharp decline at the contact exit. To maintain continuity of the flowing lubricant and to compensate for the loss of viscosity at the contact exit, a constriction is formed close to the exit. The minimum film thickness at the constriction is $h_o$.

• A pressure peak usually larger than the maximum Hertzian pressure, $p_o$, is generated just upstream of the constriction. The size and steepness of this pressure peak is dependent on the pressure-viscosity characteristics of the lubricant.

• Downstream of the constriction the pressure is lower than predicted by Hertzian theory.

Approximate expressions of the central and minimum film thicknesses for an elliptical contact can for instance be found in Stachowiak and Batchelor [12]. The elastohydrodynamic pressure profile is valid for pressures up to 3-4 GPa, which is sufficient for most machine elements [12].

3.4. Surface Topology

Commonly, the roughnesses of the two contacting surfaces are added to form the so-called composite root mean square roughness, $RMS$ or $R_q$.

$$R_q = \sqrt{\frac{1}{L} \int_0^L z^2 dx},$$

where $z$ is the combined deviation from the mean surface height. The surface roughness should be investigated in combination with the lubricant film thickness. Therefore, the film thickness ratio, $\Lambda$, is calculated from the composite roughness height and the EHD central film thickness:

$$\Lambda = \frac{h_c}{R_q}.$$

According to Ioannides and Harris [13] good lubricant conditions exists in bearings for $\Lambda = 10$ (full film lubrication) and marginal lubrication with partial lubricant film generation exists for $\Lambda = 1.5$. The range $\Lambda = 0.4$ to $3.5$ is normal in field applications, see Tallian [14]. It should be noted that other asperity characteristics such as traction coefficient, slope and spectrum width are also of importance.

Some combinations of load, slip, lubrication and surface topology may lead to wear. Fan et al. [15] noted that the initial contact fatigue crack under some conditions might be halted or removed by wear.

3.5. Lubricant Contamination

The presence of contamination in the lubricant is one important parameter influencing the life in field applications. Contaminating particles will create plastically formed indents on the contact surfaces when rolled over by the contacting bodies. The depth and steepness of these deformations will determine the reduction in application life. Examples exist where lubricant contaminations have reduced the expected time to failure with up to 90%, leaving only 10%
remaining life, see Jacobson [16]. The three important particle parameters: size, shape and hardness may determine the severeness of a particular contamination.

- Larger particles will obviously leave larger deformations on the contact surfaces. The damage is, however, dependent on particle breakthrough of the elastohydrodynamic film. If the contamination size is smaller than the film thickness, then the particle will pass through the contact without leaving any plastic deformation and there will be no life limiting effects, see Sayles and MacPherson [17].

- Compact or spherical particles are more damaging than flat particles, which, if thin enough, may pass through the contact without leaving indents.

- When passing the contact, soft particles are deformed to a thin lens shape that creates substantially less damage than spherical particles. For a hard particle on the other hand the resulting permanent indent will be deeper.

An experiment by Sayles and Macpherson [17] with large contamination particles present for the initial 30 test minutes followed by filtration to below the film thickness for the following test period showed little improvement in life compared to a test with large contamination present throughout the test. This suggests that the indent damage caused by the contamination in the initial phase of the testing is the cause of failure rather than the continued presence of contamination.

In [18], Lorösch performed tests on fine surfaces subjected to comparable low loads. When full EHL film separation was introduced through increased viscosity the life was increased with, at most, a factor 4. In a second step the lubricant was thoroughly filtered to remove all contaminants larger than the film thickness. This increased component life at least 128 times. For lower viscosity and less clean lubricants, the spalling had been surface initiated, originating from noticeable surface damages. For higher viscosity and clean lubricants the spalls were sub-surface initiated, indicating the presence of another failure process [18].

3.6. Residual Surface Stresses

If an unhardened steel surface were used with the Hertz contact pressures of rolling contact applications ($p_o = 1.5−2.5$ GPa), then severe plastic deformation would appear. Therefore, the material is case hardened, which gives increased surface hardness combined with residual compressive stresses present at the surface. The surface hardness of the case hardened gear material used in the following analysis is 700–800 HV and the residual compressive surface stresses are 180–250 MPa, see Fig. 9. It should be noted that although the hardness and the residual stresses from heat treatment are results of the same process, the material hardness is only dependent on the local material state whereas the residual stress distribution is dependent on the geometry through the equilibrium of expanded and unaffected material. An increasing resistance against surface fatigue with hardness was documented already by Way [19] in 1935. It was also found that the spalls became smaller for a hardened surface.
Figure 9. Residual hardening properties of the plane test specimens used in Papers A through D. a) The local hardness profile below the surface of three materials hardened on two different occasions. b) The residual bi-axial stress profiles throughout 5 and 10 mm thick specimens. The stresses are computed from measurements of transformation strain presented in Paper A.
3.7. Mode I and II Crack Loading

Fracture mechanics are used to model the macro-crack propagation phase. If the relative size of the plastic zone at the crack-tip is small compared to the structure, then Linear Elastic Fracture Mechanics (LEFM) may be used. Hereby, the critical crack length, growth rate and direction of growth may be estimated given the structure load. In general a crack tip can be subjected to three types of loads, or modes, see Fig. 10. In the case of spalling the crack tip loading is a combination of mode I and II, symmetric and anti-symmetric in the crack plane.

For quasi static LEFM under mode I loading, crack growth may appear when the stress intensity factor

\[ K_1 = \lim_{x \to 0} \sigma_{yy}(x, 0) \sqrt{2\pi x} \]  

reaches the critical value \( K_{IC} \). Corresponding criteria can be used for mode II, III or even mixed-mode loading. In the spalling process this criteria will determine when the final branching of the macro-crack towards the surface will happen.

Prior to final crack branching fatigue crack propagation determines the spalling process. In mode I, crack propagation is driven by the stress intensity factor range, \( \Delta K_1 \). If \( \Delta K_1 \) is above a threshold value, \( \Delta K_{th} \), then the empirically found Paris law

\[ \frac{da}{dN} = C(\Delta K_1)^n, \]  

where \( da/dN \) is the crack growth rate, \( C \) and \( n \) are material parameters, is valid.

Combined mode I and II fatigue loading, as exist during the spalling macro-crack propagation phase, has been reviewed by Bold et al. [20] and by Qian et al. [21]. These reviews indicate that the possibility to propagate a crack in another direction than mode I is dependent on material, load magnitude and loading sequence. It requires a ductile material subjected to large loads under symmetric load sequence. Thus, for the current situation with brittle steel sub-

Figure 10. Definition of fracture mechanic modes.
jected to mode I and non-symmetric mode II load, other directions of crack propagation than that of mode I is unlikely.

Gao et al. [22] performed mixed-mode I and II fatigue propagation experiments on steel. When plotting the normalised stress intensity factor ranges $\Delta K_{II}/\Delta K_{th}$ against $\Delta K_{I}/\Delta K_{th}$, three different areas of crack growth were found (limited by two curves in Fig. 11). Below the first curve no growth took place, between the two curves growth was strongly influenced by crack face friction which resulted in crack arrest, and finally above the second curve the crack turned to a growth direction with predominantly mode I loading. Thus, the effect of a mode II field is two-fold: Firstly, the crack will turn towards the pure mode I direction and secondly, the mode II field reduces the mode I threshold, $\Delta K_{th}$, and therefore promote crack propagation.

3.8. Material Inclusions

At the bottom of a sub-surface initiated spall an originating sub-surface defect may remain visible. These defects are located near, but generally above, the depth of the maximum alternating Hertzian shear stress [2]. It is also widely accepted that cleaner roller bearing steel provides an improvement in bearing life [4].

The stress field around inclusions in the area of maximum Hertzian shear stress has been investigated numerically by a number of researchers. Different types of inclusions, such as stiff
spherical particles, spherical cavities, hard $\text{Al}_2\text{O}_3$, calcium oxide, soft MnS or plastic inclusions have been modelled. Liu and Farris a [23] among others found the critical point at the interface between bulk material and defect. A limited plastic zone develops at the inclusion, typically in the $45^\circ$ angle to the rolling direction, see Takemura and Murakami [24]. The plastic zone will create tensile stresses in a limited area around the inclusion, which can be used as the driving force for crack propagation computations. Takemura and Murakami also concluded that hard inclusions such as $\text{Al}_2\text{O}_3$ are more detrimental than soft inclusions such as MnS. Longching et al. [25] concluded that the volume with increased stress level from plastic deformation around an inclusion is small in size, and that the crack can not continue to propagate outside this volume without the presence of another crack driving mechanism.

3.9. Local Structural Changes below the Contact Surface

For high contact pressures the maximum Hertzian shear stress below the contact results in local structural changes and weakening of the material. At loads above 3 GPa, Zwirlein et al. [11] describes the structural changes as formation of Dark Etching Areas (DEA) after $10^5$ to $10^8$ load cycles followed by White Bands (WB) after $10^8$ to $10^{10}$ load cycles. The DEAs are associated with the migration of carbon atoms toward heavily dislocated regions and the WBs are formed by martensite decay which also result in a local drop in hardness and residual stresses. The location and direction of WBs correspond to the location and direction of the maximum shear stresses. As the process continues, the material weakening from martensite decay will eventually lead to the development and propagation of cracks without further presence of material flaws [11]. If, however, impurities are present, then they are believed to speed up the decay process.

Another structural change process due to cyclic rolling contact is the decomposition of retained austenite, see for instance Voskamp [4]. The decomposition of austenite is associated with a volume expansion in the order of 3% of the decomposed volume, which will generate internal compressive stresses beneath the contact surface. The process of austenite decomposition is earlier than the formation of DEAs and WBs. It is noticeable after fewer cycles and for Hertzian contact pressures as low as 1.9 GPa.

The resulting change in residual stresses, size and sign, will depend on the relative volume fraction of decomposed retained austenite and martensite in each case together with the application geometry of affected and unaffected material.

4. Mechanical Modelling of Surface Fatigue

The development of spalls in cylindrical or spherical surfaces in rolling contact under heavy contact loads have been observed and investigated for decades. An early problem definition and explanation attempt of the spalling process was completed in 1935 by Way [19]. Since then numerous attempts have been made to develop models for initiation of surface contact fatigue cracks and their evolution to spalls. In spite of all these attempts, as is concluded by
Blake and Draper [26], Bower [27] and Tallian [2], the mechanism causing rolling contact fatigue to initiate and propagate is still not fully understood.

A spalling model should be able to explain a number of experimentally found observations. Based on experiments, Way [19] pointed out that in order to produce spalling damage it is necessary that

• oil is present,
• the oil viscosity is below a critical value and
• the surface roughness is above a critical value.
• It was also noticed that a hardened surface has higher resistance towards spalling than an unhardened surface.

Bower [27] concluded that apart from explaining the effect of oil presence, the spalling model should also explain

• the forward direction of crack propagation and
• the life dependence on presence and direction of frictional forces.

In the continued search for a rolling contact fatigue model it is of interest to answer the following questions:

• Why do the surface distress micro-cracks turn into the two different ranges of angles to the surface?
• Under what loading condition will a micro-crack continue to propagate in the typical direction and shape of spalls?
• Do the same or two different processes govern surface and sub-surface initiated spalls?

4.1. Existing Models

The traditional models for the development of rolling contact fatigue are all based on the two-dimensional line contact. Although a two-dimensional damage process with material decay and plastic deformation is possible no tensile surface stresses will be present, as seen in the two-dimensional Hertz solution. For this reason fatigue crack propagation in mode II has been suggested, but as discussed in Section 3.7 no independent experiments has been produced with mode II crack propagation in a hardened steel.

A mechanism based on oil entering surface cracks has been proposed. Bower [27] evaluates three possible ways for oil to affects crack propagation:

• The oil reduces friction between crack surfaces and will thereby simplify shear movement. Thus, the crack propagates in shear mode (mode II) driven by cyclic shear stresses from the rolling contact. Bower [27] concludes that, it would explain the dependence on slip and crack angle, but the crack is likely to lock up due to the compressive stresses. It is also likely that the fatigue crack would branch in ±70° to produce a mode I crack.
• The fluid may be forced into the crack by the load with pressure on the crack faces to generate mode I stress intensities at the crack tip. This would give rise to large stress intensities and rapid crack growth. The model, however, unable to account for the spalling life dependence on movement and slip direction [27].

• Fluid may be trapped in the crack and pushed towards the tip creating a hydrostatic pressure with a combination of mode I and II stresses present. This model is consistent with the spalling life dependence on movement and slip direction. Among these three models, this was favoured by Bower [27]. However, it was concluded that the model predicts a complicated pattern of combined mode I and II loading, which makes it difficult to forecast the direction and growth of the crack, [27].

Thus, mechanisms based on oil entering cracks are unsatisfying while they do not explain how the micro-cracks appear in the first place. Furthermore, they can not predict the spalling crack path without introducing severe restrictions on the load pattern.

5. A Mechanism for Rolling Contact Fatigue

Returning to the gear spalling craters in Fig. 1a, it is noted that all spalls are located in a band at the pitch line. In the detailed top view of a spall, Fig. 3a, the pitch line is visible as a horizontal shadow across the surface. Thus, the spalling crack initiated below the pitch line, in the area of small negative slip. It then propagated upward in the figure and finally the crater was formed. Each crater is located at a distinct axial position, indicating that a local phenomenon was responsible for crack initiation.

As described for the bearing and gear applications, the rolling contacts are often regarded as two-dimensional. A real contact surface is however not smooth. It inevitable contains surface asperities on a small-scale level. When an asperity of one surface comes into contact with the other surface a local three-dimensional contact is created. Around a local axi-symmetric contact the radial stress component is tensile, as is seen in Eq. (5). Thus, an asperity acts as a local surface stress-raiser. This tensile surface stress is the foundation of the current mechanism for surface initiated contact fatigue.

![Figure 12. The SCF set-up including the axi-symmetric ring/cone and lateral cracks.](image-url)
5.1. Standing Contact Fatigue

The result of the Standing Contact Fatigue (SCF) experiment shows that a local contact can produce tensile radial stresses sufficiently large to create surface fatigue cracks, see Paper A. The test comprises a through-hardened steel sphere that is repeatedly pressed onto a case-hardened test specimen with a pulsating normal force, see Fig. 12. After some initial load cycles elastic steady-state is reached.

If the total load is above an endurance level, then contact fatigue cracks will eventually appear. In Fig. 13a, a ring formed surface crack has developed. In a cut view, Fig. 13b, the crack turns outward creating a truncated cone, hence the name ring/cone crack. The ring/cone cracks initiate perpendicular to the surface and as they propagate into the material they turn to an asymptotic crack angle, $\beta = 20^\circ$–$24^\circ$. If the test is continued after ring/cone crack development a second, sub-surface, lateral crack will initiate after another 10 times as many load cycles. All lateral cracks are opened when unloaded and show the same characteristic size and shape, see Fig. 13c. The numerical computations in Paper B show that both crack types develop normal to the principal stress with the largest stress range that is also tensile sometime during the load cycle.

Figure 13. Axi-symmetric SCF crack results. a) Top partial view of contact mark with ring/cone crack. b) Cut view of a ring/cone crack with $21^\circ$ asymptotic angle to the surface. c) Cut view of lateral crack.
5.2. The Asymptotic Crack Angle

Given an initial surface perpendicular crack outside a point load such as an asperity or an SCF contact, how will it propagate? A qualitative answer to this question can be found, using some approximations [6].

In Paper A it is shown that outside an axi-symmetric contact the surface stresses are independent of the actual contact pressure distribution. Thus, the total load, \( P \), and the radial crack position, \( r_o \), give the surface stress state. Close to the surface and away from the contact the normal and tangential asperity loading may therefore be assumed to be of a point type, and the stress solutions of Boussinesq and Cerruti can be used. Furthermore, in Paper B it is noted that the presence of a short surface ring/cone crack has little effect on the compliance of the contact surface. Therefore the direction of mode I crack tip loading and the fatigue crack path is assumed to follow the principal stress trajectory, from the initiation point at the free surface, Fig. 14. Using polar co-ordinates, with the origin at the point load and \( z \) as the depth co-ordinate, the trajectories in the symmetry plane are given by the differential equations

\[
\frac{dr}{dz} = \frac{1}{U \pm \sqrt{U^2 + 1}}, \quad \text{where} \quad U = \frac{\sigma_z - \sigma_r}{2\tau_{rz}}. \tag{11}
\]

Eq. (11) was solved by Olsson in [8] for the combined Boussinesq and Cerruti stress field in the region where \( r \gg z \). This gave the asymptotic inclination angle as

\[
\beta = \arctan\left[\left(\frac{1-2\nu}{3\mu} + \frac{2(1+\nu)}{3}\right)^{1/2}\right]. \tag{12}
\]

The values for steel, \( \nu = 0.3 \) and \( 0.3 < \mu < 1 \), inserted into Eq. (12) gives \( \beta = 45^\circ - 49^\circ \). If Eq. (11) is solved numerically for the full Boussinesq and Cerruti stress field in combination with SCF test data (\( \mu = 0.3 \)), then \( \beta = 38^\circ \) is found for relevant crack lengths, see Fig. 14. For

---

Figure 14. Point load model with combined Boussinesq normal and Cerruti traction loads. From a surface perpendicular initiation site at \( r_o \), the largest principal stress trajectories of three load cases are followed.
pure rolling, or the SCF test, only Boussinesq loading is present, here the stress field is completely different in the region $r \gg z$. Eq. (11) gives

$$\beta = \arctan \left( \frac{1 - 2\nu}{3} \right)^{1/3},$$

(13)

which with $\nu = 0.3$ gives $\beta = 27^\circ$.

The full Boussinesq stress trajectory solution is evaluated numerically to $\beta = 23^\circ$, for relevant crack lengths, see Fig. 14. If compressive residual stresses are included, then $\beta \approx 12^\circ$ for relevant crack lengths and asymptotically $\beta \to 0$, see Fig. 14. These full trajectory solutions are compared to the SCF crack path in Fig. 13b. The asymptotic angle of the pure Boussinesq (without residual stresses) loading agrees with the SCF crack, indicating that SCF trajectory computations should be performed without residual stresses. A motivation for this is found in the stress computations in Papers A and B where it is noted that the residual stresses have been altered substantially outside the contact.

The spalling entry angle ($\beta = 20^\circ$–$24^\circ$) and the surface distress angle at pure rolling ($\beta = 18^\circ$–$28^\circ$) correspond well with the asymptotic principal stress angle from a Boussinesq contact load ($\beta = 23^\circ$). Similarly, with traction present the asymptotic principal stress angle ($\beta = 45^\circ$–$49^\circ$) correspond to the surface distress angle with negative slip ($\beta = 41^\circ$–$50^\circ$).

5.3. Asymmetric Effect of the EHD Contact

The rolling contact is a combination of nominal and asperity contacts [6]. The nominal contact gives substantial compressive stresses below the contact, but away from the contact it does not create any surface stresses. Consider an asperity entering or exiting the nominal contact, Fig. 15a. On the side away from the nominal contact, the axi-symmetric asperity contact will give a tensile radial stress. On the other side of the asperity, the large nominal compressive stress field cancels the tensile radial stress, Fig. 15b. Thus, the combined asperity and nominal contact would only create the outward directed part of the ring/cone crack.

Lubricants are used to reduce friction and wear. The pressure sensitivity of lubricants leads to the asymmetric elastohydrodynamic (EHD) contact, where the nominal contact surfaces approach each other in a shallow angle at the entry region, but diverge rapidly at the exit. This asymmetry shifts the intense asperity contacts to the nominal contact entry side, Fig. 15c. Thus, the asymmetry of the lubricated contact favours crack initiation on the entry side. As the cracks propagate down into the material on the entry side of the contact, they will all turn outward, that is forward, as is found in all contact fatigue applications.

Negative slip will introduce extra tensile stresses at the entry side of the rolling contact and thus further enhance the initiation of contact fatigue cracks. Positive slip on the other hand will counteract contact fatigue crack initiation. Thus, the contact fatigue damage should primarily be found in areas with negative slip.
5.4. Crack Propagation

The typical surface distress crack length is between 10 and 30 µm. In [6], this crack length was used to estimate surface distress asperity loading and size. The Stress Intensity Factor (SIF) of a presumed crack following the principal stress trajectory was estimated by applying the principal stress distribution to the faces of an edge crack. Equating the SIF to a reduced threshold (effect of mode II by Gao et al. [22]) gives the required surface distress asperity load. If the same load carrying ability is assumed for the surface distress asperity as found in the SCF simulations in Paper A, then the surface distress asperity radius equals the distance between two surface distress cracks in Fig. 2b. Thus, the properties of surface distress are self-generating, the micro-cracks creating micro-asperities which carries sufficient load to propagate the crack to the typical surface distress crack length.

The micro-crack in front of a macro-asperity will however continue to grow down into the material to form the macro-scale spalling entry crack. As the crack length increases the mode I crack loading from the point contact will decrease. Instead the influence of the nominal line contact will increase. The line contact subjects the crack tip to a sequence of mode II loading, positive (+II) before reaching the crack entrance and negative (−II) after passing the crack. The mode II loading will give four effects: Firstly, it will reduce the mode I threshold, increasing the asperity influenced crack length. Secondly, the combined mode I+II at small crack lengths will turn the crack from the entrance angle to a surface parallel direction. Thirdly, the positive mode II part (+II) for long cracks will kink the crack down into the material creating the wing cracks visible at the bottom of the spall in Fig. 3b. At this part of crack propagation the symmetry of the loading sequence (+II−II) will keep the crack growing in the main asperity mode I direction, that is surface parallel. Fourthly, the negative mode II loading will eventually be

![Figure 15. The asymmetric effect of EHD lubrication on spalling crack development. a) The two-dimensional pressure distribution from the nominal contact approaches an asperity. b) The asperity enters the contact before the neighbouring material creating a point load with tensile radial stresses. c) The asperity leaves the EHD pressure spike together with the neighbouring material.](image-url)
large enough to create a wing crack that reaches the surface. Once this happens the spall will be lost to the surface and the crater is completed.

Way [19] reported smaller spalls for case hardened surfaces. The current model explains this by giving smaller spalling crack entry angles, \( \beta \approx 12^\circ \), and a shallower crack path as is visible in Fig. 14 when negative residual stresses are present. A shallower crack is easier kinked to the free surface by the post contact mode II loading.

5.5. Sea-Shell Opening Angle

The top view opening angle, \( \alpha \), varies between papers: Olsson [8] reports \( \alpha \approx 100^\circ \) whereas Bastias et al. [7] found \( \alpha \approx 50^\circ \). The spall in Fig. 3a has \( \alpha = 105^\circ \). This spalling property depends on how large portion of the ring/cone crack is cancelled by the line contact as the asperity enters the contact, see Fig. 15b.

5.6. Sub-Surface Initiated Spalls

Experiments by Lorösch [18] and others on smooth surfaces and clean lubricants, \( i.e. \) without asperity contact, show a significant increase in life and a different, sub-surface initiated spalling damage. The SCF crack result supports this finding by supplying the lateral crack below the contact after substantially more load cycles. This crack is produced by tensile residual stresses from unloading after plastic deformation, thus a different mechanism. However, tensile stresses normal to the crack face are responsible for both surface and sub-surface initiated contact fatigue.

5.7. Summary of the Spalling Mechanism

The SCF test shows that the tensile stresses from a local three-dimensional contact can be sufficiently large to produce contact fatigue cracks in hardened gear application steel. If the principal stress trajectory is followed from the initiation point at the surface, then a crack profile is formed which corresponds to both the entry crack of spalls and to the profile of surface distress cracks. The first, surface perpendicular, part of the crack may be removed by wear. The \( K_I \) surface distress asperity load reaching the mixed-mode threshold value explains the typical crack length found for surface distress. Macro-scale asperities continue to propagate the spalling macro-crack with the aid of the reduced threshold. The mode II load from the two-dimensional contact increases with crack length and is responsible for the wing cracks at the spall bottom and for final crack breakage to the surface. Compressive residual stresses give a smaller inclination angle, \( \beta \), to the surface and thus a shallower and smaller spalling damage. A qualitative explanation to the observation that all cracks are directed forward is found in the asymmetrical interaction between the local asperity contact and the nominal EHD contact.
6. Summary of Appended Papers

The research work presented in Paper A to D of this thesis introduces and covers some features of the Standing Contact Fatigue (SCF) test. Behind the development of the SCF test lies a desire to study some fundamental aspects of contact fatigue. Surface contact fatigue of components is inherently difficult to analyse due to the complexity of many interacting parameters. Therefore, the number of influencing parameters is kept at a minimum in the SCF test, increasing the possibility of isolating the damage mechanism and finding analytical and numerical solutions. In particular, the surface irregularities such as asperities are believed to be of vital importance to rolling contact fatigue. Thus, the spherical indenter simulates a single asperity in a real application. The goal of the SCF test may be stated as:

- The production of cracks with a stationary normal contact fatigue load will show that lubricants and moving loads are not prerequisites for contact fatigue.
- The crack results will support the importance paid to surface irregularities such as asperities in initiating rolling contact fatigue. The similarities of SCF cracks and micro- and macro-cracks in real rolling contact fatigue surfaces are of importance.
- In particular the produced ring/cone crack verifies the early part of the point-load spalling model described in the thesis.
- The test should serve as a direct way to compare the relative resistance to surface contact fatigue initiation of different materials.

The research work on standing contact fatigue is presented in Papers A to D:

**Paper A**, “Standing Contact Fatigue”, introduces the experimental method where a stationary sphere subjects a test specimen to a cyclic compressive load. If the test load level is above an endurance load limit, then a ring/cone crack is found outside the remaining contact mark after test termination. The experimental series are used to compare standing contact fatigue resistance of three standard gear materials. The test materials are case-hardened, which leads to graded material properties, increased surface hardness and residual surface stresses. The plastically graded specimen materials are evaluated and used in numerical simulations of plastic deformation, crack producing surface stresses and surface strains. The surface stress range is compared to the analytical expression. The crack path is compared to that of rolling contact fatigue damage.

**Paper B**, “Standing Contact Fatigue Testing of a Ductile Material: Surface and Subsurface Cracks”, discusses four different contact fatigue cracks: ring/cone, radial, lateral and median cracks. Fatigue endurance limits and initiation laws are compared for ring/cone and lateral cracks. A compliance based crack detection method is used in combination with the test. Compliance results for ring/cone and lateral cracks are compared to numerical results. Crack initiating stresses are computed numerically for the crack producing load levels. The stresses in different directions and locations are evaluated in order to explain why a particular crack type develops at a certain load.
In Paper C, “Initiation and growth of Standing Contact Fatigue Cracks”, fatigue crack propagation of the ring/cone and lateral crack is investigated. For each crack tip position along the crack paths, the stress intensity factors are determined from numerically computed $J_1$ and $J_2$ integrals. The initial crack lengths are separated from positions where fatigue crack propagation has taken place by the aid of a threshold criterion and a direction criterion. As the cracks propagate, they orient in the direction with mode II crack tip loading close to zero.

Paper D, “Applying Multiaxial Fatigue Criteria to Standing Contact Fatigue” evaluates the ability of the Sines, Haigh principal stress, Findley, Mc Diarmid and Dang Van multiaxial fatigue criteria to predict the initiation of standing contact fatigue cracks. The material fatigue parameters of each criterion are determined based on independent bending and torsion fatigue tests. Each criteria is included as a subroutine in the numerical computation of stresses close to the contact. The predicted positions and directions of ring/cone and lateral cracks are compared to experimental results. Finally, the mean and spread in radial ring/cone crack position are evaluate statistically by a weakest link assumption in combination with a three-parameter Weibull distribution. Although none of the criteria can explain all aspects of the experimental SCF crack results, the Findley and Haigh principal stress criteria appear to have the best potential to explain contact fatigue initiation.

The differences visible in the Haigh diagrams in Papers B and D are due to the use of kinematic and isotropic hardening rules, respectively.

7. Suggestions for Future Work

Based on the presented mechanisms for rolling contact fatigue, which is supported by the SCF results, some qualitative design recommendations can be given. Most recommendations are known from extensive phenomenological studies. The current analysis adds an explanation to what happens and why some parameters are of particular importance.

- Surface irregularities such as asperities serve as local stress raisers in the surface. A smooth surface in combination with good lubrication conditions will prevent surface initiation of spalls.

- It is possible to imagine other surface irregularities. For instance, the unloaded surface may be smooth but contain areas with different stiffnesses. In this case, a stiffer surface part would act as an asperity when loaded. Thus, a homogeneous surface material should be preferred.

- Wear particles and other contaminations can destroy an initially smooth surface. An indentation leading to an asperity or adhered contamination will be more detrimental than a small crater. Note that once contaminations have been present and created surface irregularities the surface resistance to spalling is decreased.

- Compressive residual surface stresses counteract the local tensile stresses from asperities and will therefore give extra protection against surface initiated spalling.
Plastic deformation below the surface from the maximum Hertzian shear stress leads to tensile residual stresses perpendicular to the surface. These stresses will eventually be responsible for sub-surface initiated spalling. Hardened surfaces with high compressive residual surface stresses can better withstand plastic deformation below the surface, given that the case depth is sufficiently large.

Sub-surface material inclusions serve as local stress raisers and crack initiation sites. If plastic deformation from the maximum Hertzian shear stress is present below the surface, then the residual tensile stresses from this macro-scale plastic deformation will continue to propagate the micro-cracks initiated at the inclusions.

Case-hardening gives compressive residual surface stresses and tensile residual stresses in the core. For a convex surface these residual stresses alone lead to tensile stresses perpendicular to the surface in the layer between case and core.

When analysing multiaxial fatigue risk at a cyclic elastic contact in a hardened steel material the Findley criterion is recommended. The Haigh principal stress criterion can also be considered, due to its simplicity.

During the course of the current research a number of areas for the continued research have been identified.

The SCF test could be refined and used to study further aspects of contact fatigue. 1) A smaller indenter means reduced load and bulk plastic deformation. It can be used for a closer study of asperity contacts. The influence of the case properties will also increase. 2) A plane indenter gives a constant contact size and could be used to further study ring/cone crack propagation. 3) The real rolling contact often contains a combination of normal and traction loading. A load that is inclined to the surface could include this.

The effect of mechanical properties on contact fatigue requires further investigation. For instance, the plastic properties of the contacting materials influence the risk of “piling-up” or “sinking-in” around an indent. Combining this property with Rolling Contact Fatigue (RCF) tests will indicate which type of material is the most sensitive to lubricant contaminations. This research will further verify the current tensile stress mechanism for contact fatigue.

Rolling contact fatigue in bearings and gears have been investigated extensively. The damage in cam surfaces have been less investigated, in particular with the current mechanisms for surface and sub-surface initiated contact fatigue.

The combined nominal and asperity contact has been studied qualitatively. A quantitative investigation will verify the forward directed surface distress and spalling macro-crack. Also the spalling opening crack, $\alpha$, could be investigated.

The propagation phase of the spalling crack is controlled by the sequential mode I+II–II loading. This type of mixed-mode crack propagation in hardened steel is not fully investigated, in particular with the predicted spalling sequence. This includes independent mixed-mode crack propagation experiments and simulation.
• The surface layer could be further investigated. The possibility to control surface crack initiation by varied transformation strain including gradients is not fully understood. One may also investigate the mechanical effects of different material layers on contact fatigue.

• Paper D covers high cycle multiaxial fatigue in the SCF test. It is noted that a single criterion that can explain all aspects of the SCF cracks is still missing. Furthermore, in some cases of contact fatigue, cyclic plastic deformation is present. It would be interesting to compare low cycle fatigue criteria under cyclic contact loading.

8. References


